NATIONAL ACADEMY OF SCIENCES REPUBLIC OF ARMENIA
INSTITUTE OF MOLECULAR BIOLOGY

RECONSTRUCTION OF HUMAN NEOLITHIC MIGRATIONS IN THE ARMENIAN HIGHLAND BASED ON GENETIC DATA

BY

ANAHIT HOVHANNISYAN

DISSERTATION

SUBMITTED FOR THE DEGREE
OF CANDIDATE OF BIOLOGICAL SCIENCES (PhD)
(03.00.15 – Genetics)

Supervisor: Prof. Levon Yepiskoposyan

YEREVAN - 2016

TABLE OF CONTENTS

ABBREVIATIONS

aDNA – ancient DNA

AMH – anatomically modern humans

BC – before Christ

CDM – cultural diffusion model

CHG – Caucasus hunter-gatherers

CV – cross-validation index

DDM – demic diffusion model

EF – early farmers

EHG – Eastern hunter-gatherers

HVS-I – hypervariable segment I

indels – insertion/deletion polymorphisms

KYA – kilo years ago

LPK – Linear Pottery Culture

LD – linkage disequilibrium

LGM – Last Glacial Maxium

LL – log-likelihood

MDS – Multidimensional Scaling

mtDNA – mitochondrial DNA

NGS – next-generation sequencing

NRY – non-recombining region of the Y chromosome

PCA - principal component analysis

PCoA – principal coordinate analysis

PPNB – Pre-Pottery Neolithic B

SNPs – single nucleotide polymorphisms

STRs – short tandem repeats

TMRCA – time to the most recent common ancestor

UEP – unique event polymorphisms

WHG – Western hunter-gatherers

INTRODUCTION

Problem statement. The origin of farming and the nature of the Neolithic agricultural migration as a primary issue in the modern human colonization of the globe have been a major research focus for scholars of various disciplines. The keen scientific interest to this issue could be explained by a number of profound changes that accompanied the transition from hunter-gathering to farming [Renfrew C, 1996; Diamond J and Bellwood P, 2003; Bellwood P, 2013]. In particular, the adoption of agriculture drove dramatic human population growth and caused a human dispersal out of the regions of agricultural origin [Diamond J and Bellwood P, 2003].

The Fertile Crescent, a region of southwest Asia comprising the valleys of the Tigris, Euphrates and Jordan rivers with their adjacent hilly flanks, is widely recognized as the cradle of agriculture, where first settled farming communities are thought to have originated around 10 KYA [Ammerman AJ and Cavalli-Sforza LL, 1984; Riehl S, et al., 2013]. From the region, agriculture spread towards various destinations within Eurasian continent, including Europe and the Caucasus. In particular, whether this advancement to Europe was due to the migration of the Near-Eastern farmers remains the subject of intensive debates. At present, the only consensus has beed achieved on the complex nature of such a historical event, considering it as one of the crucial demographic processes in the peopling of Europe [Diamond J and Bellwood P, 2003; Barbujani G and Chikhi L, 2006; Bellwood P, 2013].

Numerous archaeological, archaeobotanical and linguistic studies have been conducted to understand the mechanisms and chronology by which the changeover from foraging to food production had occurred [Ammerman AJ and Cavalli-Sforza LL, 1971; College S, et al., 2004; Bellwood P, 2013]. With the recent rapid development of molecular techniques, genetic studies based on both modern and ancient DNA data have been extensively applied to disentangle such a crucial question in human population history [Menozzi P, et al., 1978; Cavalli-Sforza LL, et

al., 1994; Sjödin P and François O, 2011]. So far, much uncertainty is associated with the reconstruction of the routes of migration for the first farmers from the Near East. In the context of Europe, it has been suggested that the initial farmer dispersal from the Fertile Crescent reached Anatolia about 8 KYA [Bellwood P, 2013]. Later, along the Anatolian coast or by maritime routes from the Levant, the agricultural migration extended to Crete and the mainland Greece (6,5-7 KYA), from where it spread all over Europe [Bintliff J, 2012; Bellwood P, 2013]. However, these studies did not consider the populations from the eastern regions of modern Turkey and the South Caucasus, which roughly correspond to the boundaries of the Armenian Highland. Moreover, hitherto, the question of where the bifurcation of divergently directed paths of agricultural movement from the Near East towards Europe and the Caucasus took place has not been estimated and remains least understood.

In this context, hospitable climatic conditions and the key geographic position of the Armenian Highland as the central point of connection for various ancient civilizations, suggest that this region may have served as a conduit for several waves of expansion of the first agriculturalists from the Near East to Europe and the North Caucasus [Lang DM, 1980; Dolukhanov P, et al., 2004]. Furthermore, previous genetic studies on Armenians revealed that during the last several millennia the population was genetically isolated from external influences [Hellenthal G, et al., 2014; Haber M, et al., 2015], which justifies considering this indigenous ethnic group as an appropriate representative of the region for population genetic studies.

Aim and objectives. The main aim of the work is to reveal the role of the Armenian Highland in the human Neolithic migrations from the Near East towards Europe and the Caucasus and reconstruct possible migration routes via the region based on genetic data.

The specific objectives of the work are:

(1) To assess patrilineal genetic structure of populations from the region covering the possible northward and westward directions of Neolithic migrations from the Near East.

(2) To identify the Y-chromosomal haplogroups closely associated with the genetic differentiation of the populations studied.

(3) To determine the distribution pattern of the Y-chromosomal haplogroups, specific markers of the Neolithic migration, in Armenian and comparative populations from the Near East, North Caucasus, Anatolia and Europe.

(4) To reveal the genetic makeup of Armenians and other populations of the region using genome-wide autosomal SNP panel data acquired from modern and ancient DNA samples.

(5) To reconstruct the possible waves and directions of Neolithic migrations via the Armenian Highland based on combined profiles of Y-chromosomal and genome-wide SNP markers.

Scientific and practical significance of the results. The role of the Armenian Highland, a region at the crossroads of Europe and the Middle East, in the spread of Neolithic farmers from the Fertile Crescent was examined using the Y-chromosomal and genome-wide autosomal markers. Based on the results of analysis of both uniparentally and biparentally transmitted genetic markers the possible routes and waves of Neolithic migrations taken place via the Armenian Highland towards Europe and the North Caucasus were identified.

A new insight into the different itineraries and waves of Neolithic migration of the first agriculturalists through the Armenian Highland was suggested. The patterns of the Y-chromosomal structure in spatially different Armenian populations allowed stating that the Armenian plateau served as a transitional corridor for at least two distinct pathways of migration for the Neolithic farmers from the Levant westward and northward. The westward movement took place mainly through the western area of the Armenian Highland and Anatolia alongside the coastline of the Mediterranean Sea. The northward migration occurred predominantly across the central and eastern regions of the Armenian Highland. A distinct Neolithic wave of bidirectional movement to Europe and the North Caucasus was also identified.

The results of analysis of genome-wide autosomal markers revealed three ancestral components contributed to the genetic composition of all Armenian populations – the Near Eastern, Caucasus, and European. The Armenians are the only ethnic group in the Caucasus that has a substantial portion of the Near Eastern ancestry. Based on autosomal genome-wide markers modern Armenians occupy an intermediate position between the populations from the Levant, from one side, and the ethnic groups from Europe and the Caucasus, from another side. Moreover, this inference is supported by the analysis of biparentally transmiited markers of Bronze Age Armenians.

Approbation. Proceedings of the dissertation have been presented at: XIX International conference of students and young scientists “Lomonosov-2012”, Moscow, Russia; International Young Scientists Conferences "Perspectives for Development of Molecular and Cellular Biology-3 and -4", Yerevan, 2012, 2013; the EMBO International Scientific Conference “Human evolution in the genomic era: Origins, populations and phenotypes”, Leicester, UK, 2014; “9th ISABS conference on Forensic and Anthropologic Genetics”, Bol, Croatia, 2015; the meetings of the Academic Council of the Institute of Molecular Biology NAS RA, 2013, 2014, 2015, 2016.

Publications. The main results of the dissertation are reflected in 5 scientific papers and 11 abstracts of presentations at local and international scientific conferences.

Structure. The dissertation comprises 106 pages of computer-formatted English text, including 11 tables and 20 figures, and consists of the following sections: Introduction, Literature Review, Materials and Methods, Results and Discussion, Conclusion, Inferences, and References (including 168 sources).

CHAPTER 1. LITERATURE REVIEW

1.1. Geographic, geological and paleoclimatic description of the Armenian Highland

The Armenian Highland is a landlocked mountainous plateau stretching from the Eastern Caucasus to Anatolia and from the Black Sea to Mesopotamia. It covers approximately 400,000 km² and rises to an average elevation of 1,700 m above sea level. The Armenian Highland is considerably higher than the adjacent Iranian and Anatolian plateaus, with Mount Ararat (5,137 m) being the highest peak. Currently, it encompasses the whole territory of present-day Armenia, southern Georgia, western Azerbaijan, northwestern Iran and eastern Turkey. In ancient times, the Armenian Highland almost entirely corresponded to the lands of historical Armenia, which extended approximately 375,000 km².

The geological structure of the Armenian Highland is complex, with a long and varied geological history spanning the pre-Cambrian to the Quaternary epoch [Lang DM, 1980]. Formerly, more than 25 million years ago, the land of present-day Armenia was entirely covered by oceanic waters, receding in the Upper Tertiary period as a result of uplifts of the earth's crust. These crustal rises have subsequently led to the formation of extensive mountain ranges in the region, with most of the landscape shaped during the Upper Pleistocene [Lang DM, 1980; Gasparyan B, 2010]. Lying on the Anatolian fault, the Armenian plateau is located on a seismic hazard zone, with one active and several dormant volcanic peaks throughout the highland. Volcanism has served as a specific palaeogeographic factor in the formation of the natural environment. Plio-Pleistocene volcanic activity has led to the appearance of volcanic massifs (e.g., Ararat and Aragats), mountainous plateaus (e.g., Gegham, Vardenis, Javakhetian), and lakes, currently preserved as highland plains or intermountain depressions (e.g., the Lori, Ararat and Shirak depressions) [Gasparyan B, 2010]. Volcanic tufa stone, obsidian, basalt, granite, limestone and quartz are the main natural materials of the region, having been exploited for the lithic industry from the Acheulian to the Iron Age [Любин В и Беляева Е, 2006]. There are also

substantial reserves of other minerals in the terrain, including copper, iron, zinc, molybdenum, lead, platinum, silver, and gold. The oceanic origins of the region resulted in the deposition of sedimentary rocks necessary for industry and farming, such as clays, gypsum, lime and dolomite [Lang DM, 1980].

The Armenian Highland has given rise to numerous rivers, namely the Euphrates, Tigris, Yeraskh, Halis, Gaylget, Kur, and Chorokh. Therefore, in the second millennium BC, the territory of the plateau was called Nairi by ancient Assyrians, which meant "country, or land, of rivers". Likewise, there are a number of notable lakes in the highland, the largest of which are Van, Urmia and Sevan.

Paleoclimatic reconstructions indicate that from the remote past, substantial climate oscillations took place in the territory of the Armenian Highland [Ollivier V, et al., 2010; Gasparyan B, 2010]. During the Late Sarmat (11,2-9,3 million years ago) the region was marked by subtropical landscapes and a typical Mediterranean climate, with hot, arid summers and cool, humid winters. The Pliocene and Early Pleistocene epochs had frequent cyclical changes in landscape and climatic conditions. In arid periods, steppe-like vegetation appeared in the region, while in periods of humidification, forest-steppe and forest zones were prevalent. On the whole, the recorded Pleistocene climate of the Armenian Highland consisted of repeated significant cooling and warming, coinciding with global glacial and interglacial cycles [Саядян Ю, 2009]. Currently, the region has enormous climatic variety and diverse ecological niches, largely dependent upon elevation. In most parts of the region, climate is markedly continental, with sufficiently hot summers and cold winters. In spite of severe winter temperatures, the plateau's fertile volcanic soil, rich in useful minerals, is suitable for farming. Thereby, the Armenian Highland has been prominent for the cultivation of fruit, vegetables and cereals. Furthermore, the climatic variance has resulted in copious diversity of floral and faunal species. The richest palaeofloristic and palaeofaunistic remains were found in the basin of the Vorotan River, south-east Armenia [Ollivier V, et al., 2010]. Besides, horses, swine,

sheep, goats, cattle, donkeys, and different types of birds have been bred in the region from the remote past [Lang DM, 1980].

1.2. Paleoanthropological evidence of ancient dwellings on the Armenian Highland

The peculiar geographic location of the Armenian Highland has generated a great amount of scientific interest in studying the plateau for its links to the dispersal of ancient Middle Eastern populations, the Indo-European languages, and the first farmers from the Fertile Crescent [Weale ME, et al., 2001; Гамкрелидзе Т и Иванов В, 1984; Renfrew C, 1990].

Archaeological researches, through the stratigraphic analysis of ancient dwellings, indicate an extensive human presence in the territory of the Armenian Highland during the first half of the Quaternary [Паничкина М, 1950]. The earliest evidence of hominid radiation out of Africa was discovered in the Georgian site of Dmanisi, located in the northern part of the Armenian Highland [Gabunia L, et al., 2001; Presnyakov SL, et al., 2012]. The stone tools excavated from the site were attributed to the Oldowan industry, dated to 1,8 million years ago [de Lumley H, et al., 2002]. Since the Pliocene, savannas with plenty of water reservoirs prevailed upon the territory of the Armenian Highland, which could have first attracted the early human migrants [Любин В, 1989]. It has been suggested that the advancement of early humans from the Near East occurred via the Armenian Highland at the time of the appearance of the Dmanisi site, when the climate and abundant natural resources in the region were favorable for Paleolithic human settlement [Любин В и Беляева Е, 2010]. Paleolithic artifacts in the territory of present-day Armenia were first discovered by the French explorer Jacques de Morgan in 1909. He uncovered stone tools on the western slope of Mount Aragats, dated to the Late Paleolithic period [de Morgan JJM, 1918]. A lot of archaeological sites were found in the Ararat valley (Satani-Dar, Arzni, Erevan-I, etc.), in Karabakh (Azokh, Tagavard, etc.), and in the Lori plateau (Metsavan-3, Privolnoye-1) [Саядян Ю, 2009]. Without taking

into consideration the habitable gaps between glaciations, it was claimed that the Armenian and Iranian plateaus were unoccupied by human beings and covered with thick sheets of ice until the Mesolithic or Early Neolithic periods [de Morgan JJM, 1918]. However, changing climatic conditions of inter-glacial periods were most likely conducive to human settlement [Ghukasyan R, et al., 2010]. Subsequently, a number of Acheulean, Mousterian and Post-Palaeolithic deposits were found across the Armenian Highland [Паничкина М, 1950; Сардарян С, 1954; Любин В, 1989; Ерицян Б, 1975; King T, et al., 2003; Pinhasi R, et al., 2008]. During the excavations in Azokh Cave, located in the Karabakh region, hominid remains identified as *H. heidelbergensis*, stone tools and deposits rich in fauna were also uncovered [Fernández-Jalvo Y, et al., 2010].

Until recently, it was thought that the Early Paleolithic was solely represented in the territory of present-day Armenia by surface finds of Late Acheulean obsidian tools [Беляева Е и Любин В, 2012]. However, from 2003, the Armenian-Russian archaeological expedition collected numerous Early Acheulean artifacts in the Lori district of northern Armenia [Асланян С, и др., 2007; Любин В и Беляева Е, 2010]. These discoveries indicate the significance of the Armenian Highland in relation to early human dispersals from Africa towards the Eurasian landmass [Dolukhanov P, et al., 2004; King T, et al., 2003; Gasparian B, et al., 2010]. Nevertheless, few paleoanthropological fossils have been excavated from the plateau [King T, et al., 2003].

Despite the presence of archaeological records of human activity in the territory of the Armenian Highland dating to the Paleolithic, the severe climatic fluctuations during the Last Glacial Maxium (LGM) probably resulted in the lack of continuous habitation [Dolukhanov P, et al., 2004]. This glaciation began at least 26 KYA [Akçar N, et al., 2007] and has led to significant changes in species' diversity, not only in the Highland but throughout the globe. Due to the subsequent glaciations, the understanding of past climatic variations is especially complicated in mountainous regions [Ehlers J, et al., 2011]. However, the recent research in the Lesser Caucasus

has shown that the glaciers, in general, were restricted to the highest massifs, especially in the eastern ranges (2,500–2,600 m) [Gobejishvili R, et al., 2011]. Archaeological, phylogeographic and paleoecological studies indicate that there were some climatically favorable geographic regions, described as refugia, where species have contracted and persisted during glaciations [Stewart JR, et al., 2010]. It has been revealed that these Quaternary refuge areas played significant role in driving the evolution of refugial human populations [Stewart JR and Stringer CB, 2012].

As ice receded and the climate became steadily milder, people expanded into un-inhabited terrain, resulting in the prerequisites for human population growth [Hewitt GM, 1996]. The current distribution of humans in the northern regions is believed to reflect the post-glacial recolonization from refugial zones at the end of the LGM, dating to approximately 11,5 KYA [Hewitt GM, 2001; Pala M, et al., 2012]. The favorable climatic conditions for the growth of numerous species of wild plants, the variety of fauna, the diversity of landscapes, the richness of soil fertility and abundant water supply indicate that the Armenian Highland was one of the most suitable zones for permanent human dwellings during the postglacial period [Tarasov PE, et al., 2000; Dolukhanov P, et al., 2004].

1.3. Archaeological evidence on human Neolithic migrations from the Near East

A large-scale transition from hunter-gathering to farming, known as the Neolithic Revolution, is broadly recognized as one of the crucial demographic events in human prehistory [Ammerman AJ and Cavalli-Sforza LL, 1971; Diamond J and Bellwood P, 2003; Bellwood P and Oxenham M, 2008]. It is considered that the advent of the *Neolithic lifestyle*, which is characterized by the dominance of settlement sedentism and the domestication of wild animals and plants, led to obvious advantages of farmers over hunter-gatherers and, in particular, drove dramatic human population growth and dispersal [Renfrew C, 1996; Diamond J and Bellwood P, 2003; Bellwood P, 2013]. Furthermore, the adoption of farming during Holocene is widely assumed to have resulted to series of social, economic, and political changes,

including the reorganization of the processes of human social interactions, permanent aggregation of people into large villages, spread of language families, new material cultures and food-producing economies [Renfrew C, 1990; Bellwood P, 2013]. In general, the agricultural innovation eventually led to three major advantages of farming over hunter-gathering [Diamond J and Bellwood P, 2003]. First, in comparison with hunter-gathering, farming let people have a far higher food production output and a steady food supply all year long, thereby supporting denser population. Second, larger sedentary lifestyle and, in turn, potential ability to store food surpluses served as preconditions for technological progress, stratified society, development of centralized states and professional armies [Diamond J and Bellwood P, 2003, Brown TA, et al., 2009]. A third superiority of agricultural society was acquisition of some immunity to epidemic infection diseases, a problem closely related to increased population density, unsanitary practices and shifting of domestic animal diseases to man [Armelagos GR, et al., 1996; Diamond J and Bellwood P, 2003]. In contrast, unexposed hunter-gatherers did not have any resistance to these diseases that promoted their replacement by early farmers during agricultural expansion.

As such a key episode, the origin of farming and an understanding the mechanisms and conditions by which the changeover from foraging to food production occurred have been a major research focus for archaeologists, anthropologists, linguists and human population geneticists [Ammerman AJ and Cavalli-Sforza LL, 1971; College S, et al., 2004; Bellwood P, 2013]. However, the nature and chronology of this important transition in various parts of the world remain controversial and least understood.

Archaeological researches have uncovered the independent emergence of agricultural homelands in many parts of the world at different subsequent times, initially ranging between approximately 10 and 5 KYA [Diamond J and Bellwood P, 2003; Bellwood P and Oxenham M, 2008]. The regions of agricultural origin include

the Middle East, central China, West Africa, New Guinea highlands, Mesoamerica, central Andes, and the eastern woodlands of the USA (Fig. 1.1).

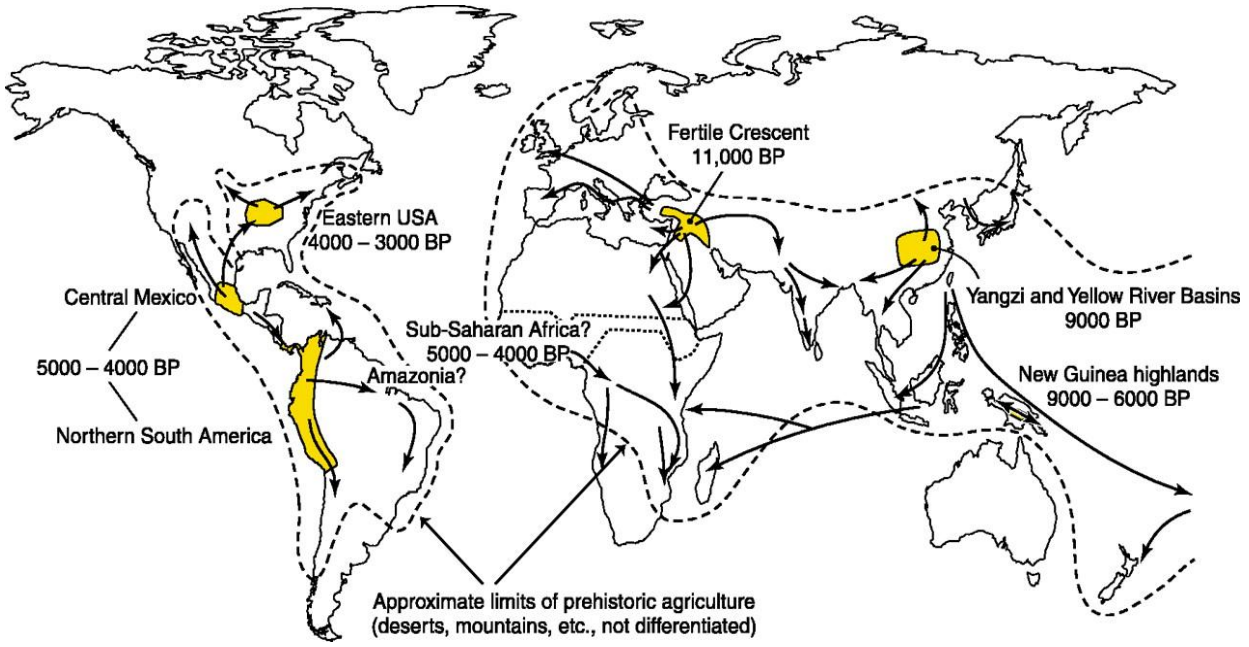


Figure 1.1. Archaelogical map of regions of agricultural origin with approximate radiocarbon dates. (From Bellwood P and Oxenham M, 2008).

In terms of chronology, the Fertile Crescent, the region in the Middle East, spanning the Zagros Mountains of Iran and Southern Mesopotamia northward to Southeast Anatolia, is widely recognized as the earliest farming center where agriculture is known to have originated, dating to around 10 KYA [Ammerman AJ and Cavalli-Sforza LL, 1984; Riehl S, et al., 2013]. The Mediterranean climate, coupled with fertile soils and the diversity of geographic altitudes, may have been preferable for the development of agriculture among hunter-gatherers in the region. It has been assumed that the Fertile Crescent is the domestication centre for three cereals, two pulses, flax, bitter vetch, chickpea, sheep, goats, pigs and taurine cattle [Brown TA, et al., 2009; Bellwood P, 2013]. Finds of large quantities of bioarchaeological remains as phenotypically wild progenitors of domesticated species have pointed to the multiregional nature of the emergence of agriculture in the wide area of the Fertile Crescent [Tanno KI and Willcox G, 2012]. Recent identification of pre-domestication cultivation and morphologically domestic cereals

demonstrated the important contribution of the eastern region of the Fertile Crescent to the management of wild plants [Riehl S, et al., 2013]. From this region, human populations, with their cultural resources and languages, migrated towards various destinations, including Europe, currently the most thoroughly investigated region by archaeologists and geneticists [Bellwood P, 2013; Barbujani G and Chikhi L, 2006].

According to archaeological sources, Europe was first occupied by anatomically modern humans (AMH) in Paleolithic approximately 40 KYA [Otte M, 2000] and underwent a re-expansion out of southern refugia after the end of the Last Glacial Maximum (LGM) (starting from approximately 16 KYA). The number, location, and distribution of these refugial regions remain still uncertain; however, some authors claim that Palaeolithic hunter-gatherers have survived in the Iberian, Italian and Balkan peninsulas, in the territory of present-day Ukraine, and in regions directly towards the East [Oppenheimer S, 2006; Stewart JR and Stringer CB, 2012]. Moreover, recent study considered the parts of the Near East, such as the Levant and coastal southern Turkey, as potential refugias for post-LGM expansion that has been underestimated in previous works [Pala M, et al., 2012]. Thus, Mesolithic gene flow was caused by migration from several centers and, as a result, did not lead to the continent-wide genetic clines. On the contrast, Neolithic spread of farming, starting to about 10 KYA, have had profound effects on the genetic diversity of Europeans and, consequently, traditionally considered as a major demographic process in the region [Diamond J and Bellwood P, 2003; Bellwood P and Oxenham M, 2008]. It has been suggested that the initial farmer spread from the Fertile Crescent reached Anatolia during Pre-Pottery Neolithic B (PPNB), about 8 KYA [Bellwood P, 2013]. From Anatolian coast or by sea from the Levant, the agricultural migration extended to Crete and, later, the mainland Greece. The earliest archaeological records for agricultural economy in South-Eastern Europe were the findings of a fully-fledged farming communities at Knossos on the island of Crete (7 KYA) and dated slightly later ones in the Northwestern Peloponnese of mainland Greece (6,5 KYA) [Bintliff J, 2012] . The Balkans and northern Europe were occupied from Anatolia since 6,2

KYA and, eventually, the Fertile Crescent food production complex reached Britain and southern Scandinavia about 4 KYA [Bellwood P, 2013]. Additionally, in the context of crucial demographic events for Europe, the Bronze age, a period that lasted roughly three thousand years (3,000-1,000 BC), was a highly dynamic period involving a large-scale migration of populations from the east [Allentoft ME, et al., 2015]. To summarize, each of these chronological episodes (depicted in Fig. 1.2) in prehistory of Europe left its genetic legacy that remains as a sedimentary layer in the genetic makeup of present-day Europeans [Soares P, et al., 2010].

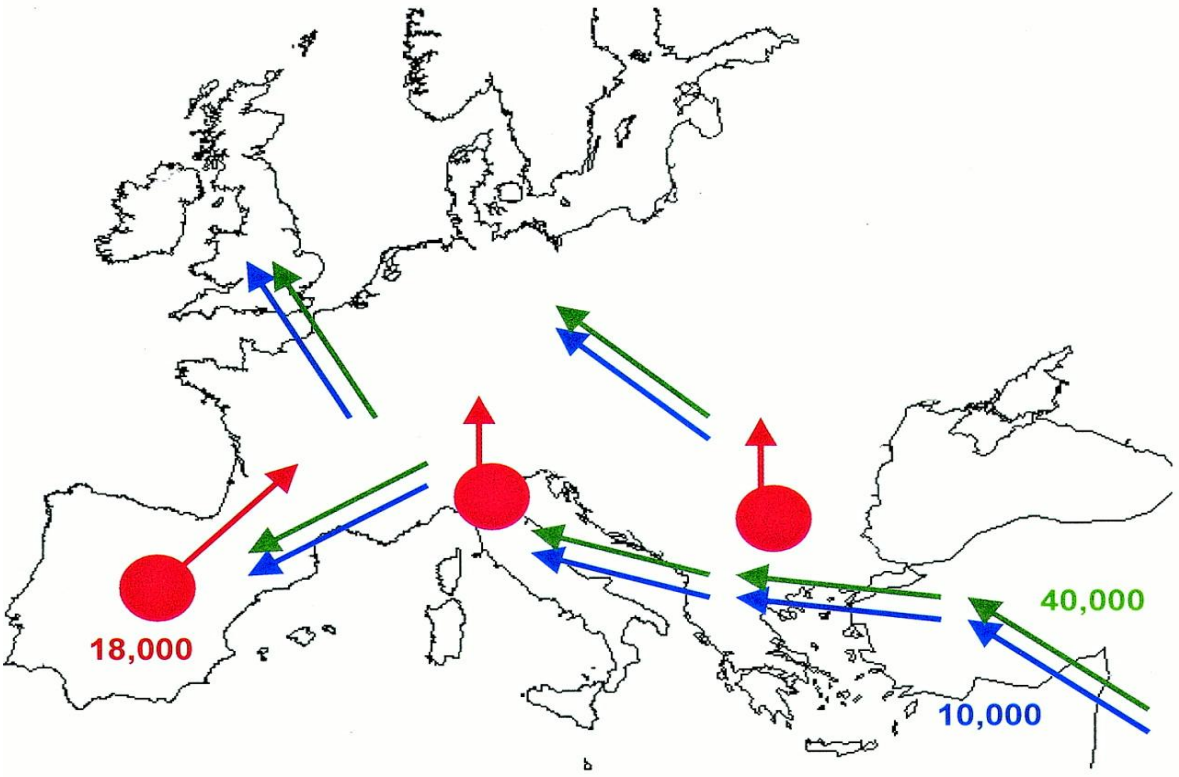


Figure 1.2. Scheme of the main migration events documented in the archaeological record of Europe. (From Barbujani G and Bertorelle G, 2001).
Profound changes in European continent supposed to have occurred during the Paleolithic first colonization (green arrows), Mesolithic reexpansions (the approximate location of glacial refugia are represented by red circles), and during the Neolithic demic diffusion (blue arrows). Numbers are approximate dates, in years before the present.

