

Paper 5.17: Report for Haplogroup R.

Abstract: Haplogroup R-M207 is found across Western Eurasia and in North Africa. Several genetic studies support the *steppe hypothesis* of Indo-European origins based on palaeogenomic modeling. This paper delivers a *de novo* review of the evolutionary history of the R-M207 haplogroup to assess the strength of their arguments. According to a triangulated Y-chromosome perspective, palaeogenomic modeling is not reliable. The *early farming dispersal hypothesis* still remains a very robust model for Afro-Asiatic and Indo-European. Before the arrival of Indo-European speaking farmers, R-M207 mutations were already present in South Asia and Europe. Similarly, R-M207 predates the arrival of Afro-Asiatic speaking farmers in North Africa. This suggests that language contact played an important role in the initial expansion of both language families.



Posted: October 15, 2019
Genetic-Linguistic Interface
Dr. Michael St. Clair
Stuttgart, Germany
mstclair@genlinginterface.com

Please cite as: St. Clair, Michael R. 2019c. “Paper 5.17: haplogroup R-M207.” *The Genetic-Linguistic Interface*. Series 5 Papers: language prehistory from a Y-Chromosome perspective. Posted October 15, 2019. <https://genlinginterface.com/>

Table of Contents for Paper 5.17 (Haplogroup R)

SECTION	PAGE
Section 1. Problematic Palaeogenomic Modeling of Indo-European Languages.	1
Section 2. The Contemporary Distribution of Haplogroup R-M207: an Overview.	2
Section 3. The Evolution of Haplogroup R-M207.	2
Section 4. The Expansion and Diversification of R1-M173.	3
Section 5. Diversification of R1b-M343 on the East European Plain.	4
Section 6. The Expansion of R1b-M269 into Western Europe.	4
Section 7. The Expansion of R1b-V88 into Africa.	6
Section 8. The DNA of Steppe Nomads.	7
Section 9. Diversification of R1a-M420 on the East European Plain.	7
Section 10. The Evolutionary History of R2-M479.	8
Section 11. Conclusions for R-M207.	9
Bibliography for Report and Tables	9
Figure 5.17.1: Part A. Initial Diversification of Haplogroup R-M207.	19
Figure 5.17.1: Part B. Linguistically Informative R1a-M420 Mutations.	20
Figure 5.17.1: Part C. Linguistically Informative R1b-M343 Mutations.	21
Figure 5.17.2. The East European Plain.	22
Figure 5.17.3. The North European Plain.	23
Important Comments Concerning Tables.	24
Table 5.17.1. Distribution of Mesolithic R1b-M343 Mutations.	25
Table 5.17.2. Distribution of Mesolithic R1a-M420 Mutations.	26
Table 5.17.3. Contemporary Distribution of R1b1a1a1-M73 (Regional View).	27
Table 5.17.4. Contemporary Distribution of R1b1a1a2a1a1-U106 (Language View).	28
Table 5.17.5. Contemporary Distribution of R1b1a1a2a1a2a -DF27 (Language View).	29
Table 5.17.6. Contemporary Distribution of R1b1a1a2a1a2b-U152 (Language View).	30
Table 5.17.7. Contemporary Distribution of R1b1a1a2a1a2c1-M529 (Language View).	31
Table 5.17.8. Contemporary Distribution of R1b1a2-V88 (Regional View).	32
Table 5.17.9. Contemporary Distribution of R1b1a2-V88 (Language View).	37
Table 5.17.10. R1b1a1a2a2-CTS1078 aDNA.	42
Table 5.17.11. Contemporary distribution of R1b1a1a2a2-CTS1078 (Regional View).	44
Table 5.17.12. Contemporary Distribution of R1a1a1b1a-Z282 (Regional View).	48
Table 5.17.13. Contemporary Distribution of R1a1a1b2-Z93 (Regional View).	50
Table 5.17.14. Contemporary Distribution of R1a1a1b1a1-M458 (Language View).	57
Table 5.17.15. Contemporary Distribution of R1a1a1b1a2b-M558 (Language View).	58
Table 5.17.16. Contemporary Distribution of R1a1a1b2- Z93 (Language View).	60
Table 5.17.17. Contemporary distribution of R2a-M124 (Regional View).	65
Table 5.17.18. R2a-M124 aDNA.	68
Table 5.17.19. Contemporary Distribution of R2a-M124 (Language View).	69



Paper 5.17. Haplogroup R-M207.

The Genetic-Linguistic Interface.

Section 1. Problematic Palaeogenomic Modeling of Indo-European Languages.

“Palaeogenomic modeling” often appears in whole genome studies. This represents an attempt to model human population history by employing statistical analysis of ancient DNA, and more specifically, autosomal DNA markers that are inherited from both parents. Statistical methods are utilized in order to overcome the problem of recombination. This “reshuffling” of genetic traits can distort and erase evolutionary relations that are needed to decipher human population history. Of course, Y-chromosome data avoid this problem as they are gathered from a non-recombining region of the human genome (see Paper 5.1, Intro, Section 3).

We also need to discuss models that attempt to decipher the origin and expansion of the Indo-European language family, a topic that was initially presented in Paper 5.7, Hg. G. The so-called “steppe nomad” hypothesis associates the spread of Indo-European with a Bronze Age expansion of steppe nomads from Eastern Europe or Central Asia. It should be emphasized that this approach is archeologically weak (e.g. Renfrew 2007; Frachetti 2012). Rather, the hypothesis is largely based on phonological reconstructions (for a recent discussion, see Anthony and Ringe 2015).

The alternate model of Indo-European origins, the *early farming dispersal hypothesis*, suggests that this language family evolved in the Middle East and expanded out of the region roughly nine thousand years ago during the Neolithic. This hypothesis was proposed by the archaeologist Peter Bellwood in his 2005 book *First Farmers: the Origins of Agricultural Societies*. Bellwood supports his model with archeological evidence. Additional support comes from the observation that Neolithic expansions explain the prehistory of other large language families, such as Afro-Asiatic, Niger-Congo, Sino-Tibetan, and Austronesian. Taking this a step further, Indo-European is simply not an exception to the rule.

In 2015 Haak et al. published a palaeogenomic study in the journal *Nature* that takes sides in the dispute surrounding the origins and expansion of Indo-European languages. The study endorses the steppe hypothesis of Indo-European language origins. Based on their statistical analysis of ancient autosomal markers, they assert a “massive” expansion of steppe nomads associated with the prehistoric Yamnaya culture of Eastern Europe and Central Asia. The study times the expansion to the Bronze Age, roughly five thousand years go. According to the study this expansion was so massive that the steppe nomads replaced seventy-five percent of the pre-existing farmer genes in Central Europe. The study further asserts that the R1a-M420 and R1b-M343 mutations are the Y-chromosome relics of this massive invasion of steppe nomads. The basis of this assertion is not exactly clear. However, it is inconsistent with studies that map the evolutionary history of both markers with phylogenetic analysis (Myres et al. 2011; Underhill et al. 2015).

The position taken by Haak et al. (2015) has been endorsed by several subsequent palaeogenomic studies (e.g. Allentoft et al. 2015; Jones et al. 2015; Cassidy et al. 2016; Jones et al. 2017; Olalde et al. 2018; Narasimhan et al. 2019). Based on the fact that these studies are often published in respected science journals, and the fact that they are often cited in the scientific literature and news media, linguists might also be tempted to endorse palaeogenomic modeling to decipher language prehistory. Accordingly, Sections 3 to 10 (below) provide a *de novo* review of the evolutionary history of the R-M207 haplogroup.



Paper 5.17. Haplogroup R-M207.

The Genetic-Linguistic Interface.

Section 2. The Contemporary Distribution of Haplogroup R-M207: an Overview.

Haplogroup has two main divisions within its phylogeny, R1-M173 and R2-M479. The R2-M479 branch is found mostly among populations living in South Asia. The R1a-M420 and R1b-343 mutations define the two main divisions within R1-M173. R1a-M420 mutations are found in Scandinavia, Eastern Europe, the Baltic Region, South Asia, Central Asia and Northern Eurasia. R1b-M343 mutations are mostly found in Western Europe and the Sahel region of Africa.

Section 3. The Evolution of Haplogroup R-M207.

Around 29 thousand years ago Marine Isotope Stage 3 ended. This marks the beginning of Marine Isotope Stage 2. Shortly thereafter, roughly 25 thousand years ago, the Ice Age glaciers reached their maximum southern extension across Eurasia. The literature refers to this event as the Last Glacial Maximum (see Clark et al. 2009 for a more detailed discussion). Within this climatological context, haplogroups R-M207 and Q-M242 diverged from the P1-M45 mutation roughly 30 thousand years ago (Poznik et al. 2016: Supplementary Table 10).

Haplogroups R-M207 and Q-M242 represent the genetic signature of those who survived the Last Glacial Maximum. Around the time of the Last Glacial Maximum, glaciation pushed human populations across Eurasia southwards into what the literature (e.g. Gavashelishvili and Tarkhnishvili 2016) describes as “refugia.” The ancient DNA data, along with the archeological and climatological data suggest that haplogroups R-M207 and Q-M242 co-evolved in a south-central Siberian refugium in the Altai-Sayan region, where the borders of contemporary Russia, Kazakhstan, China and Mongolia converge. Ancient DNA evidence comes from the so-called “Mal’ta boy.” He died about 24 thousand year ago near Lake Baikal, which is the general vicinity of the Altai-Sayan region. As previously noted in Paper 5.16, Hg. Q, Section 3, researchers place the child’s Y-chromosome somewhere near the root of the R-M207 haplogroup, where haplogroups R-M207 and Q-M242 diverge from P1-M45 (see, also, Raghavan et al. 2014).

The climatological record suggests that the Altai-Sayan region and Eastern Europe were spared from the extreme glaciation that had occurred in Western Europe (e.g. Velichko et al. 2009; Binney et al. 2016). Cold and arid climatic conditions within a “hyperzonal open-type” landscape supported large mammal food resources for human populations across a vast geographical expanse. This so-called “hyperzone” included South-Central Siberia and much of Eastern Europe about three hundred kilometers south of the ice glaciers, which equates to around 55 degrees north for both regions (Velichko et al. 2009). According to Kuzmin (2008), because of the mega-faunal (large mammal) food resources that were supported by the hyperzone, human population density in South-Central Siberia remained relatively stable. Furthermore, the researcher asserts that other factors contributed to successful cold weather adaptation for those living in this region during the Last Glacial Maximum: micro-blade tools, suitable dwellings and clothing, and the availability of bones for fuel.

Meanwhile the glacial ice sheet in Western Europe pushed much further south during the Last Glacial Maximum, almost to 40 degrees north. Kuzmin (2008) attributes greater



Paper 5.17. Haplogroup R-M207.

The Genetic-Linguistic Interface.

glaciation in Western Europe to a shift in the Atlantic storm track that brought more moisture to the region. As a result, Scandinavia and much of Western Europe became depopulated and human populations retreated to the Iberian Peninsula. Interestingly, in contrast to refugia in South-Central Siberia and Eastern Europe, human population density within the Iberian refugium may well have declined as the result of a decline in reindeer populations (Jochim et al. 1999; Morein 2008).

The comparison of Iberian and Siberian population densities during the Last Glacial Maximum, as presented in the previous paragraph, potentially explains why NO-M214 and C-M130 mutations disappeared from the European gene pool. As previously discussed in Paper 5.6, Hg. C, Section 4 and 5.14, Hg. N, Section 2, human remains and ancient DNA data suggest that both mutation were very much part of the European gene pool during Marine Isotope Stage 3. Thus, a population bottleneck may explain their demise, and this explains why haplogroup I-M170 remains the sole uncontested founder lineage among contemporary Europeans. Taking this a step further, greater reproductive success during the last Glacial Maximum may partially explain the ubiquitous presence of R-M173 mutations among the contemporary population of Eurasia and Africa.

Section 4. The Expansion and Diversification of R1-M173.

Haplogroups Q-M242 and R-M207 are so-called “sister” clades. This close phylogenetic relationship suggests that the previous discussion of Q-M242 might help to decipher the evolutionary history of the “sister.” As detailed in Paper 5.16, Section 3, Q-M242 expanded out of South-Central Siberia into the Americas beginning about 16 thousand years ago. This expansion was fueled by a dramatic increase and expansion of woolly mammoth populations. Then around the beginning of the Holocene, roughly 12 thousand years ago, the mammoths and other large “mega-faunal” food resources of the Americas disappeared. At this point human populations in the Western Hemisphere diversified their subsistence strategy. Initially, the hunter-gatherer exploited food resources available at the regional level, such as marine mammals along the Pacific Coast, or alpacas in the Andean highlands. Later, many of the indigenous peoples turned to agriculture as part of their subsistence strategy.

Turning now to R1-M173, the evolutionary history of this mutation is strikingly similar to that of Q-M242 mutations: the mammoth steppes; the collapse of mega-faunal food resources; and the transition to agriculture. The R1-M173 mutation arose about 28 thousand years (Poznik et al. 2016: Supplementary Table 10). The contemporary distribution of R2a-M124 mutations (see Section 10), along with the climatological and archeological evidence, as presented previously in Section 3, suggests that this occurred in South-Central Siberia. Then, about 23 thousand years ago, R1a-M420 and R1b-M343 diverged from R1-M173 (Poznik et al. 2016: Supplementary Table 10). Underhill et al. (2015) suggest, based on analysis of contemporary Y-chromosome data, the split occurred in the Middle East in the vicinity of Iran. However, archaeological and ancient DNA evidence suggests the East European Plain (see Figure 5.17.2), in the vicinity of Kiev in the Ukraine. Archaeological support for the European Plain stems from Abramova et al. and their 2001 report of Upper Paleolithic sites in the middle Dnieper River basin. These sites date from about 12 to 25 years thousand ago, and as such, they extend back in time to the Last Glacial Maximum. Furthermore, these sites are found along the same latitude as the Altai-Sayan refugium.



Paper 5.17. Haplogroup R-M207.

The Genetic-Linguistic Interface.

Finally, ancient DNA Y-chromosome data from the Baltic region and Eastern Europe support this idea (see Tables 5.17.1 and 5.17.2).

