

REVIEWS
AND THEORETICAL ARTICLES

The Impact of Genetics Research on Archaeology and Linguistics in Eurasia

J. Mallory^{a, *}, A. Dybo^{b, c, d}, and O. Balanovsky^{e, f, g, **}

^aQueen's University, Belfast, Northern Ireland, BT7 1NN United Kingdom

^bInstitute of Linguistics, Russian Academy of Science, Moscow, 125009 Russia

^cInstitute for Oriental and Classical Studies, Higher School of Economics, Moscow, 105066 Russia

^dLaboratory of Linguistic Anthropology, Tomsk National State University, Tomsk, 634050 Russia

^eVavilov Institute of General Genetics, Russian Academy of Science, Moscow, 119991 Russia

^fResearch Centre for Medical Genetics, Moscow, 115522 Russia

^gBiobank of North Eurasia, Moscow, 115201 Russia

*e-mail: j.mallory@qub.ac.uk

**e-mail: balanovsky@inbox.ru

Received July 17, 2019; revised July 31, 2019; accepted August 1, 2019

Abstract—This article attempts to outline the current impact that genetics is having on the fields of archaeology and historical linguistics across the Eurasian continent. It positions the relationship between all three disciplines by reviewing the earlier history of their interactions. In the area of archaeology, there has been a long history of research into the subject of human migration. We briefly review the application of such techniques as craniometry, pigmentation, dermatoglyphics, classical markers and the retrospective reconstruction of population movements from the modern DNA of human populations. We then turn to the revolution created by the application of ancient DNA in three separate areas: Early Man dispersals and legacies, the spread of agriculture and the massive expansion of populations during the Early Bronze Age. Examples are provided of how aDNA is impacting on the study of the origin and dispersals of ethno-linguistic groups. In addition to human migrations, genetics is also impacting on the reconstruction of past lifeways and examples are drawn from research on palaeodiet, palaeopathology and palaeodemography. Genetics is also contributing to major issues of historical linguistics involving the origins and dispersals of the major Eurasian language families. It provides evidence that helps distinguish between instances involving significant migration from those effected by language shift with a minimal genetic trail. Two cases, the Proto-Indo-European and the Proto-Altai homelands are reviewed along with some of the methodological problems of synchronizing genetic and linguistic evidence.

Keywords: population genetics, ancient DNA, archaeology, historical linguistics, physical anthropology, Indo-Europeans, Altaic

DOI: 10.1134/S1022795419120081

The results of population genetics research over the past century have impacted on a number of other disciplines studying human populations, particularly those with an interest in reconstructing past societies, such as archaeology (which here will also include physical anthropological analysis of human remains) and historical linguistics. The primary impacts have concerned attempts to trace both the origins and dispersals of human population groups and their languages across the entire continent of Eurasia and on into the Americas. In addition, the data have sometimes been employed to reconstruct past lifeways, palaeo-demography in its widest sense. The chronological framework extends, so far, from the entry of *Homo erectus* into Eurasia up until the recent past and the establishment of all of the historically known pop-

ulation groups of the Continent. Research has proceeded along two different methodological paths: 1) retrospection, i.e. the reconstruction of the past from the known genetic data on present populations; and 2) direct examination of the past, i.e. the application of genetic research to material retrieved from past (prehistoric and historic) contexts whose chronology has been established.

The intersection of genetics with linguistics primarily involves the search for the homelands of the various major language families of Eurasia, e.g. Indo-European, Uralic, Altaic, Sinitic, and the patterns of their dispersals. In fact, much of the archaeological discourse over the 19th and 20th centuries was driven by scholars seeking to establish the homeland of the

Indo-European language family or that of its constituent branches, e.g. Celtic, Germanic, Slavic. In this way the interface between genetics and linguistics has been largely secondary to the combination of genetic and archaeological evidence. Nevertheless, there are still some areas where genetics has had a direct impact on linguistics, including locating in both time and space the branching of populations when it may correspond to the branching of languages.

Although the term *archaeogenetics*, i.e. “the study of the human past using the techniques of molecular genetics” [1], was only coined twenty years ago, archaeologists have employed genetic evidence in their reconstructions of the past since the 19th century. The traditional progression, both methodological and chronological, between the use of Classical Markers, modern DNA evidence, and ancient DNA (aDNA) evidence was preceded by a period in which other “pre-molecular” techniques such as craniometry and the geographical plotting of a variety of phenotypic traits, e.g. dermatoglyphics, pigmentation, were regarded as reliable evidence that could be employed to trace population origins and migrations. These traits describe phenotype rather than genotype and are therefore classified as physical anthropology rather than population genetics. However, the rationale for using these traits for reconstructing the human past was that these traits are inherited ones (though influenced also by environmental factors), and therefore could be considered as a proxy of genes. Moreover, there is no clear boundary between the phenotypic and genotypic characteristics. Even classical (biochemical) genetic markers are actually the activity of human enzymes, which is, strictly speaking, a part of the phenotype. And traits, such as the ability to recognize the bitter flavor of phenictiocarbamide (controlled by a single gene) can be equally classified as physical anthropological and genetic markers.

Some of these techniques are still employed and one of the major questions is to what extent they will be ultimately replaced by full sequencing of modern genomes and aDNA or whether they will still be regarded as useful adjunct approaches. Although all of the earliest techniques are clearly associated to some degree with inherited traits, the approach to the data did not involve any of the laboratory techniques associated with molecular genetics. In fact, it has been argued that except for the Hirschfelds’ work during World War I, physical anthropology, at least in the West, “generally ignored genetics” until the 1960s because so much of early human genetics, was devoted to eugenics [2]. In the East – mainly in the Soviet Union – anthropogenetics was initially a branch of physical anthropology, and some famous anthropologists, like Viktor Bunak and Yakov Roginsky, undertook research in population genetics as well. When genetics was dismissed as a false science in the Soviet Union (during the so-called Lysenko period) human population genetics partly survived within physical

anthropology and re-expanded from there when genetic research became possible again.

In this review we consider a few points of how genetic research – starting with physical anthropology (“pre-molecular genetics”) and ending with aDNA – has impacted on the archeological understanding of human migrations and past lifeways and influenced historical linguistics.

MIGRATIONS

The study of prehistoric migrations is the main area of overlapping interests between archaeology and population genetics. Genetic and wider – biological – evidence has been widely explored to trace human migrations employing both retrospective and direct approaches, and using a variety of physical anthropological and genetic traits.

Craniometry

The longest employed anthropological approach to defining human population groups and trace their origins and dispersals involves craniometry, the measurement of the human skull. In the first half of the 19th century Anders Retzius devised the cephalic index which divided skulls into brachycephalic broad-heads and dolichocephalic long-heads. These classes, augmented later by mesocephalic ‘medium’ heads, became reified into separate classes of human populations which were imagined to have had distinct origins and varying histories of admixture. By the early 20th century the degree to which the cephalic index was governed by environmental factors rather than solely by genes was being vigorously challenged although the terminology was still employed in more complicated attempts to describe human populations based on geographical variants, e.g. Nordic, Alpine, Dinaric, Mediterranean. By the 1970s, those physical anthropologists who believed that the cranium did provide a useful index of genetic relatedness employed a suite of select measurements to which were applied multivariate analysis and principal component analysis. At least here there was sufficient evidence that indicated there was a correlation between craniometry and genetics [3]. These techniques are still employed today although reception of their conclusions has been varied across Eurasia with far greater skepticism, usually indicated by archaeologists completely ignoring such data, in western Europe, especially among Anglophone scholars, and far more attention in Eastern Europe and Asia.

Recent advances in aDNA have addressed some of the major issues that were tackled a generation ago by physical anthropologists and it is instructive to compare the results of the two approaches. As an example we take Roland Menk’s attempt [4] to resolve the issue of whether there was a major migration from the Pontic-Caspian steppe *c* 3000 BCE into Central Europe as

was suggested in a popular solution to the problem of Indo-European origins. Menk employed 26 cranial measurements across a sample of 1842 skulls to produce a principal components map (Fig. 1). The map recognized four main groups with several outliers. His Palaeo-European group comprised samples from two cultures, one of which is now separated from his other groups by geneticists being primarily Eastern European Hunter-Gatherer (EHG). He recognized a Kurgan core, a blanket label for the Yamnaya and closely related later steppe cultures which geneticists now identify [5] as primarily an admixture of EHG and Caucasian Hunter-Gatherers (CHG) (but this would also include one of his Proto-European cultures, the Sredni Stog). Beyond that he identified cultures that had been influenced by the “Kurgan” expansions. Some of these, including the Afanasievo culture of the Altai-Yenisei region, are now seen to be genetically identical to the Yamnaya culture [6] and so both craniometric and genetic evidence converge to attest a migration of *c* 1500 km. It also includes the Andronovo culture which also has a steppe core [6] but also contains a substantial additional admixture of both Anatolian Farmers (AF) and Western Hunter-Gatherers (WHG). His third group, labeled ‘Old Europe’, here a cover label for what today we would call Anatolian Farmerid contain cultures that aDNA describes as AF (e.g., Linear Ware, Gumelnitsa, Russe, TRB, Tripolye) but it also contains most of his Corded Ware samples and the descendant Unetice culture, both of which are distinguished by a principle steppe component (EHG + CHG) although mitigated by AF and WHG which may explain why they have been pulled into the craniometric orbit of Neolithic Farmers rather than steppe pastoralists. This led Menk [4] to dismiss a steppe migration into Central Europe although the aDNA evidence finds 75% of Corded Ware burials sampled possessing steppe ancestry [5]. Also, Beaker skulls were placed in a totally separate category although all those sampled should have been also genetically grouped with the steppe-derived Corded Ware culture although also heavily admixed with AF and WHG. It is also noteworthy that two of his isolates, the Siberian Neolithic sample and the Comb Ware sample might be explained genetically as belonging to the West Siberian Hunter-Gatherers and WHG groups, again showing some agreement between craniometrics and genetics. Figure 1 visualizes the similarity of the craniometrics and genetic results. Most importantly, the green area – predominantly “blue” steppe component but with some admixture with “yellow” Anatolian Farmers component – is also intermediate between these two clusters on the craniometric plot. Therefore, the discrepancy is mainly about the interpretation of the plot – should the green area be interpreted as part of the “Old Europe” cluster or as a bridge between “Old Europe” and “Kurgan influence” clusters?