The Neolithic peopling of Europe and the mechanism of agricultural dispersal have become the subjects of intensive debates. There still are major disagreements among archaeologists and geneticists on the nature of the agricultural transition happened in this region, and on how archaeological and genetic data should be

interpreted [Diamond J and Bellwood P, 2003; Chikhi L et al., 2002; Richards M, et al., 2000]. As a first approximation, two opposing scenarios have been proposed to account for the spread of farming in Europe: the Demic Diffusion model (DDM) and the Cultural Diffusion model (CDM). According to the DDM, spread of agriculture was due to the migration of Near-Eastern farmers determining a substantial gene flow into the indigenous populations [Ammerman AJ and Cavalli-Sforza LL, 1984, Menozzi P, et al., 1978]. The contrasting, CDM suggests that the farming technologies and ideas were transmitted to existing hunter-gatherers, without significant population movement and major genetic changes [Zvelebil M and Zvelebil KV, 1988; Haak W, et al., 2005]. In this context, the 'wave of advance' model was proposed by Ammerman and Cavalli-Sforza in 1971 to emphasize the role of demic diffusion in the spread of agriculture in Europe. According to this model, local population growth and migratory activity have led to the relatively gradual process of demographic expansion of farmers outwards in all directions [Ammerman AJ and Cavalli-Sforza LL, 1971]. In order to assess the average rate at which farming spread, the authors selected the radiocarbon dates from 53 available Neolithic sites, representatives of the early farming in different regions of Europe, and used regression techniques. Four archaeological sites located in the Near East were taken as possible centers of agriculture and fifth center was considered as the center of gravity of their four sites. Further, the great-circle distance from each of European sites to the Near Eastern centers was calculated. The reported average diffusion rate for all probable centers of agriculture was about one kilometer per year, which sounds fairly plausible in the context of the agricultural development considering that the correlation coefficient between geographic distances and radiocarbon ages was high ($R>0.8$). However, authors stressed that, in principle, the observed rate could represents the consequence of cultural or demic diffusion - or even some combination of the two processes. Moreover, average speed of agricultural diffusion estimated by Pinhasi et al is in the range of 0,6-1,3 km/y ($R>0.8$) [Pinhasi R, et al., 2005], which is in agreement with previously obtained value by Ammerman and Cavalli-Sforza

(1971). Further, population dynamic models of the process of Neolithisation in Europe aimed to reveal the role of major waterways and coastlines in the advance of farming from the Near East [Davison K, et al., 2006]. It was shown that sea vessels of that time were well-developed for travel at least 10 km off the coast. Recent study based on the computations of speed map using dates of 918 Early Neolithic archaeological sites in Europe and Near East suggested that the Neolithic transition had cultural nature in Northern Europe, the Alpine region and west of the Black Sea, while the demic diffusion took place in other regions, in particular, the Balkans and Central Europe [Fort J, 2015].

Neolithic migrations from the Near East took place to different directions, including that to the Caucasus, a geographic region between the Caspian and Black Seas. The area is characterized with notable cultural and linguistic diversity, and is considered to be crucial for understanding human ancient migration and genetic diversity of Eurasia [Yunusbayev B, et al., 2012]. The Caucasus has always been characterized by a rich diversity of flora and fauna, which could attract ancient humans for constant habitation. From the remote past, the Caucasus Mountains have been a conduit between the Near East and east European plains, in particular, serving as a transition bridge for migration of Near Eastern agricultural farmers towards the north and northeast [Zohary D, et al., 2012]. Archaeological data revealed that initial expansion of AMH into the region was since Early Upper Paleolithic, at least 42,000 years ago [Adler DS, et al., 2008; Adler DS, et al., 2014]. Further glaciations has led to the uninhabitable condition on the most part of the Caucasus region and its re-population after ice sheet melting [Tarasov PE, et al., 2000]. The continuous human occupation of the area is confirmed by numerous archaeological sites since about 10 KYA. Settlement sites are encountered in the North Caucasus and in the West Transcaucasia primarily in caves and rock shelters. At the time of Neolithic, the Caucasus was characterized by a transition from relatively cool climate to a warmer one with increased humidity [Kiguradze T, 2001]. Fewer Neolithic sites are found on the territory of the Northern Caucasus, while the majority of aceramic and ceramic

sites are situated in western parts of Georgia. One of the earliest Neolithic sites outside of the Near East was discovered at Chokh in the eastern part of Daghestan, dating to approximately 8,000 BC. [Гаджиев М, и др., 1996]. It was suggested that in parallel with the development of agriculture and animal husbandry, the rapid growth of population in the Neolithic Caucasus had occurred [Kiguradze T, 2001]. In particular, in western parts of Georgia a number of ancient sites of several hectares are found, which probably indicates to habitation of several kin groups together. In the context of the peopling by early farmers, it is logical to suggest that the population of the Caucasus came to this region from the South, via the central and eastern parts of the Armenian Highland, as it is adjacent to the Fertile Crescent where the earliest domesticated plants and animals are found.

1.4. Linguistic data associated with the Near Eastern first migrants

The question of the Neolithic dispersal became intertwined with that of the spread of the Indo-European languages, as it was proposed by Renfrew that the Proto-Indo-European language advanced through the continent with early farmers who entered southeastern Europe from Asia Minor about 7,000 BC [Renfrew C, 1990]. In agreement with Anatolian homeland, the reconstruction of Proto-Indo-European vocabulary reveals that it was typical vocabulary for food-producing society [Гамкрелидзе Т и Иванов В, 1984]. Moreover, both language-tree divergence times and inferred root location support the Anatolian theory of Indo-European origin indicating that the initial Indo-European divergence had taken place of between 8,000 and 9,800 years ago [Bouckaert R, et al., 2012]. Specifically, the establishment of Indo-European languages in the European continent was thought to be associated with the migrations of first farmers from the neighboring regions of the Armenian Highland [Гамкрелидзе Т и Иванов В, 1984; Renfrew C, 1990]. This hypothesis was based on the fact that Armenian, an old independent branch within the Indo-European language group, has similarities to the ancestral Proto-Indo-European languages. Furthermore, the Indo-European speaking populations of

Europe have close linguistic relationships with Armenian ethnic groups. Alternatively, it was suggested that Armenian language originated during the Bronze Age, when Indo-European speaking tribes from the Balkans and Greece migrated towards Anatolia and Transcaucasia and spread their culture and language [Greppin JA, et al., 1991]. However, genetic surveys of Armenian paternal gene pool pointed to the scarce presence of Y chromosomal lineages frequently encountered in the Balkan region [Herrera KJ, et al., 2012; Hovhannisyan A, et al., 2014]. Futher genetic studies with the application of maternal and genome-wide autosomal markers could disentangle this long-standing question.

An alternative theory on the ethnogenesis of Indo-European languages claims that they were spread by Kurgan horsemen from the Pontic steppe region north of the Caspian Sea about 6 KYA [Gimbutas M, 1970]. It was shown that words for supposed Kurgan technological innovations are relatively consistent across widely divergent Indo-European sub-families. Moreover, Kurgan hypothesis finds a strong support with the recent genetic survey suggested that the spread of the Indo-European languages was accompanied with the Bronze Age migration of Yamnaya people from the Eurasian Steppe [Haak W, et al., 2015].

In general, the two main competing theories of Indo-European homeland face severe difficulties. Linguistic argument for the Kurgan hypothesis is based on only limited evidence for a few Proto-Indo-European word forms [Renfrew C, 1990]. Furthermore, it is hard to understand how nomadic pastoralists could have imposed their language on so much of Europe west of the steppe region [Diamond J and Bellwood P, 2003]. The ‘Anatolian farming’ hypotheses has become less plausible in the light of the recent genetic evidence and archaeological criticism [Zvelebil M and Zvelebil KV, 1988; Haak W, et al., 2015]. Thus, over the past centuries, the question of how Indo-European languages had spread is very controversial to be exactly resolved and there is no general consensus on this matter till now.

1.5. Archaeological and archaeobotanical evidence on Neolithic farming spread through the Armenian Highland

Located at the crossroads of Europe and the Middle East, the Armenian Highland was a conduit for major waves of prehistoric and historic migrations [Dolukhanov P, et al., 2004], as well as a cradle for various ancient civilizations [Lang DM, 1980]. Over the centuries the plateau linked major trade ways from the Ancient Near Eastern, Iranian, Greek and Eurasian worlds, thus having the great strategic significance [Redgate AE and Redgate AE, 1998; Lang DM, 1980]. The unique geographic location of the plateau has garnered a great deal of scientific interest as a potential link between eastern and western Eurasian populations. Moreover, the variable climatic diversity and proximity to the Fertile Crescent likely contributed to the post-Last Glacial Maximum (LGM) Neolithic resettlements of the Armenian plateau, particularly by the first farmers from the Near East [Dolukhanov P, et al., 2004; Вавилов Н, 1960; Herrera KJ, et al., 2012].

As early as in 1926, A. Potapov excavated Neolithic tools near Vagharshapat [Пиотровский Б, 1949]. Later, dozens of artifacts related to agriculture and animal husbandry were discovered from the territory of the Armenian Highland, being consistent with the previous suggestion of post LGM resettlement within the region [Kushnareva KK, 1997]. According to the current data, agricultural techniques on the plateau have been well-developed and practiced from prehistoric times [Gandilyan PA, 1998; Овсепян Р и Овсепян Г, 2009]. Stone, wooden digging tools and hoes were in use for tilling the soil, while for the cultivation of harvest early farmers in Armenia applied saddle querns and pestles [Lang DM, 1980]. Furthermore, there is a written evidence of the flourishing agricultural industry in the Armenian Highland [Gandilyan PA, 1998]. In particular, based on what he had seen in the Kingdom of Van, ancient manuscripts attest the favorable impression made on the Assyrian king, Sargon II, regarding the agricultural industry and specifically the abundance of wheat and barley in the plateau [Wright EM, 1943].

Archaeobotanic data provide an opportunity to directly reveal the origin and development of farming in the region. The first investigation of plant remains in the territory of the Armenian Highland was carried out by Tamamshyan in 1935. The presence of emmer wheat remnants was found during exploration in the archaeological site of ancient Vagharshapat [Тамамшян М, 1935]. Since then, many scientists have investigated archaeobotanical materials from the Armenian Highland [Гулканян В, 1966; Лисицына Г и Прищепенко Л, 1977; Gandilyan PA, 1998; Овсепян Р и Овсепян Г, 2009]. Owing to the wide presence of wild and cultivated grains, as well as the carbonized remains of cereals in many Neolithic sites, Armenia and Georgia are suggested as being among the original homelands of a number of cereal plants [Cole SM, 1970]. Archaeological finds from the sites of Artashen and Aknashen, located in the Ararat valley, represent the earliest known evidence for Late Neolithic farmers' economy in present-day Armenia [Hovsepyan R and Willcox G, 2008; Овсепян Р, 2011]. The study of charred remains and chaff impressions in mudbrick from the two sites identified naked and hulled barley, possibly naked wheat, emmer, small-seeded lentils, pulses, bitter vetch and pips of wild vine [Hovsepyan R and Willcox G, 2008]. Almost the same plant samples were excavated from many Neolithic sites of the Caucasus and the Near East, the only difference being that both naked barley and wheat are rare finds in the Near East [Лисицына Г и Прищепенко Л, 1977; Brown TA, et al., 2009; Zohary D, et al., 2012]. The discovery of high quantities of *Alyssum* and *Camelina*, two uncommon cruciferous species for this agricultural period, has led to the assumption that these plants may have served as oil sources [Hovsepyan R and Willcox G, 2008]. In total, about 25 cultivated cereals, pulses and oil-producing plants were found during excavations of prehistoric sites in the region [Овсепян Р и Овсепян Г, 2009]. In addition to the plant remains, many crop processing artifacts were found at the sites of Artashen and Aknashen. The world's earliest known ancient winery was found in Armenia, in the "Areni-1" cave in Vayots Dzor, dated to about 6,000 years ago [Barnard H, et al., 2011]. Archaeologists revealed a wine press, fermentation jars and wine-soaked pottery

shards at this site. This discovery, combined with finds of carbonized or petrified grape-pips, led to the proposition that the region was among the earliest centers for viticulture and winemaking [Lang DM, 1980]. On the basis of archaeological surveys of ancient settlements in Armenia, the following dating schemes have been suggested to distinguish the major periods in the development of agriculture in the region: Neolithic (sites of Artashen, Aknashen, 6-5 KYA); Chalcolithic (sites of Teghut, Areni-I, Godedzor, 5-4 KYA); Bronze age and the beginning of the Iron Age (sites of Gekharot, Aparan-III, Tsaghkasar, Voskevaz, Mokhlablur, Lorut, Aygevan, Elar, Shengavit, NerkinNaver, Agavnatun, Shagat-I, Aragaciberd, Tsaghkahovit, Orom, Hnaberd, Uyc, Teishebaini, Jujevan, Metsamor, Yenokavan-II, Dvin, 4-1 KYA); and the development of agriculture associated with the Kingdom of Van (sites of Aramus, Argishtikhinili, Orom, Karmir Blur, Shagat-III, Tagavoranist, Sev-sevqareri blur, Tsaghkahovit, Shagat-I, 9-6 centuries BC, 9-6 centuries BC) [Овсепян Р, 2011].

Thus, archaeological and archaeobotanical artifacts have indicated that agriculture was adopted in Armenia around 6,5 KYA. Although the area within the Armenian Highland is currently being studied by archaeologists, as yet there is a lack of convincing data to enable to describe the generalized pattern of Neolithic migrations through this region. However, it is possible to bridge this gap by applying the genetic study of populations indigenous to this geographic area.

1.6. Genetic evidence on the Neolithic migration from the Near East

Since the advance of highly sensitive genetic markers and the development of molecular techniques, genetics studies have been extensively applied to disentangle long-standing questions in population genetics, archeology and anthropology [Menozzi P, et al., 1978; Cavalli-Sforza LL, et al., 1994; Sjödin P and François O, 2011]. Furthermore, in the last two decades, the most powerful tools to reconstruct past demographic events have become the methods of molecular population genetics based both on modern [Rootsi S, et al., 2012; Karmin M, et al., 2015] and ancient

(aDNA) data [Brandt G, et al., 2013; Lazaridis I, et al., 2014; Mathieson I, et al., 2015; Allentoft ME, et al., 2015].

Due to the patchiness and paucity of the archaeological record, the majority of recent studies on the nature of the spread of agriculture from the Fertile Crescent are based on biological (genetic and phenotypic) data analysis [Menozzi P, et al., 1978; Cavalli-Sforza LL, et al., 1994; Dupanloup I, et al., 2004; Sjödin P and François O, 2011]. Under the demic diffusion model [Ammerman AJ and Cavalli-Sforza LL, 1984; Menozzi P, et al., 1978; Chikhi L, et al., 2002], the extant genetic diversity of modern Europeans would have resulted mainly from the genetic pool of the Near Eastern Neolithic farmers, while conversely, the cultural diffusion model asserts that European lineages would have been expected to have descended from indigenous hunter-gatherers [Zvelebil M and Zvelebil K, 1988; Haak W, et al., 2005; Morelli L, et al., 2010]. In general, genetic studies based on different nuclear, mitochondrial, Y-chromosomal markers, and ancient DNA analysis differ considerably in their evaluation of the contribution of Paleolithic hunter-gatherers and Neolithic farmers to the composition of the modern European gene pool [Pinhasi R, et al., 2005; Fernández E, et al, 2014]. Overall, previous studies highlight the entanglement and complexity of such historical events as farming dispersal and, ultimately, the peopling of Europe. The intricacies of these migratory events with varying patterns of cultural and demographic diffusion in different regions require the development of relevant models reflecting the process of Neolithic dispersal throughout Eurasia [Barbujani G and Chikhi L, 2006].

The majority of early genetic studies aimed to reveal the genetic makeup of present-day European and Near Eastern populations based on the analysis of classical genetic markers, namely, blood groups, allozymes and histocompatibility alleles. The first principal components analysis (PCA) based on 39 classical genetic markers revealed a spatial frequency cline from the south-east to the north-west of Europe, which was interpreted as the signature of the demic diffusion model [Menozzi P., et

al., 1978]. The results of this work were further confirmed by Cavalli-Sforza and collaborators using over 130 genetic markers [Cavalli-Sforza LL, et al., 1994].

Although maternally inherited mitochondrial DNA and paternally inherited non-recombining region of the Y chromosome (NRY) capture only a narrow spectrum of human genetic pool, these haploid systems have been widely used over the past decade for the reconstruction of human demographic history since the out-of-Africa migration [Cavalli-Sforza LL and Feldman MW, 2003]. Furthermore, significant disagreements between the mitochondrial DNA (mtDNA) and NRY distribution patterns indicate the possibility that the demographic histories of females and males were different. For instance, patrilocality practices in early farmers lead to major differences in terms of genetic diversity within and between populations when comparing mtDNA and NRY data [Rasteiro R, et al.2012].

Maternal inheritance, high copy number per cell, apparent lack of recombination and high substitution rate are the features of mtDNA which have offered the opportunity to explore female-specific aspects of the demographic history of human populations. The patterns of mitochondrial variation has been used to elucidate the processes of ancient human colonization of the Old World, whereas regional differences in variance allow addressing questions of a more recent time scale [Kivisild T, 2015]. In the context of the agricultural migration, mitochondrial DNA (mtDNA) analyses mainly support the CDM [Richards M, et al., 2000; Richards M, et al., 2002], however some critics suggest that it actually indicates the DDM [Barbujani G and Chikhi L, 2006]. A principal component analysis of mitochondrial hypervariable segment I (HVS-I) typing data revealed insignificant east-west gradient that accounted for 23% of the variation [Cavalli-Sforza LL and Minch E, 1997]. It was proposed that the values of the time to the most recent common ancestor (TMRCA) reflect a Paleolithic origin of European maternal genetic pool, and the Neolithic input of farmers to the European mtDNA lineage is within 20% [Richards M, et al., 2000;Torroni A, et al., 2001]. Moreover, Richards et al suggested that the current mtDNA picture of Europe has resulted from multiple diffusion waves in the

Upper Paleolithic, and the phenomenon, known as the founder effect or bottleneck, has led to reduced preexisting diversity during the Last Glacial Maximum [Richards M, et al., 2000]. The analysis of more than 2,600 sequences of the first hypervariable mitochondrial control region for the samples from Europe, the Near East, and the Caucasus have identified homogeneous pattern of European specimens, with the only clinal differentiation in the region of the Mediterranean Sea [Simoni L, et al., 2000]. Based on the results of these studies, it was further suggested that selection process may had significant influence on the mtDNA to be a suitable demographic marker system, or that female gene flow had been extremely high within Europe [Barbujani G and Chikhi L, 2006]. However, the more detailed analysis with sufficiently large sample size and various subclades of mtDNA indicated significant geographical gradients that are similar to those ones based on Y-chromosomal markers and classical genetic systems [Richards M, et al., 2002]. Recent advent of complete mitochondrial DNA sequence data stepped over to a new stage of human evolutionary studies. Currently, even limited complete mtDNA sequences allow identifying the critical polymorphisms that define subclades within each particular mitochondrial haplogroups. To test whether the presence of signals of population expansion in present-day mitochondrial gene pool is detectable, a large dataset of 1.151 complete mtDNAs from modern Europeans were used in more recent study [Fu Q, et al., 2012]. The analysis of mtDNA haplogroup distribution patterns revealed the association of the haplogroup H with hunter-gatherers that adopted farming practices and admixed with immigrant farming populations. Worth to note previous high-resolution genetic survey where it was revealed that the haplogroup H, which is characterized by an extremely wide geographic distribution in Europe and a very high frequency in most of its range, entered Europe from the Near East around 20,000–25,000 years ago, at the time of the Last Glacial Maximum, and some of its subclades further re-expanded from an Iberian refugia when the glaciers retreated 15,000 years ago and conditions for habitation became more favorable [Pereira L, et al., 2005]. However, on the contrast, another high resolution survey of the haplogroup H with

the inclusion of datasets from the Caucasus, the Near and Middle East – regions where most of European genetic lineages, including mtDNA haplogroup H, have likely originated – indicated the limited post-LGM maternal gene flow between Europe on the one hand, and Near East and South Caucasus, on the another hand [Roostalu U, et al., 2007]. Moreover, it was additionally suggested that mitochondrial gene pool of the populations of the North Caucasus has received an influx of haplogroup H variants from the Ponto-Caspian/East European region. It is obvious that the question on the emergence of the haplogroup H in Europe still is controversial. Meanwhile, the haplogroup U was found to be typical for early farmers and represents the evidence of a population expansion related to the spread of animal husbandry and agricultural practices [Fu Q, et al., 2012]. On the contrast to the majority of previous mtDNA studies, the results of the analysis conducted by Fu et al. suggested the Neolithic expansion into Europe and pointed to the spread of agricultural techniques to Europe with parallel migration of early farmers followed by the eventual assimilation of resident hunter-gatherers. Overall, it was shown that in the case of only limited ancient genetic data, studying mitochondrial genome variation in modern Europeans at the complete sequence resolution could be applied in order to resolve the questions concerning ancient population histories. Interestingly, the results of the examination whether major human expansions started after the advent of agriculture, i.e. in the Neolithic time, revealed that the major population expansions occurred after the Last Glacial Maximum, with the appearance of amiable environments, but before the Neolithic [Zheng HX, et al., 2012]. For the analysis the authors used a comprehensive dataset of whole sequences of human mtDNA to identify the expansion of the lineages and to reconstruct various historical demographic events. Moreover, it was hypothesized that the continuous growth of populations subsequently generated a need for the introduction of agriculture, being a driving force for its further development process. Thus, although the phylogenetic resolution of mitochondrial diversity has greatly improved as a result of increased sequencing of mtDNA genomes, it has become evident that such a complex process

in human prehistory as the Neolithic Revolution is needed to be considered not only from the maternal perspective, but taking into account the results based on both paternal and biparental genetic systems.

Due to the presence of convenient polymorphic markers, the Y-chromosomal non-recombining region still remains an indispensable and relatively simple tool for the patrilineal study of complex historic migration events that influenced modern-day Europeans` genetic diversity [Jobling MA and Tyler-Smith C, 1995; Jobling MA and Tyler-Smith C, 2003; Underhill PA and Kivisild T, 2007]. In particular, relatively stable (in evolutionary terms) biallelic unique event polymorphisms (UEP) that include single nucleotide polymorphisms (SNPs) with Y-chromosomal haplogroup defining characteristics and certain insertion/deletion polymorphisms (indels), provide an ability to identify descendants of deep common ancestor [Weale ME, et al., 2001]. More rapidly mutating short tandem repeats (STRs) on the NRY locus, with increased or decreased number of repeats over a generation, are used in population genetic surveys for the detection of relatively recent demographic events [Morelli L, et al., 2010]. Furthermore, among the useful features of the Y chromosome for human population research is its high level of geographic stratification and diversification, providing more specific inferences concerning population movement in comparison with other genetic systems [Rosser ZH, et al., 2000; Lippold S, et al., 2014]. This phenomenon can be explained by genetic drift and predominance of the cultural practice of patrilocality, which means that women rather than men are more likely to move after marriage and, as a consequence, most men live close to their birthplaces [Seielstad MT, et al., 1998]. In addition to the frequency of classical genetic markers, the distribution of Y-chromosomal haplogroups shows broad clines across Europe, which was characterized as one of the main features of the European genetic landscape and regarded as evidence for the DDM [Ammerman AJ and Cavalli-Sforza LL, 1984]. Moreover, previous studies of Y-chromosomal haplogroup distribution reveal that the majority of contemporary European lineages fall into the haplogroups E, G, I, N, and R [Jobling MA and Tyler-

Smith C, 2003; Myres NM, et al., 2011; Semino O, et al., 2004]. Further, it has been suggested that some Y-chromosomal haplogroups serve as specific markers of the Neolithic migration involving the first farmers from the Fertile Crescent, namely, E1b1b1-M35, J2-M172, G-M201, and R1b1a2-M269 lineages [Rosser ZH, et al., 2000; Myres NM, et al., 2011; Semino O, et al., 2004]. In particular, haplogroup R1b1a2-M269 is the most common Y-chromosomal lineage in Europe, encountered in 110 million European men, and increases in frequency westward [Chiaroni J, et al., 2009; Balaresque P, et al., 2010]. Lately, the question of whether its origins were in the Paleolithic or Neolithic periods has become the subject of intense debate. In this context, Busby et al. claim that the existing data and methods are not capable of unambiguously estimating the age of its origin and the directions of its migration [Busby GB, et al., 2012]. However, in some recent works, the observed explicit frequency cline of the haplogroup R1b1a2-M269 from Anatolia to Western Europe and its associated haplotype diversity cline in the opposite direction suggest that the lineage may have spread towards Europe with the migration of Neolithic farmers from the Near East [Myres NM, et al., 2011; Balaresque P, et al., 2010]. Conversely, the Y-chromosomal haplogroups G-M201 and J2-M172 are widely distributed in populations of the Caucasus, Near/Middle East, and Southern Europe, with the highest frequency in the North Caucasus [Balanovsky O, et al., 2011; Yunusbayev B, et al., 2012]. Bearing this in mind, Morelli et al. examined the distribution of the putative Y-chromosomal tracers of agriculture in Anatolia and Sardinia, the regions where there is significant archeological, genetic and paleontological evidence of pre-Neolithic human colonization and quite high frequency rates of these haplogroups [Morelli L, et al., 2010]. The results have shown that the distribution of the considered Y-chromosomal haplogroups in population of Sardinia markedly differs from that of Anatolia. In particular, the estimated variability of the TMRCA values and STR content of R-M269 haplogroup indicated that Sardinians and other Western European populations underwent more ancient Upper Paleolithic expansion in comparison with Anatolians and other Eastern Europeans populations. Hence, the

results of this work support a pre-Neolithic spread of Y-chromosomal clades from a common ancestral gene pool and argue against demic diffusion model. Moreover, based on Y-chromosomal data Semino et al suggested that contribution of male immigrant farmers to European genetic pool is only about 20% and the distribution of haplotypes in Europe trace the migration of Paleolithic hunter-gatherers from different refugees during the end of the Last Glacial Maximum [Semino O, et al., 2004]. In contrast, observed clinal patterns of the NRY diversity within Europe have been interpreted in favor of demic diffusion model for the spread of farming and some genetic studies revealed significantly high rates of Neolithic contribution, at least 50% [Chikli et al., 2002; Dupanloup I, et al., 2004]. As for mtDNA, studies of Y-chromosomal DNA have shifted to the large-scale complete sequencing of chosen lineages or populations. However, in comparison with the mitochondrial DNA, complete sequencing of Y-chromosomal DNA has a few limitations. As Y-chromosome has 3,000-fold greater length and richer in repeated sequences than any other chromosome, accurately mapping of repetitive regions becomes particularly problematic for sequencing process [Wei W, et al., 2013]. However, next-generation sequencing (NGS) technologies have now developed further and allow overcoming this matter. The structure of phylogenetic tree resulting from the Y-chromosomal sequences is an informative method to shed light to the process of the haplogroup expansion, whereas branch lengths are in principle proportional to time, avoiding dating problems associated with STRs, namely choosing their appropriate mutation rate and possible long-term mutation saturation [Batini C, et al., 2015]. In particular, analysis of the major Y-chromosomal European lineage, the haplogroup R1b, based on complete sequencing, supported a Neolithic origin for the node, dating it to 4-13 thousand years ago [Wei W, et al., 2013]. A markedly starlike structure of the R1b phylogeny indicates that the lineage experienced an extremely rapid and extensive increase as soon as it entered the European continent, therefore only few mutations occurred during the expansion. These results were further confirmed by the study of Sikora et al [Sikora MJ, et al, 2013]. Another population-based NGS studies across

Europe have pointed to more recent coalescent times (3,5-7,3 KYA) for three major European Y-chromosomal lineages (I1,R1a and R1b) and to a recent widespread male-specific phenomenon that may be an evidence of social selection during Bronze Age [Lippold S, et al., 2014; Batini C, et al., 2015]. Moreover, based on comparison of NRY and mtDNA variation it was proposed that the effective population size of females has been larger than that of males throughout human history [Lippold S, et al., 2014]. Currently, the number of studies based on complete sequencing of Y-chromosomal and mitochondrial DNA has been growing over the time providing new insights into the paternal and maternal histories of human populations.

Although single-locus studies continue to provide valuable insights into human origins and migration events, it is clear that these systems represent only a small portion of human genome. With rapid development of new sequencing technologies in recent years, genome-wide data and whole sequences of genomes have enabled a more detailed insight into a genetic diversity of human populations, being particularly useful for inferring population divergence and admixture times, directions of migrations, and other aspects of human demographic history.

Genome-wide data generally provide more reliable insight into the population history because they are based on analysis of many independent loci. Recently, a range of genome-wide studies of present-day populations have been conducted. In one of the published papers on this topic [Yunusbayev B, et al., 2012] it was shown that irrespective to remarkable linguistic diversity, populations from the Caucasus are characterized by relatively high level of autosomal uniformity. In another survey, genome-wide diversity of populations from the Levant pointed to the recent genetic stratification of the region, which had been driven by religious affiliations of the populations [Haber M, et al., 2013].