Section 5. Diversification of R1b-M343 on the East European Plain.

Dolukhanov (2009) provides a discussion of hunter-gatherer cultures at the time of the Paleolithic-Mesolithic transition of the Eastern European, about 12 thousand years ago. From a climatological perspective, this transition marks the beginning of the Holocene. Connected to this change in climate was the southward expansion of human population from northeastern Europe to Southeastern Europe. Taking this a step further, ancient DNA, phylogenetic relationships and the contemporary distribution of modern-day Y-chromosome variation support the following idea: initial diversification of R1b-M343 variation occurred in northeastern Europe and southeastern Europe on the East European Plain. Ancient DNA support for this position comes from Table 5.17.1. As shown by the table, at the onset of the Holocene R1b-M343 mutations had already spread across the East Europe Plain along a north to south axis.

At this point the reader is invited to review Figure 5.17.1: Part C. The R1b-M335 marker was found in eight thousand-year-old remains in Latvia. Given the unique position of R1b-M335 within R1b-M343 phylogeny, we have ancient DNA evidence that support diversification of R1b-M343 in Northeastern Europe. Southeastern European diversification of R1b-M343, on the other hand, is supported by the contemporary distribution of R1b-M73 (see Table 5.17.3), R1b-V88 (see Section 7) and R1b-CTS1078 (see Section 8).

Section 6. The Expansion of R1b-M269 into Western Europe.

We now present a model for explaining the contemporary distribution of R1b-M269 mutations in Europe. R1b-M269, like R1b-M335, evolved in northeastern Europe. Around the onset of the Holocene, R1b-M269 expanded westward across the North European Plain (see Figure 5.17.3) into Western Europe. The mutation was carried by reindeer hunters. Secondary expansions then carried the mutation into Scandinavia, the British Isles, Iberia and the Mediterranean.

Turning now to the archaeological record, during the Last Glacial Maximum human populations in Western Europe retreated to southern refugia including the Iberian Peninsula (for additional details, see Section 3). Based on radio-carbon estimates (Housley et al. 1997), about 14 thousand years ago people and reindeer began their expansion northwards. The glaciers slowly retreated. Tundra then appeared in previous glaciated area. The tundra eventually succumbed to the forests, and the tundra line eventually retreated to its present location above the Arctic Circle. The reindeer followed the retreating tundra line, and hunter-gatherers followed the reindeer (see Sommer et al. 2014 for additional details). By around 12 thousand years ago, the reindeer arrived in Denmark, and 9 thousand years ago, they disappeared into the arctic region of Scandinavia (Aaris-Sorensen et al. 2007).

The archaeological evidence reflects that around the beginning of the Holocene (12 thousand years ago) almost all of the mega-faunal food resources of northeastern Europe, such as mammoths, had disappeared (e.g. Puzachenko and Markova 2019). The last



Paper 5.17. Haplogroup R-M207.

The Genetic-Linguistic Interface.

remaining mega-faunal resource at this time was reindeer (see Dolukhanov 2009: 32 and the discussion of the Swiderian culture). Thus it appears as though hunter-gatherers on the East European Plain were drawn to Western Europe because of the abundance reindeer in this region and the demise of mega-faunal resources elsewhere.

Based on genetic evidence, at the onset of the Holocene hunter-gatherers intercepted the northward migration of reindeer in Western Europe and Scandinavia along two different trajectories. As explained previously in Paper 5.9, haplogroup I-M170 represents the genetic signature of the south-north trajectory. The R1b-L51 mutation, on the other hand, represents the genetic signature of the east-west trajectory. The R1b-L51 mutation is a downstream variant of the R1b-M269 mutation. According to Myres et al. (2011), this mutation evolved in Western Europe.

Sometime during the early Mesolithic, R1b-U106 and R1b-S116 diverged from R1b-L51. R1b-U106 reflects genetic diversification of R1b-L51 on the North European Plain, the British Isles and Scandinavia (see Myres et al. 2011 and Table 5.17.4). The R1b-S116 mutation, on the other hand, reflects diversification of R1b-L51 in Iberia (Valverde et al. 2016).

R1b-S116 has three informative downstream mutations: R1b-DF27, R1b-U152, and R1b-M529. Valverde et al. (2016) suggest based on their analysis of the genetic evidence that these three mutations diverged from R1b-S116 on the Iberian Peninsula around 12 thousand years ago. They also take the position that the R1b-DF27 mutation remained on the peninsula. R1b-U152, on the other hand, expanded eastwards onto the Italian Peninsula and northwards, perhaps through the Alps into Germany. Like R1b-U152, R1b-M529 also expanded out of the Iberian Peninsula. However, this mutation expanded along a different trajectory to the British Isles. This expansion model is supported by the contemporary distribution of all three mutations (see Tables 5.17.5-5.17.7) and the observation that the highest frequency of S116 mutations are found in Spain and Portugal (Valverde et al. 2016).

Dating estimates provided by Solé-Morata et al. (2017) suggest an expansion of R1b-U152 and R1b-M529 from the Iberian Peninsula at around four thousand years ago. In contrast to this estimate, coalescent times provided Myres et al. (2011: Table S2) suggest the beginning of the Neolithic transition on the Iberian Peninsula. These estimates from Myres et al. (2011) are more consistent with the archeological record. According to Martins et al. (2015) agriculture on the Iberian Peninsula dates to around 7.5 thousand years ago. This is based mostly on radio-carbon results taken from the remains of domesticated sheep and goats, an exceptionally reliable data source for delineating the Neolithic/Mesolithic transition. Their data further suggest that the Neolithic transition drove rapid population growth within the region. Thus it seems that population pressure during the Early Iberian Neolithic resulted in an expansion of R1b-U152 and R1b-M529 from the peninsula.

One interesting observation from recent Iberian studies (Valverde et al. 2016 and Solé-Morata et al. 2017) is that the highest frequency of R1b-DF27 mutations is found among the Basque people. This is a salient point for linguists because many researchers consider the Basque language isolate a linguistic relic of Europe from a time that predates the arrival of Indo-European languages. In Paper 5.9, Section 7, genetic support was provided for this position based on data from the I2a-M26, H2-P56 and C-M130 mutations. The R1b-S116 mutation,



Paper 5.17. Haplogroup R-M207.

The Genetic-Linguistic Interface.

and its arrival in the Basque region around the beginning of the Holocene, provides additional support.

Section 7. The Expansion of R1b-V88 into Africa.

The highest frequencies of the R1b-V88 mutation are found in the Sahel region of Africa, a zone that divides North Africa and Sub-Saharan Africa (see Table 5.17.8). The archaeological and genetic data, however, suggest that the mutation evolved in southeastern Europe during the Mesolithic. Support for this position comes from a recent study (D'Atanasio et al. 2018) that suggests that posits the evolution of R1b-V88 about 12 thousand years ago in Europe. Then about eight thousand years ago the mutation was carried southwards into northeastern Africa. From this location, the mutation expanded southwards into the Sahel.

The R1b-V88 expansion into Africa stands as genetic relic of the North African Mesolithic and the so-called “humid phase.” The previous discussion of the A-M13 mutation in Paper 5.2, Section 3 reports that about 10 thousand years ago the climate of northern Africa underwent a dramatic transformation. Holocene climate change brought monsoon rain to the region. The Sahara Desert became a savannah with numerous lakes and rivers. Within the complex system of rivers and lakes, hippos, crocodiles and fish proliferated. Archeological evidence demonstrates that human hunted these animals.

Abundant food resource appears to have pulled the R1b-V88 mutation into Northern Africa. Several lines of evidence support such a position. Interestingly, the oldest R1b-M343 sample comes from the Villabruna remains found in the Dolomite Mountains region of Northern Italy (see Table 5.17.1 and data for the Mediterranean region). The remains are from a man who died around 14 thousand years ago. Researchers determined that he belongs to R1b1-L278, and as such, he becomes a potential genetic ancestor of R1b-V88 populations. Partial support for this position comes from the previous discussion in Section 6, which suggests that Mesolithic R1b-V88 mutations in Italy were replaced by the Neolithic R1b-U152 mutation. A similar position was taken by D'Atanasio et al. (2018). The same study also reports that R1b-V88 mutations among contemporary Sardinians are older than African Rb-V88 mutations. Furthermore, the Sardinian samples date to the beginning of the North African humid phase.

Kuper and Kröpelin in their 2006 study suggest that during the humid phase the Nile River was a marshland and people were not able live in this region. As such, a Mesolithic expansion of R1b-V88 via the Middle East and northeastern Africa would have been difficult. On the other hand, palaeohydrological data taken from satellite imagery (Drake et al. 2011: Figures 1 and 2) show numerous river systems that were present in North Africa during the humid phase. Taking this a step further, hunter-gatherers made a water crossing from Sardinia to the North Africa. From Mediterranean coastline of North Africa, these rivers facilitated a southward migration to the Mesolithic food resources in the vicinity of Lake Chad.

Chadic is a branch of the Afro-Asiatic language family. In a 2010 study, Cruciani et al. suggested that R1b-V88 is a genetic signature for Chadic languages based on the high frequency of the mutations among the Chadic-speaking populations of Africa. Such a position



Paper 5.17. Haplogroup R-M207.

The Genetic-Linguistic Interface.

seems problematic for several reasons. Haber et al. (2016) disputed the conclusion from Cruciani et al. (2010) asserting that the oldest R1b-V88 mutations are found among the Laal-speakers, a language isolate. Data for contemporary African populations (see Table 5.17.9) also show that R1b-V88 is also found among speakers of Semitic and Berber languages, which are also branches of the Afro-Asiatic family. Additionally, R1b-V88 is well represented among several populations that speak either a non-Bantoid Niger-Congo languages or Nilo-Saharan languages. Finally, the E-M34, E-M81 and J1-M67 mutations, along with the archaeological data, suggest that Afro-Asiatic expanded into Africa from the Middle East during the Neolithic (see Paper 5.5, Hg. E, Section 3 and Paper 5.10, Hg. J, Section 3), whereas R1b-V88 is a Mesolithic relic among African populations.

Section 8. The DNA of Steppe Nomads.

As the result of palaeogenomic studies Section 1, Bronze Age steppe nomads of Eastern Europe have taken center stage in the debate surrounding the prehistory of Indo-European languages. According to the ancient DNA data, the genetic signature of these nomads is the R1b-CTS1078 mutation (see table 5.17.10). Based on contemporary data (see Table 5.17.11), and phylogenetic relationships (see Figure 5.17.1: Part C), the R1b-CTS1078 mutation appears to have evolved in southeastern Europe. Frequency results from Table 5.17.11 clearly rule out a massive migration of the steppe nomads into Western Europe during the Bronze Age.

Section 9. Diversification of R1a-M420 on the East European Plain.

As previously explained in Section 4, R1a-M420 and R1b-M343 diverged from R1-M173 about 23 thousand years ago somewhere on the East European Plain. Like R1b-M343 (see Section 5), the initial diversification of R1a-M420 variation occurred in northeastern Europe and southeastern Europe. R1a-M420 has two main downstream variants, R1a-Z282 and R1a-Z93 (see Figure 5.17.1: Part B). Based on contemporary population data (see Table 5.17.12) it appears as though the R1a-Z282 mutation represents diversification of R1a-M420 in northeastern Europe. Limited support for this position comes from ancient DNA data (see table 5.17.2).

R1a-Z282 has three informative downstream markers: R1a-Z284, R1a-M458, and R1a-M558. R1a-Z284 is confined almost exclusively to Scandinavia where it attains a frequency of around twenty percent among Norwegians (Underhill et al. 2015). Based on coalescent time estimates (Underhill et al. 2015: Table S5) it appear as though the evolutionary history of R1a-Z284 appear to be similar to that of R1b-U106 (see Section 6). Specifically, both are genetic relics of an early Mesolithic expansion of reindeer hunters.

The R1a-M448 and R1a-M558 mutations extend across the northern East European Plain. Based on contemporary data, R1a-M448 appears to have a higher frequency among West Slavic speakers (see Table 5.17.14) and R1a-M558 seems to peak among East Slavic populations (see Table 5.17.15). Thus, it should be emphasized that these data results should not be associated with the historical Slavic expansion. Rather, mainstream linguistic opinion considers language contact theory rather than population expansion as the theoretical



Paper 5.17. Haplogroup R-M207.

The Genetic-Linguistic Interface.

approach to explaining the contemporary distribution of Slavic languages (e.g. Brackney 2007).

Contemporary data suggest that R1a-Z93 represents diversification of R1a-M420 mutations in southeastern Europe or perhaps Central Asia (Table 5.17.13). R1a-Z93 is an especially significant component of South Asian populations. Underhill et al. (2015: Table S4) suggest a frequency of about eighteen percent for the region. At this point the reader is invited to examine Table 5.17.16. The table sorts frequency data for R1a-Z93 according to language family or language branch. These data reflect that R1a-Z93 is a significant marker among populations that speak Iranian, Indo-Aryan, Dravidian and Turkic languages. According to Pamjav et al. (2012), R1a-Z93 diverged from R1a-M420 about 10 thousand years ago. These data along with the archeological record support a model that associates the evolutionary history of R1a-Z93 with Holocene climate change. Misra (2001) reports that population density was low during the Upper Paleolithic in India. Arid and cold weather had limited the availability of food resources. The Mesolithic, however, brought monsoon rains. Increased moisture produced more food resources that ultimately drove higher population density.

Section 10. The Evolutionary History of R2-M479.

As noted previously in Section 2, R1-M173 and R2-M479 form the two main downstream divisions of the R-M207 haplogroup. Data from Poznik et al. (2016: Supplementary Table 10) suggest that the R2-M479 mutation evolved roughly 28 thousand years ago, around the time of the Last Glacial Maximum. As suggested in Section 4, this appears to have occurred in South-Central Siberia. Contemporary data for the mutation consists almost entirely of frequency results for the R2a-M124 downstream mutation. Table 5.17.17 reports R2a-M125 frequencies from a regional perspective. As shown by the table, almost all the reported data for the mutation comes from South Asia.

