



Transavasya Dillerinin Tarımsal Yayılma Nirengi Desteklemesi¹

MARTINE ROBBEETS*

(Çev. Fatma DAĞ** & Çev. Muhammed Göktürk KORKMAZ***)

Öz

Bu makalede Japonca, Korece, Tunguzca, Moğolca ve Türkçe gibi Transavasya dillerini konuşanların kökeni ve ilk yayılma alanları konularına yönelik sorular genetik, arkeolojik ve dil bilimi çerçevesinde birleşik bir perspektifte 'üçgenleştirerek' ele alınmaktadır. Transavasya dillerinin ortak soylarının Neolitik Çağ'ın başından itibaren Kuzeydoğu Asya'da hareket eden ilk çiftçilere kadar izlenebileceğini ve bu dilleri konuşanların erken yayılmasının tarım tarafından yönlendirildiğini görüyoruz. Bu dillerin yayılması, tarımın ve genlerin dağılmasını yansıtan iki ana aşamayı içeriyordu. İlk aşama, Amur ile ilişkili genlerle darı çiftçilerinin Batı Liao Nehri'nden çevre bölgelere yayıldığı Erken-Orta Neolitik'e kadar uzandığı aşamadır. İkinci aşama, önemli Amur atalarına sahip darı çiftçilerinin Sarı Nehir, Batı Avrasya ve Jomon nüfuslarıyla yavaş yavaş karıştığı ve pirincin de Batı Avrasya bitkileri ve hayvancılık dahil tarımsal paketi genişlettiği Geç Neolitik, Tunç ve Demir Çağlarından kalma olduğu aşamadır. Bu araştırma, antik DNA'dan yeni kanıtlar geliştirerek, Japon ve Kore nüfuslarının Batı Liao Nehri atalarına sahip olduğuna dair son bulguları doğrularken, Transavasya dil ailesinin genetik bir karşılığı olmadığına dair önceki iddialarla çelişiyor. Daha önce yapılan bazı araştırmaların Transavasya bölgesini tarım bölgesinin dışında görmesine rağmen, bulgularımız çitçilik/dil dağılımı hipotezinin Avrasya nüfus dağılımlarını anlamak için hala geçerli bir model olduğunu göstermektedir. Dilbilimi, arkeoloji ve genetik üçgenlemesi, pastoralist ve tarımsal hipotezleri pastoralist ve tarımsal hipotezler arasındaki rekabeti çözüyor ve Transavasya dillerini konuşanların erken yayılmasının tarım tarafından yönlendirildiği sonucuna varıyor.

Anahtar Kelimeler: Japonca, Korece, Tunguzca, Moğolca, Türkçe, Darı çiftçiliği, Amur, Jomon, Avrasya Dillerinin kökeni,

Agricultural Expansion Triangulation Support of Transeurasian Languages Summary

Abstract

In this article, questions about the origin and initial spread of speakers of Trans-Eurasian languages such as Japanese, Korean, Tungusic, Mongolian and Turkish are discussed in a unified perspective within the framework of genetics, archaeological and linguistics, by 'triangulating'. We find that the common ancestry of the Transeurasian languages can be traced back to the first farmers moving in Northeast Asia from the early

¹Triangulation supports agricultural spread of the transeurasian languages, 10 Kasım 2021, <https://www.nature.com/articles/s41586-021-04108-8>

* Mainz Üniversitesi, Almanya.

** Erciyes Üniversitesi, Sosyal Bilimler Enstitüsü, Yeni Türk Dili Yüksek Lisans Öğrencisi, E-Posta: fatmadag1989@gmail.com / ORCID: 0000-0003-2843-1024

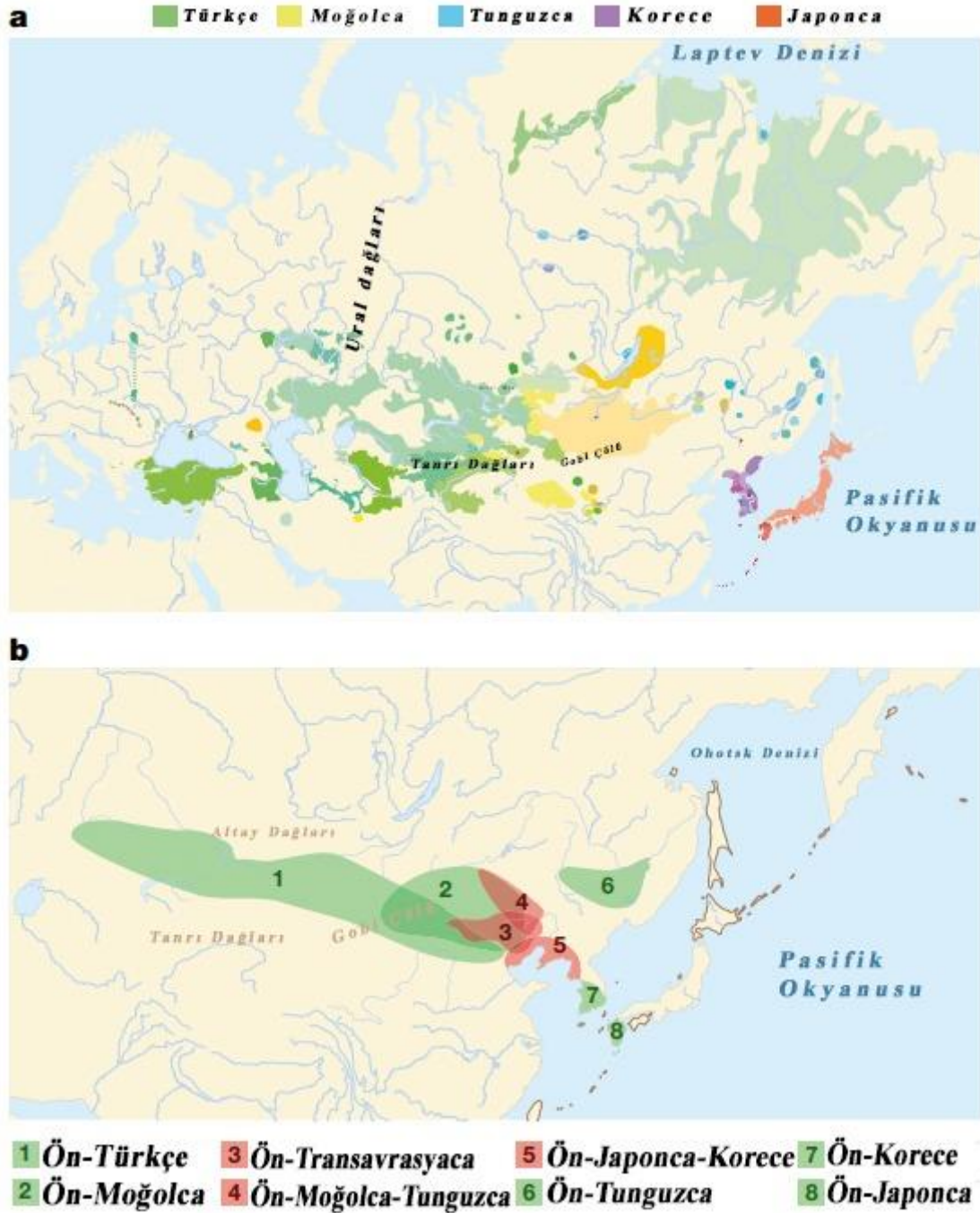
*** Erciyes Üniversitesi, Sosyal Bilimler Enstitüsü, Avrasya Araştırmaları Ana Bilim Dalı Yüksek Lisans Öğrencisi, E-posta: g.korkmaz552@gmail.com / ORCID: 0000-0002-3999-903x

Neolithic Age, and that the early spread of speakers of these languages was driven by agriculture. The diffusion of agriculture and genes coincided with the spread of these languages. The first stage is when Amur-associated genes moved from the Western Liao River to the surrounding territories during the Early-Middle Neolithic. The second phase dates from the Late Neolithic through the Bronze and Iron Ages, when millet farmers with significant Amur heritage mixed with Yellow River, Western Eurasian, and Jomon people, and rice expanded the agricultural package to include Western Eurasian crops and cattle. By advancing new evidence from ancient DNA, our research thus confirms recent findings that Japanese and Korean populations have West Liao River ancestry, whereas it contradicts previous claims that there is no genetic correlate of the Transeurasian language family. Although some previous research regarded the Transeurasian zone as beyond the area suitable for farming, our research confirms that the farming/language dispersal hypothesis remains an important model for understanding Eurasian population dispersals. Triangulation of linguistics, archaeology and genetics resolves the competition between the pastoralist and farming hypotheses and concludes that the early spread of Transeurasian speakers was driven by agriculture.

Key Words: *Japanese, Korean, Tungusic, Mongolian, Turkish, Millet Farming, Amur, Jomon, Origin of Eurasian Languages.*

Transvrasya dillerini konuşanlarının kökeni ve ilk yayılma alanları yani Japonca, Korece, Tunguzca, Moğolca ve Türkçe Avrasya halk tarihinin en çok tartışılan sorunları arasındadır¹⁻³. Anahtar bir sorun, dilbilimsel dağılımlar, tarımsal açılımlar ve nüfus hareketleri arasındaki ilişkidir⁴⁻⁵. Burada genetik, arkeoloji ve dil bilimi birleşik bir perspektifte ‘üçgenleştirerek’ bu soruyu ele alıyoruz. Kapsamlı bir Transavrasya agropastoral ve temel kelime dağarcığı da dâhil olmak üzere, bu disiplinlerden geniş kapsamlı veri kümelerini rapor ediyoruz; Kuzeydoğu Asya’dan 255 Neolitik-Bronz Çağı yerleşim yerinin arkeolojik veri tabanı; ve Kore’ den antik kalıtların bir koleksiyonu tanımlayan, Ryukyu, adalarından ve Japonya’daki erken tahıl çiftçilerinden, Doğu Asya’dan daha önce yayınlanmış kalıtların bir koleksiyonunu tanımlamaktadır. Geleneksel ‘pastoralist hipoteze’⁶⁻⁸ meydan okuyarak Transavrasya dillerinin ortak soylarının ve birincil dağılımlarının Neolitik Çağ’ın başından itibaren Kuzeydoğu Asya’da hareket eden ilk çiftçilere kadar izlenebileceğini ancak bu ortak mirasın Tunç Çağı’ndan bu yana kapsamlı kültürel etkileşimle maskelendiğini gösteriyoruz. Üç ayrı disiplinde kayda değer bir ilerleme kaydetmenin yanı sıra, onların yakınlaşan kanıtlarını birleştirerek, Transavrasyaca konuşanların erken yayılmasının tarım tarafından yönlendirildiğini görüyoruz.

Antik DNA dizilemesindeki son gelişmeler, Avrasya’daki insani, dilsel ve kültürel açılımlar arasındaki bağlantıları yeniden düşünmemizi sağladı. Batı Avrasya⁹⁻¹¹ ile karşılaştırıldığında, Doğu Avrasya hala tam olarak anlaşılammıştır. Kuzeydoğu Asya—İç Moğolistan, Sarı, Liao ve Amur Nehri havzalarını, Rus Uzak Doğusunu, Kore yarımadasını ve Japon adalarını kapsayan geniş bölge— son literatürde özellikle yetersiz temsil edilmeye devam ediyor. Ağırlıklı olarak genetiğe odaklanan¹²⁻¹⁴ veya mevcut veri kümelerini⁴ gözden geçirmekle sınırlı birkaç istisna dışında, Kuzeydoğu Asya’ya gerçekten disiplinlerarası yaklaşımlar azdır.



Şekil.1. Transavasya dillerinin geçmişteki ve günümüzdeki dağılımı.a, 98 Transavasya dil çeşitliliğinin coğrafi dağılımı bu çalışmaya dâhil edilmiştir. Çağdaş diller renkli yüzeylerle, tarihi çeşitlilik kırmızı noktalarla temsil edilir. Açıklama için bkz. Genişletilmiş Veri Şekil 1.b. Neolitik (kırmızı) ve Tunç Çağı boyunca ve sonrasında (yeşil) konuşulan Transavasya atalarının dillerinin yeniden inşa edilmiş yerleri. Ayrıntılı anavatan tespiti için, Ek Veri 4’e bakınız. Tahmini zaman derinliği, Ek Veri 24’te sunulan Bayes çıkarımına dayanmaktadır.

“Altay” olarak da bilinen Transavasya dillerinin dilsel akrabalığı, tarih öncesi dilbilim alanındaki en tartışmalı konulardan biridir. Bu beş grubun tek bir ortak atadan gelip gelmediği sorusu, mirasın destekçileri ve alıntı taraftarları arasında uzun süredir devam eden bir

tartışmanın konusu olmuştur.Son değerlendirmeler gösteriyor ki, pek çok ortak bu diller arasındaki özellikler gerçekten de ödünçlemeden kaynaklansa bile¹⁵⁻¹⁷, yine de Transavasya'nın geçerli bir soy grubu olarak sınıflandırılması için güvenilir bir kanıt çekirdeği vardır^{1,2,18,19}. Ancak bu sınıflandırmayı kabul etmek zaman derinliği, konumu, kültürel kimliği ve Transavasya konuşma topluluklarının atalarının dağılma yolları hakkında yeni sorulara yol açar. Burada, günümüzden önce dördüncü binyılda doğu bozkırında başlayan (GÖ)⁶⁻⁸'de göçebe genişlemelerle Transavasya dillerinin birincil dağılımlarını tanımlayan geleneksel "pastoralist hipotez"e,' tarım/dil dağılma hipotezi' kapsamındaki dağılımları^{5,20,21} ileri sürerek meydan okuyoruz.Bu sorunlar dilbilimin çok ötesine geçtiğinden, onları 'üçgenleme' olarak adlandırılan tek bir yaklaşımda arkeoloji ve genetiği entegre ederek ele alıyoruz.

Dilbilimi

Lehçeler ve tarihsel çeşitler dâhil olmak üzere 98 Transavasya dili için 254 temel kelime kavramını temsil eden 3.193 ortak kökenli yeni bir veri seti topladık (Ek Veri 1).Transavasya dillerinin tarihli bir filo genetiğini çıkarmak için Bayes yöntemlerini uyguladık (Ek Veri 24).Sonuçlarımız, ailenin Ön-Transavasya kökü için 9181 GÖ'lik bir zaman derinliğini (5595–12793 %95 en yüksek olasılık yoğunluğu (%95 HPD)) göstermektedir; Türk, Moğol ve Tunguz dillerinin birliği olan Ön-Altay için 6811 GÖ (4404–10166 %95 HPD); Moğol-Tunguzca için 4491 GÖ (2599–6373 %95 HPD); ve Japon-Korece için 5458 GÖ (3335–8024 %95 HPD) (Şekil 1b).Bu tarihler, belirli bir dil ailesinin birden fazla temel alt gruba ilk bölünmesinin zaman derinliğini tahmin eder.

Transavasya dillerinin alanda genişlemesini modellemek için sözcüksel veri setimizi kullandık (Ek Veri 3, 4).Leksikoistatistiksel, bio-çeşitlilik etkin nokta ilkesi ve kültürel yeniden yapılandırma 1–3,8 gibi klasik yaklaşımları tamamlamak için Bayesci filo-coğrafyayı uyguladık. Daha önce önerilen anavatanların aksine, Altaydan⁶⁻⁸Sarı Nehir'e²², Büyük Kingan Dağları'na²³, Amur havzasına²⁴, Erken Neolitik'te Batı Liao Nehri bölgesinde Transavasya kökenli bir destek buluyoruz. Ailenin Neolitik'te birincil olarak dağılmasından sonra, Geç Neolitik ve Tunç Çağı'nda daha fazla dağılma meydana geldi. Moğol dillerinin atası kuzeye Moğol Platosu'na genişledi, Ön-Türkçe doğu bozkırı üzerinden batıya ve diğer dallar doğuya doğru hareket etti: Amur-Ussuri-Khanka bölgesine Ön-Tunguzca, Ön-Korece Kore Yarımadası'na ve Ön-Japonca (Şekil 1b) Kore üzerinden Japon adalarına.Ayrıca belirli bir zamanda belirli bir bölgede atadan kalma konuşma toplulukları için kültürel olarak tanısal olan öğeleri belirledik. Ortak ata dilleri Ön-Transavasyaca, Ön-Altayca, Ön-Moğolca-

Tunguzca ve Ön-Japonca-Korece gibi Neolitik'te ayrılan, ekim ile ilgili küçük bir kalıtsal kelime çekirdeğini yansıtır (“tarla”, “ekmek”, “bitki”, “büyümek”, “yetiştirmek”, “kürek”); darı, pirinç veya diğer mahsuller hariç (‘darı tohum’, ‘darı lapası’, ‘ahır darı’); gıda üretimi ve muhafazası (“fermente etme”, “öğütme”, “ezerek hamur haline getirme”, “demleme”); yerleşik hayata işaret eden yabancı yiyecekler (“ceviz”, “meşe palamudu”, “kestane”); tekstil üretimi (“dikiş”, “kumaş dokuma”, “tezgâhla dokuma”, “döndürme”, “kumaş kesme”, “rami”, “kenevir”); ve tek evcil hayvanlar olarak da domuzlar ve köpekleri gösterir.