Verily new era in population genetic studies inaugurated only a few years ago, with the availability of ancient DNA samples for genetic analysis. In comparison to the studies of genetic makeup of present-day populations, aDNA analysis offers an opportunity to directly investigate specimens from specific temporal layers of

interest. In particular, aDNA of the earliest farmers has been used for understanding whether or not the diffusion of farming in Europe was accompanied by substantial influx of people [Sampietro ML, et al., 2007; Olalde I, et al., 2014]. However, as the DNA from ancient samples are rare well-preserved, only few genetic analyses could be performed on excavated Neolithic remains, and they are often limited to the study of genetic variation of their maternal gene pool [Haak W, et al., 2005; Sampietro ML, et al., 2007; Bramanti B, et al., 2009]. In relation to the debate surrounding the phenomenon of Neolithisation in Europe, Haak et al analyzed mtDNA of 24 Neolithic skeletal remains in Central Europe, dated to the Linear Pottery Culture (LPK) (5,500–4,900 calibrated B.C.) and found that 25% of the specimen had the N1a haplogroup [Haak W, et al., 2005]. In comparison, modern mitochondrial genetic pool of Europeans consists of only 0.2% in frequency of the lineage. This pattern was interpreted as a signature for a little genetic influence of Neolithic immigrants to the extant European female lineages and, thus, supports the cultural model for agricultural diffusion. Further, the results of Haak et al have been criticized for limited sample size and generalization of the inferences to the whole Europe [Barbujani G and Chikhi L, 2006]. On the other hand, a genetic survey of 11 aDNA samples from a Middle/Late Neolithic site (5,5–5 KYA) on the Iberian Peninsula did not reveal significant differences from present-day populations, indicating that Neolithic transition in Iberia occurred with a quite different scenario [Sampietro ML et al., 2007]. Likewise, the results of patrilineal and matrilineal analysis of 29 ancient samples from a French Mediterranean region (5 KYA) [Lacan M, et al., 2011] are consistent with those published by Sampietro et al. The presence of mtDNA haplogroup, which are typical for modern European populations, and absence of N1a lineage support a demic diffusion model in Mediterranean area of Europe. Moreover, the results of the study based on the investigation of 22 mitochondrial DNA sequences from central and northern European post-LGM hunter-gatherer skeletons allow proposing the absence of genetic continuity between hunter-gatherers and first farmers and between hunter-gatherers and present-day Europeans [Bramanti B, et al.,

2009]. The authors have come to the conclusion that early farmers of central Europe were not the descendants of local hunter-gatherers and, on the contrast, immigrated into the region from the Near East. More recent genetic survey by Haak et al confirms this suggestion, revealing shared genetic affinity of LPK samples with the modern-day populations of the Near East and Anatolia [Haak W, et al., 2010]. Thus, the results indicate a significant genetic input from this area during the agricultural transition in Europe.

It is notable that unique genetic features were found in the LPK populations which were interpreted as a signature for major demographic events taken place in Europe after the early Neolithic [Haak W, et al., 2010]. Furthermore, a model of random rather than clinal dispersal of agricultural migration suggested by Hervella et al., presumed different genetic impact of Neolithic farmers on the various geographic regions and in different periods of time. Thus, more complex process of Neolithisation in Europe was proposed [Hervella M, et al., 2012]. In particular, the mitochondrial DNA analysis of hunter-gatherers and first farmers from Northern Spain suggests different scenario in Central Europe, Mediterranean Europe and the Cantabrian fringe. In regard to Central Europe, the key role in the formation of modern mtDNA diversity of the region was suggested for late Neolithic cultures [Brandt G, et al., 2013]. Additionally, the analysis of 15 mitochondrial DNA data of the original Near-Eastern Neolithic communities (8,700–6,600 BC) revealed genetic proximity between these samples and the present-day populations of Cyprus and Crete, which could be interpreted in favor of the introduction of Neolithic into Europe through the pioneer seafaring colonization [Fernández E, et al., 2014]. Concerning the question of different waves of migration, the study of remains from archaeological sites in Romania, spanning a time-period from the Early Neolithic to the Late Bronze Age, aimed to reveal the genetic structure of the various prehistoric human populations that spread throughout Europe from Anatolia and the steppes north of the Black Sea [Hervella M, et al., 2015]. The findings allow proposing the hypothesis that second wave of Neolithic migration from Anatolia had the highest

impact on the genetic composition of the European populations. Application of NGS technologies to the ancient DNA data have led to a larger capturing of the genetic information contained in ancient genomes. The first complete sequenced ancient DNA was 5,300-year- old Tyrolean Iceman (Late Neolithic or 'Copper Age'), who, surprisingly, had more genetic similarity to modern Sardinians than to the present-day population inhabiting the region where he probably lived [Keller A, et al., 2012]. The results suggested that crucial demographic changes have occurred in Europe after the Neolithic period. Further complexity in the understanding of European genetic composition has come with recent discoveries indicating a third population, the Northern Eurasians that contributed their genetic legacy to modern Europeans [Lazaridis I, et al., 2014]. Authors sequenced the ancient European genomes of eight 8,000-year-old hunter-gatherers and 7,000-year-old farmer, comparing them further with genome-wide data of other ancient and modern samples. Additionally, it was revealed that early European farmers had about 44% of ancestral legacy from a 'basal Eurasian' population, that split before the variegation of other non-African lineages. Demic diffusion model was further refined with the study of 69 European genome-wide ancient DNA data that revealed Yamnaya related ancestry persisted in present-day Europeans, which is, in particular, lower in southern and higher in northern parts of Europe [Haak W, et al., 2015]. It was suggested that the arrival of the Near Eastern early farmers during the Early Neolithic was followed by a massive migration from the Eurasian Steppe (approximately 4,5 KYA) involving Yamnaya population. This migration afterwards has led to the profound changes in European Y-chromosomal haplogroup composition, giving the predominance to the haplogroups R1a and R1b instead of the G2a lineage. Thus, ancient genomes from Eurasia indicate tree ancestral populations with various impacts to the modern European gene pool: Western hunter-gatherers who had inhabited Europe since Paleolithic times; Early farmers from the Levant, that shaped the European genetic makeup during Neolithic migration; Pontic steppe herders, who arrived during the Bronze Age [Lazaridis I, et al., 2014; Haak W, et al., 2015] (Fig. 1.3). In its turn, genetic source for Steppe

herders and, in particular, for Yamnaya population was attributed to Near Easterners and eastern Eurasian hunter-gatherers. Recent genome-wide study of 230 ancient DNAs, spanning time period from 6,500 and 300 BC, revealed that Anatolian Neolithic farmers were members of the population that was the source of European first farmers [Mathieson I, et al., 2015]. Moreover, ancient whole genome from Ireland possessed predominantly Near Eastern origin, which suggestes a substantial influx of early farmers to the island [Cassidy LM, et al., 2016].

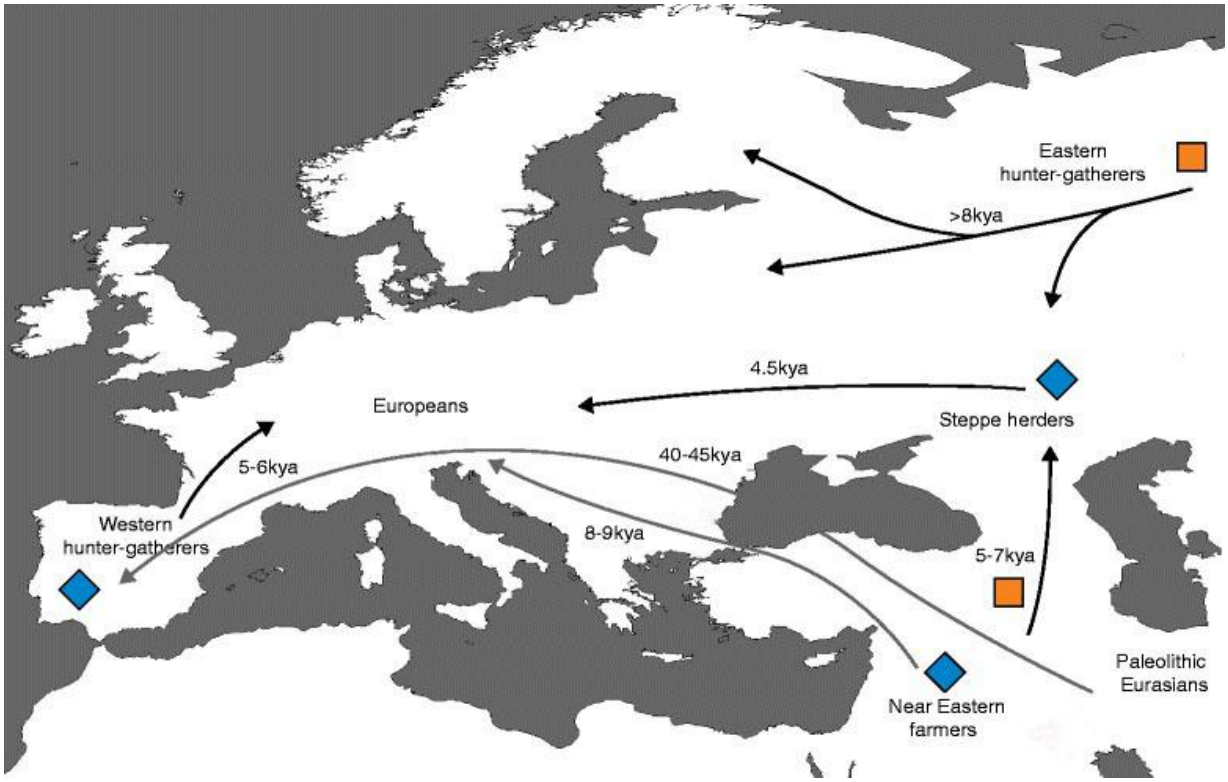


Figure 1.3. Peopling of Europe based on the results of ancient and modern DNA analysis. (From Haber et al., 2016; modified).
Three ancient populations had an impact on the modern Europeans genetic makeup (diamonds). Eastern Eurasian hunter-gatherers and Near Easterners are represented by squares. Grey arrows indicates the model for populating Europe inferred from modern DNA analysis, while black ones represents additional layers of information inferred based on aDNA data.

To summarize, ancient DNA studies support a complex nature for the diffusion of agriculture, providing convincing evidence for the spread of farmers at the beginning of the Neolithic and indicating the presence of genetic impact hunter-gatherers on the European gene pool. In this respect, further discoveries of ancient

samples combined with the modern data from indigenous populations of considered regions could disentangle this matter.

As the only present-day population assumed autochthonous for the Armenian Highland is represented by Armenians, the population is considered for the study of genetic makeup of the region. The genetic structure of Armenians was analyzed based on paleoanthropological, dermatoglyphics, isoserological data and using various genetic markers, such as the NRY, mtDNA, Alu insertion polymorphisms, and autosomal markers [Nasidze I and Stoneking M, 2001; Weale ME, et al., 2001; Nasidze I, et al., 2004; Harutyunyan A, et al., 2009; Herrera KJ, et al., 2012, Rootsi S, et al., 2012; Hovhannisyan A, et al., 2014]. However, the results of uniparental genetic analyses are relatively different, suggesting the presence of gender-specific population events [Nasidze I, et al., 2004]. The comparison of Y-chromosomal and mtDNA diversity within Armenian populations has identified more expressed geographic pattern for paternal rather than maternally inherited markers. This indicates the significant impact of the cultural tradition of patrilocality on population structure [Harutyunyan A, et al., 2009]. Moreover, results of a previous study pointed to the marked genetic divergence of populations from the mountainous southern (Syunik, Karabakh) and eastern areas of the Armenian Highland compared to those of the Ararat valley and the northern, western and central parts of historical Armenia [Weale ME, et al., 2001]. In addition, the HLA data survey confirmed a significant geographic stratification of the Armenians and supported their indigenous origin [Margaryan A, et al., 2011]. A recent high resolution survey of the patrilineal genetic structure of four Armenian populations of Ararat valley, Gardman, Lake Van and Sasun, which encompass the full extent of historical Armenia, revealed the greatest prevalence of Y-chromosomal haplogroups associated with the arrival of the Neolithic farmers from the Near East, including the J2a-M410-, R1b1b1*-L23-, G2a-P15- and J1-M267-derived lineages [Herrera KJ, et al., 2012]. It has been proposed that the wide presence of the Neolithic genetic markers, coupled with the scarcity of Paleolithic archaeological artifacts, indicate a repopulation of the Armenian Highland

by early agriculturalists from the Near East. Moreover, due to the presence of high sub-haplogroup diversity and deep phylogenetic branches the haplogroup G is thought to have originated nearby eastern Anatolia, Armenia or western Iran [Rootsi S, et al., 2012]. High-throughput sequencing of complete Armenian mtDNA genomes combined with the same data of the populations of Azeri, Georgians, Turks and Iranians revealed extraordinary high genetic diversity in the groups [Schönberg A, et al., 2011]. Moreover, Indo-European speaking Armenians, Turkic-speaking Azeri and Caucasian-speaking Georgians are appeared to be in one cluster on the PCA plot, indicating a greater impact of geographic features rather than linguistic relationships on the population interactions. These results are in agreement with previous studies of mitochondrial and Y-chromosomal variation [Nasidze I and Stoneking M, 2001; Nasidze I, et al., 2004].

Currently, few genome-wide association studies of modern Armenian population have been performed [Behar DM, et al., 2010; Yunusbayev B, et al., 2012; Haber M, et al., 2015]. It was shown that Armenians represents the only population in the Caucasus that share Near/ Middle Eastern component and, at the same time, follows clustering patterns similar to the Caucasus and the Levantine ones [Yunusbayev B, et al., 2012; Haber M, et al., 2015]. Statistical analysis of modern genetic data indicated the absence of any significant admixture events for Armenians over the past 4 KYA [Hellenthal G, et al., 2014] and, thus, justifies using the population as a reference group for addressing the issue of Neolithic migration from the Near East. More recent massive genetic survey of 173 Armenians revealed that the population coupled with other Near Eastern genetic isolates (Cypriots, Near Eastern Jews and Christians) have higher genetic affinity to Neolithic European farmers rather than to modern Near Eastern populations [Haber M, et al., 2015]. Thus, in comparison to other present-day Near Eastern populations, the genetic isolates preserve the features of ancient regional genetic landscape. More specifically, it was assessed that Armenians have 29% of ancestry from source population, represented by Neolithic Europeans. In general, these results points to the importance for genetic studies of continually

isolated populations, such as Armenians. It was shown that the period of Bronze Age was the most crucial one for the genetic structuring of Armenians and since this era (3,000 year ago) the population was genetically isolated from other neighboring populations. The only study of ancient Armenian data was performed by Allentoft et al, who sequenced eight Bronze Age samples from the eastern parts of the Armenian Highland [Allentoft ME, et al., 2015].

Summarizing the literature review, it can be argued that the Neolithic was a crucial period not only in terms of invention and spread of agriculture, but also for mass human migration from the Near East to Europe and the North Caucasus. In this context, the Armenian Highland, located on the crossroads of the major waves of prehistoric and historic migrations, could have served as a conduit for Neolithic human spread. The only appropriate population of the region is represented by Armenians, whose gene pool might preserve the signals of Neolithic migration through the Armenian Highland.

CHAPTER 2. MATERIALS AND METHODS

2.1. Description of the material

The study was performed at the Institute of Molecular Biology of the Armenian National Academy of Sciences (IMB NAS RA). The study protocol was approved by the Ethics Committee of the IMB NAS RA (IORG number 0003427, Assurance number FWA00015042, IRB number 00004079). The genotyping of the Armenian samples analysed here was conducted in the Lebanese American University, Beirut, Lebanon; University of Arizona, Tucson, Arizona, and the Estonian Biocentre, Tartu, Estonia.

Samples for Y-chromosomal analysis

Buccal swabs were collected with informed consent from a total of 757 unrelated (at the paternal grandfather level) self-identified ethnic Armenian males, representing four geographically distinct Armenian regions of the historical expanse of Armenia. These regions include Salmast (n=199), eastern (Karabakh and Syunik) (n=210), central (Alashkert and Bayazet) (n=200), and western (n=148) parts of the Armenian plateau. All subjects were notified about the aim of the study and gave their consent to participate. Further, in order to roughly encompass the whole region for analysis, we used the previously published data from the populations of Van (n=103), Sasun (n=104), the Ararat Valley (n=110), and Gardman (n=96) [Herrera KJ, et al., 2012], with the latter two, along with Karabakh and Syunik, then included in one group representing the eastern part of the Armenian Highland (Fig. 2.1). To assess the frequency and diversity distribution of encountered Y-chromosomal haplogroups, we combined our data with previously published comparative datasets representing the Near East, the North Caucasus, and Europe.

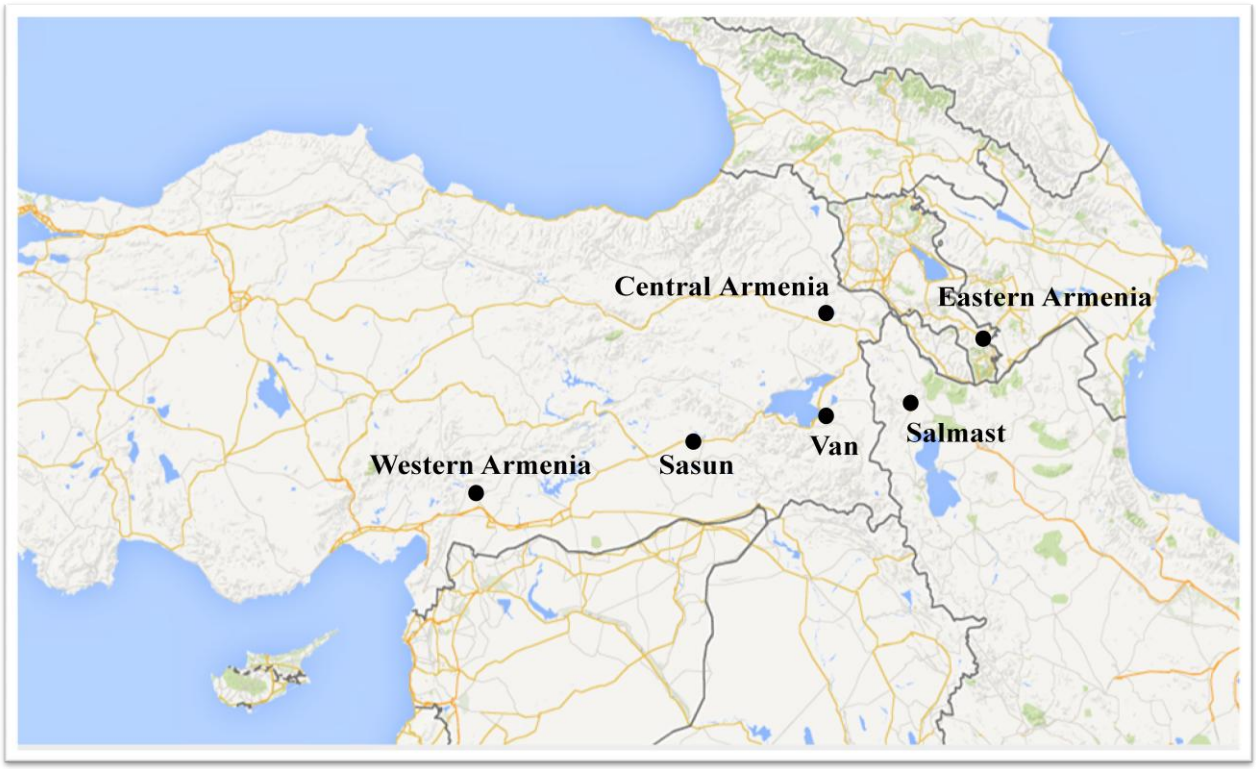


Figure 2.1. Geographic locations of the Armenian populations used for Y-chromosomal analysis.

Y-SNP and Y-STR genotyping

The genotyping was performed in a hierarchical manner for the Y-chromosomal binary (SNP) markers and for STRs. The samples of western and central Armenia, Karabakh, and Syunik were genotyped at the Lebanese American University for 32 SNPs and 17 STRs. The genotyping of Salmast specimens was performed at the University of Arizona for 44 SNPs and 14 STRs. Nomenclature of Y-chromosomal haplogroups was assigned in accordance with ISOGG (http://www.isogg.org). In order to unify the number of haplogroups and STR markers while doing comparative analysis, we used 24 haplogroups for analysis within the Armenian populations (Fig. 3.1), nine haplogroups – in comparison with other ethnic groups (Table 2.1) and the following eight common STRs for all other cross-comparisons: DYS19, DYS389I, DYS389b, DYS390, DYS391, DYS392, DYS393, and DYS439.

Table 2.1.
List of populations analyzed for the Y-chromosomal markers.

Geographic region/ population name	Abbr.	N	Reference	Available data	Region of origin
Armenian Highland					
Western Armenia	WA	148	Present study	SNPs, STRs for hgs R1b1a2, J2, G	Armenia
Central Armenia	CA	200	Present study	SNPs, STRs for hgs R1b1a2, J2, G	Armenia
Eastern Armenia	EA	416	Present study; Herrera et al, 2012	SNPs, STRs for hgs R1b1a2, J2, G	Armenia
Van	VAN	103	Herrera et al, 2012	SNPs, STRs for hgs R1b1a2, J2, G	Armenia
Sasun	SAS	104	Herrera et al, 2012	SNPs, STRs for hgs R1b1a2, J2, G	Armenia
Salmast	SAL	199	Present study	SNPs, STRs for hgs R1b1a2, J2, G	Armenia
Caucasus					
Georgia	GEO	66	Yunusbayev et al, 2012; Rootsi et al, 2012	SNPs, STRs for hg G	Georgia
Ossets-Iron	OS-I	230	Balanovsky et al, 2011	SNPs, STRs for hg J2	Russia
Ossets-Digor	OS-D	127	Balanovsky et al, 2011	STRs for hg R1b1a2	Russia
Abkhazians	ABH	162	Yunusbayev et al, 2012; Rootsi et al, 2012; Balanovsky et al, 2011	SNPs, STRs for hg G	Abkhazia
Lezghins	LEZ	81	Balanovsky et al, 2011	STRs for hg R1b1a2	Russia
Cherkessians	CHK	126	Yunusbayev et al, 2012; Rootsi et al, 2012	STRs for hg G	Russia
Chechen (Dagestan)	CHR	100	Balanovsky et al, 2011	STRs for hg J2	Russia
Near East					
Lebanon	LBN	916	Zalloua et al, 2008a	SNPs, STRs for hgs R1b1a2, J2, G	Lebanon
Syria	SYR	200	Zalloua et al, 2008b	SNPs, STRs for hgs J2, G	Syria
Palestine	PSE	290	Zalloua et al, 2008b	SNPs, STRs for hgs G	Palestine
Iran					
Iran	IRN	324	Haber et al, 2011	STRs for hg J2, G	Iran
Anatolia					
Turkey	TUR	617	Rootsi et al, 2012 Cinnioğlu et al, 2004	SNPs, STRs for hgs R1b1a2, J2, G	Turkey

Europe

Cyprus	CYP	165	Zalloua et al, 2008b	SNPs, STRs for hgs J2, G	Cyprus
Malta	MLT	187	Zalloua et al, 2008b	STRs for hg J2	Malta
Sicily	SIC	236	Di Gaetano et al, 2009	SNPs, STRs for hg J2	Italy
Crete	CRE	193	King et al, 2008	SNPs, STRs for hg J2	Greece
Greece	GRC	171	King et al, 2008	STRs for hg J2	Greece
Slovakia	SLK	276	Myres et al, 2010	STRs for hg R1b1a2	Slovakia
Romania	ROM	330	Myres et al, 2010	STRs for hg R1b1a2	Romania
Basques	BSQ	168	Alonso et al, 2005	SNPs	Spain
Belgium	BEL	1028	Larmuseau et al, 2014	SNPs	Belgium
England	GBR	102	Myres et al, 2010	STRs for hg R1b1a2	England
France	FRA	93	Myres et al, 2010	STRs for hg R1b1a2	France
Netherlands	NLD	87	Myres et al, 2010	STRs for hg R1b1a2	Netherlands
Switzerland	CHE	175	Myres et al, 2010	STRs for hg R1b1a2	Switzerland
Germany	DEU	321	Myres et al, 2010	STRs for hg R1b1a2	Germany
Denmark	DNK	110	Myres et al, 2010	STRs for hg R1b1a2	Denmark
Italy	ITA	282	Myres et al, 2010	STRs for hg R1b1a2	Italy
Ireland	IRL	99	Myres et al, 2010	STRs for hg R1b1a2	Ireland

Samples for genome-wide autosomal analysis.

Blood samples were collected in 4 ml EDTA-coated vacutainers with informed consent from a total of 27 unrelated (at the paternal and maternal grandfather and grandmother level) self-identified ethnic Armenians, representing three geographically distinct Armenian regions: western (n=4), eastern (n=14), and central (n=9) parts of the Armenian Highland. As comparative datasets from the region, 8 Georgian and 10 Yezidis samples were also genotyped (Fig. 2.2).

DNA extraction for genome-wide analysis

3-4 ml of blood sample was added in a 15 ml Falcon tube and centrifuged at 3,000 rpm for 10 min. After removal of supernatant 14 ml TE buffer was added and centrifuged at 3,000 rpm for 10 min. This step was repeated until a clear white pellet was obtained. 25 µl of Proteinase K (20 mkg/ml) and 5 ml of WBC lysis buffer were added and the pellet was disturbed. The tubes were incubated at 56ºC for 2.5-3 hours. Then 2 ml 6 M NaCl was added, the tubes were shaking periodically for 20 minutes (3-4 times) and centrifuged at 3,000 rpm for 25-30 min. This step was repeated until a

clear supernatant and rigid pellet were obtained. To precipitate DNA aqueonus phase was added into the 50 ml transparent clear tubes with 35-40 ml 96% ethanol. The tubes were gently shaked until the DNA medusa was generated. Transferred in 1.5 ml sterile eppendorf tubes, the DNA medusa was washed in 1 ml 70% ethanol. After drying 350 µl TE buffer (pH=8.0) was added and the tubes with DNAs were put at 37ºC overnight to solve DNAs. DNA samples were kept at -20 ºC.

Figure 2.2. Geographic location of the Armenian and comparative populations (Georgians and Yezidis) used for genome-wide autosomal analysis.

DNA purification for genome-wide analysis

Ethanol precipitation was used to concentrate and purify nucleic acids. 1/10 volume of 7.5 M ammonium acetate was added to 200-300 mkl of DNA samples. Then 2.5–3.0X volume (calculated after addition of ammonium acetate) of at least 95% ethanol was added and tubes were freezed overnight in -20. Samples were centifuged at > 14,000 rpm for 20 minutes, supernatant was discarded and 1,000 mkl of 70% ethanol was added. After 5 min of centrifugation supernatant was decanted, the pellet was dried and 50 mkl of H_2O was added in the tubes to dissolve the pellet.

Genome-wide typing

DNA concentration was adjusted to approximately 70 ng/µl. Before genotyping the quality of the samples was checked by electrophoresis (0.8% gel). DNA samples were genotyped with the Illumina 710 K single nucleotide polymorphism (SNP) array and used for genome-wide autosomal analysis together with comparative ancient and modern samples obtained from the literature (Tables 2.2 and 2.3).

Table 2.2.
List of ancient samples analyzed in this study for genome-wide autosomal markers.

Group	Description	N	Reference	Region
Late Neolithic – Bronze Age				
Yamnaya	Yamnaya (Bronze Age)	5	Haak et al 2015	Russia
aArm	Armenians (Bronze Age)	8	Allentoft et al 2015	Armenia
Early – Middle Neolithic				
Hungary_EF	Early Farmers	2	Haak et al 2015; Gamba et al 2014	Hungary
LBK_EF	Early Farmers	2	Haak et al 2015	Germany
Stuttgart_EF	Early Farmer	1	Lazaridis et al 2014	Germany
Pleistocene/Holocene Hunter-Gatherer				
EHG	Eastern Hunter-Gatherers	2	Haak et al 2015; Mathieson et al 2015	Russia
CHG	Caucasus Hunter-Gatherers	2	Jones et al 2015	Georgia

Table 2.3.
List of modern populations analyzed for genome-wide autosomal markers.

Geographic region/population	Abbr.	N	Reference	Chip	Analysis	Country of origin/Collection
Armenian Highland						
Armenians (general)	Arm	19	Behar et al 2010	610k	ADMIXTURE	Armenia
Armenians (general)	Arm	16	Yunusbayev et al 2012	660k	ADMIXTURE	Armenia
Central Armenia	CA	9	Present study	710k	ADMIXTURE, PCA	Armenia
Eastern Armenia	EA	14	Present study	710k	ADMIXTURE, PCA	Armenia

Western Armenia	WA	4	Present study	710k	ADMIXTURE, PCA	Armenia
Caucasus						
Abkhasians	Abh	20	Yunusbayev et al 2012	610k	ADMIXTURE	Georgia
Chechens	Che	20	Yunusbayev et al 2012	610k	ADMIXTURE, PCA	Russia
Lezgins	Lez	18	Behar et al 2010	610k	ADMIXTURE	Russia
North Ossetians	OsN	15	Yunusbayev et al 2012	660k	PCA	Russia
Georgians	Geo	8	Present study	710k	ADMIXTURE, PCA	Georgia
Anatolia						
Turks	Tur	19	Behar et al 2010	610k	ADMIXTURE, PCA	Turkey
Iran						
Azeris	Az	10	Yunusbayev et al 2015	660k	PCA	Iran
Iranians	Irn	20	Behar et al 2010	610k	ADMIXTURE, PCA	Iran
Central Asia						
Kurds	Krd	5	Yunusbayev et al 2012	660k	ADMIXTURE, PCA	Kazakhstan
Uzbeks	Uzb	15	Behar et al 2010	610k	ADMIXTURE	Uzbekistan
Turkmens	Trk	15	Yunusbayev et al 2012	610k	ADMIXTURE	Turkmenistan
Tajiks	Taj	15	Yunusbayev et al 2012	610k	ADMIXTURE	Tajikistan
Near East						
Jordanians	Jor	20	Behar et al 2010	660k	ADMIXTURE	Jordania
Yezidis	Yez	10	Present study	710k	ADMIXTURE, PCA	Armenia
Syrians	Syr	16	Behar et al 2010	610-660k	ADMIXTURE, PCA	Syria
Lebanese	Leb	8	Behar et al 2010; Behar et al 2013	610k	PCA	Lebanon
South Asia						
Pathan	Pat	23	Li et al 2008	650k	ADMIXTURE	Pakistan
Gujaratis	Guj	20	Hapmap3	Hap map	ADMIXTURE	India
Siberia						
Mongolians	Mng	9	Rasmussen et al 2010	660k	ADMIXTURE	Mongolia
Buryats	Bur	19	Rasmussen et al 2010	660k	ADMIXTURE	Russia
Europe						
Ukranians	Ukr	20	Yunusbayev et al 2012	660k	ADMIXTURE	Ukraine
Greeks	Grc	10	Behar et al 2013	660k	ADMIXTURE	Greece
France	Fr	20	Li et al 2008	650k	ADMIXTURE, PCA	France
Belarusians	Bel	22	Behar et al 2010; Behar et al 2013	610-660k	ADMIXTURE	Belarusia
Germans	Ger	13	Yunusbayev et al 2015	660k	PCA	Germany
Estonians	Est	20	Raghavan et al 2014	650k	ADMIXTURE	Estonia

Sicilians	Sic	13	Behar et al 2013	660k	ADMIXTURE	Italy
North Italians	Ita	13	Li et al 2008	650k	ADMIXTURE	Italy
Tuscans	Tus	8	Li et al 2008	650k	ADMIXTURE	Italy
Bulgarians	Bul	13	Yunusbayev et al 2012	660k	PCA	Bulgaria

2.2. Data analysis

Methods for Y-chromosomal analysis

The values of pairwise genetic distances (F_{ST}) were calculated using the software package Arlequin 3.5 [Excoffier L, et al., 2005]. We also estimated the intra-population locus-specific variance, V_L, and the intra-population genetic variance, V_P, according to the formulae given in Kayser et al. (2001):

$$V_L = [1/(n-1)]\Sigma_{i=1}^{n}(Xi - X)^2$$

$$V_P = (1/m)\Sigma_{j=L}^{m} V_{Lj}$$

where $X = (1/n)\Sigma_{i=1}^{n} X_i$

n – number of chromosomes sampled for the population

m – number of loci.