Ancient DNA helps to resolve the evolutionary history of R2a-M124. As shown by Table 5.17.18, R2a-M124 was extracted from a 10 thousand year old sample found at Ganj Dareh in northwestern Iran. This data supports the idea that climate change drove R2a-M124 into South Asia during the Mesolithic, a model that is similar to that of R1a-Z93 (see Section 9 above).

Sengupta et al (2006) suggest that about nine percent of Indian and seven percent of Pakistani males have the R2a-M124 mutation. At this point the reader is directed to Table 5.17.19 which reports R2a-M124 data from a language perspective. The mutation appears to be an especially informative marker for deciphering the population history among those that speak Dravidian and Indo-Aryan languages. Like R1a-Z93, R2a-M124 was present in South Asia before the arrival of Indo-European speaking population (see Sections 3 and 5 for a more detailed discussion).



Paper 5.17. Haplogroup R-M207.

The Genetic-Linguistic Interface.

Section 11. Conclusions for R-M207.

The R-M207 haplogroup represent a Paleolithic component for populations in Northern Eurasia, the Eastern Europe, and the Baltic Region. For populations in South Asia, Western Europe, and Northern Africa, it is a Mesolithic component. The available data, on the other hand, do not allow us to make a conclusion at this time for Central Asia and the Caucasus.

For linguists, haplogroup R-M207 and its downstream variants provide especially useful makers for deciphering the prehistory of the Indo-European, Dravidian, Afro-Asiatic, and Nilo-Saharan language families as well as the Basque and Laal language isolates. One interesting observation is that language contact appears to partially explain the prehistoric expansion of Indo-European and Afro-Asiatic language families. R-M207 mutations were already in South Asia and Europe when Indo-European speaking farmers arrived in these regions. Similarly, R-M207 was in northern Africa when Afro-Asiatic speaking farmers arrived.

Our *de novo review* of the evolutionary history of R-M207 and its downstream variants (Sections 3 to 10) demonstrates the following: palaeogenomic modeling produces conclusions that are flagrantly inconsistent with the archaeological record. It is tempting to suggest that this problem could be resolved by improving the statistical methods. However, the root of the problem is that autosomal markers are simply not useful for deciphering human population history because of recombination. This reshuffling of the genetic “cards” seriously distorts evolutionary relationships and statistical methods will not overcome this deficiency. Thus the message to linguists seems clear. The *early farming dispersal hypothesis* still remains a very robust model of Indo-European prehistory. Finally, we need a different model for deciphering the prehistory of language. Palaeogenomic modeling is not reliable. Triangulated Y-chromosome modeling is an alternative that is empirical, transparent and far more reliable.

Bibliography for Report and Tables

Aaris-Sorensen, Kim et al. 2007. “The Scandinavian reindeer (*Rangifer tarandus* L.) after the last glacial maximum: time, seasonality and human exploitation.” *Journal of Archaeological Science* 34(6): 914-923.

Abramova et al. 2001. “The age of Upper Paleolithic sites in the middle Dnieper River basin of Eastern Europe.” *Radiocarbon* 43(2b): 1077-1084.

Abu-Amro, Khaled K. et al. 2009 . “Saudi Arabian Y-chromosome diversity and its relationship with nearby regions.” *BioMed Central Genetics* 10:59.

Agdzhoyan, A. T. et al. 2016. “Gene pool of Siberian Tatars: five ways of origin for five subethnic groups.” *Molecular Biology* 50(6): 860–873.

Allentoft, M. E. et al. 2015. “Population genomics of Bronze Age Eurasia.” *Nature* 522: 167-172.



Paper 5.17. Haplogroup R-M207.

The Genetic-Linguistic Interface.

Alonso, Santos et al., 2005. "The place of the Basques in the European Y-Chromosome diversity landscape." *European Journal of Human Genetics* 13: 1293-1302.

Anthony, David W. and Don Ringe 2015. "The Indo-European homeland from linguistic and archaeological perspectives." *Annual Review of Linguistics* 1: 199-219.

Arunkumar, Ganesh Prasad et al. 2012. "Population differentiation of southern Indian male lineages correlates with agricultural expansions predating the caste system." *Public Library of Science One* 7(11): e50269.

Balanovsky, Oleg et al. 2008. "Two sources of the Russian patrilineal heritage in their Eurasian context." *American Journal of Human Genetics* 82: 236-250.

Balanovsky et al. 2011. "Parallel evolution of genes and languages in the Caucasus region." *Molecular Biology and Evolution* 28(10): 2905-2920.

Balanovsky, O. et al. 2017. "Genetic differentiation between upland and lowland populations shapes the Y-chromosomal landscape of West Asia." *Human Genetics* 136: 437-450.

Bellwood, Peter 2005. *First Farmers: the Origins of Agricultural Societies*. Malden, MA; Oxford, UK; Victoria, Australia: Blackwell Publishing.

Binney, Heather et al. 2016. "Vegetation of Eurasia from the Last Glacial Maximum to present: key biogeographic patterns." *Quaternary Science Reviews* 157: 80-97.

Boattini, Alessio et al. 2013. "Uniparental markers in Italy reveal a sex-biased genetic structure and different historical strata." *Public Library of Science One* 8(5): e65441.

Borkar, Minal et al. 2011. "Paleolithic spread of Y-chromosomal lineage of tribes in eastern and northeastern India." *Annals of Human Biology* 38(6): 736-746.

Bosch, Elena et al. 2003. "High level of male-biased Scandinavian admixture in Greenlandic Inuit shown by Y-chromosome analysis." *Human Genetics* 112: 353-363.

Brackney, Noel C. 2007. *The Origins of Slavonic: Language Contact and Language Change*. Munich: LINCOM Europa.

Busby, George et al. 2012. "The peopling of Europe and the cautionary tale of Y chromosome lineage R-M269." *Proceedings of the Royal Society B*. 279: 884-992.

Brunelli, Andrea et al. 2017. "Y chromosomal evidence on the origin of northern Thai people." *Public Library of Science One* 12(7): e0181935.

Bučková, Jana et al. 2013. "Multiple and differentiated contributions to the male gene pool of pastoral and farmer populations of the African Sahel." *American Journal of Physical Anthropology* 151: 10-21.



Paper 5.17. Haplogroup R-M207.

The Genetic-Linguistic Interface.

- Cadenas, Alicia M. et al. 2008. "Y-chromosome diversity characterizes the Gulf of Oman." *European Journal of Human Genetics* 16: 374-386.
- Cassidy, Lara M. et al. 2016. "Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome." *Proceedings of the National Academy of Sciences of the United States of America* 113(2): 368-373.
- Chennakrishnaiah, et al. 2013. "Indigenous and foreign Y-chromosomes characterize the Lingayat and Vokkaliga populations of Southwest India." *Gene* 526: 96-106.
- Cinnioglu, Cengiz et al. 2004. "Excavating Y-chromosome haplotype strata in Anatolia." *Human Genetics* 114: 127-148.
- Cruciani, Fulvio 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.
- Cruciani, Fulvio et al 2011. "Strong intra- and inter-continental differentiation revealed by Y chromosome SNPs M269, U106 and U152." *Forensic Science International: Genetics* 5: e-49-e52.
- Clark, Peter U. et al. 2009. "The Last Glacial Maximum." *Science* 325: 710-714.
- Csányi, B. et al 2008. "Y-chromosome analysis of ancient Hungarian and two modern Hungarian-speaking populations from the Carpathian Basin." *Annals of Human Genetics* 72: 519-534.
- Damgaard, P. B. et al. 2018. "137 ancient human genomes from across the Eurasian steppes." *Nature* 557: 369-374.
- D'Atanasio et al. 2018. "The peopling of the last Green Sahara revealed by high-coverage resequencing of trans-Saharan patrilineages." *Genome Biology* 19: 20.
- Debnath, Monojit et al. 2011. "Y-chromosome haplogroup diversity in the sub-Himalayan Terai and Duars populations of East India." *Journal of Human Genetics* 56: 765-771.
- Di Cristofaro, Julie et al. 2013. "Afghan Hindu Kush: where Eurasian Sub-Continent gene flows converge." *Public Library of Science One* 8(10): e76748.
- Dolukhanov, P. 2009. "The Mesolithic of the East European Plain." In: *The East European Plain on the Eve of Agriculture*. Edited by Pavel M. Dolukhanov, Graeme R. Sarson and Anvar M. Shukurov. BAR International Series 1964. Oxford, UK: Archaeopress, pp. 23-34.
- Drake, Nick A. et al. 2011. "Ancient watercourses and biogeography of the Sahara explain the peopling of the desert." *Proceedings of the National Academy of Sciences of the United States of America* 108(2): 458-462.



Paper 5.17. Haplogroup R-M207.

The Genetic-Linguistic Interface.

Dulik, Matthew C. et al. 2012. “Mitochondrial DNA and Y chromosome variation provide evidence for a recent common ancestry between Native Americans and indigenous Altaians.” *American Journal of Human Genetics* 90: 229-246.

Fedorova, Sardana A. et al. 2013. “Autosomal and uniparental portraits of the native populations of Sakha (Yakutia): implications for the peopling of Northeast Eurasia.” *BioMed Central Evolutionary Biology* 13:127.

Fornarino, Simona et al. 2009. “Mitochondria and Y-chromosome diversity of the Tharus (Nepal): a reservoir of genetic variation.” *BioMed Central Evolutionary Biology* 9:154.

Frachetti, Michael D. 2012. “Multiregional emergence of mobile pastoralism and nonuniform institutional complexity across Eurasia.” *Current Anthropology* 53(1): 2-38.

Fu, Q. et al. 2016. “The genetic history of Ice Age Europe.” *Nature* 534: 200-205.

Gayden, Tenzin et al. 2007. “The Himalayas as a directional barrier to gene flow.” *American Journal of Human Genetics* 80: 884-894.

Gavashelishvili, Alexander and David Tarkhnishvili 2016. “Biomes and human distribution during the last Ice Age.” *Global Ecology and Biogeography* 25: 563-574.

Gazi, Nurun Nahar et al. 2013. “Genetic structure of Tibeto-Burman populations of Bangladesh: evaluating the gene flow along the sides of Bay-of-Bengal.” *Public Library of Science One* 8(10): e75064.

Grugni, Viola et al. 2012. “Ancient migratory events in the Middle East: new clues from the Y-Chromosome variation of modern Iranians.” *Public Library of Science One* 7(7): e41252.

Gubina, M. A. et al. 2013. “Haplotype diversity in mtDNA and Y-chromosome in populations of Altai–Sayan Region.” *Russian Journal of Genetics* 49(3): 329-343.

Haak, Wolfgang et al. 2015. “Massive migration from the steppe was a source for Indo-European languages in Europe.” *Nature* 522: 207-211.

Haber, Marc et al. 2012. “Afghanistan’s ethnic groups share a Y-chromosomal heritage structured by historical events.” *Public Library of Science One* 7(3): e34288.

Haber, Marc et al. 2016. “Chad genetic diversity reveals an African history marked by multiple Holocene Eurasian migrations.” *American Journal of Human Genetics* 99: 1316–1324.

Hassan, Hisham Y. et al. 2008. “Y-chromosome variation among Sudanese: restricted gene flow, concordance with language, geography and history.” *American Journal of Physical Anthropology* 137: 316-323.

Helgason, Agnar et al. 2000. “Estimating Scandinavian and Gaelic ancestry in the male settlers of Iceland.” *American Journal of Human Genetics* 67: 697-717.



Paper 5.17. Haplogroup R-M207.

The Genetic-Linguistic Interface.

Housley, R.A. et al. 1997. "Radiocarbon evidence for the late glacial human recolonisation of Northern Europe." *Proceedings of the Prehistoric Society* 63: 25-54.

Hovhannisyan, Anahit et al. 2014. "Different waves and directions of Neolithic migrations in the Armenian Highland." *Investigative Genetics* 5:15.

Jochim, Michael et al. 1999. "The Magdalenian colonization of southern Germany." *American Anthropologist* 101:129-142.

Jones, Eppie R. et al. 2015. "Upper Palaeolithic genomes of modern Eurasians." *Nature Communications* 6: 8912.

Jones, Eppie R. et al. 2017. "The Neolithic transition in the Baltic was not driven by admixture with early European farmers." *Current Biology* 27: 576–582.

Karachanak, Sena et al. 2013. "Y-chromosome diversity in modern Bulgarians: new clues about their ancestry." *Public Library of Science One* 8(3): e56779.

Karafet, Tatiana et al. 2016. "Coevolution of genes and languages and high levels of population structure among the highland populations of Daghestan." *Journal of Human Genetics* (2016) 61: 181-191.

Kharkov, V. N. et al. 2007. "Gene pool differences between Northern and Southern Altaians inferred from the data on Y-chromosomal haplogroups." *Russian Journal of Genetics* 43(5): 675-687.

Kharkov, V.N et al. 2011. "Genetic diversity of the Khakass gene pool: subethnic differentiation and the structure of Y-chromosome haplogroups." *Molecular Biology* 45(3): 404-416.

Kharkov, V. N. et al. 2013. "Gene pool structure of Tuvinians inferred from Y-chromosome marker data." *Russian Journal of Genetics* 49(12) 1236-1244.

Khurana, Priyanka et al. 2014. "Y chromosome haplogroup distribution in Indo-European speaking tribes of Gujarat, western India." *Public Library of Science One* 9(3): e90414.

King, R.J. et al. 2008. "Differential Y-chromosome Anatolian influences on the Greek and Cretan Neolithic." *Annals of Human Genetics* 72: 205-214.

Kuper, Rudolph and Stefan Kroepelin 2006. "Climate-controlled Holocene occupation in the Sahara: motor of Africa's evolution." *Science* 313: 803-807.

Kushniarevich, Alena et al. 2013. "Uniparental genetic heritage of Belarusians: encounter of rare Middle Eastern matrilineages with a Central European mitochondrial DNA pool." *Public Library of Science One* 8(6): e66499.



Paper 5.17. Haplogroup R-M207.

The Genetic-Linguistic Interface.

