Ancient Paleo-DNA of Pre-Copper Age North-Eastern Europe: Establishing the Migration Traces of R1a1 Y-DNA Haplogroup Part 2. Baikal Episode and Indo-Uralic Framework

Alexander S. Semenov*, Vladimir V. Bulat

* Deep Dive Research Group, Russian Federation

Abstract

The work considers the problems of paleogenetics and anthropology connected with pre-Copper Age after-Glacial repopulation process of the North-Eastern Europe. The unified data, obtained in various laboratories in 2010–2016, collects a certain amount of the ancient mt-DNA and Y-DNA haplogroup samples of the considered period, what allows establishing the connection between some of them. The first part of the paper (Semenov, Bulat, 2016) showed the possible connection of R1a1 Y-haplogroup dispersion in Neolithic time with Comb Ware pottery. The second part shows that the recently found R1a1’s in Neolithic Baikal cultures also fit the hypothesis and can go along with Indo-Uralic language superfamily hypothesis.

The paper makes an attempt to build a picture of the population of North-Eastern Europe in pre-Copper Age time and to systemize the pale DNA genotyping results into clusters corresponding to different migration waves. The paper can be of use in biomedical purposes also, as some correlations between diseases and haplogroups were noticed in different medical works.

Keywords: Y-DNA haplogroup, R1a1, J2b, mtDNA haplogroups U4 and U5a1, Yuzniy Oleni Ostrov, Khvalynsk, Serteya, Kitoi, Indo-Uralic, paleogenetics, paleolinguistics, subclades.

1. Introduction

The interest in the origin and early localization of carriers of Y-DNA haplogroup R1a1 is serious, since this Y-DNA subclade is inherent to the significant percentage of the population of Central and Eastern Europe, India, Middle East. It is widely recognized and already proven in terms of archeology and paleogenetics that a significant concentration of R1a1 Y-DNA haplogroup was inherent to the population of European Corded Ware culture (authors note that there also existed less famous East Asian Corded Ware culture group (Semenov, Bulat, 2015). However, the location of R1a1 bearers of Pre-Corded Ware horizons causes debates due to a lack of data. However, over the last year such data emerged, allowing formulating a data-based hypothesis. New data from Kitoi also go along with it.

* Corresponding author
E-mail addresses: semyonov1980@mail.ru (A.S. Semenov)
2. Materials and Methods

The main materials for the research are data from paleogenetic samples described in other works, and catalogized in the aggregated list (Ancestral Journeys).

Table 1. R1a1 in pre-Corded Ware sites

<table>
<thead>
<tr>
<th>Sample</th>
<th>Y-DNA from the site</th>
<th>Mitochondrial DNA from the site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yuzniy Oleni Ostrov burial № 125, 5500-5000 BCE. Indirect indicators of comb ware possible presence.</td>
<td>R1a1*-M459+, M198-, J</td>
<td>C1g (formerly C1f), U4, U2e, U5a, H</td>
</tr>
<tr>
<td>Serteya archeological site, middle of V-IV mill. BCE (Comb ware presence)</td>
<td>R1a1-M17</td>
<td>H2</td>
</tr>
<tr>
<td>Khvalynsk-II burial 5200-4000 BCE. (Comb ware presence in preceding layers, partial continuity of population)</td>
<td>R1a1*-M459+, M198-, R1b1, Q</td>
<td>U5a1i, U4, H2</td>
</tr>
<tr>
<td>Early Neolithic Kitoi culture burials (Baikal, Irkutsk area) (Comb ware presence, reference later here) 6000-4500 BCE (Moussa, 2015; Moussa et al, 2016)</td>
<td>R1a1-M17, C3-M217, K-M9</td>
<td>U5a, A, C, D,F</td>
</tr>
<tr>
<td>Satsurblia burial (Georgia), Upper Paleolithic</td>
<td>J</td>
<td>K3</td>
</tr>
<tr>
<td>Kotias burial (Georgia), Mesolithic</td>
<td>J2a</td>
<td>H13c</td>
</tr>
</tbody>
</table>

The main research method of this paper is the interpretation of recently obtained genetic data, which are compared with archaeological cultures distribution.