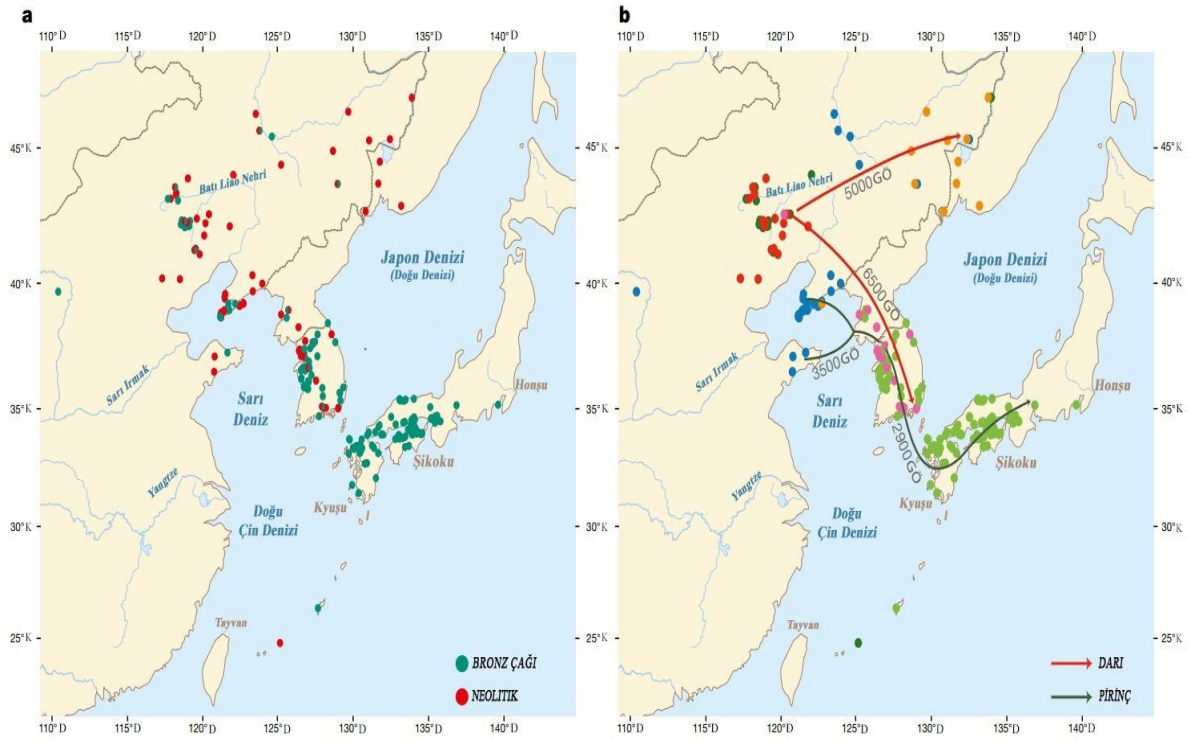
Buna karşılık, Tunç Çağı'nda ayrılan bireysel alt aileler, Türkçe, Moğolca, Tunguzca, Korece ve Japonca gibi, pirinç, buğday ve arpa ekimi ile ilgili yeni geçim terimleri ekledi; mandıracılık, sığır, koyun ve at gibi evcilleştirilmiş hayvanlar, çiftçilik veya mutfak aletleri ve ipek gibi tekstiller (Ek Veri 5). Bu kelimeler Çeşitli Transavasya dili ve Transavasya olmayan dilleri konuşan Tunç Çağı popülasyonları arasında dilsel etkileşimden kaynaklanan ödünçlemelerdir.

Özetle, yaş, vatan, özgün tarımsal söz varlığı ve Transavasya ailesinin temas profili, çiftçilik hipotezini destekler ve pastoralist hipotezi hariç tutar (Ek Veri 5).

Arkeoloji

Neolitik Kuzeydoğu Asya yaygın olarak geniş bitki ekimi ile karakterize edilmiş olsada²⁵, Transavasya için en önemli süpürge darısının ekiminin 9000 GÖ'lerde²⁶⁻²⁹ başladığı batı BatıLiao havzasında tahıl tarımı çeşitli yerleşik merkezlerden genişletildi. Yayınlanmış literatürden veri çıkararak 255 Neolitik ve Tunç Çağı alanı için 172 arkeolojik özellik (Ek Veri 6, Şekil 2a) puanladık ve Kuzey Çin, Primorye, Kore ve Japonya'dan karbon-14 tarihli erken mahsul kalıntılarını (Ek Veri 9) doğrudan 269 envanterine derledik.

255 alanı kültürel benzerliğe göre kümeleyen Bayes analizimizin ana sonuçları (Ek Veri 25) Şekil 2b'de görselleştirilmiştir. BatıLiao havzasında darı çiftliği ile ilişkili iki dalın bulunduğu neolitik bir kültür kümesi buluyoruz: Kore Çulmun dalı ve Neolitik kültürlerin bir kolu Amur, Primorye ve Liaodong'u kapsıyor. Bu, darı tarımının Kore'ye 5500 GÖ'ye ve Amur üzerinden Primorye'ye 5000 GÖ'ye kadar yayılmasıyla ilgili önceki bulguları doğrulamaktadır^{30,31}.



Şekil.2|Arkeolojik veri tabanında yer alan alanların mekansal-zamansal dağılımı ve kümelenmesi.a,Neolitik (kırmızı) ve Tunç Çağı'ndan (yeşil) 255 alanın coğrafi dağılımı. **b,** Renkli noktalar, incelenen alanları, Ek Veriler 25'teki Bayes analizi ile uyumlu olarak kültürel benzerliğe göre kümeler ve zaman ve mekanda darı ve pirincin yayıldığıının göstergesidir. Şekil 2'deki Arkeolojik sit alanlarının dağılımı Görsel.1'deki çağdaş dillerden daha küçüktür çünkü Neolitik ve Tunç Çağı'nda dilsel alt grupların erken dağılımına ve tarımın doğuya yayılması ile dil dağılımı arasındaki bağlantılara odaklanıyoruz.

Analizimiz, Batı Liao bölgesindeki Tunç Çağı sitelerini Kore'deki Mumun siteleri ve Japonya'daki Yayoi siteleri ile daha da kümeler. Bu, GÖ'nin dördüncü bin yılında Liaodong-Şandong bölgesinin tarımsal paketine pirinç ve buğdayın nasıl desteklendiğini yansıtıyor. Bu ürünler Kore Yarımadası'na Erken Tunç Çağı'nda (3300–2800 GÖ) ve oradan da 3000 GÖ'den sonra Japonya'ya taşınmıştır (Şekil 2b).

Nüfus hareketleri tek yönlü hareketlerle bağlantılı olmasa da arkeolojik kültürler, Kuzeydoğu Asya'daki Neolitik tarım genişlemeleri, ekim ve hasat için taş aletler ve tekstil teknolojisi gibi bazı teşhis özellikleriyle ilişkilendirildi³² (Ek Veri 7). Evcilleştirilmiş hayvanlar ve sütçülük Batı Avrasya'daNeolitiğin yayılmasında önemli bir rol oynamıştır ancak köpekler ve domuzlar dışında, veri tabanımız Tunç Çağın'dan önce Kuzeydoğu Asya'da hayvan evcilleştirmesi için çok az kanıt göstermektedir (Ek Veri 6). Tarım ve nüfus göçleri arasındaki bağlantı, Kore ile batı Japonya arasındaki seramik, taş aletler ve ev ve mezar mimarisi arasındaki benzerliklerden özellikle açıktır³³.

Önceki çalışmalara dayanarak, çalışmamızda bölgeler arasında darı çiftçiliğinin tanıtılmasıyla ilişkili demografik değişikliklere genel bir bakış sunuyoruz (Genişletilmiş Veri Şekil 3). Ayrıntılı çeltik tarlalarına yatırım yapan ıslak pirinç çiftçileri, nüfus artışını fazladan emekle emerek tek bir yerde kalma eğilimindeyken, darı çiftçileri tipik olarak daha genişleyici bir yerleşim modelini benimsedi³⁴. Geç Neolitik^{35,36}de bir nüfus çökmesinden önce Kuzeydoğu Asya'da Neolitik nüfus yoğunlukları arttı. Tunç Çağı daha sonra Çin, Kore ve Japonya' da üstel nüfus artışları gördü.

Kalıtımbilim

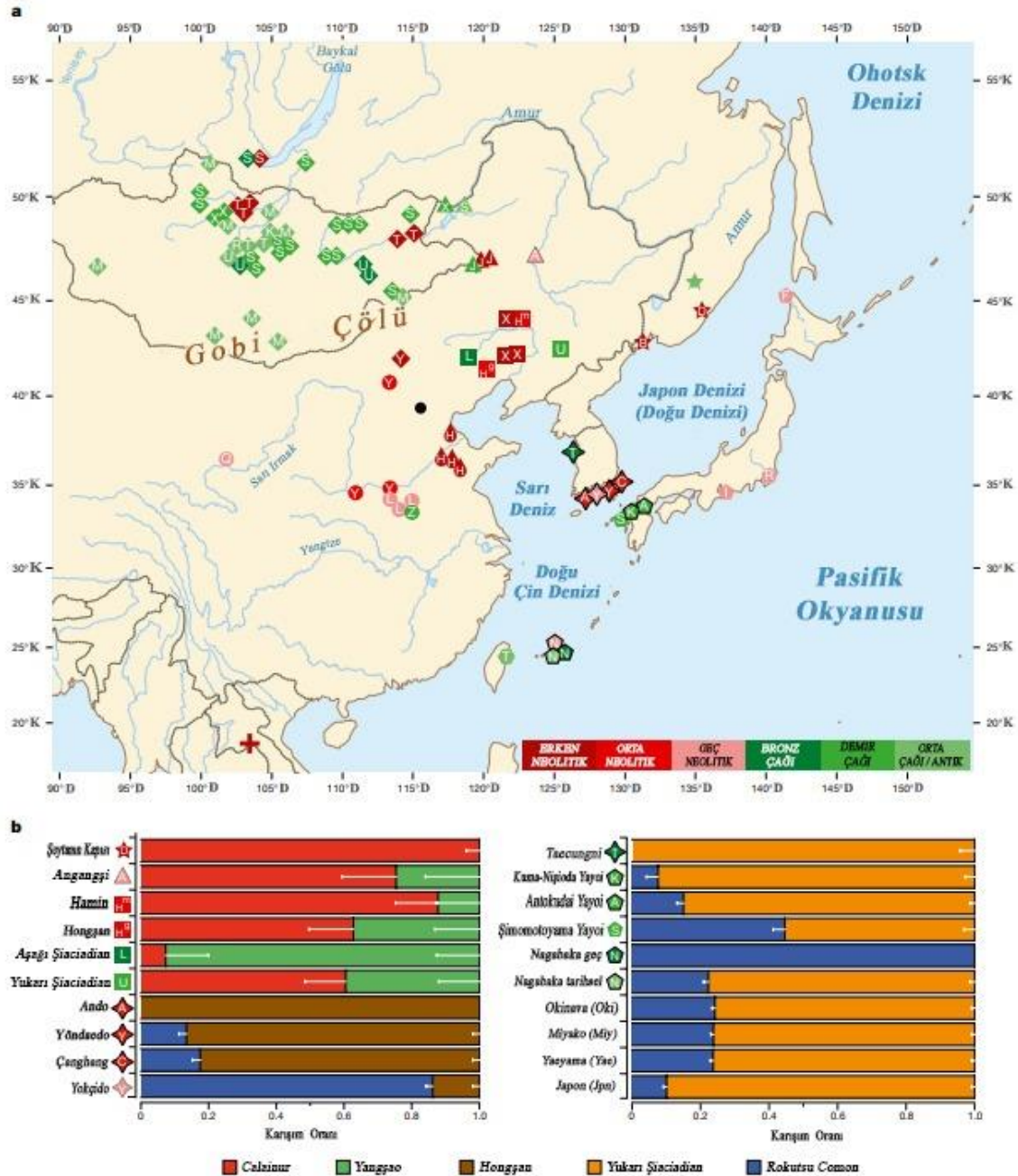
Amur, Kore, Kyuşu ve Ryukyus'tan kimliği doğrulanmış 19 antik bireyin genomik analizlerini rapor ediyor ve bunları doğu bozkır, Batı Liao, Amur ve Sarı Nehir bölgeleri, Liaodong, Şandong, Primorye ve Japonya'yı 9500 GÖ yılları arasında kapsayan yayınlanmış genomlarla birleştiriyoruz. ve 300 GÖ (Şekil 3a, Genişletilmiş Veri Şekil 4, Ek Veri 11, 13, 17). Bunları günümüzün 149 Avrasya nüfusuna ve 45 Doğu Asya nüfusunun temel bileşen analizine (PCA) yansıttık(Genişletilmiş Veri Şekil 5-8).Şekil 3b, kilit antik nüfuslarımızı , Calainur'un Amur, Yangşao Sarı Nehri ve RokutsuJomon genomunu temsil ettiği beş genetik bileşenin bir karışım olarak modellerken, Batı Liao Nehri'ndeki Hongşan ve Yukarı Şiaciadian, Sarı Nehir ve Amur genomlarından oluşur (Ek Veri 16'daki çeşitli Doğu Asya genetik bileşenlerinin qpAdm karışımı).

Amur'daki Çağdaş Tunguzca ve Nihiv konuşanlar sıkı bir küme oluşturmaktadır¹³ (Genişletilmiş Veri Şekil 5). Baykal, Primorye ve güneydoğu bozkırlarından gelen neolitik avcı-toplayıcıların yanı sıra Batı Liao ve Amur'dan gelen çiftçilerin tümü bu küme içinde yansıtılır (Genişletilmiş Veriler Şekil 8-10).

Geç Neolitik Angangşi (Ek Veri 12) Amur benzeri ataların yüksek bir oranını gösterirken Batı Liao Neolitik darı çiftçileri, zamanla Sarı Nehir genomuna doğru kademeli geçiş¹²ile önemli bir orandaAmur benzeri bir ataya sahip olduklarını gösterir. Batı Liao Nehri'nde erken Neolitik genomlardan yoksun olsak da bu nedenle Amur benzeri ataların, Baykal, Amur, Primorye, güneydoğu bozkır ve Batı Liao'yu kapsayan yerli Neolitik öncesi (veya geç Paleolitik) avcı-toplayıcıların orijinal genetik profilini temsil etmesi muhtemeldir.Bu, Moğolistan ve Amur'dan gelen antik genomlarda Sarı Nehir etkisinin yokluğunun Transavasya dil ailesinin Batı Liao genetik korelasyonunu desteklemediği sonucuna varan yakın tarihli bir genetik çalışmayla¹³ çelişmektedir.

PCA (Genişletilmiş Veriler Şek. 8-10) Moğolistan'dan gelen Neolitik bireylerin, batı Avrasya'dan Bronzdan Orta Çağ'a kadar geniş gen akışı ile yüksek Amur benzeri soy içermesi için genel bir eğilim göstermektedir³⁷. Oysa Türkçe konuşan Hunlar³⁸, Eski Uygurlar ve Türkler son derece dağınıktır. Moğolca konuşan³⁹ Demir Çağı Siyenpileri, Amur kümesine Şiveyler, Cücenler, Hıtaylar ve Antik Çağ ve Orta Moğol Hanlığından daha yakındır.

MAKALE



Şekil.3|Eski genomların mekânsal-zamansal dağılımı ve karışımı.a,Zaman ve mekanda bulunan eski genomlar. Detaylı gösterge için bakınız genişletilmiş bilgi görsel 4. **B**, Bu çalışmadan 20 önemli antik nüfusun QpAdm proksimal katkı modellenmesi. X eksen, y

eksenindeki hedef nüfuslar için soy oranı tahminlerini gösterir; hata çubukları, 5 cM blok çakma ile tahmin edilen ± 1 s.e.m. aralığını temsil eder.

Amur ile ilgili ataların izi Japonca ve Korece¹³ konuşanlara kadar izi sürülebildiğinden, tüm Transavasya dili konuşanlar tarafından ortak olan özgün genetik bileşen gibi görünmektedir. Eski Kore genomlarını analiz ederek Jomon soyunun Yarımada'da (Peninsula) 6000 GÖ (Şekil 3b, Ek Veri 13) mevcut olduğunu bulduk (Ek Veri 12).