- Kuzmin, Yaroslav V. 2008. "Siberia at the Last Glacial Maximum: environment and archaeology." *Journal of Archaeological Research* 16: 163-221.
- Lacau, Harlette et al. 2012. "Afghanistan from a Y-chromosome perspective." *European Journal of Human Genetics* 20: 1063-1070.
- Lappalainen, T. et al. 2008. "Migration waves to the Baltic Sea region" *Annals of Human Genetics* 72: 337-348.
- Lazaridis, I. et al. 2016. "Genomic insights into the origin of farming in the ancient Near East." *Nature* 536: 419-424.
- Lee, Eun Young et al. 2014. "Analysis of 22 Y chromosomal STR haplotypes and Y haplogroup distribution in Pathans of Pakistan." *Forensic Science International: Genetics* 11: 111-116.
- Lipson, M. et al. 2017. "Parallel palaeogenomic transects reveal complex genetic history of early European farmers." *Nature* 551: 368-372.
- Luca, F. et al. 2007. "Y-chromosomal variation in the Czech Republic." *American Journal of Physical Anthropology* 132: 132-139.
- Luis, J.R. et al. 2004. "The Levant versus the Horn of Africa: evidence for bidirectional corridors of human migrations." *American Journal of Human Genetics* 74: 532-544.
- Malyarchuk, Boris et al. 2011. "Ancient links between Siberians and Native Americans revealed by subtyping the Y chromosome haplogroup Q1a." *Journal of Human Genetics* 56: 583-588.
- Malyarchuk, Boris et al. 2013. "Y-chromosome diversity in the Kalmyks at the ethnical and tribal levels." *Journal of Human Genetics* 58: 804-811.
- Marjanovic, D. et al. 2005. "The peopling of modern Bosnia-Herzegovina: Y-chromosome haplogroups in the three main ethnic groups." *Annals of Human Genetics* 69: 757-763.
- Martinez, Laisel et al. 2007. "Paleolithic Y-haplogroup heritage predominates in Cretan highland plateau." *European Journal of Human Genetics* 15: 485-493.
- Martins, Haidé et al. 2015. "Radiocarbon Dating the Beginning of the Neolithic in Iberia: New Results, New Problems." *Journal of Mediterranean Archaeology* 28(1): 105-131.
- Mathieson, I. et al. 2015. "Genome-wide patterns of selection in 230 ancient Eurasians." *Nature* 528, 499-503.
- Mathieson, Iain et al. 2018. "The genomic history of southeastern Europe." *Nature* 555: 197-203.



Paper 5.17. Haplogroup R-M207.

The Genetic-Linguistic Interface.

Mielnik-Sikorska, Marta et al. 2013. "Genetic data from Y-chromosome STR and SNP loci in Ukrainian population." *Forensic Science International: Genetics* 7: 200-203.

Misra, Virendra Nath 2001. "Prehistoric human colonization of India." *Journal of Biosciences* 26(4):491-531.

Morein, Eugene 2008. "Evidence for declines in human population densities during the early Upper Paleolithic in western Europe." *Proceedings of the National Academy of Sciences of the United States of America* 105(1): 48-53.

Myres, Natalie M. et al. 2011. "A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe." *European Journal of Human Genetics* 19: 95-101

Narasimhan, Vagheesh M. et al. 2019. "The formation of human populations in South and Central Asia." *Science* 365: eaat7487.

Novackova, Jana et al. 2015. "The place of Slovakian paternal diversity in the clinal European landscape." *Annals of Human Genetics* 42(6): 511-522

Olalde, Inigo et al 2018. "The beaker phenomenon and the genomic transformation of Northwest Europe." *Nature* 555: 190-196.

Otoni, Claudio et al. 2011. "Deep into the roots of the Libyan Tuareg: a genetic survey of their paternal heritage." *American Journal of Physical Anthropology* 145: 118-124.

Pamjav, Horolma et al. 2012. "New Y-chromosome binary markers improve phylogenetic resolution within haplogroup R1a1." *American Journal of Physical Anthropology* 149: 611-615.

Peng, Min-Sheng et al. 2014. "Retrieving Y-chromosomal haplogroup tree GWAS data." *European Journal of Human Genetics* 22: 1046-1050.

Peričić, Marijana 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(10): 1964-1975.

Pimenoff, Ville N. 2008. "Northwest Siberian Khanty and Mansi in the junction of West and East Eurasian gene pools as revealed by uniparental markers." *European Journal of Human Genetics* 16: 1254-1264.

Poznik, G. David et al. 2016. "Punctuated bursts in human male demography inferred from 1,244 worldwide Y-chromosome sequences." *Nature Genetics* 48(6): 593-600.

Puzachenko, A.Yu. and A.K. Markova 2019. "Evolution of mammal species composition and species richness during the Late Pleistocene - Holocene transition in Europe: a general view at the regional scale." *Quaternary International* in press.



Paper 5.17. Haplogroup R-M207.

The Genetic-Linguistic Interface.

Raghavan, Maanasa et al. 2014. “Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans.” *Nature* 505: 87-91.

Ramos-Luis, E. et al. 2014. “Y-chromosomal DNA analysis in French male lineages.” *Forensic Science International: Genetics* 9: 162-168.

Rębała, Krzysztof et al. 2013. “Contemporary paternal genetic landscape of Polish and German populations: from early medieval Slavic expansion to post-World War II resettlements.” *European Journal of Human Genetics* 21: 415-422.

Regueiro, M. et al. 2006. “Iran: tricontinental nexus of Y-chromosome driven migration.” *Human Heredity* 61: 132-143.

Regueiro, Maria et al. 2012. “High levels of Paleolithic Y-chromosome lineages characterize Serbia.” *Gene* 498: 59-67.

Renfrew, Colin 1987. *Archaeology and Language: the Puzzle of Indo-European Origins*. Cambridge; New York: Cambridge University Press.

Saag, L. et al. 2017. “Extensive farming in Estonia started through a sex-biased migration from the steppe. *Current Biology* 27: 2185-2193.

Sarac, Jelena et al. 2016. “Genetic heritage of Croatians in the southeastern European gene pool - Y chromosome analysis of the Croatian continental and island population.” *American Journal of Human Biology* 28:837–845.

Semino, Ornella et al. 2000. “The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective.” *Science* 290: 1155-1159.

Sengupta, Sanghamitra et al. 2006. “Polarity and temporality of high-resolution Y-chromosome distributions in India identify both indigenous and exogenous expansions and reveals minor genetic influence of central Asian pastoralists.” *American Journal of Human Genetics* 78: 201-221.

Solé-Morata, Neus et al. 2017. “Analysis of the R1b-DF27 haplogroup shows that a large fraction of Iberian Y-chromosome lineages originated recently in situ.” *Scientific Reports* 7: 7341.

Sommer, Robert S. et al. 2014. “Range dynamics of the reindeer in Europe during the last 25,000 years.” *Journal of Biogeography* 41: 298-306.

Thangaraj, Kumarasamy et al. 2010. “The influence of natural barriers in shaping the genetic structure of Maharashtra Populations.” *Public Library of Science One* 5(12): e15283.

Thanseem, Ismail et al. 2006. “Genetic affinities among the lower castes and tribal groups of India: inference from Y chromosome and mitochondrial DNA.” *BioMed Central Genetics* 7: 42.



Paper 5.17. Haplogroup R-M207.

The Genetic-Linguistic Interface.

- Trofimova, N. V. et al. 2015. “Genetic characterization of populations of the Volga-Ural region according to the variability of the Y-chromosome.” *Russian Journal of Genetics* 51 (1): 108–115.
- Underhill, Peter A. et al. 2010. “Separating the post-Glacial coancestry of European and Asian Y chromosomes within haplogroup R1a.” *European Journal of Human Genetics* 18: 479-484.
- Underhill, Peter A. et al. 2015. “The phylogenetic and geographic structure of Y-chromosome haplogroup R1a.” *European Journal of Human Genetics* 23: 124-131.
- Valverde, Laura et al. 2016. “European paternal lineage M269: dissection of the Y-SNP S116 in Atlantic Europe and Iberia.” *European Journal of Human Genetics* 24: 437-441.
- Varzari, Alexander et al. 2013. “Paleo-Balkan and Slavic contributions to the genetic pool of Moldavians: insights from the Y chromosome.” *Public Library of Science One* 8(1): e53731.
- Velichko, A.A. et al. 2009. “Initial human settlement of East European Plain.” In: *The East European Plain on the Eve of Agriculture*. Edited by Pavel M. Dolukhanov, Graeme R. Sarson and Anvar M. Shukurov. BAR International Series 1964. Oxford, UK: Archaeopress, pp. 17-21.
- Villalba-Mouco et al. 2019. “Survival of Late Pleistocene Hunter-Gatherer Ancestry in the Iberian Peninsula.” *Current Biology* 29: 1-9.
- Völgyi, Antonia et al. 2008. “Haplogroup distribution of Hungarian population and the largest minority group.” *Forensic Science International Genetics Supplement Series* 1: 383-385.
- Voskarides, Konstantinos et al. 2016. “Y-chromosome phylogeographic analysis of the Greek-Cypriot population reveals elements consistent with Neolithic and Bronze Age settlements.” *Investigative Genetics* 7:1.
- Wang, Chuan-Chao et al. 2014. “Genetic structure of Qiangic populations residing in the Western Sichuan Corridor.” *Public Library of Science One* 9(8): e103772.
- Wilson, James F. et al. 2001. “Genetic evidence for different male and female roles during cultural transitions in the British Isles.” *Proceedings of the National Academy of Sciences of the United States of America* 98(9): 5078-5083.
- Yunusbayev, Bayazit et al. 2012. “The Caucasus as an asymmetric semipermeable barrier to ancient human migrations.” *Molecular Biology and Evolution* 29(1): 359-365.
- Zalloua, Pierre A. et al. 2008. “Y-chromosome diversity in Lebanon is structured by recent historical events.” *The American Journal of Human Genetics* 82: 873-882.
- Zhabagin, Maxat et al. 2017. “The connection of the genetic, cultural and geographic landscapes of Transoxiana.” *Scientific Reports* 7: 3085.



Paper 5.17. Haplogroup R-M207.

The Genetic-Linguistic Interface.

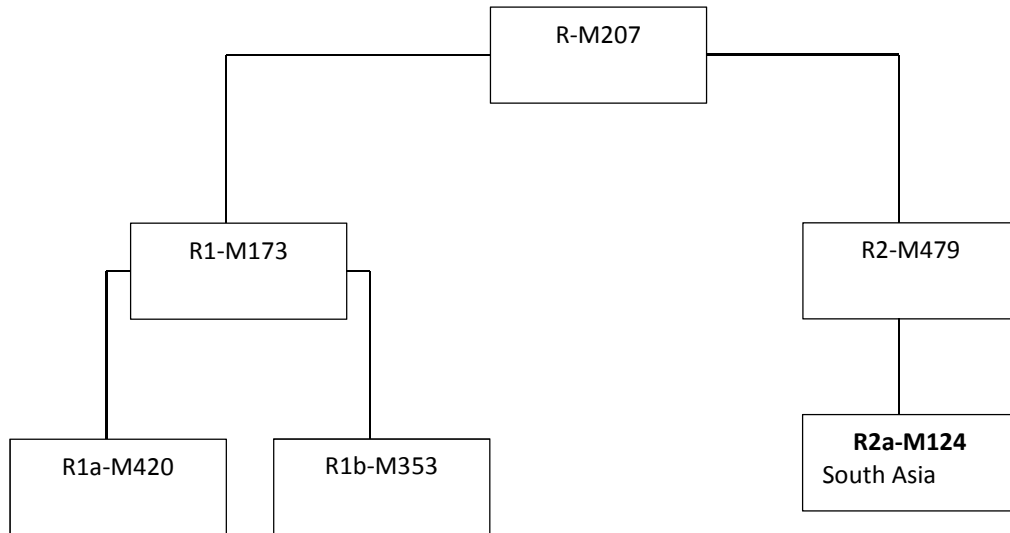
Zupan, Andrej et al. 2013. “The paternal perspective of the Slovenian population and its relationship with other populations.” *Annals of Human Biology* 40(6): 515-526.



Paper 5.17. Haplogroup R.

The Genetic-Linguistic Interface.

Figure 5.17.1: Part A. Initial Diversification of Haplogroup R-M207.



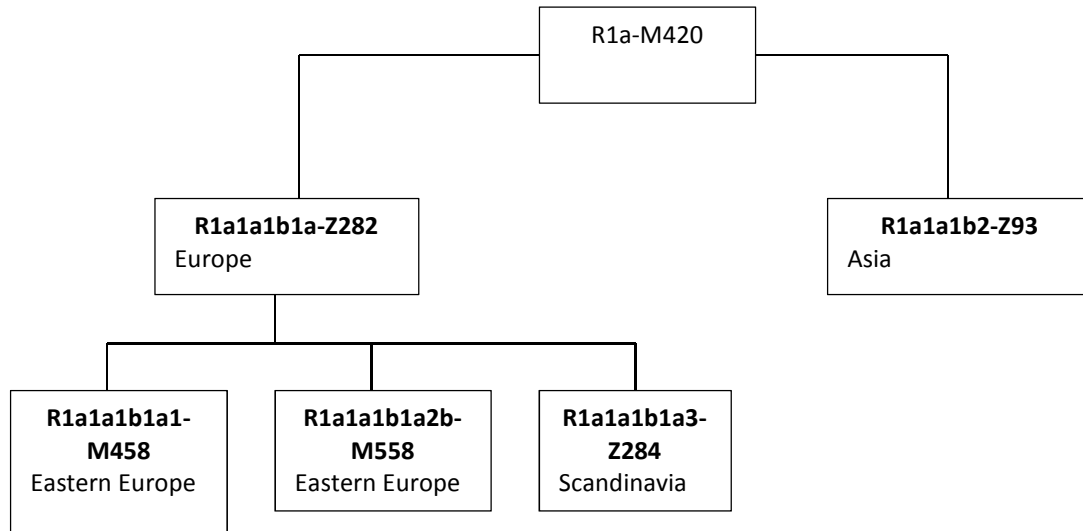
Nomenclature follows ISOGG 2017.



Paper 5.17. Haplogroup R.

The Genetic-Linguistic Interface.

Figure 5.17.1: Part B. Linguistically Informative R1a-M420 Mutations.