3. Discussion and Results

In the previous paper we conclude that the presence of haplogroup R1a1 is strongly probable in the cultures of Comb Stamp Ware in the Neolithic context of Eastern Europe, but though the origin and the trace of R1a1 migration is still unclarified and more data are needed. This year the unexpected and striking data from Baikal were obtained, and, as we show, that is in accordance with our previous results and, moreover, that gives some insights on early history of Y-DNA haplogroup R1a1 bearers. So the new Baikal findings should be examined according to our view.

We base our previous findings on the D. L. Gaskevych’s view on the origins of Comb Ware. In his long article «North Pontic Impresso: the origin of the Neolithic Pottery with Comb Decoration in the South Eastern Europe» In (Gaskevych, 2010: 246-247) he proposed the origin of this type of pottery in the northern Black Sea coast. «The absolute data collected over the last 15 years in Kiev Radiocarbon Laboratory, have revealed that such ware appeared in the North-Pontic region earlier than in Upper Dnieper, Volga Region, Kama basin, Trans-Urals. However, in the steppe Pontic region it appeared earlier than in forest-steppe. All these data have proved unreliability of abovementioned hypothesis. As an alternative, the author suggests considering the Pontic region Neolithic area with comb ceramic ornamentation as a part of Neolithic cultures with Impresso ware from the Mediterranean region»

This way, according to D. Gaskevych, we see in the Eastern Europe only a small episode of a big process, which took place from Sahara to Trans-Urals and from Morrocco to the Levant. As D.L. Gaskevich refers to the initial spread of the Neolithic within the Eastern Europe, we should consider the issue of the genetic reflection of this process and specify the genetic map of the North Black Sea Region and the adjust territories in the period, preceding the Mediterranean ware adoption (though, there is another possible address of this initial spread, namely Elshanskaya culture and its derivates up to the Crimea), and the earliest time of intrusion was 6500 BC.
Gaskevych emphasizes the pre-Neolithic Bug-Dniester culture as the transmitter of Comb tradition. In our paper (Semenov, Bulat, 2016:46) we noticed, that this culture was developed on the basis of the Kukrek one. The latter belongs to Epi-Gravettian tradition and the Epi-Gravettian tradition show the presence of J (Satsurblia and Kotias burial grounds) and R1b1 (Villabruna) (Ancestral Journeys).

In our previous paper we discussed the traces of J in Native Eastern European populations (especially Finno-Ugric) and the possible Comb Cultures (Karelia) and conclude that they may have Epi-Gravettian ancestors. And if R1a1 went throughout Eastern Europe Comb Ware areas (as findings of paleoDNA show) from the Epi-Gravettian area, it can be also present in the Epi-Gravettian itself, maybe somewhere near Bug-Dniester area.

First step of this research is to see more precise origins of Corded Ware R1a1 population. According to (Kozintsev, 2016:2) «Traditional marker of Indo-European migrations is the Corded ornament on vessels. As it is presumed, this tradition started in Dereivka culture of the end of V-beginning of IV mil BC (around 4000 BC), and then entered to Cucuteni-Tripolye areal and more to the South. The most ancient variant of the corded ornament, the caterpillar, is seen on Cucuteni C ceramics. The first its appearance in Balkans is the same, the early Eneolithic. But at first it is met rarely and the wide dispersal of it starts from the end of IV mil. BC». We notice, that Serteya’s pre-Corded Ware R1a1 belongs also to this time period.

According to (Genofond Thesaurus), the ceramics of Dereivka is hand-made, with the addition of the shells. It has comb ornament, or also stroked, pearled and pitted. Also some corded ware is noticed and the similarity with Funnel Beakers of Germany and Poland by the geometrical shape is also observable (Genofond Thesaurus). So, we see that Dereivka may be the part of the initial area of spread of «Indo-European» subclades of R1a1 (maybe R1a1-M417 and downstream) and has connections and relations with Funnel Beakers. It should be noted that one stage of Serteya culture development is also viewed as the Eastern Branch of Funnel Beakers. Hence, the area of initial spread of R1a1-M417 could be the wide area from Central Europe (Funnel Beakers) to the Black Sea, what correlates with the initial spread of Corded Ware in Early Bronze Age (Fig. 2).