Frequencies and microsatellite variances of the haplogroups were visually displayed using Surfer 10 (Golden Software) by the gridding method. Latitude and longitude values were calculated for the geographic centers of the sampling regions.

Principal coordinate analysis (PCoA) was performed on distance matrices based on F_{ST} genetic distances using Genstat software.

The phylogenetic relationships among eight loci haplotypes of equal number of individuals from different populations within the haplogroups R1b1a2, J2, and G were ascertained using the NETWORK 4.6.1.0 (available at http://www.fluxus-engineering.com) and Network Publisher softwares. Median-joining networks were generated by processing haplotypes with the reduced median algorithm, followed by the median-joining method, and with weighted STR loci tabulated to be proportional to the inverse of the repeat variance. The star contraction and maximum parsimony options were used for further simplification of the networks.

GENE-E software was used to graphically represent genetic similarities between populations by color coding pairwise F_{ST} values on a heatmap.

To estimate differences in the haplogroup composition of the regions, correspondence analysis was conducted using SPSS ver. 19 software package (SPSS Inc.).

Methods for genome-wide autosomal analysis.

Data management and quality control. We used PLINK 1.05 software [Purcell S, et al., 2007] to merge our genotyped samples with the comparative dataset consisted of different modern and ancient populations that were genotyped with various arrays in previous genetic studies. Quality control and linkage disequilibrium (LD) pruning were also performed using the corresponding software. Sex-linked and mitochondrial SNPs were removed for all subsequent analysis.

Genetic ancestry analysis. Only SNPs with genotyping success >95% were included for the analysis. In general, 164,778 autosomal SNPs were used to infer individual genomic ancestry components in modern population samples by ADMIXTURE software [Alexander DH, et al., 2009]. The program was run assuming 2 to 8 (K = 2 to K = 8) genetic clusters or "ancestral populations" in 10 replicates and assessed convergence between individual runs. Individual ancestries were estimated by efficiently computing maximum likelihood estimates in a parametric model. On the basis of the cross-validation error distribution, the model at K=6 was assessed as having better predictive accuracy than at other K values. In order to evaluate the convergence of conducted runs at each K value, the difference in the distribution of log likelihood (LL) scores for the runs was considered. The low level of variation in LLs within each runs was reached at K=6, that verify the usage of this particular K as the best single representation of the ADMIXTURE genetic structure for our data.

***Principal component analysis*.** In order to prepare our data of modern population samples for principal component analysis (PCA), quality control (minor allele frequency >1%; genotyping success >97%; per-individual genotyping threshold <3%) and LD-pruning (the method assumes that considered markers are

unlinked) were conducted. LD score (r^2) was calculated for each pair of SNPs in 200-SNP windows, excluding one of a pair of SNPs if the LD is greater than 0.5. Thus, the filtering steps resulted in 178878 SNPs within modern population data remaining for downstream analyses. PCA was performed for the autosomal genome-wide dataset using the smartpca program of the EIGENSOFT package with the outlier removal option off. [Patterson N, et al., 2006]. Data of ancient DNA samples were projected onto the pattern containing the modern ones. No standardization or transformation of genotypes was performed before running the software. The R package [Team RC, 2013] was used to graphically represent genetic similarities between the populations. In order to present the results at the population level, population medians for PC coordinates were assessed. Plot of the first and second components of the PCA was visualized.

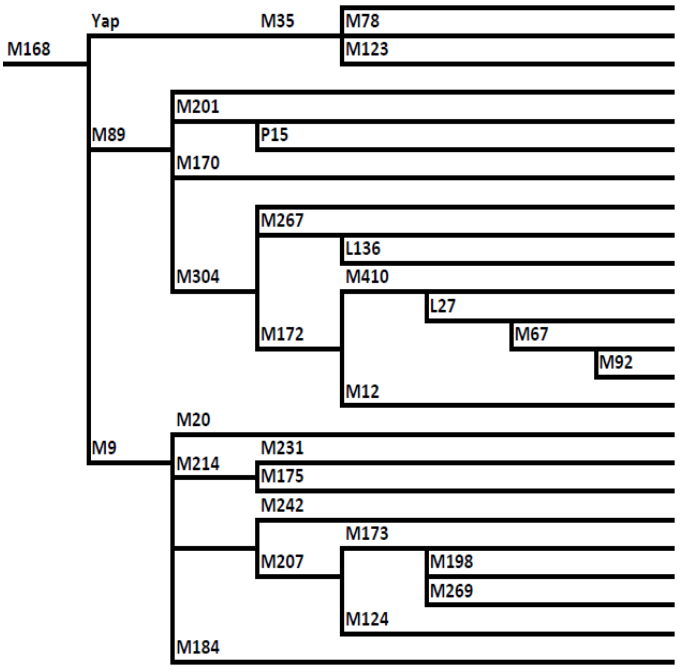
CHAPTER 3. RESULTS AND DISCUSSION

3.1. Y-haplogroup frequency distribution

The phylogenetic relationships of the Ychromosomal markers and frequency distribution of the defined 24 haplogroups in the six Armenian populations are shown in Figure 3.1. Analysis of haplogroup distribution revealed that the lineage R1b1a2-M269 is the most frequently encountered subclade in all Armenian samples, except Sasun, which differs from others due to the predominance of the haplogroup T (20%) [Herrera KJ, et al., 2012]. Of the lineages within the haplogroup R, the subclade R1a1a-M198 is linked to the spread of the Indo-Aryan languages [Underhill PA, et al., 2010] and detected with low frequencies or even absent in the analyzed populations. The majority of the J-M304 samples within Armenian populations belongs to its J2-M172 branch, though in the population of Salmast, there is a nearly equal frequency distribution of J2 and J1 lineages. An aggregated frequency of all J-M304 sublineages encompasses 27-43% of the Y-chromosomes, in which 51-70% account for the haplogroup J2. The haplogroup G is also observed at relatively high frequencies in all Armenian samples (Fig. 3.1). Worth mentioning that Y-chromosomal haplogroups frequently encountered in Balkans populations (the lineages E-M78, I-M170, and R-M198) are not among the prevalent ones in the Armenian paternal gene pool. Moreover, the haplogroup N, which is considered as a genetic marker of Mongolian ethnic groups, is presented only in one sample from the population of Gardman, while being absent in all other Armenian groups.

On the whole, the results of analysis of patrilineal lineages revealed the prevalence of the Y-chromosomal haplogroups associated with the arrival of Neolithic farmers from the Near East. Three prospective genetic markers of agricultural migration, namely, the haplogroups J2, G, and R-M269, represent the most common lineages in all six Armenian populations, together accounting for 49-70% of the sampled groups. It has previously been proposed that the wide presence of genetic markers attributed to agriculturalists, coupled with Neolithic archaeological artifacts, indicates continuous habitation of the Armenian Highland since the

Neolithic [Dolukhanov P, et al., 2004; Herrera KJ, et al., 2012]. It was suggested that the majority of Paleolithic genetic markers detected in the modern populations of the Armenian Highland represent genetic influx from continuously settled regions [Herrera KJ, et al., 2012]. This inference is confirmed by time estimations of the haplogroup T, which is considered to have originated in the Near East during the Paleolithic (20 KYA) and later introduced by migrants in the territory of the Armenian Plateau (12-13 KYA).



	WA (n=148)	CA (n=200)	EA (n=416)	Van (n=103)	Sasun (n=104)	Salmast (n=199)
E1b1b1	-	-	-	-	-	0.0101
E1b1b1a1	0.0270	0.0200	0.0096	0.0388	-	0.0704
E1b1b1b2a	0.0135	0.0250	0.0361	0.0388	0.0288	0.0402
F	-	0.0050	-	-	-	-
G	0.0203	0.0150	0.0144	-	-	0.0201
G2a	0.1149	0.1450	0.0721	0.0777	0.1250	0.0553
I	0.0405	0.0350	0.0361	0.0194	-	0.0151
J	-	-	-	-	-	0.0101
J1	0.0608	0.0800	0.0697	0.0777	0.0673	0.0704
J1a2	0.0608	0.0900	0.0625	0.0583	0.0288	0.1256
J2a	0.1959	0.1450	0.1274	0.1942	0.0962	0.0050
J2a1	0.0068	-	-	-	-	0.1206
J2a1b	0.0676	0.0700	0.0625	0.0680	0.0769	0.0503
J2a1b1	-	-	0.0409	0.0097	-	0.0251
J2b	0.0068	0.0400	0.0144	0.0194	-	0.0101
L	0.0203	0.0050	0.0096	0.0194	0.0385	0.0201
N	-	-	0.0024	-	-	-
O	0.0068	-	-	-	-	-
Q	-	-	0.0048	0.0097	-	0.0151
R1	-	0.0050	0.0024	-	-	0.0050
R1a1a	0.0541	0.0300	0.0240	-	0.0096	0.0553
R1b1a2	**0.2230**	**0.2150**	**0.3606**	**0.3204**	0.1538	**0.2362**
R2a	0.0338	0.0150	0.0048	0.0097	0.1731	0.0151
T	0.0473	0.0600	0.0457	0.0388	**0.2019**	0.0251

Figure 3.1. Phylogenetic relationships and Y-chromosome haplogroup frequencies in six Armenian populations.

To get deeper insight into the issue of the directions of the agriculturalists' migration from the Near East, we used the PCoA approach to visualize the F_{ST} genetic distances (based on absolute haplogroup frequencies, summarized in Table 3.1) between the Armenian and comparative datasets from the Near East, the North Caucasus, and Europe (Table 3.2).

Table 3.1.

Absolute frequency distribution of haplogroups in the studied populations. Frequencies of the main Y-chromosomal haplogroups in the 18 populations include in the PCoA and correspondence analysis of Figures 3.2 and 3.3.

HG	POPULATION																	
	WA n=148	CA n=200	EA n=416	VAN n=103	SAS n=104	SAL n=199	LBN n=916	SYR n=200	PSE n=290	CYP n=165	TUR n=523	OS-I n=230	ABH n=58	CRE n=193	GEO n=65	BSQ n=168	BEL n=1028	SIC n=236
E1b1b1-M35	6	9	19	8	3	24	148	27	71	51	56	1	1	17	1	6	52	43
ExE1b1b1	0	0	0	0	0	0	13	7	3	2	3	0	0	0	0	1	0	0
G-M201	20	32	36	8	13	15	60	11	25	22	57	171	33	21	33	1	38	14
I-M170	6	7	15	2	0	3	44	9	13	1	28	0	0	25	2	5	201	18
J2-M172	41	51	102	30	18	42	237	52	56	25	127	42	8	59	19	1	44	61
JxJ2	18	34	55	14	10	41	184	48	101	38	48	3	0	16	0	2	10	9
L-M20	3	1	4	2	4	4	48	10	2	3	22	2	2	1	0	0	4	1
R1b1a2-M269	33	43	150	33	16	47	67	8	5	0	76	6	7	33	1	144	626	58
RxR1b1a2	13	10	13	1	19	15	44	17	5	20	50	2	6	17	9	4	42	13
Other	8	13	22	5	21	8	71	11	9	3	56	3	1	4	0	4	11	19

The PCoA plot reveals strong regional clustering indicating the separation of the populations from the Near East and Eastern Europe from those of the North Caucasus and Western Europe (Fig. 3.2). In this context, populations of the Armenian Highland, the Near East, and Eastern Europe appear to be in one extensive cluster with a clear geographic gradient from the Levant towards the northwest. In fact, the closest population to the Near Eastern groups is Cyprus, the region settled by Neolithic farmers from the mainland shortly after the emergence of agriculture [Vigne JD, et al., 2012]. Archaeological discoveries at Klimonas indicated that the use of cereals introduced from the Levant and appearence of Neolithic villages on the island dates to approximately 10-11 KYA. The Cretan population within the cluster is centrally located between the populations of the Near East and Europe. It was shown that Crete hosts one of the oldest Neolithic settlements of Europe and underwent an agricultural transition from either the Anatolian coast or by sea from the Levant approximately 7–8 KYA [Evans JD, et al., 1994; Bellwood P, 2013]. In general, our pattern is in accordance with previously found genetic affinity between human remains from the Neolithic sites (based on aDNA data) and the modern populations of Cyprus and Crete, suggesting the leading role of pioneer seafaring colonization in the expansion towards the rest of Europe [Davison K, et al., 2006; Fernández E, et al., 2014]. Specifically, our results of the PCoA analysis support a key role for Crete in the spread of Neolithic farmers through maritime routes from the Near East to Europe, which is also confirmed by pairwise F_{ST} value comparisons based on haplogroup frequencies (Table 3.2). The plot on Figure 3.2 clearly separates the western European and North Caucasus populations from each other and bidirectionally from the Armenian cluster. These overall results further bolster the Armenian Highland as a corridor between the two aforementioned regions and the Near East.

Table 3.2.

Pairwise F_{ST} genetic distances between the populations studied based on the all haplogroup frequencies (P<0.05).

Pops.	POPULATION																	
	WA	**CA**	**EA**	**VAN**	**SAS**	**SAL**	**LBN**	**SYR**	**PSE**	**CYR**	**TUR**	**OS-I**	**ABH**	**CRE**	**GEO**	**BSQ**	**BEL**	**SIC**
WA	-																	
CA	-0.003	-																
EA	0.014*	0.016*	-															
VAN	0.006	0.008	-0.002	-														
SAS	0.008	0.017*	0.045*	0.041*	-													
SAL	0.009*	0.007*	0.018*	0.009	0.021*	-												
LBN	0.030*	0.028*	0.066*	0.045*	0.044*	0.019*	-											
SYR	0.037*	0.035*	0.082*	0.059*	0.043*	0.025*	0.000	-										
PSE	0.090*	0.074*	0.127*	0.106*	0.102*	0.053*	0.022*	0.020*	-									
CYR	0.086*	0.080*	0.136*	0.115*	0.075*	0.057*	0.029*	0.029*	0.019*	-								
TUR	0.004	0.010*	0.035*	0.022*	0.011*	0.013*	0.016*	0.022*	0.067*	0.053*	-							
OS-I	0.286*	0.263*	0.318*	0.362*	0.299*	0.321*	0.273*	0.327*	0.342*	0.323*	0.246*	-						
ABH	0.130*	0.124*	0.183*	0.190*	0.116*	0.168*	0.174*	0.189*	0.228*	0.184*	0.126*	0.045*	-					
CRE	0.004	0.013*	0.035*	0.023*	0.024*	0.023*	0.025*	0.032*	0.085*	0.078*	0.002	0.289*	0.145*	-				
GEO	0.108*	0.108*	0.180*	0.177*	0.100*	0.157*	0.141*	0.150*	0.201*	0.161*	0.100*	0.082*	0.017	0.110*	-			
BSQ	0.346*	0.336*	0.215*	0.319*	0.406*	0.312*	0.349*	0.436*	0.457*	0.481*	0.310*	0.668*	0.582*	0.361*	0.611*	-		
BEL	0.174*	0.181*	0.104*	0.141*	0.204*	0.168*	0.237*	0.279*	0.316*	0.317*	0.186*	0.458*	0.334*	0.187*	0.367*	0.084*	-	
SIC	0.020*	0.030*	0.027*	0.015*	0.044*	0.022*	0.039*	0.056*	0.099*	0.084*	0.014*	0.330*	0.179*	0.013*	0.159*	0.302*	0.147*	-

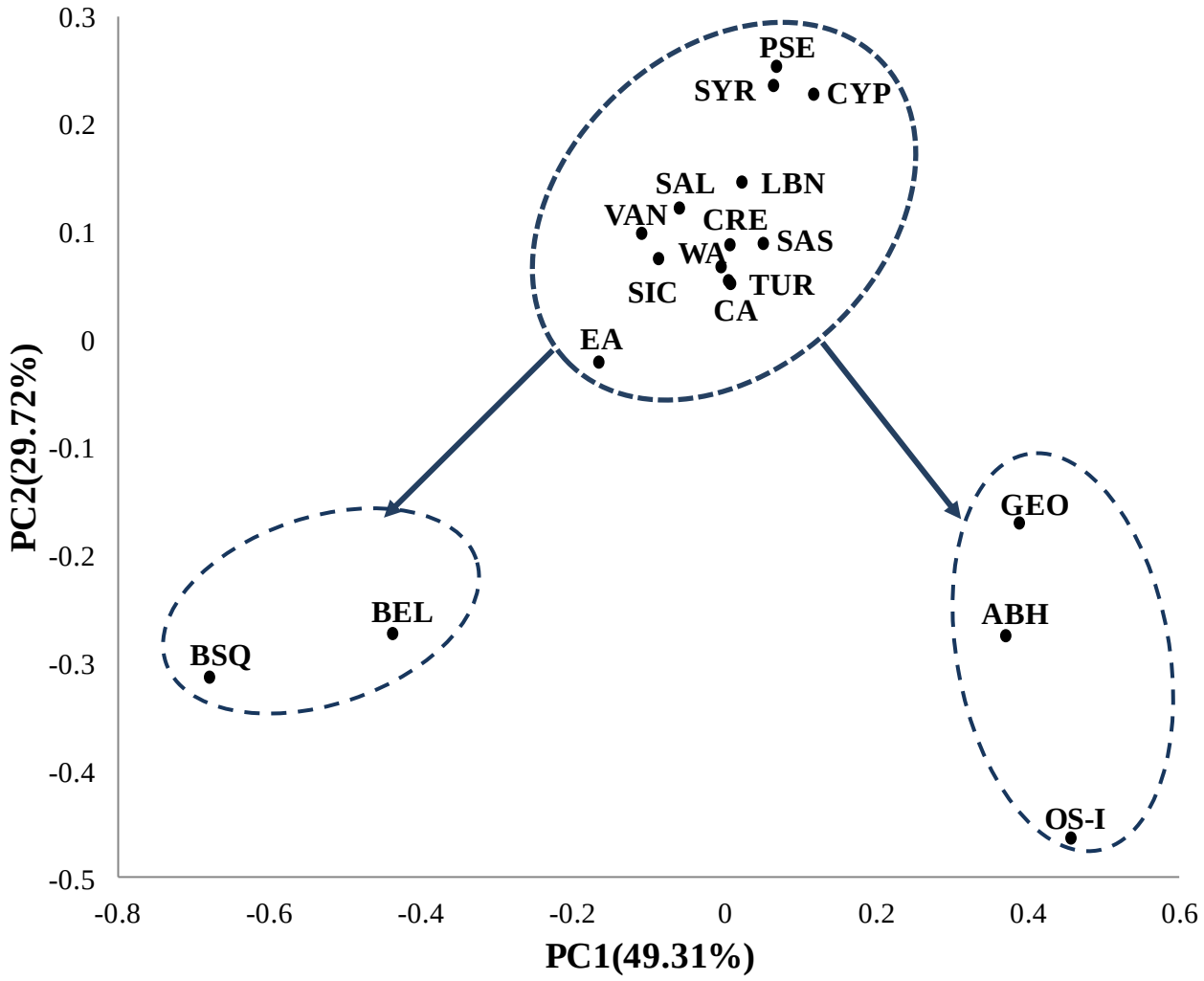


Figure 3.2. PCoA plot based on pairwise F_{ST} genetic distances calculated from haplogroup frequencies in the populations of this study.
The plot is based on F_{ST} pairwise genetic distances calculated from frequencies of nine common Y-chromosomal haplogroups [E1b1b1-M35, E(xE1b1b1), G-M201, I-M170, J2-M172, J(xJ2), L-M20, R1b1a2-M269, R(xR1b1a2)].

In order to provide a potential genetic explanation for the classification presented in Figure 3.2, we have conducted a correspondence analysis (Fig. 3.3) on the haplogroup frequency data in the populations studied (Table 3.1). On the whole, the patterns of population distribution for the correspondence analysis and PCoA are nearly identical.

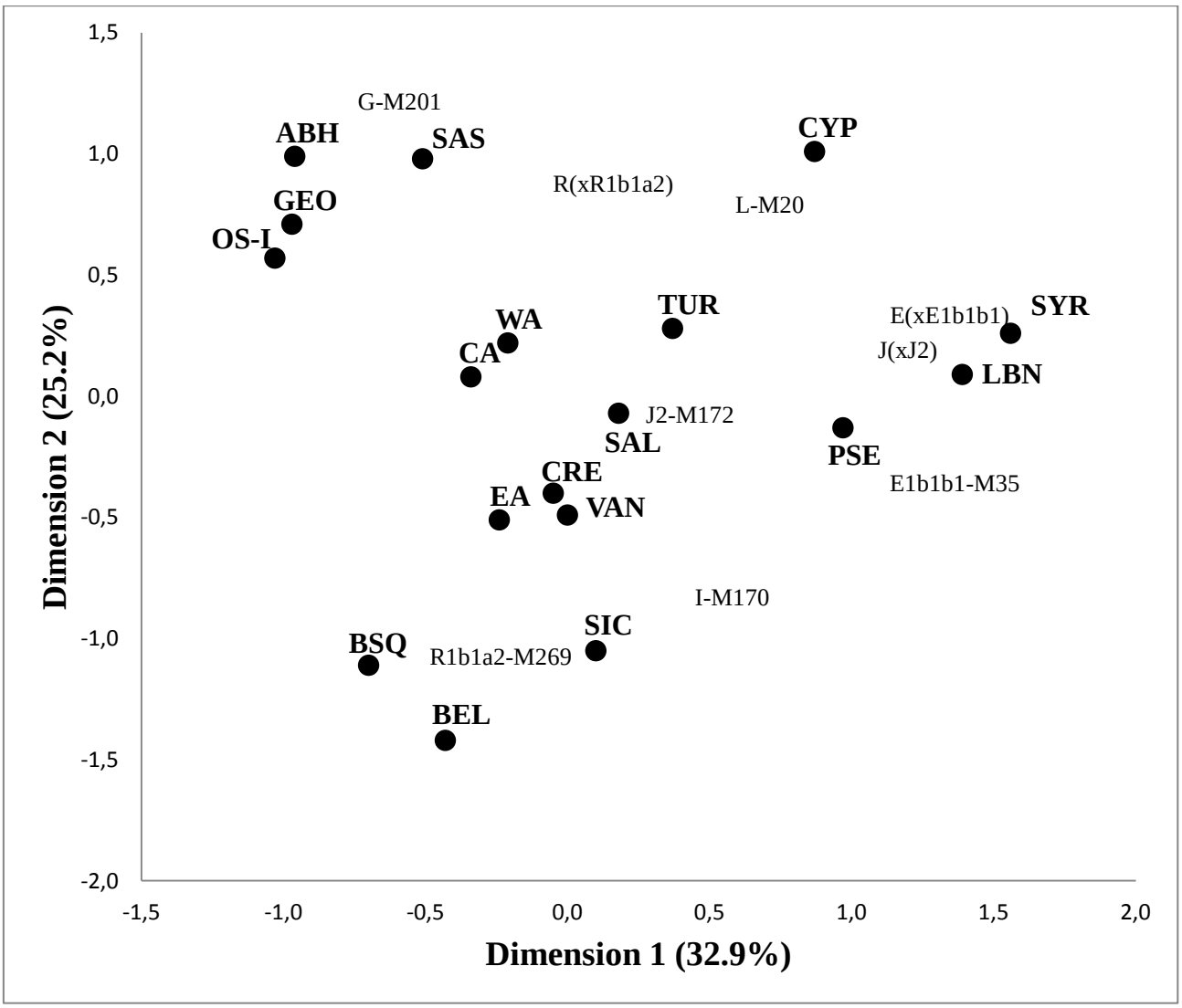


Figure 3.3. Correspondence analysis plot based on the haplogroup frequency data in the populations studied.

It was shown that the European cluster, containing the Basques, Sicilians, and Belgians, is associated with the haplogroups R1b1a2-M269 and I-M170, both widely spread in Europe, and the former being a marker for the Neolithic migration. The Caucasus cluster, comprising Abkhazians, Georgians, and Ossetians, is found to be connected to the haplogroup G-M201, which is also a marker for the Neolithic migration. The presence of the outlying Armenian population of Sasun in the vicinity of the Caucasus cluster could be explained by the geographic peculiarities of this high-mountainous group which resulted in its genetic isolation from other Armenians during the intervening centuries. Completing the analysis of the haplogroups

associated with the Neolithic agriculturalists, the lineage J2-M172 appears in between the European and Caucasus clusters.

The results of the PCoA and correspondence analysis show that the haplogroup composition of the Near Eastern populations is very similar to that found for the populations from the Anatolian and Armenian plateaus, as well as those of the Mediterranean islands. Apparently, the population migration of the first farmers from the Levant could have been both by land to Anatolia and the North Caucasus, and by maritime routes via eastern Mediterranean islands towards continental Europe. This scenario is supported by the result of the comparison of F_{ST} genetic distance values based on the frequencies of all haplogroups identified (Table 3.2), showing that the populations of the Armenian Highland display an intermediate position between the Near East and Europe, and the Near East and the North Caucasus. Though previous work based on 15 autosomal STR loci from four Armenian populations (Ararat Valley, Gardman, Sasun, and Van) [Lowery RK, et al., 2011] derived a potential Balkan origin for one of these locations (Van), the results of our analysis not only support the transition zone model of the Armenian Highlands but also the potential gene flow of some Neolithic markers, shared among Armenians and the Balkan populations, from the Near East through this region.

Further, in order to obtain deeper insight into the relationships between the populations observed and to analyze possible routes of expansion, we separately assessed the distribution patterns of putative Y-chromosomal tracers of the spread of the first agriculturalists, namely the haplogroups R1b1a2-M269, J2-M172, and G-M201.

3.2. Distribution patterns of the haplogroup R1b1a2-M269

The most frequently encountered subclade of the core Y-chromosomal lineage R-M207, haplogroup R1b1b-M269, is thought to undergone a rapid spread in the European continent, which resulted in present-day carriage of the haplogroup by 110 million European man. Interestingly, the spatial distribution of the main western

European Y-chromosomal lineage shows a significant frequency cline from 7% in Lebanon to 82% in Ireland [Myres NM, et al., 2011, Zalloua PA, et al., 2008a], though it is also present in trace amounts in the majority of the North Caucasus populations [Balanovsky O, et al., 2011].

Among the Armenian samples, the haplogroup R1b1b-M269 is one of the most common lineages, which is frequently encountered in the eastern part of the Armenian Highland and Van. The rates of frequency and genetic variance within the haplogroup among considered populations are provided in Table 3.3.

Table 3.3.
Frequency and varience distributions of the haplogroup R1b1a2.

Population	N	Freq	Vp
Armenian Highland			
WA	148	0.2230	0.4936
CA	200	0.2150	0.3509
EA	416	0.3606	0.3650
Van	103	0.3204	0.2912
Sasun	104	0.1538	0.3948
Salmast	199	0.2362	0.4712
Caucasus			
Lezghins	81	0.2963	0.1397
Ossets-Digor	127	0.1575	0.0444
Near East			
Lebanon	916	0.0731	0.3930
Anatolia			
Turkey	523	0.1453	0.3543
Europe			
Romania	330	0.0970	0.3246
Slovakia	276	0.1630	0.2654
England	102	0.7157	0.2610
France	93	0.5161	0.2877
Netherlands	87	0.5057	0.2075
Switzerland	175	0.4743	0.2937
Germany	321	0.4361	0.2888
Denmark	110	0.3636	0.2326
Italy	282	0.3723	0.2545
Ireland	99	0.8182	0.2561

In contrast, the analysis of haplogroup distribution revealed a decreasing cline of microsatellite variance from the Levant towards northwest and northeast. Furthermore, in comparison with all analyzed populations from the Near East, Europe, and Anatolia, the haplogroup R1b1a2-M269 occurs with the highest genetic variances in the western parts of the Armenian plateau, in Sasun and Salmast (Fig. 3.4).