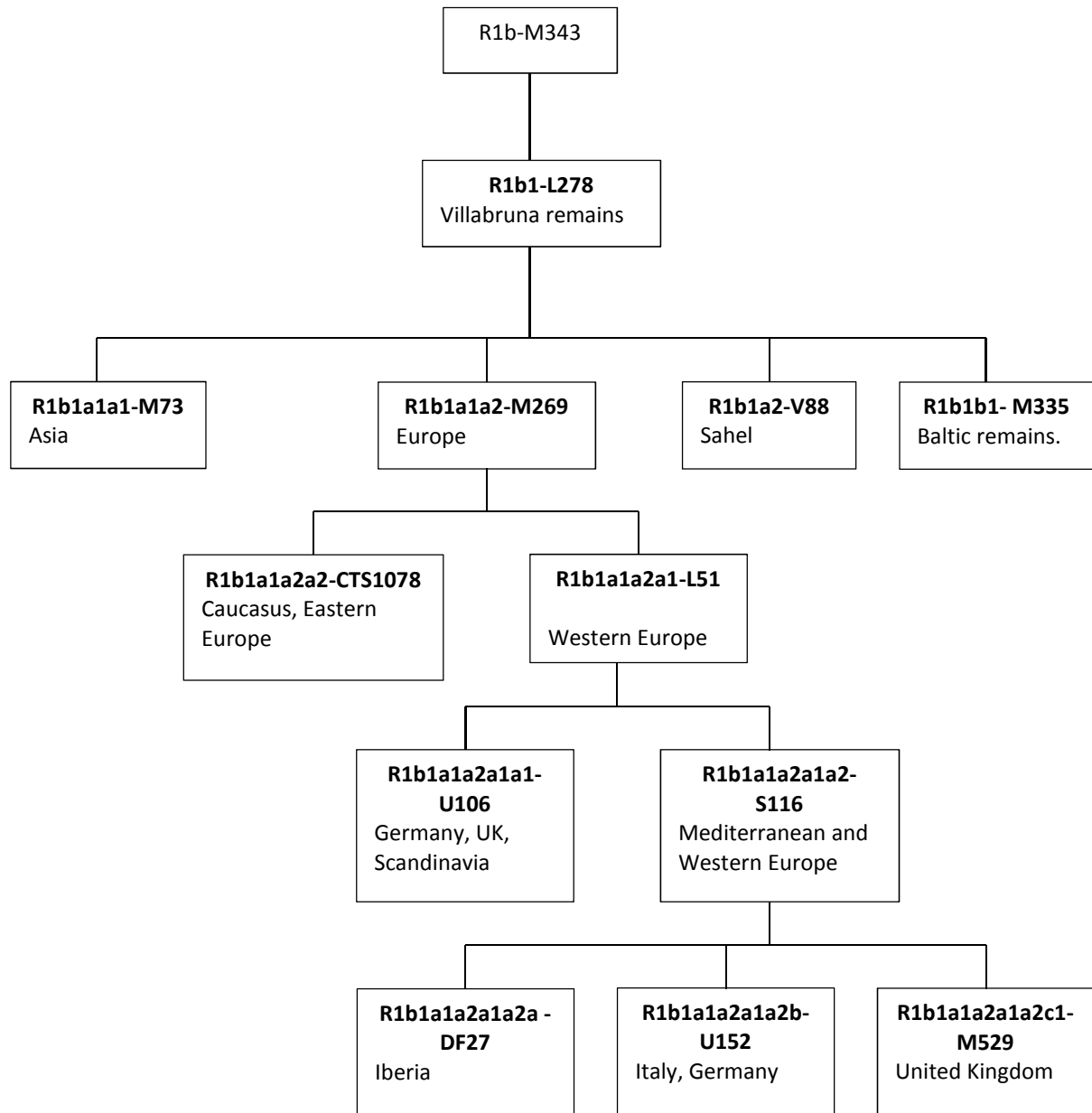
Nomenclature follows ISOGG 2017.



Paper 5.17. Haplogroup R.

The Genetic-Linguistic Interface.

Figure 5.17.1: Part C. Linguistically Informative R1b-M343 Mutations.
Nomenclature generally conforms to ISOGG 2017.





Paper 5.17. Haplogroup R.

The Genetic-Linguistic Interface.

Figure 5.17.2. The East European Plain.



Source: Wikipedia

https://en.wikipedia.org/wiki/East_European_Plain#/media/File:Europe_landforms_East_European_Lowlands.svg

In the public domain.



Paper 5.17. Haplogroup R.

The Genetic-Linguistic Interface.

Figure 5.17.3. The North European Plain.



Source: Wikipedia

https://en.wikipedia.org/wiki/North_European_Plain#/media/File:Europe_landforms_-_North_European_Plain.svg

In the public domain.



Paper 5.17. Haplogroup R-M207.

The Genetic-Linguistic Interface.

Important Comments Concerning Tables

Table 5.17.1. Distribution of Mesolithic R1b-M343 Mutations.

“Death” refers to the age of the sample in years.

Table 5.17.2. Distribution of Mesolithic R1a-M420 Mutations.

“Death” refers to the age of the sample in years.

Table 5.17.8. Contemporary Distribution of R1b1a2-V88 (Regional View).

Upstream R1b mutations reported by Hassan et al. (2008) are assumed to be R1a-V88 as suggested by Cruciani et al. (2010).

Table 5.17.9. Contemporary Distribution of R1b1a2-V88 (Language View).

Upstream R1b mutations reported by Hassan et al. (2008) are assumed to be R1a-V88 as suggested by Cruciani et al. (2010).

Table 5.17.10. R1b1a1a2a2-CTS1078 aDNA.

“Death” refers to the age of the sample in years.

Table 5.17.11. Contemporary distribution of R1b1a1a2a2-CTS1078 (Regional View).

Data from Myres et al. (2011) and Busby et al. (2012) are R1b-M269 (xL51) as suggested by Balanovsky et al. (2017). Frequency results for CTS1078 could be overinflated, especially for Western Europe.

Table 5.17.13. Contemporary Distribution of R1a1a1b2-Z93 (Regional View).

Previously reported upstream R1a mutations for South Asia are assumed to be R1a-Z93, as suggested by Underhill et al. (2015).

Table 5.17.16. Contemporary Distribution of R1a1a1b2- Z93 (Language View).

Previously reported upstream R1a mutations for South Asia are assumed to be R1a-Z93, as suggested by Underhill et al. (2015).

Table 5.17.18. R2a-M124 aDNA.

“Death” refers to the age of the sample in years.

Table 5.17.1. Distribution of Mesolithic R1b-M343 Mutations.

St Clair Ref	Reference No. from Study	Death	Region	Location	Mutation	Reference
72	I4630	9222	Baltic	Zvejnieki, Latvia	R1b1a1a-P297	Mathieson et al. 2018
2	Latvia_Hg2	8000	Baltic	Zvejnieki, Latvia	R1b1b-M335	Jones et al. 2017
67	I5235	10835	Eastern Europe	Padina, Serbia	R1b1a-L761	Mathieson et al. 2018
71	I4081	9335	Eastern Europe	Ostrovul Corbului, Romania	R1b1a-L761	Mathieson et al. 2018
73	I1734	9202	Eastern Europe	Vasil'evka, Ukraine	R1b1a-L761	Mathieson et al. 2018
76	I4916	8763	Eastern Europe	Hadučka Vodenica, Serbia	R1b1a-L761	Mathieson et al. 2018
77	I5411	8600	Eastern Europe	Schela Cladovei, Romania	R1b-M343	Mathieson et al. 2018
78	I5772	8450	Eastern Europe	Vlasac, Serbia	R1b1a-L761	Mathieson et al. 2018
79	I4666	8017	Eastern Europe	Lepenski Vir, Serbia	R1b1a-L761	Mathieson et al. 2018
33	I0124	7549	Eastern Europe	Lebyanzhinka, Sok River, Samara, Russia.	R1b1a1a-P297	Mathieson et al. 2015
84	I5408	7450	Eastern Europe	Ostrovul Corbului, Romania	R1b1a-L761	Mathieson et al. 2018
1	Villabruna	14000	Mediterranean	Dolomite Mountains, northern Italy	R1b1-L278	Fu et al. 2016

Table 5.17.2. Distribution of Mesolithic R1a-M420 Mutations.

St. Clair Reference No.	Standard aDNA Reference	Death	Region	Location	Mutation	Reference
32	I0061	8280	Baltic	Yuzhnyy Oleni Ostrov, Karelia, Russia.	R1a-M420	Mathieson et al. 2015
69	I1819	10643	Eastern Europe	Vasil'evka, Ukraine	R1a-M420	Mathieson et al. 2018
74	I5876	8822	Eastern Europe	Dereivka, Ukraine	R1a-M420	Mathieson et al. 2018

Table 5.17.3. Contemporary Distribution of R1b1a1a1- M73 (Regional View).

Region	Location	Population	Language Famil	Sub Family 1	Sample Size	Percent R1b-M73	Reference
Caucasus	Caucasus Russia	Nogais, Kara	Turkic		76	16%	Yunubayev et al. 2012
Caucasus	Caucasus Russia	Balkars	Turkic		135	10%	Yunubayev et al. 2012
Caucasus	Northwest Caucasus	Balkars	Turkic		136	10%	Myres et al. 2011
Eastern Europe	Bashkortostan , Russia	Bashkir	Turkic		586	13%	Myres et al. 2011
Eastern Europe	Russia, Volga- Ural Region	Mordvins	Uralic		59	10%	Trofimova et al. 2015
Northern Eurasia	Siberia, Russia	Shors	Turkic		38	13%	Malyarchuk et al. 2011
Northern Eurasia	Siberia, Russia	Teleuts	Turkic		44	11%	Malyarchuk et al. 2011
South Asia	North Pakistan	Hazara	Indo-European	Iranian	25	32%	Sengupta et al. 2006
South Asia	Pakistan	Hazara	Indo-European	Iranian	25	32%	Di Cristofaro et al. 2013

Table 5.17.4. Contemporary Distribution of R1b1a1a2a1a1-U106 (Language View).

Language Family	Sub Family 1	Region	Population	Sample Size	Percent R1b-U106	Reference
Indo-European	Celtic	Western Europe	Scots	168	12%	Busby et al. 2012
Indo-European	Celtic	Western Europe	Welsh	129	10%	Busby et al. 2012
Indo-European	Germanic	Western Europe	Frisians	94	43%	Busby et al. 2012
Indo-European	Germanic	Western Europe	Dutch	87	37%	Myres et al. 2011
Indo-European	Germanic	Western Europe	Germans	102	27%	Busby et al. 2012
Indo-European	Germanic	Western Europe	English	102	25%	Myres et al. 2011
Indo-European	Germanic	Western Europe	Germans	349	24%	Rebala et al. 2013
Indo-European	Germanic	Scandinavia	Danes	35	23%	Cruciani et al. 2011b
Indo-European	Germanic	Western Europe	Austrians	18	22%	Myres et al. 2011
Indo-European	Germanic	Western Europe	English	337	22%	Busby et al. 2012
Indo-European	Germanic	Western Europe	Germans	321	21%	Myres et al. 2011
Indo-European	Germanic	Western Europe	Germans	77	21%	Cruciani et al. 2011b
Indo-European	Germanic	Scandinavia	Norwegians	138	20%	Busby et al. 2012
Indo-European	Germanic	Scandinavia	Danes	110	18%	Myres et al. 2011
Indo-European	Germanic	Scandinavia	Danes	174	18%	Valverde et al. 2016
Indo-European	Germanic	Western Europe	Orkney Islanders	112	17%	Busby et al. 2012
Indo-European	West Slavic	Eastern Europe	Poles	40	10%	Cruciani et al. 2011b

Table 5.17.5. Contemporary Distribution of R1b1a1a2a1a2a -DF27 (Language View).

Language Family	Sub Family 1	Region	Population	Sample Size	Percent R1b-DF27	Reference
Indo-European	Italic	Western Europe	Spanish	644	45%	Valverde et al. 2016
Indo-European	Italic	Western Europe	Portuguese	110	41%	Valverde et al. 2016
Indo-European	Italic	Western Europe	French	145	17%	Valverde et al. 2016
Isolate - Basque		Western Europe	Basques	341	63%	Valverde et al. 2016

Table 5.17.6. Contemporary Distribution of R1b1a1a2a1a2b-U152 (Language View).

Language Family	Sub Family 1	Region	Population	Sample Size	Percent R1b-U152	Reference
Indo-European	Germanic	Western Europe	Germans	102	18%	Busby et al. 2012
Indo-European	Germanic	Western Europe	Germans	77	12%	Cruciani et al. 2011b
Indo-European	Germanic	Western Europe	Germans	321	10%	Myres et al. 2011
Indo-European	Italic	Mediterranean	Corsicans	24	42%	Busby et al. 2012
Indo-European	Italic	Mediterranean	Italians	34	35%	Cruciani et al. 2011b
Indo-European	Italic	Mediterranean	Corsicans	28	32%	Cruciani et al. 2011b
Indo-European	Italic	Western Europe	French	43	21%	Cruciani et al. 2011b
Indo-European	Italic	Mediterranean	Italians	282	20%	Myres et al. 2011
Indo-European	Italic	Mediterranean	Italians	884	19%	Boattini et al. 2013
Indo-European	Italic	Mediterranean	Italians	262	18%	Cruciani et al. 2011b
Indo-European	Italic	Mediterranean	Italians	432	17%	Busby et al. 2012
Indo-European	Italic	Western Europe	French	93	15%	Myres et al. 2011
Indo-European	Italic	Western Europe	French	558	14%	Ramos-Luis, E. et al. 2014
Indo-European	Italic	Mediterranean	Sicilians	258	14%	Busby et al. 2012
Indo-European	Italic	Western Europe	French	482	13%	Busby et al. 2012
Indo-European	Italic	Mediterranean	Sicilians	86	12%	Cruciani et al. 2011b
Indo-European	Italic	Mediterranean	Sardinians	270	10%	Cruciani et al. 2011b

Table 5.17.7. Contemporary Distribution of R1b1a1a2a1a2c1-M529 (Language View).