**Fig. 1.** The time of arrival of the Cardial ceramics, BC (Gaskevych, 2010: 239)
Fig. 2. The initial spread of Corded Ware

So, the new Baikal findings are to be examined according to this starting view, and we here show that they may fit well to the picture.

Fig. 3. The place of Dereivka culture, 4500–4000 BC (Genofond Thesaurus)

Of course, the broad range of dating Kitoi’s R1a1’s (6000–4500 BC) gives certain ambiguity. But we will try to show that the scenario of western migration of Kitoi R1a1 bearers is possible and
fits well with our previous conclusions (Semenov, Bulat, 2016: 51). The first argument of the Western migration to Kitoi area is the Comb Ware character of Kitoi culture (Oxford Archaeology, 1996: 644). The second and stronger is the presence of U5a mtDNA haplogroup, which is found in Paleo-, Meso and Neolithic of Europe to Urals and West Siberia (we describe it later in this paper).

As the possible Kitoi dating is earlier than Serteya-Dereivka, we should look what happened earlier on the Eastern European plain in the epoch prior to Indo-European (4500 BC). In (Kozintsev, 2016: 1, 3) it is mentioned, that the spilt of initial Indo-European had happened in the V mil. BC. (the precise time of Dereivka). So, to understand what happened earlier, we should use the hypothesis of the possible ancestors of the Indo-European language as tools. Of course, Indo-European is considered as a Nostratic branch, but this superfamly is too wide, so we should use narrower grouping. One of that is Indo-Uralic framework (sub-family, or the language union). It is not totally accepted, but starts becoming a working tool for prominent linguists. The Academician of RAS Vyach.Vs. Ivanov, in his analysis of the names of seasons mentions Uralic materials and Indo-Uralic correspondences (Ivanov, 2006: 1-5). In (Kloekhorst, 2008: 94) it is shown that unique for Hittite languages merger the roots with possessive pronouns has the parallels in Uralics. According (Napolskikh, 2015: 4) the etymology for the Indo-European numeral ‘4’ may be viewed as an extension of root ‘2’ present in Uralic languages.

Let us regard the recent works with dating of language splits. The Dutch linguist Frederik Kortlandt supports a model of Indo-Uralic in which the original Indo-Uralic speakers lived north of the Caspian Sea, and the Proto-Indo-European speakers began as a group that branched off westward from there to come into geographic proximity with the Northwest Caucasian languages, absorbing a Northwest Caucasian lexical blending before moving farther westward to a region north of the Black Sea where their language settled into canonical Proto-Indo-European (2002). Allan Bomhard suggests a similar schema in Indo-European and the Nostratic Hypothesis (1996). Alternatively, the common protolanguage may have been located North of the Black Sea, with Proto-Uralic moving northwards with the climatic improvement of post-glacial times. Continuing this, in the work in (Romanchuk, Semenov, 2014; Klyosov, 2014) it was proposed, that the bearers of Y-DNA haplogroup R used to speak Sino-Caucasian languages, and some of them (including R1a1 bearers) developed proto-Indo-European language later, in the process of migrations.

According to (Kortlandt, 2002: 1) «C.C. Uhlenbeck made a distinction between two components of Proto-Indo-European, which he called A and B. The first component comprises pronouns, verbal roots, and derivational suffixes, and may be compared with Uralic, whereas the second component contains isolated words, such as numerals and most underived nouns, which have a different source. Though Uhlenbeck objects to the term “substratum” for his B complex, I think that it is a perfectly appropriate denomination». So, dealing with two substrata, the first is can be «Indo-Uralic», and the second could be «Sino-Caucasic».

Also Kortlandt’s link Indo-Uralic and Altaic (Kortlandt, 2004: 4) show the following dating, «We may conclude that Proto-Indo-Uralic and Proto-Altaic may have been contemporaries (6000−5500 BC), that Proto-Uralic and Proto-Uralo-Yukagir may have been the same thing and contemporaneous with Proto-Indo-Hittite (4500−4000 BC), and that Proto-Finno-Ugric and nuclear Proto-Indo-European may again have been contemporary languages (3500−3000 BC). This puts the dissolution of the Uralo-Siberian language family in the 7th millennium. It now becomes attractive to identify the latter with the abrupt climate change of 8200 BP or 6200 BC, when severe cold struck the northern hemisphere for more than a century. The catastrophic nature of this disastrous event agrees well with the sudden dispersal and large-scale lexical replacement which are characteristic of the Uralo-Siberian languages».