The concordance between physical anthropological and genetic patterns seems to be a rule, rather than an exception. This is not surprising, because both systems describe the biological variation of human populations. Another example of such concordance comes from data on modern populations of Eastern Europe. Figure 2A presents one of the largest studies based on 22 phenotypic traits, including height, eye and hair pigmentation, and face measurements of 648 populations from Eastern Europe and the Caucasus [7]. Figure 2B presents the PC plot for the populations from exactly the same geographic area constructed using genome-wide data available via the GG-base [8]. Both, physical anthropological (Fig. 2A) and genetic (Fig 2B) data agree on the main patterns: the abrupt genetic difference between East European and (geographically adjacent) Caucasus groups, gradual difference between East European and adjacent Volga-Ural groups, the similarity of West and East Slavs with Baltic-speaking groups, while South Slavs cluster with Balkan populations. There are also some minor points where physical anthropological and genetic data disagree – for example, on the genetic plot Belorussians cluster with Poles/Lithuanians while on the anthropological plot Belorussians, like Russians, are intermediate between Poles/Lithuanians and Volga-Ural groups.

What becomes clear from such studies is that modern advances in the identification of specific cranial measurements and the manipulation of data does indicate that differences in the human phenotypes (crania, bodies, and other physical anthropological traits) do carry a significant genetic signal but that aDNA offers a far more sensitive tool for evaluating migrations and admixture in populations. But although geneticists, again at least in the West, have been largely content to ignore the craniometric evidence, there may still be areas of common interest. It should be noted that the sample size for craniometric datasets is much larger than that for aDNA, and even with the rapid growth of aDNA data there will always be many skulls which can be described craniometrically but not genetically because of the poor preservation of DNA. It would be also helpful – though not necessary for productive cooperation – if genetics can isolate out the genes that control variability in cranial measurements.

Pigmentation

In 1878 the German-American anthropologist, Theodor Poesche, published his study of Indo-European origins which employed the textual evidence of ancient literature to argue that the earliest Indo-Europeans (e.g., Greeks, Romans) had light hair and blue eyes. Confusing a high incidence of albinism with the center of the blond ‘race’, Poesche located the Indo-European homeland in the Pinsk Marshes [9]. Poesche’s theory was soon replaced by Karl Penka [10, 11] who assembled far more textual evidence on the

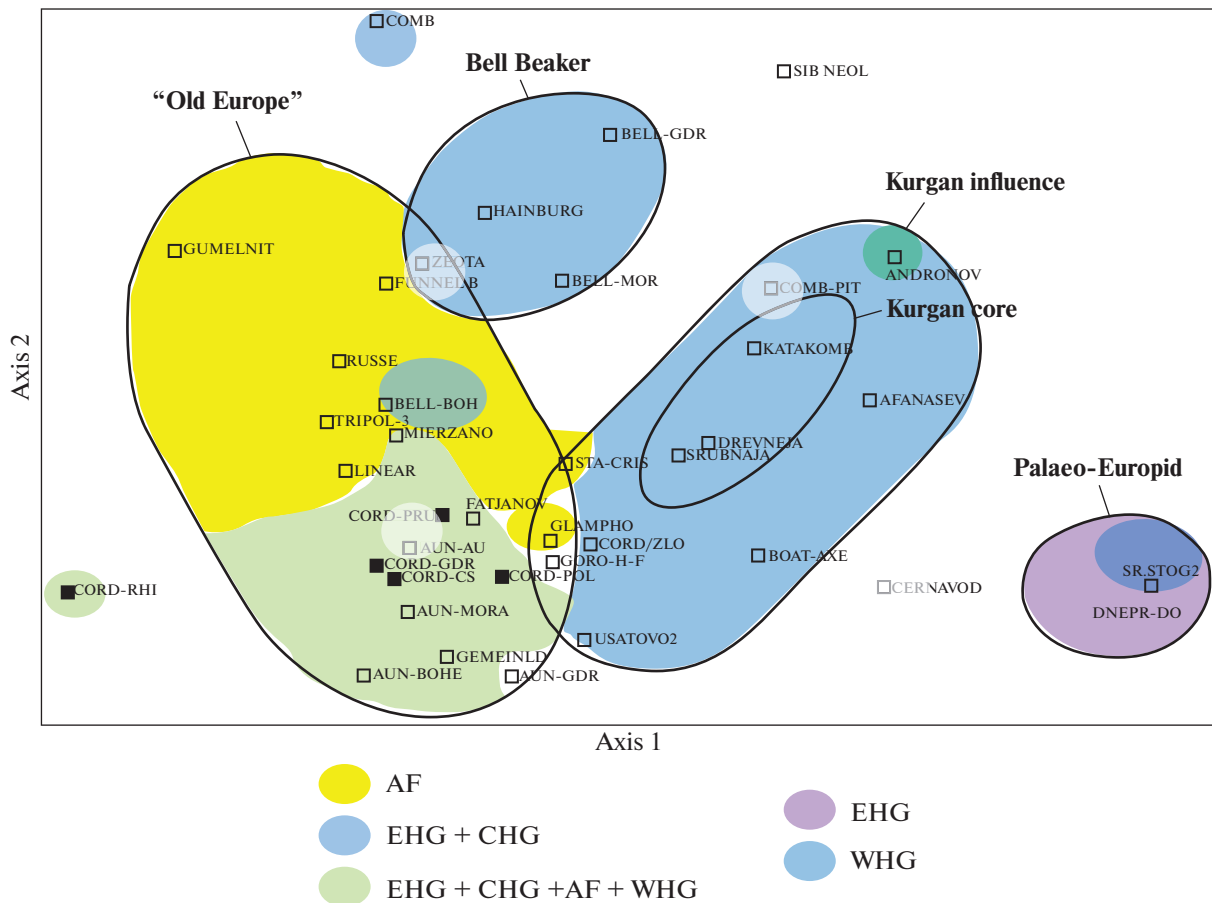


Fig. 1. Correspondence between craniometrics and ancient DNA analysis of Neolithic/Bronze Age Europe. The plot represents similarities between ancient European populations based on craniometrics data (modified from [4]). Black lines designate clusters of populations identified in [4]. Colors designate which genetic component(s) – revealed by aDNA – predominates in the same populations. The green color marks the mix of the blue (steppe) with the yellow (Anatolian farmers) component.

pigmentation of the ancient Indo-Europeans to conclude that the homeland lay in Northern Europe. This launched a frenzy of publications of increasingly dubious value regarding the ‘original’ phenotype of Indo-European speakers. The political misuse of these observations is too well known to recount and discussion of pigmentation, whether drawn from the accounts of ancient literature or from the artwork of various peoples of Eurasia, receded from most scholarly journals after 1945. Nevertheless, the subject has been returned to more recently where more dispassionate scholars have indeed concluded (based on non-genetic evidence) that “early IE-speaking elite groups often had light hair and light eyes, even in regions where the mass of people had dark hair and dark eyes” [12].

Now aDNA provides the opportunity to test both the literary and artistic depictions of ancient populations. The pigmentation and eye-color of a Mesolithic Iberian [13] has been associated with alleles consistent with dark skin pigmentation, dark hair but blue eyes, while an Irish Neolithic woman had black hair and

brown eyes, while another individual, associated with the shift in the main western European genotype in the Early Bronze Age had alleles indicative of light hair and (possibly) blue eye color [14]. More importantly for those who have sought light hair and eye color among ancient Greeks, aDNA has revealed that Bronze Age samples of Greeks have revealed them to resemble their depiction in contemporary frescoes, i.e. dark hair and eye color [15]. In short, research in aDNA is not only revolutionizing our understanding of the genotype of past populations but it is now shedding light on their phenotype, on the actual appearance of prehistoric populations.