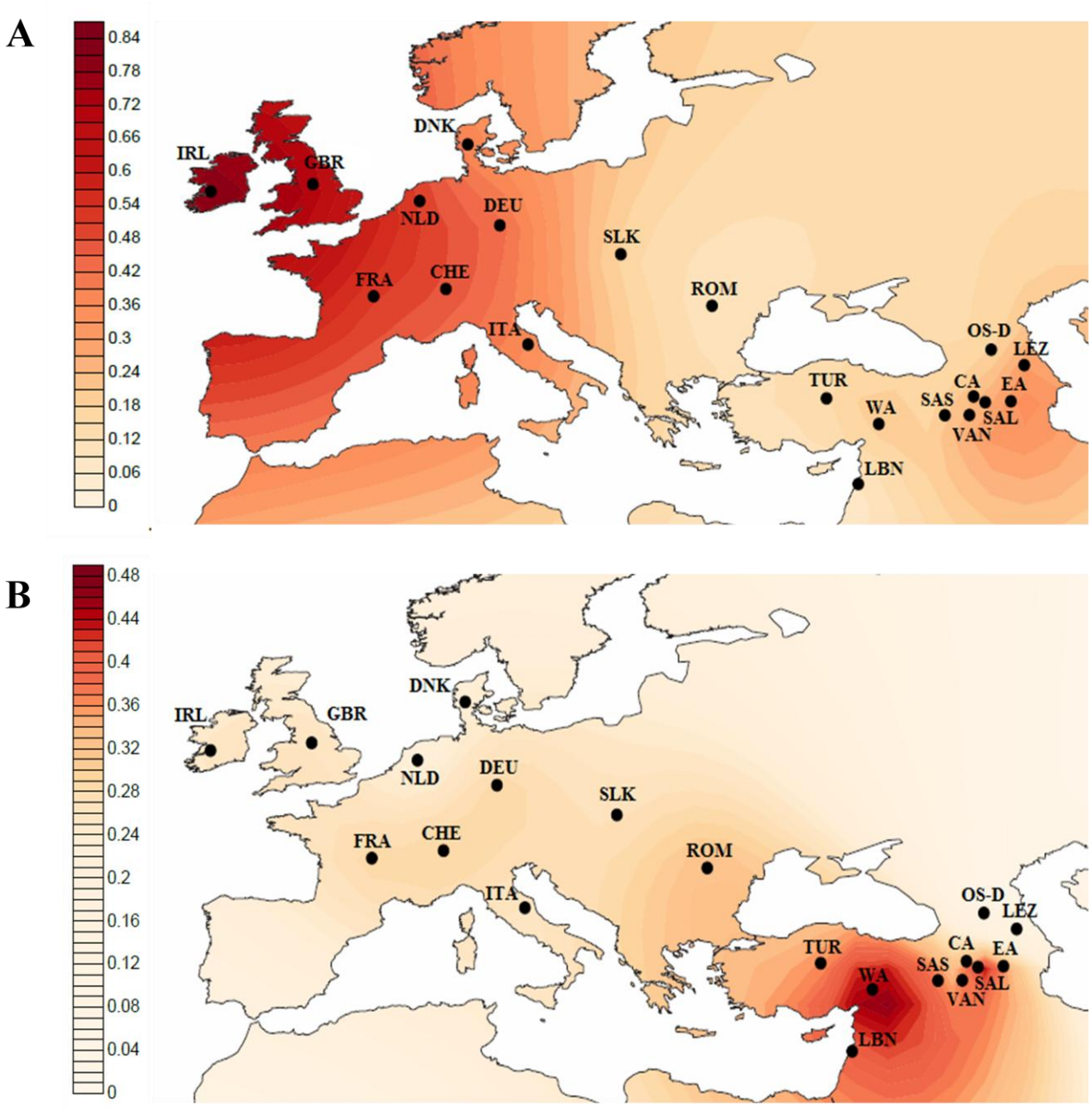


Figure 3.4. Geographical distribution maps of haplogroup R1b1a2-M269 frequencies(A) and genetic variances (B).

A heatmap plot of F_{ST} genetic distances in the frame of the haplogroup R1b1a2-M269 (Fig.3.5) reveals two large clusters with close genetic affinities. The first represents a genetic homogeneity of European populations, while the second encompasses all populations of the Near East. Generally, only the population of Sasun is slightly different within the last group, likely due to the long centuries of its aforementioned isolation by geographic barriers. Moreover, in contrast to other populations of the Near Eastern cluster, the populations of the western part of the Armenian Highland, Van, Turkey, and Lebanon show a moderate level of genetic affinity to the central European populations. Indeed, the actual estimates of the F_{ST} values for the haplogroup R1b1a2 place the western region of the Armenian Highland in a transitional position between the Near East and Europe (Table 3.4). Previous data on the limited Y-chromosomal and autosomal sharing among the Armenian and European populations [Yunusbayev B, et al., 2012; Herrera KJ, et al., 2012] should be considered as a consequence of the absence, in their Armenian datasets, of populations from the western region of the Armenian Highland.

Previously, genetic relationships between Armenians and comparative populations were considered mostly on the level of the haplogroup R1b-M343 [Herrera KJ, et al., 2012]. In this context, it was shown that the expansion time estimates in the Armenian populations (Sasun, Gardman, Ararat Valley, Van) for the haplogroup R1b-M343 (calculated using the genealogical mutation rates) are intermediate to those of the Near East and Europe [Herrera KJ, et al., 2012]. Moreover, it was detected that despite the lowest frequency rates among all considered Armenian populations, the level of microsatellite variance within the mentioned lineage in the population of Sasun is the highest, being inferior only to those for the Near Eastern populations of Jordan and Lebanon. Interestingly, Multidimensional Scaling Analysis (MDS) plot based on haplogroup frequencies revealed close proximity of the Armenian populations of Gardman and Van to that from Eastern Turkey, while the populations of Sasun and Ararat Valley are appeared to be as outliers in comparison to other studied populations. Though the ethnic groups

from the western part of the Armenian Highland were not concidered, in general, the data obtained in the study pointed to the presence of the haplogroup R1b-M343 on the territory of the Armenian Highland at least since the Neolithic. In this context, our observations of the distribution patterns for its downstream lineage, haplogroup R1b1a2-M269, support these results.

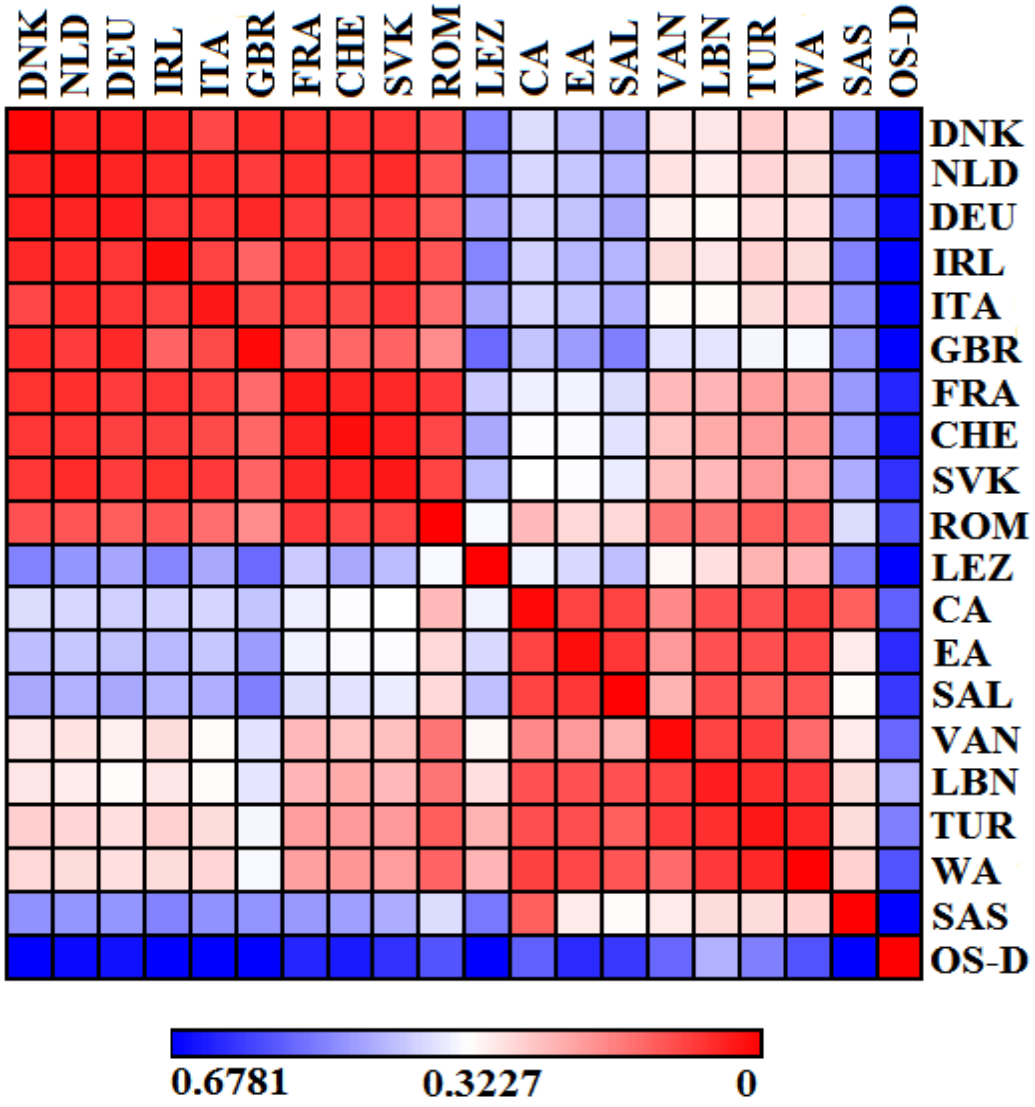


Figure 3.5. Heatmap of pairwise F_{ST} genetic distances, ranging from low (red) to high (blue), calculated for the haplogroup R1b1a2.

Table 3.4.

Pairwise F_{ST} genetic distances between the populations studied based on the haplogroup R1b1a2 (P<0.05).

Pops.	POPULATION																		
	CA	**DNK**	**EA**	**GBR**	**FRA**	**DEU**	**IRL**	**ITA**	**LBN**	**LEZ**	**NLD**	**OS-D**	**ROM**	**SAL**	**SAS**	**SLK**	**CHE**	**TUR**	**VAN**	**WA**
CA	-																			
DNK	0.228*	-																		
EA	0.037*	0.164*	-																	
GBR	0.233*	-0.002	0.168*	-																
FRA	0.239*	0.003	0.162*	0.005	-															
DEU	0.240*	-0.004	0.181*	0.000	0.013*	-														
IRL	0.236*	0.030*	0.181*	0.008	0.026*	0.013*	-													
ITA	0.202*	0.012	0.121*	0.009*	0.013*	0.018*	0.026*	-												
LBN	0.084*	0.164*	0.027*	0.161*	0.153*	0.173*	0.180*	0.125*	-											
LEZ	0.202*	0.339*	0.158*	0.316*	0.334*	0.294*	0.290*	0.248*	0.178*	-										
NLD	0.256*	0.007	0.215*	0.020*	0.053*	0.003	0.032*	0.058*	0.218*	0.371*	-									
OS-D	0.382*	0.541*	0.279*	0.491*	0.515*	0.480*	0.525*	0.453*	0.374*	0.678*	0.583*	-								
ROM	0.125*	0.038*	0.066*	0.040*	0.041*	0.047*	0.061*	0.016*	0.065*	0.192*	0.087*	0.397*	-							
SAL	0.025*	0.264*	0.036*	0.253*	0.270*	0.258*	0.254*	0.197*	0.099*	0.228*	0.307*	0.448*	0.150*	-						
SAS	0.049*	0.321*	0.154*	0.313*	0.339*	0.313*	0.320*	0.313*	0.167*	0.354*	0.317*	0.581*	0.226*	0.166*	-					
SLK	0.186*	0.013	0.114*	0.014*	0.022*	0.022*	0.031*	-0.001	0.133*	0.292*	0.056*	0.469*	0.027*	0.188*	0.302*	-				
CHE	0.182*	0.014	0.123*	0.005	0.010	0.018*	0.016*	0.001	0.132*	0.267*	0.051*	0.440*	0.025*	0.185*	0.284*	-0.003	-			
TUR	0.034*	0.143*	0.008*	0.148*	0.145*	0.157*	0.154*	0.100*	0.020*	0.118*	0.194*	0.341*	0.047*	0.033*	0.153*	0.099*	0.100*	-		
VAN	0.024*	0.291*	0.036*	0.280*	0.277*	0.291*	0.282*	0.226*	0.118*	0.263*	0.344*	0.430*	0.151*	0.015	0.179*	0.218*	0.207*	0.049*	-	
WA	0.024*	0.152*	0.016	0.155*	0.155*	0.157*	0.147*	0.105*	0.057*	0.123*	0.191*	0.398*	0.052*	0.028*	0.144*	0.096*	0.100*	0.002	0.039*	-

To assess the relationship between the haplotypes, we have conducted a median-joining network analysis within the haplogroup R1b1a2-M269 for the populations of Lebanon, the western part of the Armenian Highland, Italy, and Ireland, roughly approximating the path of human Neolithic migrations (Fig. 3.6). The haplotypes of western Armenian origin are widely scattered and mainly associated with haplotypes from the Near Eastern (Lebanese) population. In addition, there are four haplotypes shared between Armenians and Europeans (Ireland and Italy), which was not revealed in Herrera et al. 2012.

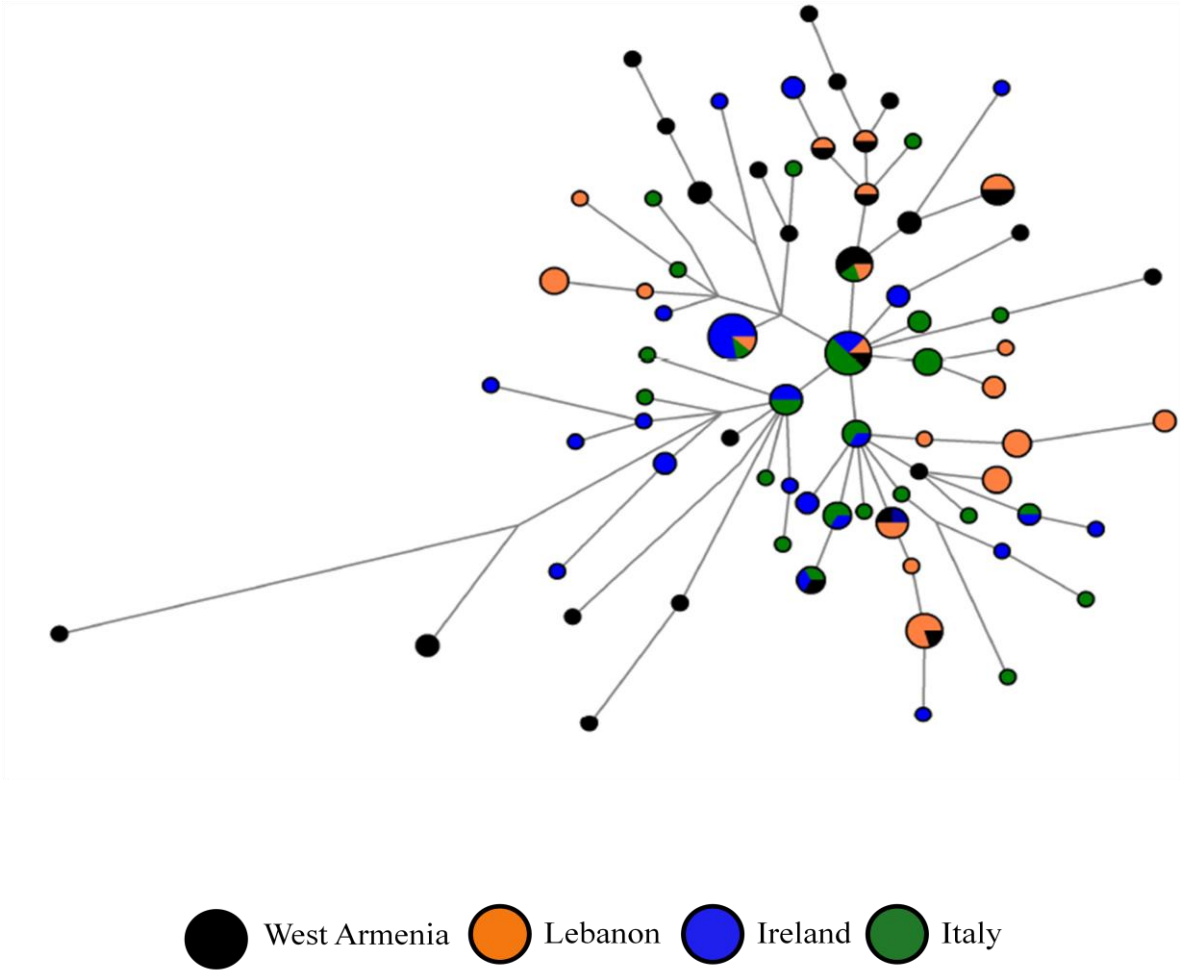


Figure 3.6. Median-joining network of microsatellite haplotypes within the haplogroup R1b1a2.
Circles represent microsatellite haplotypes, the areas of the circles are proportional to haplotype frequency (smallest circle corresponds to one individual), and population is indicated by color.

Similarly, the constructed network for the haplogroup R1b-M343 pointed to the wide distribution pattern of the haplotypes from all four considered Armenian populations and their genetic association with Near Eastern ones [Herrera KJ, et al., 2012].

3.3. Distribution patterns of the haplogroup J2-M172

Though both haplogroups J1 and J2 are encountered in the Near East, lineage J1-M267 is more typical for East African and Arabian populations, having a decreasing frequency northwards. Alternatively, the majority of European J-M304 lineages belong to its J-M172 subclade that radiated from the Near East. The spatial distribution of the haplogroup indicates the highest encountered frequencies (>15%) in the areas between the Near East and the Mediterranean littoral [Semino O, et al., 2004; Di Giacomo F, et al., 2004]. Similarly, this lineage is also one of the most common haplogroups in the Caucasus [Balanovsky O, et al., 2011; Nasidze I, et al., 2004]. In particular, the lineage comprises 59% of the Y chromosomes in the Chechen population and occurs with the lowest STR variance (14%), likely representing a strong founder effect signal [Balanovsky O, et al., 2011]. Moreover, the distribution pattern of the haplogroup is consistent with a Levantine/Anatolian dispersal route to southeastern Europe and the Caucasus [Semino O, et al., 2004]. By this definition, the notion of 'Anatolia,' taken from Cinnioğlu et al. [Cinnioğlu C, et al., 2004], actually includes the western and central areas of the Armenian Highland.

The frequency analysis of the haplogroup J2-M172 data within the Armenian populations shows that it is the most commonly encountered clade in the western and central areas of historical Armenia (27.7% and 25.5%, respectively) (Fig. 3.7). Further, the western and eastern parts of the Armenian Highland have relatively high values of genetic variances, while the highest level among all populations was detected in Syria, in accordance with the suggested Near Eastern origin of this haplogroup (Table 3.5) [Semino O, et al., 2004].

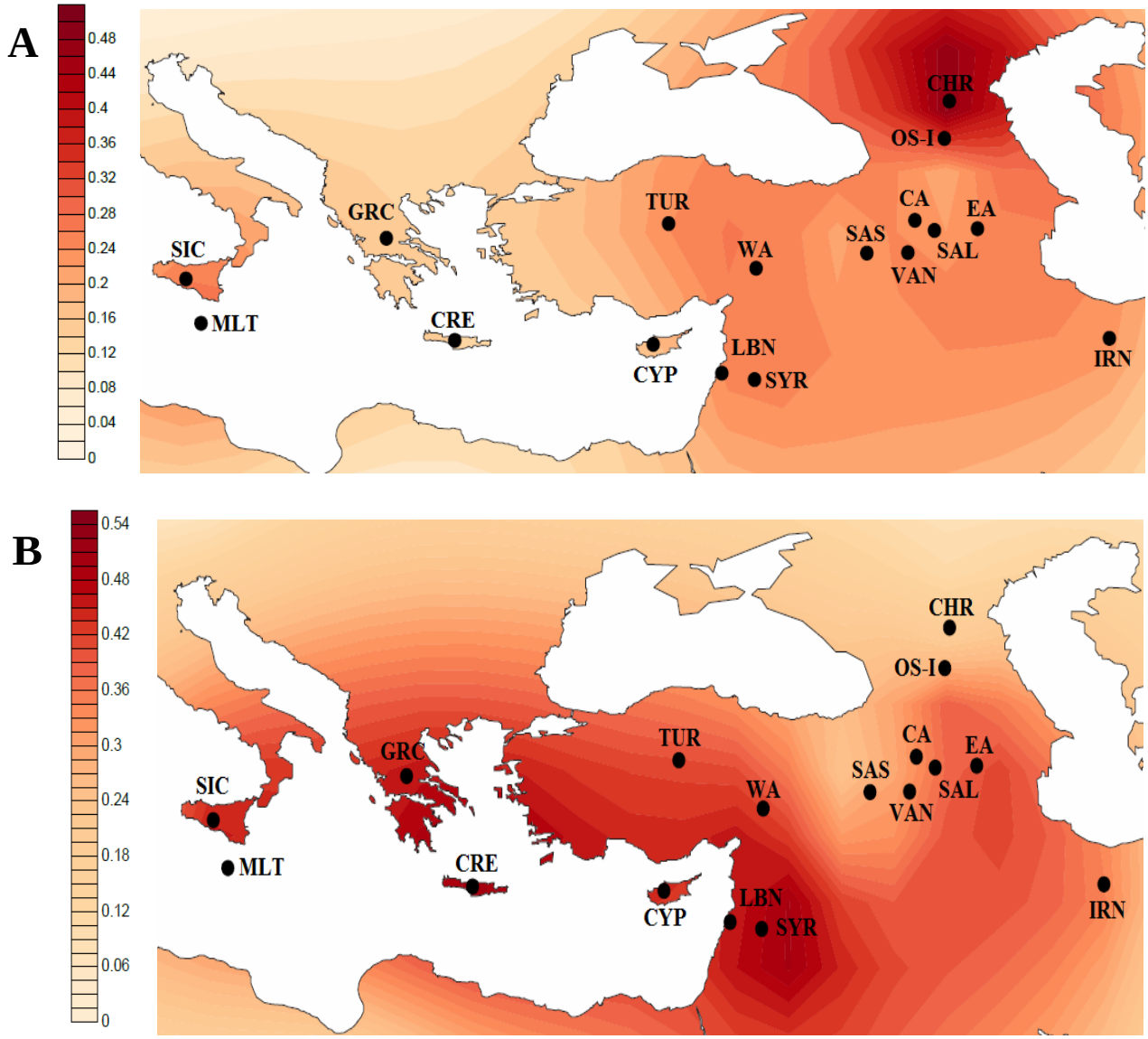


Figure 3.7. Geographical distribution maps of haplogroup J2 frequencies (A) and genetic variances (B).

The heatmap plot of the F_{ST} values (Fig. 3.8) within the haplogroup separates a distinct cluster of western Asian populations (Armenians, Turks, Lebanese, and Iranians). It also demonstrates a moderate level of genetic similarity between the majority of Armenian geographic groups (except Sasun) and the European populations. Our findings also indicate that western Armenians rather than eastern Armenians have a slightly closer genetic affinity with Greeks and Cretans based on the absolute values of pairwise F_{ST} distances. This result contradicts Herrera et al. 2012, who demonstrated a segregation of Armenians from the European populations mentioned.

Table 3.5.
Frequency and varience distributions of the haplogroup J2

Population	N	Freq	Vp
Armenian Highland			
WA	148	0.2770	0.4534
CA	200	0.2550	0.3563
EA	416	0.2452	0.4317
Van	103	0.2913	0.3819
Sasun	104	0.1731	0.1614
Salmast	199	0.2111	0.3503
Caucasus			
Ossets-Iron	230	0.1826	0.3895
Chechen(Dagestan)	100	0.5900	0.1357
Near East			
Syria	200	0.2600	0.5644
Lebanon	916	0.2587	0.4210
Iran			
Iran	324	0.2377	0.3615
Anatolia			
Turkey	523	0.2430	0.4085
Balkans and Southern Europe			
Malta	187	0.3155	0.4131
Sicily	236	0.2585	0.4821
Crete	193	0.3057	0.5218
Cyprus	165	0.1515	0.4363
Greece	171	0.1462	0.4629

In addition, eastern Armenians rather than western Armenians display closer genetic proximity to Ossets (relying on the F_{ST} values). On the whole, the comparison of F_{ST} genetic distances for the haplogroup J2 indicates that the western Armenian population occupies an intermediate position between the Near East and Balkans on one side, and Southern Europe on the other, while eastern Armenia serves as a genetic bridge between the Levant and the North Caucasus (Table 3.6).

According to the genetic survey of Herrera et al. (2012), the expansion time esimates for the haplogroup J2-M172 in the Armenian populations generally corresponds to those for the R1b-M343. On the MDS plot based on R_{ST} genetic distances for the haplogroup J2 the Armenian ethnic groups of Gardman, Sasun and Van are appeared to be among the populations of Turkey, Iran and the Near East,

while Greeks and Cretans located further from the Armenians. The highest level of haplotype varience within the haplogroup J2-M172 among considered Armenian populations was observed in the population of Gardman (based on 8 STR markers).

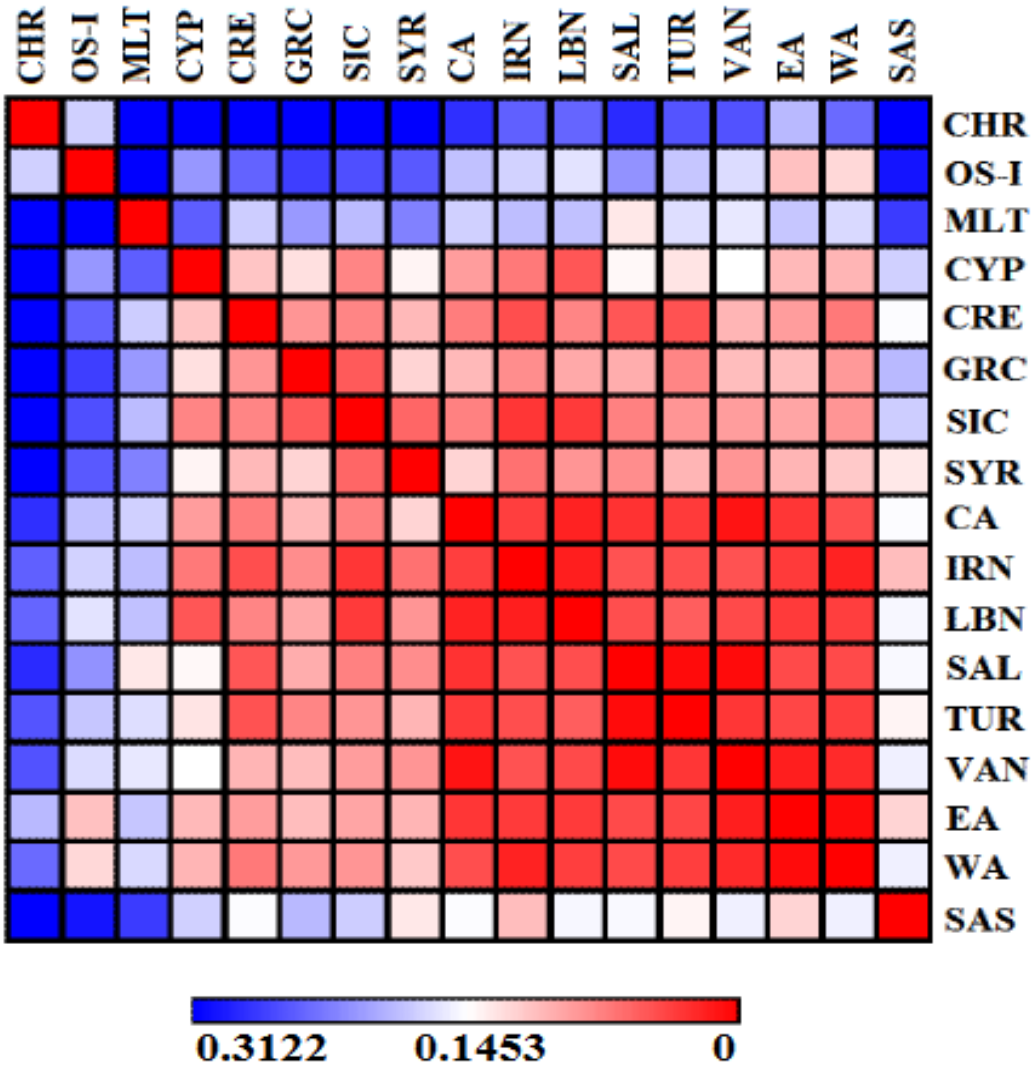


Figure 3.8. Heatmap of pairwise F_{ST} genetic distances, ranging from low (red) to high (blue), calculated for the haplogroup J2.

Median-joining network analysis within the haplogroup J2 for the populations of Syria, western and eastern parts of the Armenian Highland, Crete, and Chechens reflects the bifurcation of the haplogroup J2 to two distinct directions from the Near East: westward and northward, mainly connecting western Armenia to Europe and eastern Armenia to the North Caucasus (Fig. 3.9).

Table 3.6.

Pairwise F_{ST} genetic distances between the populations studied based on the haplogroup J2 (P<0.05).