Our previous hypothesis based on Gaskevych’s works, and the Indo-Uralic view of Kortlandt and Bomhard can be matched together and linked with Kitoi results according to the old work of A.Kh. Khalikov (Khalikov, 1967: 27-35). The possibility of application of Indo Uralic-Theory is the findings of Khalikov that Bug-Dniestr culture influenced eastwards. It should be mentioned that Khalikov applies it to Uralic only and the his dating is as latter as possible. He also mentions Indo-Uralic hypothesis, but in earlier and very careful forms, preceding Kortlandt’s. And we will show, that it matches well with current mtDNA results.

In (Khalikov, 1967: 27-35) the following scheme of the genesis of Uralic tribes genetics was outlined. Before IV mil. BC the Proto-Uralic tribes with stroked pottery lived in the territories of Volga and Kama, both sides of Urals, from Oka to Ob (and even more eastwards). In cultural and
ethnic point of view they were close related to the tribes of Dnieper-Donets culture, who were either the branch of Uralic or some extinct branch between Indo-European and Finno-Ugric. Khalikov also mentions the remark of G.F. Debets who noticed the Siberian influence on Dnieper-Donets burials. He also mentions that the ceramics looking similarly to Dnieper-Donets was found near Yenisei.

Khalikov also stresses that the Uralic community in IV mil. BC was intersected near Ural into two parts by the tribes similar to Kelteminar culture bearers. Then the Samodian tribes moved eastwards to Yenisei and Finno-Ugric stayed. Then proto-Saami may move northwards to the West.

The more distant roots of Finno-Ugric and Uralic Khalikov places into the mesolithic community in Volga-Kama and Ural regions connected with both Eastern Europe and Siberia. Khalikov avoids the precise dating but the framework supposes VI–V mil. BC. Now we know the mtDNA from Uralic (geographically) burials Chekalino, Lebyazinka IV, and they are predominantly U5a, as some of Kitoi’s (the reference is later in this paper).

He mentions that the basics of this culture have the roots in Siberia, but also shows the contacts between the proto-Uralic tribes and the ancestors of Indo-Europeans, and he mentions the early Indo-Uralic theories directly. The connection between Indo-Europeans and Proto-Uralics can be traced by the inclusion the Western microlithic technique in Uralic cultures. So, he points on two substrates in Uralic cultures – Paleo-Siberian and Paleo-European.

Khalikov also mentions that the tribes from Volga and Kama (where he notices Comb pottery) and Ural borrowed the production of pottery from the South (maybe only some of the technologies, as it is seen now) via Dnieper-Donets culture. And the latter borrowed the pottery from Bug-Dniester culture, which, as Khalikov mentions, could be one of the basic of Indo-Europeans (and it was itself influenced by Starchevo-Krish culture).

Now we know the mtDNA haplogroups from Dnieper-Donets circle of cultures and they include U5a (as unique western mtDNA haplogroup present in Kitoi). Though Khalikov mentions stroked pottery only, in fact it can be both comb and stroke as it happens in Dereivka, and we saw the presence of the comb pottery in Kitoi.

The continuity of the space from Baikal to Black Sea in Meso- and Neolithic may also result in possible transfer of Corded Ware techniques to the West (Semenov, Bulat, 2015), possibly with the bearers of mtDNA haplogroup C, as we show earlier (C is present in Dnieper-Donets circle).

So, Khalikov’s movement of Uralic people to the East could be that of Kitoi’s R1a1 bearers. Also, if we admit the scenario that it could be two waves, one of proto-Altaic and another of Uralic, Kitoi’s also could be the first, Altaic wave (the scenario that they are the ancestor of modern Altaic people is mentioned in (Turov, 2003), but there is also shown some arguments against). We stress that (Turov, 2003) also cites G. F. Debets, who admitted that Kitoi burials of the Neolithic were differentiated anthropologically and possessed a substantial European admixture. So their bearers could not be the direct ancestors of Tungusic people. This gives an additional argument for the Uralic migration version.