ProksimalqpAdm modellemesi (Ek Veri 13) şunu önerir: Neolitik Ando tamamen Hongşan ile ilgili bir atadan türetilebilirken, Yöndaedo'nun yalnızca sınırlı ayrımı olmasına rağmen Jomon'un yüksek oranda Hongşan soyuna sahip bir karışımı Yöndaedo ve Çanghang olarak modellenenebilir. Kore'nin güney yakasındaki Yokçido yaklaşık %95 Jomon soyu içerir. Genetik analizimiz Bronz Çağı Taecungni için olası Doğu Asya ataları arasında ayrım yapmasa da Tunç Çağı tarihi göz önüne alındığında, en iyi Üst Şiaciadian olarak modellenenebilir; olası bir minör Jomon katkısı istatistiksel olarak önemli değildir ($P = 0.228$; Ek Veri 16). Bu nedenle, günümüz Korelilerinde göz ardı edilebilir bir Jomon katkısı ile gösterildiği gibi, Neolitik Korelilerde (%0-95) Jomon soyunun heterojen bir varlığını ve zaman içinde nihai olarak ortadan kalktığını gözlemliyoruz. Genetik verilerimizin şu anda sınırlı örneklem büyüklüğü ve kapsama alanı nedeniyle bu hipotezi test etme çözünürlüğü olmasa da Taecungni'de önemli bir Jomon bileşeninin olmaması, günümüz Korelileriyle bağlantılı saptanabilir bir Jomon atası olmayan erken nüfusların, pirinç çiftçiliği ile Kore yarımadasına göç ettiğini ve Neolitik nüfusları bir miktar Jomon katkısı ile değiştirdiğini gösterir. Bu nedenle, çiftçiliğin Kore'ye yayılmasını, Hongşan tarafından Neolitik darı çiftçiliğinin tanıtımı için ve Yukarı Şiaciadian tarafından Bronz Çağı'nda pirinç tarımının eklenmesi için modellenen farklı Amur ve Sarı Nehir gen akışı dalgalarıyla ilişkilendiriyoruz.

Yayoi çiftçilerinden alınan genomların analiz edilmesi (Ek Veri 12), Taecungni gibi onların da Jomon ataları, Bronz Çağı Üst Şiaciadian ataları ile karıştırılmış yerliler olarak modellenenebileceğini bulduk. Sonuçlarımız, Bronz Çağında Kore'den Japonya'ya yoğun göçü destekliyor.

Miyako Adası'ndan gelen Nagaoka genomları (Ek Veriler 12), bildiğimiz kadarıyla, Ryukyus'tan gelen ilk genom çapında verileri temsil eder. Holosen nüfuslarının Tayvan'dan⁴⁰ güney Ryukyusa ulaştığı önceki bulguların aksine, sonuçlarımız tarih öncesi Hagabaka nüfusunun kuzeydeki Jomon kültürlerinden kaynaklandığını göstermektedir (Genişletilmiş

Veriler Şek. 7). Erken modern dönemden önce, Jomon'danYayoi benzeri ataya genetik geçiş, tarımın ve Ryukyuan dillerinin bu bölgeye geç gelişini yansıtıyor.

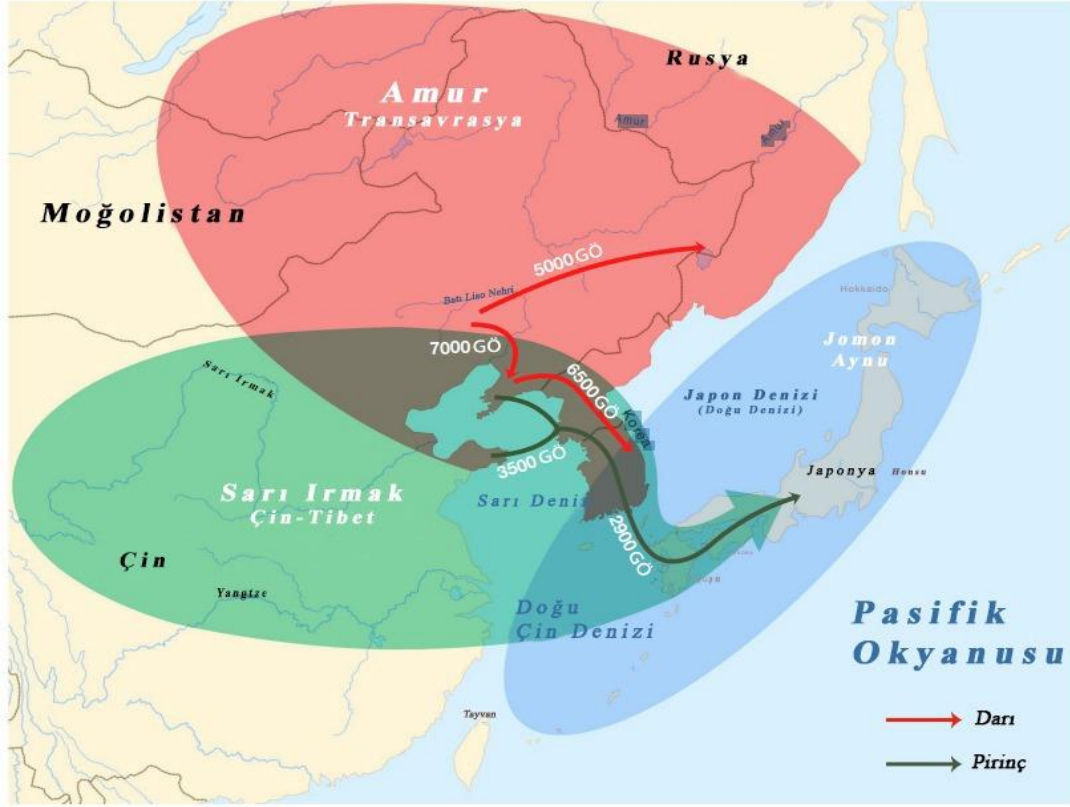
Tartışma

Dilsel, arkeolojik ve genetik kanıtların üçgenleştirilmesi, Transavasya dillerinin kökenlerinin, Neolitik Kuzeydoğu Asya'daki darı yetiştiriciliğinin başlangıcına ve erken Amur gen havuzuna kadar izlenebileceğini göstermektedir. Bu dillerin yayılması, tarımın ve genlerin dağılmasını yansıtan iki ana aşamayı içeriyordu (Şek. 4). Transavasya ailesindeki birincil bölünmelerle temsil edilen ilk aşama, Amur ile ilişkili genlerle ilişkili darı çiftçilerinin Batı Liao Nehri'nden bitişik bölgelere yayıldığı Erken-Orta Neolitik'e kadar uzanır. Beş kardeş dal arasındaki dilsel temaslarla temsil edilen ikinci aşama, önemli Amur atalarına sahip darı çiftçilerinin Sarı Nehir, Batı Avrasya ve Jomon nüfuslarıyla yavaş yavaş karıştığı ve pirincin de dahilbatı Avrasya bitkileri ve hayvancılık dahil tarımsal paketi genişlettiği Geç Neolitik, Tunç ve Demir Çağlarından kalmadır.

Mekân-zamansal ve geçim kalıplarını birleştirdiğimizde üç disiplin arasındaki bariz bağlantıları buluyoruz (Ek Veri 26).BatıLiao bölgesinde dokuzuncu bin yıl civarında darı ekiminin başlaması, önemli Amur-bağlı atalarla ilişkilendirilebilir ve ataların Transavasya konuşma topluluğu ile zamanı ve mekanı örtüşebilir.

8000 GÖ^{41,42} olarak tahmin edilen Çin-Tibet ailesi ile Yukarı ve Orta Sarı Nehir'den^{13,14} gelen Neolitik çiftçiler arasındaki yakın tarihli ilişkilere uygun olarak, bulgularımız Kuzeydoğu Asya'daki iki büyük dil ailesinin kökenleri arasında bağlantı kuruyor: Sarı Nehrindeki Çin-Tibet ve Batı Liao Nehrindeki Transavasya. Antik Transavasya dilinde ve genlerinde Sarı Nehir etkisine dair kanıt eksikliği, arkeobotaniğin²⁸ de bahsi geçen çok merkezli darı tarımı önerisiyle tutarlıdır.

Darı evcilleştirmesinin dokuzuncu ila yedinci bin yıllık GÖ'deki ilk aşamalarına nüfus artışı eşlik etmesi (Genişletilmiş Veri Şekil 3), Batı Liao bölgesinde çevresel veya sosyal olarak ayrılmış alt grupların oluşumuna ve Altay ve Japon-Korece konuşanlar arasındaki bağlantının kopmasına yol açıyor.



Şekil.4|Kuzeydoğu Asya’da dilsel, tarımsal ve genetik yayılımların entegrasyonu.Amur soyu kırmızı, Sarı Nehir soyu yeşil ve Jomon soyu mavi olarak işaretlenmiştir. Kırmızı oklar, Neolitik dönemde darı çiftçilerinin doğuya göçlerini göstermekte ve belirtilen bölgelere Korece ve Tunguzca dillerini getirmektedir. Yeşil oklar, Geç Neolitik ve Tunç Çağı’nda pirinç tarımının entegrasyonunu işaret ederek Japon dilini Kore üzerinden Japonya’ya getiriyor.

Altıncı bin yılın ortalarına doğru, bu çiftçilerden bazıları doğuya, Sarı Deniz’in etrafından Kore’ ye ve Primorye’ye doğru kuzeydoğuya göç etmeye, bu bölgelere Korece ve Tunguz dillerini, Primorye’ye Batı Liao bölgesinden ilave Amur soylarını, Kore’ye ise karışık Amur-Sarı Nehir soylarını getirmeye başlamışlardır. Yeni analiz edilen Kore genomlarımız Japonya dışındaki Jomon ile ilgili soyların varlığına ve bunlarla karıştığına tanıklık etmesi açısından dikkate değerdir.

Geç Tunç Çağı boyunca, Avrasya bozkırında büyük bir kültürel değişim yaşandı ve bu da Batı Liao bölgesinden ve Doğu bozkırından gelen popülasyonların batı Avrasya genetik soylarıyla melezlenmesiyle sonuçlandı. Dilbilimsel olarak bu etkileşim, özellikle buğday ve arpa çiftçiliği, hayvancılık, mandracılık ve at kullanımı ile ilgili olarak, Ön-Moğolca ve Ön-Türkçe konuşanlar tarafından agropastoral terminolojinin ödünç alınmasında gösterilmektedir.

Liaodong-Şandong bölgesinden çiftçiler, darı tarımına pirinç, arpa ve buğday ekleyerek 3300 GÖ civarında Kore yarımadasına taşındılar. Bu göç, Kore’ den Bronz Çağı örneğimizde Yukarı Şiaciadian olarak modellenen genetik bileşenle uyumludur ve Japon ve Kore dilleri arasındaki erken ödünçlemelere yansır.Arkeolojik olarak, özellikle Yukarı Şiaciadian maddi kültürüyle sınırlı olmaksızın daha büyük Liaodong-Şandong bölgesindeki tarımla ilişkilendirilebilir.

Bu tarımsal paket, üçüncü milenyum GÖ’ de Kyuşu’ya taşındı ve tam ölçekli çiftçiliğe geçişi, Jomon’danYayoi kökenine genetik bir geçişi ve Japonca’ ya dilsel bir geçişi tetikledi.GüneyRyukyus’dakiNagabaka’dan benzersiz örnekler ekleyerek, Transavasya dünyasının sınırlarına kadar tarım/dil dağılımını izledik.Jomon atalarının MiyakoAdası kadar güneye uzandığını gösteren bulgularımız, Tayvan’dan kuzeye doğru genişleyen Avustralavasya gruplarının daha önceki fikirlerini çürütüyor. Bulgularımız, Kore, Yokçido’da keşfedilen Jomonprofili ile, Jomon genomlarının ve materyal kültürünün her zaman örtüşmediğini gösteriyor.

Araştırmamız, antik DNA’dan yeni kanıtlar geliştirerek, Japon ve Kore nüfuslarının Batı Liao Nehri atalarına sahip olduğuna dair son bulguları doğrularken, Transavasya dil ailesinin¹³ genetik bir karşılığı olmadığına dair önceki iddialarla çelişiyor.

Daha önce yapılan bazı araştırmaların Transavasya bölgesini tarım bölgesinin²⁰ dışında görmesine rağmen, bulgularımız çiftçilik/dil dağılımı hipotezinin Avrasya nüfus dağılımlarını²¹ anlamak için hala geçerli bir model olduğunu göstermektedir.

Dilbilimi, arkeoloji ve genetik üçgenlemesi, pastoralist ve tarımsal hipotezleri pastoralist ve tarımsal hipotezler arasındaki rekabeti çözüyor ve Transavasya dillerini konuşanların erken yayılmasının tarım tarafından yönlendirildiği sonucuna varıyor.

Çevrimiçi içerik

Herhangi bir yöntem, ek referanslar, Nature Research raporlama özetleri, kaynak veriler, genişletilmiş veriler, ek bilgiler, teşekkür, meslektaş incelemesi bilgileri; yazar katkılarının ayrıntıları ve rekabet eden çıkarlar ve veri ve kod kullanılabilirliği bildirimleri <https://doi.org/10.1038/s41586-021-04108-8> adresinde bulunabilir.

1. Starostin, S., Dybo, A. & Mudrak, O. *Etymological Dictionary of the Altaic Languages*Vol. I– III (Brill, 2003).

2. Blažek, V. *Altaic Languages. History of Research, Survey, Classification and a Sketch of Comparative Grammar* (Masaryk Univ. Press, 2019).
3. Robbeets, M. in *The Oxford Guide to the Transeurasian Languages* (eds Robbeets, M. & Savel'ev, A.) 772–783 (Oxford Univ. Press, 2020).
4. Mallory, J., Dybo, A. & Balanovsky, O. The impact of genetics research on archaeology and linguistics in Eurasia. *Russ. J. Genet.* 55, 1472–1487 (2019).
5. Bellwood, P. & Renfrew, C. (eds) *Examining the Farming/Language Dispersal Hypothesis* (McDonald Institute for Archaeological Research, 2002).
6. Menges, K. Dravidian and Altaic. *Anthropos* 72, 129–179 (1977).
7. Miller, R. A. Archaeological light on Japanese linguistic origins. *Asian Pac. Quart. Soc. Cult. Affairs* 22, 1–26 (1990).
8. Dybo, A. Language and archeology: some methodological problems. 1. Indo-European and Altaic landscapes. *J. Language Relationship* 9, 69–92 (2013).
9. Haak, W. et al. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* 522, 207–211 (2015).
10. Allentoft, M. et al. Population genomics of Bronze Age Eurasia. *Nature* 522, 167–172 (2015).
11. Damgaard, P. et al. The first horse herders and the impact of early Bronze Age steppe expansions into Asia. *Science* 360, eaar7711 (2018).
12. Ning, C. et al. Ancient genomes from northern China suggest links between subsistence changes and human migration. *Nat. Commun.* 11, 2700 (2020).
13. Wang, C. C. et al. Genomic insights into the formation of human populations in East Asia. *Nature* 591, 413–419 (2021).
14. Yang, M. A. et al. Ancient DNA indicates human population shifts and admixture in northern and southern China. *Science* 369, 282–288 (2020).
15. Francis-Ratte, A. & Unger, J. M. in *The Oxford Guide to the Transeurasian Languages* (eds Robbeets, M. & Savel'ev, A.) 705–714 (Oxford Univ. Press, 2020).
16. Anderson, G. in *The Oxford Guide to the Transeurasian Languages* (eds Robbeets, M. & Savel'ev, A.) 715–725 (Oxford Univ. Press, 2020).

17. Vajda, E. in *The Oxford Guide to the Transeurasian Languages* (eds Robbeets, M. & Savel'ev, A.) 726–734 (Oxford Univ. Press, 2020).
18. Robbeets, M. *Is Japanese related to Korean, Tungusic, Mongolic and Turkic?* (Harrassowitz, 2005).
19. Robbeets, M. *Diachrony of Verb Morphology: Japanese and the Transeurasian languages* (Vol. 291 in *Trends in Linguistics. Studies and Monographs*) (Mouton de Gruyter, 2015).
20. Heggarty, P. & Beresford-Jones, D. in *Encyclopedia of Global Archaeology* (ed. Smith, C.) 1–9 (Springer, 2014).
21. Bellwood, P. *First Farmers: The Origins of Agricultural Societies* (Blackwell, 2005).
22. Starostin, S. in *Past Human Migrations in East Asia: Matching Archaeology, Linguistics and Genetics* (eds Sanchez-Mazas, A. et al.) 254–262 (Routledge, 2008).
23. Ramstedt, G. J. A Comparison of the Altaic Languages with Japanese. *Trans. Asiatic Soc. Japan Second Ser.* 7, 41–54 (1924).
24. Kämpfer, E. *De Beschryving van Japan, benevens eene Beschryving van het Koningryk Siam* (Balthasar Lakeman, 1729).
25. Crawford, G. W. in *Handbook of East and Southeast Asian Archaeology* (eds Habu, J., Lape, P. V. & Olsen, J. W.) 421–435 (Springer, 2018).
26. Stevens, C. & Fuller, D. The spread of agriculture in eastern Asia: archaeological bases for hypothetical farmer/language dispersals. *Lang. Dyn. Chang.* 7, 152–186 (2017).
27. Leipe, C. et al. Discontinuous spread of millet agriculture in eastern Asia and prehistoric population dynamics. *Sci. Adv.* 5, eaax6225 (2019).
28. Stevens, C. et al. A model for the domestication of *Panicum miliaceum* (common, prosoor broomcorn millet) in China. *Veget. Hist. Archaeobot.* 30, 21–33 (2021).
29. Shelach-Lavi, G. et al. Sedentism and plant cultivation in northeast China emerged during affluent conditions. *PLoS ONE* 14, e0218751 (2019).
30. Lee, G. A. in *Handbook of East and Southeast Asian Archaeology* (eds Habu, J., Lape, P. & Olsen, J.) 451–481 (Springer, 2017).