Dermatoglyphics

The analysis of the patterns in both finger and palm prints has been employed to reconstruct past population movements on the understanding that broad patterns of frequency are governed by genetics. Retrospective studies have been undertaken across Eurasia [16–18], from which the SE-NW clines within Europe

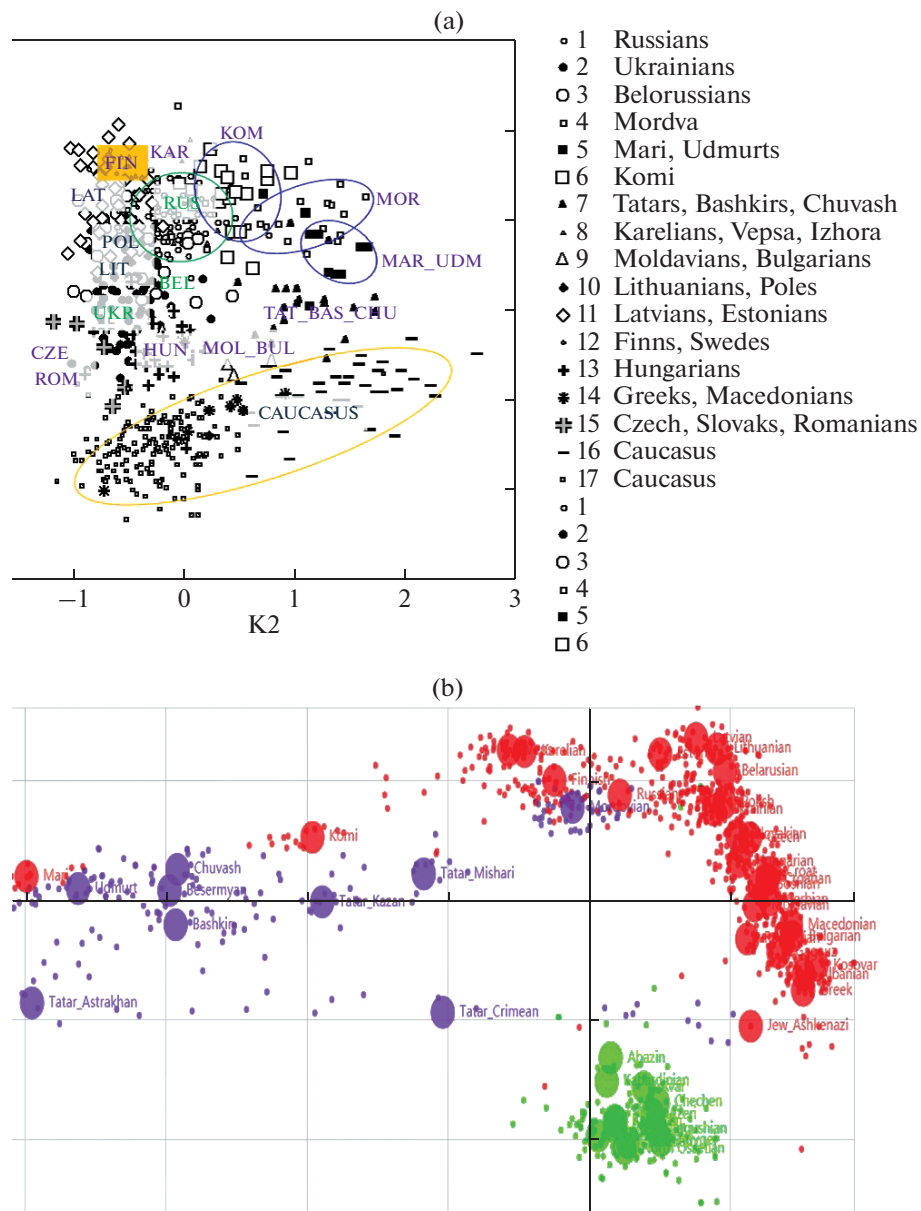


Fig. 2. Correspondence between physical anthropological and genetic analysis of East European and Caucasus populations. *a* – analysis of physical anthropological data (modified from [7]); *b* – analysis of genome-wide data (PC plot).

have been used to support the hypothesis that the European Neolithic was a product of demic diffusion from SW Asia. The application of such techniques to prehistoric populations is extremely limited to finger impressions on prehistoric pottery and clay tablets recovered in the SE Aegean where the frequency patterns (between arches, loops and whorls) have been published revealing a marked shift between the frequencies found in Minoan Crete and Mycenaean Greece [19]. But very little could be made of this as the Mycenaean pattern is also reflected in several modern Central and South African populations.

Classical Markers

The preparation of principal component maps of classical genetic markers, e.g., ABO blood group, Acid phosphatase 1, Adenosine deaminase, by Luca Cavalli-Sforza and his associates had a major impact on both archaeological and linguistic discourse in Eurasia and elsewhere throughout the world. His maps for Europe, for example, illustrated a 1st PC (28.1%), which was usually interpreted as having been formed by the expansion of the first farmers across Europe in the Neolithic (*c* 7000–4000 BCE) while his 3rd PC (10.6%), which was centered on the Black and Caspian seas, was associated with later expansions of

the Yamnaya culture and its descendants westwards. The problem with such maps based on modern gene distributions was the uncertainty as to whether they actually mapped past genetic 'events'. Critics, for example, argued that the 1st PC may have also included evidence of the earlier peopling of Europe during the Palaeolithic or later east to west movements associated with the Roman Empire. Similarly, the 3rd PC, interpreted as a result of an Eneolithic migration might very well have also been a product of a series of later historical migrations from the steppelands into Central Europe.

A very similar set of cartographic methods for analyzing genetic structure was independently developed in the East by Soviet researchers. These methods have been applied not only for variation of classical markers but also for variation of archeological characteristics [21, 22]. In one study, 186 Upper Paleolithic sites (26–16 ky BP) from across the USSR were characterized by 116 features of material culture – for example, presence/absence of bones of a given species at the site, presence/absence of specific tools, etc. This dataset resulted in a map of the 1st PC (Fig. 3A) revealing a genetic contrast between West and East Eurasians with a very narrow intermediate zone. Applying the same methodology to the Final Paleolithic period (15–12 ky BP, Fig. 3B) it identified a wider intermediate zone [21, 22], while analysis of the modern gene pool (Fig. 3C) demonstrated that the same West-to-East trend became smooth and gradual [23].

These and many other studies exemplified that geographic maps of PCsworks as an interdisciplinary method, revealing the stable patterns of a genetic landscape, however, it is often not clear which migrations shaped this landscape.

Modern DNA

At the beginning of the 21st century there were frequent attempts to retrospectively establish patterns of migration employing frequency distribution maps of DNA markers in modern populations. Haplogroups of both non-recombining chromosomes – mitochondrial DNA and the Y-chromosome – became the most popular tools. Some of these studies crossed from specifically scientific journals to the genre of popular science where haplogroups were made more reader-friendly by being assigned personal names, e.g. Ursula (mt U), Helena (mt H), [24] or Ruslan (Y R) and Ruisko (R1b) [25]. Potentially, this phylogeographic approach is able to locate a haplogroup in both space and time, i.e. estimate the place and date of a haplogroup's origin and the directions of its dispersal by human migrations. However, this approach, when employed to tackle some of the major questions of linguistics and ethnic origins, suffered from a number of methodological problems, and some proposed expansions were later contradicted by the evidence of

aDNA. For example, in such studies the center of distribution of R1b, the dominant male haplogroup of western Europe, was placed in Iberia (hence the use of a Basque personal name, Ruisko), and it was presumed that its distribution should be explained by the repopulation of western Europe after the last glaciation. Some critics thought it might be better explained by the expansion of the earliest farmers from the Near East [26]. But the evidence of aDNA has now indicated that R1b was only carried westwards to Atlantic Europe as part of the migration from the steppelands at the transition between the Late Neolithic and Early Bronze Age [5, 6, 27].

There are also examples, when retrospective-based conclusions match the aDNA evidence, like the basal position of Upper Paleolithic Ust-Ishim within the Y-chromosomal NO clade, or concordance between phylogeography of haplogroup Q variants and the aDNA evidence of its presence in the Americas. In cases of relatively minor discrepancies between aDNA and retrospective phylogeographic approaches, aDNA – due to data scarcity – does not necessarily provide the better estimate. For example, the appearance of haplogroup N3a3 in the Baltic region is phylogenetically dated back to the 2nd millennium BC, while aDNA data available so far identified the earliest N3a3 one thousand years later [28]. Normally, data on phylogeography of modern haplogroups play in orchestra with the aDNA and helps to resolve the picture better.

Ancient DNA

The greatest current impact of genetics on archaeology has been in its ability to provide creditable evidence for human migrations and the deeper origins of mixed populations within verifiable time depths. Examples can be taken from a cross section of the many research projects that have sought to trace human dispersals in Eurasia across time.

