Home	Sources	Genetics	Ogur and Oguz	English-Türkic Lexicon	Alan Dateline	Kimak Dateline
Back	Roots	Geography	Lingo-Ethnical Tree	Türkic in Romance	Avar Dateline	Kipchak Dateline
In Russian	Tamgas	Archeology	Indo-European, Aryans, Dravidian, and Rigveda	Alans in Pyrenees	Besenyo Dateline	Khazar Dateline
Contents Genetics	Alphabet	Religion	Scythian Ethnic Affiliation	Türkic in Greek	Bulgar Dateline	Kyrgyz Dateline
Contents Pazyryk Genetics	Writing	Coins	Anti-Iranian Arguments	Türkic-Sumerian	Huns Dateline	Sabir Dateline
Contents Amerind Genetics	Language	Wikipedia	Foundation of the Scythian-Iranian theory	Türkic-Etruscan	Karluk Dateline	Seyanto Dateline

Overview of Türkic genetics

Anatole A. Klyosov

Ancient History of the Arbins, Bearers of Haplogroup R1b, from Central Asia to Europe, 16,000 to 1,500 Years before Present

Contents Содержание

Links

http://www.scirp.org/journal/aa/
http://aklyosov.home.comcast.net

Editorial Introduction

This is seminal. At the SciRP, the number of downloads for the series relates as 2/2.5/8.5/1 for R1a/R1b/Out-of-Africa/Ave counts, with the *Ave counts* reflecting counts for the next few runner-ups; the understanding of the Dr. A. Klyosov's analysis sweeps the day. The Y-DNA markers R1a and R1b allow to make visible the veins of the Eurasian history like the iodine agents allow to inspect a cardiovascular system. And not unlike the iodine agents, these markers identify demographical bottlenecks and founder events, allowing to learn of the demographic splits. Not too much can be added: the cautious euphemistic Arbins left their traces not only in the veins of the modern and ancient people, but also in the numerous ethnic groups that are united by their origin and history, the most explicit of which is the multitude of the Türkic people and their close and remote siblings, with their geography paralleling the spread of the euphemistic Arbins from one end of the Eurasia to the other, with excursions to the north and to the south. The 16,000±1400 years were more than enough to do some travel, intermixing, waning and ascensions while transitioning from the Holocene to the Nuclear Age. More than half of that period, for about 9,000 years, the euphemistic Arbins carried around their Kurgan burial culture, sending their deceased off on a travel to Tengri for incarnation, equipped with all the travel paraphernalia including food supplies as best as they could afford. The Kurgan burial tradition extended well into the literate history, leaving us not only with the physical relicts from China to Ireland, but with numerous eyewitness accounts of the rites and interpretations. The Kurgan economy was largely unaffected by the sedentary agricultural producing economy, it managed to turn their initial hunter-gatherer economy into an animal husbandry producing economy, and carry that economy well into the present. The Kurgan nomadic culture stimulated many innate effects, like military aptitude, gender equality, independence and egalitarianism, techniques of domination over chattel and humans, and unequalled self-reliance. It produced etiology that not only left deep trails in the past religious doctrines, but in numerous morphed forms is still with us. It affected biological make-up of its peoples, weeding out those unable to adjust to the milk-and-meat diet, to the mobile lifestyle, to the social and ethnical stresses, and to yet not understood maladies. Numerous examples of superior adaptability are reflected in our recent historical memory: the Huns moved their state from the Ordos bend to the Aral and on to the Pannonia, the Kangars moved their state from the lake Balkhash to the Adriatic, the Tele's Seyanto moved their state from Mongolia to the Kimak Kaganate in the South Siberia and to the Kipchak Khanate in the Eastern Europe, the Suvars moved their state from the Caucasus to the Ukraine and to the Itil-Kama interfluve, the Oguz Turks moved their state from the Balkhash to Aral and on to the Mesopotamia and Anatolia, to list just the few of the many more.

This is still another mighty breakthrough analysis of Dr. A. Klyosov that will open floodgates for new explorations and discoveries in Turkology. The linguistic scenery has flexive IE languages, and the agglutinative Fennic, Türkic, and Chinese languages, their interactions are yet terra incognita. Of the ca. 10-8 mill. BC "Arans", "Aryans", or "non-Aryans", the R1a migrants that did or did not venture to the Balkans include numerous Türkic- and Fennic-speaking ethnicities:

Modern or formerly Türkic-speaking Southern Altaians (53%, largely Tele/South Siberia), Pashtuns (51% R1a, blend of Abdaly/Durrani/Ephtalite Huns and Saka, South Siberia), Kirgiz (40%, descendants of the Siberian Enisei Kirgiz, South Siberia), Northern Altaians (38%, largely Kipchks, South Siberia), Croats/Hrvati (35%, formerly Bechen/Patsinak tribe Charaboi, Middle Asia), Chuvash (30%, largely Hunno-Bulgars, South Siberia), Gagauzes (30%), Rajput (28%, Abdaly/Ephtalite Huns tribe, South Siberia), Karachai (27%, to some degree Ases/Alans/Masguts), Tatars (25%, largely Hunno-Bulgars, South Siberia), Uzbeks (25%, largely Karluks/Uigurs, South Siberia), Abdaly (20%, Abzeli of Bashkortstan/Ephtalite Huns, South Siberia), Uigurs (20%, South Siberia), Karakalpaks (18%), Bashkirs (9%); Mari (45%/Eastern Europe, Fennic), Komi (30%/Eastern Europe, Fennic), Mordvins (25%, Burtases/Eastern Europe, Fennic), Hungarians (20%/Eastern Europe, Fennic); the list is far from exhaustive, it is more a pointer than an inventory.

The Farsi Tajiks (16-64%, Turco-Perso-Arabic conflation), Indic speakers Punjabi (36-64%, with still surviving Türkic Alats/Khalaches), Hindustanis (35-70%). A very partial map of R1a hints to the scope of the R1a spread:

Map 3 Map of South-Eastern branch, with birthplaces of furthermost ancestors recorded in available databases (Klyosov and Rozhanskii, 2012)
Magenta pins denote a parent Z93 branch
red and green pins correspond to Bashkir and Kyrgyz branches, respectively. The rest of the branch (L342.2 positive) is deep-blue.
The most represented Ashkenazy Jewish branch (not shown) covers densely an area of historical Polish-Lithuanian Commonwealth

And numerous populations carry both the 20,000 years-old R1a1 line and the 16,000 years-old R1b line:

Croats/Hrvati (35% +15% respectively), Komi (30%+15%), Gagauzes (30%+15%), Karachai (27%+15%), Mordvins (25%+15%), Uzbeks (25%+10%), Hungarians (20%+20%), Uigurs (20%+20%), Abdaly (20%+65%, Abzeli of Bashkortstan), Karakalpaks (18%+9%), Kumyks(15%+21%), Balkars (13%+13%), Bashkirs (9%+86%).

These examples demonstrate cases of retention the original language (Abdaly of Bashkortstan, Hungarians, Mordvins, Komi, Mari) and of the language switch (Abdaly of Pashtuns, Croats/Hrvati, Abdaly of Rajputs, Abdaly of Jats, Russians of R1a1 and R1b lines). In the Eastern Europe after the 10th c. AD, the Ugro-Fennic speakers in the north and the Turkic speakers in the south were linguistically nearly completely Slavicized, and after the 16th c. AD in Siberia were linguistically Slavicized numerous ethnicities (just in a century between 1900 and 2000 in Russia were wiped out about 100 native languages, nearly all of them agglutinative non-IE). In the Central Europe were Germanized Sarmats and the Hunnic tribes, in Spain and France were Romanized Gauls, Alans and Goths.

They could be re-blending of the early R1a1 migrants from the period 20,000 ybp to 16,000 ybp with the later R1a1+ R1b migrants, or the post-16,000 ybp migrants that retained traces of their mixture at the time of separation into distinct groups. No paleogroup never knew who has what Y-DNA markers, and no political, economic, geographical, or religious division of society could occur along the borders of unknown Y-DNA markers. Only repeated stochastic filtering with statistically insufficient number of outcomes could lead to a purely random division of society into detectable sectors with singular Y-DNA markers: five coin tosses on the table may accidentally put 3 eagles on one half and two tails on another, but 500 coin tosses can not put 300 eagles on one and 200 tails on the other half. This model gives us the original picture: people are few, groups and producers are scattered, complete loss of the entire group is a commonplace, and recombination of the fractions is less frequent than their fatality. But by the time of the 4th mill. bp, the society of the 16th-10th mill. bp did not exist, the stochastic filtering would not work with larger and more cohesive groups, recombination of the fractions is way more frequent than their utter destruction, and no noble isolation to maintain invisible marker could have lasted. The Aryans (from ar = men) and non-Aryans (from non-ar = non-men, "an-iran") lived interspersed and mutually dependent, and the modern concept of ethnicity as people united by acceptance of common biological, historical, cultural, and linguistic origin can't be projected backwards across 4 millenniums based on any selected biological marker. The dashed line connecting the modern R1a Brahman caste with their R1a ancestors 4 mill. bp needs much more meditation to be sustained. The Earth is not flat, the composition of the modern ethno-national conglomerates is not single-dimensional, and their multi-dimensional nature can be reduced to 2-dimentional approximation only adhering to the rules of the differential calculus. Neither a linear R1a1 depiction, nor a flat depiction with R1a1 and R1b ordinates would produce adequate description. We can trace the development and movement of the markers, but the correlation of the markers with the

peoples and languages is a very multi-dimensional phenomenon.

Posting clarifications, comments, and additional *subheadings* are *(in blue italics)* and *blue boxes*. Page numbers are shown at the beginning of the page in blue. Most of the references on the author's work can be found on the author's homepage http://aklyosov.home.comcast.net/~aklyosov/. The adjective Türkic and the noun Türk are used to denote the global world for the Türkic community that includes Turkish and Turks as one of the constituents; Türk is a noun of which the Turkic and Türkic are adjectival derivatives, it is needed for translation from Russian, which has four distinct designations for four phenomena. The semantics of the above terminology in English vs. Russian is a result of their national histories.

Anatole A. Klyosov

Ancient History of the Arbins, Bearers of Haplogroup R1b, from Central Asia to Europe, 16,000 to 1,500 Years before Present

Summary

This article aims at reconstructing the history of R1b ancient migrations between 16,000 and 1500 years before present (ybp). Four thousand four hundred eight (4408) haplotypes of haplogroup R1b (with subclades) were considered in terms of base (ancestral) haplotypes of R1b populations and the calculated time to their common ancestors. The regions considered are from South Siberia/Central Asia in the east (where R1b haplogroup arose ~16,000 ybp) via the North Kazakhstan, South Ural to the East European Plain and further west to Europe (the northern route entering Europe around 4500 ybp); from the East European Plain south to the Caucasus (6000 ybp), Asia Minor (6000 ybp) and the Middle East (6000 - 5500 ybp) to the Balkans in Europe (the southern route, entering Europe around 4500 ybp); along North Africa and the Mediterranean Sea (5500 - 5000 ybp) via Egypt to the Atlantic, north to Iberia (the North African route with arrival to the Pyrenees 4800 ybp). The Arbins (bearers of R1b haplogroup) along their migration route to the Middle East and South Mesopotamia apparently have established the Sumer culture (and the state), moving westward to Europe (5000 - 4500 ybp) carrying mainly the R-M269 subclade and its downstream L23 subclade. This last subclade was nearly absent along the North African route, and/or did not survive the migration to Iberia or evidenced later. At the arrival to Iberia (4800 ybp) the M269 subclade split off M51 and soon thereafter the L11 downstream subclades. These populations became known as the Bell Beakers and moved north, along with the newly arisen subclades of P312 and L21 (which split off within a few centuries after P312). Those subclades and their downstream clades have effectively, without major interruptions, populated Europe (the smooth haplotype trees demonstrate the near non-stop proliferation of R1b haplotypes in Europe). They are evidenced from the Atlantic eastward to the Balkans, Carpathian Mountains, present day Poland to the western border of the East European Plain and up to the Baltic Sea.

The Isles had a different history of R1b migrations. The bearers of L11, P312 and L21 moved to the Isles by land and sea concurrently with those Arbins who were populating Europe between 4000 and 2500 ybp and formed the respective "local" subclades of P314, M222, L226, which largely populated the Isles. As a result, a significant part of the Isles is populated almost exclusively by the Arbins, whose frequency reaches 85% - 95% among the current population. In general, the frequency of Arbins in Western and Central Europe, reaches — albeit not uniformly — some 60% of the population. This study essentially presents an example of application of DNA genealogy in studying the history of mankind.

Introduction

Due to lack of a common name for the bearers of R1b haplogroup (with subclades) and their languages in ancient times, which were carried for millennia and eventually brought to Europe as non-Indo European languages, I refer to them as the Arbins (from R1b), both the people and their original languages, on the analogy with the Aryans *(from Arans, same format as Arbins)*, who essentially belonged to haplogroup R1a (Klyosov & Rozhanskii, 2012a).

The origin and history of haplogroup R1b, bearers I refer to as the Arbins, currently populate nearly 60% of Western and Central Europe. Additionally, Arbins populate significant parts of the Caucasus, Anatolia and Asia Minor, Middle East, and many locations in Central Asia, including South Siberia, Altai, Tuva, North-Western China, Middle Asia, some Ural and Middle Volga regions with ethnic groups and populations such as the Bashkirs, Tatars, Chuvash, and other.

History of R1b was significantly distorted from the beginning of "genetic genealogy" at the end of 1990s, when it was claimed, quite groundlessly, that R1b arose in Europe some 30,000 years before present (ybp). Groundlessly because, indeed, the claims were based on no data. Such data never existed. Nevertheless, statements and claims such as "*Around* 30,000 *years ago, a descendant of the clan making its way into Europe gave rise to marker M*343, *the defining marker of haplogroup R1b. These travelers are direct descendants of the people who dominated the human expansion into Europe, the Cro-Magnon*" (Spencer Wells, "Deep Ancestry", 2006). This and similar claims, such as R1b (and its M269 subclades) were "*well established throughout Paleolithic Europe*", "*contemporaneous with Aurignacian culture*", "*the earliest expansion into Europe, during the Upper Paleolithic ~30,000 years ago*" by Wells, Semino, Underhill, Cavalli-Sforza, Cinnioglu, Kivisild, Wiik and many others (e.g., Semino et al., 2000; Wells et al., 2001; Cinnioglu et al., 2004; Wiik, 2008) were essentially based on "thoughts" that if people lived in Europe some 30,000 years ago, they necessarily were of the R1b haplogroup, and not of I, G, J, E, F or any other haplogroups. Were any haplotypes analyzed? Their mutations counted? Any chronological evaluations? There was nothing of the above.

87

"Paleolithic origin" of R1b in Europe, or their "Paleolithic migrations" to Europe around 10,000 - 8000 ybp are still claimed in recent academic papers, such as (Myres et al., 2010; Balaresque et al., 2010; Morelli et al., 2010). These assumptions and resulting calculations are based, typically, on "population mutation rates" (Zhivotovsky et al., 2004; Hammer et al., 2009; Underhill et al., 2009), which commonly exaggerate the chronological estimates of migrations and events by 200% - 400%, since they are based on crude, artificial and unrealistically naïve and generalized reasoning (for critique, see Klyosov, 2009a, 2009b, 2009c; Rozhanskii & Klyosov, 2011). The "academic papers", placing origin of R1b in Asia Minor or nearby, have not considered regions east of Asia Minor as well as R1b haplotypes of those eastern regions. In short, the whole story of R1b migrations and their history is in disarray, and the "population geneticists" continue to advance misleading conclusions due to their methodology.

This study's methodology/analysis includes considerations of Y chromosome extended 67 and 111 marker haplotypes, when available. The methodology was described in detail in the preceding papers in this journal (Rozhanskii & Klyosov, 2011; Klyosov & Rozhanskii, 2012a) and elsewhere (e.g., Klyosov, 2009a, 2009c, 2009d), and in **Materials and Methods** section of this article.

As described (Klyosov & Rozhanskii, 2012b), Europeoids (Caucasoids) appeared ~58,000 ybp. They gradually branched to downstream haplogroups and their subclades, and migrated to the north, west, south and east. Haplogroup NOP, which was among them, arose ~48,000 ybp, and moved eastward,

presumably towards South Siberia and/or adjacent regions. Haplogroup P split off ~38,000 ybp, presumably in South Siberia, and gave rise to haplogroup R and then R1 ~30,000 - 26,000 ybp (see the diagram in Klyosov and Rozhanskii, 2012b). Haplogroup R1b arose ~16,000 ybp, as it will be shown further in this paper.

Fig 3 from Klyosov, A. & Rozhanskii, I. (2012). Re-Examining the "Out of Africa" Theory and the Origin of Europeoids (Caucasoids) in Light of DNA Genealogy
Advances in Anthropology, 2, 80-86
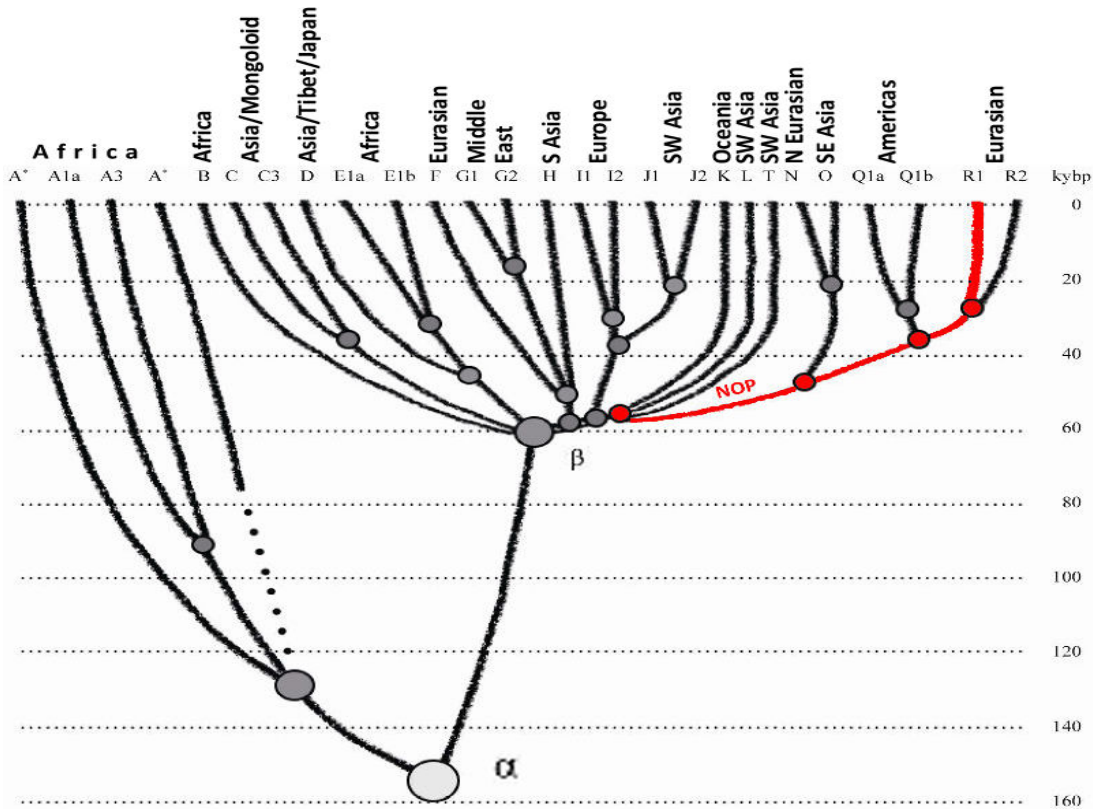
The timing may be reconstructed from a series of R1b haplotypes, made available from the databases (see **Appendix**). The most distant R1b haplotypes (those exhibiting the greatest mutational differences) from the European R1b haplotypes, were found in Siberia and Middle Asia (a part of Central Asia) populations. Central Asian R1b haplotype bearers have most ancient common ancestors with the European R1b bearers, and those ancient common ancestors lived ~16,000 ybp in Central Asia. We do not know as yet whether in South Siberia or Middle Asia; however, the evidence will demonstrate that it was somewhere in that vast region.