31. Li, T. et al. Millet agriculture dispersed from Northeast China to the Russian Far East: integrating archaeology, genetics and linguistics. *Archaeol. Res. Asia* 22, 100177(2020).
32. Nelson, S. M. et al. Tracing population movements in ancient East Asia through the linguistics and archaeology of textile production. *Evol. Hum. Sci.* 2, e5 (2020).
33. Hudson, M. J. *Ruins of Identity: Ethnogenesis in the Japanese Islands* (Univ. Hawai'i Press, 1999).
34. Qin, L. & Fuller D. Q. in *Prehistoric Maritime Cultures and Seafaring* (eds Wu, C. & Rolett, B.) 159–191 (Springer, 2019).
35. Hosner, D. et al. Spatiotemporal distribution patterns of archaeological sites in China during the Neolithic and Bronze Age: an overview. *Holocene* 26, 1576–1593 (2016).
36. Hudson, M. J. & Robbeets, M. Archaeolinguistic evidence for the farming/language dispersal of Koreanic. *Evol. Hum. Sci.* 2, e52 (2020).
37. Jeong, C. et al. A dynamic 6,000-year genetic history of Eurasia's Eastern Steppe. *Cell* 183, 890–904 (2020).
38. Savelyev, A. & Jeong, C. Early nomads of the Eastern Steppe and their tentative connections in the West. *Evol. Human Sci.* 2, e20 (2020).
39. Janhunen, J. in *The Mongolic languages* (ed. Janhunen, J.) 1–29 (Routledge, 2003).
40. Hudson, M. J. in *New Perspectives in Southeast Asian and Pacific Prehistory* (eds Piper, P., H. Matsumura, H. & Bulbeck, D.) 189–199 (ANU Press, 2017).
41. Sagart, L. et al. Dated language phylogenies shed light on the ancestry of Sino-Tibetan. *Proc. Natl Acad. Sci. USA* 116, 10317–10322 (2019).
42. Zhang, H. et al. Dated phylogeny suggests early Neolithic origin of Sino-Tibetan languages. *Sci. Rep.* 10, 20792 (2020).

Yayıncının notu Springer Nature, yayımlanan haritalardaki ve kurumsal bağlantılardaki yargı iddiaları konusunda tarafsız kalır.

Açık erişim Bu makale, orijinal yazar(lar)a ve kaynağa uygun bir şekilde atıfta bulunduğunuz ve bir Creative Commons lisansına bağlantı verin ve değişiklik yapıp yapılmadığımı belirtin. Bu makaledeki görseller veya diğer üçüncü şahıs materyalleri, materyalin kredi limitinde aksi

belirtilmedikçe, makalenin Creative Commons lisansına dahildir. Materyal, makalenin Creative Commons lisansında yer almıyorsa ve kullanım amacınıza yasal düzenleme izin vermiyorsa veya izin verilen kullanımı aşarsa, doğrudan telif hakkı sahibinden izin almanız gerekir. Bu lisansın bir kopyasını görüntülemek için <http://creativecommons.org/licenses/by/4.0/> adresini ziyaret edin.

© Yazar(lar) 2021

Metotlar

Dilbilim

Bayesfilogenetiği. Bayes çıkarımını kullanan bir filogenetik türüdür, sözlük arama ve saha çalışmasının bir kombinasyonunu kullanarak, çağdaş ve tarihi çeşitler de dahil olmak üzere 98 Transavasya dili için 254 temel kelime kavramını temsil eden 3.193 veri noktasından oluşan karşılaştırmalı bir veri kümesi derledik (Ek Veri 1). Bu kavramlar, Leipzig-Jakarta'nın birleşmesine dayanmaktadır. 200 (ref. ⁴³) ve Jena 200 (ref. ⁴⁴) listeleri (Ek Veri 2). Dahil edilen Türkçe ve Tunguzca için temel sözcük dağarcığı, yakın zamanda yayımlanan veri kümelerinin ^{45,46} yeniden gözden geçirilmesine dayanmaktadır. Aynı kökten kodlama, Ek Veri 2'de sunulan Transavasya dilleri arasında temel kelime etimolojileri ve sağlam yazışmaların bir envanteri ile desteklenir.

İkili veri⁴⁷ olarak kodlanmış aynı kökenliler ile bir Bayesfilogenetik analizi gerçekleştirdik. Veriler en az bir soydaş bulunabilecek şekilde toplandığından, tamamı sıfır olan hiçbir yerleşim yerinin bulunmadığı belirlendi. Bunu karşılamak için tespit düzeltmesi uygulandı⁴⁷.

Ağaç dalları boyunca aynı kökenlilerin evrimsel sürecini yöneten aşağıdaki ikame modellerini düşündük: sabit bir mutasyon oranı varsayan sürekli zamanlı Markov zinciri (CTMC); bu iki durum arasında geçiş yapan model ile yavaş ve hızlı bir mutasyon oranı varsayan kovaryon; ve bir soydaşın yalnızca bir kez ortaya çıkabileceği, ancak birçok kez kaybolabileceği gibi Dollo yasasına dayanan sözde Dollo kovaryon modeli. Ayrıntılı açıklamaları CTMC ve kovaryon modelleri⁴⁷ ve sözde Dollokovaryon modeli⁴⁸ literatürde mevcuttur. Farklı kelimelerin evrim oranları arasındaki varyasyonu yakalamak için, tüm modellerde her anlam sınıfının kendi görelî hızına sahip olduğunu varsayıyoruz.

Dil ortalama olarak sabit bir hızla gelişse de oranların bir ağaçtaki dallar arasında önemli ölçüde değişebileceğini görüyoruz^{47,48}. Bu tür bir varyasyon⁴⁹, oranların log-normal olarak dağıldığını varsayarsak ilişkisiz gevşetme kullanılarak yakalanabilir.

Bir doğum ölüm modeli dil oluşturma işleminin üretimsel sürecini tanımlamak için kullanılır. Veriler, geçerli dillerin atası olabilecek eski dilleri içerdiğinden, ağaçtaki atalara ait düğümlere izin veririz. Bu tür atalara izin veren fosilleşmiş bir doğum ölüm modeli⁵⁰ düğümler, ağaçta önceki gibi kullanılır. Dil ailesi düğüm yaşları yaş geçmişlerine göre bilgilendirildi (Japon 2100 GÖ ± 175, Kore 800 GÖ ± 175, Türk 2100 GÖ ± 175, Moğol 750 GÖ ± 50, Tunguz 1900 GÖ ± 275). Bu kalibrasyonlar önerilen kronolojik tahminlerle dil bilimsel literatürde desteklenir. Bu düğüm yaşı önceliklerinin, kök yaş dağılımındaki belirsizliği biraz azaltmaya yardımcı olduğunu bulduk.

İç içe örnekleme⁵¹ (Ek Veri 18) kullanarak marjinal olasılıkları tahmin ederek farklı modellerin uyumunu karşılaştırdık ve rahat saatli sözde Dollokovaryon modelinin en iyi uyuma sahip olduğu ve rahat saatli kovaryonun bir sonraki en uygun olduğu sonucuna vardık. Her iki model de uyumlu zaman tahminleri üretir, ancak kovaryon tahminleri daha büyük belirsizliğe sahip olma eğilimindedir (yani, daha büyük% 95 HPD aralıklarına sahiptir). CTMC modelinin rahat saatli zaman tahminleri hala uyumludur, ancak daha da geniştir ve daha yüksek bir ortalamaya sahip olma eğilimindedir.

Tüm önceki tahminler, uyarlanabilir bağlı Markov zinciri Monte Carlo (MCMC)⁵³ kullanılarak BEAST v.2.6⁵² kullanılarak yapıldı. Çalıştırmak için kullanılan modellerin, önceliklerin, hiper önceliklerin ve ayarların ayrıntılı özellikleri bu modeller BEAST XML dosyalarında bulunabilir (Ek Veriler 19). Bayes analizimizin sonuçları, Transavasya dillerinin tarihli bir filogenetik ağacı olarak görselleştirilmiştir (Ek Veriler 24).

Bayesfilocoğrafyası. İnsanların Avrasya'ya dağılmasının rastgele bir yürüyüş olarak tanımlanabileceğini ve bunu yakalamanın en iyi yolunun bir küre⁵⁴ üzerindeki yayılmanın olduğunu varsaydık. Böyle bir modelle kökenleri bulmadaki belirsizlik hakkında bir izlenim elde etmek için, sözcüksel analizden arka ağaç kümesini kullanarak post hoc bir analiz yaptık. MCMC aracılığıyla coğrafi parametreleri tahmin ederken uçlara nokta konumları ve arkadan rastgele örneklenmiş ağaçlar atadık. Bu nispeten kısıtlı kurulumda bile, kök konumundaki belirsizlik farklı coğrafi köken hipotezlerini ayırt etmemize izin verir. Analizimizin sonuçları bir harita üzerinde sunulmuştur (Ek Veri 3). Bayesfilocoğrafyasının birtakım sınırlamalarla mücadele etmesi gerektiğinden^{55,56}, Transavasya ailesinin kök ve düğümlerinin anavatanları için dengeli bir konuma ulaşmak için dilsel paleontoloji ve çeşitlilik hotspot ilkesi gibi diğer anavatan tespit yöntemleriyle tamamladık (Ek Veriler 4).

Dilbilimsel paleontoloji. Her Transavasya alt ailesi için karşılaştırmalı agropastoral sözlükler derledik: Türkçe (Ek Veri 5a), Moğolca (Ek Veri 5b), Tunguzca (Ek Veri 5c),

Korece (Ek Veri 5d) ve Japonca (Ek Veri 5e). Bir dilin sahipsiz bir atadan kalma durumunu, daha sonraki bir dönemden elde edilebilecek verilerin kanıtı üzerine, karşılık gelen kelimelere çıkarmak için bir prosedür olan dilsel yeniden yapılandırmayı uyguladık (Ek Veriler 5).

Devralınan ve ödünç alınan yazışmaları ayırt etme kümeler, sesbilim, anlambilime dayalı standart ölçütler kullandık, ek Veriler 5'te belirtildiği gibi, söz konusu kelimenin morfolojisi ve dağılımı. Veri setimizi kalıtsal ve ödünç alınan geçimlik kelime dağarcığına bölerek, her kategori için kendine özgü mekânsal-zamansal ve kültürel kalıpları belirledik (Ek Veriler 5).

Dilbilimsel yeniden yapılanmalarımızı arkeolojiden bu kelimeleri kullanan eski konuşma topluluklarının kültürü hakkındaki bilgilerle ilişkilendirerek insan tarih öncesini incelememizi sağlayan tarihsel karşılaştırmalı bir yöntem olan geçim sözlüğümüze dilbilimsel paleontoloji uyguladık. Bu şekilde, Neolitik ve Tunç Çağı'ndaki farklı Transavasya ön-dillerinin konuşmacıları için mevcut geçim stratejileri hakkında çıkarımlar yaptık (Ek Veriler 5) ve ilgili eski konuşma topluluklarının anavatanı için makul bir yer belirledik (Ek Veriler 4).

Çeşitlilik etkin nokta prensibi. Katılan eski konuşma topluluklarının yerini tahmin etmek için, Bayesfilocoğrafyası ve dilsel paleontolojiyi çeşitlilik etkin nokta prensibi ile birleştirdik. İlke, dil ailesinin anavatanının en büyük çeşitliliğe ve en derin alt gruplarına göre en yakın olduğu varsayımına dayanmaktadır. Bu alanları harita üzerinde konumlandırdık ve belirli bir ön-dilin çeşitlenmeye başladığı alanın bir yaklaşımı olarak aldık (Ek Veriler 4). Bu yöntemin bazı sınırlamaları olmasına rağmen (Ek Veri 4), burada tartışılan anavatan konumu için diğer tekniklerle birleştirildiğinde, bize eski bir dil topluluğunun konumu hakkında makul derecede doğru bir tahmin sağlayabilir.

Arkeoloji

Arkeolojik veritabanı. Batı Liao nehri havzası (36), Amur (Cilin, Heilongciang ve iç Liaoning) (32), Primorye (4), Liaodong yarımadası (37), doğu bozkırlarından 255 Neolitik–Tunç Çağı arkeolojik alanı veya evresi için 172 kültürel özellik puanladık. (1), Şandong yarımadası (4), Sarı Nehir havzası (2), Kore yarımadası (58) ve Japon adaları (85).

Birkaç önemli kültürel aşamaya sahip yerleşim yerleri ayrı ayrı puanlandı. Yerler 8400–1700 GÖ'den kalmaz ve Erken Neolitik'ten Bronz Çağı'na kadar olan dönemleri içerir. Kuzeydoğu Çin'de yaş, Orta Neolitik Çağ'da Zaisanovka kültürü Primorye, Orta-Geç Neolitik Çulmun ve Tunç Çağı Mumun Kore'deki kültürler ve Geç Neolitik-Bronz Çağı Son Jomon ve Batı Japonya'daki Yayoi kültürleri. Puanlanan kültürel özellik kategorileri şunları

içeriyordu seramik (70), taş aletler (38), binalar (9), bitki ve hayvan kalıntıları (26), kabuk ve kemik eserleri (17) ve mezarlar (12). Puanlanan özelliklerin tanımları Ek Veri 6'da (sayfa 2) bulunur ve puanlama yöntemlerinin daha fazla tartışılması Ek Veri 7'de bulunabilir. Yayımlanan site raporları veya diğer literatürlerin ardından tüm özellikler mevcut (1) veya yok (0) olarak puanlandı.

Veri tabanı, dağıtımdaki değişiklikleri analiz etmek için kullanıldı. Özellikle Neolitik ve Tunç Çağı eserleri ile ilgili olarak zaman içinde kuzeydoğu Asya'da tarım sistemlerinin yayılması (Ek Veri 7).

Ayrıca arkeolojik veri tabanımızdaki kültürel veriler Bayesfilogenetik yöntemleri kullanılarak analiz edilmiştir. Arkeolojik verilerle büyük miktarda filogenetik çalışma var⁵⁷, bazıları parsimoniye-dayalı⁵⁸, diğerleri mesafeye-dayalı⁵⁹. Bayesci yaklaşımların yararı, model tabanlı olmaları, olasılık teorisinde tüm tahminler etrafındaki belirsizliği tahmin etmemize izin veren sağlam resmi matematiksel temellere sahip olmaları ve çeşitli kaynaklardan (eş kökenli ve coğrafi veriler gibi) gelen verilerin olasılık teorisine dayalı tek bir yerde analizde entegrasyonuna izin vermeleridir. BEAST, özellikle köklü zaman ağaçlarını ve zaman tahminlerinin belirsizliğini çıkarmayı amaçlamaktadır, bu da onu köklenmemiş ağaçları hedef alan diğer Bayes paketlerinden ayırmaktadır. Ayrıca, BEAST şu anda diğer paketlerde bulunmayan modelleri, dolayısıyla bu paketin kullanımını desteklemektedir.

Kültürel veriler ikili bir hizalama olarak kodlanmıştır ve sözcüksel verilerle aynı ikame ve saat modellerini uyguladık. Rahat saatli sözde Dollo modeli verilere en iyi şekilde uyar (Ek Veriler 20). Gevşemiş saatin varyasyon katsayısı 1'i aştığından, bu da önemli miktarda varyasyona işaret ettiği için, analizi 1'de sınırlandırılmış standart sapma ile de çalıştırdık, bu da zaman tahminlerini biraz etkiledi.