Early Man

The impact of genetics on our understanding of the archaeological ancestors of modern humans has been profound. It has resolved a number of major issues that have been long debated among archaeologists investigating the Palaeolithic. For example, the contribution of Neanderthals to modern Homo sapiens has long been disputed among physical anthropologists. With the recovery of Neanderthal mitochondrial genome it was concluded there was no admixture at all. In a few years, however, the same group of researchers led by Svante Paabo managed to sequence the nuclear Neanderthal genome and demonstrated that Neanderthals did breed with biologically modern humans and that their genes contribute somewhere between 1 and 3% of the genome of non-African humans. Moreover, due to the analysis of a finger bone from the Denisova cave in the Altai Mountains,

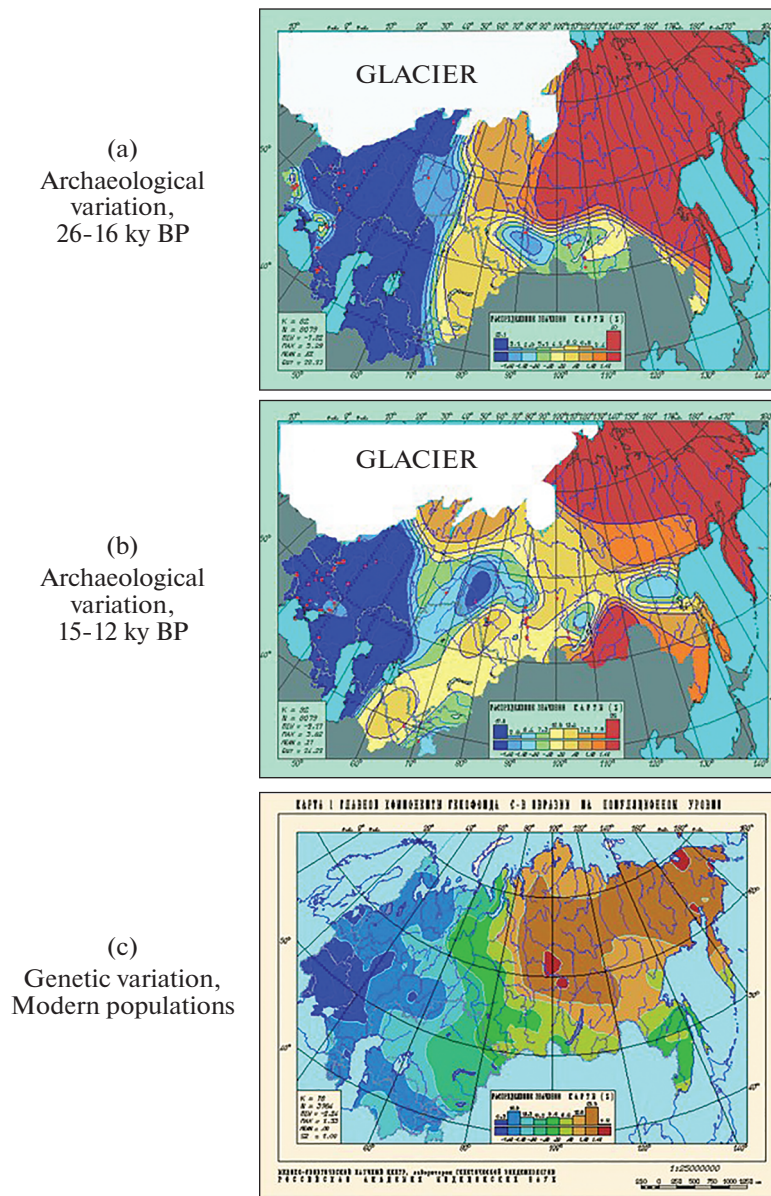


Fig. 3. Correspondence between archeological and genetic variation. a – map of the 1st PC of archeological data on Upper Paleolithic (26–16 ky BP) (modified from [21, 22]); b – map of the 1st PC of archeological data on Final Paleolithic (15–12 ky BP) (modified from [21, 22]); c – map of the 1st PC of genetic data on modern populations (modified from [23]).

aDNA has revealed that the Neanderthals were not the only distant cousin to breed with modern Homo sapiens, but that genetically different Denisovans also contributed to populations in East Asia and Oceania where they provide somewhere between 5 and 8% of the genes of the people of New Guinea. One of the more astonishing finds was that of a young woman who had a Denisovan father and a Neanderthal mother. In any event, archaeologists and physical anthropologists now recognize that there can be no coherent story of the expansion of the human race unless all three strands of their information can be brought together.

Spread of Agriculture

While archaeological evidence had long shown that farming was introduced into Europe from Southwest Asia, primarily Anatolia, there had been long serious disputes as to what extent it was solely the result of migrating farmers. As Europe was already populated by hunter-gathering Mesolithic communities, archaeologists sought to determine to what extent the spread of farming was due to generations of migrants of Anatolian ancestry and to what extent it involved acculturation by local Mesolithic populations. The standard model presumed that migration was the primary vector in Southeast Europe and possibly central Europe

but that the local component increased all around the periphery of Europe. The results of aDNA have largely confirmed this model with Balkan Neolithic populations closely resembling Northwest Anatolian [29], and the earliest Central European farmers were likewise primarily of Anatolian origin [30], and, perhaps more unexpected, the mitochondrial DNA of the earliest farmers in northern Europe (the TRB culture) were more closely related to the farmers who migrated from Anatolia than the local hunter-gatherers [31]. In France, however, there was increased evidence of WHG in the Paris Basin in the same area that Neolithic migration streams along the Danube and along the Mediterranean merged [32].

On the other hand, although we have seen aDNA research routinely refuting the retrospective extension of modern DNA patterns, the discrepancies have also stimulated more focused research. For example, the model of the earliest farming communities being entirely from Anatolia has been challenged where evidence drawn from both aDNA and modern DNA have suggested that some of the supposed markers of Anatolian farmer ancestry may have already been established in post-glacial Mediterranean Europe before the spread of farming [33].

Early Bronze Age Migrations

One of the most sensational discoveries of recent aDNA research has concerned the role of migrations from the steppelands of the Pontic-Caspian region both east as far as the Yenisei and Altai and as far west as Ireland. The samples were gathered to help resolve a two-hundred year old controversy as to the location of the homeland of the Indo-European language family. Although potential homelands had been identified all over Eurasia, most discourse had concerned three potential dates and locations: Anatolia, where it was associated with the dispersal of farming *c* 7000 BC, Greater Armenia where it was dated to *c* 5000 BC and largely argued on the basis of linguistics rather than archaeology, and the European steppelands *c* 4500–3000 BC where it was tied to a series of expansions both east across western Siberia and west into Central and eventually Western Europe. Although arguments involved many archaeological cultures, the two most prominent were the Corded Ware culture that extended from the Netherlands east to the Urals and the Yamnaya culture that stretched from the Urals west to the Danube. Those who supported a Steppe Homeland argued that the Corded Ware culture had its origin from the steppe cultures. While there were certain somewhat generic similarities between the two cultures, e.g., single grave burial under a kurgan, pastoral economy, use of corded decoration on pottery, even supporters of the Steppe Homeland did not argue for a major invasion but rather some form of acculturation to a new Yamnaya socio-economic package of strategies and values [34].

In 2015 two articles appeared to vindicate the Steppe model [5, 6]. They revealed that 75% of Corded Ware individuals carried the genetic signature identified in the Yamnaya culture which argued for a source population somewhere in the steppelands. Further research found that the steppe signature was known as far west as Britain [27] and Ireland [14] where it was recovered from burials of the Beaker culture that covered much of western Europe after the Corded Ware culture.

Prehistorians are still trying to deal with the impact of these results because they far exceed the expectations of archaeologists concerning their ability to identify migrations in the archaeological record. Stefan Burmeister has even suggested that it has become “obvious that archaeology has lost its previous methodological hold on investigating migration.” [35]. To explain the patterns of aDNA some have argued for a more protracted period of interactions between the steppelands and Central Europe rather than a massive migration while others have sought other explanations to deal with a situation in which they believe there is absolutely no archaeological evidence for either a migration from the steppe or for the derivation of the Corded Ware culture from the Yamnaya [36–38]. The genetic implications of aDNA on Britain and Ireland are even more extreme as an archaeological community, that had rejected migration as a serious vector for the appearance of the Beaker culture on the two islands, are now attempting to understand not how it could have happened but also how it resulted in a major replacement of the male genome in a period of only a few centuries. As Volker Heyd has remarked: “the aDNA results force us to reconsider; to question our own evidence and the methodology we apply...and to re-focus our interpretations” [36]. Archaeologists are having to come to grips with living in ‘interesting times’.

Origin of Ethno-linguistic Groups

Working with archaeologists and linguists, geneticists are now providing critical information on the origin of ethno-linguistic groups across the world. These can vary in complexity, especially when an ethnic group occupies a territory that has seen many passing footprints. A good example would be the Hungarians. Although they are one of the more recent peoples to arrive in their historical seats, settling there according to both historical and archaeological evidence *c* 900 CE, their long trek from the eastern Urals carried them across much of the same ground as previous settlers in their territory, the Huns and Avars, the last of whom are generally agreed to have made a more significant contribution to the peopling of Hungary than the later Magyars. Applying aDNA to the burials from the various earlier population groups as well as the earliest Magyar graves is slowly providing greater clarity to the diverse ethnic composition of each of Hungary’s pre-

vious settlers [39–41]. Mixing with the more numerous previous inhabitants of what is now Hungary has resulted in the dissolution of the genetic ancestry of the Magyars, so that modern Hungarians resemble the populations that currently surround them. However, the ancient mtDNA from the period right after the Magyar invasion finds its closest parallels among the Bashkirs— a Turkish-speaking people of the Urals.

PAST LIFEWAYS

Determination of Sex

The determination of the biological sex of an individual by physical anthropologists is often a probabilistic determination depending on the skeletal material available and the age of the individual. Ancient DNA now routinely recovers the sex through molecular data and should expand our ability to determine the sex of samples, especially among pre-adolescents who are usually too young to exhibit diagnostic sexual characteristics. It also draws attention to misidentifications made by physical anthropologists. In a recent publication [42] the sex of 132 individuals was determined by both skeletal and genetic evidence and this revealed 22 cases (17%) of conflicting identifications. Although physical anthropologists have recognized that distinguishing older females (post-menopause) from males is problematic, the sample did not reveal any bias in the identification of older females. Determination of sex, as with age, is one of the critical variables employed by archaeologists to assess the social organization of a society by burial remains.