The following is a primer on the history of the R1b haplogroup taken from Klyosov A. (2011), *Haplotypes of R1b1a2-P312 and related subclades: origin and "ages" of most recent common ancestors// Proceedings of the Russian Academy of DNA Genealogy, vol. 4, No. 6, 2011*. The primer is on haplogroups, timing, migrations, and language:
1130

HISTORY OF R1b HAPLOGROUP

from Central Asia to Europe
Appearance of R1b in Central Asia (Altai, E.Turkestan, North-Western China), its ancient migrations to Europe; Bashkirs, Caucasian people, Anatolia, Sumer, North Africa, Iberia, the Basques, continental Europe, British Isles), Bell Beakers as R1b1a2-P312 and –U106 and their upstream (L51, L11/L151/P310/P311) and downstream subclades in Europe, and non-Bell Beakers R1b1a2-M269, -L23, and their downstream subclades in Europe

It was shown in a series of more than 30 papers (Klyosov, 2008-2011, see above and the list of references at the end of the paper) that haplogroup R1b1, as well as its brother haplogroup R1a1, arose in Central Asia (Southern Siberia or adjacent regions, such as Altai, E.Turkestan including North-Western China). R1a arose around 21,000 years before present (ybp), R1b1 arose next, about 16,000 ybp. Since the very slow (with respect to mutations) 22 marker European R1a1-M17 base (aka deduced ancestral) haplotype [DYS 426, 388, 392, 455, 454, 438, 531, 578, 395S1a, 395S1b, 590, 641, 472, 425, 594, 436, 490, 450, 617, 568, 640, 492]

12 12 11 – 11 11 – 11 – 11 8 17 17 8 10 8 12 10 12 12 8 12 11 11 12 **(R1a)**

and R1b1a2-M269 base 22 marker haplotype

12 12 **13** – 11 11 – **12** – 11 **9 15 16** 8 10 8 12 10 12 12 8 12 11 11 12 **(R1b1a2)**

differ by 7 mutations, as marked above (Klyosov, 2011a), which gives 7/0.006 = 1167 generations without correction for back mutations, or 1380 generations with the correction (the correction factor in this particular case equals 1.187, see pp. 1280-1281 in this issue), that is

34,600 years between their common ancestors, THEIR common ancestor lived approximately (34600+5000+7000)/2 = 23,300 years before present. This is a common ancestor of R1 haplogroup. Here 0.0060 is the mutation rate constant for the 22 marker haplotype (in mutation/haplotype/generation of 25 years) [Klyosov, 2011b, and the paper in this issue, pp. 1239-1256].

Indeed, if the R1 common ancestor lived 23,300 ybp, then the R1a1-M17 base haplotype in the Eastern Europe appeared (2,300 + 16,000) = 18,300 years after the R1 appearance, and the R1b1b2-M269 arose (7,300 + 9,000) = 16,300 years after the R1 appearance; hence, the difference between the two base haplotypes above equals to (18,300 + 16,300) = 34,600 years, exactly as the two 22 marker base haplotypes show as the seven mutation difference between them.

After haplogroup R1 arose about 23,000 years ago, and its downstream R1b arose in Central Asia 16,000 ago, the R1b had migrated across North Kazakhstan (including the later Botai archaeological culture, 6500-5500 ybp), through the territory of the present-day Bashkirs (13-11-8 thousand ybp), established Seroglazovo archaeological culture (13-11 thousand ybp) and then Middle Volga Basin archaeological cultures (Srednevolga culture 8-7 thousand ybp; Samara culture 7 thousand ybp; Srednestog culture 7500-5500 ybp; Khvalyn culture 7-6 thousand ybp, Kama culture 7-5 thousand ybp), that is the cultures which were (tentatively) named "Pre-Kurgan cultures"; R1b1a2-M269 subclade arose around 7,000 ybp, and then R1b1a2-L23, around 6,500 ybp. Both of them came to the Caucasus region around 6,500-6,000 ybp, and possibly earlier. Those R1b1b2 spoke non-Indo-European language(s), which can be vaguely traced now under various names. Some call them Proto-Türkic language, some Sino-Caucasian language, some Paleo-Eurasian language, some call them the Erbin language (because of R1b haplogroup of their bearers), some call them Western Caucasian or North-Western Caucasian languages, some call them the Basque language, and find some similarities between the Caucasian and Basque languages, etc.
1131

In fact, those similarities do exist, albeit in a rather weak form because thousands of years passed since then. Still now most of R1b1a2 haplotypes in the Caucasus (in Armenia, Dagestan, Georgia) belong to the ancient L23 subclade (with a common ancestor in the Caucasus of around 6,000 ybp), and have a characteristic DYS393=12 allele, unlike DYS393=13 in most of European R1b1b2 haplotypes.

From the Caucasus, R1b1a2-L23 and R1b1a2-M269 bearers went south over the mountains, to Anatolia (a common ancestor of 6,000 ybp), and then split into three major routes. One went further south, to Tigris and Euphrates rivers, and became the Sumers. Many present-day Assyrians, descendants of ancient Sumers, still have their R1b1a2 haplotypes. Another went westward, across Asia Minor, and came to Europe, to the Balkans and Mediterranean Sea region around 4500 ybp. The third group went across Northern Africa and Egypt (and, incidentally, might have left some R1b1b2 Pharaohs there) to the Atlantic and went across Gibraltar to the Iberian Peninsula around 4800 ybp. They became the Bell Beakers, and moved up north into the continental Europe. The Bell Beaker culture in Europe had lasted between about 4400 and 3800 ybp.

R1b1a2 bearers came to the Pyrenees, apparently, as mainly L51/M412 and/or L11/L151/P310 subclades, and soon split there into U106 and P312 downstream subclades, which went to the continental Europe as said Bell Beakers with downstream subclades such as U198, U152, L2, L20, L4, L21, M222, L226, SRY2627, etc. They left the Basques behind, which still maintain the ancient Erbin (R1b) language, similar in kind with some Caucasian languages, and with the Sumer language, and having some elements of Proto-Türkic languages back to Asia and further to Siberia.

Until the beginning of the 1st millennium BC the R1b1a2 language in Europe was predominantly (or only) non-Indo-European *(non-IE)*. There is not a single solid evidence of otherwise, that R1b1a2 in Europe spoke a IE language before the 1st millennium BC. Celtic language, as well as "Proto-Celtic" language, for instance, is placed by linguists not earlier than 800-900 BC, and all more earlier placements are typically groundless. At the same time, we know that it were R1a1 bearers who brought their IE language to India around 3500 ybp, and it is being dated in Europe to at least 6,000 ybp. There were no R1b1a2 in Europe those times. It mush have been R1a1 bearers who spoke proto-IE language(s) in Europe 6,000 ybp and some earlier.
1132

The roots of the Bell Beakers can be traced back by comparing 67 marker base haplotypes of their present-day descendants of "parallel" subclades U106 and P312:

13 **23** 14 11 11 14 12 12 12 13 13 29 – 17 9 10 11 11 25 15 19 29 15 15 17 17 – 11 11 19 23 **16** 15 **17** 17 **37 39** 12 12 – 11 9 15 16 8 10 10 8 10 10 12 23 23 16 10 12 12 15 8 12 22 20 13 12 11 13 11 11 **13** 12 **(U106)**

13 24 14 11 11 14 12 12 12 13 13 29 – 17 9 10 11 11 25 15 19 29 15 15 17 17 – 11 11 19 23 15 15 18 17 36 38 12 12 – 11 9 15 16 8 10 10 8 10 10 12 23 23 16 10 12 12 15 8 12 22 20 13 12 11 13 11 11 12 12 **(P312)**

There are six mutational differences between them (marked in the base U106 base haplotype above), which corresponds to 6/0.12 = 50 → 53 "conditional" generations (25 year per generation), that is 1,325 year time span between their common ancestors. The sign → means a correction for back mutations (see below). Since a common ancestor of U106 and P312 lived 4175±430 ybp and 4100±415 ybp respectively (see below, also 3950±400 ybp, Klyosov, 2010a, which is practically the same figure within the margin of error), THEIR common ancestor lived (1325+4175+4100)/2 = 4800 years ago. This fits to the Bell Beakers' timing quite well.

GENETIC SUMMARY

A detailed analysis of 2299 67-marker haplotype dataset provided by Michael Walsh, administrator of R1b1a2-L21 site http://www.familytreedna.com/public/R-L21/default.aspx?section=yresults has allowed to identify base (ancestral) 67 marker haplotypes and timespans to their common ancestors as follows (the scheme is deficient in Asiatic subclades):

• • • • • **R1b1a2**	L265, M269, M520, S3, S10, S13, S17 **~ 7000 ybp**	Asia
• • • • • • **R1b1a2a**	L23/S141, L49.1 **~ 6200 ybp**	Asia
• • • • • • • **R1b1a2a1a**	L51/M412/S167 **5300±700 ybp**	Asia

• • • • •	•		**R1b1a2a1a1**		L11, P310, P311	**~ 4800 ybp (alt. 4575±580 ybp)**	Europe
• • • • •	•	•	**R1b1a2a1a1a**		M405/S21/U106	**4175±430 ybp**	Europe
• • • • •	•	•	•	**null mutation U106**		**3325±450 ybp**	Europe
• • • • •	•	•	•	**R1b1a2a1a1b**	P312/S116	**~ 4800 ybp**	Europe
• • • • •	•	•	•	**null mutation P312**		**3575±400 ybp**	Europe
• • • • •	•	•	•	**R1b1a2a1a1b1** *M65*		**~1800 ybp**	Europe
• • • • •	•	•	•	**R1b1a2a1a1b2**	M153	**~ 3640 ybp**	Europe
• • • • •	•	•	•	**R1b1a2a1a1b3**	S28/U152	**4125±450 ybp (alt. 3800±380)**	Europe
• • • • •	•	•	•	**null mutation U152**		**3525±460 ybp**	Europe
• • • • •	•	•	•	**R1b1a2a1a1b3c**	L2/S139	**4025±410 ybp**	Europe
• • • • •	•	•	•	**R1b1a2a1a1b3c1•**	L20	**3650±400 ybp**	Europe
• • • • •	•	•	•	**R1b1a2a1a1b3d**	L4/S178	**1275±290 ybp**	Europe
• • • • •	•	•	•	**R1b1a2a1a1b4**	L21/M529/S145	**3750±380 ybp**	Europe
• • • • •	•	•	•	**null mutation L21**		**3025±460 ybp**	Europe
• • • • •	•	•	•	**null mutation L21**		**1500±325 ybp**	Europe
• • • • •	•	•	•	**Unidentified clade**		**1650±175 ybp**	Europe
• • • • •	•	•	•	**R1b1a2a1a1b4b**	M222	**1450±160 ybp**	Europe
• • • • •	•	•	•	**R1b1a2a1a1b4e**	L144	**~ 4000 ybp**	Europe
• • • • •	•	•	•	**R1b1a2a1a1b4f**	L159.2	**1775±200 ybp**	Europe
• • • • •	•	•	•	**R1b1a2a1a1b4g**	L193	**1275±170 ybp**	Europe
• • • • •	•	•	•	**R1b1a2a1a1b4h**	L226	**1500±170 ybp**	Europe
• • • • •	•	•	•	**R1b1a2a1a1b4i**	P314.2	**2225±300 ybp**	Europe
• • • • •	•	•	•	**R1b1a2a1a1b5**	L176.2/S179.2	**3675±560 ybp**	Europe
• • • • •	•	•	•	**R1b1a2a1a1b5a**	SRY2627	**3150±320 ybp**	Europe
• • • • •	•	•	•	**null mutation SRY**		**~ 2800 ybp**	Europe
• • • • •	•	•	•	**R1b1a2a1a1b5b**	L165/S68	**~ 3000 ybp**	Europe
• • • • •	•	•	•	**R1b1a2a1a1b6**	L238/S182	**~ 1000 ybp**	Europe

In this endeavor both terms, "haplogroup" and "subclade", are employed as near equivalents because all haplogroups are essentially subclades of other upstream haplogroups, and usage of one or the other of these terms is suggested by the context. This is shown in the following diagram (ISOGG-2012, a fragment, http://www.isogg.org/tree/ISOGG_HapgrpR.html) which is related to the most ancient subclades of R1b haplogroup:

|	
R1b	M343
R1b1	L278, M415, P25_1, P25_2, P25_3
R1b1*	-
R1b1a	P297, L320
R1b1a1	M73, M478
R1b1a2	L265, M269, M520, S3, S10
R1b1b	M335
R1b1c	V88
R1b1c*	-
R1b1c1	M18
R1b1c4	V69

A Bird's-Eye View at R1b Haplotypes, and the Most Ancient R1b Populations

Figure 1 presenting an overview of an R1b haplotype tree includes 338 haplotypes in a short 8 marker format. The purpose of the presentation is to show the complex pattern of the R1b haplogroup and identify the most ancient branch on the tree.

The haplotype tree as shown next is arranged by a computer program which combines branches based on the similarity of their alleles in the respective markers (or loci) in Y chromosome, and dynamics of their alleles (Klyosov, 2009c and references therein). Haplotypes identical to each other and prevailing in the dataset are sitting at the top of the tree; in relatively recent datasets they typically represent the "base", or "ancestral" haplotype, from which all other haplotype are derived through mutations. Branches which are located close to the "trunk" of the tree contain not many mutations from the base haplotype, hence, they are relatively "young". The most ancient branches are those which shoot away from the trunk since they contain the most mutations in their haplotypes.

Figure 1. The 8-marker 338 haplotype tree for R1b1 haplotypes, composed from data listed in (Zhong et al., 2010).
The tree includes haplotypes of subclade R1b* (8 haplotypes), R1b1* (2 haplotypes), R1b1c (2 haplotypes), R1b (140 haplotypes), R1b1a1 (12 haplotypes), R1b2b (21 haplotypes), R1b1a2 (153 haplotypes). Regarding ethnicity and/or location, the haplotypes belong to 40 Western Europeans, 17 Iberians, 84 Turks, 48 Albanians, Macedonians and Romanians, 40 Italians, 30 Cretans, 13 Greeks and Macedonians, 9 Romanians from a different region, 13

Pakistanis, 4 Indians, 2 Japanese, 2 Mongolians (one of them, No. 176, belongs to R1b1b1-M73 subclade and stands out in the lower part of the tree), 2 Tibetans, 1 Kyrgyz from Xinjiang, 7 Han Chinese, 16 Uigur (12 from the Xinjiang Province, six of them belong to R1b* subclade, three R1b1b1 and three R1b1b2); four Chinese (three R1b2b and one R1b1b2), 2 Chinese of the Hui tribe *(mostly Uigurs, since Chinese use the ethnonym Hui = Uigur to denote generic Moslems)*, and 6 Chinese of other ethnicities (Hazak R1b1b1, Tu R1b2b, and four Naxi R1b2b)
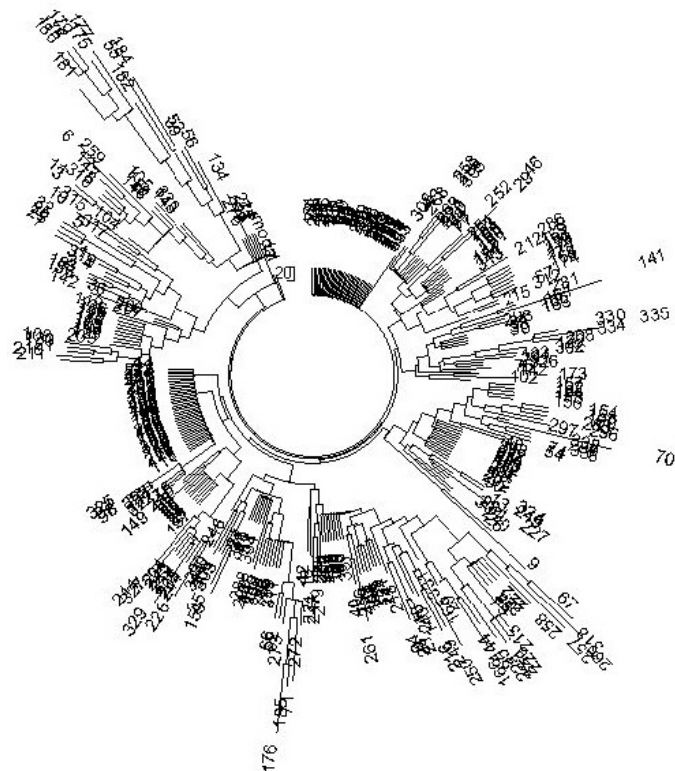
In **Figure 1** that rather recent and prevailing R1b branch has its base haplotype

13 24 14 11 X X X 12 X 13 13 29.

Thirty-five of those identical haplotypes sit on the top of the tree. Here X stand for missing alleles in the haplotype presented in the standard twelve marker FTDNA format (for definitions see **Materials and Methods**). This base haplotype belongs to the most populous European base haplotypes of the subclade R-M269, which in turn includes subclade R-P312 and its many downstream subclades. For example, the base 67 marker haplotype of P312 subclade is as follows (Klyosov, 2011b):

13 24 14 11 11 14 12 **12** 12 **13 13 29** - 17 9 10 11 11 25 15 19 29 15 15 17 17 - 11 11 19 23 15 15 18 17 36 38 12 12 - 11 9 15 16 8 10 10 8 10 10 12 23 23 16 10 12 12 15 8 12 22 20 13 12 11 13 11 11 12 12 (**P312**)

The alleles identical to those of the apparent "base" haplotype in the tree in **Figure 1** are marked in bold. Indeed, the subclade R1b1a2-M269 takes the largest part of the tree in **Figure 1**, since"R1b" as indicated in the legend is largely R-M269 as well.

The analysis of the tree shows that many of the Asian R1b haplotypes, particularly all the Indian, most of Pakistani, a quarter of Uigurs, one Japanese, one Tibetan, and two Chinese Hans are not the "indigenous" haplotypes but those which are closely associated with European haplotypes and share with them the same branches (Klyosov, 2010a). In other words, they are "returns" to the regions with Europeans, or having the European R1b origin. They are indistinguishable from the European R1b haplotypes.

Some of haplotypes, however, form separate branches. The most remarkable branch is most remote from the trunk, hence, the most ancient, is located in the upper left part of the tree. It contains 12 haplotypes which all belong to R1b1a1-M73 and R1b-M343 subclades, and provided by Uigurs and the close tribes of the Naxi, Han and Tu (Zhong et al., 2010). All the 12 haplotypes are derived from the base haplotype

13 24 15 11/10 X X X 12 X 14/13 12 30,

and contain collectively 65 mutations from it. Using the "linear" formula, we obtain 65/12/.013 = 417 → 619 conditional generations (the arrow here is a correction for back mutations), or 15,475±2500 years from their common ancestor (the calculation is explained in the **Materials and Methods** section). This date is in a fair agreement with 16,000±1400 years from a common ancestor of R1b haplotypes (Klyosov, 2008a, 2009d) and will be additionally confirmed later in this paper. It also indicates that the R1b haplogroup arose in Central Asia, and, apparently, in the Altai region in South Siberia, where their upstream haplogroups NOP → NO + P; P → R + Q; R → R1 + R2; R1 → R1a + R1b all migrated and lived there during the time period between 52 and 20-15 thousand years before present (Klyosov & Rozhanskii, 2012b).

Siberian, Bashkir, and Central Asian R1b-M73 Haplotypes

Although the M73 subclade is positioned a few steps from the top of the above diagram, it contains the most distant available haplotypes from the

Europeans, both geographically and by their mutations. **Figures 2** and **3** show two M73 haplotype trees composed from two different datasets; one was obtained directly from the researchers and published earlier (Klyosov, 2008a), the second was collected from the R1b1b1 Haplogroup Project (see **Figure 3** and references). Both of them have principally the same shape, with two primary branches in **Figure 2** and three in **Figure 3**.