Language Family	Sub Family 1	Region	Population	Sample Size	Percent R1b-M529	Reference
Indo-European	Celtic	Western Europe	Irish	377	0.74	Busby et al. 2012
Indo-European	Celtic	Western Europe	Irish	99	0.59	Myres et al. 2011
Indo-European	Celtic	Western Europe	Irish	146	0.54	Valverde et al. 2016
Indo-European	Celtic	Western Europe	Scots	168	0.52	Busby et al. 2012
Indo-European	Celtic	Western Europe	Welsh	129	0.46	Busby et al. 2012
Indo-European	Germanic	Western Europe	Orkney Islanders	112	0.34	Busby et al. 2012
Indo-European	Germanic	Western Europe	English	102	0.25	Myres et al. 2011
Indo-European	Germanic	Western Europe	English	337	0.15	Busby et al. 2012
Indo-European	Italic	Western Europe	French	145	0.52	Valverde et al. 2016
Indo-European	Italic	Western Europe	French	482	0.16	Busby et al. 2012

Table 5.17.8. Contemporary Distribution of R1b1a2-V88 (Regional View).

Region	Location	Population	Language Family	Sub Family 1	Sample Size	Percent R1b-V88	Reference
Eastern Europe	Bulgaria	Sephardic Bulgarians	Indo-European	South Slavic	20	5%	D’Atanasio et al. 2018
Mediterranean	Italy	Sardinians	Indo-European	Italic	375	1%	D’Atanasio et al. 2018
Middle East	Jordan: Amman & Dead Sea	Jordanians	Afro-Asiatic	Semitic	146	14%	Myres et al. 2011
North Africa	Sudan	Hausa	Afro-Asiatic	Chadic	32	41%	Hassan et al. 2008
North Africa	Egypt	Egyptian Berbers from Siwa	Afro-Asiatic	Semitic	93	26%	D’Atanasio et al. 2018
North Africa	Sudan	Copts	Afro-Asiatic	Egyptian	33	15%	Hassan et al. 2008
North Africa	Algeria	Mozabite Berbers	Afro-Asiatic	Berber	67	12%	Cruciani et al. 2010b
North Africa	Sudan	Nubians	Nilo-Saharan		39	10%	Hassan et al. 2008
North Africa	Libya	Taureg	Afro-Asiatic	Berber	47	6%	Ottoni et al. 2011
North Africa	Egypt	Northern Egyptians	Afro-Asiatic	Semitic	52	6%	D’Atanasio et al. 2018
North Africa	Egypt	Egyptians from Baharia	Afro-Asiatic	Semitic	41	5%	D’Atanasio et al. 2018
North Africa	Baharia, Egypt	Egyptians	Afro-Asiatic	Semitic	41	5%	Cruciani et al. 2010b
North Africa	Sudan	Beja	Afro-Asiatic	Cushitic	42	5%	Hassan et al. 2008
North Africa	Egypt	Northern Egyptians	Afro-Asiatic	Semitic	49	4%	Cruciani et al. 2010b
North Africa	Morocco	Ouarzazate Berbers	Afro-Asiatic	Berber	32	3%	Cruciani et al. 2010b
North Africa	Morocco	Ouarzazate Berbers	Afro-Asiatic	Berber	31	3%	D’Atanasio et al. 2018

Region	Location	Population	Language Family	Sub Family 1	Sample Size	Percent R1b-V88	Reference
North Africa	Southern Egypt	Egyptians	Afro-Asiatic	Semitic	69	3%	Cruciani et al. 2010b
North Africa	Siwa, Egypt	Egyptian Berbers	Afro-Asiatic	Berber	93	2%	Cruciani et al. 2010b
North Africa	Algeria	Mozabite Berbersb	Afro-Asiatic	Berber	67	2%	D’Atanasio et al. 2018
North Africa	Libya	Libian Arabs	Afro-Asiatic	Semitic	71	1%	D’Atanasio et al. 2018
Sub-Saharan Africa	Northern Cameroon	Mafa	Afro-Asiatic	Chadic	8	88%	Cruciani et al. 2010b
Sub-Saharan Africa	Cameroon	Mada	Afro-Asiatic	Chadic	17	82%	Cruciani et al. 2010b
Sub-Saharan Africa	Cameroon	Uldeme	Afro-Asiatic	Chadic	22	64%	Cruciani et al. 2010b
Sub-Saharan Africa	Southern Nigeria	Yoruba	Niger-Congo	Non-Bantoid	21	48%	Cruciani et al. 2010b
Sub-Saharan Africa	Northern Cameroon	Shuwa Arabs	Afro-Asiatic	Semitic	5	40%	Cruciani et al. 2010b
Sub-Saharan Africa	Chad	Toubou	Nilo-Saharan		71	34%	Haber et al. 2016
Sub-Saharan Africa	Niger	Fula Zinder	Niger-Congo	Non-Bantoid	16	31%	Bučková et al 2013
Sub-Saharan Africa	Niger	Fula Diffa	Niger-Congo	Non-Bantoid	13	31%	Bučková et al 2013
Sub-Saharan Africa	Northern Cameroon	Mundang	Niger-Congo	Non-Bantoid	21	29%	Cruciani et al. 2010b
Sub-Saharan Africa	Northern Cameroon	Masa	Afro-Asiatic	Chadic	7	29%	Cruciani et al. 2010b
Sub-Saharan Africa	Chad	Laal	Isolate - Laal		26	23%	Haber et al. 2016
Sub-Saharan Africa	Northern Cameroon	Giziga	Afro-Asiatic	Chadic	9	22%	Cruciani et al. 2010b
Sub-Saharan Africa	Chad	Sara	Nilo-Saharan		61	20%	Haber et al. 2016
Sub-Saharan Africa	Chad	Ngambay	Nilo-Saharan		11	18%	Cruciani et al. 2010b
Sub-Saharan Africa	Niger	Fulbe	Niger-Congo	Non-Bantoid	7	14%	Cruciani et al. 2010b

Region	Location	Population	Language Family	Sub Family 1	Sample Size	Percent R1b-V88	Reference
Sub-Saharan Africa	Northern Cameroon	Kanuri	Nilo-Saharan		7	14%	Cruciani et al. 2010b
Sub-Saharan Africa	Northern Cameroon	Gidar	Afro-Asiatic	Chadic	9	11%	Cruciani et al. 2010b
Sub-Saharan Africa	Northern Cameroon	Daba	Afro-Asiatic	Chadic	19	11%	Cruciani et al. 2010b
Sub-Saharan Africa	Nigeria	Hausa	Afro-Asiatic	Chadic	10	10%	Cruciani et al. 2010b
Sub-Saharan Africa	Niger	Fula Ader	Niger-Congo	Non-Bantoid	20	10%	Bučková et al 2013
Sub-Saharan Africa	Northern Cameroon	Fali	Niger-Congo	Non-Bantoid	48	8%	Cruciani et al. 2010b
Sub-Saharan Africa	Burkina Faso	Fula Balatungur	Niger-Congo	Non-Bantoid	13	8%	Bučková et al 2013
Sub-Saharan Africa	Northern Cameroon	Fulbe	Niger-Congo	Non-Bantoid	18	6%	Cruciani et al. 2010b
Sub-Saharan Africa	Niger	Fula Abala	Niger-Congo	Non-Bantoid	21	5%	Bučková et al 2013
Sub-Saharan Africa	Cameroon	Fula Tcheboua	Niger-Congo	Non-Bantoid	22	5%	Bučková et al 2013
Sub-Saharan Africa	Niger	Berber	Afro-Asiatic	Berber	22	5%	Cruciani et al. 2010b
Sub-Saharan Africa	Chad	Fula Bongor	Niger-Congo	Non-Bantoid	25	4%	Bučková et al 2013
Sub-Saharen Africa	Cameroon (North)	Ouldeme	Afro-Asiatic	Chadic	24	96%	D’Atanasio et al. 2018
Sub-Saharen Africa	Cameroon (North)	Mafa	Afro-Asiatic	Chadic	9	88%	D’Atanasio et al. 2018
Sub-Saharen Africa	Cameroon (North)	Podowko	Afro-Asiatic	Chadic	5	80%	D’Atanasio et al. 2018
Sub-Saharen Africa	Cameroon (North)	Guiziga	Afro-Asiatic	Chadic	9	77%	D’Atanasio et al. 2018
Sub-Saharen Africa	Cameroon (North)	Toupouri	Niger-Congo	Non-Bantoid	31	71%	D’Atanasio et al. 2018

Region	Location	Population	Language Family	Sub Family 1	Sample Size	Percent R1b-V88	Reference
Sub-Saharan Africa	Cameroon (North)	Mada	Afro-Asiatic	Chadic	10	70%	D’Atanasio et al. 2018
Sub-Saharan Africa	Cameroon (North)	Guidar	Afro-Asiatic	Chadic	9	67%	D’Atanasio et al. 2018
Sub-Saharan Africa	Cameroon (North)	Moundang	Niger-Congo	Non-Bantoid	21	67%	D’Atanasio et al. 2018
Sub-Saharan Africa	Cameroon (North)	Daba	Afro-Asiatic	Chadic	29	51%	D’Atanasio et al. 2018
Sub-Saharan Africa	Cameroon (North)	Other Chadic	Afro-Asiatic	Chadic	8	50%	D’Atanasio et al. 2018
Sub-Saharan Africa	Cameroon (North)	Mandara	Afro-Asiatic	Chadic	82	43%	D’Atanasio et al. 2018
Sub-Saharan Africa	Chad	Moundang	Niger-Congo	Non-Bantoid	5	40%	D’Atanasio et al. 2018
Sub-Saharan Africa	Chad	Toupouri	Niger-Congo	Non-Bantoid	13	39%	D’Atanasio et al. 2018
Sub-Saharan Africa	Cameroon (North)/Chad	Shuwa Arabs	Afro-Asiatic	Semitic	6	33%	D’Atanasio et al. 2018
Sub-Saharan Africa	Cameroon (North)	Massa	Afro-Asiatic	Chadic	7	28%	D’Atanasio et al. 2018
Sub-Saharan Africa	Chad	Sara	Nilo-Saharan		12	25%	D’Atanasio et al. 2018
Sub-Saharan Africa	Cameroon (North)	Fali	Niger-Congo	Non-Bantoid	105	22%	D’Atanasio et al. 2018
Sub-Saharan Africa	Chad	Other Chadic	Afro-Asiatic	Chadic	5	20%	D’Atanasio et al. 2018
Sub-Saharan Africa	Cameroon (North)/Chad	Fulbe	Niger-Congo	Non-Bantoid	79	19%	D’Atanasio et al. 2018
Sub-Saharan Africa	Chad	Ngambai	Nilo-Saharan		12	17%	D’Atanasio et al. 2018
Sub-Saharan Africa	Cameroon (North)	Kanuri	Nilo-Saharan		7	14%	D’Atanasio et al. 2018
Sub-Saharan Africa	Chad	Goulaye	Nilo-Saharan		7	14%	D’Atanasio et al. 2018

Region	Location	Population	Language Family	Sub Family 1	Sample Size	Percent R1b-V88	Reference
Sub-Saharen Africa	Niger	Fulbe	Niger-Congo	Non-Bantoid	7	14%	D’Atanasio et al. 2018
Sub-Saharen Africa	Nigeria (North)	Hausa	Afro-Asiatic	Chadic	9	11%	D’Atanasio et al. 2018
Sub-Saharen Africa	Chad	Gor	Nilo-Saharan		9	11%	D’Atanasio et al. 2018
Sub-Saharen Africa	Cameroon (North)	Tali	Niger-Congo	Non-Bantoid	22	10%	D’Atanasio et al. 2018
Sub-Saharen Africa	Cameroon (North)	Ngambai	Nilo-Saharan		11	9%	D’Atanasio et al. 2018
Sub-Saharen Africa	Chad	Massa	Afro-Asiatic	Chadic	11	9%	D’Atanasio et al. 2018
Sub-Saharen Africa	Chad	Madjingay	Nilo-Saharan		15	7%	D’Atanasio et al. 2018
Sub-Saharen Africa	Nigeria (South)	Yoruba	Niger-Congo	Non-Bantoid	20	5%	D’Atanasio et al. 2018
Sub-Saharen Africa	Niger	Tuareg	Afro-Asiatic	Berber	22	5%	D’Atanasio et al. 2018
Sub-Saharen Africa	Cameroon (South)	Ewondo	Niger-Congo	Bantoid	31	3%	D’Atanasio et al. 2018

Table 5.17.9. Contemporary Distribution of R1b1a2-V88 (Language View).