Palaeodiet

Although archaeologists and palaeo-ecologists have a wide range of techniques for recovering past diets, e.g., palaeo-ethnobotany, stable isotope analysis, analysis of ancient DNA and other ancient molecules is now uniquely placed to help resolve some major issues of Eurasian palaeo-diets. This is particularly interesting with reference to the exploitation of livestock for secondary products. A popular model of European prehistoric economies argued that in the Early Neolithic livestock were raised exclusively for meat and that dairying was a later development of a so-called ‘Secondary Products’ Revolution [43]. Lipid analysis of Early Neolithic ceramics from Anatolia [44] to Ireland [45] has revealed that dairy products were routinely being exploited from the very beginning of the Neolithic. On the other hand, aDNA has so far shown no evidence of genes remitting lactose intolerance in the Neolithic and even in the Early Bronze Age cultures that are traditionally associated with increased pastoralism, the presence of genes that would enhance consumption of milk products has been found to occur at only very small levels [6, 14, 27, 46]. Of course, the processing of milk into cheese greatly reduces the problems of lactose intolerance but

the contrast between the archaeological and genetic evidence does raise interesting issues regarding how and why natural selection for the genes permitting milk consumption developed against a background of archaeological evidence that demonstrates the practice long preceded the rise of the appropriate genetic mutations.

Palaeopathology

Research into the evidence of pathogens from archaeological sites has greatly augmented our knowledge of ancient diseases, previously recoverable only when they left skeletal traces. Some recent examples include the recovery of evidence for bovine tuberculosis (*Mycobacterium bovis*) from Iron Age individuals from southern Siberia [47]; the discovery of traces of leprosy (*Mycobacterium leprae*) from Japan [48] to Ireland [49]; and the earliest emergence of plague (*Yersinia pestis*) in Late Neolithic Ukraine that appeared to spread with a massive incursion of steppe pastoralists into Central Europe [50].

Palaeodemography

The theoretical potential of aDNA to provide estimates of background populations, calculated on the basis of a measure of heterozygosity, the lower the value the more likely that the individual comes from a restricted breeding population that would be typical of hunter-gatherers. The time transect in Central Europe was the first study which clearly demonstrated that mitochondrial haplotype diversity (analog of heterozygosity which is applicable for haploid chromosomes) indeed increased with the start of the Neolithic [51]. The higher heterozygosity is typical for farmer societies because these have a much larger population size than hunter-gatherers, and thus are less affected by genetic drift (genetic drift often converts variation within a population – heterozygosity – into genetic variation between populations). The genetic drift depends on effective population size (N_e), and genetic methods allow an estimated N_e from the modern genomes. A well-known example is that size estimate of a population which came out of Africa was just a few thousand. Even though the absolute numbers might depend on the demographic parameters used for calculations, the relative numbers are more reliable and from genetic data we can see in which epochs/regions populations grew fast, or grew slow, or experienced a decay or were stable in size.

Knowing the population size is of course important for reconstructing the past lifeways, but one should take into account the non-trivial relation between total population size (which archeology is interested in) and the effective population size (estimated from genetic data). Though for many populations their effective size comprises from one quarter to one third of a total size, this conversion coefficient becomes different

when population size changes over generations (the bottleneck effect), or when only some males contribute their genes to the next generation, or when different families have quite unequal numbers of children. The detailed analysis of Eurasian Y-chromosomal lineages [52] determined that a huge portion of them disappeared around 8–4 ky BP and a number of new ones expanded right after this period. As there are no reasons to expect a dramatic population decline in the late Neolithic and Bronze Age (actually human population continued to grow in number during these periods), and most other factors affecting the genetic drift could be ruled out, the analysis concluded that it was a substantial change in social organization in the Bronze Age which resulted in the dramatic differences in the number of surviving children across different families/clans/demes: “A combination of culturally driven increased male variance in offspring number within demes and an increased male-specific variance among demes” [52]. This study exemplifies, that a genetic estimate of the effective population size does not necessarily describe the total population size but it can provide even more interesting evidence of the past demography.

It is possible that the greatest impact of genetics on archaeology may well lie in its future contribution to the reconstruction of the social behavior of past societies. Already there have been a number of studies that have targeted sites that can address longstanding archaeological problems. The Upper Palaeolithic site of Sungir, Russia, for example has yielded evidence that Upper Palaeolithic societies may have engaged in the same forms of reproductive strategies (exogamy, avoidance of inbreeding) as seen in modern hunter-gathering populations [53]. Similarly, a study concerning the emergence of the Yamnaya culture relies heavily on a discourse focusing on marriage networks indicated by aDNA [54].

On a more local level, aDNA analysis of a group of Corded Ware burials from Eulau, Germany, have verified the frequent assumption that people buried together are from the same family [55]. Another example from Koszyce, Poland, involved a mass grave belonging to the Globular Amphora culture in which all 15 individuals had been killed by a blow to the skull. The aDNA evidence indicated that the deceased consisted of four nuclear families and that mothers and their children had been buried adjacent to one another, obviously by people who recognized their relationship. Moreover, the fact that the graves contained males who were related but females who were not provided evidence that the society was organized along patrilineal lines [56]. In contrast to these family groups, aDNA from five burials from the same chamber of a megalithic tomb at Carrowkeel, Ireland, revealed no evidence of close kinship between the individuals [57]. And in another genetics-aided examination of a large Neolithic cemetery at Gurgy, France, where mtDNA indicated that the deceased

were derived from three different sources (Danubian Farmers, Mediterranean Farmers, and local Hunter-Gatherers) no correlations with the variety of mortuary traits could be determined, other than perhaps a greater tendency for those of HG descent to be buried on the periphery and with more exotic grave goods [32].

These are isolated and somewhat fortuitous examples of the future potential of genetics to inform on the social organization of past societies. Much will depend on to what extent the genetic profiles of ancient populations will be able to go beyond immediate kinship relations and provide a broader picture of the complex of breeding networks that that might help to define the more widespread social relationships that existed between political entities.

LINGUISTICS

The impact of genetics on linguistics is closely tied to archaeological narratives of population dispersals but here there is probably a subtle difference. It should be universally acknowledged that the spread of languages and the spread of genes may be two entirely unassociated processes [58] because the language one speaks can be entirely independent of one’s genetic history (you can choose your language but not your genes). On the other hand, there is perhaps an even larger rejection of any one-to-one association between a language and material culture so while genetics may be an unsuitable proxy for language affiliation, it may still be regarded as a better substitute than what is offered by archaeology.

Massive Migration or Language Shift: Genetics Judges

Language dispersals are generally explained by two mechanisms: migration of people where a significant replacement of a prior population results in the spread of the immigrants’ target language. The second process involves language shift, the adoption by a local population of an intrusive target language. Since the evidence of genetics not only can reveal rather clear evidence that a migration has taken place but it can also reveal the creation of new marriage networks that might enhance the spread of the target language by language shift, genetic evidence may often be more convincing than the type of cultural evidence available to an archaeologist.

One of the genetic methods used to distinguish between the demic spread of languages and language shift is a comparison whether a genetic pattern is more correlated with a linguistic one or simply with geography. To perform such a test, one needs a matrix of linguistic distances between languages (lexicostatistics provides most reliable ones though attempts were also made to use formalized grammar databases) and a matrix of genetic distances between the same populations. The third matrix – geographic distances – is the easiest one to obtain. In many regions of Eurasia,

the pattern of genetic similarities between populations better correlates with their geographic proximity rather than with the relatedness of their languages. But there are a few regions with the opposite trend, for example areas populated by North Caucasian and Tungusic speakers. In the North Caucasus, the genetic matrix perfectly correlates with the linguistic one, and even the genetic tree of population relatedness is nearly the same as the tree of their languages [59]. Such a coincidence is possible when the demographic history of population splits was equally reflected in splits of languages and splits of gene pools – a realistic assumption provided that after the split the daughter populations have not genetically mixed with external groups, or all populations received the same degree of admixture. This was likely the case in the narrow and isolated mountainous area of the North Caucasus. However, this assumption is certainly not correct for the Slavonic-speaking populations which rapidly spread across half of Europe and assimilated the previous inhabitants of the Balkans, Northeast Europe and the Volga area. As a result, the genetics of Slavic populations reflects largely the genetics of pre-Slavic populations and exhibit therefore only a weak correlation with the tree of Slavonic languages [60]. However, even in the case of language shift by elite dominance this elite group spread their language along with their genes, and this small proportion of genes that arrived with the language can be potentially traced. For example, one can identify the set of genetic markers (IBD haplotypes) that spread along with the Turkic languages [61].

The aforementioned studies were done on modern populations, and aDNA is at least equally promising. Still the current impact of aDNA research on language dispersals is largely experienced through archaeological models and linguists are only just coming to grips with some of the wider implications. Here are a few.