Pops.	POPULATION																
	CA	**CHR**	**CRE**	**CYR**	**EA**	**GRC**	**IRN**	**LBN**	**MLT**	**OS-I**	**SAL**	**SAS**	**SIC**	**SYR**	**TUR**	**VAN**	**WA**
CA	-																
CHR	0.211*	-															
CRE	0.041*	0.252*	-														
CYR	0.052*	0.289*	0.068*	-													
EA	0.014*	0.130*	0.052*	0.063*	-												
GRC	0.062*	0.296*	0.050*	0.077*	0.064*	-											
IRN	0.018*	0.182*	0.023*	0.040*	0.016*	0.047*	-										
LBN	0.007	0.179*	0.045*	0.027*	0.016*	0.058*	0.006	-									
MLT	0.117*	0.312*	0.117*	0.184*	0.122*	0.148*	0.127*	0.126*	-								
OS-I	0.125*	0.117*	0.180*	0.150*	0.065*	0.202*	0.116*	0.104*	0.248*	-							
SAL	0.013	0.214*	0.026*	0.086*	0.022*	0.059*	0.025*	0.024*	0.079*	0.153*	-						
SAS	0.090*	0.274*	0.090*	0.117*	0.072*	0.130*	0.064*	0.093*	0.203*	0.227*	0.091*	-					
SIC	0.042*	0.279*	0.043*	0.043*	0.055*	0.028*	0.015*	0.017*	0.128*	0.192*	0.042*	0.118*	-				
SYR	0.073*	0.250*	0.063*	0.084*	0.061*	0.073*	0.036*	0.050*	0.163*	0.186*	0.046*	0.081*	0.033*	-			
TUR	0.016*	0.189*	0.025*	0.078*	0.021*	0.043*	0.024*	0.029*	0.108*	0.122*	-0.002	0.085*	0.050*	0.061*	-		
VAN	0.002	0.190*	0.061*	0.089*	0.006	0.064*	0.025*	0.021*	0.102*	0.109*	-0.001	0.098*	0.053*	0.049*	0.014	-	
WA	0.023*	0.176*	0.039*	0.061*	-0.002	0.052*	0.008	0.018*	0.111*	0.075*	0.022*	0.098*	0.050*	0.069*	0.017*	0.010	-

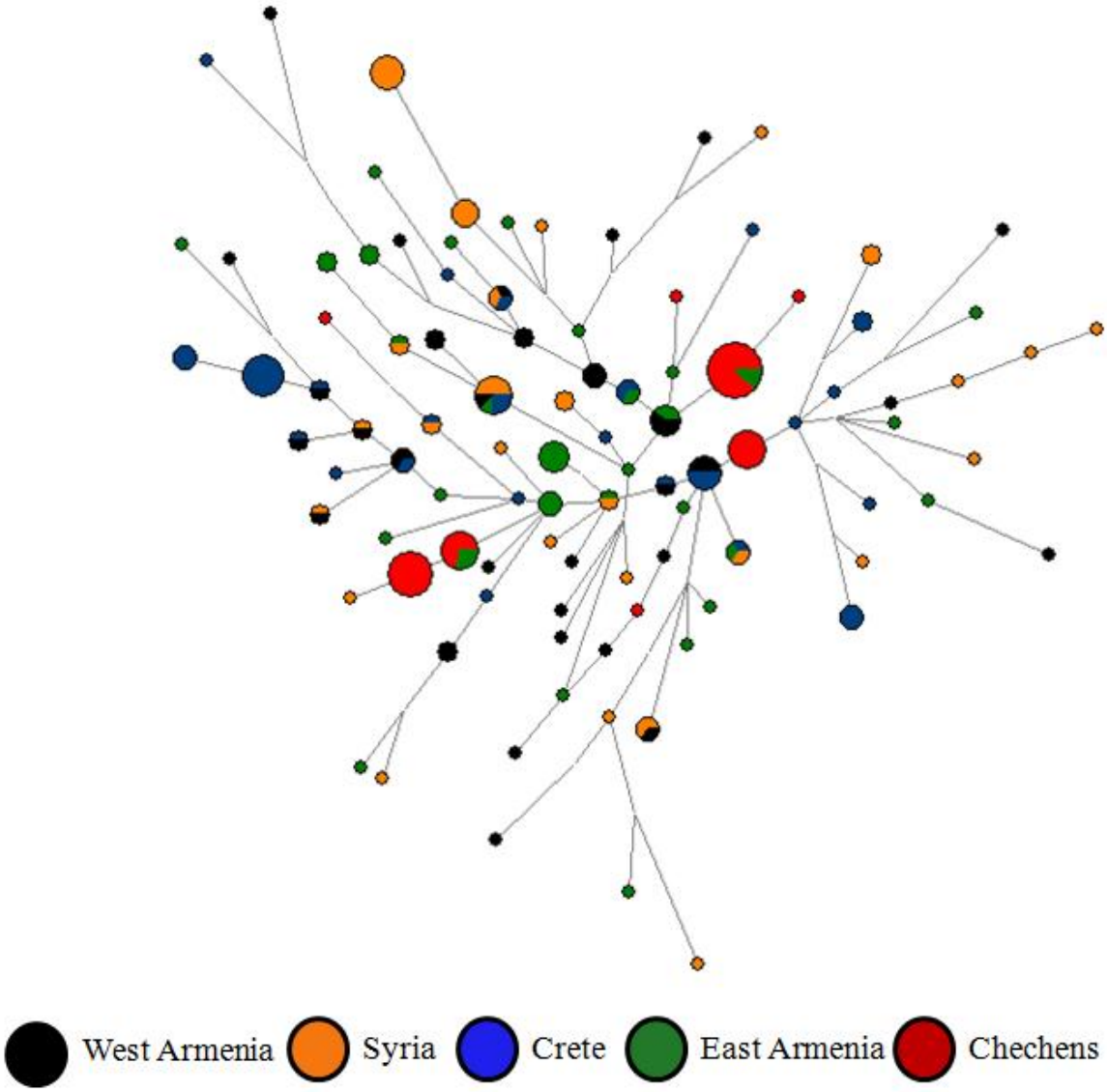


Figure 3.9. Median-joining network of microsatellite haplotypes within the haplogroup J2.
Circles represent microsatellite haplotypes, the areas of the circles are proportional to haplotype frequency (smallest circle corresponds to one individual), and population is indicated by color.

Previously, it was shown that in the constructed J2-M172 network samples from the populations of Sasun and Ararat Valley encountered only with a few haplotypes, while the haplotypes from the populations of Van and Sasun are widely scattered and mainly associated with those from the Near Eastern region [Herrera KJ, et al., 2012].

3.4. Distribution patterns of the haplogroup G-M201

The Y-chromosomal haplogroup G-M201 is widely distributed in the populations of the Caucasus, the Near East, and Southern Europe, with the highest frequencies occurring in the North Caucasus (Fig. 3.10) [Balanovsky O, et al., 2011; Yunusbayev B, et al., 2012]. However, it has a decreasing frequency gradient towards the Balkans and northern Europe.

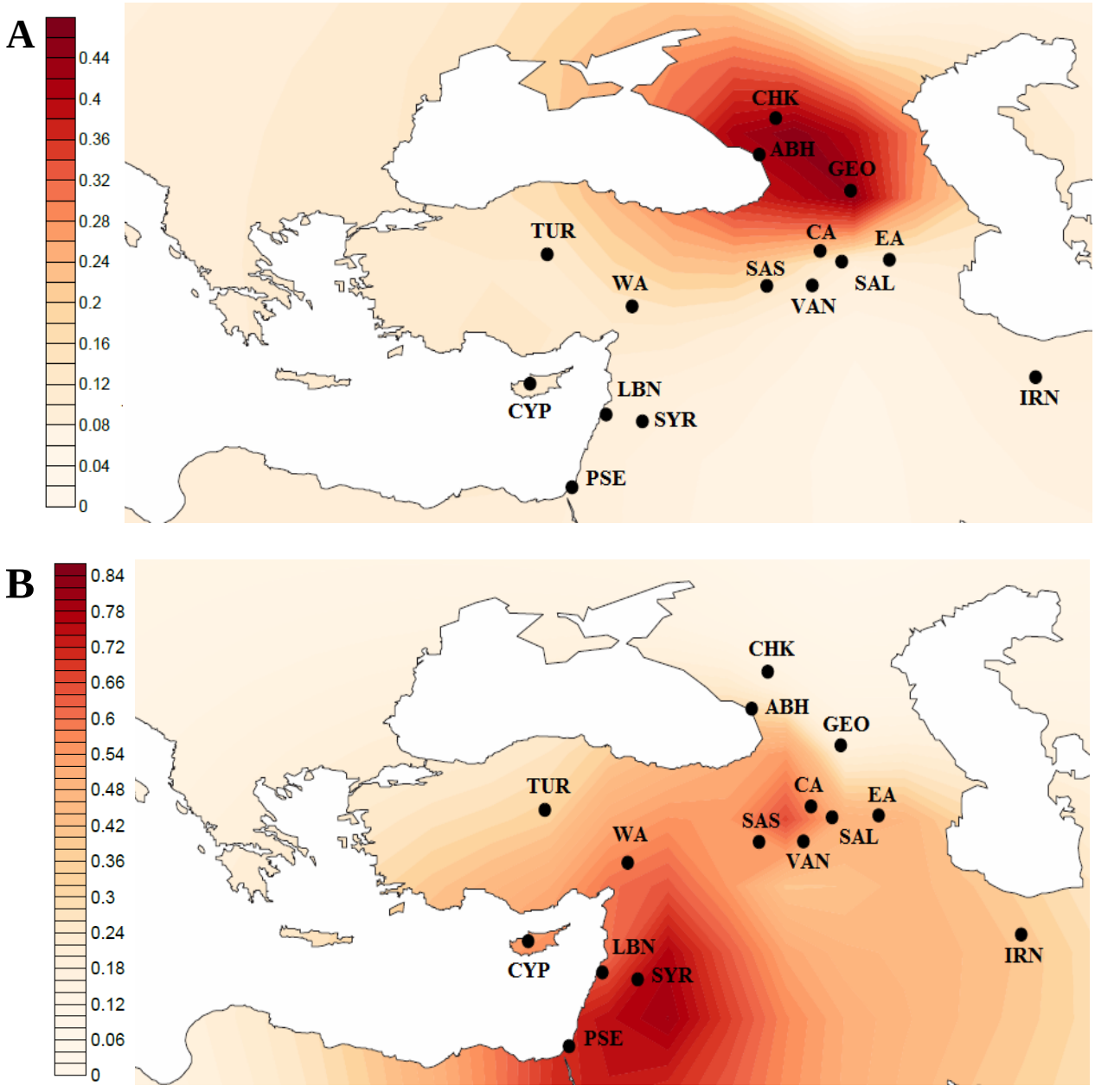


Figure 3.10. Geographical distribution maps of haplogroup G frequencies (A) and genetic variances (B).

Our observations indicate that in the central part of the Armenian Highland, the haplogroup occurs with a relatively high frequency (16%), being inferior by this rate only to the populations of the North Caucasus. At the same time, the Armenian sample from the central region of the Armenian Highland has a comparable value of haplotype diversity (74.5%) with that of the Near Eastern populations of Syria (88.6%) and Palestine (79.3%) (Table 3.7). Thus, our results support the recently published data on the origin of this haplogroup in the neighboring areas of eastern Anatolia, Armenia, and Western Iran [Rootsi S, et al., 2012]. In the previous genetic survey of the haplogroup distribution patterns it was shown that the majority of the G-M201samples within considered four Armenian populations belongs to its G2-P287 branch [Herrera KJ, et al., 2012].

Table 3.7.
Frequency and varience distributions of the haplogroup G.

Population	N	Freq	Vp
Armenian Highland			
WA	148	0.1351	0.6352
CA	200	0.1600	0.7453
EA	416	0.0865	0.5375
Van	103	0.0777	0.3661
Sasun	104	0.1250	0.4744
Salmast	199	0.0754	0.5155
Caucasus			
Georgia	66	0.5000	0.2476
Abkhazians	162	0.4753	0.4038
Cherkessians	126	0.4524	0.2365
Near East			
Palestine	290	0.0862	0.7929
Lebanon	916	0.0655	0.5462
Syria	200	0.0550	0.8864
Iran			
Iran	324	0.0679	0.4372
Anatolia			
Turkey	617	0.1151	0.3896
Southern Europe			
Cyprus	165	0.1333	0.5679

The heatmap plot of F_{ST} values for the haplogroup G (Fig.3.11) does not identify distinct clusters of western Asian or European populations. Though the comparison of F_{ST} values does not conclusively indicate the intermediate position of the central part of the Armenian Highland for the Neolithic migration from the Near East to the North Caucasus, it does not reject this possibility either (Table 3.8).

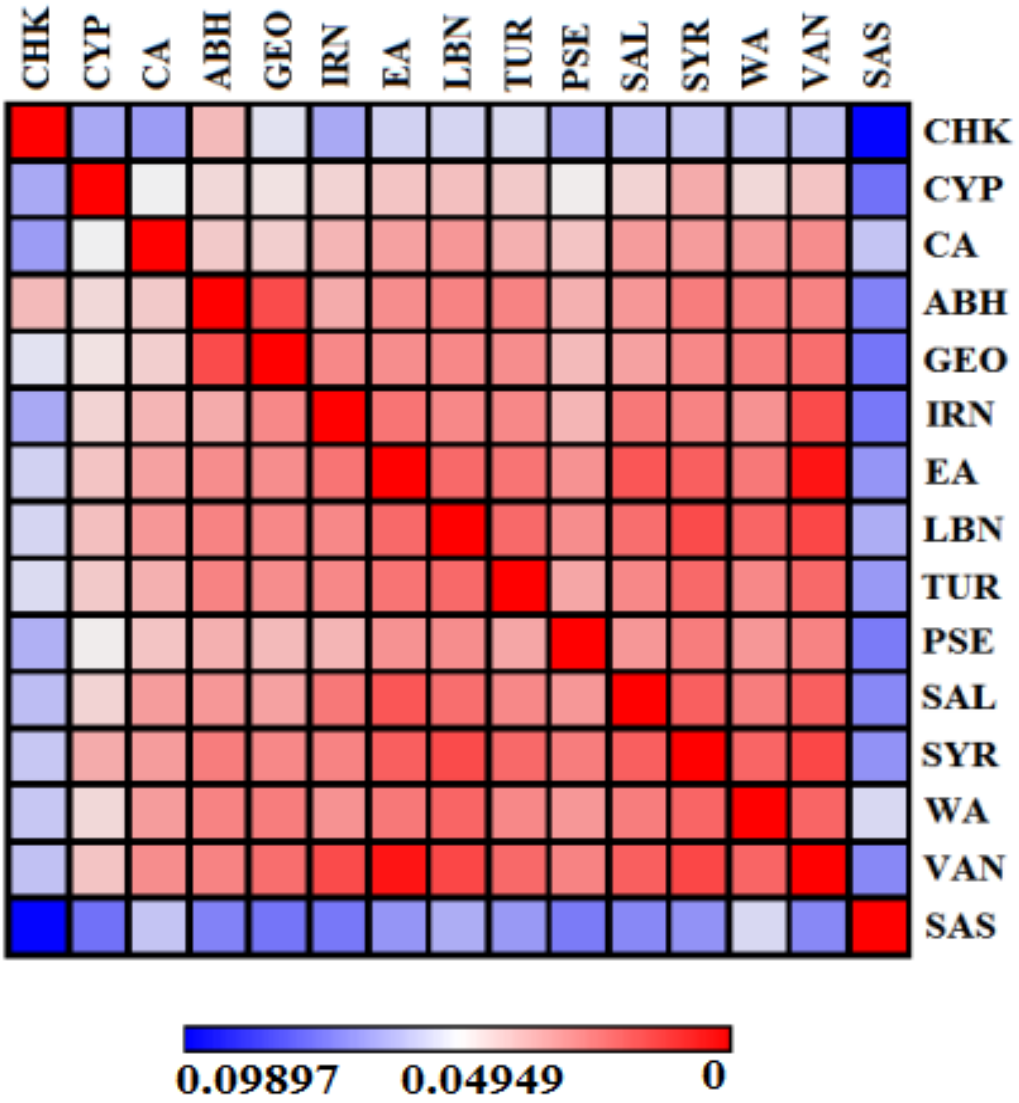


Figure 3.11. Heatmap of pairwise F_{ST} genetic distances, ranging from low (red) to high (blue), calculated for the haplogroup G.

Table 3.8.

Pairwise F_{ST} genetic distances between the populations studied based on the haplogroup G.

Pops.	Population														
	ABH	**CA**	**CHK**	**CYP**	**EA**	**GEO**	**IRN**	**LBN**	**PSE**	**SAL**	**SAS**	**SYR**	**TUR**	**VAN**	**WA**
ABH	-														
CA	0.024*	-													
CHK	0.022*	0.056*	-												
CYP	0.028*	0.033*	0.052*	-											
EA	0.013*	0.017*	0.040*	0.024*	-										
GEO	0.001	0.026*	0.036*	0.029*	0.013*	-									
IRN	0.018*	0.021*	0.052*	0.026*	0.008	0.012	-								
LBN	0.011*	0.015*	0.039*	0.023*	0.007*	0.012*	0.012*	-							
PSE	0.020*	0.024*	0.051*	0.032*	0.014*	0.022*	0.021*	0.013*	-						
SAL	0.015*	0.016*	0.047*	0.027*	0.003	0.017*	0.010	0.008*	0.015*	-					
SAS	0.063*	0.045*	0.099*	0.068*	0.058*	0.068*	0.066*	0.051*	0.065*	0.062*	-				
SYR	0.010	0.016	0.043*	0.019	0.004	0.012	0.011	0.001	0.011	0.005	0.059*	-			
TUR	0.011*	0.019*	0.038*	0.024*	0.009*	0.013*	0.012*	0.007*	0.018*	0.013*	0.057*	0.007	-		
VAN	0.011	0.013	0.045	0.024	-0.010	0.007	0.000	0.000	0.011	0.005	0.061*	0.000	0.006	-	
WA	0.012*	0.016*	0.043*	0.027*	0.009*	0.010	0.014*	0.005	0.015*	0.010	0.039*	0.005	0.012*	0.006	-

The constructed median-joining network within the haplogroup G (Fig. 3.12) reveals the highest level of scattering for the central Armenian haplotypes among various neighboring populations (Palestinians, Cherkessians, Iranians), which is expected under the assumption of local origin of this lineage. Furthermore, the network clearly shows the presence of a founder effect among the North Caucasian Cherkessian population who shares their ancestral haplogroup with Armenians.

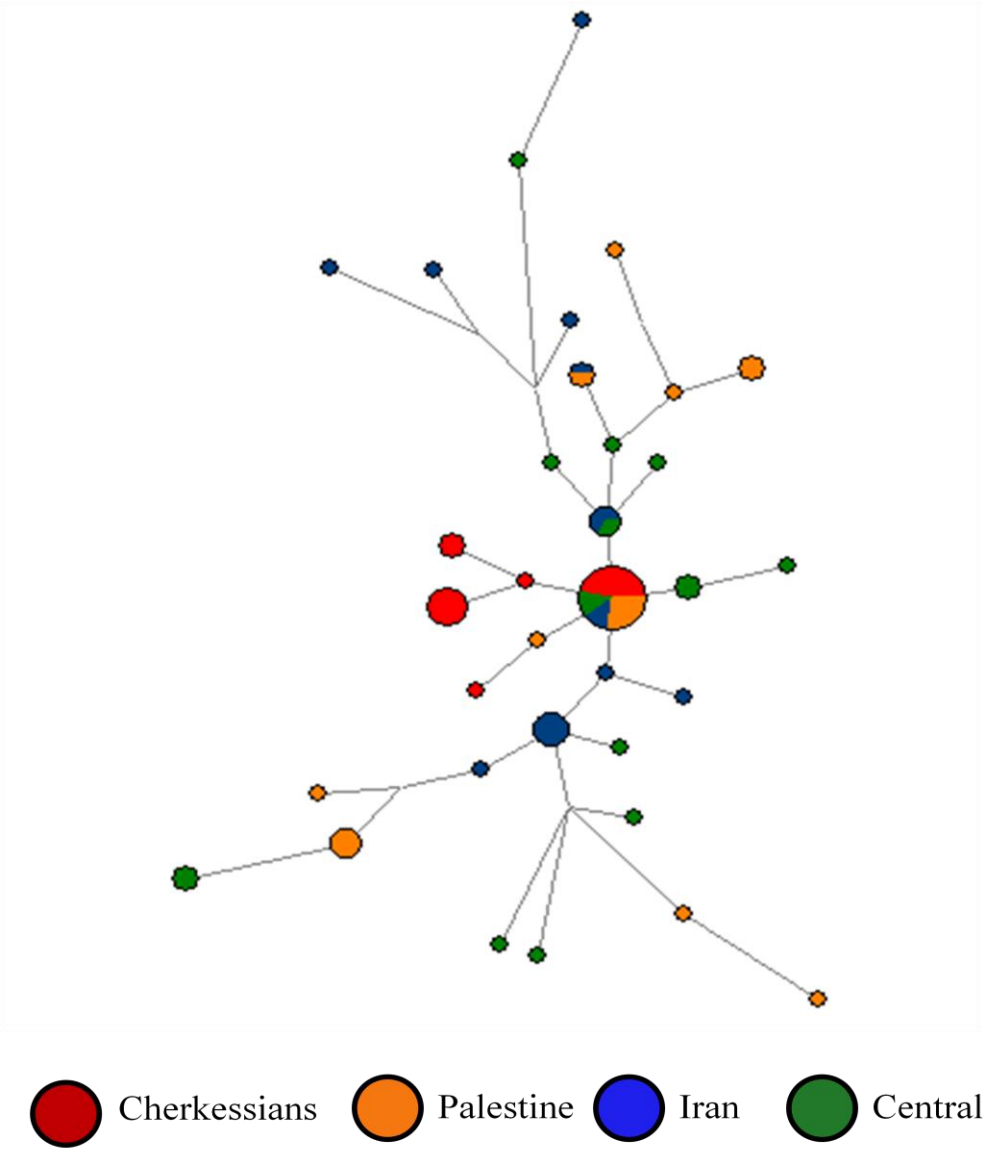


Figure 3.12. Median-joining network of microsatellite haplotypes within the haplogroup G.

Circles represent microsatellite haplotypes, the areas of the circles are proportional to haplotype frequency (smallest circle corresponds to one individual), and population is indicated by color.

3.5. Ancestral components of population`s gene pool based on genome-wide autosomal data

Genome-wide autosomal data of Armenians and reference populations were analyzed using a structure-like [Pritchard JK, et al., 2000] clustering approach by inferring the most likely number of genetic clusters and mixing proportions. The assessed genetic structure of populations from ADMIXTURE with K=6 identified that Armenians have three ancestry components differently contributed to their genetic makeup, namely the Near Eastern (shown in light orange, Fig. 3.13), Caucasus (orange), and European (blue) ones. In agreement with previous genetic studies [Behar DM, et al., 2010; Yunusbayev B, et al., 2012], it was revealed that Armenians are the only population in the Caucasus that have the widest presence of the major ancestry component of the Near Eastern populations, which is almost absent in the populations of Chechens and Lezgins. Similarly to the Near Eastern component, the Caucasus ancestry is also ubiquitously present among all considered Armenian populations, while the European one is encountered only with a trace frequency rates.

In the previous genetic study it was indicated that Armenians share North and East African ancestry component, the appearance of which was explained as the introduction of a genetic component novel to the region [Yunusbayev B, et al., 2012]. However, the results of our ADMIXTURE analysis did not support this conclusion. Also we detected the absence of genetic ancestry component from the Central Asia in all our Armenian samples.

.

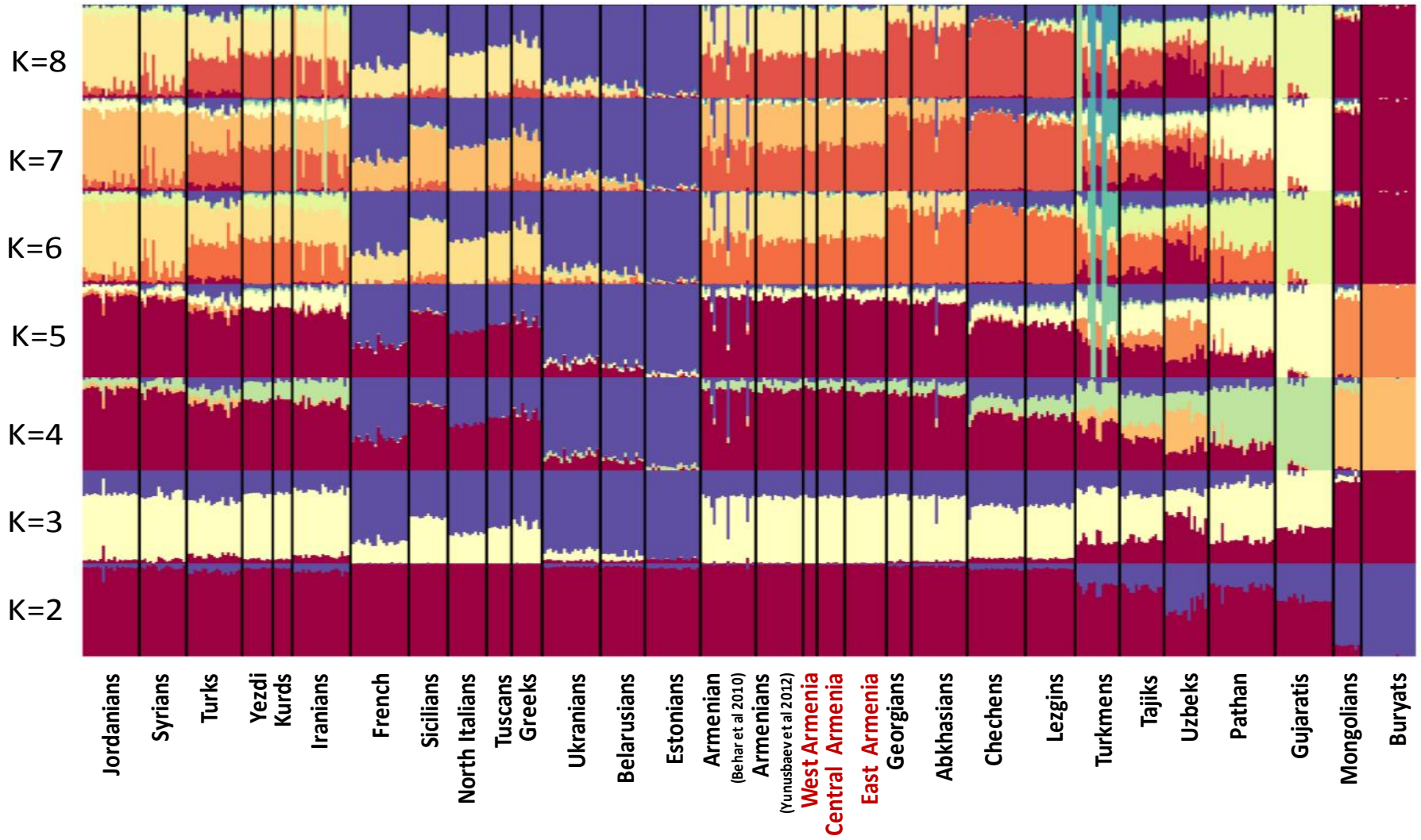


Figure 3.13. Population structure inferred using ADMIXTURE analysis. Each individual is represented by a vertical stacked column indicating the proportions of ancestry in K constructed ancestral populations. Armenian populations are shown in red.

3.6. Armenians' relationship to modern and ancient world populations

In agreement with the results based on Y-chromosomal analysis, the autosomal PCA plot placed present-day Armenians in the intermediate position between modern Near Eastern populations and oppositely directed European and Caucasian clusters (Fig. 3.14). Futhermore, the same pattern was detected for the ancient Armenian samples, thus indicating the genetic continuity of the populations of East Armenia since the Bronze Age. Moreover, on the genetic landscape, all Armenian populations appeared to be in one cluster, which supports the results of the ADMIXTURE analysis.

In general, the populations of Kurds, Turks, Azeris, and Iranians are located in close proximity to each other, flanked by Armenians. This pattern is a consequence of absence in our dataset populations from the South and Central Asia, which appeared to be closer than Armenians to the abovementioned populations [Yunusbayev B, et al., 2012]. In general, the pattern is in accordance with the results of previously conducted PCAs on the Armenian data [Behar DM, et al., 2010; Yunusbayev B, et al., 2012] with the only difference that in the former paper the analysed dataset of Armenians is not structurally homogeneous.

Interestingly, Yamnaya samples from Samara appeared to be very close to ancient Armenian Bronze age sample that confirms previous results of common ancestry for Yamnaya and Armenian samples [Haak W, et al., 2015; Allentoft ME, et al., 2015]. It was revealed that the only admixture pattern in Yamnaya population is detectable when using the Bronze Age Armenians and Upper Paleolithic Mal'ta as sourse populations [Allentoft ME, et al., 2015]. Within the plot, the Yamnaya samples are located in an intermediate position between the Caucasus (CHG) and Eastern (EHG) hunter-gathereres. It was estimated that the split between CHG and early farmers (EF) was approximately 20–30 KYA [Jones ER, et al., 2015]. Moreover, it was suggested that the Caucasus hunter-gatherers belong to a disticnst ancient clade that split from western hunter-gathereres (WHG) about 45 KYA. Thus,

WHG and CHG are the descendants of two different ancient populations that separately survived the Last Glacial Maximum.