It is interesting to notice that though S. Starostin gives the common etymology for pot, vessel to all Borean languages we see that the most reliable etymologies are in Indo-European, Uralic, Altaic, and Sino-Caucasian (Starling). (Austric analog mentioned by Starostin also can be explained by ‘Boreal’ wave of creamics moving westwards from the Far East). So that can be cultural borrowing.

- Indo-European: *pod-
- Altaic: *p’a’di
- Uralic: *pata
- Proto-Sino-Caucasian: *phātV

So that gives another argument that Altaic can be the first wave on Bug-Dniester influence eastwards. The common roots for ‘pot’ in Altaic and Uralic can be the sigh that the Khalikov’s scenario can take place in two waves. We see that first appearance of Cordial pottery in Black Sea cost happened is in 7 mill BC, just before Kortlandt’s Indo-Uralo-Altaic separated, and both Gaskevych and Khalikov show the migration of pottery cultures deeply eastwards.
Table 2. Early Neolithic of Eurasia (IV mil. BC according to Khalikov, 1967: 33)'

<table>
<thead>
<tr>
<th>Culture (name according to Khalikov, 1967)</th>
<th>Known Y-DNA</th>
<th>Known mtDNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Volga-Kama (Uralic)</td>
<td>R1b1 (Lebyazhinka-IV)</td>
<td>U5a (Lebyazhinka-IV)</td>
</tr>
<tr>
<td>1a. Moving Eastwards (Samoedic)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1b. Sperrings (moving westwards)</td>
<td>R1a1* - M459+, M198-, J (if Yuzhni Oleni Ostrov is part of it or direct ancestor, question is risen in Semenov, Bulat, 2016:43)</td>
<td>C1g (formely C1f), U4, U2e, U5a, H (the same)</td>
</tr>
<tr>
<td>2. Kelteminar</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Dniepro-Donets</td>
<td></td>
<td>U5a1a C, C4a, H, U3 (including Dereivka)</td>
</tr>
<tr>
<td>4. Early Pit-Comb (Volga-Oka)</td>
<td>R1a1, N1c1 (outflow to Zhizhitsa, Smolensk area)</td>
<td>H2</td>
</tr>
<tr>
<td>5. Dassang, Rakushechny Yar</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Narva</td>
<td></td>
<td>U5a</td>
</tr>
</tbody>
</table>

Of course, the bearers of Proto-Indo-Uralic-Altaic language can also have R1a1 and R1b taken from the Epi-Gravettian. And our main conjecture is that if the Indo-Uralic language family or commonwealth existed and dissolved along with Khalikov’s scenario, it can be correlated by the U5a mtDNA dispersion from Central Europe to Kitoi. As we see C in Dnieper-Donets and R1a1 and U5a in Kitoi we can make a hypothesis on connectivity (at least, existence of connections) between

---

* Some references of findings are hypothetical and shows the opinion of authors according to time and geographical proximity and archeological data.
the areas of Dnieper to Baikal. As G.F. Debets notices Siberian admixture in Dneper-Donets, and European in Kitoi, that looks possible.

The modern distribution of U5a mtDNA haplogroup

The modern distribution of U5a fits well with this idea, as it has vector to the East and goes through areals of Uralic people (Gentis).

U5 is the mitochondrial group, typical to the Upper-Paleolithic cultures like Gravettian one (Fu et al., 2013) and the dispersal routes of U5a bearers across Europe are concerned with the Central Europe (U5b prevailed in the west of Europe in Paleolithic-Mesolithic), rather than with the Western one. U5a seems to mark some migration (unknown to the archeologists) of the final Paleolithic or early Mesolithic from the Danube (possibly where Dolni-Vestonice site was located, although this location dates back to the XXV millennium BC) (ibid) to Central Germany, where the route of the U5a bearers divides into two subclades U5a1 and U5a2.