Çok sayıda örnekleme tarihi ve eksik kültürlerin sayısındaki belirsizlik, fosilleşmiş doğum ölümünün önceden uygulanmasını zorlaştırdı, bu nedenle bunun yerine⁶⁰esnek Bayesci silüet grafiğini seçtik. Zamanlama bilgileri arkeolojik buluntuların örnekleme tarihlerine dayanmaktadır. Bu bulguların tarihlendirilmesinde belirsizlik olduğundan, MCMC sırasında bu aralıklarda ipucu tarihleri tek tip olarak örneklenmiştir. Önceki arkeolojik çalışmalara⁶¹⁻⁶³ uygun olarak, 'Şinglongva-Jabaogou-Hongşan' ve 'Yabuli-Primorye' dallarını monofiletik olacak şekilde sınırladık (Ek Veri 8). Tüm analizler, BEAST v.2.6'da⁵² uyarlamalı eşleştirilmiş MCMC⁵³ kullanılarak gerçekleştirilmiştir. Modeller, öncelikler, hiperönceler ve ayarlarla ilgili ayrıntılar BEAST XML' de bulunabilir (Ek Veri 21). Bayes analizimizin

sonuçları, Kuzeydoğu Asya'daki (Ek Veri 25) arkeolojik kültürlerin filogenetik bir ağacı olarak görselleştirilir ve Ek Veri 8' de yorumlanır.

Arkeobotanik veri tabanı. Arkeolojik özellikler veri tabanına ek olarak, Kuzeydoğu Asya'nın her bölgesinden doğrudan radyokarbon ile tarihlenen en eski mahsul kalıntılarının bir listesini derledik (Ek Veri 9). Bu liste 269 örnek içermektedir (Çin, 82; Primorye, 12; Kore, 31; Japonya (Ryukyus hariç), 120; Ryukyu Adaları, 24). Bu veri tabanındaki radyokarbon tarihleri, OxCal v.4.4 kullanılarak yeniden kalibre edildi. Bu veri tabanındaki tahılların yayılmasını zaman içinde planlamak için çekirdek yoğunluğu haritalamasını kullandık Ek Veri 7). Veri tabanlarımız, fauna kalıntıları^{64,65}, dolmenler⁶⁶ ve ağırşaklar⁶⁷ için yayınlanmış veri kümeleri ile desteklenmiştir.

Genetik

Laboratuvar prosedürleri DNA özütleme ve kütüphane hazırlığı da dahil olmak üzere eski DNA ıslak laboratuvar çalışması, Max Planck İnsanlık Tarihi Bilim Enstitüsündeki (MPI-SHH) özel bir antik DNA temiz oda tesisinde ve Jilin Üniversitesi'ndeki eski bir DNA laboratuvarında yerleşik protokolleri izleyerek⁶⁸ gerçekleştirildi. Hem P5 hem de P7 Illumina adaptörlerinde 8-mer indeks dizileriyle çift sarmallı bir kitaplık oluşturuldu. Jilin'de karakterize edilen Çin'den dört kişi, yeterli bir kapsama alanı elde etmek için 150-bp çift uçlu dizileme tasarımında Illumina HiSeq X10 cihazında doğrudan çifte-diziledi. Japonya ve Kore'den 33 kişi için seksen üç çift telli kütüphane çifte-dizilimi veya yaklaşık 1,2 milyon bilgilendirici nükleer tek nükleotid polimorfizminde çözünme yakalama yoluyla (SNP'ler) MPI-SHH'de üretildi ve karakterize edildi. Bu kitaplıkların korunmasının ilk taramasından sonra, örneklerden daha fazla endojen DNA elde etmeyi amaçlayan 108 tek zincirli kitaplık daha inşa edildi ve yine bu kitaplıklar doğrudan çifte sıralandı ve yaklaşık 1,2 milyon SNP'de çözelti içinde tutuldu (Ek Veri 17) ve üreticinin protokollerini izleyerek Illumina HiSeq 4000 platformunda sıralanmıştır.

Dizi veri işleme. Ham sıralama okumaları, EAGER v.1.92.55 programı⁶⁹ ile otomatik bir iş akışı tarafından işlendi. Illumina adaptör dizileri, dizileme verilerinden kırpıldı ve örtüşen çiftler AdapterRemoval v.2.2.0⁷⁰ ile birleştirildi. BWA v.0.7.12⁷¹ kullanarak birleştirilmiş okumaları minimum 30 bp ile insan referans genomuna (hs37d5; GRCh37 tuzak dizili) eşledik. PCR kopyalarını DeDup v.0.12.2⁶⁰ ile kaldırdık. Otopsi DNA hasarının genotipleme

üzerindeki etkisini en aza indirmek için bamUtils v.1.0.13'de⁷² trimbam işlevini kullanarak okuma başına UDG olmayan kitaplıklar için 2 bp ve yarım UDG kitaplıkları için 10 bp maskeledik. Hem temel kalite (Phred ölçeği kalitesi) hem de haritalama kalitesi (Phred ölçeği eşleme kalitesi) ile 30'un üzerinde temizlenmiş okumalar, mpileup işleviyle SAMtools 1.3⁶⁰ tarafından biriktirildi. Rastgele haploid çağrı modu altında '1240k' panelinde^{73,74} SNP'lere karşı pileupCaller programını (<https://github.com/stschiff/sequenceTools>) kullanarak sözde diploid genotipleri aradık. C/T ve G/A SNP'leri için maskelenmiş BAM dosyalarını kullandık; geri kalanı için orijinal maskesiz BAM dosyalarını kullandık.

Referans veri kümeleri. Antik bireylerimizi, biri Affymetrix İnsan Kökenleri Aksiyom Genom Çapında İnsan Kökenleri 1 dizisine ('HumanOrigins'; 593,124 otozomal SNP'ler)⁷⁵, '1240k' panel⁷³ ve 'Illumina'ya dayanan dünya çapındaki üç genotip paneli 'veri seti'⁷⁶ ile karşılaştırdık. Simons Genom Çeştliliği Panelini⁷⁷ ekleyerek bu veri kümelerini genişlettik ve eski genomları yayınladık (Ek Veri 11).

Antik DNA kimlik doğrulaması. Kore ve Japonya'dan yeni yayınlanan antik genomların geçerliliğini doğrulamak için birden fazla kriter uyguladık. İlk olarak, mapDamage v.2.0.6⁷⁸'i kullanarak antik DNA için karakteristik olan ölüm sonrası kimyasal modifikasyonları karakterize ettik. İkinci olarak, Schmutzi v.1.5.1⁷⁹ kullanan tüm bireyler için mitokondriyal kirlenme oranlarını tahmin ettik. Üçüncü olarak, ANGSD v.0.910⁸⁰'de uygulanan X kromozomu verileri temelinde erkeklerde nükleer genom bulaşma oranını ölçtük. Erkeklerde X kromozomunun yalnızca tek bir kopyası bulunduğundan, dizileme hatası seviyesinin ötesinde, aynı polimorfik konuma hizalanmış bazlar arasındaki uyumsuzluklar bulaşma kanıtı olarak kabul edilir. Dördüncüsü, PMDtools⁸¹'de PMD puanı en az 3 olan tek zincirli kitaplıklarda mevcut tüm okumalar ve hasar kısıtlanmalı okumalarla potansiyel Batı Avrasya kontaminasyonunu değerlendirdik ve bir Avrasya PCA'daki konumlarını yalnızca tüm okumalar ve hasarlı okumalarla karşılaştırdık. Beşinci olarak, Batı Avrasya kirlenmesini kaynak olarak Sintashta_MLBA veya LBK_EN gibi Batı Avrasya karakteristik gruplarıyla daha fazla karakterize etmek için kişi başına qpAdm⁷⁴ uyguladık (ayrıntılar için bkz. Ek Veri 17, 22).

Nüfus yapısı analizi. 'HumanOrigins' veri kümesinden ve 'Isqproject: YES' ve 'shrinkmode: YES' seçeneğiyle '1240kIllumina' veri kümesinden 2.077 günümüz Avrasya bireyi kümesini kullanarak smartpca v.1.6000⁸² ile bir PCA gerçekleştirdik. Afrika dış grubundan ayrılmalarından bu yana iki popülasyon arasındaki genetik yakınlığın bir ölçümünü elde

etmek için dış grup- f_3 istatistiklerini^{83,84} kullandık. Admixtools31'deki 'f4mode: YES' işleviyle f_4 istatistiklerini hesapladık. Hem f_3 hem de f_4 istatistikleri, admixtools paketinde qp3Pop v.435 ve qpDstat v.755 kullanılarak hesaplandı.

Genetik cinsiyet ayrımı ve tek ebeveynli haplogrup ataması. Eski örneklerimizin moleküler cinsiyetini, X ve Y kromozomu kapsamalarının otozomlara oranını karşılaştırarak belirledik⁸⁵. Kadınlar için, X' inotozom kapsamına yaklaşık olarak eşit bir oranı ve 0'lık bir Y oranı bekleriz. Erkekler için, otozomlara kıyasla X ve Y'deki kapsamın kabaca yarısını bekleriz.

qpAdm ile karışım modelleme. Admixtools v.5.1 paketindeki⁷⁴ qpWave/qpAdm çerçevesini (qpWave v.410 ve qpAdm v.810) kullanarak bu çalışmada eski bireyleri modelledik. Aşağıdaki 7 popülasyonu '1240k' veri setlerinde dış grup ('OG') olarak kullandık: Mbuti, Onge, Iran_N, Villabruna, Karitiana, Naxi ve FunadomariJomon. Bu set bir Afrika dış grubu (Mbuti), Andamanese adalıları (Onge), Tepe GanjDareh bölgesinden (Iran_N) erken Neolitik İraniiler, geç Pleistosen Avrupa avcı-toplayıcıları (Villabruna), Brezilya'dan yerli Karitiana, Tibet-Burman konuşan bir grup içerir. Güney Çin'den (Naxi) ve Japonya'dan eski avcı-toplayıcılardan (FunadomariJomon) (Ek Veri 13, 16).

Üçgenleme

"Üçgenleme" terimi, diğer iki farklı noktadan alınan ölçümlerin yakınsaması ile uzayda tek bir noktayı belirleyen bir seyir tekniğinden ödünç alınmıştır. Nitel araştırmada, üç farklı bilimsel disiplinden gelen kanıtları kullanarak aynı olgunun farklı boyutlarını yakalamak için kullanılan bir yöntemi belirtir. Tartışmada döngüsellikten kaçınmak için, veri toplama, analizler ve sonuçlar, diğer ikisinden bağımsız olarak her bir disiplinin sınırları içinde gerçekleştirilir veya bunlara ulaşılır. Üçgenleştirme sürecinin yalnızca son aşamasında, fenomeni tanımlayan bir dizi değişken karşılaştırılarak üç disiplinin yaptığı çıkarımlar birbiri üzerine haritalanır. Üçgenleştirmenin amacı, üç disiplinden elde edilen kanıtların ne ölçüde birleştiğini değerlendirerek ve araştırılan fenomenler üzerindeki farklı bakış açıları arasındaki korelasyonları, tutarsızlıkları, belirsizlikleri ve olası önyargıları belirleyerek sonuçların güvenilirliğini ve geçerliliğini artırmaktır.

Antropolojideki daha önceki üçgenleme uygulamalarına dayanarak⁸⁶, fenomenin daha iyi anlaşılmasına katkıda bulunmak için dilbilimi, arkeolojiyi ve genetiği entegre ederek yöntemi Transavasya dillerinin dağılımına uyguladık. Farklı veri kümeleri topladık ve konum, kronoloji, göç dinamikleri, sürekliliğe karşı yayılma ve geçim gibi bir dizi değişkenle ilgili

bağımsız çıkarımlar yapmak için yukarıda açıklanan yöntemleri uyguladık (Ek Veri 26). Her disiplin, doğrudan veya dolaylı kanıtlara dayalı olarak, niteliksel veya niceliksel olsun, kendi alanına dahil olan araçların uygulanması temelinde bu değişkenleri içeren en cimri modeli çıkarmıştır. Tek başına ele alındığında, tek bir disiplin tek başına çiftçilik/dil dağılımları hakkındaki soruyu kesin olarak çözemez, ancak üç disiplin birlikte alındığında bu senaryonun güvenilirliğini ve geçerliliğini artırır. Üç disiplin tarafından sunulan kanıtları hizalayarak, Transavryasya göçü hakkında, üç disiplinin her birinin bize tek tek sağlayabileceğinden daha dengeli ve daha zengin bir anlayış kazandık.

Raporlama özeti

Araştırma tasarımı hakkında daha fazla bilgi, bu makaleyle bağlantılı Doğa Araştırma Raporlama Özetinde mevcuttur.

Veri kullanılabilirliği

Dilbilimsel ve arkeolojik veri kümelerine Ek Bilgiler aracılığıyla erişilebilir. Kullanım gerektiren dosyalar FigShare'ye yüklendi. FigShare'e bağlantılar aşağıdaki gibidir: Ek Veri 3: Transavryasya dillerinin uzaysal-zamansal genişlemesini modelleyen Bayesfilocoğrafik analizi (<https://figshare.com/s/b9c67ca3ea47faf51d48>); Ek Veri 19: Dilsel veri tabanının analizlerini çalıştırmak için kullanılan modelleri, öncelikleri, hiperönceleri ve ayarları belirten BEAST XML dosyaları (<https://figshare.com/s/748bf751fe3ba7752046>); Ek Veri 21: Arkeolojik veri tabanının analizlerini çalıştırmak için kullanılan modelleri, öncelikleri, hiperönceleri ve ayarları belirten BEAST XML dosyaları (<https://figshare.com/s/99f5aab9a2e43eb2ffd4>); Ek Veri 24: Transavryasya dillerinin tarihli Bayesfilogenisi (<https://figshare.com/s/709f239fa45982911b87>); ve Ek Veri 25: Arkeolojik veri tabanının Bayesfilogenetik analizi (<https://figshare.com/s/65615dddc0817bc0184f>). İncir ağacı uygulamasının bağlantısı: <https://github.com/rambaut/figtree/releases/tag/v1.4.3> Genetik veri kümelerimiz için, bu yazıda bildirilen DNA dizileri Avrupa Nükleotid Arşivinde (ENA) depolanmıştır. Katılım PRJEB46162 altında. '1240k' panelindeki bu çalışmadaki antik bireylerin haploidgenotip verileri aşağıdaki bağlantıdan EIGENSTRAT formatında mevcuttur: <https://edmond.mpdl.mpg.de/imeji/collection/59JGAaOpSxRb96Vh>

Kod kullanılabilirliği

Okuyucular, Ek Bilgiler aracılığıyla Bayesian dilsel ve kültürel veri kümeleri analizlerimizin altında yatan koda erişebilir. Ek Veriler 19'daki dosyalar dillerle ve Ek Veriler 21'deki

dosyalar kültürlerle ilgilidir. Web bağlantıları şunlardır: Ek Veri 19: Dilsel veri tabanının analizlerini çalıştırmak için kullanılan modelleri, öncelikleri, hiperönceleri ve ayarları belirten BEAST XML dosyaları (<https://figshare.com/s/748bf751fe3ba7752046>); Ek Veri 21: Arkeolojik veri tabanının analizlerini çalıştırmak için kullanılan modelleri, öncelikleri, hiperönceleri ve ayarları belirten BEASTXML dosyaları (<https://figshare.com/s/99f5aab9a2e43eb2ffd4>).

43. Haspelmath, M. & Tadmor, U. *Loanwords in the World's Languages: a Comparative Handbook* (Mouton de Gruyter, 2009).
44. Heggarty, P. & Anderson, C. *Cognacy in Basic Lexicon (CoBL)*, <https://www.shh.mpg.de/dlce-research-projects/ie-cor-database> (Max Planck Institute for the Science of Human History, 2015).
45. Savelyev, A. & Robbeets, M. Bayesian phylolinguistics infers the internal structure and the time-depth of the Turkic language family. *J. Lang. Evol.* 39–53 (2019).
46. Oskolskaya, S., Koile, E. & Robbeets, M. A Bayesian approach to the classification of Tungusic languages. *Diachronica* <https://doi.org/10.1075/dia.20010.osk> (2021).
47. Bouckaert, R., Bowern, C. & Atkinson, Q. D. The origin and expansion of Pama–Nyungan languages across Australia. *Nat. Ecol. Evol.* 2, 741–749 (2018).
48. Bouckaert, R. & Robbeets, M. Pseudo Dollo models for the evolution of binary characters along a tree. Preprint at <https://doi.org/10.1101/207571> (2018).
49. Drummond, A. J. et al. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4, e88 (2006).
50. Gavryushkina, A. et al. Bayesian inference of sampled ancestor trees for epidemiology and fossil calibration. *PLoS Comput. Biol.* 10, e1003919 (2014).
51. Maturana, P. M. et al. Model selection and parameter inference in phylogenetics using nested sampling. *Syst. Biol.* 68, 219–233 (2019).
52. Bouckaert, R. et al. BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.*, 15, e1006650 (2019).
53. Mueller, N. F. & Bouckaert, R. Adaptive parallel tempering for BEAST 2. Preprint at <https://doi.org/10.1101/603514> (2020).