The Indo-European Case

Just as the genetic evidence for a steppe homeland appeared to weaken a popular theory (among archaeologists more than linguists) that the Indo-European languages spread from an Anatolian homeland with the spread of farming and the AF genetic signature, a new complication arose: the steppe signal that is found from Ireland to the Yenisei comprises an admixture of EHG and CHG. Such an admixture would appear to involve two deep sources that should have developed separately over the course of thousands of years; in short, there is no reason to believe that the two components spoke closely related languages or even belonged to the same language families. Such a model suggested that Proto-Indo-European may have originated out of the merger of two very different language families, a theory that had once had been suggested by several linguists but had never attained anything

remotely resembling consensus [62]. If one does not accept an “admixture language” then the natural question remains: did Proto-Indo-European evolve out of language spoken by EHG or out of language spoken by CHG? So genetics has pushed the current homeland debate into several camps: those who seek the homeland either in the southern Caucasus or Iran (CHG) and those who locate it in the steppelands north of the Caucasus and Caspian Sea (EHG). Those who prefer a southern homeland look to proposed contacts between Indo-European and the Kartvelian and Semitic languages [63–65] while those who prefer the steppe hypothesis point to the evidence that Indo-European is most closely related to Uralic, which should pitch its origin nearer the Urals [54, 66, 67]. The matter is complicated enormously by the fact that CHG cannot possibly represent the signature of a single language family as it is found over a broad area from the Caucasus to the southern Zagros that encompassed the territory of a whole series of other language families, e.g., Hurro-Urartian, Elamite, Kartvelian, and one must devise a credible model of how the Indo-European-speaking segment of CHG found its way north of the Caucasus [68].

The Altaic Case

The somewhat similar interrelation between genetic and linguistic evidence can be seen in another half of Eurasia where Altaic languages are spoken. Though the limited number of aDNA samples available to date from this region does not permit one to verify the principal migrations, there is a large area of so-called “Amur basin ancestry” with a supposed proto-Altaic segment within it.

The Altaic language family includes the Turkic, Mongolic, Tungus-Manchu, Korean and Japonic language groups. Though this view is not generally accepted, most issues become resolved when ancient borrowings between the proto-languages of these groups are identified [69, 70]. After all, Indo-Europeanists are not confused by the significant layer of borrowings between Germanic and Celtic, or Germanic and Slavic protolanguages.

At present, two groups of linguists – in Moscow and in Jena – are actually involved in the detailed study of the Altaic family and its constituent groups and the chronology of the collapse of groups obtained on glottochronological grounds practically coincides. For example, Proto-Altaic split into Korean-Japanese, Manchu-Tungusic, and Turko-Mongolic families c8 – 7 ky BP; Turkic and Mongolic diverged about 6.5 ky BP; the Bulgar branch split from Turkic 2.5 ky BP. Some of these dates are confirmed by borrowings to/from well dated historical stages of ancient and middle Chinese.

As for the ancestral homeland for Proto-Altaic, attempts to locate it were made on the basis of the

reconstruction of the vocabulary of the natural environment [71–76]. It is interesting that applying the same approaches to Proto-Altaic and Proto-Indo-European resulted in different environments for these proto-populations. Proto-Altaic speakers lived in (not very high) mountains, arid and dusty steppes. There are many terms for quick-flowing rivers, words for large rivers (but not for the sea), coniferous trees, oaks, a number of narrow-foliated trees, and wild cereals, which altogether indicate the border zone of the southern taiga and the steppes [71, 72]. A careful consideration of the agricultural terminology does not confirm the hypothesis of Proto-Altaic semi-pastoralism; instead, there are names for pigs and cultural (or pre-cultural) plants, including millet.

Like the hypothesis linking the expansion of the Indo-European languages with the spread of farming in West Eurasia, Martine Robbeets ties the distribution of Altaic languages to the dispersal of millet farming from Manchuria. “In the area of southern Manchuria and eastern Inner Mongolia, the predominant basis of life since the 7th millennium BC has been millet agriculture, supplemented by fishing, hunting and gathering in the surrounding woodlands. In the western part of this region, which is ecologically transitional towards Mongolia, nomadic pastoralism developed as an innovation in the first millennium BC” [75]. Specifically, she connects the Altaic family with two subsequent cultures – the Xinglongwa (ca. 6200–5400 BC) and the Zhaobaogou (ca. 5400–4500 BC). This is the area of the Liao river valley, and that is highly suitable for the landscape reconstructed from the Proto-Altaic vocabulary [72]; moreover, this culture used two species of millet and began to domesticate pigs.

Here, in contrast to the Indo-European case, there is not much debate about either the date nor the place of the proto-Altaic homeland. However, the mechanisms for the spread of Altaic languages over the area comparable with the Indo-European one are still unclear, and genetics has a potential to shed light on this.

On the general overview, the distribution of the genetic features of Altaic speakers is much more related to the geographical proximity of population groups than to the genealogical classification of their languages. This coincides well with the migration model through language replacement (elite dominance), when only a few genes remain from the original proto-language-speaking group, migrating along with the languages [75]. However, when the Turkic group – for which many examples of language replacement are documented historically and genetically – is omitted from the analysis, the genome-wide variation within the remaining Mongolic and Tungusic groups starts to better resemble linguistics rather than geography (Balanovsky, Dybo, et al., in preparation).

The gene pool of Mongolic, Tungusic, and Turkic speakers forms a Trans-Eurasian cline of mixed West and East Eurasian ancestry, in which the Eastern Eurasian component is nearly identical to the contemporary populations from the Amur river basin [77]. Notably, this genetic component was present in this area for at least 7 thousand years [78] and likely predates the first split of the Altaic family. This might mean either that the Amur basin groups shifted their language to Tungusic or that there was classical demic diffusion of Tungusic speakers but the ancestries of incomers and aboriginals were too similar to be distinguished with current genetic techniques.

The Farming/Language Dispersal Hypothesis posits that a language family owes its dispersal to the growth of a speaker population through the practice of agriculture; the population growth steadily pushed the new farmers and their language into wider territories, displacing the languages of preexisting hunter-gatherer populations. This model does not fit very well, because farming was not an important component in most areas populated by Altaic speakers. One can suppose that only the beginning of the Altaic dispersal was initiated by agricultural development, and later the language was transferred to the West by equestrian elites; further genetic data from both, ancient and modern DNA may help to reconstruct the migrations in more detail.

Methodological Cases

The integration of linguistic and genetic studies is partly driven by some methodological similarities. For example, both linguistics and genetics compute distances between languages or populations, while for archeology it is much more problematic to quantitatively measure similarities between cultures. Bayesian methods which are popular in genetics for phylogenetic tree construction are now widely used by linguists for creating trees of languages.

Genetics and linguistics are also similar in the way they date events: by linguistic diversity between modern languages (assuming a stable word substitution rate) or by genetic diversity between modern genetic lineages (assuming a stable mutation rate). Coincidence of genetic and linguistic dates make the reconstruction of the demographic history much more reliable, while the lack of coincidence indicates a likely language shift event. In addition to these retrospective approaches, the application of ancient DNA allows us to use the dated episodes of genetic migration to estimate the various dates of language branching. For example, as the Corded Ware culture expanded into Northern Europe by c 3000 BC and this region is the presumed homeland of the Germanic languages, a presumption is sometimes made that some form of Germanic was spoken in south Scandinavia from c 3000 BC onwards. However, linguists have rarely imagined that the Germanic language itself came into

existence much before 500 BC leaving a 2500 year gap between the putative genetic event and the presumed formation of Proto-Germanic after which it expanded over northern Europe. While one might feel confident that the genetic evidence indicates that some form of Indo-European was spoken in Northern Europe by c 3000 BC, what that form was and whether it was in any way directly ancestral to Proto-Germanic seems questionable and it is entirely possible that Proto-Germanic arose elsewhere and only assumed its geographical position and archaeological identity at a later date. The same goes for the presumption that the genomic change associated with the Beaker culture is equivalent to the expansion of the Celtic languages. Here again we find that setting Proto-Celtic to c 2500 BC over an area that extended from northern Italy to Ireland hardly fits most models for either the date of Proto-Celtic nor the incredible similarity between the earliest Irish inscriptions and those of southern France 2500–3000 years later. The problems here is that while aDNA can provide a firm chronology of gene flow and migrations, the linguistic entities that one attempts to associate with these events may not be chronologically compatible. In the case of Germanic and Celtic, the stimulus for the formation of the target language may well be much further downstream of the initial genetic events and may involve processes of language shift that carry little if any clear genetic trails.

Linguistic, genetic, and archeological evidences come together when considering the spread of dairy pastoralism to the eastern Eurasian steppe. The archeological data and molecular analysis of ceramics indicated the arrival of dairy technologies as early as the 14th century BC with the invasion of Western cultures, but ancient DNA indicated that it was not associated with a massive migration and population replacement [79]. This agrees well with the linguistic data where the Turkic vocabulary related to dairy farming is mainly Eastern Iranian but except for some borrowing, the autochthonous languages of the eastern steppe remained the same and have not been replaced [80].

Integration

In recent years, there is a tendency of closer collaboration between linguists, archeologists, and geneticists. The numerous joint projects, interdisciplinary conferences (two of which were named LAG – Linguistics, Archaeology, and Genetics) and joint papers – like the present review written by archeologists, linguists, and geneticists – gives a hope for further successful interdisciplinary studies.

Nevertheless, the keyinterdisciplinary problems, like Indo-European (Altaic, Uralic) homeland and dispersal, gene-language coevolution, or demographic history have been baffling linguists, archeologists, and anthropologists for over two centuries now and while genetics has now provided a powerful new tool, a final solution still remains to be found.