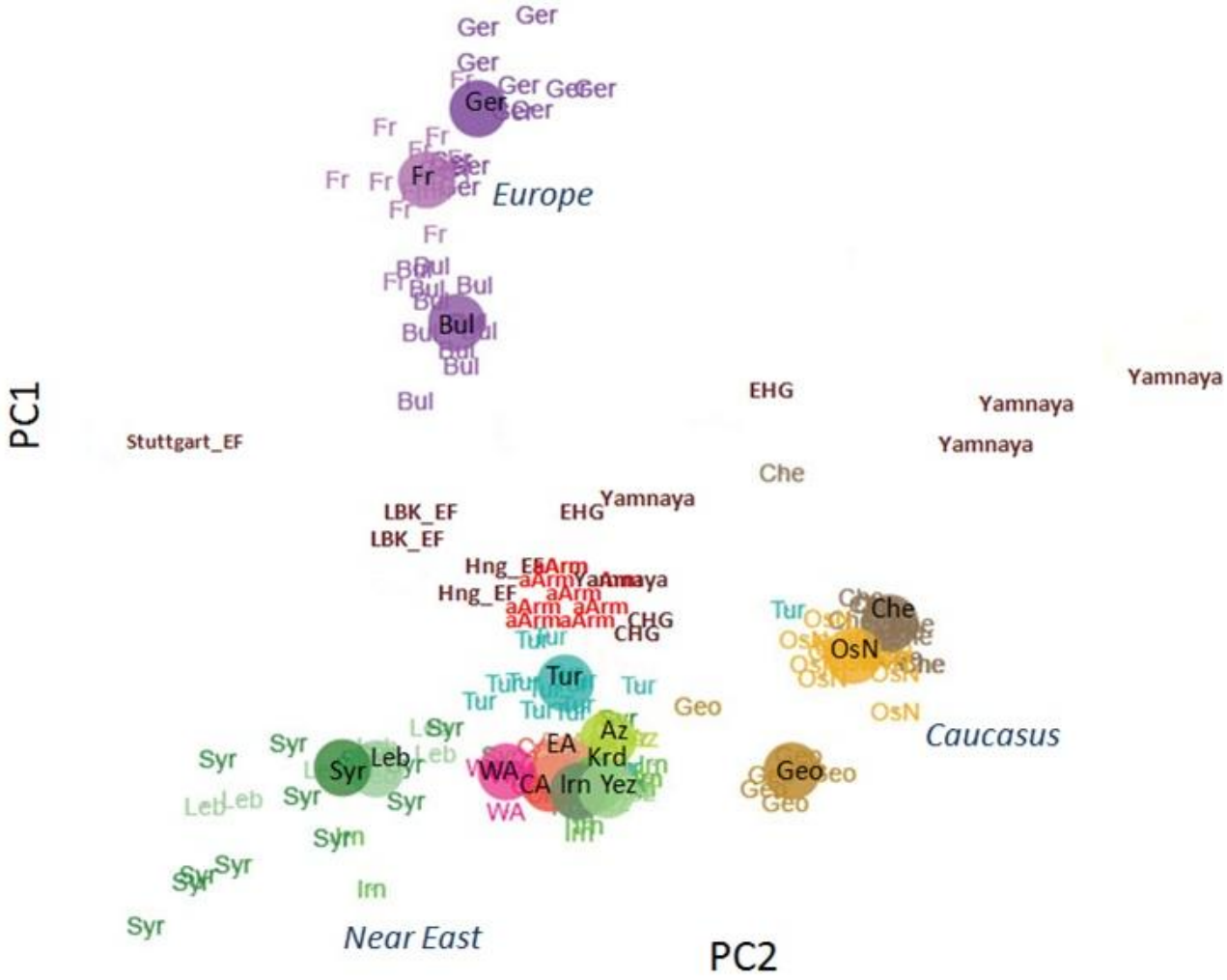


Figure 3.14. Plot of the first and second components of the PCA of the Armenian and comparative populations.
Analysis based on genome-wide autosomal data, with the median representation for all populations.

It was suggested that CHG left their imprint on the genomes of Yamnaya people who migrated to Europe around 3,000 BC. While considering the modern populations, CHG markedly contributed to the genetic makeup of populations from the Caucasus and Central and South Asia. Thus, through the Yamnaya steppe herders, the CHG ancestral are the divergent fourth strand of Europeans.

Moreover, ancient Armenian samples from eastern part of the Armenian Highland appeared to be close to European early farmers and CHG samples, which confirm our suggestion on the role of the region in Neolithic human migration towards Europe and the North Caucasus.

CONCLUSION

The transition from hunting and gathering to agricultural communities was one of the most important landmarks in prehistory of mankind, which profoundly altered human societies around the world [Ammerman AJ and Cavalli-Sforza LL, 1971; Diamond J and Bellwood P, 2003; Bellwood P and Oxenham M, 2008]. From the Fertile Crescent, the region where agriculture is thought to have originated some 10 KYA, this technology spread westward into Europe and northward into the Caucasus, during what has become known as the Neolithic Revolution [Ammerman AJ and Cavalli-Sforza LL, 1984; Riehl S, et al., 2013]. However, the nature and chronology of this transition is a matter of continuing scientific debate in archaeology, anthropology, and human population genetics. Specifically, the question of the routes for agricultural migration, including different directions and waves, remains controversial and requires further comprehensive study to be answered. The only widely accepted idea is that the actual scenario was undoubtedly multifacet and complex.

In this work, we consider the role of the Armenian Highland, a region at the crossroad of Europe and the Middle East, in the spread of Neolithic farmers from the Fertile Crescent. Within this framework, we have assessed the patrilineal genetic makeup of Armenians and other populations of the region roughly covering the area of possible directions of Neolithic farmers` migrations from the Fertile Crescent. To fulfill one of our main objectives, we have identified the Y-chromosomal lineages with the highest contribution to the differentiation of populations from the Near East, the North Caucasus, Anatolia and Europe. While comparing the distribution patterns of these specific paternally inherited haplogroups, which are the genetic markers of Neolithic spread, in the populations studied, we managed to trace back the possible migration routes of the first agriculturalists. As biparentally transmitted genetic markers enable to get more detailed insight into a human population history, we characterized genetic variation of modern and ancient populations from the

considered region by analyzing genome-wide high-density genotype data. Based on the results of analysis of both uniparentally (paternal) and biparentally (autosomal) inhereted traits we have reconstructed the plausible directions and waves of Neolithic movements that had taken place via the Armenian Highland towards Europe and the North Caucasus.

To address these questions, we generated the Y-chromosomal and genome-wide data from self-identified ethnic Armenians, roughly representing the whole area of the Armenian Highland. Our choice of Armenians as an indigenous ethnic group of the region is justified by the results of recent genetic surveys [Hellenthal G, et al., 2014; Haber M, et al., 2015], where the absence of any significant admixture events for Armenians over at least the past 4 KYA was detected. Having included in our datasets the DNA samples acquired from human fossils we intended to check the existence of genetic continuity between the ancient and modern populations of the region. Specifically, we used ancient Armenian samples, dated to Bronze Age and excavated in the eastern part of the Highland.

On the whole, Y-chromosomal haplogroup composition and calculated genetic distance values based on the frequencies of all haplogroups pointed to the intermediate position of the populations from the Armenian Highland between the Near East and Europe, and the Near East and the North Caucasus. Moreover, a significant prevalence of the Y-chromosomal haplogroups, associated with the advance of Neolithic farmers from the Near East, was revealed in all Armenian populations. In particular, the haplogroups R-M269, J2, and G, which are considered as putative genetic markers of the Neolithic migration, represent the most common lineages in the all Armenian populations, together accounting for 50–70% of the sampled groups. This pattern, coupled with dozens of Neolithic archaeological artifacts, indicates the continuous habitation of the Armenian Highland since the dawn of the Neolithic [Dolukhanov P, et al., 2004; Herrera KJ, et al., 2012]. Visualization of genetic distances between the observed ethnic groups revealed that the populations of the Armenian Highland, the Near East, and Eastern Europe appear

to be in one extensive cluster with a clear geographic gradient from the Near East towards the northwest. Besides, the populations of Western Europe and the North Caucasus are separated from each other and bidirectionally from the Armenian populations. The same pattern of populations distribution was obtain while using another multivariate approach (correspondence analysis), which allowed concluding that the haplogroup composition of the Near Eastern populations is very similar to that found for the populations from the Anatolian and Armenian plateaus, as well as those of the Mediterranean islands. This result possibly points to a long-term genetic continuity, persistent since at least the Neolithic. Concerning the routes of agricultural dispersal, we could state that the seafaring played an important role in the colonization of Europe. This conclusion was based on the result of genetic comparisons which indicated the intermediate position of the populations from the Mediterranean islands, namely Crete and Cyprus, between the Near Eastern and European ones.

Considering the question of various haplogroup contributions to the populations' differentiation, it was shown that the European cluster, comprising Basques, Sicilians, and Belgians, is connected to the haplogroups R1b1a2-M269 and I-M170, which are frequently encountered in Europe, and the former being a marker for the Neolithic migration. In this context, the Caucasus cluster, containing Abkhazians, Georgians, and Ossetians, is found to be associated with another genetic marker of Neolithic migration, namely the haplogroup G-M201. Meanwhile the third genetic tracer of agricultural dispersal, the lineage J2-M172, appears in between the European and Caucasus clusters.

More detailed analysis on the distribution patterns of the genetic markers of Neolithic migration revealed that the haplogroup R1b1a2-M269 is the most frequently encountered lineage in all Armenian samples, except Sasun, which differs from others due to the predominance of haplogroup T. On the whole, the spatial occurrence of the main western European Y-chromosomal lineage, R1b1a2-M269, shows a significant positive frequency cline from the Near East towards Europe. In

contrast, a negative cline of genetic variation within the haplogroup is detected from the Levant towards northwest and northeast. Moreover, in comparison with all considered populations from the Near East, Europe, and Anatolia, the haplogroup R1b1a2-M269 occurs with the highest genetic variances in the populations of the Armenian Highland, namely from the western parts of the area, Sasun and Salmast. Worth mentioning that comparison of genetic distances based on the haplogroup variation values places the western part of the Armenian Highland in a transitional position between the Near East and Europe.

The Y-chromosomal haplogroup J2-M172 it widely distributed in the populations of Mediterranean littoral and the Caucasus, while the highest levels of genetic variances are detected in the Near East that is in accordance with the suggested origin of the lineage in this region [Semino O, et al., 2004]. It was also shown that the J2-M172 clade is the most commonly encountered one in the populations of western and central parts of the Armenian Highland, and the relatively high values of its genetic variances are detected in the western and eastern regions of the plateau. In addition, the comparison of genetic distances based on haplogroup diversity estimates pointed to an intermediate position of the western Armenian population between the Near Eastern and Southern European ones, while the eastern Armenian population serves as a genetic link between the Levant and the North Caucasus.

The spatial distribution of the haplogroup G-M201 shows that its relatively high frequency values are encountered in the populations of central part of the Armenian Highland, being inferior by this rate only to the populations of the North Caucasus. The highest estimates of the haplogroup diversity values were found in the Armenian and Near Eastern populations, which support recently published data on the origin of the lineage in the neighboring areas of eastern Anatolia, Armenia, and Western Iran [Rootsi S, et al., 2012]. The comparison of genetic distances based on the microsatellite diversity within the lineage does not reject the possibility of an

intermediate position of the central part of the Armenian Highland for the Neolithic spread from the Levant towards the North Caucasus.

Summarizing, our observation on the Y-chromosomal structure in geographically different Armenian populations suggests that the Armenian Highland served as a transitional corridor for at least two distinct pathways of migration for Neolithic farmers from the Near East westward and northward. The movement to Europe took place predominantly via the western region of the Armenian Highland alongside the coastline of the Mediterranean Sea, which is supported by the spatial distribution pattern of the haplogroup R1b1a2-M269. The migration to the North Caucasus occurred mainly across the central and eastern regions of the Armenian Highland, which is shown by the geographical distribution of haplogroup G-M201. In addition, we identified a distinct Neolithic wave of bidirectional expansion to Europe and the North Caucasus associated with the haplogroup J2-M172. Thus, based only on Y-chromosomal markers, at the initial stage of the Neolithic migration from the Levant, different directions and waves of population movement could be identified in the Armenian Highland (Fig. 3.15).

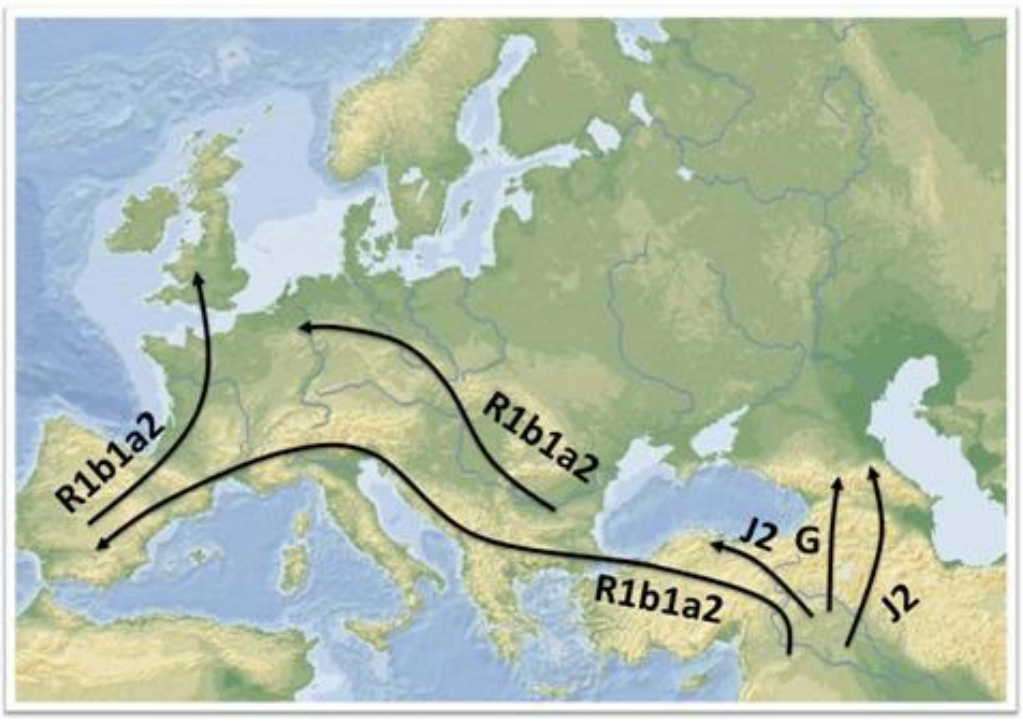


Figure 3.15. Different waves and directions of Neolithic migration from the Fertile Crescent.

Genetic ancestry analysis based on genome-wide autosomal patterns of Armenians and reference populations has identified that all populations of the Armenian Highland have three ancestry components contributed to their genetic makeup, namely the Near Eastern, Caucasus, and European ones. It was shown that Armenians, unlike the other populations of the Caucasus, have the widest presence of the Near Eastern ancestry component, which is in agreement with the previous studies [Behar DM, et al., 2010; Yunusbayev B, et al., 2012]. In all Armenian populations, the Caucasus ancestry component has almost the same rate of contribution as the Near Eastern one, while the European component is encountered in a much less amount. Similarly to the multivariate analysis based on Y-chromosomal data, the results based on autosomal genome-wide markers placed modern Armenians in the intermediate position between the modern Near Eastern populations, on the one side, and oppositely directed European and Caucasian clusters, on another side. On the genetic map, the Bronze Age Armenian samples also occupy a position in between the populations of the Near East and Europe/Caucasus. Interestingly, Yamnaya samples from Samara appeared to be very close to ancient Armenian Bronze age samples, that confirms previous results of common ancestry for Yamnaya and Armenian samples [Haak W, et al., 2015].

Thus, relying on combined results of analysis of paternallly and biparentally transmitted genetic markers, it was detected that at least, two principal Neolithic migratory directions: (1) westward – alongside the coastline of the Mediterranean Sea and (2) northward – to the North Caucasus. Moreover, we assumed that the population movement of the first farmers from the Levant could have taken place both by land through Anatolia and the North Caucasus via the Armenian Highland, and by maritime routes via eastern Mediterranean islands towards continental Europe.

INFERENCES

1. Y-chromosomal haplogroup composition of geographically different Armenian groups indicates their close genetic proximity to the Near Eastern populations. The Europeans and North Caucasians, being genetically distant from the Levant populations, occupy different positions from the Armenian cluster, which in its turn has an intermediate location between the Levant and Europe/Caucasus.
2. Majority of the paternal genetic pool of Armenians is presented by Neolithic Y-chromosomal haplogroups that have the highest contribution to the genetic differentiation of populations from the Middle East, Europe and the North Caucausus.
3. Distribution patterns of Neolithic Y-chromosomal haplogroups indicate an intermediate position of the Armenian populations between the Near East and Europe/Caucasus, pointing to the transitional role of the Armenian Highland for first farmers' migration from the Levant.
4. Armenians` genetic makeup based on genome-wide autosomal SNP data mainly consists of the Near Eastern and Caucasus components, of which the former has relatively higher contribution in comparison with the populations of the North Caucasus. This pattern is in congruence with that based on the Y-chromosomal markers.
5. Transitional role of the Armenian Highland in the migration of the Neolithic farmers is further supported by the genome-wide autosomal structure of Bronze Age Armenians, which occupy an intermediate position between the populations of the Near East and Europe/Caucasus.
6. Armenian Highland served as a transitional corridor for at least two distinct pathways of migrations for the Neolithic farmers from the Near East westward and northward. The movement to Europe took place predominantly via the western region of the Highland, while the direction to the North Caucasus occurred mainly across the central and eastern regions of the Armenian

plateau. A distinct Neolithic wave of bidirectional movement to Europe and the North Caucasus was also identified.

REFERENCES

1. Асланян С, Беляева Е, Колпаков Е, Любин В, Саркисян Г, Суворов А. Работы Армяно-Российской археологической экспедиции в 2003-2006 гг. В кн.: Носов Е (ред.). Записки ИИМК РАН, Санкт-Петербург. 2007; 142-154.
2. Беляева Е, Любин В. Первые данные о ранних этапах ашеля на Кавказе. Краткие сообщения Института археологии. 2012; 227:28-37.
3. Вавилов НИ. Проблемы селекции. Роль Евразии и Нового Света в происхождении культурных растений. АН СССР, Москва-Ленинград; 1960. -520 с.
4. Гаджиев М, Давудов О, Шихсаидов А. История Дагестана с древнейших времён до конца XV века. ДНЦ РАН, Махачкала; 1996. -462 с.
5. Гамкрелидзе Т и Иванов В. Индоевропейский язык и индоевропейцы. Реконструкция и историко-типологический анализ праязыка и протокультуры. Изд-во Тбилисского университета, Тбилиси; 1984. -1328 с.
6. Гулканян В. О некоторых сельскохозяйственных культурах из Аргиштихинили. Историко-филологический журнал. 1966; 4(35):102-116.
7. Ерицян Б. Новая нижнепалеолитическая пещерная стоянка Лусакерт 1 (Армения). КСИА. 1975; 141:41-50.
8. Лисицына Г, Прищепенко Л. Палеоэтноботанические находки Кавказа и БлижнегоВостока. Наука, Москва; 1977. -128 с.
9. Любин В, Беляева Е. Новые данные о раннем палеолите Армении. В кн.: Васильев С, Щелинский В (ред.). Древнейшие обитатели Кавказа и расселение предков человека в Евразии. Востоковедение, Санкт-Петербург. 2010; 107-126.
10. Любин В, Беляева Е. Ранняя преистория Кавказа. Востоковедение, Санкт-Петербург; 2006. -108 с.
11. Любин В. Палеолит Кавказа. В кн.: Палеолит Кавказа и Северной Азии (Палеолит мира). Наука, Ленинград. 1989; 7-142.

12. Овсепян Р. Основные этапы предыстории земледелия на территории Армении по палеоэтноботаническим данным. Международная научная конференция «Археология, этнология, фольклористика Кавказа», Тбилиси-Гори-Батуми, 27-30 сентября 2010 г., Сборник кратких содержаний докладов, Тбилиси, Издательство «Меридиани». 2011; 191-192.
13. Овсепян Р, Овсепян Г. Археоботанические исследования полевых культур в Армении. Бюллетень Национального аграрного университета Армении. 2009; 4:28-31.
14. Паничкина М. Палеолит Армении. Наука, Ленинград; 1950. -107 с.
15. Пиотровский Б. Археология Закавказья. Издательство Ленинградского государственного ордена Ленина университета Им. А. А. Жданова, Ленинград. 1949. -134 с.
16. Сардарян С. Палеолит Армении. АН Армянской СССР, Ереван; 1954. –204 с.
17. Саядян Ю. Новейшая геологическая история Армении. Гитутюн, Ереван; 2009. -357 с.
18. Тамамшян С. “Ископаемая” пшеница из раскопок кургана близ Вагаршапата. В кн.: Калантар А. Раскопки Древнего Вагаршапата. Фонд Мелконян, Ереван. 1935; 87-95.
19. Adler DS, Bar-Yosef O, Belfer-Cohen A, Tushabramishvili N, Boaretto E, Mercier N, Valladas H, Rink WJ. Dating the demise: Neanderthal extinction and the establishment of modern humans in the southern Caucasus. Journal of Human Evolution. 2008; 55(5):817-833.
20. Adler DS, Wilkinson KN, Blockley S, Mark DF, Pinhasi R, Schmidt-Magee BA, Nahapetyan S, Mallol C, Berna F, Glauberman PJ, Raczynski-Henk Y, Wales N, Frahm E, Jöris O, MacLeod A, Smith VC, Cullen VL, Gasparian B. Early Levallois technology and the Lower to Middle Paleolithic transition in the Southern Caucasus. Science. 2014; 345(6204):1609-1613.

21. Akçar N, Yavuz V, Ivy-Ochs S, Kubik PW, Vardar M, Schlüchter C. Paleoglacial Records from Kavron Valley, NE Turkey: Field and Cosmogenic Exposure Dating Evidence. Quaternary International. 2007; 164:170-183.
22. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome research. 2009; 19(9):1655-1664.
23. Allentoft ME, Sikora M, Sjögren KG, Rasmussen S, Rasmussen M, Stenderup J, Damgaard PB, Schroeder H, Ahlström T, Vinner L, Malaspinas As, et al. Population genomics of Bronze Age Eurasia. Nature. 2015; 522(7555):167-172.
24. Alonso S, Flores C, Cabrera V, Alonso A, Martín P, Albarrán C, Izagirre N, de la Rúa C, García O. The place of the Basques in the European Y-chromosome diversity landscape. European journal of human genetics. 2005; 13(12):1293-1302.
25. Ammerman AJ and Cavalli-Sforza LL. Measuring the rate of spread of early farming in Europe. Man. 1971; 1:674-688.
26. Ammerman AJ and Cavalli-Sforza LL. Neolithic Transition and the Genetics of Populations in Europe. Princeton University Press, Princeton; 1984. -194 pp.
27. Armelagos GJ, Barnes KC, Lin J. Disease in human evolution: the re-emergence of infectious disease in the third epidemiological transition. AnthroNotes. 1996; 18(3):1-7.
28. Balanovsky O, Dibirova K, Dybo A, Mudrak O, Frolova S, Pocheshkhova E, Haber M, Platt D, Schurr T, Haak W, Kuznetsova M, Radzhabov M, Balaganskaya O, Romanov A, Zakharova T, Soria Hernanz DF, Zalloua P, Koshel S, Ruhlen M, Renfrew C, Wells RS, Tyler-Smith C, Balanovska E, Genographic Consortium. Parallel evolution of genes and languages in the Caucasus region. Mol Biol Evol. 2011; 28:2905-2920.
29. Balaresque P, Bowden GR, Adams SM, Leung HY, King TE, Rosser ZH, Goodwin J, Moisan JP, Richard C, Millward A, Demaine AG, Barbujani G, Previderè C, Wilson IJ, Tyler-Smith C, Jobling MA. A predominantly neolithic origin for European paternal lineages. PLoS Biol. 2010; 8:e1000285.

30. Barbujani G and Bertorelle G. Genetics and the population history of Europe. Proceedings of the National Academy of Sciences. 2001; 98(1):22-25.
31. Barbujani G and Chikhi L. DNAs from the European Neolithic. Heredity. 2006; 97:84-85.
32. Barnard H, Dooley AN, Areshian G, Gasparyan B, Faull KF. Chemical evidence for wine production around 4000 BCE in the Late Chalcolithic Near Eastern highlands. Journal of Archaeological Science. 2011; 38(5):977-984.
33. Batini C, Hallast P, Zadik D, Delser PM, Benazzo A, Ghirotto S, Arroyo-Pardo E, Cavalleri GL, de Knijff P, Dupuy BM, Eriksen HA, et al. Large-scale recent expansion of European patrilineages shown by population resequencing. Nature communications. 2015; 6 (7152).
34. Behar DM, Metspalu M, Baran Y, Kopelman NM, Yunusbayen B, Gladstein A, Tzur S, Sahakyan H, Bahmanimehr A, Yepiskoposyan L, Tambets K. No evidence from genome-wide data of a Khazar origin for the Ashkenazi Jews. Human biology. 2013; 85(6):859-900.
35. Behar DM, Yunusbayev B, Metspalu M, Metspalu E, Rosset S, Parik J, Rootsi S, Chaubey G, Kutuev I, Yudkovsky G, Khusnutdinova EK. The genome-wide structure of the Jewish people. Nature. 2010; 466(7303):238-242.
36. Bellwood P and Oxenham M. The expansions of farming societies and the role of the Neolithic demographic transition. In: Bocquet-Appel JP, Bar-Yosef O (Eds.). The Neolithic Demographic Transition and its Consequences. Springer, Netherlands. 2008; 13-34.
37. Bellwood P. First Migrants: Ancient Migration in Global Perspective. Wiley-Blackwell, Chichester; 2013. -326 pp.
38. Bintliff J. The complete archaeology of Greece: from hunter-gatherers to the 20th century AD; John Wiley & Sons, Chichester; 2012. -544 pp.
39. Bouckaert R, Lemey P, Dunn M, Greenhill SJ, Alekseyenko AV, Drummond AJ, Gray RD, Suchard MA, Atkinson QD. Mapping the origins and expansion of the Indo-European language family. Science. 2012; 337(6097):957-960.

40. Bramanti B, Thomas MG, Haak W, Unterlaender M, Jores P, Tambets K, Antanaitis-Jacobs I, Haidle MN, Jankauskas R, Kind CJ, Lueth F. Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. Science. 2009; 326(5949):137-140.
41. Brandt G, Haak W, Adler CJ, Roth C, Szécsényi-Nagy A, Karimnia S, Möller-Rieker S, Meller H, Ganslmeier R, Friederich S, Dresely V, et al. Ancient DNA reveals key stages in the formation of central European mitochondrial genetic diversity. Science. 2013; 342(6155):257-261.
42. Brown TA, Jones MK, Powell W, Allaby RG. The Complex Origins of Domesticated Crops in the Fertile Crescent. Trends in Ecology & Evolution. 2009; 24(2):103-109.
43. Busby GB, Brisighelli F, Sánchez-Diz P, Ramos-Luis E, Martinez-Cadenas C, Thomas MG, Bradley DG, Gusmão L, Winney B, Bodmer W, Vennemann M, Coia V, Scarnicci F, Tofanelli S, Vona G, Ploski R, Vecchiotti C, Zemunik T, Rudan I, Karachanak S, Toncheva D, Anagnostou P, Ferri G, Rapone C, Hervig T, Moen T, Wilson JF, Capelli C. The peopling of Europe and the cautionary tale of Y chromosome lineage R-M269. Proc Biol Sci. 2012; 279:884-892.
44. Cassidy LM, Martiniano R, Murphy EM, Teasdale MD, Mallory J, Hartwell B, Bradley DG. Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome. Proceedings of the National Academy of Sciences. 2016; 113(2):368-373.
45. Cavalli-Sforza LL and Feldman MW. The application of molecular genetic approaches to the study of human evolution. Nature Genetics. 2003; 33:266-275.
46. Cavalli-Sforza LL and Minch E. Paleolithic and Neolithic lineages in the European mitochondrial gene pool. American journal of human genetics. 1997; 61(1):247.
47. Cavalli-Sforza LL, Menozzi P, Piazza A. The History and Geography of Human Genes. Princeton University Press, Princeton; 1994. -432 pp.

48. Chiaroni J, Underhill PA, Cavalli-Sforza LL. Y chromosome diversity, human expansion, drift, and cultural evolution. Proc Natl Acad Sci U S A. 2009; 106:20174-20179.
49. Chikhi L, Nichols RA, Barbujani G, Beaumont MA. Y genetic data support the Neolithic demic diffusion model. Proc Natl Acad Sci U S A. 2002; 99:11008-11013.
50. Cinnioğlu C, King R, Kivisild T, Kalfoğlu E, Atasoy S, Cavalleri GL, Lillie AS, Roseman CC, Lin AA, Prince K, Oefner PJ, Shen P, Semino O, Cavalli-Sforza LL, Underhill PA. Excavating Y-chromosome haplotype strata in Anatolia. Hum Genet. 2004; 114:127-148.
51. Cole SM. The Neolithic Revolution (No. 541). British Museum (Natural History), London; 1970. -70 pp.
52. College S, Conolly J, Shennan S. Archaeobotanical evidence for the spread of farming in the Eastern Mediterranean. Curr Anthropol. 2004; 45: S35-S58.
53. Davison K, Dolukhanov P, Sarson GR, Shukurov A. The role of waterways in the spread of the Neolithic. J Arch Sci. 2006; 33:641-652.
54. de Lumley H, Lordkipanidze D, Féraud G, Garcia T, Perrenoud C, Falguères C, Gagnepain J, Saos T, Voinchet P. Datation par la Méthode 40Ar / 39Ar de la Couche de Cendres Volcaniques (Couche VI) de Dmanissi (Géorgie) qui a Livré des Restes d'Hominidés Fossiles de 1,81 Ma. Comptes Rendus Palevol. 2002; 1(3):181-189.
55. de Morgan JJM. The History of the Armenian People, from the Remotest Times to the Present Day. Hairenik Press, Boston; 1918. -430 pp.
56. Di Gaetano C, Cerutti N, Crobu F, Robino C, Inturri S, Gino S, Guarrera S, Underhill PA, King RJ, Romano V, Cali F. Differential Greek and northern African migrations to Sicily are supported by genetic evidence from the Y chromosome. European Journal of Human Genetics. 2009; 17(1):91-99.
57. Di Giacomo F, Luca F, Popa LO, Akar N, Anagnou N, Banyko J, Brdicka R, Barbujani G, Papola F, Ciavarella G, Cucci F, Di Stasi L, Gavrila L, Kerimova

MG, Kovatchev D, Kozlov AI, Loutradis A, Mandarino V, Mammi' C, Michalodimitrakis EN, Paoli G, Pappa KI, Pedicini G, Terrenato L, Tofanelli S, Malaspina P, Novelletto A. Y-chromosomal haplogroup J as a signature of the post-neolithic colonization of Europe. Hum Genet. 2004; 115:357-371.