54. Bouckaert, R. Phylogeography by diffusion on a sphere: whole world phylogeography. *PeerJ*, 4, e2406 (2016).
55. Wichmann, S. & Rama, T. Testing methods of linguistic homeland detection using synthetic data. Preprint at <https://doi.org/10.1101/2020.09.03.280826> (2020).
56. Neureiter, N., Ranacher, P., van Gijn, R., Bickel, B. & Weibel, R. 2021 Can Bayesian phylogeography reconstruct migrations and expansions in linguistic evolution? *R. Soc. Open Sci.* 8, 201079 (2021).
57. Mace, R., Holden, C. & Shennan, S. *The Evolution of Cultural Diversity—a Phylogenetic Approach* (UCL Press, 2005).
58. O'Brien, M. J. & Lyman, R. L. Evolutionary archeology: current status and future prospects. *Evol. Anthropol.* 11, 26–36 (2002).
59. Allaby, R. G., Fuller, D. Q. & Brown, T. A. The genetic expectations of a protracted model for the origins of domesticated crops. *Proc. Natl Acad. Sci. USA* 105, 13982–13986 (2008).
60. Drummond, A. J. et al. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* 22, 1185–1192 (2005).
61. Shelach, G. & Teng, M. in *A Companion to Chinese Archaeology* (ed. Underhill, A.) 37–54 (Wiley–Blackwell, 2013).
62. Miyamoto, K. The initial spread of early agriculture into Northeast Asia. *Asian Archaeol.* 3, 11–26 (2014).
63. Li, T., Ning, C., Zhushchikhovskaya, I. S., Hudson, M. J. & Robbeets, M. Millet agriculture dispersed from Northeast China to the Russian Far East: integrating archaeology, genetics and linguistics. *Archaeol. Res. Asia* 22, e100177 (2020).
64. Kōmoto, M. in *A Study on the Environmental Change and Adaptation System in Prehistoric Northeast Asia* (ed. Kōmoto, M.) 8–34 (Kumamoto Univ., 2007).
65. An, S. (ed.) *Nongōbūi kogohak* (Sahoep'yōngnon, 2013).
66. Nishitani, T. (ed.) *Higashi Ajia ni okeru shisekibo no sōgōteki kenkyū* (Kyushu Univ., 1997).
67. Furusawa, Y. in *A Study on the Environmental Change and Adaptation System in Prehistoric Northeast Asia* (ed. Kōmoto, M.) 86–109 (Kumamoto Univ., 2007).

68. Dabney, J. et al. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl Acad. Sci. USA* 110, 15758–15763 (2013).
69. Peltzer, A., Herbig, A. & Krause, J. EAGER: efficient ancient genome reconstruction. *Genome Biol.* 17, 60 (2016).
70. Schubert, M., Lindgreen, S. & Orlando, L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res. Notes* 9, 88 (2016).
71. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079 (2009).
72. Jun, G. et al. An efficient and scalable analysis framework for variant extraction and refinement from population-scale DNA sequence data. *Genome Res.* 25, 918–925 (2015).
73. Mathieson, I. et al. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* 528, 499–503 (2015).
74. Haak, W. et al. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* 522, 207–211 (2015).
75. Jeong, C. et al. The genetic history of admixture across inner Eurasia. *Nat. Ecol. Evol.* 3, 966–976 (2019).
76. Jeong, C. et al. Bronze Age population dynamics and the rise of dairy pastoralism on the eastern Eurasian steppe. *Proc. Natl Acad. Sci. USA* 115, E11248–E11255 (2018).
77. Mallick, S. et al. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* 538, 201–206 (2016).
78. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F. & Orlando, L. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* 29, 1682–1684 (2013).
79. Renaud, G., Slon, V., Duggan, A. T. & Kelso, J. Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biol.* 16, 224 (2015).
80. Korneliussen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics* 15, 356 (2014).

81. Skoglund, P. et al. Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. *Proc. Natl Acad. Sci. USA* 111, 2229–2234 (2014).
82. Patterson, N., Price, A. L. & Reich, D. Population structure and eigen analysis. *PLoS Genet.* 2, e190 (2006).
83. Raghavan, M. et al. Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* 505, 87–91 (2014).
84. Patterson, N. et al. Ancient admixture in human history. *Genetics* 192, 1065–1093 (2012).
85. Fu, Q. et al. An early modern human from Romania with a recent Neanderthal ancestor. *Nature* 524, 216–219 (2015).
86. Kirch, P. V. & Green, R. *Hawaiki, Ancestral Polynesia: An Essay in Historical Anthropology* (Cambridge Univ. Press, 2001).
87. Oh, Y., Conte, M., Kang, S., Kim, J. & Hwang, J. Population fluctuation and the adoption of food production in prehistoric Korea: using radiocarbon dates as a proxy for population change. *Radiocarbon* 59, 1761–1770 (2017).
88. Hosner, D., Wagner, M., Tarasov, P. E., Chen, X. & Leipe, C. Spatiotemporal distribution patterns of archaeological sites in China during the Neolithic and Bronze Age: an overview. *Holocene* 26, 1576–1593 (2016).
89. Koyama, S. Jomon subsistence and population. *SENRI Ethnol. Stud.* 2, 1–65 (1978).

Teşekkürler Bu sonuçlara ulaştıran araştırma M.R'nin izniyle Avrupa Birliği'nin Ufku 2020 araştırma ve yenilik programı (646612 numaralı hibe sözleşmesi) altında Avrupa Araştırma Birliği tarafından finanse edilmiştir. R.B. Yeni Zelanda Kraliyet Cemiyeti'nden bir Marsden hibesi 18-UOA-096 tarafından desteklendi. N. Adachi, T. Kakuda, E. Savelyeva, W. Lawrence, S. Wichmann, C. Wang, M. Burri, N. Klyuev, I. Zhushchikhovskaya, M. Byington, H. Miyagi, Y. Vostretsov, A. Jarosz, J.-O. Svantesson, M. Levy, J. Lefort, M. Miller, K. Mishchenkova, E. Perekhvalskaya, I. Nikolaeva, P. Czerwinski, N. Aralova, A. Francis-Ratte, I. Joo, R. Máté, T. Pellard'a ve Kore Ulusal Müzesi'ne verileri derlemeye, analiz etmeye veya yorumlamaya yardımcı oldukları için teşekkür ederiz.

Yazar Katkıları Araştırma, M.R. tarafından kavramsallaştırılmıştır. Dilsel veri kümeleri A.S., J.D., S.O., B.D., R. Bjørn, S.R., K.-D.A., I.G., O.M., J.R.B. ve M.R. tarafından toplanmıştır. Dilsel veri tabanı M.R. tarafından puanlandı ve M.R. ve R. Bouckaert tarafından analiz edildi. Etimolojiler M.R. tarafından oluşturulmuştur. Arkeoloji veri tabanı T.L., M.C.,

T.K., G.K., J.U. ve L.G. puanlanmıştır ve M.J.H., R. Bouckaert, M.R., M.C. ve I.R.B. tarafından analiz edilmiştir. Nagabaka sitesi T.K. ve K.-Y.Y. M.J.H.'nin yönetiminde M.K.'nin tavsiyesi ile ve H.I. tarafından kazılmıştır. Nagabaka'dan gelen malzemelerin kazı sonrası analizleri K.-Y.Y., T.K., N.S., H. Tomita, H. Takamiya, J.U., P.R., R.F. ve M.Y. tarafından analiz edildi. Y.C. Angangxi verilerini , D.I.-A. ve J.-H.K. eski Kore verileri, K.I.S. Yayoi verilerini ve H.I., R.K., T.S. ve H.O. modern Ryukyu verileri paylaştı. Kore ve Japonya'dan gelen antik DNA verileri için ıslak laboratuvar çalışmaları R.A.B. ve M.H. tarafından yapıldı. Genetik bilgi analizleri F.Z VE H.K.-K.'nin verileriyle C.N tarafından sürdürüldü. Yazım işlemi ise M.R., M.J.H. ve C.N tarafından yapıldı.

Finansman Max Planck Derneği tarafından sağlanan açık erişim finansmanı.

Çakışan Çıkarlar Yazarlar, rekabet eden çıkarlar olmadığını beyan eder.

Ek Bilgiler

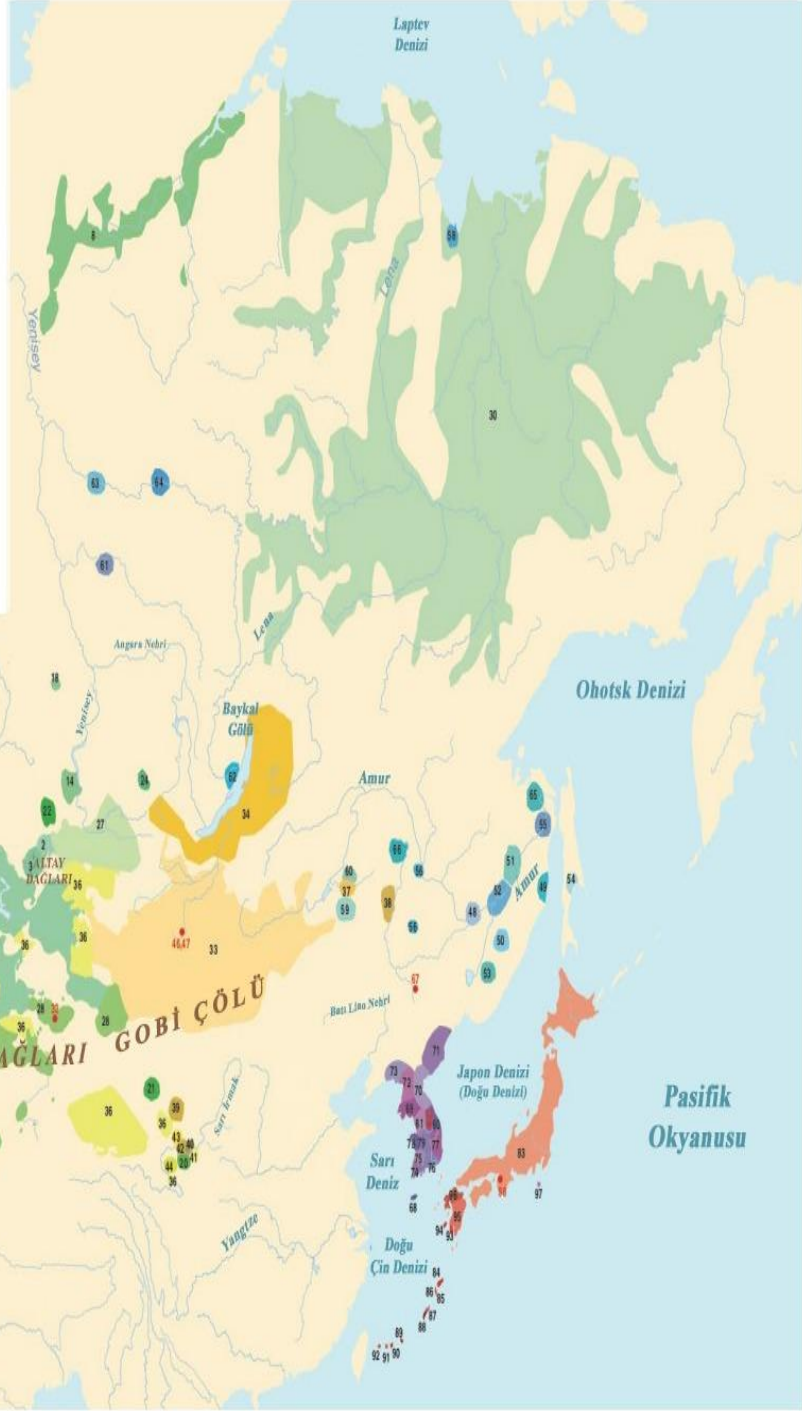
Ek bilgiler Çevrimiçi sürüm, aşağıdaki adreste bulunan ek materyalleri içerir:
<https://doi.org/10.1038/s41586-021-04108-8>

Yazışmalar ve materyal talepleri Martine Robbeets, Mark J. Hudson veya Chao Ning'e yöneltilmelidir.

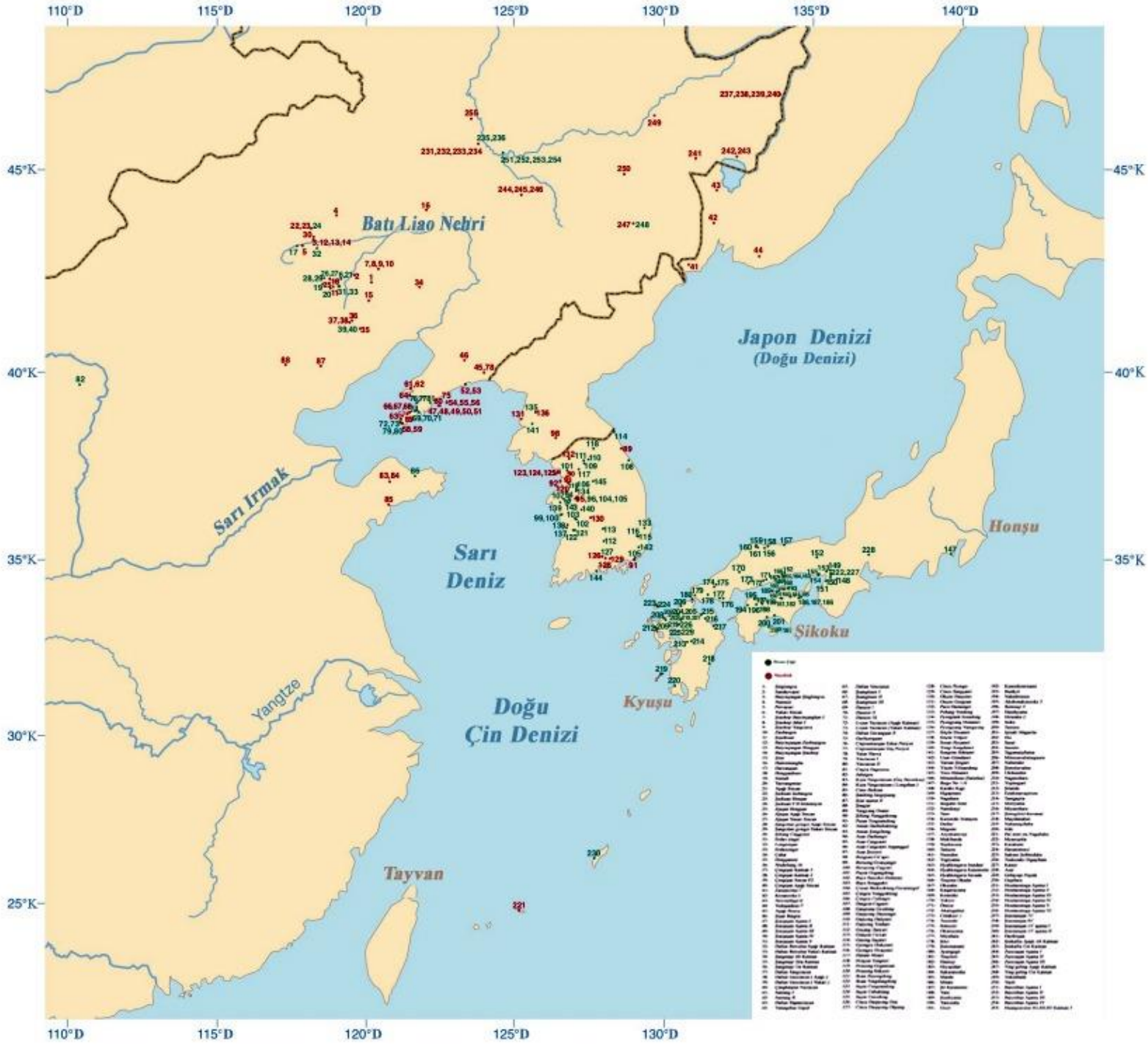
Hakem değerlendirmesibilgisi *Nature* bu çalışmanın akran değerlendirmesine katkılarından dolayı Vaclav Blažek, Dorian Fuller, Carles Lalueza-Fox ve diğer anonim gözden geçirenlere teşekkür eder. Hakem raporları mevcuttur.