Haplogroup U5a1 is the most probable for the U5a bearer in Yuzniy Oleniy Ostrov. The bearers of U5a spread far beyond Scandinavia and Baltics. Bearers of this haplogroup can be found in Chekalino culture about 7800 BC (Bramanti et al., 2009), Lebyazhinka IV culture of the VI millennium BC (ibid, Ancestral Journeys) at the Volga and Lokomotiv burial ground of the Kitoi culture at the Angara (6100–4900 BC) (Moore et al., 2006; Moussa, 2015; Moussa et al., 2016; Ancestral Journeys). The latter enables to affirm that the Kitoi culture, which some researchers consider as the ancestral for the Altaic-language tribes or at least for the part of them (Turov, 2003), or the Uralic-language tribes, may be connected with some Mesolithic or Neolithic migrations from the more westward regions. The analysis of the burial grounds of Bolshoi Oleniy Ostrov (Sarkissian et al., 2013: 4), Lebyazhinka IV and the later cultures shows the existence of subclade U5a1. The work (Sarkissian et al., 2013: 3) shows similarity of populations of Yuzniy and Bolshoi Oleniy Ostrov. This haplogroup was also found in Mesolithic Sweden (Motala burial). The latter two arguments make the existence of U5a1 in Yuzniy Oleniy Ostrov the most possible.

Swedish burial ground Motala contains burials with a bunch of mitochondrial DNA: U2e, U5a1, U5a2, U5a2d (Ancestral Journeys; Lazaridis et al., 2013). The fixed subclade U5a2 in Les Closeaux (Rueil-Malmaison) location, dated back to 8870 BC (Posth et al., 2016) and German subclade U5a2c3 in Blätterhöhle, dated back to 8638 BC (Bolongino et al., 2013) could be considered as the similar for Motala subclades. It is possible to consider that U5a subclades were brought from the Western or Central Europe in the period, preceding to Yuzniy Oleniy Ostrov burial and the center of it dispersal was continental Western-Central Europe, and then bearers of U5a mover far to the East.
So, according to wide distribution of U5a throughout Mesolithic and Early Neolithic cultures of Eurasia we can make a conjecture that the initial bearers of Indo-European-Uralic-Altaiic unity of 7th millennium BC included the bearers of mtDNA U5a, and in Early Neolithic Time R1a1-M17 could be present in proto-Altaiic, and proto-Uralic populations, as well as pre-Indo-European. Nevertheless, R1a1-M17 in Proto-Altaiic and proto-Uralic populations may reflect the influences from «Old Europe». Also, Uralic-like sister languages could be also present in Meso- and Neolithic Europe, where U5a mitochondrial haplogroup was detected.

4. Conclusion

We can conclude that the Early Neolithic time is characterized by the strong connection of Bug-Dniester and Dnieper-Don culture. The latter was connected with both cultures of European Neolithic and proto-Uralic cultures of Easter Europe and Urals (and, very probably, Altaic). This continuum existed since the late Mesolithic and can be characterized by mtDNA haplogroup U5a presence. Comb Ware cultures could be considered the culture of Indo-Uralic language community. Indo-Uralic community also could have sister branches in Europe, in areas, where U5a was present.

According to F. Kortlandt, the dual nature of Indo-European language can be explained by the mix of the Epi-Gravett Black Sea inhabitants initially speaking Sino-Caucasian dialects (maybe some of them connected with Zarzian Groups) and having R1a1, possibly R1b1, J Y-DNA haplogroups and Paleolithic European populations having U5a mtDNA haplogroup. Y-haplogroup R1a1-M17 in Proto-Altaiic and proto-Uralic populations may reflect the influences from «Old Europe» and Black Sea.

References

Ancestral Journeys – Aggregator http://www.ancestraljourneys.org


Genofond Thesaurus – www.репонфонд.рф. The Thesaurus of Archaeological Cultures http://xn--ciaaveaafa1c.xn--p1ai/?page_id=9312

Gentis – The Mitochondrial Haplogroup Description http://gentis.ru/info/mtdna-tutorial/hg-u/u5a


Posth et al., 2016 – Posth, C. et al. (2016). Pleistocene Mitochondrial Genomes Suggest a Single Major dispersal of non-Africans and a Late Glacial population turnover in Europe. DOI: http://dx.doi.org/10.1016/j.cub.2016.01.03


Starling – Etymological Database starling.rinet.ru