Figure 2. The 25-marker 14 haplotype tree for R1b1-M73 haplotypes in Central Asia (Uzbeks, Tajiks, Tuva, Uigurs, Kazakhs) (Klyosov, 2008a)
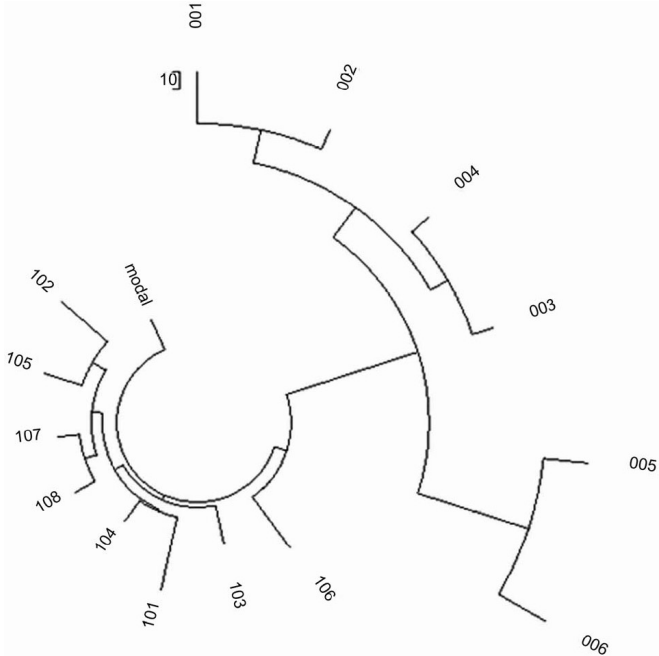
Figure 3. The 25-marker 25 haplotype tree for R1b1a1-M73 haplotypes:
the right-hand side branch — five Russians, three Kazakhs, one Chinese; the lower branch — four Uzbeks, two Tajiks, one Turk, one from Luxembourg; the left-hand branch — two English, two Italians, one French and one Spanish.
The data are taken from the FTDNA Project, in which many haplotypes are available in the 67 marker format:
http://www.familytreedna.com/public/R1b1b1/default.aspx?section=yresults

The tight 8-haplotype branch close to the "trunk" of the tree in **Figure 2** and the 9-haplotype right-hand branch in **Figure 3** both have the same base haplotype as follows:

13 19 14 11 13 13 12 12 13 14 13 30 - 17 9 9 11 11 23 15 19 33 12 15 15 16 - 10 10 19 25 15 16 16 17 30 36 12 10 - 11 8 16 16 8 10 10 8 11 10 12 23 23 16 10 12 12 16 8 12 24 21 13 12 11 13 11 11 12 11 (**M73, branch**)

The remarkable feature of this base haplotype (and every haplotype in the both branches) is the second allele, DYS390 = 19. In the 25 marker format, the branches contain 15 and 17 mutations respectively, which gives the same 1075±280 years to the common ancestor in both instances.

It should be mentioned here that a similar series of ten R-M73 haplotypes (with DYS 390 = 19) were obtained from the Bashkir population (Myres et al.,

2010) near the border between Europe and Asia, on both sides of the Ural Mountains. Their base haplotype was

13 19 14 10 X X X 12 13 14 13 29 - 10

(the last marker is DYS461), and all hundred alleles in the ten haplotypes contained only three mutations; this gives 3/10/.018 = 17 generations, or 425±250 years to their common ancestor. However, as it will be shown below, a common ancestor of these and other R-M73 haplotypes lived more that 10 thousand years ago. What we see in these haplotypes are essentially "tips of the iceberg"; they are tremendously distant from each other and must be very remote from their common ancestor in time. It seems that the "base" of the proverbial iceberg vanished without leaving their ancient haplotypes behind. In other words, we see here a series of present-day haplotypes which either arrived from elsewhere a few centuries earlier, or went through a severe bottleneck(s) of the population.

The Bashkir's (similar or slightly mutated) R-M73 haplotypes were identified among Tatars and Mari populations, in the adjacent regions (ibid.). In the same study, the R-M73 were found also among Kabardino-Balkars, Mengrels, Turks, with all speaking Turkic or Uralic languages, which belong to the Altai language family. Another name for Mari people is Cheremis, meaning "people from the East" in a neighboring Komi language. Their base haplotype

13 22 14 11 X X X 12 13 14 13 32 - 10

differs by 7 mutations (!) in ten markers from the Bashkir base L23 haplotype, which places their common ancestor by 9700 ybp (Klyosov, 2011a), and if haplotypes with DYS390 = 19 are included, their common ancestor moves deeper in time, to at least 16,000 years before present.

Revisiting the tree in **Figure 3**, the lower branch in the tree has the following base haplotype (when extended to 67 markers)

13 22 14 11 13 17 12 12 12 13 13 30 - 15 9 9 11 11 23 15 20 30 12 15 15 17 - 11 10 19 24 15 14 17 17 35 39 12 10 - 11 8 16 16 8 10 10 8 10 10 12 23 23 15 10 12 12 18 9 12 25 20 13 12 11 13 11 11 12 12 (**M73, branch**)

and its common ancestor lived 3525±600 ybp. It corresponds to the right branch in **Figure 2**.

The left-hand side branch in **Figure 3** has the following base haplotype:

12 25 14 11 13 14 12 12 12 14 13 29 - 16 9 10 11 11 22 15 20 32 12 15 16 17 - 11 10 19 25 15 17 17 18 34 36 14 10 - 11 8 16 16 8 11 10 8 11 10 12 22 23 17 10 12 12 15 8 13 22 20 14 12 11 14 11 11 12 12 (**M73, branch**)

and its common ancestor lived 3050±600 ybp.

All three branches are rather "young", however, they all are very distant from each other. Their pair-wise mutational differences are 38, 38 and 42, which translates into the separation time between their common ancestors as 11,500 - 13,200 years of the mutational evolution in their haplotypes (we call it the "lateral time" between two base haplotypes). All three M73 base haplotypes differ from each other collectively by 62 mutations, which place their common ancestor at ~7750 ybp. However, their mutational difference is even more pronounced when compared with the base haplotype of R1b1a2-P312 subclade, ancestral to many European subclades of R1b1a2 haplotypes. For the above M73 base haplotypes it amounts to 48, 35 and 42 mutations, or up to 16,000 years of the mutational difference between the two base haplotypes. This places their common ancestor at 10,600 ybp.

Another series of 26 of 12 marker M73 haplotypes was listed in (Malyarchuk et al., 2011), and were collected among the Siberian populations: Shors, Teleuts, Kalmyks, Khakass, Tuva, and Altaians. Those haplotypes split into two branches. The first one fits exactly to the first 67 marker base haplotypes with their available 12 markers; the second deviates from the base haplotype of the respective right-hand branch in **Figure 2** in two alleles (marked):

13 19 14 11 13 13 X X 13 14 13 30 - 15 10
13 22 14 11 13 **16** X X **13** 13 13 30 - 15 10

The first 12 marker base haplotype is a mixed Siberian-Ural-Caucasian population (if to consider their current bearers—Bashkirs, Kalmyks, Tuva, Tatar, Kabardin, Russian, and Altaian). All 11 haplotypes of this branch have collectively 20 mutations, which gives 20/11/.024 = 76 → 83 conditional generations, or 2075±510 years to a common ancestor. The logarithmic method gives [ln(11/2)]/.024 = 71 → 77 generations, or 1925 years, which is nearly the same (2 haplotypes are identical in their base haplotypes among 11 haplotypes in the dataset. .024 is the mutation rate constant for the 12 marker haplotypes). The same but in 67 markers, Middle Asian M73 haplotypes showed 1075±280 years to the common ancestor. Thus, the Siberia-Ural-Caucasian dataset is somewhat "older" compared to the Middle Asian dataset.

The second 12 marker base haplotype was largely from Siberian populations (Shors, Teleuts, Khakasses), plus a singular Mari, Turk and Tatar each. All 15 haplotypes of this branch had 22 mutations, which gives 22/15/.024 = 61 → 65 conditional generations; or 1625±380 years to a common ancestor. The logarithmic method gives [ln(15/5)]/.024 = 46 → 48 generations, or 1200±580 years, which is the same within the margin of error. The 67 marker base haplotype, however showed 3525±600 years to the common ancestor. Hence, now the Middle-Asian M73 branch is "older" compared to the Siberian. Overall, the Siberian and Middle Asian haplotypes were in about the same estimate of "age".

The two 12 marker haplotypes differ by 8 mutations, which sets their two common ancestors apart by 8/.024 = 333 → 492 generations, or 12,300 years. It gives (12,300+2075+1625)/2 = 7985 ybp to their common ancestor. One can see that the data obtained with the 67 and 12 marker haplotypes of the M73 subclade from Siberia and Middle Asia are fairly reproducible, and point to a common ancestor of Central Asian M73 haplotypes nearly 8000 ybp. It should be kept in mind that these data are related to currently living descendants of those common ancestors. We also need to be mindful that the 67 marker M73 Central Asian haplotype examples a distance up to 48 mutations from the European R1b1a2 haplotypes, which is up to 16,000 years of the mutational difference between the two base haplotypes. This places their (R-M73 in Central Asia and R-M269 in Europe) common ancestor at 10,600 years before present.

90

R1b1-L278/R1b1* Subclade Haplotypes

The most ancient subclades of R1b accounted by the Project R1b1 (xP297); xP297 does not include subclades M73 and M269, discussed in the

preceding section of this article. All 68 haplotypes of the Project, available in the 67 marker format, are shown in the tree in **Figure 4**. The tree includes haplotypes of subclade R1b1*, a paragroup of R1b1-L278 (see the R1b diagram shown above), V88, its down-stream subclade V69, and a series of haplotypes R1b1a2-M269 which apparently are in the Project by mistake (however, in the tree it is useful as a reference branch).

A detailed analysis of the tree was performed in (Klyosov, 2011a). Haplotypes of the R1b1* paragroup are represented in the Project by either Ashkenazi Jews (from Hungary, Russia, Belarus) with their very recent common ancestors a few centuries ago, or by very assorted geography — from Italy, Puerto-Rico, Germany, Armenia, Turkey.

R1b1c4-V88												R1b1c4-V69	R1b1a2	R1b1a1-M73
1-11	12	13	14	15	16	17-19	20-22	24-39	40	41	42-47			
Ashk.	Ind.	Arm.	Uzb.	Iraq	Bahr.	Germ.	Ashk.	Sefard.	SauAr	Sud.	-		Europe	Kaz./Uig./Uzb./Taj/Tuv.

Formally, a triple branch on the right of **Figure 4** (haplotypes 1 - 10) has a common ancestor who lived about 5800 ybp. It does not carry any certain historical meaning judging from their composition except, the time span corresponds to that of R-L23 subclade in the Caucasus and Anatolia (Klyosov, 2010d); R-L23 was certainly not alone there. Those assorted haplotypes could have been brought by their ancestors from Central Asia along with R-M269 followed by 67 marker haplotype tree of 68 haplotypes of subclades R1b1* (1 - 17), R1b1c-V88 (18 - 41), R1b1c4-V88-V69 (42 - 47), R1b1a2 (48 - 68). The tree was composed from data of the FTDNA Project http://www.familytreedna.com/public/R1b1Asterisk/default.aspx?section=yresults.

The subclade R1b1-L278/R1b1* (1 - 17) sheds light on one of the most interesting and puzzling pages of the European history. The Hebrew Ashkenaz (pl. Ashkenazim) referred to the Scythians Askuzes/Ishkuzes, the name which renders the Türkic As-kiji (As-people) or As-guzes (As-tribe), a transparent designation of the Türkic tribe of Ases (Ch. Yuezhi/Yueji 月支, possible compound of Hui/Yu + zhou 肉; the Ch. appellation Sui/Hui/Yui/Yu denotes Uigurs and was suggested to denote the Usun/Wusun ancestors of the modern Uisyns). However, in Biblical Hebrew, Ashkenaz is a generic name for the Türkic tribes independent of the tribal affiliation, and as such it was applied to the Khazars.

Life in the Khazar Kaganate was far from carefree, and religious turmoil was not any better than the political, ethnic, or military turmoils. Initially Tengrian, the population has divided into interspersed religious fractions, with the descendents of the converts to Judaism finding themselves alien and defenseless after the fall of the Kaganate and the start of the ethnic-religious persecution. The Jewish Türks largely dispersed, joining various established Jewish communities, and establishing a category called in Hebrew Ashkenazim across the Medieval Europe. Fortunately for them, Europe was divided into uncounted number of independent and quasi-independent principalities, and some of them welcomed refugees as monetary and intellectual capital. Naturally, the Ashkenazim boys carried the Y-DNA of their ultimately Türkic forefathers, but their closer forefathers came from a hodgepodge of the Khazar Kaganate Türkic tribes that spun off millenniums earlier, and had millenniums of independent development, admixtures, and linguistic and historical evolution.

Khazars were an offshoot of Bulgars, possibly sharing common genetics. The Khazars, like Bulgars, were identified with the ancient tribe of Ases. On the genetic tree, Khazars and Bulgars would likely constitute few distinct paired branches. A separate main branch of Khazars would be that of the Suvars/Savirs/Severyans, a second-largest Kaganate's majority, who had their own history and likely distinct genetic profile. The specific Bulgar clades should have survived among the the Türkic ethnicities in the Danube Bulgaria; the specific Savir clades should have a demographical peak in the vicinity of the Desna river, from Chernigiv to Kursk, among the Ukrainian and Russian populations. These subclades should show up on the exclusively Ashkenazi DNA tree, with allocation between Bulgar and Savir branches in accordance with the Danube Bulgaria/Chernigiv results.

On a deeper level, the initial appearance of the Bulgars and Huns in the Eastern Europe were likely unrelated events, the Huns migrating from the north, from the Aral and upper Itil/Volga area, and Bulgars migrating from the south, from the Bactria/Tocharistan, or the modern eastern Turkmenistan and southern Uzbekistan. The Hunno-Bulgars they became already in the Eastern Europe. A presence of the identical subclades in the southern Uzbekistan and the dispersal east and west of it would help to identify the Bulgar branch, and thus help to identify the Savir branch. The Savir branch should have elevated level of remnants in the vicinity of Dagestan. The number 5800 ybp is quite significant in that respect, separating the Bulgar and Savir branches 2000 years prior to the Savirs/Subars' showing up in the Babylonian records in 23rd c. BC. 28 centuries later, when Bulgars and Savirs reunited, they spoke languages of the same Ogur family, apparently still mutually comprehensible.

Not all Ashkenazim originated as Khazar subjects. Some may have quite different origin, like those that joined the Jewish communities from other Türkic lands (N.-E. Europe, Balkans, etc.), later newcomers from the Kipchak Khanate, Persian Jews via Caucasus, and the like, or may have gained their nickname by their association with the Ashkenazi communities..

Figure 4. L23 haplotypes, eventually dispersed across Europe
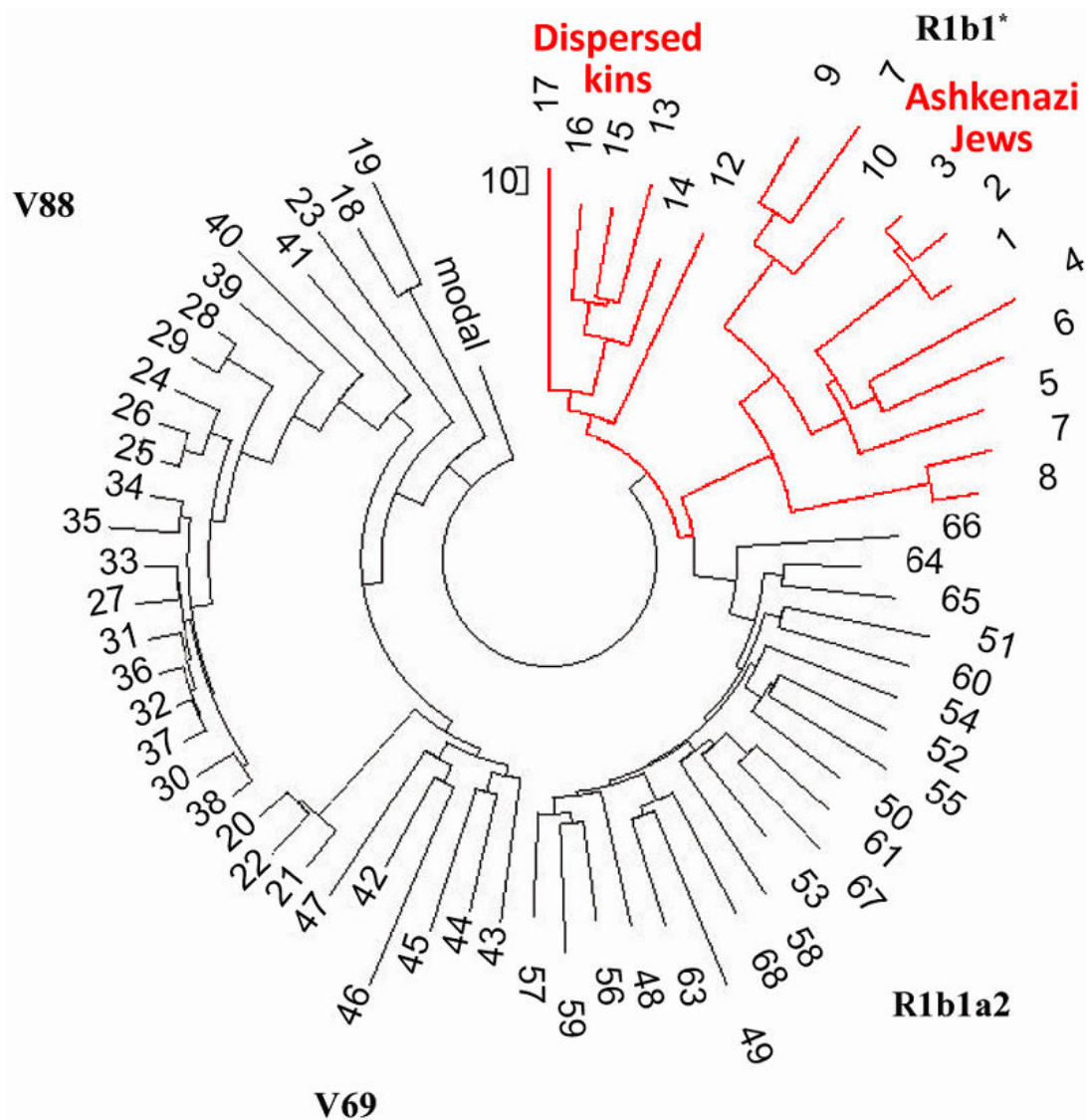
The next small branch of six haplotypes (12 - 17) includes those from Central Asia (Uzbek), Armenia, Middle East (Iraq and Bahrain), India, and Germany (the last being relatively close to the Uzbek haplotype, within 4900 years of the mutational evolution). The Indian haplotype has DYS390 = 18, which brings it closer to the Central Asian haplotypes with DYS390 = 19, considered above. Others have DYS390 = 24 or 25. Their mutational deviations (without inclusion of the very distant haplotype) from base R1b1* haplotype

13 24 15 10 12 13/14 12 13 11/12 13 13 29 - 14 9 9 11 11 26 14/15 18 30 13/14 14 14/15 16 - 11 11 18 23 15 16 18 16 36 36 13 11 - 11 8 15 16 8 10 10 8 11 10 12 19 21 16 10 12 12 14/15 8 12 24 20/21 14 11 11 13 12 11 12 12 (**R1b1*, branch**)

in the first 25, 37, and 67 markers suggests 238, 217, and 235 generations from their common ancestor (without a correction for back mutations), that is 230±11 generations, a fairly reproducible value. The most reliable value, for the 67 marker format, gives ~8400 years to their common ancestor. However, the Indian haplotype differs from the above base haplotype by 62.5 mutations (!), with its DYS390 = 18 contributing only 6 mutations into the difference. This places the timespan to the R1b1* common ancestor (of the available haplotypes) at 16,600 years before present. It certainly points to Central Asia, since the European-Middle Eastern R1b1* haplotypes are only ~8400 years "old".