This study was carried out within the state assignment of Ministry of Science and Higher Education of the Russian Federation for the Research Center for Medical Genetics (impact of genetics on archaeology) and by the Russian Science Foundation project 17-14-01345 (impact of genetics on linguistics) and by the Russian Science Foundation project 18-18-00501 (collection and analysis of the Altaic linguistic materials).

This paper does not include experiments involved animals.

This paper does not contain experimental data on humans.

Authors declare no conflict of interests.

REFERENCES

1. Renfrew, C., *Archaeogenetics: Towards a Population Prehistory of Europe*, *Archaeogenetics: DNA and the Population Prehistory of Europe*, Renfrew, C. and Boyle, K., Eds., Cambridge: 2000.
2. Marks, J., The origins of anthropological genetics, *Curr. Anthropol.*, 2012, vol. 53, suppl. 5, pp. S161–S172.
3. Howells, W.W., *Analyses of patterns of variation in crania of recent man, The Functional and Evolutional Biology of Primates*, Tuttle, R., Ed., Chicago: 1972, pp. 123–151.
4. Menk, R., A synopsis of the physical anthropology of the Corded Ware complex on the background of the expansion of the Kurgan cultures, *J. Indo-Eur. Stud.*, 1980, vol. 8, nos. 3–4, pp. 361–392.
5. Haak, W., Lazaridis, I., Patterson, N., et al., Massive migration from the steppe was a source for Indo-European languages in Europe, *Nature*, 2015, vol. 522, no. 7555, pp. 207–211. <https://doi.org/10.1038/nature14317>
6. Allentoft, M.E., Sikora, M., Sjögren, K.G., et al., Population genomics of Bronze Age Eurasia, *Nature*, 2015, vol. 522, no. 7555, pp. 167–172. <https://doi.org/10.1038/nature14507>
7. Deryabin, V.E., *Etnicheskaya antropologia sovremennih slavyanskikh narodov Vostochnoy Evropy* (Physical Anthropology of Modern Slavic Populations from East Europe), Moscow: VINITI, 2002.
8. <http://www.gg-base.org>. Accessed June 1, 2019
9. Poesche, T., *Die Arier: ein Beitrag zur historischen Anthropologie*, Jena, 1878.
10. Penka, K., *Origines Ariacae: linguistisch-ethnologische Untersuchungen zur ältesten Geschichte der arischen Völker und Sprachen*, Vienna, 1883.
11. Penka, K., *Die Herkunft der Arier: neue Beiträge zur historischen Anthropologie der europäischen Völker*, Vienna, 1886.
12. Day, J., *Indo-European Origins: the Anthropological Evidence*, Washington, DC, 2001.
13. Olalde, I., Allentoft, M.E., Sánchez-Quinto, F., et al., Derived immune and ancestral pigmentation alleles in a 7000-year-old Mesolithic European, *Nature*, 2014, vol. 507, no. 7491, pp. 225–228. <https://doi.org/10.1038/nature12960>

14. Cassidy, L. M., Martiniano, R., Murphy, E. M., et al., Neolithic and Bronze Age migration to Ireland and the establishment of the insular Atlantic genome, *Proc. Natl. Acad. Sci. U.S.A.*, 2016, vol. 113, no. 2, pp. 368–373.
15. Lazaridis, I., Mittnik, A., Patterson, N., et al., Genetic origins of the Minoans and Mycenaeans, *Nature*, 2014, 2017, vol. 548, no. 7666, pp. 214–218. .
<https://doi.org/10.1038/nature23310>
16. Sokal, R. R. and Livshits, G., Geographic variation of six dermatological traits in Eurasia, *Am. J. Phys. Anthropol.*, 1993, pp. 393–407.
17. Sokal, R. R., Jantz, R. L. and Thomson, B. A., Dermatological variation in Europe, *Am. J. Phys. Anthropol.*, 1996, pp. 35–47.
18. Jantz, R. L., Anthropological dermatoglyphic research., *Annu. Rev. Anthropol.*, 1987, vol. 16, pp. 161–177.
19. Åstrom, P. and Eriksson, S.A., Fingerprints and the Indo-Europeans in Greece, *Acta of the 2nd International Colloquium on Aegean Prehistory: the First Arrival of the Indo-European Elements in Greece*, Arditis, E., Ed., Athens, 1983, pp. 72–75.
20. Cavalli-Sforza, L.L., Menozzi, P., Piazza, A., *The History and Geography of Human Genes*, Princeton: Princeton University Press, 1994.
21. Grekhova, L.V., Balanovska, E.V., and Rychkov, Yu.G., Evaluation of the computer technology of regional atlases: the late Paleolithic of Northern Eurasia, in *Gumanitarnaya nauka v Rossii: Sorosovskiy laureaty (istoriya. archeologiya. kulturnaya antropologiya i etnografiya)* (Humanitarian Science in Russia: Soros Laureates (History, Archeology, Cultural Anthropology, and Ethnography)), Moscow: 1996, pp. 286–304.
22. Balanovska, E.V., Grekhova, L.V., and Rychkov, Yu.G., Computer gene geography and archeology: methods for cartographic modeling of the material culture dispersal, in *Gorizonty antropologii* (Horizons of Anthropology), Moscow: Russkiimir, 1997, pp.54–62.
23. Rychkov, Yu.G. and Balanovskaia, E.V., Gene pool and gene geography of USSR populations, *Russ. J. Genet.*, 1992, vol. 28, pp.52–75.
24. Sykes, B., *Seven Daughters of Eve*, Norton, 2001.
25. Oppenheimer, S., *The Origins of the British*, Constable, 2006.
26. Balaresque, P., Bowden, G.R., Adams, S.M., et al., A predominantly Neolithic origin for European paternal lineages, *PLoS Biol.*, 2010, 8, no. 1. e1000285.
<https://doi.org/10.1371/journal.pbio.1000285>
27. Olalde, I., Brace, S., Allentoft, M.E., et al., The Beaker phenomenon and the genomic transformation of northwest Europe, *Nature*, 2018, vol. 555, no. 7695, pp. 190–196.
<https://doi.org/10.1038/nature25738>
28. Saag, L., Laneman, M., Varul, L., et al., The arrival of Siberian ancestry connecting the eastern Baltic to Uralic speakers further East, *Curr. Biol.*, 2019, vol. 29, no. 10, pp. 1701–1711.
<https://doi.org/10.1016/j.cub.2019.04.026>
29. Mathieson, I., Alpaslan-Roodenberg, S., Posth, C., et al., The genomic history of southeastern Europe, *Nature*, 2018, vol. 555, no. 7695, pp. 197–203.
<https://doi.org/10.1038/nature25778>
30. Haak, W., Balanovsky, O., Sanchez, J. J., et al., Ancient DNA from European Early Neolithic farmers reveals their Near Eastern affinities, *PLoS Biol.*, 2010, vol. 8, no. 11. e1000536.
31. Malmström, H., Linderholm, A., Skoglund, P., et al., Ancient mitochondrial DNA from the northern fringe of the Neolithic farming expansion in Europe sheds light on the dispersion process, *Philos. Trans. R. Soc., B*, 2015, vol. 370, no. 1660, p. 20130373.
<https://doi.org/10.1098/rstb.2013.0373>
32. Le Roy, M., Rivollat, M., Mendisco, F., et al., Distinct ancestries for similar funerary practices? A GIS analysis comparing funerary, osteological and DNA data from the Middle Neolithic necropolis Gurgy “Les Noisats” (Yonne, France), *J. Archaeol. Sci.*, 2016, vol. 73, pp. 45–54,
<https://doi.org/10.1016/j.jas.2016.07.003>
33. Pereira, J.B., Costa, M.D., Vieira, D., et al., Reconciling evidence from ancient and contemporary genomes: a major source for the European Neolithic within Mediterranean Europe, *Proc. Biol. Sci.*, 2017, vol. 284, no. 1851, pii: 20161976.
<https://doi.org/10.1098/rspb.2016.1976>
34. Harrison, R. and Heyd, V., The transformation of Europe in the Third Millennium BC: the example of ‘Le Petit-Chasseur I + II’ (Sion, Valais, Switzerland). *Prähistorische Z.*, 2007, vol. 82, pp. 129–214.
35. Burmeister, S., Archaeological research on migration as a multidisciplinary challenge, *Medieval Worlds*, 2016, vol. 4, pp. 42–64.
36. Heyd, V., Kossina’s smile, *Antiquity*, 2017, vol. 91, pp. 348–359.
37. Kristiansen, K., Allentoft, M.E., Frei, K.M., et al., Rethorising mobility and the formation of culture and language among the Corded Ware culture in Europe, *Antiquity*, 2017, vol. 91, pp. 334–347.
38. Ivanova, S.V., Nikitin, A.G., and Kiosak, D.V., Pendulum migrations in Circum-Pontic steppe and Central Europe during the Paleometal Epoch and the problem of the Pit-Grave Culture origin, *Arkheol. Davnya Ist. Ukr.*, 2018, vol. 1, pp. 101–146.
39. Neparáczki, E., Maróti, Z., Kalmar, T., et al., Y-chromosome haplogroups from Hun, Avar and conquering Hungarian period nomadic people of the Carpathian Basin, *Preprint of bioRxiv*. <https://www.biorxiv.org/content/10.1101/597997v1>. Accessed July 15, 2017.
40. Post, H., Németh, E., Klima, L., et al., Y-chromosomal connection between Hungarians and geographically distant populations of the Ural Mountain region and West Siberia, *Sci. Rep.*, 2019, vol. 9, p. 7786.
<https://doi.org/10.1038/s41598-019-44272-6>
41. Csősz, A., Szécsényi-Nagy, A., and Csákyová, V., Maternal genetic ancestry and legacy of 10th century AD Hungarians, *Sci. Rep.*, 2016, vol. 6, p. 33446.
<https://doi.org/10.1038/srep33446>