Language Family	Sub Family 1	Region	Population	Location	Sample Size	Percent R1b-V88	Reference
Afro-Asiatic	Berber	North Africa	Mozabite Berbers	Algeria	67	12%	Cruciani et al. 2010b
Afro-Asiatic	Berber	North Africa	Taureg	Libya	47	6%	Otoni et al. 2011
Afro-Asiatic	Berber	Sub-Saharen Africa	Tuareg	Niger	22	5%	D’Atanasio et al. 2018
Afro-Asiatic	Berber	Sub-Saharan Africa	Berber	Niger	22	5%	Cruciani et al. 2010b
Afro-Asiatic	Berber	North Africa	Ouarzazate Berbers	Morocco	32	3%	Cruciani et al. 2010b
Afro-Asiatic	Berber	North Africa	Ouarzazate Berbers	Morocco	31	3%	D’Atanasio et al. 2018
Afro-Asiatic	Berber	North Africa	Egyptian Berbers	Siwa, Egypt	93	2%	Cruciani et al. 2010b
Afro-Asiatic	Berber	North Africa	Mozabite Berbersb	Algeria	67	2%	D’Atanasio et al. 2018
Afro-Asiatic	Chadic	Sub-Saharen Africa	Ouldeme	Cameroon (North)	24	96%	D’Atanasio et al. 2018
Afro-Asiatic	Chadic	Sub-Saharen Africa	Mafa	Cameroon (North)	9	88%	D’Atanasio et al. 2018
Afro-Asiatic	Chadic	Sub-Saharan Africa	Mafa	Northern Cameroon	8	88%	Cruciani et al. 2010b
Afro-Asiatic	Chadic	Sub-Saharan Africa	Mada	Cameroon	17	82%	Cruciani et al. 2010b
Afro-Asiatic	Chadic	Sub-Saharen Africa	Podowko	Cameroon (North)	5	80%	D’Atanasio et al. 2018
Afro-Asiatic	Chadic	Sub-Saharen Africa	Guiziga	Cameroon (North)	9	77%	D’Atanasio et al. 2018
Afro-Asiatic	Chadic	Sub-Saharen Africa	Mada	Cameroon (North)	10	70%	D’Atanasio et al. 2018
Afro-Asiatic	Chadic	Sub-Saharen Africa	Guidar	Cameroon (North)	9	67%	D’Atanasio et al. 2018

Language Family	Sub Family 1	Region	Population	Location	Sample Size	Percent R1b-V88	Reference
Afro-Asiatic	Chadic	Sub-Saharan Africa	Uldeme	Cameroon	22	64%	Cruciani et al. 2010b
Afro-Asiatic	Chadic	Sub-Saharen Africa	Daba	Cameroon (North)	29	51%	D’Atanasio et al. 2018
Afro-Asiatic	Chadic	Sub-Saharen Africa	Other Chadic	Cameroon (North)	8	50%	D’Atanasio et al. 2018
Afro-Asiatic	Chadic	Sub-Saharen Africa	Mandara	Cameroon (North)	82	43%	D’Atanasio et al. 2018
Afro-Asiatic	Chadic	North Africa	Hausa	Sudan	32	41%	Hassan et al. 2008
Afro-Asiatic	Chadic	Sub-Saharan Africa	Masa	Northern Cameroon	7	29%	Cruciani et al. 2010b
Afro-Asiatic	Chadic	Sub-Saharen Africa	Massa	Cameroon (North)	7	28%	D’Atanasio et al. 2018
Afro-Asiatic	Chadic	Sub-Saharan Africa	Giziga	Northern Cameroon	9	22%	Cruciani et al. 2010b
Afro-Asiatic	Chadic	Sub-Saharen Africa	Other Chadic	Chad	5	20%	D’Atanasio et al. 2018
Afro-Asiatic	Chadic	Sub-Saharan Africa	Gidar	Northern Cameroon	9	11%	Cruciani et al. 2010b
Afro-Asiatic	Chadic	Sub-Saharen Africa	Hausa	Nigeria (North)	9	11%	D’Atanasio et al. 2018
Afro-Asiatic	Chadic	Sub-Saharan Africa	Daba	Northern Cameroon	19	11%	Cruciani et al. 2010b
Afro-Asiatic	Chadic	Sub-Saharan Africa	Hausa	Nigeria	10	10%	Cruciani et al. 2010b
Afro-Asiatic	Chadic	Sub-Saharen Africa	Massa	Chad	11	9%	D’Atanasio et al. 2018
Afro-Asiatic	Cushitic	North Africa	Beja	Sudan	42	5%	Hassan et al. 2008
Afro-Asiatic	Egyptian	North Africa	Copts	Sudan	33	15%	Hassan et al. 2008
Afro-Asiatic	Semitic	Sub-Saharan Africa	Shuwa Arabs	Northern Cameroon	5	40%	Cruciani et al. 2010b

Language Family	Sub Family 1	Region	Population	Location	Sample Size	Percent R1b-V88	Reference
Afro-Asiatic	Semitic	Sub-Saharen Africa	Shuwa Arabs	Cameroon (North)/Chad	6	33%	D’Atanasio et al. 2018
Afro-Asiatic	Semitic	North Africa	Egyptian Berbers from Siwa	Egypt	93	26%	D’Atanasio et al. 2018
Afro-Asiatic	Semitic	Middle East	Jordanians	Jordan: Amman & Dead Sea	146	14%	Myres et al. 2011
Afro-Asiatic	Semitic	North Africa	Northern Egyptians	Egypt	52	6%	D’Atanasio et al. 2018
Afro-Asiatic	Semitic	North Africa	Egyptians from Baharia	Egypt	41	5%	D’Atanasio et al. 2018
Afro-Asiatic	Semitic	North Africa	Egyptians	Baharia, Egypt	41	5%	Cruciani et al. 2010b
Afro-Asiatic	Semitic	North Africa	Northern Egyptians	Egypt	49	4%	Cruciani et al. 2010b
Afro-Asiatic	Semitic	North Africa	Egyptians	Southern Egypt	69	3%	Cruciani et al. 2010b
Afro-Asiatic	Semitic	North Africa	Libian Arabs	Libya	71	1%	D’Atanasio et al. 2018
Indo-European	Italic	Mediterranean	Sardinians	Italy	375	1%	D’Atanasio et al. 2018
Indo-European	South Slavic	Eastern Europe	Sephardic Bulgarians	Bulgaria	20	5%	D’Atanasio et al. 2018
Isolate - Laal		Sub-Saharan Africa	Laal	Chad	26	23%	Haber et al. 2016
Niger-Congo	Bantoid	Sub-Saharen Africa	Ewondo	Cameroon (South)	31	3%	D’Atanasio et al. 2018
Niger-Congo	Non-Bantoid	Sub-Saharen Africa	Toupouri	Cameroon (North)	31	71%	D’Atanasio et al. 2018
Niger-Congo	Non-Bantoid	Sub-Saharen Africa	Moundang	Cameroon (North)	21	67%	D’Atanasio et al. 2018
Niger-Congo	Non-Bantoid	Sub-Saharan Africa	Yoruba	Southern Nigeria	21	48%	Cruciani et al. 2010b

Language Family	Sub Family 1	Region	Population	Location	Sample Size	Percent R1b-V88	Reference
Niger-Congo	Non-Bantoid	Sub-Saharen Africa	Moundang	Chad	5	40%	D’Atanasio et al. 2018
Niger-Congo	Non-Bantoid	Sub-Saharen Africa	Toupouri	Chad	13	39%	D’Atanasio et al. 2018
Niger-Congo	Non-Bantoid	Sub-Saharan Africa	Fula Zinder	Niger	16	31%	Bučková et al 2013
Niger-Congo	Non-Bantoid	Sub-Saharan Africa	Fula Diffa	Niger	13	31%	Bučková et al 2013
Niger-Congo	Non-Bantoid	Sub-Saharan Africa	Mundang	Northern Cameroon	21	29%	Cruciani et al. 2010b
Niger-Congo	Non-Bantoid	Sub-Saharen Africa	Fali	Cameroon (North)	105	22%	D’Atanasio et al. 2018
Niger-Congo	Non-Bantoid	Sub-Saharen Africa	Fulbe	Cameroon (North)/Chad	79	19%	D’Atanasio et al. 2018
Niger-Congo	Non-Bantoid	Sub-Saharan Africa	Fulbe	Niger	7	14%	Cruciani et al. 2010b
Niger-Congo	Non-Bantoid	Sub-Saharen Africa	Fulbe	Niger	7	14%	D’Atanasio et al. 2018
Niger-Congo	Non-Bantoid	Sub-Saharan Africa	Fula Ader	Niger	20	10%	Bučková et al 2013
Niger-Congo	Non-Bantoid	Sub-Saharen Africa	Tali	Cameroon (North)	22	10%	D’Atanasio et al. 2018
Niger-Congo	Non-Bantoid	Sub-Saharan Africa	Fali	Northern Cameroon	48	8%	Cruciani et al. 2010b
Niger-Congo	Non-Bantoid	Sub-Saharan Africa	Fula Balatungur	Burkina Faso	13	8%	Bučková et al 2013
Niger-Congo	Non-Bantoid	Sub-Saharan Africa	Fulbe	Northern Cameroon	18	6%	Cruciani et al. 2010b
Niger-Congo	Non-Bantoid	Sub-Saharen Africa	Yoruba	Nigeria (South)	20	5%	D’Atanasio et al. 2018
Niger-Congo	Non-Bantoid	Sub-Saharan Africa	Fula Abala	Niger	21	5%	Bučková et al 2013
Niger-Congo	Non-Bantoid	Sub-Saharan Africa	Fula Tcheboua	Cameroon	22	5%	Bučková et al 2013
Niger-Congo	Non-Bantoid	Sub-Saharan Africa	Fula Bongor	Chad	25	4%	Bučková et al 2013
Nilo-Saharan		Sub-Saharan Africa	Toubou	Chad	71	34%	Haber et al. 2016

Language Family	Sub Family 1	Region	Population	Location	Sample Size	Percent R1b-V88	Reference
Nilo-Saharan		Sub-Saharen Africa	Sara	Chad	12	25%	D’Atanasio et al. 2018
Nilo-Saharan		Sub-Saharan Africa	Sara	Chad	61	20%	Haber et al. 2016
Nilo-Saharan		Sub-Saharan Africa	Ngambay	Chad	11	18%	Cruciani et al. 2010b
Nilo-Saharan		Sub-Saharen Africa	Ngambai	Chad	12	17%	D’Atanasio et al. 2018
Nilo-Saharan		Sub-Saharan Africa	Kanuri	Northern Cameroon	7	14%	Cruciani et al. 2010b
Nilo-Saharan		Sub-Saharen Africa	Kanuri	Cameroon (North)	7	14%	D’Atanasio et al. 2018
Nilo-Saharan		Sub-Saharen Africa	Goulaye	Chad	7	14%	D’Atanasio et al. 2018
Nilo-Saharan		Sub-Saharen Africa	Gor	Chad	9	11%	D’Atanasio et al. 2018
Nilo-Saharan		North Africa	Nubians	Sudan	39	10%	Hassan et al. 2008
Nilo-Saharan		Sub-Saharen Africa	Ngambai	Cameroon (North)	11	9%	D’Atanasio et al. 2018
Nilo-Saharan		Sub-Saharen Africa	Madjingay	Chad	15	7%	D’Atanasio et al. 2018

Table 5.17.10. R1b1a1a2a2-CTS1078 aDNA.

St Clair Ref	aDNA Ref	Region	Location	Death	Culture	Mutation	Reference
21	RISE555.SG	Eastern Europe	Stalingrad Quarry, Russia	4627	Bronze Age	R1b1a1a2a2-CTS1078	Allentoft et al. 2015
26	RISE547.SG	Northern Eurasia	Temrta, Russia	4711	Yamnaya	R1b1a1a2a2-CTS1078	Allentoft et al. 2015
27	RISE548.SG	Northern Eurasia	Temrta, Russia.	4650	Yamnaya	R1b1a1a2a2-CTS1078	Allentoft et al. 2015
38	I0429	Eastern Europe	Lopatino, Sok River, Samara, Russia	5079	Yamnaya	R1b1a1a2a2-CTS1078	Mathieson et al. 2015
41	I0370	Eastern Europe	Ishkinovka, Eastern Orenburg, Pre-Ural steppe, Samara	4950	Yamnaya	R1b1a1a2a2-CTS1078	Mathieson et al. 2015
43	I0231	Eastern Europe	Ekaterinovka, Southern Steppe, Samara	4792	Yamnaya	R1b1a1a2a2-CTS1078	Mathieson et al. 2015
44	I0440	Eastern Europe	Lopatino, Sok River, Samara, Russia	4727	Yamnaya	R1b1a1a2a2-CTS1078	Mathieson et al. 2015
47	I0126	Eastern Europe	Kutuluk River, Samara, Russia	4627	Bronze Age	R1b1a1a2a2-CTS1078	Mathieson et al. 2015
48	I0374	Eastern Europe	Nikolaevka III, Samara River, Samara, Russia	4450	Bronze Age	R1b1a1a2a2-CTS1078	Mathieson et al. 2015
95	I5884	Eastern Europe	Dereivka, Ukraine	4743	Neolithic	R1b1a1a2a2-CTS1078	Mathieson et al. 2018
96	I3499	Eastern Europe	Beli Manastir-Popova Zemlja, Croatia	4725	Neolithic	R1b1a1a2a2-CTS1078	Mathieson et al. 2018
101	I5270	Northern Eurasia	Elo, Russia	5081	Afanasievo	R1b1a1a2a2-CTS1078	Narasimhan et al. 2019
102	I2071	Northern Eurasia	Altai Mountains, Russia	4968	Afanasievo	R1b1a1a2a2-CTS1078	Narasimhan et al. 2019
103	I5278	Northern Eurasia	Lower Tyumechin, Altai Mountains, Russia	4948	Afanasievo	R1b1a1a2a2-CTS1078	Narasimhan et al. 2019
104	I7671	Eastern Europe	Samara Oblast, Volga River Valley, Bobrovka, Russia	4733	Bronze Age	R1b1a1a2a2-CTS1078	Narasimhan et al. 2019

St Clair Ref	aDNA Ref	Region	Location	Death	Culture	Mutation	Reference
106	I6712	Northern Eurasia	Middle Yenisei region, Afanasieva Gora, Russia	4632	Afanasievo	R1b1a1a2a2-CTS1078	Narasimhan et al. 2019

Table 5.17.11. Contemporary distribution of R1b1a1a2a2-CTS1078 (Regional View).