R1b1c-V88 Subclade in Europe and Africa

Haplotypes V88 occupy the left-hand side of the tree in **Figure 4**. Again, a number of their sub-branches are quite "young". For example, a part of the branch belongs to Ashkenazi Jews (Russia, Ukraine, Hungary, Germany, France) with a common ancestor of only 350 ybp. Haplotypes of England and Scotland have a common ancestor who lived 650 ybp. However, a common ancestor of the two sub-branches (Jews and the Isles) lived 6875 ybp, at the very bottom of the V88 subclade. Another series of V88 haplotypes identifies their route from Saudi Arabia (6225 ybp) via Spain to the Jewish community (1525 ybp, in the middle of the 1st millennium CE).

R1b1c4-V88

1-11   12   13   14   15   16   17-19 20-22 24-39 40   41
Ashk. Ind. Arm. Uzb. Iraq Bahr. Germ. Ashk. Sefard. SauAr Sud.

The base haplotype of the V88 branch in **Figure 4** is as follows:

13 24 16 10 13 14 12 12 12 14 13 29 - 16/17 9 10 11 12 26 14 19 29 12 12 15 15 - 11 12 21 23 15 15 18 19 33 35/36 12 11 - 12 8 15 16 8 10 10 8 9 11 12 22 23 15 11 12 12 15 8 12/13 23 20 13 12 11 13 11 12 12 12 (**V88**)

and its bearer lived 6575±700 ybp.

A rather extended series of 72 of 11 marker V88 haplotypes from Africa was provided in (Cruciani et al., 2010). They split into two rather distinct branches and an assorted series of haplotypes (Klyosov, 2010b). One branch of 37 haplotypes has the following base haplotype (in the format — 413a,b 460, 461, GATA A10, YCAIIa,b after the first standard 12 markers, some of them are missing):

13 X 15 11 X X X X 12 X X X - 23 21 11 10 13 21 23

All of them contain 111 mutations, which gives 111/37/.02 = 150 → 176 conditional generations, or 4400±610 years to their common ancestor. This African V88 base haplotype does not fit with either of the Eurasian V88 base haplotypes, however, it fits rather well to the R1b1c4-V69 base haplotype (see the next section), whose common ancestor lived 4300±600 ybp. The paper (Cruciani et al, 2010) suggested one more subclade named R1b1a4, with 29 haplotypes. Their base haplotype is identical with that shown above; all 29 haplotypes contain 76 mutations, and gives 76/29/.02 = 131 → 150 generations, or 3750±570 ybp, which is within the margin of error with the above date.
[91]

Another branch of V88 African haplotypes is recent with a common ancestor of 750±290 ybp. All 72 African haplotypes have the same base haplotype as shown above with 201 mutations from it, which gives 201/72/.020 = 140 → 163 generations, or 4075±500 ybp, within the margin of error of that for the main V88 branch; however, the first is more accurate since it was calculated for the distinct branch.

This date, 4300±600 ybp, fits well with that for the migration route of R1b bearers from the Middle East westward along the Mediterranean Sea coast to the Atlantic, which took place between 5500 and 4800 ybp (see below). It seems that the R1b-V88-V69 tribe may have split altering migration direction south, to Cameroon and Chad, where R1b-V88 bearers live today.

R1b1c4-V69 Subclade

This subclade is represented by six haplotypes at the bottom of the tree in **Figure 4** *(No 42-47)*, with the base haplotype

13 23 15 11 13 15 12 12 13 13 13 29 - 16/17 9 10 11 12 27 14 19 29 12 12 14 15 - 11 12 21 23 15 16 17 19 33 36/37 12 12 - 12 8 15 16 8 10 10 8 10 11 12 21 24 15 11 12 12 14 8 12/13 22 20 13 12 11 13 11 12 12 13 (**V69**)

All of them collectively have 106 mutations from the base haplotype, which gives 106/6/.12 = 147 → 172 generations, or 4300±600 years from a common ancestor. This is in agreement with the "age" of the upstream V88 subclade, of ~6575 ybp. The V69 base haplotype differs by 16 mutations from the base V88 haplotype (see above; some mutations are fractional), which sets apart their common ancestors by 16/.12 = 133 → 154 generations, or 3850 years, and places their common ancestor at (3850 + 6575 + 4300)/2 = 7400±800 ybp. This is within the margin of error with 6575±700 ybp for the V88 upstream subclade.

The analysis demonstrates that R-V88 is a rather "young" subclade in comparison with the entire R1b, and R-V69 is "younger". It may have arisen on the R1b general migration route to the Middle East, for example, north-east or east of the Caspian Sea, in the Western Iran.

R1b1a2-M269 Subclade

The last branch on the tree in **Figure 4** in the lower right-hand side *(No 48, 56-57, 59)* has the base haplotype

13 24 14 11 11 14 12 12 12 13 13 29 - 17 9 10 11 11 25 15 19 29 15 15 17 17 - 11 11 19 23 15 15 18 17 37 38 12 12 - 11 9 15 16 8 10 10 8 10 10 12 23 23 16 10 12 12 15 8 12 22 20 13 12 11 13 11 11 12 12 (**M269, a mix**)

It is exactly the base haplotype of subclades R-P312 and R-L21. They are typical and widespread European haplotypes. All 20 haplotypes of the branch contain 365 mutations, which gives 365/20/.12 = 152 → 179 generations, or 4475±505 years to their common ancestor. This, again, is a typical "age" for a mix of haplotypes of subclades R-P312 and R-L21 and their downstream subclades.

The M269 subclade along with its downstream subclades were analyzed in many publications (Klyosov, 2008a, 2009d, 2010c, 2010d, 2010e, 2011a, 2011b). In short, this subclade arose around 7000 ybp in Central Asia or in the eastern area of the East European Plain — its immediate downstream subclade R-L23 (arose ~6000 ybp) is widespread among the Bashkirs in the South Urals, North Kazakhstan and adjacent regions, and can be seen in Russia. It migrated south to the Caucasus and beyond, to Anatolia and the Middle East. Nearly all principal European and Middle Eastern subclades of the R1b haplogroup (currently more than 80 subclades) are derived from the R1b1a2-M269 subclade.

It should be mentioned here, that two brother subclades, R1a and R1b, have been migrating from Central Asia westward by two quite different routes. While R1a were moving along the southern route from the Altai region across the Himalayas, Hindustan, the Iranian plateau, Anatolia and the balance of Asia Minor to the Balkans (Klyosov & Rozhanskii, 2012), R1b were moving along the northern route, from the same region across South Urals, Middle Asia, North Kazakhstan, Middle Volga, the Caucasus, and then split between the southward and westward directions. This migration pattern also explains why R1a, but not R1b bearers were found in the Eastern Himalayas (Kang et al., 2011; see also Klyosov & Rozhanskii, 2012).

The 37 and 67 marker haplotype trees of the R1b1a2*-M269 subclade are shown in **Figures 5** and **6**, respectively. The trees contain two distinct branches, one rather complex and obviously "older", with the base haplotype

12 25 14 11 11 14 11 12 12 13 13 29 - 17 9 10 11 11 25 15 19 29 15 15 16 17 - 10 11 19 23 15 15 19 17 35 38 12 12 - 11 9 15 16 8 10 10 8 11 10 12 23 23 16 10 12 12 15 8 13 22 20 13 12 11 13 11 11 12 12 (**M269**)

and another, a tight, flat, "younger", and largely Jewish branch, with the base haplotype.

12 **24** 14 **10** 11 14 11 12 12 13 **14** 29 - 17 9 10 11 11 25 15 19 **30** 15 15 16 **16** - **11 10** 19 23 **17 16 17** 17 **38 39 13** 12 - 11 9 15 16 8 10 10 8 11 10 12 **21** 23 16 10 12 12 **14** 8 **12** 22 20 13 12 11 13 **10** 11 12 12 (**M269, Jewish**)

Figure 5. The 37 marker 25-haplotype tree for R1b1a2*-M269 haplotypes.
The older (~6225 ybp) branch on the left and the bottom present haplotypes from Turkey, Armenia, Italy, France, England. The younger (1000 ybp) branch in the upper right side represents Ukraine, Belarus, Poland, and Kazakhstan. http://www.familytreedna.com/public/ht35new/default.aspx
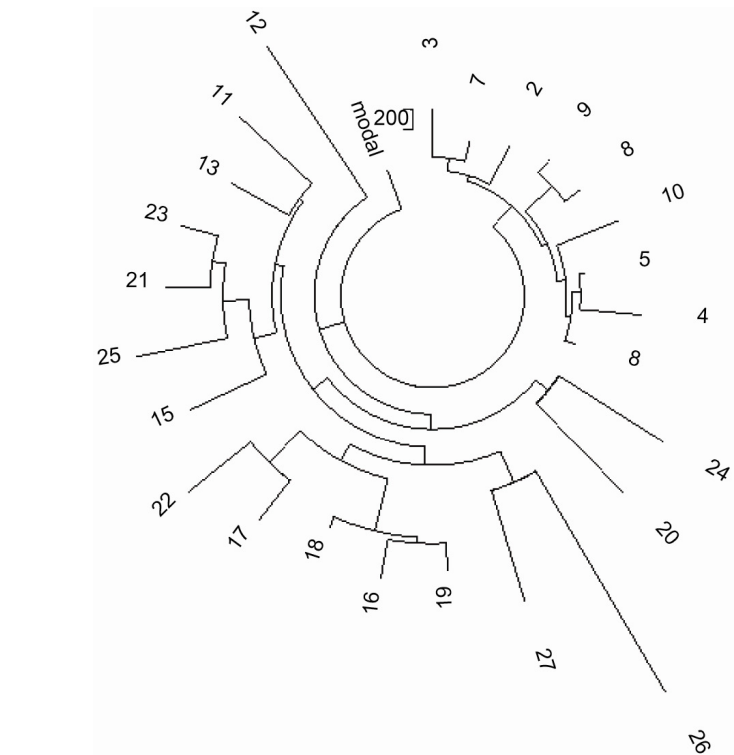
Figure 6. The 67 marker haplotype tree for R1b1a2*-M269 haplotypes,
the same dataset as shown in Figure 5, however, with fewer haplotypes available
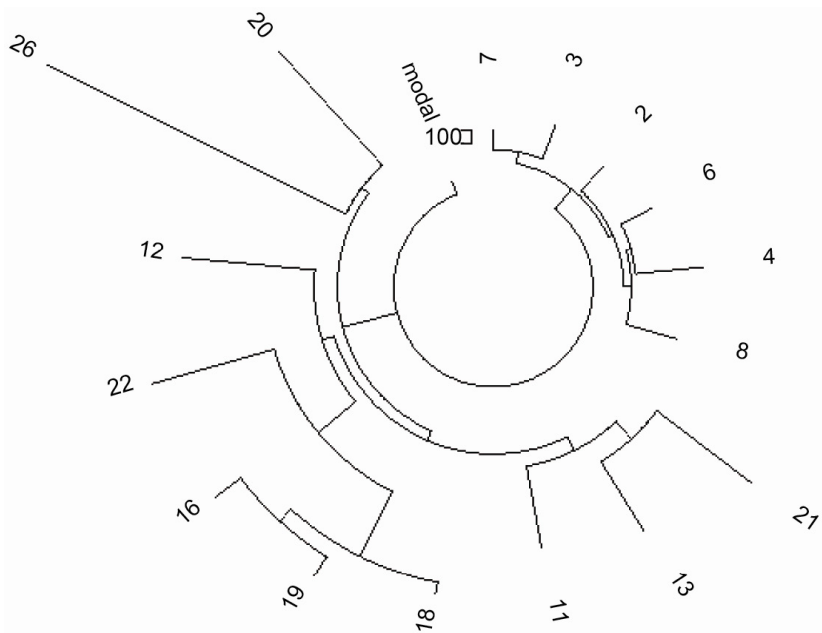
Mutation deviations between the two are shown in bold. The younger branch has 26 mutations in all six 67 marker haplotypes, which gives 26/6/.12 = 36 → 37 conditional generations, or 925±200 years to a common ancestor. The 37 marker branch contains 31 mutation in nine haplotypes, which gives 31/9/.09 = 38 → 40 generations, or 1000±200 years — essentially the same. This Jewish population *("younger", **Fig. 5**)* exemplified with 14 of 37 marker haplotypes was analyzed earlier (Klyosov, 2008b) and determined that their common ancestor lived 1100±250 ybp; the base haplotype is exactly as that of the "young" branch listed above.

> The bottleneck timing of the Jewish R1b1a2*-M269 haplotypes also coincides with the Ashkenazim dispersion after the fall of the Khazar Kaganate.

The older branch has a common ancestor who lived 6200±900 ybp. The two base haplotypes differ by 22 mutations which separate them by 22/.12 = 183 → 224 generations, or 5600 years, and a common ancestor of both branches lived 6400±900 ybp.

This description gives us a general picture of migration of the Arbins from Central Asia (~16,000 ybp) westward during the next 10 - 9 thousand years, to about 7000 - 6000 ybp.

Haplogroup R1b (mainly R1b1a2-L23) among Bashkirs, and in the Caucasus, Anatolia, Middle East

R-L23 apparently arose on the eastern side of the East European Plain, where Europe meets Asia, ~6200 ybp and migrated to the Caucasus and further South, to Anatolia and the Middle East. Another branch of L23 went westward, to Europe, approximately 4500 ybp.

The most eastern population with the prevailing R-L23 subclade is the Bashkirs, a Türkic-language people who live largely on both sides of the Ural Mountains and in North Kazakhstan. Frequency of R1b haplogroup reaches 84% in the Perm Bashkirs, 81% among Baimak Bashkirs, and lower figures in other Bashkir tribes (Lobov, 2009). 29 of 10 marker haplotypes of subclade R-L23 of the Bashkirs were published (Myres et al., 2010), and their base haplotype is

12 24 14 10 X X X 12 12 13 13 30 - 10.

This is a typical albeit slightly mutated L23 haplotype with its characteristic first allele DYS393 = 12. 26 haplotypes of those 29 were identical, as shown above, and the whole Bashkir L23 branch has a common ancestor who lived only 575±175 ybp. However, this base haplotype differs from the European R-L23 base haplotype

12 24 14 10 X X X 12 12 13 14 29 - 11

by three mutations, which sets these two base haplotypes apart by 3/.018 = 167 → 200 conditional generations — 5000 years, and places their common ancestor at 5500 ybp. Other Bashkir haplotypes belonged to R-M269 (one haplotype), R-M73 (10 haplotypes), and R-U152 (8 haplotypes). The last series of haplotypes were all identical to each other, and thus are derived from a very recent common ancestor, who certainly had a European origin [U152 arose in Europe 4125±450 ybp (Klyosov, 2011b)].

> Marija Gimbutas dated the Third Kurgan Wave, the intrusion of the N.Pontic "Pit Grave" Kurgans into East-Central Europe by c. 3000 BC ~ 5000 ybp

> **R1b haplotypes in Myres et al., 2010, Supplementary Table S3, for nations with strong Türkic presence**
> Without ethnical identification of the samples, the ethnical attribution is impossible (e.g. Turkey)
> For multi-ethnic nations, like Bashkirs, Kumyks, Turkey, Iran (Persia), the sampling is woefully inadequate

Bashkir			Turkey		
L23	14 12 13 17 24 10 13 12 12 10	N=21 44%	L23	14 12 13 16 24 11 13 12 12 10	N=6 8%
M73, M478	14 12 14 16 19 10 13 13 13 10 8	17%	L23	14 12 13 16 24 11 13 12 13 10 4	5%
U152	15 12 13 17 24 11 13 13 12 11 8	17%	M269	14 12 13 16 25 11 13 12 11 10 3	4%
L23	14 12 13 17 24 10 13 12 11 10 2	4%	L23	14 12 14 16 24 11 13 13 11 11 2	3%
L23	14 12 13 17 24 10 13 12 13 10 2	4%	L23	14 12 13 16 23 10 13 12 12 10 2	3%
L23	14 12 13 17 23 10 13 12 12 10 1	2%	L23	14 12 13 15 24 11 14 12 12 10 2	3%
L23	14 12 13 17 25 10 13 12 12 10 1	2%	L23	14 12 13 15 24 11 14 12 13 10 2	3%
L23	14 12 13 16 23 11 13 12 13 10 1	2%	L23	14 12 13 16 24 10 13 12 12 12 2	3%
L23	14 12 13 17 24 10 13 13 12 10 1	2%	L23	14 12 12 16 24 11 13 13 14 10 2	3%
L23	15 12 13 17 24 10 13 12 11 10 1	2%	L23	15 12 12 15 24 10 13 12 11 11 1	1%
M73, M478	14 12 12 16 19 10 13 13 13 10 1	2%	L23	13 12 13 16 25 11 13 12 12 10 1	1%
M73, M478	14 12 14 16 19 10 13 13 12 10 1	2%	L23	14 12 12 15 24 10 14 12 12 11 1	1%
Balkar			L23	14 12 13 16 24 11 13 13 12 11 1	1%
M73, M478	14 12 14 19 22 11 13 13 12 10	N=3 30%	L23	14 12 13 15 24 11 15 12 13 10 1	1%
M73, M478	14 12 14 19 23 11 13 13 13 10 2	20%	L23	14 13 14 16 23 10 13 12 11 10 1	1%
M73, M478	14 12 14 18 22 11 13 13 12 10 2	20%	L23	15 12 13 16 24 11 13 12 12 10 1	1%
L23	14 13 13 17 25 11 13 12 11 10 1	10%	L23	14 12 13 16 21 10 13 12 12 10 1	1%
M73, M478	14 12 15 19 22 11 13 13 13 10 1	10%	L23	15 12 13 16 21 10 13 12 12 10 1	1%
M73, M478	14 12 14 18 22 11 13 13 13 10 1	10%	L23	14 12 13 17 25 10 13 12 12 10 1	1%
M73, M478	14 14 13 19 22 11 13 13 13 10 1	10%	L23	15 12 12 15 23 10 14 12 12 10 1	1%
Kumyk			L23	14 12 12 15 24 10 14 12 12 10 1	1%
L23	14 12 14 16 24 11 13 12 14 10	N=1 33%	L23	14 12 13 15 24 10 15 12 12 10 1	1%
L23	14 12 13 15 24 11 15 12 12 10 1	33%	L23	14 12 13 16 24 10 13 12 13 11 1	1%
L23	14 12 13 16 24 11 13 12 12 10 1	33%	L23	14 12 13 16 24 11 13 12 11 10 1	1%
Kaitak (i.e. Kayi-tag, Kayitag), R1b1b2, (Balanovsky et al., 2012)			L23	13 12 13 18 25 11 13 12 11 10 1	1%
M269	13 28 23 16 14 11 14 17 15 13 19 13 10 13 12 13 23 12 12		L23	14 12 13 15 24 10 14 12 13 10 1	1%

M269 13 29 24 17 14 11 14 16 15 12 19 12 11 13 12 12 23 12 12

Tatar

M73, M478 14 12 14 16 19 10 13 13 12 10 N=1 50%

M73, M478 14 12 13 17 22 11 13 13 13 10 1   50%

Dargin

L23 14 12 14 17 23 11 12 12 12 10 N=1 50%

L23 15 12 13 16 23 11 13 12 13 11 1   50%

Avar

L23 14 12 12 16 24 11 13 12 13 10 N=1 100%

Megrel

M73, M478 14 12 14 19 21 11 13 12 13 10 N=1 50%

L23   14 12 14 17 24 11 13 12 12 10 1   50%

N.Ossetian (Digor or Ardon)