Yeniden baskılar ve izinler hakkında bilgi şu adreste bulunabilir:
<http://www.nature.com/reprints>

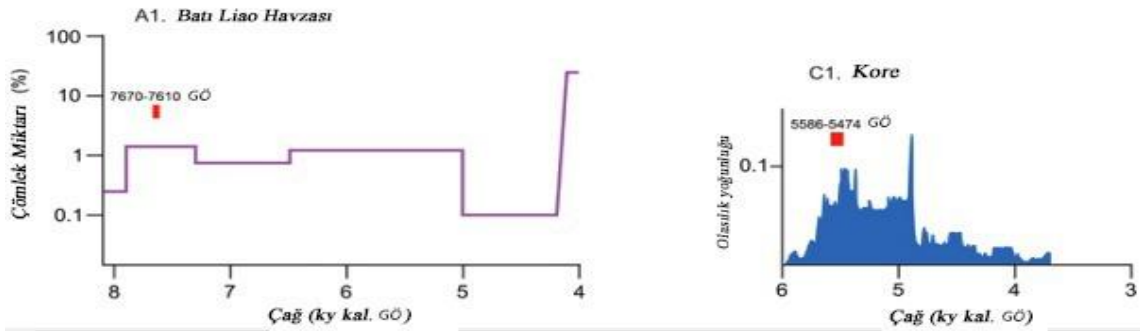
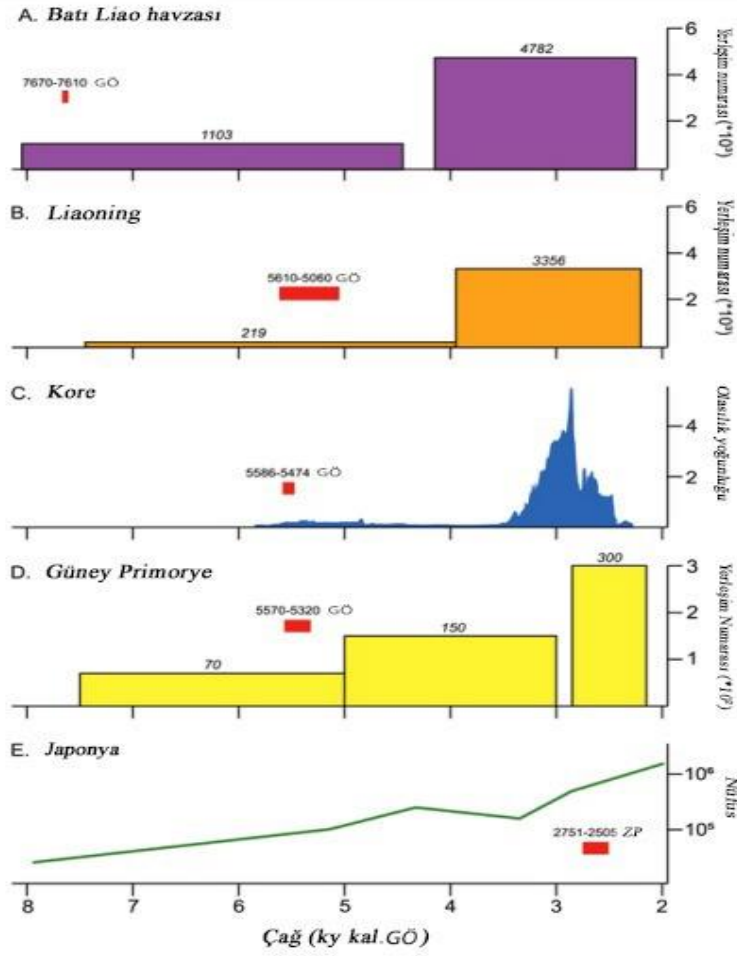
Türkçe	Moğolca	Tunguzca	Korece	Japonca
1-Bareba Tatar Türkçesi (Sibiryalı)	28-Uygur Türkçesi	55-Ulçu Tunguzcası	81-Gangwon Korecesi	
2-Kuzey Altay Türkçesi	29-Özbek Türkçesi	56-Mançu Tunguzcası	82-Geç Orta Korece	
3-Güney Altay Türkçesi	30-Yakut Türkçesi(Saba Türkçesi)	57-Sibe Tunguzcası	83-Japonca	
4-Azerbaycan Türkçesi	31-Kuman Türkçesi	58-Dren Tunguzcası	84-Tamotokuma Japoncası	
5-Bağkurt Türkçesi	32-Baki Türkçe	59-Salon Tunguzcası	85-Tokusojima Japoncası	
6-Cuvay Türkçesi	33-Halba Moğolcası	60-Brenk Tunguzcası(Humigan)	86-Toron Japoncası	
7-Kırım Tatar Türkçesi(Batı)	34-Baryat Moğolcası	61-Güney Brenk Akrahu Tunguzcası(Tunguzca'nın ayağı)	87-Yonamine Japoncası	
8-Dolgan Türkçesi	35-Kalmık Moğolcası	62-Güney Brenk Tunguzcası	88-Şuri Japoncası	
9-Gagaus Türkçesi	36-Oyrat Moğolcası	63-Kuzey Brenk Tunguzcası(Tutongcası)	89-İrabu Japoncası	
10-Kara-Kalpak Türkçesi	37-Hemnigan Moğolcası	64-Kuzey Brenk Tunguzcası(Tura)	90-Işigaki Japoncası	
11-Karaçay Balkar Türkçesi	38-Dağur Moğolcası	65-Negidal Tunguzcası	91-Hatoma Japoncası	
12-Karaim Türkçesi	39-Şera Yugar(Sarı Uygur)Moğolcası	66-Oropen Tunguzcası	92-Tongami Japoncası	
13-Kazak Türkçesi	40-Baoun(Bonun) Moğolcası	67-Carpen Tunguzcası	93-Kagayima Japoncası	
14-Hakas Türkçesi	41-Dongjian (Santla) Moğolcası	68-Ceun Korecesi	94-Kuşiki Adaları Japoncası	
15-Halac Türkçesi	42-Minhe Mongar Moğolcası(Mangğuer Moğolcası)	69-Hwanghae Korecesi	95-Kumamoto Japoncası	
16-Kırız Türkçesi	43-Haju Mongar Moğolcası(Mongğul Moğolcası)	70-Güney Hamgyong Korecesi	96-Fukuoka Japoncası	
17-Kumuk Türkçesi	44-Kangxia Moğolcası	71-Kuzey Hamgyong Korecesi	97-Hajio Japoncası	
18-Çulim Türkçesi	45-Moğolca	72-Güney Pyeongyang Korecesi	98-Baki Japonca	
19-Nogay Türkçesi	46-Orta Moğolcası(Oiizi Tarih)	73-Kuzey Pyeongyang Korecesi		
20-Salar Türkçesi	47-Orta Moğolcası(Mahaddimeth)'1-edeb	74-Güney Jeolla Korecesi		
21-Batı Uygur Türkçesi(Sarı Uygur)	48-Heshe Tunguzcası	75-Kuzey Jeolla Korecesi		
22-Şor Türkçesi	49-Oroç Tunguzcası	76-Güney Gyeongsang Korecesi		
23-Kazan Tatar Türkçesi (Merkezi)	50-Udihe Tunguzcası	77-Kuzey Gyeongsang Korecesi		
24-Tofa Türkçesi(Karagunlar)	51-Naxai Tunguzcası (Orta Amur)	78-Güney Jeunggyeong Korecesi		
25-Türkiye Türkçesi	52-Kar-Urmi Naxai Tunguzcası	79-Kuzey Jeunggyeong Korecesi		
26-Türkmen Türkçesi	53-Nanay Tunguzcası	80-Korece(Gyeonggi)		
27-Tuva Türkçesi	54-Orok Tunguzcası			



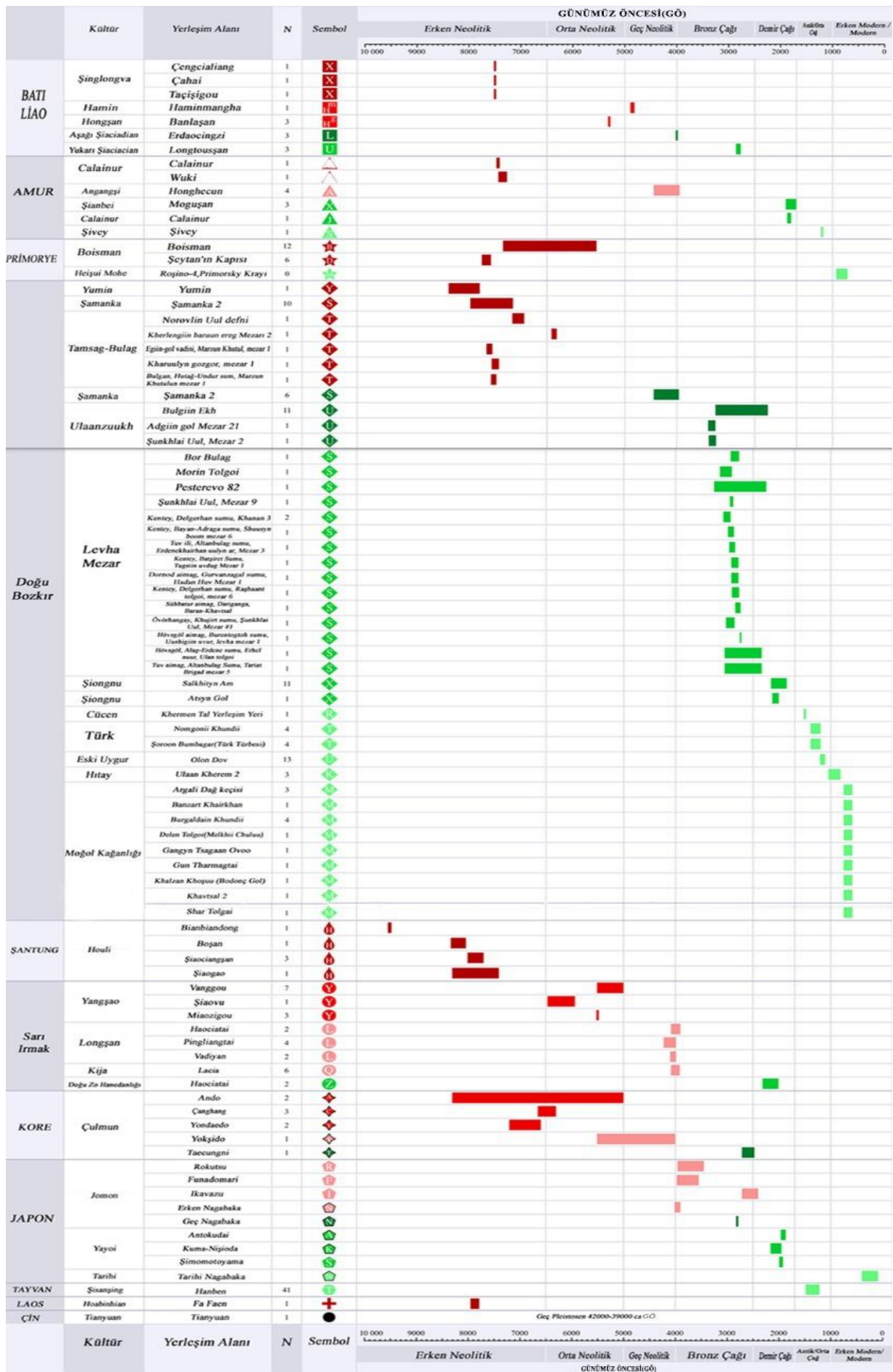
Genişletilmiş Veri Şekil 1|Şekil 1 için açıklamalar. Ana Şekil 1'e eşlik edecek ayrıntılı açıklamalar.



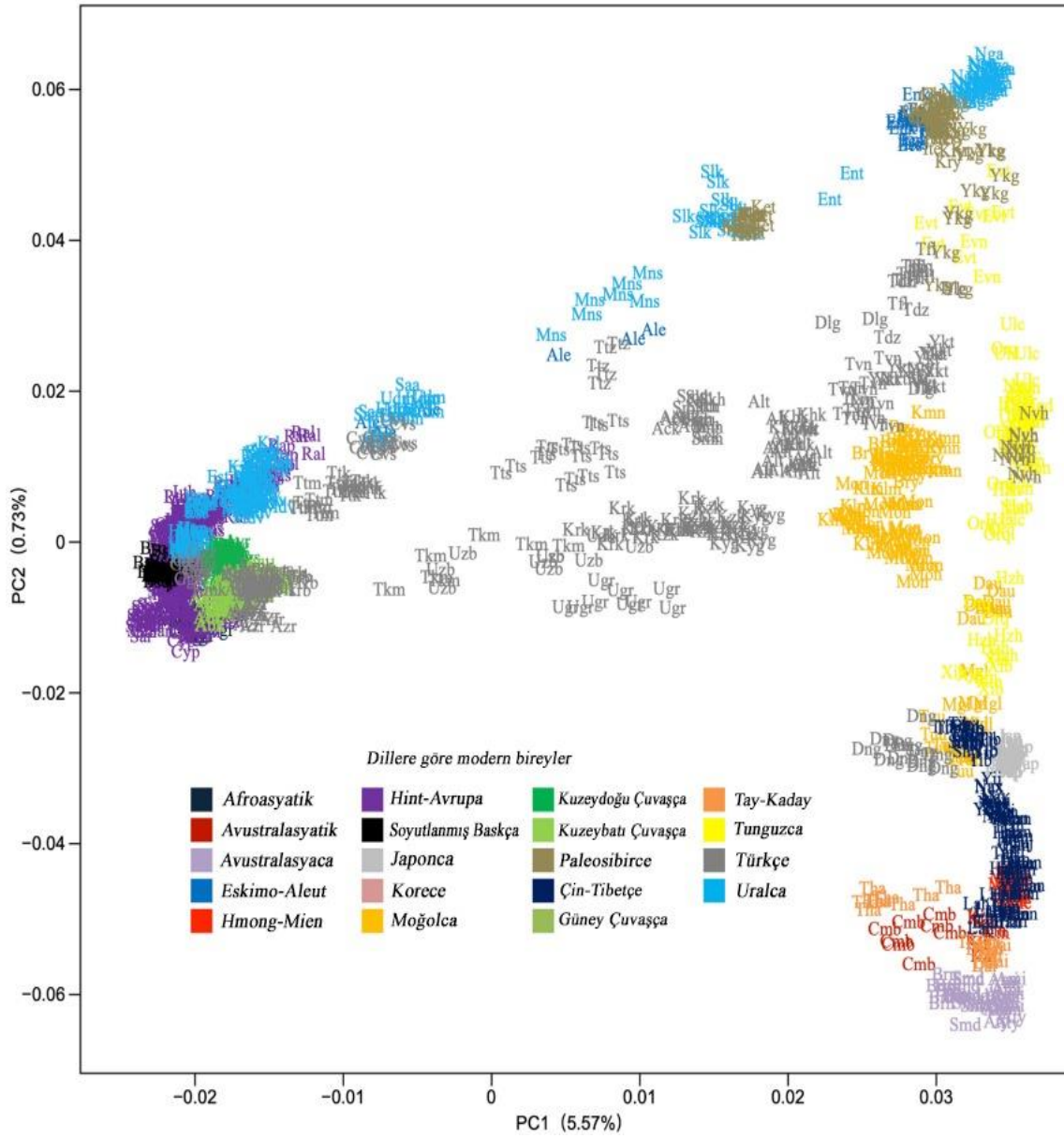
Genişletilmiş Veri Şekil 2| Şekil 2 için Açıklamalar. Ana Şekil 2'ye eşlik edecek ayrıntılı açıklamalar.