58. Diamond J and Bellwood P. Farmers and their languages: the first expansions. Science. 2003; 300:597-603.
59. Dolukhanov P, Aslanyan S, Kolpakov E, Belyayeva E. Prehistoric sites in northern Armenia. Antiquity. 2004; 78: (http://antiquity.ac.uk/projgall/ dolukhanov301/).
60. Dupanloup I, Bertorelle G, Chikhi L, Barbujani G. Estimating the impact of prehistoric admixture on the genome of Europeans. Mol Biol Evol. 2004; 21:1361-1372.
61. Ehlers J, Gibbard PL, and Hughes PD (Eds.). Quaternary Glaciations-Extent and Chronology: A Closer Look. Developments in Quaternary Science. Elsevier; UK, Netherlands. 2011. -1108 pp.
62. Evans JD. The early millennia: continuity and change in a farming settlement. In: Evely D, Hughes-Brock H, Momigliano N (Eds.). Knossos: A Labyrinth of History: Papers Presented in Honour of Sinclair Hood. British School at Athens, London. 1994; 1-20.
63. Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform. Online.2005; 1:47-50.
64. Fernández E, Pérez-Pérez A, Gamba C, Prats E, Cuesta P, Anfruns J, Molist M, Arroyo-Pardo E, Turbón D. Ancient DNA analysis of 8000 BC near eastern farmers supports an early neolithic pioneer maritime colonization of Mainland Europe through Cyprus and the Aegean Islands. PLoS Genet. 2014; 10:e1004401.
65. Fernández-Jalvo Y, King T, Andrews P, Yepiskoposyan L, Moloney N, Murray J, Domínguez- Alonso P, Asryan L, Ditchfield P, Made J, Torres T, Sevilla P,

Nieto Díaz M, Cáceres I, Allué E, Marín Monfort MD, Sanz Martín T. The Azokh Cave Complex: Middle Pleistocene to Holocene Human Occupation in the Caucasus. Journal of Human Evolution. 2010; 58:103-109.

66. Fort J. Demic and cultural diffusion propagated the Neolithic transition across different regions of Europe. Journal of The Royal Society Interface. 2015; 12(106):20150166.
67. Fu Q, Rudan P, Pääbo S, Krause J. Complete mitochondrial genomes reveal Neolithic expansion into Europe. PLoS One. 2012; 7(3):e32473.
68. Gabunia L, Antón SC, Lordkipanidze D, Vekua A, Justus A, Swisher CC. Dmanisi and Dispersal. Evolutionary Anthropology: Issues, News, and Reviews. 2001; 10(5): 158-170.
69. Gamba C, Jones ER, Teasdale MD, McLaughlin RL, Gonzalez-Fortes G, Mattiangeli V, Domboróczki L, Kővári I, Pap I, Anders A, Whittle A. Genome flux and stasis in a five millennium transect of European prehistory. Nature communications. 2014; 5:5257.
70. Gandilyan PA. Archaeobotanical evidence for evolution of cultivated wheat and barley in Armenia. Proceedings of the Harlan Symposium "The Origins of Agriculture and the Domestication of Crop Plants in the Near East". 10-14 May, 1997, Aleppo, Syria. ICARDA. 1998; 280-285.
71. Gasparyan B. Landscape organization and resource management in the Lower Palaeolithic of Armenia. Turkish Academy of Sciences Journal of Archaeology. 2010; 13:159-183.
72. Ghukasyan R, Colonge D, Nahapetyan S, Ollivier V, Gasparyan B, Monchot H, Chataigner C. Kalavan-2 (North of Lake Sevan, Armenia): A New Late Middle Paleolithic Site in the Lesser Caucasus. Archaeology, Ethnology and Anthropology of Eurasia. 2010; 38(4):39-51.
73. Gimbutas M. Proto-Indo-European culture: the Kurgan culture during the fifth, fourth, and third millennia BC. In: Cardona G, Hoenigswald H, Senn A (Eds.).

Indo-European and Indo-Europeans. University of Pennsylvania Press, Philadelphia. 1970; 155-197.

74. Gobejishvili R, Lomidze N, Tielidze L. Late Pleistocene (Wurmian) Glaciations of the Caucasus. In Ehlers J, Gibbard PL, and Hughes PD (Eds.). Quaternary Glaciations-Extent and Chronology: A Closer Look. Developments in Quaternary Science. Elsevier, UK, Netherlands. 2011; 141-148.
75. Greppin JA, Diakonoff IM. Some effects of the Hurro-Urartian people and their languages upon the earliest Armenians. Journal of the American Oriental Society. 1991;111:720-730.
76. Haak W, Balanovsky O, Sanchez JJ, Koshel S, Zaporozhchenko V, Adler CJ, Der Sarkissian CS, Brandt G, Schwarz C, Nicklisch N, Dresely V. Ancient DNA from European early neolithic farmers reveals their near eastern affinities. PLoS Biol. 2010; 8(11):e1000536.
77. Haak W, Forster P, Bramanti B, Matsumura S, Brandt G, Tänzer M, Villems R, Renfrew C, Gronenborn D, Alt KW, Burger J. Ancient DNA from the first European farmers in 7500-year-old Neolithic sites. Science. 2005; 310:1016-1018.
78. Haak W, Lazaridis I, Patterson N, Rohland N, Mallick S, Llamas B, Brandt G, Nordenfelt S, Harney E, Stewardson K, Fu Q, et al. Massive migration from the steppe was a source for Indo-European languages in Europe. Nature. 2015; 522:207-211.
79. Haber M, Gauguier D, Youhanna S, Patterson N, Moorjani P, Botigué LR, Platt DE, Matisoo-Smith E, Soria-Hernanz DF, Wells RS, Bertranpetit J. Genome-wide diversity in the Levant reveals recent structuring by culture. PLoS Genet. 2013; 9(2):e1003316.
80. Haber M, Mezzavilla M, Xue Y, Comas D, Gasparini P, Zalloua P, Tyler-Smith C. Genetic evidence for an origin of the Armenians from Bronze Age mixing of multiple populations. European Journal of Human Genetics. 2015; online publication, doi: 10.1038/ejhg.2015.206.

81. Haber M, Mezzavilla M, Xue Y, Tyler-Smith C. Ancient DNA and the rewriting of human history: be sparing with Occam's razor. Genome biology. 2016; 17(1):1-8.
82. Haber M, Platt DE, Badro DA, Xue Y, El-Sibai M, Bonab MA, Youhanna SC, Saade S, Soria-Hernanz DF, Royyuru A, Wells RS, Tyler-Smith C, Zalloua PA; Genographic Consortium: Influences of history, geography, and religion on genetic structure: the Maronites in Lebanon. Eur J Hum Genet. 2011; 19:334-340.
83. Harutyunyan A, Khudoyan A, Yepiskoposyan L. Patrilocality and recent migrations have little impact on shaping patterns of genetic structure of the armenian population. Russian journal of genetics. 2009; 45(8):987-993.
84. Hellenthal G, Busby GB, Band G, Wilson JF, Capelli C, Falush D, Myers S: A genetic atlas of human admixture history. Science. 2014; 343:747-751.
85. Herrera KJ, Lowery RK, Hadden L, Calderon S, Chiou C, Yepiskoposyan L, Regueiro M, Underhill PA, Herrera RJ. Neolithic patrilineal signals indicate that the Armenian plateau was repopulated by agriculturalists. Eur J Hum Genet. 2012; 20:313-320.
86. Hervella M, Izagirre N, Alonso S, Fregel R, Alonso A, Cabrera VM, de la Rúa C. Ancient DNA from hunter-gatherer and farmer groups from Northern Spain supports a random dispersion model for the Neolithic expansion into Europe. PLoS One. 2012; 7(4):e34417.
87. Hervella M, Rotea M, Izagirre N, Constantinescu M, Alonso S, Ioana M, Lazăr C, Ridiche F, Soficaru AD, Netea MG, de-la-Rua C. Ancient DNA from South-East Europe Reveals Different Events during Early and Middle Neolithic Influencing the European Genetic Heritage. PloS one. 2015; 10(6):e0128810.
88. Hewitt GM. Some Genetic Consequences of Ice Ages, and Their Role in Divergence and Speciation. Biological Journal of the Linnean Society. 1996; 58(3):247-276.

89. Hewitt GM. Speciation, Hybrid Zones and Phylogeography – Or Seeing Genes in Space and Time. Molecular Ecology. 2001; 10(3):537-549.
90. Hovhannisyan A and Yepiskoposyan L. Role of the Armenian Highland in the Neolithic migration of Homo sapiens from the Near East to Europe. International Journal of Anthropology. 2014; 29:171-182.
91. Hovhannisyan A, Khachatryan Z, Haber M, Hrechdakian P, Karafet T, Zalloua P, Yepiskoposyan L. Different waves and directions of Neolithic migrations in the Armenian Highland. Investigative genetics. 2014; 5(1):1.
92. Hovsepyan R and Willcox G. The earliest finds of cultivated plants in Armenia: evidence from charred remains and crop processing in Pisé from the Neolithic settlements of Aratashen and Aknashen. Veget Hist Archaeobot. 2008; 17:S63-S71.
93. Jobling MA and Tyler-Smith C. Fathers and sons: the Y chromosome and human evolution. Trends Genet. 1995; 11:449-456.
94. Jobling MA and Tyler-Smith C. The human Y chromosome: an evolutionary marker comes of age. Nat Rev Genet. 2003; 4:598-612.
95. Jones ER, Gonzalez-Fortes G, Connell S, Siska V, Eriksson A, Martiniano R, McLaughlin RL, Llorente MG, Cassidy LM, Gamba C, Meshveliani T. Upper Palaeolithic genomes reveal deep roots of modern Eurasians. Nature communications. 2015; 6:8912.
96. Karmin M, Saag L, Vicente M, Sayres MA, Järve M, Talas UG, Rootsi S, Ilumäe AM, Mägi R, Mitt M, Pagani L, et al. A recent bottleneck of Y chromosome diversity coincides with a global change in culture. Genome research. 2015; 25(4):459-466.
97. Kayser M, Krawczak M, Excoffier L, Dieltjes P, Corach D, Pascali V, Gehrig C, Bernini LF, Jespersen J, Bakker E, Roewer L, de Knijff P. An extensive analysis of Y-chromosomal microsatellite haplotypes in globally dispersed human populations. Am J Hum Genet. 2001; 68:990-1018.

98. Keller A, Graefen A, Ball M, Matzas M, Boisguerin V, Maixner F, Leidinger P, Backes C, Khairat R, Forster M, Stade B. New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. Nature communications. 2012; 3:698.
99. Kiguradze T. Caucasian Neolithic. In: Peregrine PN and Ember M. (Eds.). Encyclopedia of Prehistory. Europe. Kluwer Academic/Plenum. New York. 2001; vol 4, 55-76.
100. King RJ, Özcan SS, Carter T, Kalfoğlu E, Atasoy S, Triantaphyllidis C, Kouvatsi A, Lin AA, Chow CE, Zhivotovsky LA, Michalodimitrakis M. Differential Y-chromosome Anatolian Influences on the Greek and Cretan Neolithic. Annals of human genetics. 2008; 72(2):205-214.
101. King T, Fernandez-Jalvo Y, Moloney N, Andrews P, Melkonyan A, Ditchfield P, Yepiskoposyan L, Karapetyan S. Exploration and Survey of Pleistocene Hominid Sites in Armenia and Karabagh. Antiquity. 2003; 77(295).
102. Kivisild T. Maternal ancestry and population history from whole mitochondrial genomes. Investigative genetics. 2015; 6(1):1.
103. Kushnareva KK. The Southern Caucasus in Prehistory: Stages of Cultural and Socioeconomic Development from the Eighth to the Second Millennium BC. University of Pennsylvania Museum, Philadelphia; 1997. -296 pp.
104. Lacan M, Keyser C, Ricaut FX, Brucato N, Duranthon F, Guilaine J, Crubézy E, Ludes B. Ancient DNA reveals male diffusion through the Neolithic Mediterranean route. Proceedings of the National Academy of Sciences. 2011; 108(24):9788-9791.
105. Lang DM. Armenia: Cradle of Civilization. Allen & Unwin, London; 1980. -320 pp.
106. Larmuseau MH, Vanderheyden N, Van Geystelen A, Oven M, Knijff P, Decorte R. Recent Radiation within Y-chromosomal Haplogroup R-M269 Resulted in High Y-STR Haplotype Resemblance. Annals of human genetics. 2014; 78(2):92-103.

107. Lazaridis I, Patterson N, Mittnik A, Renaud G, Mallick S, Kirsanow K, Sudmant PH, Schraiber JG, Castellano S, Lipson M, Berger B, Economou C, Bollongino R, Fu Q, Bos KI, Nordenfelt S, Li H, de Filippo C, Prüfer K, Sawyer S, Posth C, Haak W, Hallgren F, Fornander E, Rohland N, Delsate D, Francken M, Guinet JM, Wahl J, Ayodo G, et al. Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature. 2014; 513:409-413.
108. Li JZ, Absher DM, Tang H, Southwick AM, Casto AM, Ramachandran S, Cann HM, Barsh GS, Feldman M, Cavalli-Sforza LL, Myers RM. Worldwide human relationships inferred from genome-wide patterns of variation. Science. 2008; 319(5866):1100-1104.
109. Lippold S, Xu H, Ko A, Li M, Renaud G, Butthof A, Schroeder R, Stoneking M. Human paternal and maternal demographic histories: insights from high-resolution Y chromosome and mtDNA sequences. Investigative Genet. 2014; 5:13.
110. Lowery RK, Herrera KJ, Barrett DA, Rodriguez R, Hadden LR, Harutyunyan A, Margaryan A, Yepiskoposyan L, Herrera RJ. Regionalized autosomal STR profiles among Armenian groups suggest disparate genetic influences. Am J Phys Anthropol. 2011; 146:171-178.
111. Margaryan A, Harutyunyan A, Khachatryan Z, Yepiskoposyan L. HLA Allele and Haplotype Frequencies in the Armenian Population: Evidence of Mediterranean Origin. International Journal of Antropology. 2011; 26:195-210.
112. Mathieson I, Lazaridis I, Rohland N, Mallick S, Patterson N, Roodenberg SA, Harney E, Stewardson K, Fernandes D, Novak M, Sirak K. Genome-wide patterns of selection in 230 ancient Eurasians. Nature. 2015; 528(7583):499-503.
113. Menozzi P, Piazza A, Cavalli-Sforza L. Synthetic maps of human gene frequencies in Europeans. Science. 1978; 201:786-792.
114. Morelli L, Contu D, Santoni F, Whalen M, Francalacci P, Cucca F. A comparison of Y chromosome variation in Sardinia and Anatolia is more

consistent with cultural rather than demic diffusion of agriculture. PLoS One. 2010; 5:e10419.

115. Myres NM, Rootsi S, Lin AA, Järve M, King RJ, Kutuev I, Cabrera VM, Khusnutdinova EK, Pshenichnov A, Yunusbayev B, Balanovsky O, Balanovska E, Rudan P, Baldovic M, Herrera RJ, Chiaroni J, Di Cristofaro J, Villems R, Kivisild T, Underhill PA. A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe. Eur J Hum Genet. 2011; 19:95-101.
116. Nasidze I and Stoneking M. Mitochondrial DNA variation and language replacements in the Caucasus. Proceedings of the Royal Society of London B: Biological Sciences. 2001; 268(1472):1197-1206.
117. Nasidze I, Ling EY, Quinque D, Dupanloup I, Cordaux R, Rychkov S, Naumova O, Zhukova O, Sarraf-Zadegan N, Naderi GA, Asgary S, Sardas S, Farhud DD, Sarkisian T, Asadov C, Kerimov A, Stoneking M. Mitochondrial DNA and Y chromosome variation in the Caucasus. Ann Hum Genet. 2004; 68:205-221.
118. Olalde I, Allentoft ME, Sánchez-Quinto F, Santpere G, Chiang CW, DeGiorgio M, Prado-Martinez J, Rodríguez JA, Rasmussen S, Quilez J, Ramírez O. Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. Nature. 2014; 507(7491):225-228.
119. Ollivier V, Nahapetyan S, Roiron P, Gabrielyan I, Gasparyan B, Chataigner C, Joannin S, Cornee JJ, Guillou H, Scaillet S, Munch P, Krijgsman W. Quaternary Volcano-Lacustrine Patterns and Palaeobotanical Data in Southern Armenia. Quaternary International. 2010; 223:312-326.
120. Oppenheimer S. The Origins of the British: A Genetic Detective Story. Carroll & Graf, New York; 2006. -656 pp.
121. Otte M. The history of European populations as seen by archaeology. In: Renfrew C, Boyle K (Eds.). Archaeogenetics: DNA and the population prehistory of Europe. McDonald Institute for Archaeological Research, Cambridge. 2000; 41-44.

122. Pala M, Olivieri A, Achilli A, Accetturo M, Metspalu E, Reidla M, Tamm E, Karmin M, Reisberg T, Hooshiar Kashani B, Perego UA, Carossa V, Gandini F, Pereira JB, Soares P, Angerhofer N, Rychkov S, Al-Zahery N, Carelli V, Sanati MH, Houshmand M, Hatina J, Macaulay V, Pereira L, Woodward SR, Davies W, Gamble C, Baird D, Semino O, Villems R, Torroni A, Richards MB. Mitochondrial DNA Signals of Late Glacial Recolonization of Europe from Near Eastern Refugia. The American Journal of Human Genetics. 2012; 90(5):915-924.
123. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. PLoS genet. 2006; 2(12):e190.
124. Pereira L, Richards M, Goios A, Alonso A, Albarrán C, Garcia O, Behar DM, Gölge M, Hatina J, Al-Gazali L, Bradley DG. High-resolution mtDNA evidence for the late-glacial resettlement of Europe from an Iberian refugium. Genome research. 2005; 15(1):19-24.
125. Pinhasi R, Fort J, Ammerman AJ. Tracing the origin and spread of agriculture in Europe. PLoS Biol. 2005; 3:e410.
126. Pinhasi R, Gasparian B, Wilkinson K, Bailey R, Bar-Oz G, Bruch A, Chataigner C, Hoffmann D, Hovsepyan R, Nahapetyan S, Pike AW, Schreve D, Stephens M. Hovk 1 and the Middle and Upper Paleolithic of Armenia: A Preliminary Framework. Journal of Human Evolution. 2008; 55(5):803-816.
127. Presnyakov SL, Belyaeva EV, Lyubin VP, Rodionov NV, Antonov AV, Saltykova AK, Berezhnaya NG, Sergeev SA. Age of the Earliest Paleolithic Sites in the Northern Part of the Armenian Highland by SHRIMP-II U–Pb Geochronology of Zircons from Volcanic Ashes. Gondwana Research. 2012; 21(4):928-938.
128. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000; 155(2):945-959.
129. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, De Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-

genome association and population-based linkage analyses. The American Journal of Human Genetics. 2007; 81(3):559-575.

130. Raghavan M, Skoglund P, Graf KE, Metspalu M, Albrechtsen A, Moltke I, Rasmussen S, Stafford Jr TW, Orlando L, Metspalu E, Karmin M. Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. Nature. 2014; 505(7481):87-91.
131. Rasmussen M, Li Y, Lindgreen S, Pedersen JS, Albrechtsen A, Moltke I, Metspalu M, Metspalu E, Kivisild T, Gupta R, Bertalan M. Ancient human genome sequence of an extinct Palaeo-Eskimo. Nature. 2010; 463(7282):757-762.
132. Rasteiro R, Bouttier PA, Sousa VC, Chikhi L. Investigating sex-biased migration during the Neolithic transition in Europe, using an explicit spatial simulation framework. Proceedings of the Royal Society of London B: Biological Sciences. 2012; rspb20112323.
133. Redgate AE and Redgate AE. The Armenians. Blackwell, Oxford; 1998.
134. Renfrew C. Archaeology and Language: The Puzzle of Indo-European Origins. Cambridge University Press, Cambridge; 1990. -368 pp.
135. Renfrew C. Language families and the spread of farming. In: Harris DR (Ed.), The Origins and Spread of Agriculture and Pastoralism in Eurasia. Smithsonian Institution Press, Washington DC, 1996; 70-92.
136. Richards M, Macaulay V, Hickey E, Vega E, Sykes B, Guida V, Rengo C, Sellitto D, Cruciani F, Kivisild T, Villems R. Tracing European founder lineages in the Near Eastern mtDNA pool. The American Journal of Human Genetics. 2000; 67(5):1251-1276.
137. Richards M, Macaulay V, Torroni A, Bandelt HJ. In search of geographical patterns in European mitochondrial DNA. The American Journal of Human Genetics. 2002; 71(5):1168-1174.
138. Riehl S, Zeidi M, Conard NJ. Emergence of agriculture in the foothills of the Zagros Mountains of Iran. Science. 2013; 341:65-67.

139. Roostalu U, Kutuev I, Loogväli EL, Metspalu E, Tambets K, Reidla M, Khusnutdinova EK, Usanga E, Kivisild T, Villems R. Origin and expansion of haplogroup H, the dominant human mitochondrial DNA lineage in West Eurasia: the Near Eastern and Caucasian perspective. Molecular biology and evolution. 2007; 24(2):436-448.
140. Rootsi S, Myres NM, Lin AA, Järve M, King RJ, Kutuev I, Cabrera VM, Khusnutdinova EK, Varendi K, Sahakyan H, Behar DM, Khusainova R, Balanovsky O, Balanovska E, Rudan P, Yepiskoposyan L, Bahmanimehr A, Farjadian S, Kushniarevich A, Herrera RJ, Grugni V, Battaglia V, Nici C, Crobu F, et al. Distinguishing the co-ancestries of haplogroup G Y-chromosomes in the populations of Europe and the Caucasus. Eur J Hum Genet. 2012; 20:1275-1282.
141. Rosser ZH, Zerjal T, Hurles ME, Adojaan M, Alavantic D, Amorim A, Amos W, Armenteros M, Arroyo E, Barbujani G, Beckman G, Beckman L, Bertranpetit J, Bosch E, Bradley DG, Brede G, Cooper G, Côrte-Real HB, de Knijff P, Decorte R, Dubrova YE, Evgrafov O, Gilissen A, Glisic S, Gölge M, Hill EW, Jeziorowska A, Kalaydjieva L, Kayser M, Kivisild T, et al. Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. Am J Hum Genet. 2000; 67:1526-1543.
142. Sampietro ML, Lao O, Caramelli D, Lari M, Pou R, Marti M, Bertranpetit J, Lalueza-Fox C. Palaeogenetic evidence supports a dual model of Neolithic spreading into Europe. Proceedings of the Royal Society of London B: Biological Sciences. 2007; 274(1622):2161-2167.
143. Schönberg A, Theunert C, Li M, Stoneking M, Nasidze I. High-throughput sequencing of complete human mtDNA genomes from the Caucasus and West Asia: high diversity and demographic inferences. European Journal of Human Genetics. 2011; 19(9):988-994.
144. Seielstad MT, Minch E, Cavalli-Sforza LL. Genetic evidence for a higher female migration rate in humans. Nature genetics. 1998; 20(3):278-280.

145. Semino O, Magri C, Benuzzi G, Lin AA, Al-Zahery N, Battaglia V, Maccioni L, Triantaphyllidis C, Shen P, Oefner PJ, Zhivotovsky LA, King R, Torroni A, Cavalli-Sforza LL, Underhill PA, Santachiara-Benerecetti AS. Origin, diffusion, and differentiation of Y-chromosome haplogroups E and J: inferences on the neolithization of Europe and later migratory events in the Mediterranean area. Am J Hum Genet. 2004; 74:1023-1034.
146. Sikora MJ, Colonna V, Xue Y, Tyler-Smith C. Modeling the contrasting Neolithic male lineage expansions in Europe and Africa. Investigative genetics. 2013; 4(1):1.
147. Simoni L, Calafell F, Pettener D, Bertranpetit J, Barbujani G. Geographic patterns of mtDNA diversity in Europe. The American Journal of Human Genetics. 2000; 66(1):262-278.
148. Sjödin P and François O. Wave-of-advance models of the diffusion of the Y chromosome haplogroup R1b1b2 in Europe. PLoS One. 2011; 6:e21592.
149. Soares P, Achilli A, Semino O, Davies W, Macaulay V, Bandelt HJ, Torroni A, Richards MB. The archaeogenetics of Europe.Current Biology. 2010; 20(4):R174-R183.
150. Stewart JR and Stringer CB. Human Evolution Out of Africa: The Role of Refugia and Climate Change. Science. 2012; 335(6074):1317-1321.
151. Stewart JR, Lister AM, Barnes I, Dalén L. Refugia Revisited: Individualistic Responses of Species in Space and Time. Proceedings of the Royal Society B: Biological Sciences. 2010; 277(1682):661-671.
152. Tanno KI and Willcox G. Distinguishing wild and domestic wheat and barley spikelets from early Holocene sites in the Near East. Vegetation history and archaeobotany. 2012; 21(2):107-115.
153. Tarasov PE, Volkova VS, Webb TIII, Guiot J, Andreev AA., Bezusko LG, Bezusko TV, Bykova GV, Dorofeyuk NI, Kvavadze EV, Osipova IM, Panova NK, Sevastyanov DV. Last Glacial Maximum Biomes Reconstructed from

Pollen and Plant Macrofossil Data from Northern Eurasia. Journal of Biogeography. 2000; 27(3):609-620.

154. Team RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria; 2013.
155. Torroni A, Bandelt HJ, Macaulay V, Richards M, Cruciani F, Rengo C, Martinez-Cabrera V, Villems R, Kivisild T, Metspalu E, Parik J. A signal, from human mtDNA, of postglacial recolonization in Europe. The American Journal of Human Genetics. 2001; 69(4):844-852.
156. Underhill PA and Kivisild T. Use of Y chromosome and mitochondrial DNA population structure in tracing human migrations. Annu Rev Genet. 2007; 41:539-564.
157. Underhill PA, Myres NM, Rootsi S, Metspalu M, Zhivotovsky LA, King RJ, Lin AA, Chow CE, Semino O, Battaglia V, Kutuev I, Järve M, Chaubey G, Ayub Q, Mohyuddin A, Mehdi SQ, Sengupta S, Rogaev EI, Khusnutdinova EK, Pshenichnov A, Balanovsky O, et al. Separating the post-Glacial coancestry of European and Asian Y chromosomes within haplogroup R1a. Eur J Hum Genet. 2010; 18:479-484.
158. Vigne JD, Briois F, Zazzo A, Willcox G, Cucchi T, Thiébault S, Carrère I, Franel Y, Touquet R, Martin C, Moreau C, Comby C, Guilaine J. First wave of cultivators spread to Cyprus at least 10,600 y ago. Proc Natl Acad Sci U S A. 2012; 109:8445-8449.
159. Weale ME, Yepiskoposyan L, Jager RF, Hovhannisyan N, Khudoyan A, Burbage-Hall O, Bradman N, Thomas M. Armenian Y Chromosome Haplotypes Reveal Strong Regional Structure Within a Single Ethno-National Group. Human Genetics. 2001; 109(6):659-674.
160. Wei W, Ayub Q, Chen Y, McCarthy S, Hou Y, Carbone I, Xue Y, Tyler-Smith C. A calibrated human Y-chromosomal phylogeny based on resequencing. Genome research. 2013; 23(2):388-395.

161. Wright EM. The Eighth Campaign of Sargon II of Assyria (714 BC). Journal of Near Eastern Studies. 1943; 2(3):173-186.
162. Yunusbayev B, Metspalu M, Järve M, Kutuev I, Rootsi S, Metspalu E, Behar DM, Varendi K, Sahakyan H, Khusainova R, Yepiskoposyan L, Khusnutdinova EK, Underhill PA, Kivisild T, Villems R. The Caucasus as an asymmetric semipermeable barrier to ancient human migrations. Mol Biol Evol. 2012; 29:359-365.
163. Yunusbayev B, Metspalu M, Metspalu E, Valeev A, Litvinov S, Valiev R, Akhmetova V, Balanovska E, Balanovsky O, Turdikulova S, Dalimova D. The genetic legacy of the expansion of Turkic-speaking nomads across Eurasia. PLoS Genet. 2015; 11(4):e1005068.
164. Zalloua PA, Xue Y, Khalife J, Makhoul N, Debiane L, Platt DE, Royyuru AK, Herrera RJ, Hernanz DF, Blue-Smith J, Wells RS, Comas D, Bertranpetit J, Tyler-Smith C, Genographic Consortium. Y-chromosomal diversity in Lebanon is structured by recent historical events. Am J Hum Genet. 2008a; 82:873-882.
165. Zalloua PA, Platt DE, El Sibai M, Khalife J, Makhoul N, Haber M, Xue Y, Izaabel H, Bosch E, Adams SM, Arroyo E. Identifying genetic traces of historical expansions: Phoenician footprints in the Mediterranean. The American Journal of Human Genetics. 2008b; 83(5):633-642.
166. Zheng HX, Yan S, Qin ZD, Jin L. MtDNA analysis of global populations support that major population expansions began before Neolithic Time. Scientific reports. 2012; 2(745).
167. Zohary D, Hopf M, Weiss E. Domestication of Plants in the Old World: The origin and spread of domesticated plants in Southwest Asia, Europe, and the Mediterranean Basin. Oxford University Press, UK; 2012.
168. Zvelebil M and Zvelebil K. Agricultural transition and Indo-European dispersal. Antiquity. 1988; 62:574-583.