L23 14 12 12 18 24 11 13 12 12 11 N=1 50%

L23 14 12 12 17 24 11 13 12 12 11 1   50%

Tabassaran

L23   14 13 13 16 23 11 13 12 12 10 N=2 33%

L23   14 12 12 16 23 10 13 12 12 11 1   17%

L23   14 12 13 16 23 11 13 12 14 10 1   17%

L23   14 12 13 16 23 10 13 12 13 10 1   17%

M269 15 12 13 16 25 10 13 12 11 10 1   17%

Iran

M269 14 12 14 16 24 11 12 13 12 10 N=1 100%

Mari

M73, M478 14 12 13 17 23 10 13 13 14 10 N=1 50%

M73, M478 14 12 14 16 19 10 13 13 13 10 1   50%

Pakistan

L23 14 13 13 16 24 10 13 12 12 10 N=1 20%

L23 14 12 13 17 24 11 13 14 12 11 1   20%

L23 14 12 13 15 25 11 14 12 13 9 1   20%

L23 14 12 13 16 24 11 13 12 12 10 1   20%

L23 14 12 15 15 24 11 14 12 12 11 1   20%

	Haplotype	%
L23	14 12 13 16 24 11 13 12 13 11 1	1%
L23	14 12 13 16 25 11 13 12 12 10 1	1%
L23	14 12 13 16 24 10 13 12 13 10 1	1%
L23	13 12 13 17 24 11 13 12 14 10 1	1%
L23	14 12 13 16 24 10 14 12 13 10 1	1%
L23	15 12 13 16 23 11 13 12 12 10 1	1%
L23	14 12 13 16 24 12 13 13 12 10 1	1%
L23	14 12 13 15 23 10 14 12 12 10 1	1%
L23	14 12 13 16 24 10 12 12 12 11 1	1%
L23	14 12 12 16 25 11 13 12 13 10 1	1%
L23	14 12 13 16 24 11 13 13 12 10 1	1%
L23	15 12 14 15 25 10 12 12 11 10 1	1%
L23	14 12 13 15 23 11 11 12 12 10 1	1%
L23	14 12 14 17 25 10 13 12 12 10 1	1%
L23	14 12 13 15 24 10 14 12 11 11 1	1%
L23	14 12 13 16 24 10 13 12 12 10 1	1%
L23	14 12 13 16 23 11 13 12 11 11 1	1%
L23	14 12 14 17 23 10 13 12 12 11 1	1%
L23	13 12 13 17 25 11 13 12 12 10 1	1%
M269	14 12 13 16 24 11 13 12 12 10 1	1%
M269	14 12 13 16 25 10 13 12 13 10 1	1%
M269	14 12 13 18 24 10 13 12 11 10 1	1%
M269	14 12 13 16 24 11 13 13 11 10 1	1%
M269	14 12 14 16 24 11 13 12 11 10 1	1%
M269	14 12 14 16 24 11 12 12 11 10 1	1%
M269	14 12 12 15 24 10 14 12 12 11 1	1%
M412	14 12 13 16 24 11 13 13 12 12 1	1%
M73, M478	14 12 13 17 23 11 13 13 14 10 1	1%
S116	14 12 13 16 24 11 13 13 12 12 1	1%
S116	15 12 12 17 24 10 13 13 13 11 1	1%
U106	14 12 13 15 23 11 13 13 12 11 1	1%
U106	14 12 13 16 25 11 14 13 11 12 1	1%
U152	14 12 12 16 24 11 13 14 11 10 1	1%
U152	14 12 13 16 24 11 13 14 12 12 1	1%
U152	14 12 13 16 24 11 15 13 12 12 1	1%

We therefore see today's reflection of ancient migrations of the Arbins westward from Central Asia, apparently from the South Siberian region, across the South Urals and further to the East European Plain and then the Caucasus.

Almost all R1b1a2 haplotypes in the Caucasus region belong to the subclade L23 (with its characteristic DYS393 = 12). In a recent paper (Balanovsky et al., 2012) 90 Caucasian haplotypes of R1b haplogroup were listed, and with exception of five R1b* haplotypes and a relatively "young" Abkhazian branch (**Figure 7**) 79 of 81 haplotypes (97.5%) in the dataset were of the L23 subclade (Note: the cited paper did not consider haplotype trees nor has analyzed the haplotypes in the manner presented here).

The same pattern is observed with Armenian R1b haplotypes, and with most of Anatolian R1b haplotypes (Klyosov, 2010c, 2011c). The 67 marker base R-L23 haplotype, obtained from an extended haplotype dataset from the world over (tree **Figure 8**) is as follows:

12 24 14 11 11 14 12 12 12 13 13 29 - 16 9 10 11 11 25 **15 19** 29 15 15 16 17 - 11 **11** 19 23 **16** 15 17 17 36 37 12 **12** - 11 9 15 16 8 10 10 8 10 11 12 23 23 16 10 12 12 15 8 12 22 20 13 12 11 13 11 11 12 12 (**L23**)

The L23 base short Caucasian haplotype in **Figure 7** fits exactly to the above haplotype (the matching alleles are marked in bold). A common ancestor of the L23 subclade lived ~6200 ybp (Klyosov, 2010d, 2011a). The 81 Caucasian L23 haplotypes containing 425 mutations from their base haplotype give 425/ 81/.035 = 150 → 176 conditional generations, or 4400±490 years to their common ancestor. The "younger" date (compared with the "age" of L23 of about 6200 ybp) can be explained by a detailed consideration of an extended series of 107 of R-L23 haplotypes listed in the FTDNA Project (see the legend in **Figure 8**). The tree in **Figure 8** splits into two parts. On the left are 38 haplotypes, with the base

12 24 14 11 12 15 12 12 12 13 13 29 - 16 9 10 11 11 25 15 19 30 15 15 16 18 - 11 11 19 23 16 16 18 17 36 37 12 12 - 11 9 15 16 8 10 10 8 11 11 12 23 23 15 10 12 12 16 8 12 22 20 13 12 11 13 11 11 12 12 (**L23, branch**)

93

Figure 7. The 19-marker 90 haplotype tree for R1b haplotypes in the Caucasus (Lezghins, Ossets, Avars, Ankhaz, Circassians, Chechens, Dargins, Kaitaks)
Haplotypes 1 - 5 (above on the left) represent R1b* subclade,
and haplotypes 49, 57 - 59 (above on the left) represent a "young" Abkhazian branch. Haplotypes were listed in

(Balanovsky et al., 2012)
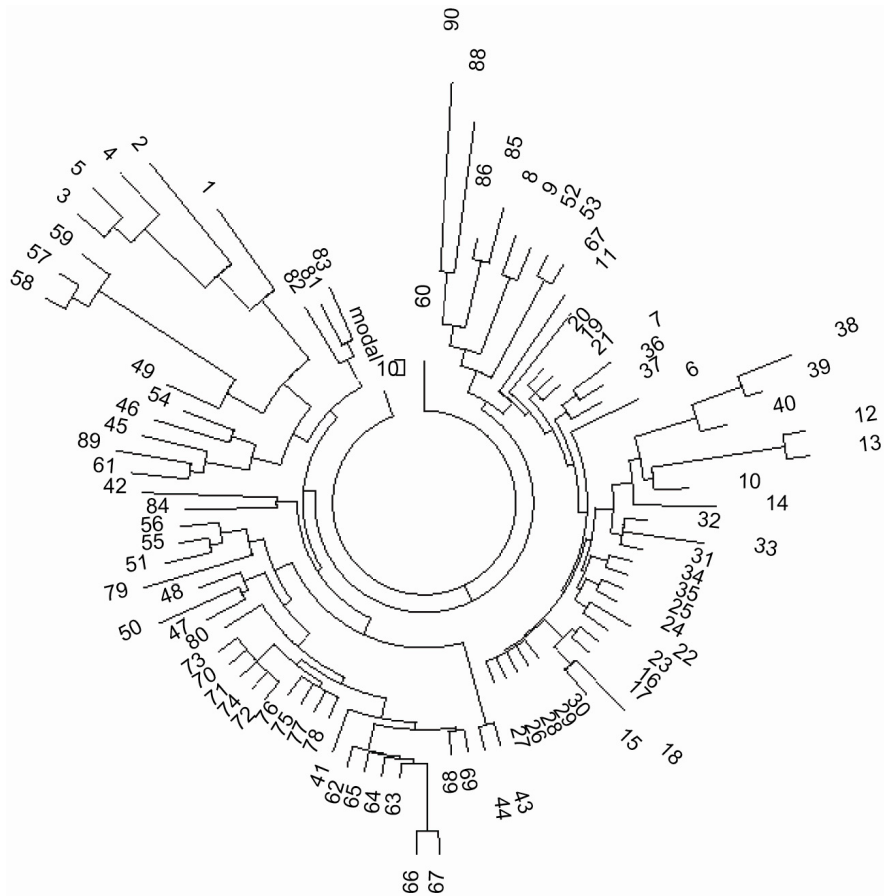
On the right are 69 haplotypes, with the base

12 24 14 11 **11 14** 12 12 12 13 13 29 - 16 9 10 11 11 25 15 19 **29** 15 15 16 **17** - 11 11 19 23 **15** 16 18 17 36 **38** 12 12 - 11 9 15 16 8 10 10 8 **10** 11 12 23 23 **16** 10 12 12 **15** 8 12 22 20 13 12 11 13 11 11 12 12 (**L23, branch**)

More than 70% of haplotypes of the Armenians and Turks from the dataset belong to the second and larger branch as well as all eight Iraqis and all five Iranians in the dataset. Presence of R-L23 of the Iranians might be a result of diffusion of the subclade from Anatolia eastward, or the migration of the Arbins might have been southward from the East European Plain east of the Caspian Sea.

Both branches descended from their ancestral R-L23 base haplotype, and are parted by 9 mutations (marked in bold). These 9 mutations are accumulated over $9/.12 = 75 \rightarrow 81$ conditional generations, or 2025 "lateral" years. The first branch split 4600±490 ybp; the second, 4200±440 ybp. Therefore, their common ancestor lived $(4600 + 4200 + 2025)/2 = 5400\pm800$ ybp. This fits within margin of error to the time when a common ancestor of the subclade R-L23 lived (~6200 ybp). Extended, 111 marker haplotypes available for the same dataset, and for the smaller branch the base haplotype is as follows:

12 24 14 11 **12 15** 12 12 12 13 13 29 - 16 9 10 11 11 25 15 19 **30** 15 15 16 **18** - 11 11 19 23 **16** 16 18 17 36 **37** 12 12 - 11 9 15 16 8 10 10 8 **11** 11 12 23 23 **15** 10 12 12 **16** 8 12 22 20 13 12 11 13 11 11 12 12 - **36** 15 9 **15** 12 **25** 26 19 12 11 13 12 10 9 12 12 10 11 **10** 30 12 13 24 13 10 10 **19** 15 19 **13** 24 **17** 12 15 24 12 23 18 10 14 17 9 11 11 (**L23, branch**)

It differs by 16 mutations from that of the larger branch (7 mutations added by the 68 - 111 extension), which separates the branches by $16/.198 = 82 \rightarrow 90$ generations, or 2250 years, and the R-L23 common ancestor lived $(4600 + 4200 + 2250)/2 = 5525\pm700$ years — similar to 5400±800 ybp, obtained above, and illustrates the consistency of calculations.

It seems that the Caucasian R-L23 haplotypes with their common ancestor of 4400±490 years belong to one of the branches in the tree in **Figure 8**. A much smaller Caucasian R1b dataset, analyzed earlier (Klyosov, 2008a) resulted with similar times — 4650±700 ybp, as the recent, larger dataset of short haplotypes (Balanovsky et al., 2012). The Caucasian R-L23 haplotypes may have experienced a population bottleneck around 5000 ybp.

120 of 17 marker Armenian haplotypes were published recently in (Herrera et al., 2011). A haplotype tree, composed from those haplotypes, is shown in **Figure 9**. It consists of three approximately equal (by number, or by "weight") branches. Five haplotypes of the M269 mini-branch *(which ones?)* are nearly equal to each other, evidencing only two mutations among their 85 alleles. Their common ancestor was 300±210 ybp.
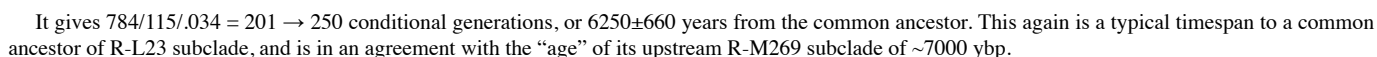
> In ca. 1600 Ottoman Empire and Safavid Persia divided Armenia, with the northwestern Armenia annexed by the Ottoman Empire; that appear to be reflected in the young M269 mini-branch of five haplotypes.

The balance of 115 haplotypes of L23 subclade, have 784 mutations from their base haplotype

12 24 14 11 11 14 X X 13 13 13 29 - 17 15 19 12 15 12 23

(the right-hand side of the haplotype corresponds to DYS 458, 437, 448, GATA H4, DYS 456, 438, 635).

94

Figure 8. The 67-marker tree composed with 107 haplotypes of the subclade R-L23
40% of the haplotypes belong to Armenians and Turks, followed by Iraqis and Italians (5% each).
The haplotypes were collected in the FTDNA Project http://www.familytreedna.com/public/ht35new/default.aspx

It gives 784/115/.034 = 201 → 250 conditional generations, or 6250±660 years from the common ancestor. This again is a typical timespan to a common ancestor of R-L23 subclade, and is in an agreement with the "age" of its upstream R-M269 subclade of ~7000 ybp.

The same base haplotype, as shown in the preceding paragraph, was found in a dataset of 238 Armenian six-marker R1b haplotypes published earlier (Weale et al., 2001) and analyzed in (Klyosov, 2008a). It included haplotypes from six regions of Armenia, Karabakh, Iran, and other areas of the Armenian diaspora. It can be presented as

12 24 14 11 X X X 12 X X 13 X.

An average "age" of the common ancestor of R1b haplotypes in all the six regions was 5750±1500 years, which is in line with other estimates for the R-L23 subclade.

The Balkan Phrygian origin of the Armenians is well-established in the Classical sources; Armenian is an agglutinative language; the Phrygian language was not mutually intelligible with the language of the Trojans and with the surrounding Anatolian languages. Ethnologically, we know Phrygian hat, identical with the Scythian hats and the traditional hats of the Türkic peoples; we also know the Phrygian shot, identical with the Scythian shot and the legendary shots of the Türkic riders; the Phrygian king Gordion was buried in a traditional Scythian and Türkic kurgan, surrounded with many other kurgans; that all points to the Kurgan expansion to the Balkans. Thus, the Kurgan R1b marker at Armenians is consistent with their origin in the Classical sources; however, the Armenians have Türkic admixtures from antiquity to the Late Middle Ages, and their R1b tree should have numerous branches originating during that period that parallel branches among numerous distinct Türkic peoples.

Considering other populations, containing a rather high share of R-L23 haplotypes, which in turn could be a tracer of ancient migration of the Arbins, one can mention Russia with their 37% of available R1b haplotypes having DYS393 = 12 (Roewer et al., 2008; Klyosov, 2010c and references therein) and the base haplotype

12 24 14 11 11 14 X X 12 13 13 29/30 - 16 15 19 (the last three alleles are of DYS 458, 437, 448),

the same as R-L23, shown above, and with a common ancestor of the Russian R1b of 6775±830 ybp (Klyosov, 2009d); Anatolia and Lebanon with their

12 24 14 11 X X X 12 13 13 29 and
12 24 14 **10** X X X 12 13 13 29 (mutated DYS391 11 → 10 is marked)

and with the respective common ancestors of 6000±820 ybp (Anatolia) and 5200±670 ybp (Lebanon) [Klyosov, 2010c; haplotypes were listed in Cinnioglu et al., 2004, and Zalloua et al., 2008].

Among Jewish R1b haplotypes (most were not typed for subclades when tested several years ago as the typing was not available), one branch remarkably resembled R-L23 haplotypes, and its 37 marker base haplotypes was as follows (Klyosov, 2008b):

12 24 14 10 11 14 12 12 12 13 14 29 - 17 9 10 11 11 25 15 19 29 15 15 16 17 - 11 11 19 23 16 16 18 17 37 38 12 12 (**L23, Jewish**)

In this particular instance the dataset contained 42 of 37 marker haplotypes, which collectively had 620 mutations; this gives 620/42/.09 = 164 → 196 generations, or 4900±530 years from a common ancestor. The base haplotype differs by 9 mutations from the "young" Jewish base haplotype above in its first 37 markers, which makes the 4900 ybp haplotype likely ancestral. The oldest identified Jewish R1b base haplotype is of 5400±500 ybp (Klyosov, 2008b). It has certainly appeared in the Middle East, during the Sumerian era.
95

Sumers, the likely Bearers of R1b1a2 Haplotypes, and Their Descendant Assyrians

Assyrians are one of the oldest surviving groups descended, as it is believed, from the historic Sumers. Among Assyrians, R1b is the major haplogroup, reaching 40% of the studied population; haplogroup J takes second place with 11% and others are in singular percentages (Lashgary et al., 2011). Such a high percentage of R1b haplogroup is very unusual in South Mesopotamia. According to the Assyrian FTDNA Project (see Appendix) half of their R1b bearers belong to R-L23 subclade with DYS393 = 12, and all nine 12 marker haplotypes have 35 mutations, which gives 35/9/.02 = 194 → 240 conditional generations, or 6000±1200 years to their common ancestor. Only 6 haplotypes were available in the 25 marker format, and they have 48 mutations; they yield 48/6/.046 = 174 → 211 generations, or 5275±930 years. Those are indeed Sumerian times (e.g., Kramer, 1971). Most of the bearers of those haplotypes now live in Iraq and Turkey. Another half of the Assyrian haplotypes in the Project, mostly from Iran, have a slightly mutated "classical" European R1b1a2 base haplotype (with DYS464 = 15 15 17 17), and a common ancestor of 850±360 ybp calculated from the first 12 markers, and 1100±280 ybp from the first 37 markers. This R1b evidence is clearly a relatively recent event in Assyrian population, brought from Europe.

There is no known mass movement of Assyrians from Europe 1000 years ago, or but there was a major movement of people 1000 years ago, connected with the Turkic takeover of the Near East, and starting with Mesopotamia, and that could be a reasonable source of the relatively recent event.

Figure 9. The 17-marker tree composed with 120 haplotypes of the R-L23 haplogroup, except five haplotypes of the M269 subclade (numbers 1 - 5)
The tree is composed based on data in (Herrera et al., 2011)
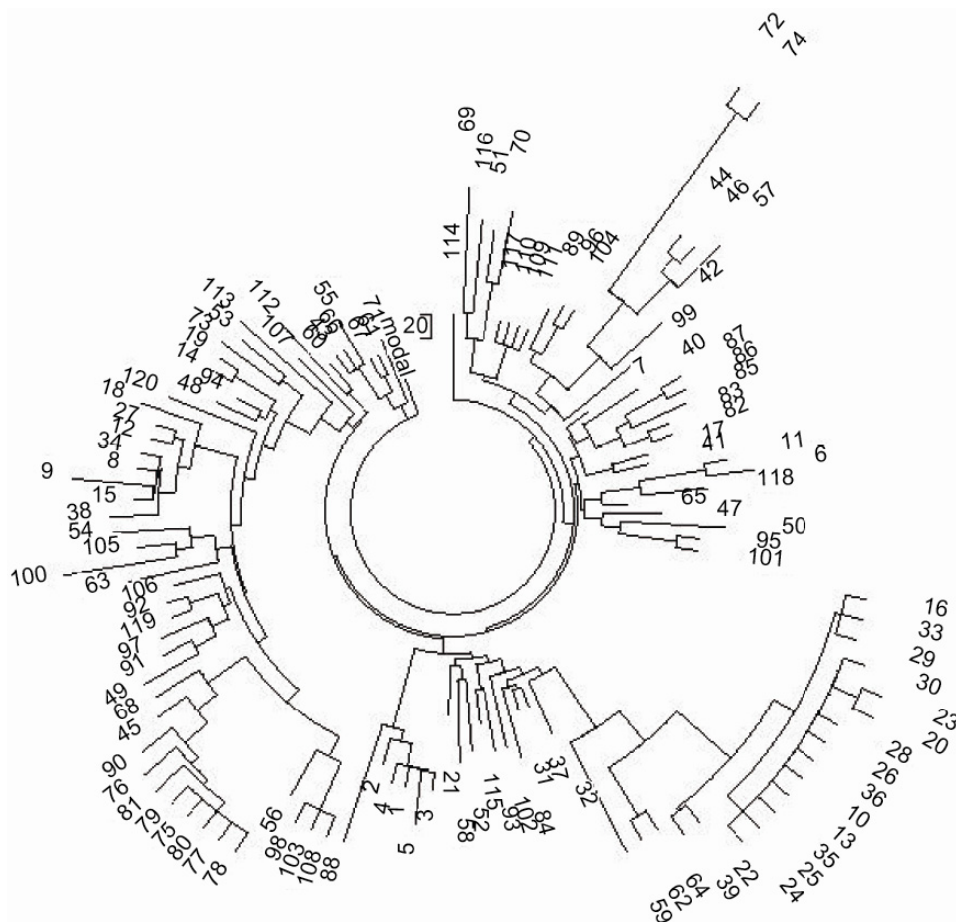
R-M269

The R-L23 subclade is clearly traced from the East European Plain south via the Caucasus, where it prevails among R1b haplotypes, and via Anatolia, where it is very pronounced; down to South Mesopotamia, where Sumers had lived between 6000 and 4500 ybp (Kramer, 1971). Since timespans to common ancestors of those R-L23 haplotypes are around 6,000 - 5,000 ybp, it is quite likely that those common ancestors lived among the Sumers and their ancestors (in the Caucasus and Anatolia). It might be an additional feature for linguists, some of whom consider Sumerian as a remnant of a subgroup of the Dene-Caucasian language superfamily (e.g., Bengtson, 1997).