42. Narasmhan, V.M., Patterson, N., Moorjani, P., et al., The formation of human populations in South and Central Asia, *Science*, 2019, vol. 365, no. 6457. <https://doi.org/10.1126/science.aat7487>
43. Sherratt, A., *Plough and pastoralism: aspects of the Secondary Products Revolution, Patterns of the Past: Studies in Honour of David Clarke*, Hodder, I., Isac, G., and Hammond, N., Eds., Cambridge, 1981.
44. Evershed, R.P., Payne, S., Sherratt, A.G., et al., Earliest date for milk use in the Near East and southeastern Europe linked to cattle raising, *Nature*, 2008, vol. 455, pp. 528–531.
45. Smyth, J. and Evershed, R., The molecules of meals: new insight into Neolithic foodways, *Proc. R. Irish Acad.*, 2015, vol. 115C, pp. 1–20.
46. Gamba, et al., 2014;
47. Murphy, E.M., Chistov, Y.K., Hopkins, R., et al., Tuberculosis among Iron Age individuals from Tyva, South Siberia: palaeopathological and biomolecular findings, *J. Archaeol. Sci.*, 2009, vol. 36, pp. 2029–2038.
48. Suzuki K., Saso, A., and Hoshino, K., Paleopathological evidence and detection of *Mycobacterium leprae* DNA from archaeological skeletal remains of Nabe-kaburi (Head-Covered with Iron Pots) Burials in Japan, *PLoS One*, 2014, vol. 9, no. 2. e88356. <https://doi.org/10.1371/journal.pone.0088356>
49. Taylor, G. M., Murphy, E. M., Mendum, T. A., et al., Leprosy at the edge of Europe—biomolecular, isotopic and osteoarchaeological findings from medieval Ireland, *PLoS One*, 2018, vol. 13, no. 12. e0209495. <https://doi.org/10.1371/journal.pone.0209495>
50. Rascovan, N., Sjögren, K-G., Kristiansen, K., et al., Emergence and spread of basal lineages of *Yersinia pestis* during the Neolithic Decline, *Cell*, 2019, vol. 176, pp. 1–11.
51. Brandt, G., Haak, W., and Adler, C.J., Ancient DNA reveals key stages in the formation of Central European mitochondrial genetic diversity, *Science*, 2013, vol. 342, no. 6155, pp. 257–261. <https://doi.org/10.1126/science.1241844>
52. Karmin, M., Saag, L., Vicente, M., et al., A recent bottleneck of Y chromosome diversity coincides with a global change in culture, *Genome Res.*, 2015, vol. 25, no. 4, pp. 459–466. <https://doi.org/10.1101/gr.186684.114>
53. Sikora, M., Seguin-Orlando, A., Sousa, V.C., et al., Ancient genomes show social and reproductive behavior of early Upper Palaeolithic foragers, *Science*, 2017, vol. 358, no. 6363, pp. 659–662. <https://doi.org/10.1126/science.aao1807>
54. Anthony, D., Archaeology, genetics and language in the steppes: a comment on Bomhard, *J. Indo-Eur. Stud.*, 2019, vol. 47, nos. 1–2, pp. 175–198.
55. Haak, W., Brandt, G., de Jong, H.N., et al., Ancient DNA, strontium isotopes, and osteological analyses shed light on social and kinship organization of the Later Stone Age, *Proc. Natl. Acad. Sci. U.S.A.*, 2008, vol. 105, no. 47. <https://doi.org/10.1073/pnas.0807592105>
56. Schroeder, H., Margaryan, A., Szmyt, M., et al., Unraveling ancestry, kinship, and violence in a Late Neolithic mass grave, *Proc. Natl. Acad. Sci. U.S.A.*, 2019, vol. 116, no. 22, pp. 10705–10710. <https://doi.org/10.1073/pnas.1820210116>
57. Kador, T., Cassidy, L., Geber, J., et al., Rites of passage: mortuary practice, population dynamics, and chronology at the Carrowkeel passage tomb complex, Co. Sligo, Ireland, *Proc. Prehist. Soc.*, 2018.
58. Campbell, L., Do languages and genes correlate? Some methodological issues, *Language Dynamics and Change*, 2015, vol. 5, pp. 202–226.
59. Balanovsky, O., Dibirova, K., Dybo, A., et al., Parallel evolution of genes and languages in the Caucasus Region, *Mol. Biol. Evol.*, 2011, vol. 28, no. 10, pp. 2905–2920. <https://doi.org/10.1093/molbev/msr126>
60. Kushniarevich, A., Utevska, O., Chuhryaeva, M., et al., Genetic heritage of the Balto-Slavic speaking populations: a synthesis of autosomal, mitochondrial and Y-chromosomal data, *PLoS One*, 2015, vol. 10, no. 9. e0135820. <https://doi.org/10.1371/journal.pone.0135820>
61. Yunusbayev, B., Metspalu, M., Metspalu, E., et al., The Genetic legacy of the expansion of Turkic-speaking Nomads across Eurasia, *PLoS Genet.*, 2015, vol. 11, no. 4. e1005068. <https://doi.org/10.1371/journal.pgen.1005068>
62. Mallory, J.P., The Caucasian substrate hypothesis, *J. Indo-Eur. Stud.*, 2019, vol. 47, nos. 1–2, pp. 1–8.
63. Gamkrelidze, T. and Ivanov, V. *Indo-European and the Indo-Europeans*, Berlin, Mouton de Gruyter, 1995.
64. Grigoriev, S., *Drevnie indoyevropeitsy* (Ancient Indo-Europeans), Chelyabinsk: Rifei, 2002.
65. Kozintsev, A., Proto-Indo-Europeans: the prologue, *J. Indo-Eur. Stud.*, 2019 (in press).
66. Kortlandt, F. *Studies in Germanic, Indo-European and Indo-Uralic*, 2010. Amsterdam: Rodopi, 2010.
67. Kassian, A., Zhivlov, M., Starostin, G., Proto-Indo-European-Uralic comparison from the probabilistic point of view, *J. Indo-Eur. Stud.*, 2015, vol. 43, pp. 301–310.
68. Mallory, J., *Speculations on the Neolithic origins of the language families of Southwest Asia, Usque ad Radices: Indo-European Studies in Honour of Birgit Anette Olsen*, Simmelkjaer, H., Ed., Copenhagen: Museum Tusulanum, 2017, pp. 503–516,
69. Dybo, A., Early contacts of Turks and Proto-Turkic reconstruction, *Tatarica*, 2014, no. 2, pp. 7–17;
70. Schönig, C., *Turko-Mongolic relations, The Mongolic Languages*, Janhunen, L., Ed., New York, 2005, pp. 405–406.
71. Dybo, A., Indo-Europeans and Altaians through the linguistic reconstruction, in: *Kinship in the Altaic World* (Proc. 48th PIAC), Wiesbaden, 2006, pp. 87–102.
72. Dybo, A., Language and archeology: some methodological problems: Indo-European and Altaic landscapes, *J. Lang. Relat.*, 2013, no. 9, pp. 69–92.

73. Starostin, S., Altaic loans in old Chinese, *Past Human Migrations in East Asia: Matching Archaeology, Linguistics and Genetics*, Sanchez-Mazas, Ed., London: Routledge, 2008, pp. 254–262.
74. Robbeets, M., Proto-Trans-Eurasian: where and when? *Man India*, 2016, vol. 95, no. 4, pp. 921–946.
75. Robbeets, M., The language of the Transeurasian farmers, in: *Language Dispersal beyond Farming*, John Benjamins, 2017, pp. 93–121.
76. Robbeets, M., Austronesian influence and Transeurasian ancestry in Japanese: a case of farming/language dispersal, in *Language Dynamics and Change*, 2017, pp. 210–251.
77. Jeong, C., Balanovsky, O., Lukiyanova, E., et al., The genetic history of admixture across inner Eurasia, *Nat. Ecol. Evol.*, 2019, vol. 3, pp. 966–976.
78. Siska, V., Jones, E.R., Jeon, S., et al., Genome-wide data from two early Neolithic East Asian individuals dating to 7700 years ago, *Sci. Adv.*, 2017, vol. 3, no. 2, e1601877. <https://doi.org/10.1126/sciadv.1601877>
79. Jeong, C., Wilkin, S., Amgalantugs, T., et al., Bronze Age population dynamics and the rise of dairy pastoralism on the eastern Eurasian steppe, *Proc. Natl. Acad. Sci. U.S.A.*, 2018, vol. 115, no. 48, pp. 11248–11255. <https://doi.org/10.1073/pnas.1813608115>
80. Dybo, A.V., *Lingvističeskie contacti rannih T'urkov, Proto-T'urkskiy period* (Language Contacts of Early Turks: the Proto-Turkic Period), Moscow: 2007.

SPELL: 1. lifeways