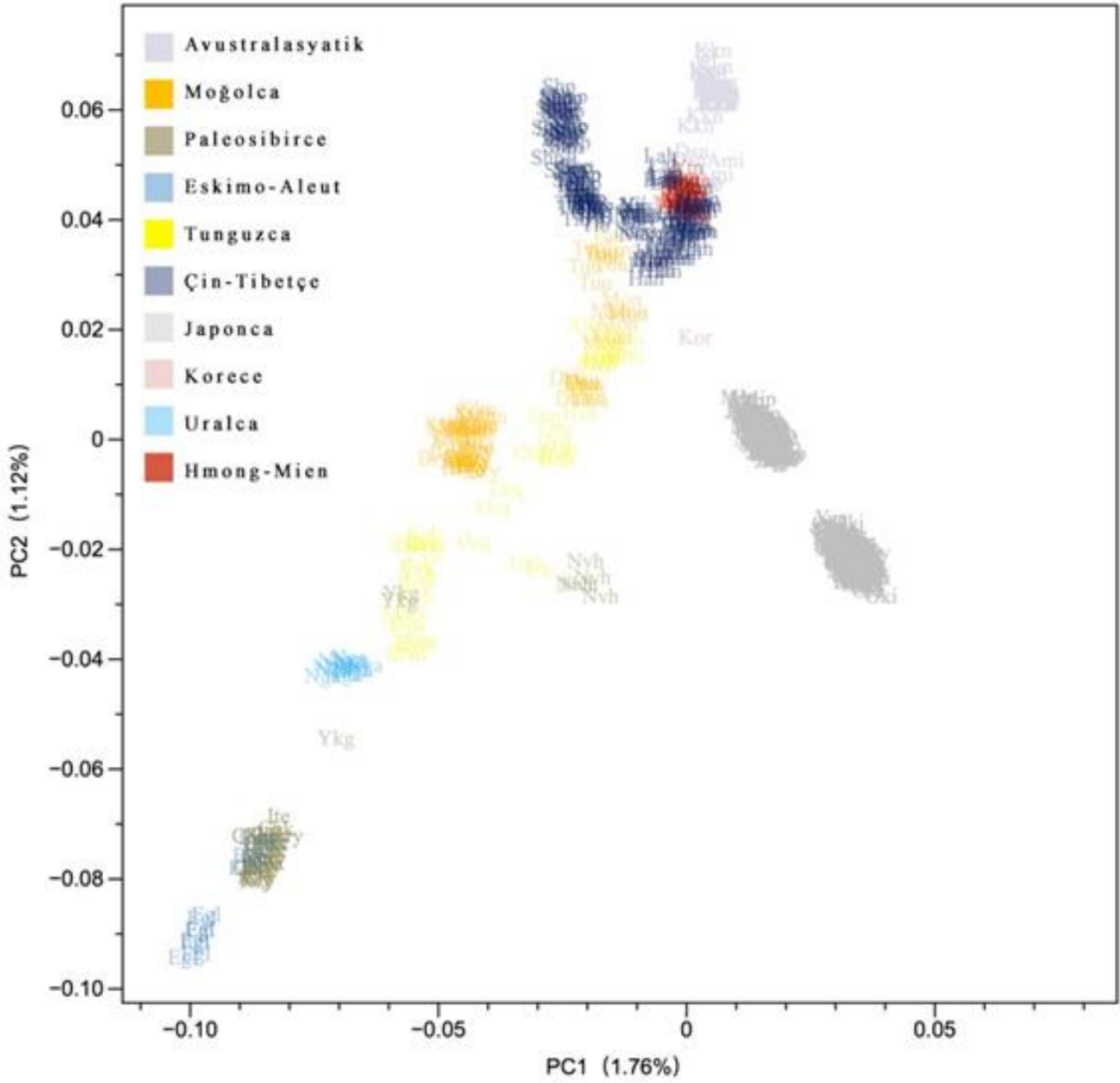
Genişletilmiş Veri Şekil 3|Neolitik ve Tunç Çağı'nda tarımla demografik değişimler. Kuzeydoğu Asya. A1 Batı Liao²⁹ için çanak çömlek miktarını kullanarak 8000-4000 GÖ civarında darı yetiştiriciliğinin benimsenmesini takiben değişiklikleri gösterir ve B2, Kore⁸⁷ için radyokarbon vekalet tarihlerini kullanarak bu değişiklikleri gösterir. Şekil A'dan E' ye, Bronz Çağı'nda darı ile pirinç, arpa ve buğdayın uyumunu takiben Şekil A'dan E' ye, Bronz Çağı'nda darı ile pirinç, arpa ve buğdayın uyumunu takiben 8000-2000 GÖ civarında uzun vadeli dinamikleri göstermektedir ve Kuzeydoğu Çin⁸⁸ için saha numaralarına, Kore⁸⁷ için radyo karbon tarihlerine ve Japonya⁸⁹ için saha numaralarına dayanmaktadır. Demografik bilgileri vekillerden elde etmek için kullanılan referanslar ve yöntemler için bkz. Ek Veriler 7.



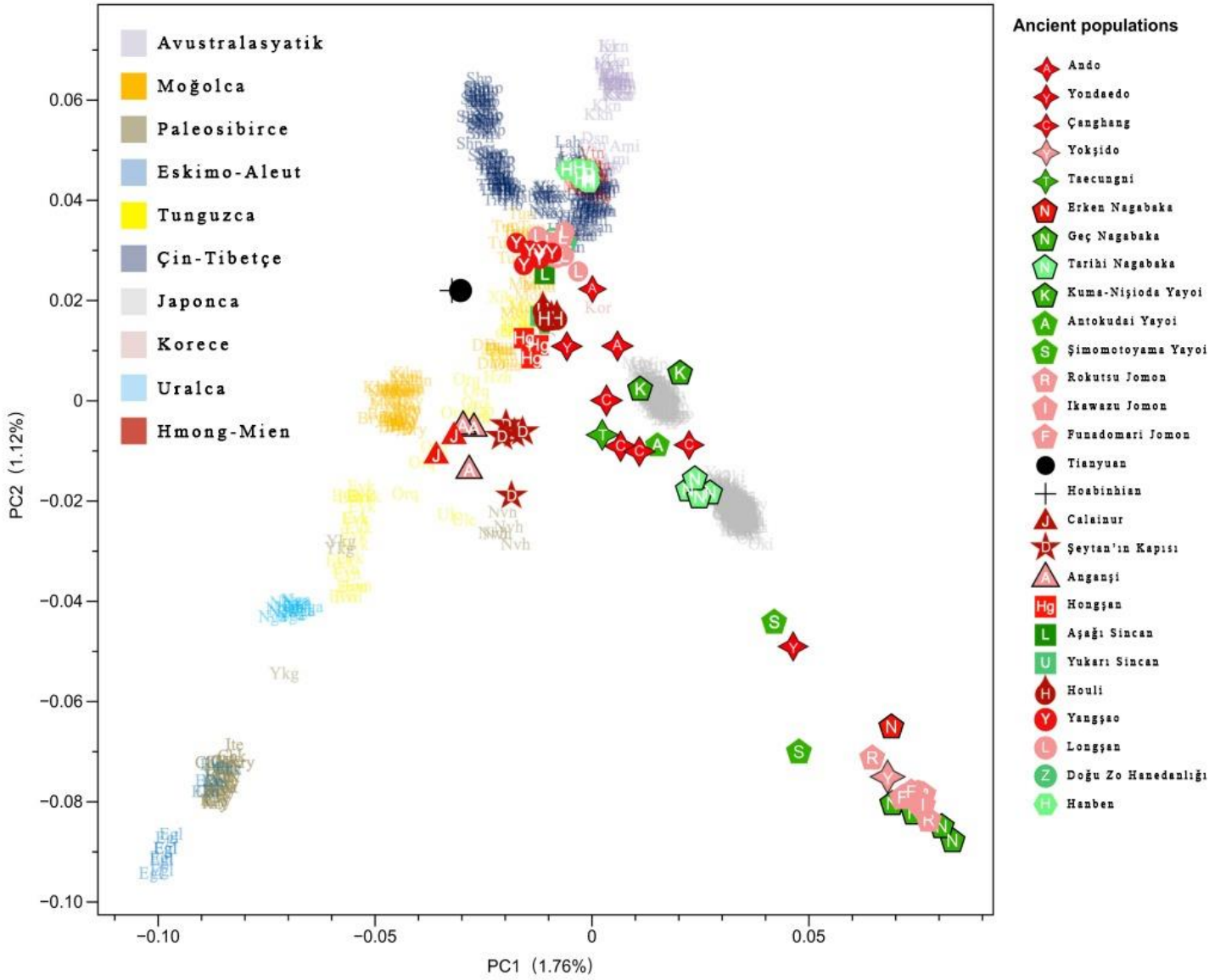
Genişletilmiş Veri Şekil 4| Zaman ve Mekanda Bulunan Antik Genomlar.Ana Şekil 3'e eşlik edecek ayrıntılı bir açıklama ve Genişletilmiş Veri Şekillerini içerir. 7-10



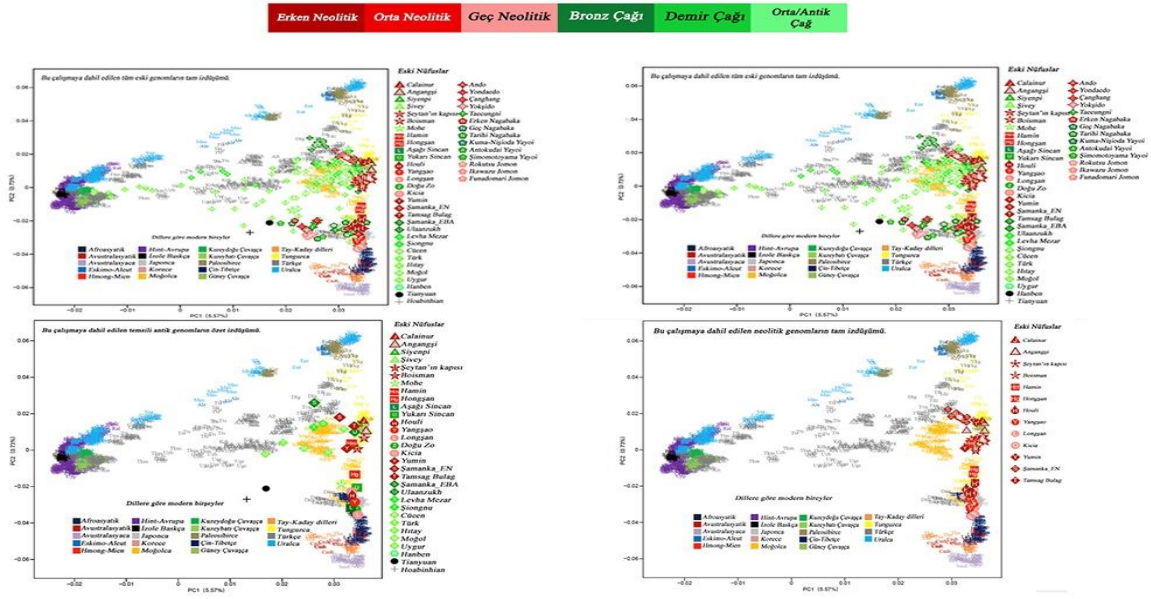
Genişletilmiş Veri Şekil 5| PCA günümüz Avrasya'sının genetik yapısını göstermektedir.PC1, Batı ve Doğu Avrasya nüfuslarını, PC2 Güney ve Kuzey Avrasya nüfuslarını birbirinden ayırır. Transavraya nüfusları alt familyaya göre renklendirilmiştir (gri Türk, turuncu Moğol, sarı Tunguz, pembe Koreli, açık gri Japon). Transavrayalı olmayan nüfuslar soylara göre renklendirilmiştir. Nüfuslar üç harfle etiketlenmiştir, kısaltmaların listesi için bkz. Ek Veriler 10.



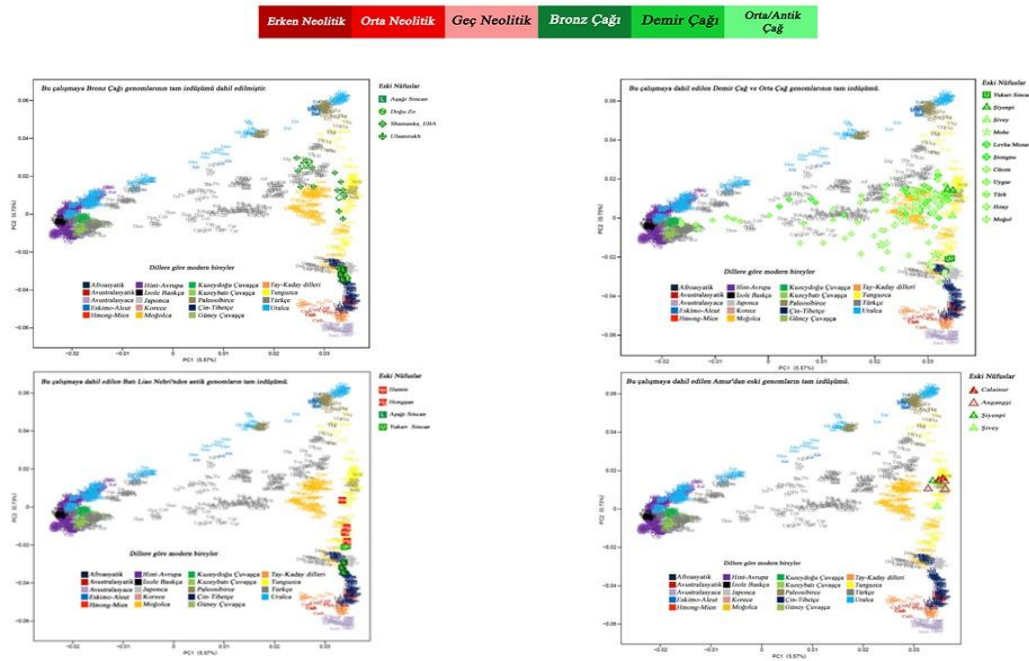
Geniřletilmiş Veri Őekil 6|PCA g¼n¼m¼z Doęu Asyalıların genetik yapısını g¼stermektedir. N¼fuslar ¼ harfle etiketlenmiřtir, kısaltmaların listesi i¼in bkz. Ek Veriler 10.



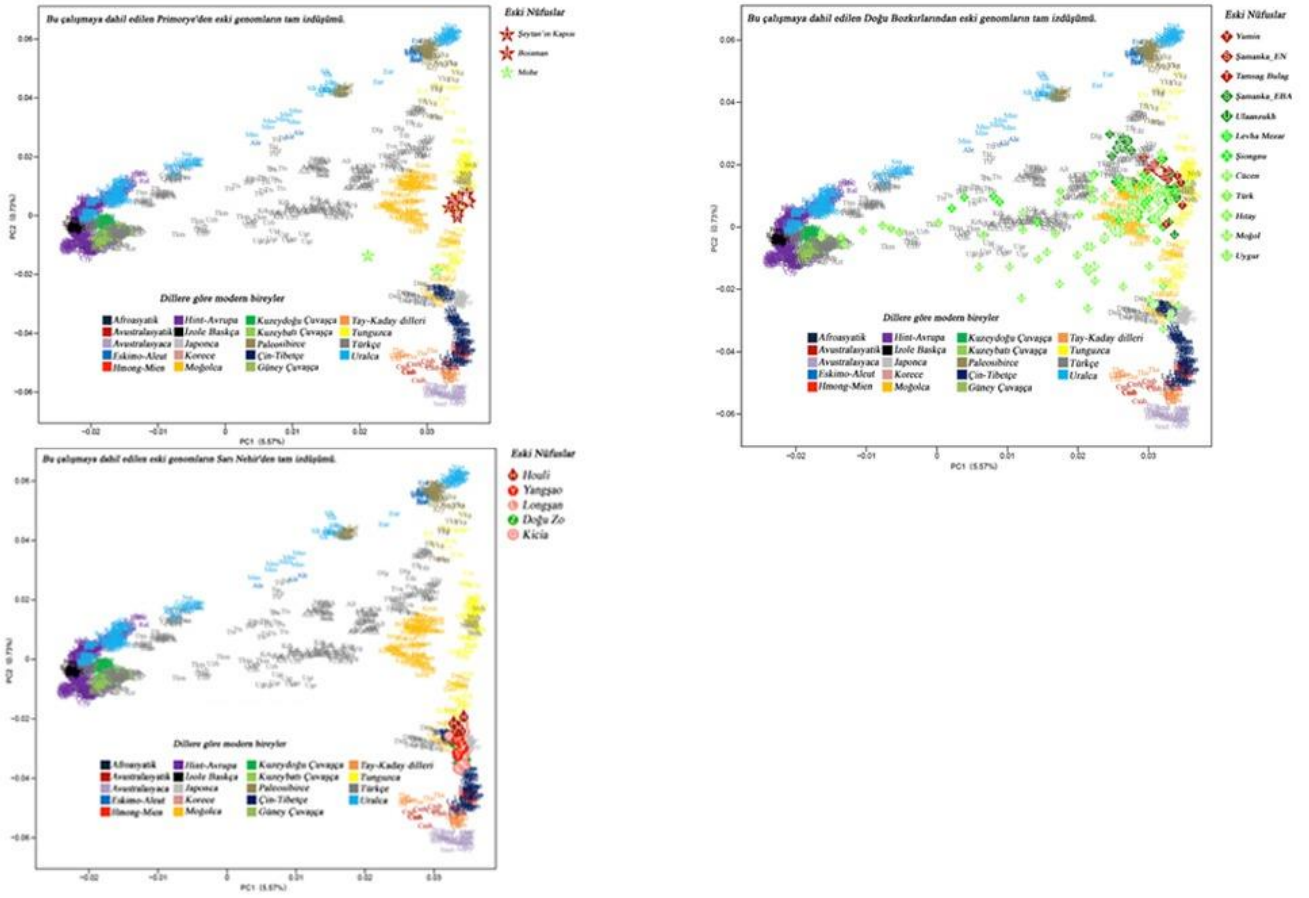
Genişletilmiş Veri Şekil 7 | Günümüz Doğu Asyalıların genetik yapısını gösteren PCA üzerinde çizilen antik genomlar. Ayrıntılı açıklama için bkz. Genişletilmiş Veri 4.



Genişletilmiş Veri Şekil 8 | PCA’da çizilen antik genomlar, günümüz Avrasyalıların genetik yapısını gösteriyor. Ayrıntılı bir açıklama için bkz. Genişletilmiş Veri Şekil 4.



Genişletilmiş Veri Şekil 9 | Tunç Çağı, Demir Çağı, Batı Lio ve Amur’da kalma antik genomlar, günümüzün Avrasyalıların genetik yapısını gösteren PCA’da çizildi. Ayrıntılı bir açıklama için. Bkz. Genişletilmiş Veri Şekil 4.



Genişletilmiş Veri Şekil 10 | Primorye doğu bozkır ve Sarı Nehirden gelen eski genomlar, günümüz Avrasyalıların genetik yapısını gösteren PCA’da çizildi.Ayrıntılı bir açıklama için. Bkz. Genişletilmiş Veri Şekil 4.