> Numerous authors connect Sumerian with Türkic, and the Sumer's endonym Kang with the Central Asian Kangar people; research suggests that Sumerian has a load of Türkic borrowings interspersed with the Sumerian substrate (see Türkic-Sumerian for a sampling of works and bibliographies); research also finds numerous Hungarian isoglosses, with also potential origin in the Türkic lexis. Thus there is concurring linguistic, onomastic, and biological evidence of Sumer - Türkic historical connection.
>
> The reference to the Dene-Caucasian language superfamily (superfamily including North Caucasian languages, Sino-Tibetan languages, and Ket as an Yeniseian language) is also significant, research suggests that the agglutinative Amerind languages have too many systemic correspondences with the Türkic agglutinative languages to surmise a random coincidence. The significance is especially momentous because the temporal separation between the American and Eurasian ancestors is guestimated at 15,000 ybp, which points to a unique stability of the agglutinative languages as a linguistic phenomenon (see Contents Amerind Turkic for a sampling of works and bibliographies).

Migrations of the Arbins from the Pontic Steppes, Asia Minor and Middle East Westward to Europe

In the descriptions above we left bearers of R1b haplogroup in the Caucasus/East European Plain/Pontic steppes, Anatolia, and Lebanon/Southern Mesopotamia around 6000 - 5000 ybp. All three areas were the relay regions for the Arbins to move further westward. Haplotypes of their present day descendants serve as the tracers of those ancient migrations.

> Their movement could not be limited to the "further westward", it is our self-centered attention that locks our focus in that particular fashion. It appears that once explored, the whole known terrain becomes permanently logged in the collective memory, and the shape of migrations is rather a sequence of boomerang flights than a flight of an arrow, as exemplified by the independent returns of the Ases and Kayis back to the Caucasus. With varying demographic scales and temporal separations, not all of the see-saw waves are detectable. Exceptional, on the today's scale, demographic paucity of the migratory groups had to cause an interference effect, with high crests and low troughs; the literary sources reflected the crests, while demographical genetics tends to see the troughs, genetical "noise" of the transitory migrations.

A dataset of R1b1a2 haplotypes on the Balkans was published in (Barac et al., 2003a, 2003b; Pericic et al., 2005, and private communications with M. Pericic). It contains a series of obviously R-L23 haplotypes with a base

12 24 14 11 11 15 X X X 13 13 29,

and a common ancestor of 4050±890 ybp (Klyosov, 2010c). Another series of the Balkan R1b1a2 has the base

13 24 14 11 11 11 X X X 14 13 29,

with a common ancestor of 4975±1300 ybp (ibid). Four mutations between them (including multi-copy mutations counted differently) place their common ancestor at ~7000 ybp to the time of ancient R-M269 haplotypes of the East European Plain's (or further on to the east). Migrations westward from the Pontic steppes are also supported by multiple archaeological evidence, which include those dating to the 3rd millennium BC (e.g., Mallory, 1989).

The haplotype tree shown in **Figure 8** provides some clues regarding possible directions of those ancient migrations. 40% of all 107 haplotypes of the tree belong to the Armenians and Turks; of those, 70% are haplotypes on the right-hand (larger) part of the tree. The upper left branch in the tree contains haplotypes from Russia, Lithuania, Poland, Croatia and Ireland.
[96]

They all have the same pattern of mutations, with the base haplotype

12 24 14 11 **11 14** 12 12 12 **12** 13 29 - 16 9 10 11 11 25 15 19 **31 14** 15 16 18 (**L23, mainly Eastern Europe**)

which differs from the base haplotype of the left-hand side 38-haplotype branch by three mutations (marked; some alleles are fractional). This places their common ancestor with the entire branch at 4600 ybp. This indicates that the above largely East-European branch split from the sizeable branch at the East European Plain in the middle of the 3rd millennium BC, and apparently migrated westward. That is one migration route (out of several described) which explains how R-L23 haplotypes appeared in Europe during those times.

Another migration route took place from the East European Plain southward, as indicated by another quite distinct branch on the opposite side (at 5 o'clock) of the haplotype tree. The branch includes haplotypes from Russia, Lithuania, Armenia (two haplotypes), Turkey (three haplotypes), Syria. The base haplotype is as follows:

12 24 14 11 11 14 12 12 12 **13** 13 30 - 16 9 10 11 11 25 15 19 **28 15** 15 16 **17** (**L23, mainly South**)

There are five full mutations between these two base haplotypes of the branches, both which includes Russian and Lithuanian haplotypes, albeit with a principally different history. One group of L23 bearers went west to Europe; another went south to the Caucasus and the Middle East. They are separated by $5/.046 = 109 \rightarrow 122$ generations, or by 3050 "lateral" years (this time is required on average to make five mutations in the 25 marker haplotype). In other words, these two branches show a fork in migration routes of R-L23 from the East European Plain, westward and southward.

One additional branch of R-L23 includes haplotypes from Armenia, Lebanon, Bulgaria, Italy, France, Spain, Germany. It is located at 7 o'clock on the haplotype tree in **Figure 8**. Their base haplotype

12 24 14 11 11 14 12 12 12 13 13 **29** - 16 9 10 11 11 25 15 19 **29** 15 **16** 16 18 (**L23, South and South-West**)

differs from the preceding discussed branch by only 2.6 mutations or by 1550 "lateral" years. This points to R-L23 movement from the Middle East westward to Europe. Its neighboring branch, including haplotypes from Turkey (three haplotypes), Greece and Albania evidences nearly the same base haplotype as the above, with only one mutation $12 \rightarrow 13$ in the first allele (DYS393) and $18 \rightarrow 17$ (17.3 on average) in the last allele. It is essentially the same migration route from Asia Minor to Europe. In fact, the entire right-hand large branch in **Figure 8** consists of sub-branches with predominantly Armenian and Turkish (some Iraqi and Iranian) R-L23 haplotypes with singular inclusions of haplotypes from Greece, Germany, Netherlands, England, Scotland. This reflects migration routes of the Arbins from Asia Minor and the Middle East to Europe mainly between 4500 and 3500 bp.

It may be expected that some migrations from the Middle East to Europe are associated with this mutation of DYS393 $12 \rightarrow 13$ in the subclade R-L23. However, other, later migrations, could have occurred from the European continent eastward, and belong to more recent subclades, such as L51, L11, P312, U106, L21, U152, etc. (Klyosov, 2011b). Indeed, most Sardinian haplotypes have DYS393=13, such as the following base haplotypes on Sardinia (calculated from data provided in [Contu et al., 2008]):

13 24 14 11 11 14 X X X 13 X 29
13 24 14 10 11 14 X X X 13 X 29
13 24 14 10 11 15 X X X 13 X 29

The first one was presented on the tree by a series of identical haplotypes which are obviously derived from the very recent common ancestor (Klyosov, 2008a, 2010c). They are indistinguishable from common European R1b1a2- M269 subclades, such as R-P312, R-L21, R-U152, R-L2, R-L20, etc. The second one has a timespan to its common ancestor of 3550±790 ybp, the third descended from a common ancestor who lived 2900±620 ybp. A common ancestor of all of them lived 5025±630 ybp, which fits the time and direction of ancient migration of the Arbins from the Middle East to Europe (however, DYS393 = 13 would be unusual for them), and from Iberia up north to the continent, and subsequently populating Europe in all possible directions.

One of the most common R1b base haplotypes in Sicily is as follows (Di Gaetano et al., 2009):

13 24 14 11 11 14 X X 12 13 13 29,

which is identical to one of the Sardinian base R1b haplotypes above, and indistinguishable from common R1b subclades in Europe. The timespan to a common ancestor for that haplotype in Sicily calculated from the cited study is 4550±1020 years. The same base haplotype is the most common in Italy (Capelli et al., 2007), with a timespan to the common ancestor of 4125±500 years calculated by the linear method, and 4300±1160 years calculated by the logarithmic method (Klyosov, 2008a).

There is one more, but very important migration route of the Arbins which does not practically include R-L23 haplotypes (only 5% of DYS393 = 12

among Iberian R1b haplotypes, it corresponds to a random mutation from the parent DYS393 = 13). It is a route from the Middle East westward along North African Mediterranean seacoast. R-L23 was either not represented or did not survive along this route. It seems that bearers of R-V88 were part of the journey, however, they split and went southward and settled in Central Africa (mainly Cameroon and Chad), mentioned previously. The migration of the Arbins from the Middle East to the Atlantic, then across the Strait of Gibraltar to the Pyrenees took place from ~ 5500 - 5200 ybp to 4800 ybp when the Arbins landed in Iberia (see supportive references to archaeological data below). Part of the way from Egypt to Iberia could have been made by sea, details are not known as yet. There are some historical reports of arrival of the Egyptian military fleet to Iberia some 5000 ybp; there are some allegedly Egyptian mummies and fragments of Egyptian tombs exhibited in the Royal Academy of History in Madrid and in the Tarragona City Museum. Their status, however, is rather vague. It seems, nevertheless, that bearers of R1b subclades, mainly R-M269, and newly formed L51 and L11 (see below) had arrived in Iberia and this was the beginning of the archaeological Bell Beaker culture.

Previous to discussing the Bell Beakers and history of the Arbins in Europe, it is worth to mention one more rather vague evidence of the R1b journey via the ancient Egypt between 5500 and 5200 ybp. It concerns the alleged R1b haplotype of Pharaoh Tutankhamun. [97]

Alleged Pharaoh Tutankhamun R1b1a1-M269* Haplotype and Its Possible History

Recently the Swiss company iGENEA has published the alleged 16 marker haplotype of the Pharaoh:

13 24 14 11 11 14 X X 10 13 13 30 - 16 14 19 10 15 12 (www.igenea.com/en/index.php?c=62).

Here the first 12 markers are shown in the FTDNA format, and the rest are DYS 458, 437, 448, GATA H4, DYS 456, 438. It is obviously not a typical European R1b1a2 haplotype, since it has DYS439 = 10, and not a common European 12. There are only about 0.5% R1b haplotypes in Europe with DYS439 = 10. The most likely and the most closely related base haplotype is that of R1b1a2-M269, shown in **Figure 5**, in which the same markers as those available in the Pharaoh haplotype are noted in bold:

12 25 14 11 11 14 11 12 **12 13 13 29 - 17** 9 10 11 11 25 **15 19** 29 15 15 16 17 - 10 **11** 19 23 **15** 15 19 17 35 38 12 **12** - 11 9 15 16 8 10 10 8 11 10 12 23 23 16 10 12 12 15 8 13 22 20 13 12 11 13 11 11 12 12 (**M269**)

There are 6.8 mutations between the two 16 marker haplotypes (some alleles in M269 haplotype are fractional), which translates to 6.8/.0315 = 216 → 274 conditional generations, or 6850 years between them. Since Tutankhamun lived 3300 ybp, and the R-M269 base haplotype is 6200±900 years "old", then a common ancestor of the two lived (6850 + 6200 + 3300)/2 = 8175 ybp. This might have been either an ancient R-M269 ancestor of the Tutankhamun from Central Asia, or the respective upstream subclade.

The main point here is that the Pharaoh haplotype is not some erroneously (or intentionally) picked European R1b1a2 haplotype; it is an archaic haplotype of R1b haplogroup, likely of the R-M269 ancient subclade.

It should be mentioned here that the founder of the Egyptian pharaohs was Narmer, his origin is not firmly known, he was a founder of the First Dynasty, he lived ca. 32nd century BC — around 5200 ybp. This date fits the migration times for the Arbins along the Mediterranean coast in North Africa. It cannot be excluded that the bearers of R1b haplogroup may have actually established the Royal lineage in ancient Egypt. It does not mean that the lineage was not interrupted through the Dynasties; however, it kept returning to the Dynasties, if the R1b origin of some Pharaohs is right.

Overview of Migratory Path of the Arbins, Bearers of R1b Haplogroup, from Central Asia to Europe

An overall map of R1b migration routes from the very "beginning", where "beginning" is undefined in detail thus far and can be estimated between 16,000 and 12,000 ybp, and until their arrival to the Pyrenees by 4800 ybp as future Bell Beakers, is shown in **Figure 10**. Their migratory path was slowly taking place from Central Asia, apparently from the Altai, South Siberia, where some very different R1b haplotypes were discovered. As it is described in the first sections of this paper, their tremendous mutational difference with European and Middle Eastern R1b haplotypes places their common ancestors at ~16,000 ybp.

Figure 10. Map of R1b haplotype locations addressed in this study, until they arrived to Iberia ~4800 ybp.
The overall migration path of the Arbins had apparently started in South Siberia, where present day R1b haplotypes of the Shors, Teleuts, Altaians, Uigurs, Khakasses, Tuva people (the right-hand side pins) are tremendously different from those in the Europe and Middle East. The migration route continued through the regions in Middle Asia and Russia, where presently live Bashkirs, Kazakhs, Uzbeks, Mari, Tatars, and Kalmyks. From the Central Russia's region the Arbins moved to the Caucasus and further South to Anatolia, Middle East, present day Iraq, Lebanon, and then westward along the Mediterranean Sea onto Northern Africa to the Atlantic. On that way R1b-V88 bearers split and went deep into Africa, where their descendants currently live in Cameroon and Chad (Cruciani et al., 2010). Their migratory path into Africa is reproduced here from the above reference. The map shows R1b locations previous to arrival to the Pyrenees ~4800 ybp, from where they moved North as Bell Beakers and populated Europe between 4800 - 3000 ybp and later.

The legend to **Figure 10** describes those Central Asian/Siberian populations. There are many Neolithic, Chalcolithic and Eneolithic archaeological cultures in the area, such as Tersek, Ural, Surtandi, Mahandzhar, Iman-Burluk, Botai, Atbasar, Kelteminar, and other archaeological Central Asian cultures in present-day Russia (e.g, Zakharov, 2010), which might be assigned to the Arbins; however, it would be premature to assign any of them to R1b or to any other haplogroup. Such a task is quite new for archaeologists. It is tempting to point at Seroglazovo, Khvalyn, Samara, Middle-Volga and adjacent archaeological cultures of 12,000 - 5000 ybp of the European east as the most likely R1b cultures. We cannot, however do it for the same reason of prematurity, and it would be irresponsible to suggest such at this time. The same may be said for Timber Grave, Catacomb and neighboring archaeological cultures of Central and South Russia, which apparently were shared by both R1b and R1a bearers, albeit at different time periods. The R1b people before 5000 ybp, and R1a people after 4500 ybp have confused archaeologists who have observed "different roots" of those cultures, spreading in different directions and at different times.

> Here we encounter an interesting thesis, that the people who separated 10,000 years before, could reunite in a single archeological culture like they parted within a life of one generation. In purely generational terms, 10,000 years constitutes 400 generations for either prong of the fork. That's 400 generations separated uncounted, and differing, technological innovations, genetical and linguistic admixtures, social developments, religious developments, etiological developments, ethnological developments, and who knows what else. And then the refugees from the European genocide and the nomadic old timers from the Asia, with whatever genetical markers in their Y-DNA chromosome, encounter each other on the fringes of the Eastern Europe, and for 500 years become indistinguishable from each other. A single culture, single technology. single society, single art, ceramics, religion, etiology, ethnology, and everything else. Like there was no Babylon Tower that separated people into warring fractions, like a return of the Edem to the Earth. And 500 years later, comes another miracle, a magic separation of indistinguishable single people into two contrasting and opposing multitudes, one clearly a continuation of the previous single nomadic animal husbandry archeological civilization, and the other purely agricultural, with a language of incompatible family, and clearly unrelated to the previous single nomadic animal husbandry archeological civilization.
>
> Genetically-wise, it is premature to tell who is who, such a task is quite new for archaeologists. Otherwise, and including biologically, it is perfectly clear who is who, and no amount of modern mislabeling can change the past situation in or on the ground.

The map shows that the current bearers of R1b spread over Central Russia up to Arkhangelsk on the White Sea. Very likely it was a relatively recent relocation, although it remains to be determined. Currently there are only about 5% of the R1b bearers in the European part of Russia.

As was described above in this study, the Arbins went South through the Caucasus to the Mesopotamia and Middle East around 6000 - 5000 ybp; they established the Sumer civilization; went westward via Egypt to the Atlantic, and across the Gibraltar Strait to the Pyrenees. On their way, some R1b-V88 bearers split; they went deep into Africa, and currently populate Cameroon and Chad in appreciable amounts (see legend to **Figure 10** and description and references above). By 4800 ybp the Arbins have reached Iberia to become the first Bell Beakers. This date was obtained for a common ancestor of haplotypes of P312 and U106 subclades (see the next Section), and is supported by archaeology data (Cardoso & Soares, 1990; Martinez et al., 1996; Cardoso, 2001; Muller & Willigen, 2001; Nocete, 2006).

Several Entry Routes of R1b1a2 Haplotypes to Europe: from the East European Plain, from Asia Minor/Middle East, from the Pyrenees

To sum up the preceding section, the Arbins were entering Europe from the east by several routes: from the East European Plain (between the Pontic Steppes and the Baltic Sea); from Asia Minor and the Middle East; and after a long way around Mediterranean Sea to Iberia, up north to the European continent. The first two principal routes were associated with bearers of the R-L23 subclade. The Iberian route was made by mainly M269 people, who at

the time of entry the Pyrenees spun off the R-L51, and immediately later the R-L11. Both are very similar and have very close timespans to their common ancestors, as is shown in the next section. In 4850 ybp L11 promptly spun off two "brother" subclades, P312 and U106 (Klyosov, 2011b), which after a long "population bottleneck", on the edge of extinction, eventually survived and expanded around 4000 - 3700 ybp, and actively populated Europe, first as Bell Beakers, between 4000 and 3000 ybp, and then toward the era of the Ancient Rome, as Gauls and Celts, to mention only the names with certain historical "milestones" present. In fact, in Europe lived dozens if not hundreds of the ancient R1b tribes.
98

Main R1b1a2 Subclades on the European Continent: Entering Europe 4800 - 4500 ybp from the East and from the Southwest

Nearly all haplotypes of the subclade R-M269 6000 - 5000 ybp belonged either to R-M269* paragroup, or to its downstream subclade R-L23. The Arbins who had migrated from the east to Europe during those times mostly carried DYS393 = 12. Currently we see only those R1b1a2 (xL23) haplotypes in Europe, they are derived from the common ancestors who lived a maximum ~4500 ybp. Either descendants of more ancient L23 ancestors have not survived into the present time, or none came from the East European Plain, Asia Minor, and the Middle east.

R1b	M343
R1b1	L278, M415, P25_1, P25_2, P25_3
R1b1a	P297, L320
R1b1a1	M73, M478
R1b1a2	L265, M269, M520, S3, S10
R1b1a2a	L23/S141, L49.1
R1b1a2a1a	L51/M412/S167
R1b1a2a1a1	L11, P310S129, P311
R1b1a2a1a1a	U106/S21
R1b1a2a1a1b	P312/S116
R1b1a2a1a1b3	U152/S28
R1b1a2a1a1b4	L21/M529
R1b1a2a1a1b4b	M222
R1b1a2a1a1b4g1	L226
R1b1a2a1a1b4h1	P314

The above diagram shows that the immediate downstream subclades of the R-L23 were L51 and then L11 (ISOGG-2012, in an abbreviated form). The dynamics of these subclades is much more understood via Iberia into the continent, where the migration of the Arbins is being identified with the Bell Beakers.

The question is — where those L51 and L11 subclades could have arisen? If they are 6000-5000 years "old", they could have split in Asia Minor, the Middle East or on the East European Plain, and enter Europe from there. The "intra-clade" haplotypes, that is only L51 or only L11 subclade, might reflect population bottlenecks, hence, they look "younger" than their actual age (in terms of mutations and respective TMRCA). However, their "inter-clade" comparison could reveal lost (due to bottlenecks) timespans to more ancient common ancestors. To analyze those subclades, a combined L51-L11 haplotype tree is shown in **Figure 11**.
99

Figure 11. The 67-marker tree with 42 haplotypes of the R-L51 subclade (15 haplotypes on the right, numbers between 1 and 23) and R-L11 (27 haplotypes on the left, numbers between 26 and 61).
Haplotypes are listed in the FTDNA public Project http://www.familytreedna.com/public/ht35new/default.aspx
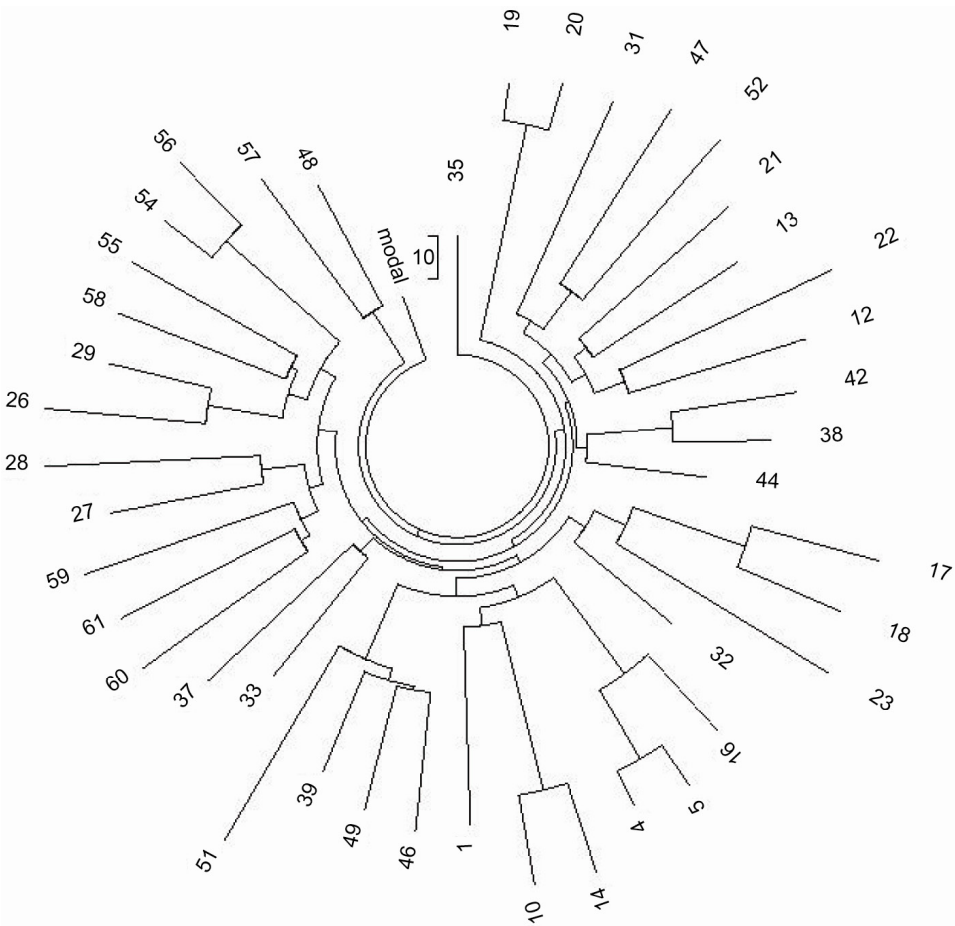
One-third of haplotypes on the tree belong to the L51 subclade, and they occupy the right side. Another two-thirds are L11 haplotypes, which are on the left and make some "insert" branches on the right. The 67 marker haplotypes of L51 and L11 subclades are very similar, therefore the tree could not distinguish them in a number of cases; hence, the mix of the branches. The tree also shows that the "age" of the two subclades is also very similar, since the "height" (which generally indicates the "age") of the branches around the tree is about the same. The base haplotype of R-L51 subclade is as follows

13 24 14 11 11 14 **13** 12 12 13 13 29 - **17** 9 10 11 11 25 15 19 **30** 15 **16 17 18** - 11 11 19 23 16 15 **18** 17 36 **38** 12 12 - 11 9 15 16 8 10 10 8 10 **10** 12 23 23 16 10 12 12 **16** 8 12 22 20 **14** 12 11 13 11 11 12 12 (**L51**)

It deviates from the base R-L23 haplotype (5400±800 ybp) by 9 mutations (12 mutations are marked above, but many of them are fractional), that is by 9/.12 = 75 → 81 generations, or ~ 2025 years of the "lateral" distance. All 15 haplotypes contain 280 mutations from the base haplotype, which gives 280/15/.12 = 156 → 184 generations, or 4600±535 years from their common ancestor, L51. In turn, it gives (2025 + 5400 + 4600)/2 = 6000 years, which is an "age" of a common ancestor of both L23 and L51 subclades. Obviously, it is the common ancestor of L23 subclade himself.

In other words, the "age" of the L51 subclade as 4600±535 years is obtained from mutations on the L51 branch on the tree (**Figure 11**) and confirmed 1) by a mutational distance from the base haplotype of the parent L23 subclade, 2) by the "age" of the L23 subclade, determined earlier, and 3) by the general phylogeny of the R1b subclade. The "age" of the L51 subclade fits well with the arrival time of the Arbins to Iberia (4800 ybp), to the distribution pattern of the current bearers of L51 in Europe (the highest frequency is in the Pyrenees and immediately up north in France and on the Isles, see [Myres et al, 2010]. It is practically absent in European South-East, East, and North-East (ibid.).

Analysis of the L11 subclade is more complicated since its haplotypes are spread around the tree by at least four branches, each with its common ancestor. They differ from each other by 32 mutations which gives 32/4/.12 = 67 → 72 generations, or 1800 years below their average "age" ~2500 years. Therefore the "age" of the R-L11 subclade is ~4300 years, which is indeed consistent with 4600±535 years for the L51 subclade. The base haplotype for the L11 subclade is as follows:

13 24 14 11 11 14 **12** 12 12 13 13 29 - 17 9 10 11 11 25 15 19 30 15 **15** 17 **17** - 11 11 19 23 16 15 18 17 36 38 12 12 - 11 9 15 16 8 10 10 8 10 10 12 23 23 16 10 12 12 **15** 8 12 22 20 **13** 12 11 13 11 11 12 12 (**L11**)

Five deviations between them (marked in bold) present in fact 3.7 mutations (since most of them are fractional), which gives 3.7/.12 = 31 → 32 generations; that is only 800 "lateral" years between them. This confirms that L51 and L11 subclades are very close to each other in time. Their common ancestor (which presumably should be L51) lived (800 + 4600 + 4300)/2 = 4850 ybp. It seems that L11 split off only ~250 years later. All of them are likely to have established the Bell Beaker archaeological culture. The oldest artifacts of the Bell Beakers were found in Portugal, dated 4800 - 4600 ybp (Cardoso & Soares, 1990; Martinez et al., 1996; Cardoso, 2001; Muller & Willigen, 2001; Nocete, 2006).
100

Main R1b1a2 Subclades on the European Continent: Population of Europe after 4800 ybp, Main Subclades on the Isles and Their Likely Origin

The phylogeny of R1b (see the **diagram** above) shows that the further downstream subclades of R-L11 are P312 and U106. Their 67 marker haplotypes are as follows:

13 24 14 11 11 14 12 12 12 13 13 29 - 17 9 10 11 11 25 15 19 29 15 15 17 17 - 11 11 19 23 15 15 18 17 36 38 12 12 - 11 9 15 16 8 10 10 8 10 10 12 23 23 16 10 12 12 15 8 12 22 20 13 12 11 13 11 11 12 12 **(P312)**

13 **23** 14 11 11 14 12 12 12 13 13 29 - 17 9 10 11 11 25 15 19 29 15 15 17 17 - 11 11 19 23 **16** 15 **17** 17 **37 39** 12 12 - 11 9 15 16 8 10 10 8 10 10 12 23 23 16 10 12 12 15 8 12 22 20 13 12 11 13 11 11 **13** 12 **(U106)**

There are 5.5 (marked in bold) mutations between them (DYS456 in P312 fluctuates between 15 and 16 in various datasets), which separates the two base haplotypes by 5.5/.12 = 46 → 48 generations, or by 1200 years. Since common ancestors of both subclades lived 4350±700 and 4175±430 ybp (Klyosov, 2011b), their common ancestor lived (1200 + 4350 + 4175)/2 = 4850±700 ybp. Indeed, it is the "age" of the subclade L11 within margin of error. Generally, it is safe to accept the "age" of both P312 and U106 as ~4200 ybp, since estimates, depending on a dataset, are fluctuating around this value. There is 1200 years between them (see above) which defines a time-span to their common ancestor, L11, as (4200 + 4200 + 1200)/2 = 4800 ybp.

As it was stated above, the four subclades, L51, L11, P312 and U106 historically represent the first wave of Bell Beaker movement from Iberia to the European continent. The highest frequency of P312* subclade in Europe is currently observed in Iberia (Myres et al, 2010), however, it is naive to judge the ancient locations of the subclades based on their current distribution. Their bejumble on the Continent has been too intense in the past millennia to make any unambiguous extrapolations. The current regional distribution of the R-M269's downstream subclades, which moved from the Pyrenees nearly five millennia ago, is emblematically non-informative, except in instances of some local subclades. One of them is R-M222.

It seems that the history of R-M222 appearance on the Isles may have began with seafaring carriers of L11* bearers from Iberia to the Isles some 4500 - 4000 ybp. However, that is just one theory. The other involves migrations of L21 bearers to the Isles by land. And then P312, a downstream clade of L11, is also intensely represented on the Isles, along with its downstream subclades, as well as its "brother" subclade U106 and its downstream clades (ibid). L21, a downstream clade of P312, is abundant on the Isles, unlike its "brother" subclade U152 (ibid). This ancient migration of the Arbins from Iberia to the Isles explains some genetics data on the "Spanish origin of the Irish".

The M222, a downstream clade of L21, is among the most represented R1b subclades on the Isles. A massive Ireland Heritage FTDNA project (see Appendix) lists several thousand of R1b (with subclades) haplotypes in the Isles, of which M222 is the largest, and includes about 25% of all R1b (Klyosov, 2010e). The most the subclade is expressed in Ireland. In Scotland it is observed mainly in the lowlands and in the central region, and is poorly present in England. The base haplotype of the M222 subclade is:

13 25 14 11 11 13 12 12 12 13 14 29 - 17 9 10 11 11 25 15 18 30 15 16 16 17 - 11 11 19 23 17 16 18 17 38 39 12 12 - 11 9 15 16 8 10 10 8 10 10 12 21 23 16 10 12 12 16 8 12 25 20 13 12 11 13 11 11 12 12 **(M222)**

It is of interest that the above base haplotype differs from the "brother" base haplotypes of P312 and U106 by 18 mutations in each case (in fact, 15 mutations, since some mutations are fractional). It shows that P312 and U106 are indeed very close (in time and in heritage) base haplotypes, and that their common ancestor, a bearer of L11 base haplotype, lived very close in time to both of them. 15 mutations place their common ancestor at 4600±500 ybp.

Multiple analyses of M222 haplotype datasets produce a time-span to a common ancestor of the currently living M222 descendants at 1450±160 years (Klyosov, 2010e). However, cross-examination of numerous lineages in the M222 subclade has shown that the M222 lineage arose not later than in the beginning of the Common Era, some 400 years earlier than the above estimate. It points out to a possible population bottleneck between the time when M222 arose, and the expansion of the subclade in the middle of the 1st millennium CE. One cannot possibly say, at least yet, where specifically the M222 subclade first appeared (that is, the G → A single nucleotide mutation occurred in certain nucleotide of Y chromosome of a certain individual, the bearer of the upstream L21 subclade), whether in the Isles, in North Western Europe, or elsewhere. We also cannot determine specifically what the R-M222 direct descendants have been struggling against for several centuries for their survival. Was it a continuous climatic event? The Roman invasions? Epidemics? Famine? or a combination? We know, however, that it was eventually the Isles where that mutation on the Y chromosome was carried on and flourished.

The phylogenic diagram shows that the M222 mutation arose concurrently with its "brother" SNP mutations that now define the subclades L226 and P314. The first one has a common ancestor who lived 1500±170 ybp, which is at the same time with that of M222, albeit in L226 the subclade-defined mutation was C → T. Its base haplotype is as follows:

13 **24** 14 11 11 **14** 12 12 **11** 13 **13** 29 - 17 **8 9** 11 11 25 15 **19 29 13 13 15** 17 - 11 11 19 23 **15 15** 18 17 **36 38** 12 12 - 11 9 15 16 8 10 10 8 10 10 12 **23** 23 15 10 12 12 **15** 8 12 22 20 13 12 11 13 11 11 12 12 **(L226)**

It differs by as many as 21 mutations from its "brother" M222 subclade, despite the observation that their common ancestors lived at the same time. It only confirms that their common ancestor, the founder of subclade L21, lived around 4100 ybp. It also shows that L21 formed almost immediately from P312 (see the phylogenic diagram) since P312 arose 4350 - 4200 ybp and L21 around 4100 ybp. Their 67 marker haplotypes are identical, within some fractional mutational difference:

13 24 14 11 11 14 12 12 12 13 13 29 - 17 9 10 11 11 25 15 19 29 15 15 17 17 - 11 11 19 23 15 15 18 17 36 38 12 12 - 11 9 15 16 8 10 10 8 10 10 12 23 23 16 10 12 12 15 8 12 22 20 13 12 11 13 11 11 12 12 **(P312)**

13 24 14 11 11 14 12 12 12 13 13 29 - 17 9 10 11 11 25 15 19 29 15 15 17 17 - 11 11 19 23 **16** 15 18 17 36 38 12 12 - 11 9 15 16 8 10 10 8 10 10 12 23 23 16 10 12 12 15 8 12 22 20 13 12 11 13 11 11 12 12 **(L21)**
[101]

The only apparent difference, in DYS456, is also fractional, since in the P312, this allele fluctuates between the values of 15 and 16.

Despite their similar chronology, for the preceding 2600 years the subclades M222 and L226 have quite different histories. Yet another "brother" subclade, P314 (more accurately, P314.2, since the same P314 SNP mutation, T → C, has occurred in haplogroup H2a), is somewhat "older", with its 2225±300 ybp; however, it could be within the same time frame when M222 actually arose. Its base haplotype

13 23 14 11 11 14 12 13 13 13 13 29 - 17 9 10 11 11 26 15 19 29 15 15 16 17 - 10 11 19 23 16 15 18 18 37 38 12 12 - 11 9 16 16 8 10 10 8 11 10 12 23 23 17 10 12 12 15 8 12 22 20 13 12 11 13 11 11 12 12 **(P314.2)**

is again standing out from the base haplotypes of the M222 and L226 subclades by 24 and 17 mutations respectively. In the L21 family this is a third, completely different lineage. The 10 full mutations from its L21 parent base haplotype produce $10/.12 = 83 \rightarrow 91$ generations, or 2275 years of the mutational difference, and place the beginning of the L21 subclade at $(2275 + 4100 + 2225)/2 = 4300$ years, which is within margin of error from 4100 years estimated above. All these base haplotypes, along with the intra-clade and inter-clade datings form a rather robust system of Arbins' DNA genealogy in Europe and Asia.

Materials and Methods

Four thousand four hundred eight (4408) of R1b haplotypes (with subclades) were collected in databases from FTDNA, YSearch, and in peer review publications.

The methodology of haplotype datasets' analysis was described in the preceding publication in this journal (Rozhanskii & Klyosov, 2011). This study employed linear and logarithmic methods; the latter when the base haplotype in the dataset was easily identifiable, as described in (Klyosov, 2009c). The mutation rate constants are listed in (Klyosov, 2009c; Rozhanskii & Klyosov, 2011), and for a number of cases are given in the text of this paper. The most widely used mutation rate constants, including those employed in this paper, are as follows (number of mutations per haplotype per conditional generation of 25 years); references show examples of the haplotypes in the format indicated:

.013, 8 marker haplotype (Zhong et al.)
.017, 8 marker (Contu et al.)
.018, 10 marker (Myres et al.)
.020, 11 marker (Cruciani et al.)
.019, 11 marker (Zalloua et al.)
.020, 12 marker FTDNA format
.024, 12 marker (Malyarchuk et al.)
.034, 17 marker Y-filer format (Herrera)
.035, 19 marker (Balanovsky et al.)
.046, 25 marker FTDNA format
.090, 37 marker FTDNA format
.120, 67 marker FTDNA format
.198, 111 marker FTDNA format

The FTDNA haplotype formats are given in http://www.familytreedna.com/faq/answers.aspx?id=9

Haplotype trees were composed using software PHYLIP, Phylogeny Inference Package program (see Klyosov, 2009c, 2009d and references therein). Corrections for back mutations were introduced as described in (Klyosov, 2009c). Margins of error were calculated as described in (Klyosov, 2009c).

Base haplotypes in the dataset were determined by minimization of mutations; by definition, the base haplotype is one which has the minimum collective number of mutations in the dataset, derived from one common ancestor. The base haplotype is the ancestral haplotype or the closest approximation to the latter.

A timespan to the common ancestor of two base haplotypes is determined as follows:
1) count the number of mutations between the two base haplotypes,
2) divide the obtained number by the mutation rate constant,
3) introduce a correction for back mutations, calculated using the following formula (Klyosov, 2009c):

$$\lambda = \lambda_{obs}/2_{obs}(1+\exp 2(\lambda_{obs}))\ \textbf{(1)}$$

where:

λ_{obs} = observed average number of mutations per marker in a dataset (or in a branch, if the dataset contains several branches/lineages),
λ = average (actual) number of mutations per marker corrected for back mutations,

4) add the obtained value, multiplied by 25 (years), which represent the "lateral" timespan between the times of appearance of the two base haplotypes, to TMRCAs for base haplotypes and divide by 2.
The result represents the TMRCA (time to the most recent common ancestor) for the examined two base haplotypes. The following examples are based on specific cases in this study.

Example 1: Calculation of a timespan to a common ancestor of the branch (the most remote on the haplotype tree shown in **Figure 1**) which contains 12 of eight marker haplotypes with collective 65 mutations from their base haplotype. Since the mutation rate constant for these haplotypes equal 0.013 per haplotype per conditional generation (25 years), we have $65/12/.013 = 417 \rightarrow 619$ generations, that is $619 \times 25 = 15{,}475$ years to a common ancestor of the branch. The arrow shows a correction for back mutations. This correction can be calculated using a formula **(1)** as follows. Since the observed number of mutations per marker is $65/12/8 = 0.677$, we employ the formula **(1)** and obtain

$$1/2(1+\exp(0.677))\lambda_{obs}$$

The obtained number of 1.484 is the correction coefficient for back mutations. Therefore, by multiplying *(uncorrected number of generations)* 417×1.484, we obtain corrected number of generations 619, that is $619 \times 25 = 15{,}475$ years. This is usually depicted as $417 \rightarrow 618$ (generations). Since for 65 mutations in the dataset the margin of error is 15.93% (calculated as explained in Klyosov, 2009c), we at last obtain a timespan to a common ancestor of the haplotypes to be equal to $15{,}475\pm2500$ years.

102

Example 2: Two 67 marker base haplotypes of the subclades R-P312 and R-U106 differ by 5.5 mutations (see the text above). Applying the same rule explained in the preceding example, we get 5.5/.12 = 46 → 48 generations, that is 1200 years from a common ancestor of the two base haplotypes (i.e, of the two subclades). Should the two base haplotypes have the same "age" (the same TMRCA), their common ancestor would have lived 600 years "deeper" in time from either one of them. However, in this particular case the two TMRCA are equal to 4350 and 4175 years, respectively. Therefore their common ancestor lived (1200 + 4350 + 4175)/2 = 4850 years ago.

Conclusion

The results of this study lend a support to the theory that haplogroup R1b arose in Central Asia, apparently in South Siberia or the neighboring regions, around 16,000 years before present. The preceding history of the haplogroup is directly related to the appearance of the Europeoids (Caucasoids) ~58,000 ybp, likely in the vast triangle that stretched from Western Europe through the East European Plain to the east and to Levant to the south, as it was suggested in (Klyosov, 2011d). A succeeding sequence of SNP mutations in the Y chromosome in the course of their migration eastward to South Siberia, with the appearance of the haplogroups NOP ~48,000 ybp and P ~38,000 ybp, eventually gave rise to the haplogroup R ~30,000 ybp and R1 ~26,000 ybp, and then to the haplogroup R1a/R1a1 ~20,000 ybp (the timeframe between the appearances of R1a and R1a1 is uncertain) and R1b ~16,000 ybp (ibid.).

Based on the respective syllables, we call bearers of R1a the Aryans *(from Arans, same format as Arbins)*, and those of R1b the Arbins. In the first case that name is justified since the bearers of the R1a haplogroup became the legendary Aryans who arrived in the Hindustan and Iranian Plateau ~3500 ybp. In other words, those Aryans belonged to R1a haplogroup, hence, the double meaning (albeit coinciding) of the term the Aryans. The Arbins is a convenient common term to avoid repetition of the: "bearers of haplogroup R1b".

At some point in time, the Arbins began migration to the west, across Central Asia, North Kazakhstan, South Urals, to the East European Plain where they have established a number of archaeological cultures between 12,000 and 4500 ybp (apparently including Seroglazovo, Khvalyn, Samaran, Middle Volga, Timber Grave, Catacomb, and also "Proto-Kugran" and/or "Kurgan" cultures, the last are largely held controversial and are not accepted by many historians; it should be emphasized that all those above suggestions regarding the archaeological cultures at present can only be viewed as very tentative attributions) *(The reason for tentativeness is a not an absence of archaeological materials that tie these cultures genetically, but a complete void in the paleogenetic studies of these archeological cultures)*. "Arbins" migrated southward, around 6000 ybp a part migrated over the Caucasus to Anatolia (leaving their R1b haplogroup and the respective haplotypes behind); to the rest of the Asia Minor, and to the Middle East. Apparently, the Arbins have established the Sumer culture and state, and by several routes migrated westward to Europe, carrying mainly the R-M269 subclade and its downstream L23 subclade. One route was a northern route, from the East European Plain to the west, ~4600 - 4400 ybp; another concurrent route with the same two subclades was westward along the Asia Minor and Middle East; and yet another route that would populate the Europe the most, was migrating along the North Africa-Mediterranean Sea via the ancient Egypt to the Pyrenees, to arrive at ~4800 ybp. On that route, the R1b-V88 tribe split off and went south, eventually to Central Africa (mainly Cameroon and Chad, judging by their present-day distribution), where a common ancestor of the current R1b-V88 haplotype lived ~4400 ybp.

At the arrival time to Iberia ~4800 ybp, the M269 subclade split off the M51 clade, and soon thereafter the L11 clade and its downstream subclades. They became the Bell Beaker tribes and moved north along with newly arisen subclades P312 and L21, the latter arising within a few centuries after P312. Those subclades and their downstream clades effectively and without major interruptions have populated the Europe from the Atlantic to the Balkans, Carpathian Mountains, present day Poland, the western border of the East European Plain, to the Baltic Sea evidenced by the smooth haplotype trees testifying to the non-stop proliferation of the R1b haplotypes.

The Isles had a different history of their R1b haplotypes and lineages. The bearers of L11, P312 and L21 moved to the Isles by land and sea concurrently with those Arbins who were populating Europe between 4000 and 2500 ybp, and formed their respective "local" subclades, such as P314, M222, L226, which largely populated the Isles. As a result, a significant part of the Isles is populated almost exclusively by the Arbins whose allele frequency reaches 92% - 96% among the population. In general, the frequency of the Arbins in Western and central Europe reaches — albeit not uniformly — some 60% of the population.

This study essentially presents an example of the application of DNA Genealogy for studies on the history of mankind. This example is a complex and challenging endeavor which also touches upon some mysterious puzzles of history and linguistics. One of those puzzles is what language or languages were spoken by the Arbins from 16,000 to 3000 ybp and it almost certainly was a continuing, in its dynamics, a non-Indo-European language. Assumptively, these languages are considered by linguists as assorted and disconnected "dead" and not so dead languages, such as proto-Turkic, Sumer, North-Caucasian, Dene-Caucasian, Basque, and many pre-Indo-European languages in Europe of 5000 - 2000 ybp, and some later languages *(M.Gimbutas termed these languages "Old Europe", and ascribed them to the sedentary farming populations; unfortunately, she did not clearly perceive that the "Old Europe" constituted an amalgam of the pre-"Old Europe" hunter-gatherer and farming populace with the Asian hunter-gatherer migrants marked with R1a1 who reached the Balkans ca. 10-8 mill. BC and along with other innovations brought over their pre-Kurgan etiology with the Pra-Mama Goddess)*. The language of the Arbins may have originally been a single language easily flowing through millennia and across Eurasia. However, this is a subject of another study.

Acknowledgements

The author is indebted to Susan Hedeen for her valuable help with the preparation of the manuscript.

REFERENCES

Balanovsky, O., Dibirova, K., Dybo, A., Mudrak, O., Frolova, S., Pocheshkhova, E. et al. (2012). Parallel evolution of genes and languages in the Caucasus region. *Molecular Biology and Evolution, 29,* 359-365.
Balaresque, P., Bowden, G. R., Adams, S. M., Leung, H.-Y., King, T. E., Rosser, Z. H. et al. (2010). A predominantly Neolithic origin for European paternal lineages. *PLoS Biology, 8,* 1000285
Barac, L., Pericic, M., Klaric, I. M., Janicijevic, B., Parik, J., Rootsi, S. et al. (2003a). Y chromosome STRs in Croatians. *Forensic Science International, 138,* 127-133.
Barac, L., Pericic, M., Klaric, I. M., Rootsi, S., Janicijevic, B., Kivisild, T. et al. (2003b). Y chromosomal heritage of Croatian population and its island isolates. *European Journal of Human Genetics, 11,* 535- 542. doi:10.1038/sj.ejhg.5200992
Bengtson, J. D. (1997). The riddle of Sumerian: A Dene-Caucasian language? *Mother Tongue, 3,* 63-74.
Capelli, C., Brisighelli, F., Scarnicci, F., Arredi, B., Caglia, A., Vetrugno, G. et al. (2007). Y chromosome genetic variation in the Italian peninsula is clinal

and supports an admixture model for the Mesolithic-Neolithic encounter. *Molecular Phylogenetics and Evolution, 44,* 228-239. doi:10.1016/j.ympev.2006.11.030

Cardoso, J. L. (2001). Le phenomene campaniformedans les basses vallees du Tage et du Sado (Portugal). In F. Nicolis (Ed.), *Bell Beak-ers Today*, Trento: Ufficio Beni Archeologici.

103

Cardoso, J. L., & Soares, A. M. (1990). Chronologia absoluta para o campaniforme da Estremadura e do Sudoeste de Portugal. *O Arqueologo Portugues, 8-10,* 203-228.

Cinnioglu, C., King, R., Kivisild, T., Kalfoglu, E., Atasoy, S., Cavalleri, G. L. et al. (2004). Excavating Y-chromosome haplotype strata In Anatolia. *Human Genetics, 114,* 127-148. doi:10.1007/s00439-003-1031-4

Contu, D., Morelli, L., Santoni, F., Foster, J. W., Francalacci, P., Cucca, F. (2008). Y-Chromosome based evidence for pre-Neolithic origin of the genetically homogeneous but diverse Sardinian population: inference for association scans. *PLoS ONE, 3.* doi:10.1371/journal.pone

Cruciani, F., Trombetta, B., Sellitto, D., Massaia, A., Destro-Bisol, G., Watson, E. et al. (2010). Human Y chromosome haplogroup R-V88: a paternal genetic record of early mid Holocene trans-Saharan connections and the spread of Chadic languages. *European Journal of Human Genetics, 18,* 800-807. doi:10.1038/eihg.2009.231

Di Gaetano, C., Cerutti, N., Crobu, F., Robino, C., Inturri, S., Gino, S. et al. (2009). Differential Greek and northern African migrations to Sicily are supported by genetic evidence from the Y chromosome. *European Journal of Human Genetics, 17,* 91-99. doi:10.1038/ejhg.2008.120

Hammer, M. F., Behar, D. M., Karafet, T. M., Mendez, F. L., Hallmark, B., Erez, T. et al. (2009). Extended Y chromosome haplotypes re-solve multiple and unique lineages of the Jewish priesthood. *Human Genetics, 126,* 707-717. doi:10.1007/s00439-009-0727-5

Herrera, K. J., Lowery, R. K., Hadden, L., Calderon, S., Chiou, C., Yepiskoposyan, L. et al. (2011). Neolithic patrineal signals indicate that the Armenian plateau was repopulated by agriculturalists. *European Journal of Human Genetics, 20,* 313-320. doi:10.1038/ejhg.2011.192

Kang, L., Lu, Y., Wang, C., Hu, K., Chen, F., Liu, K. et al. (2011). Y-chromosome O3 haplogroup diversity in Sino-Tibetan populations reveals two migration routes into the Eastern Himalayas. *Annals of Human Genetics, 76,* 92-99. doi:10.1111/j.1469-1809.2011.00690

Klyosov, A. A. (2008a). Mysteries of the "Western European" haplogroup R1b. *Proceedings of the Russian Academy of DNA Genealogy, 1,* 568-630 (in Russian).

Klyosov, A. A. (2008b). Origin of the Jews via DNA genealogy. *Proceedings of the Russian Academy of DNA Genealogy, 1,* 54-232.

Klyosov, A. A. (2009a). A comment on the paper: Extended Y chromosome haplotypes resolve multiple and unique lineages of the Jewish priesthood. *Human Genetics, 126,* 719-724. doi:10.1007/s00439-009-0739-1

Klyosov, A. A. (2009b). Once again on the "population mutation rate" of L. Zhivotovsky, or how fables are born. *Proceedings of the Russian Academy of DNA Genealogy, 2,* 1162-1181 (in Russian).

Klyosov, A. A. (2009c). DNA Genealogy, mutation rates, and some historical evidences written in Y-chromosome. I. Basic principles and the method. *Journal of Genetic Genealogy, 5,* 186-216.

Klyosov, A. A. (2009d). DNA Genealogy, mutation rates, and some historical evidences written in Y-chromosome. II. Walking the map. *Journal of Genetic Genealogy, 5,* 217-256.

Klyosov, A. A. (2010a). Haplogroup R1b1 and its subclades in Asia. *Proceedings of the Russian Academy of DNA Genealogy, 3,* 1676 - 1695 (in Russian).

Klyosov, A. A. (2010b). Haplotypes of haplogroup R1b in central Africa. *Proceedings of the Russian Academy of DNA Genealogy, 3,* 369-378 (in Russian).

Klyosov, A. A. (2010c). Haplogroup R1b. Part 2. *Proceedings of the Russian Academy of DNA Genealogy, 3,* 406-475 (in Russian).

Klyosov, A. A. (2010d). The "age" of subclade R1b1a2-M269 and its subclades (L23, L51, L11). *Proceedings of the Russian Academy of DNA Genealogy, 3,* 1310-1315 (in Russian).

Klyosov, A. A. (2010e). Irish haplotypes and haplogroups. *Proceedings of the Russian Academy of DNA Genealogy, 3,* 1029-1053.

Klyosov, A. A. (2011a). Origin of ancient subclades of haplogroup R1b — Regions and times. *Proceedings of the Russian Academy of DNA Genealogy, 4,* 2227-2245 (in Russian).

Klyosov, A. A. (2011b). Haplotypes of R1b1a2-P312 and related subclades: Origin and "ages" of most recent common ancestors. *Proceedings of the Russian Academy of DNA Genealogy, 4,* 1127-1195.

Klyosov, A. A. (2011c). Haplogroups and haplotypes in Armenia (haplogroups J2, R1b-L23, R1b-M269, and T-M184). *Proceedings of the Russian Academy of DNA Genealogy, 4,* 1985-1993 (in Russian).

Klyosov, A. A. (2011d). "Out of Africa" theory re-examined. *Proceedings of the Russian Academy of DNA Genealogy, 4,* 1908-1977 (in Russian).

Klyosov, A. A., & Rozhanskii, I. L. (2012a). Haplogroup R1a as the Proto-Indo-Europeans and the legendary Aryans as witnessed by the DNA of their current descendants. *Advances in Anthropology, 2,* 1- 13. doi:10.4236/aa.2012.21001

Klyosov, A. A., & Rozhanskii, I. L. (2012b). Re-examining the "out of Africa" theory and the origin of Europeoids (Caucasoids) in light of DNA genealogy. *Advances in Anthropology,* 2, in press.

Kramer, S. N. (1971). *The Sumerians: Their history, culture, and character*. Chicago, IL: University of Chicago Press.

Lashgary, Z., Khodadadi, A., Singh, Y., Houshmand, S. M., Mahjoubi, F., Sharma, P. et al. (2011). Y chromosome diversity among the Iranian religious groups: A reservoir of genetic variations. *Annals of Human Biology, 38,* 364-372.

Lobov, A. S. (2009). Structure of the gene pool of the Bashkir subpopulations. Ufa. http://ftp.anrb.ru/molgen/Lobov_AS.PDF

Mallory, J. P. (1989). *In search of the indo-Europeans: Language, archaeology and myth*. London: Thames and Hudson.

Malyarchuk, B., Derenko, M., Denisova, G., Maksimov, A., Wozniak, M., Grzybowski, T. et al. (2011). Ancient links between Siberians and native Americans revealed by sub-typing the Y chromosome haplogroup Q1a. *Journal of Human Genetics, 56,* 583-588

Martinez, P. V., Lull, V., & Mico, R. (1996). *El vaso campaniforme, in: Cronologia de la Prehistoria Reciente de la Peninsula Iberica y Baleares (c. 2800-900 cal. A.N.E.).* British Archaeological Reports, International Series 652, 105-110.

Morelli, L., Contu, D., Santoni, F., Whalen, M., Francalacci, P., & Cucca, F. (2010). A comparison of Y-chromosome variation in Sardinia and Anatolia is more consistent with cultural rather than demic diffusion of agriculture. *PLoS ONE, 5,* e10419, doi:10.1371/journal.pone.0010419

Muller, J., & Van Willigen, S. (2001). New radiocarbon evidence for European Bell Beakers and the consequences for the diffusion of the Bell Beaker phenomenon. In F. Nicolis (Ed.), *Bell Beakers Today*, Trento: Ufficio Beni Archeologici.

Myres, N. M., Rootsi, S., Lin, A. A., Jarve, M., King, R. J., Kutuev, I. et al. (2010). A major Y-chromosome haplogroup R1b holocene era founder effect in central and western Europe. *European Journal of Human Genetics, 19,* 95-101.

Nocete, F. (2006). The first specialized copper industry in the Iberian Peninsula: Cabezo Jure (2900-2200 BC). *Antiquity, 80,* 646-657.

Pericic, M., Lauc, L. B., Klaric, A. M. et al. (2005). High-resolution phylogenetic analysis of southeastern Europe traces major episodes of paternal gene flow among Slavic populations. *Molecular Biology and Evolution, 22,* 1964-1975. doi:10.1093/molbev/msi185

Roewer L., Willuweit S., Kruger C., Nagy M., Rychkov S., Morozowa I. et al. (2008). Analysis of Y chromosome STR haplotypes in the European part of Russia reveals high diversities but non-significant genetic distances between populations. *International Journal of Legal Medicine, 122,* 219-223.

doi:10.1007/s00414-007-0222-2

Rozhanskii, I. L., & Klyosov, A. A. (2011). Mutation rate constants in DNA genealogy (Y chromosome). *Advances in Anthropology, 1,* 26-34. doi:10.4236/aa.2011.12005

Semino, O., Passarino, G., Oefner, P. J., Lin, A. A., Arbuzova, S., Beckman, L. E. et al. (2000). The genetic legacy of paleolithic *Homo sapiens sapiens* in extant Europeans: A Y chromosome perspective. *Science, 290,* 1155-1159. doi:10.1126/science.290.5494.1155

Underhill, P. A., Myres, N. M., Rootsi, S., Metspalu, M., Zhivotovsky, M. A., King, R. J. et al. (2009). Separating the post-Glacial coancestry of European and Asian Y chromosomes within haplogroup R1a. *European Journal of Human Genetics, 18,* 479-484.

105

Weale, M. E., Yepiskoposyan, L., Jager, R. F., Hovhannisyan, N., Khudoyan, A., Barbage-Hall, O. et al. (2001). Armenian Y chromosome haplotypes reveal strong regional structure within a single ethnonational group. *Human Genetics, 109,* 659-674. doi:10.1007/s00439-001-0627-9

Wells, R. S., Yuldasheva, N., Ruzibakiev, R., Underhill, P. A., Evseeva, I., Blue-Smith, J. et al (2001). The Eurasian heartland: a continental perspective on Y-chromosome diversity. *Proceedings of the National Academy of Sciences USA, 98,* 10244-10249. doi:10.1073/pnas.171305098

Wiik, K. (2008). Where did European men come from? *Journal of Genetic Genealogy, 4,* 35-85.

Zakharov, S. V. (2010). On the origin of Botai culture. *Vestnik of Archaeology, Anthropology, and Ethnography, 1,* 49-58.

Zalloua, P. A., Xue, Y., Khalife, J., Makhoul, N., Debiane, L., Platt, D. E. et al. (2008). Y-Chromosomal diversity in Lebanon is structured by recent historical events. *American Journal of Human Genetics, 82,* 973-882. doi:10.1016/j.ajhg.2008.01.020

Zhivotovsky, L. A., Underhill, P. A., Cinnioglu, C., Kayser, M., Morar, B., Kivisild, T. et al. (2004). The effective mutation rate at Y-chromosome short tandem repeats, with application to human population divergence time. *American Journal of Human Genetics, 74,* 50-61. doi:10.1086/380911

Zhong, H., Shi, H., Qi, X.-B., Duan, Z.-Y., Tan, P.-P., Jin, L. et al. (2010). Extended Y-chromosome investigation suggests post-Glacial migrations of modern humans into East Asia via the northern route. *Molecular Biology and Evolution, 28,* 717-727.

APPENDIX

The following DNA projects were selected as primary haplotype databases:
http://www.familytreedna.com/public/R1b1b1/default.aspx?section=yresults
http://www.familytreedna.com/public/ht35new/default.aspx
http://www.familytreedna.com/public/AssyrianHeritageDNAProject/default.aspx?section=yresults
http://www.familytreedna.com/public/IrelandHeritage/default.aspx?section=yresults
http://www.familytreedna.com/public/R-L21/default.aspx?section=yresults

Reference data were selected according to SNP assignment from YSearch database: (http://www.ysearch.org) and public projects of FTDNA (http://www.familytreedna.com)