Region	Location	Population	Language Family	Sub Family 1	Sample Size	Percent CTS 1078	Reference
Baltic	Estonia	Estonians	Uralic	Finnic	210	1%	Myres et al. 2011
Caucasus	Northeast Caucasus	Bagvalals	North Caucasian		28	68%	Myres et al. 2011
Caucasus	Northeast Caucasus	Tabasarans	North Caucasian		43	40%	Myres et al. 2011
Caucasus	Russia: Adygea	Armenians Adigei	Indo-European	Armenian	49	20%	Balanovski et al. 2017
Caucasus	Northeast Caucasus	Lezgis	North Caucasian		31	16%	Myres et al. 2011
Caucasus	Georgia	Armenians Erzurum	Indo-European	Armenian	99	15%	Balanovski et al. 2017
Caucasus	Northeast Caucasus	Kumyks	Turkic		76	15%	Myres et al. 2011
Caucasus	Daghestan	Daghestani	North Caucasian		49	14%	Busby et al. 2012
Caucasus	Russia: Krasnodar	Armenians Krasnodar	Indo-European	Armenian	117	12%	Balanovski et al. 2017
Caucasus	Russia: Rostov and Myasnikovsky Regions	Armenians Don	Indo-European	Armenian	92	9%	Balanovski et al. 2017
Caucasus	Georgia	Imereti	Kartvelian		62	6%	Balanovski et al. 2017
Caucasus	Northeast Caucasus	Andis	North Caucasian		49	6%	Myres et al. 2011
Caucasus	Northwest Caucasus	Ossetians	Indo-European	Iranian	134	5%	Myres et al. 2011
Caucasus	Northeast Caucasus	Karachays	Turkic		69	4%	Myres et al. 2011
Caucasus	South Caucasus	Armenians	Indo-European	Armenian	26	4%	Myres et al. 2011

Region	Location	Population	Language Family	Sub Family 1	Sample Size	Percent CTS 1078	Reference
Caucasus	Northwest Caucasus	Abazas	North Caucasian		89	3%	Myres et al. 2011
Caucasus	South Caucasus	Abkhaz	North Caucasian		162	3%	Myres et al. 2011
Caucasus	Northwest Caucasus	Balkars	Turkic		136	3%	Myres et al. 2011
Caucasus	Northeast Caucasus	Dargins	North Caucasian		68	3%	Myres et al. 2011
Caucasus	Northwest Caucasus	Kabardians	North Caucasian		141	3%	Myres et al. 2011
Caucasus	Northeast Caucasus	Avars	North Caucasian		42	2%	Myres et al. 2011
Caucasus	South Caucasus	Megrels	Kartvelian		67	2%	Myres et al. 2011
Central Asia	Afghanistan	Afghans	Indo-European	Iranian	51	6%	Busby et al. 2012
Eastern Europe	Bashkortostan, Russia	Bashkir	Turkic		586	23%	Myres et al. 2011
Eastern Europe	Kosovo	Kosovans	Indo-European	South Slavic	114	19%	Myres et al. 2011
Eastern Europe	Hungary	Hungarians	Uralic		113	8%	Myres et al. 2011
Eastern Europe	Bashkortostan, Russia	Tatars	Turkic		39	8%	Myres et al. 2011
Eastern Europe	Romania	Romanians	Indo-European	Italic	330	7%	Myres et al. 2011
Eastern Europe	Poland	Poles	Indo-European	West Slavic	202	7%	Myres et al. 2011
Eastern Europe	Poland	Poles	Indo-European	West Slavic	174	6%	Busby et al. 2012
Eastern Europe	Czech Republic	Czechs	Indo-European	West Slavic	87	6%	Myres et al. 2011
Eastern Europe	Slovenia	Slovenes	Indo-European	South Slavic	102	6%	Myres et al. 2011
Eastern Europe	Serbia	Serbians	Indo-European	South Slavic	113	5%	Myres et al. 2011
Eastern Europe	Macedonia	Macedonians	Indo-European	South Slavic	79	5%	Myres et al. 2011
Eastern Europe	Slovakia	Slovaks	Indo-European	West Slavic	276	5%	Myres et al. 2011

Region	Location	Population	Language Family	Sub Family 1	Sample Size	Percent CTS 1078	Reference
Eastern Europe	Romania	Romanians	Indo-European	Italic	59	5%	Busby et al. 2012
Eastern Europe	Russia	Urdmurts	Uralic	Permian	54	4%	Myres et al. 2011
Eastern Europe	Belarus	Belarusians	Indo-European	East Slavic	205	3%	Myres et al. 2011
Eastern Europe	Bulgaria	Bulgarians	Indo-European	South Slavic	30	3%	Busby et al. 2012
Eastern Europe	Croatia	Croatians	Indo-European	South Slavic	108	3%	Myres et al. 2011
Eastern Europe	Herzegovina	Herzegovinans	Indo-European	South Slavic	141	3%	Myres et al. 2011
Eastern Europe	Kazan, Russia	Tatars	Turkic		80	3%	Myres et al. 2011
Eastern Europe	Russia	Russians	Indo-European	East Slavic	1037	2%	Myres et al. 2011
Eastern Europe	Ukraine	Ukrainians	Indo-European	East Slavic	504	2%	Myres et al. 2011
Eastern Europe	Ukraine	Ukrainians	Indo-European	East Slavic	55	2%	Busby et al. 2012
Eastern Europe	Russia	Chuvash	Turkic		117	2%	Myres et al. 2011
Eastern Europe	Croatia	Croatians	Indo-European	East Slavic	144	1%	Busby et al. 2012
Eastern Europe	Bosnia	Bosnians	Indo-European	South Slavic	78	1%	Myres et al. 2011
Mediterranean	Albania	Albanians	Indo-European	Albanian	158	13%	Busby et al. 2012
Mediterranean	Sicily	Sicilians	Indo-European	Italic	258	12%	Busby et al. 2012
Mediterranean	Italy	Italians	Indo-European	Italic	432	9%	Busby et al. 2012
Mediterranean	Greece	Greeks	Indo-European	Greek	185	9%	Myres et al. 2011
Mediterranean	Italy	Italians	Indo-European	Italic	282	8%	Myres et al. 2011
Mediterranean	Crete	Cretans	Indo-European	Greek	193	8%	Myres et al. 2011
Mediterranean	Greece	Greeks	Indo-European	Greeks	167	5%	Busby et al. 2012
Mediterranean	Corsica	Corsicans	Indo-European	Italic	24	4%	Busby et al. 2012
Middle East	Turkey	Turks	Turkic		322	20%	Busby et al. 2012
Middle East	Turkey	Turks	Turkic		611	14%	Myres et al. 2011
Middle East	Iran	Iranians	Indo-European	Iranian	328	10%	Busby et al. 2012

Region	Location	Population	Language Family	Sub Family 1	Sample Size	Percent CTS 1078	Reference
Middle East	Iran	Iranians	Indo-European	Iranian	150	8%	Myres et al. 2011
Middle East	Syria	Syrians	Afro-Asiatic	Semitic	72	7%	Busby et al. 2012
Middle East	Iraq	Iraqi	Afro-Asiatic	Semitic	33	3%	Busby et al. 2012
Middle East	Jordan	Jordanians	Afro-Asiatic	Semitic	76	3%	Myres et al. 2011
Middle East	Jordan: Amman & Dead Sea	Jordanians	Afro-Asiatic	Semitic	146	2%	Myres et al. 2011
Northern Eurasia	Perm Oblast, Russia	Komi	Uralic	Permian	61	12%	Myres et al. 2011
Scandinavia	Sweden South (Malmö)	Swedes	Indo-European	Germanic	139	5%	Myres et al. 2011
Scandinavia	Denmark	Danes	Indo-European	Germanic	110	4%	Myres et al. 2011
Scandinavia	Norway	Norwegians	Indo-European	Germanic	138	1%	Busby et al. 2012
Western Europe	Germans	Germans	Indo-European	Germanic	102	13%	Busby et al. 2012
Western Europe	Portugal	Portuguese	Indo-European	Italic	300	10%	Busby et al. 2012
Western Europe	France	French	Indo-European	Italic	482	5%	Busby et al. 2012
Western Europe	England	English	Indo-European	Germanic	102	4%	Myres et al. 2011
Western Europe	Spain	Spanish	Indo-European	Italic	359	4%	Busby et al. 2012
Western Europe	England	English	Indo-European	Germanic	337	4%	Busby et al. 2012
Western Europe	Germany	Germans	Indo-European	Germanic	321	4%	Myres et al. 2011
Western Europe	Orkney	Orkney Islanders	Indo-European	Germanic	112	4%	Busby et al. 2012
Western Europe	Ireland	Irish	Indo-European	Celtic	377	3%	Busby et al. 2012
Western Europe	France	French	Indo-European	Italic	93	3%	Myres et al. 2011
Western Europe	Spain	Spanish	Indo-European	Italic	454	2%	Myres et al. 2011
Western Europe	Ireland	Irish	Indo-European	Celtic	99	2%	Myres et al. 2011
Western Europe	Wales	Welsh	Indo-European	Celtic	129	2%	Busby et al. 2012

Table 5.17.12. Contemporary Distribution of R1a1a1b1a-Z282 (Regional View).

Region	Location	Population	Language Family	Sub Family 1	Sample Size	Percent R1a-Z282	Reference
Baltic	Russia	Karelians	Uralic	Finnic	140	40%	Underhill et al. 2010; Underhill et al. 2015
Baltic	Russia	Veps	Uralic	Finnic	39	36%	Underhill et al. 2010; Underhill et al. 2015
Baltic	Estonia	Estonians	Uralic	Finnic	236	31%	Underhill et al. 2010; Underhill et al. 2015
Eastern Europe	Belarus	Belarusians	Indo-European	East Slavic	364	56%	Underhill et al. 2010; Kushniarevich et al 2013; Underhill et al. 2015
Eastern Europe	Russia	Russians	Indo-European	East Slavic	433	53%	Underhill et al. 2010; Underhill et al. 2015
Eastern Europe	Poland	Poles	Indo-European	West Slavic	202	50%	Underhill et al. 2010; Underhill et al. 2015
Eastern Europe	Ukraine	Ukrainians	Indo-European	East Slavic	498	47%	Underhill et al. 2010; Underhill et al. 2015
Eastern Europe	Slovakia	Slovaks	Indo-European	West Slavic	257	46%	Underhill et al. 2010; Underhill et al. 2015
Eastern Europe	Czech Republic	Czechs	Indo-European	West Slavic	53	38%	Underhill et al. 2010; Underhill et al. 2015
Eastern Europe	Russia	Mari	Uralic		51	29%	Underhill et al. 2010; Underhill et al. 2015
Eastern Europe	Croatia	Croatians	Indo-European	South Slavic	108	29%	Underhill et al. 2010; Underhill et al. 2015
Eastern Europe	Poland	Poles	Indo-European	West Slavic	520	28%	Rebala et al. 2013
Eastern Europe	Croatia	Croatians	Indo-European	South Slavic	720	26%	Sarac et al. 2016

Region	Location	Population	Language Family	Sub Family 1	Sample Size	Percent R1a-Z282	Reference
Eastern Europe	Tatarstan, Russia	Tatars	Turkic		66	24%	Underhill et al. 2010; Underhill et al. 2015
Eastern Europe	Slovenia	Slovenes	Indo-European	West Slavic	180	20%	Underhill et al. 2010; Underhill et al. 2015
Northern Eurasia	Russia	Komi	Uralic	Permian	112	21%	Underhill et al. 2010; Underhill et al. 2015
Scandinavia	Norway	Norwegians	Indo-European	Germanic	118	21%	Underhill et al. 2015
Western Europe	Germany	Sorbs	Indo-European	West Slavic	123	57%	Rebala et al. 2013

Table 5.17.13. Contemporary Distribution of R1a1a1b2-Z93 (Regional View).

Region	Location	Population	Language Family	Branch	Sample Size	Percent R1a-Z93	Reference
Caucasus	Daghestan	Tats	Indo-European	Iranian	24	25%	Karafet et al. 2016
Caucasus	Caucasus	Balkars	Turkic		136	25%	Underhill et al. 2010; Underhill et al. 2015
Caucasus	Caucasus	Karachays	Turkic		69	25%	Underhill et al. 2010; Underhill et al. 2015
Central Asia	Afghanistan	Pashtuns	Indo-European	Iranian	84	52%	Di Cristofaro et al. 2013; Underhill et al. 2015
Central Asia	Kyrgyzstan	Kyrgyz	Turkic		232	47%	Underhill et al. 2010; Underhill et al. 2015
Central Asia	Afghanistan	Tajik	Indo-European	Iranian	140	24%	Di Cristofaro et al. 2013; Underhill et al. 2015
Central Asia	Afghanistan	Uzbeks	Turkic		125	23%	Di Cristofaro et al. 2013; Underhill et al. 2015
Northern Eurasia	Russia	Altai (Altaians, Altay)	Turkic		61	38%	Underhill et al. 2010; Underhill et al. 2015
Northern Eurasia	Russia	Khakas	Turkic		76	33%	Underhill et al. 2010; Underhill et al. 2015
South Asia	East India	West Bengal Brahmin	Indo-European	Indo-Aryan	18	72%	Sengupta et al. 2006
South Asia	Gujarat	Pavagadhi Chaudhari	Indo-European	Indo-Aryan	29	62%	Khurana et al. 2014
South Asia	Pakistan	Sindhi	Indo-European	Indo-Aryan	21	52%	Di Cristofaro et al. 2013
South Asia	South Pakistan	Sindhi	Indo-European	Indo-Aryan	21	52%	Sengupta et al. 2006
South Asia	Pakistan	Pathans	Indo-European	Iranian	270	49%	Lee et al. 2014
South Asia	West India	Koknasth Brahmin	Indo-European	Indo-Aryan	25	48%	Sengupta et al. 2006
South Asia	West Bengal	Bengali	Indo-European	Indo-Aryan	54	41%	Debnath et al. 2011
South Asia	Pakistan	Pathans	Indo-European	Iranian	20	40%	Di Cristofaro et al. 2013

Region	Location	Population	Language Family	Branch	Sample Size	Percent R1a-Z93	Reference
South Asia	North Pakistan	Pathans	Indo-European	Iranian	20	40%	Sengupta et al. 2006
South Asia	Northern India	Chamar	Indo-European	Indo-Aryan	18	39%	Sengupta et al. 2006
South Asia	Northern India	Uttar Pradesh Brahmin	Indo-European	Indo-Aryan	14	36%	Sengupta et al. 2006
South Asia	Andra Pradesh	Andh	Indo-European	Indo-Aryan	54	31%	Thanseem et al. 2006
South Asia	Northern India	Rajput	Indo-European	Indo-Aryan	29	31%	Sengupta et al. 2006
South Asia	South India	Iyengar	Dravidian		30	30%	Sengupta et al. 2006
South Asia	Pakistan	Makrani	Indo-European	Indo-Aryan	20	30%	Di Cristofaro et al. 2013
South Asia	Gujarat	Konkana	Indo-European	Indo-Aryan	24	29%	Khurana et al. 2014
South Asia	Maharashtra	Thakur	Indo-European	Indo-Aryan	48	29%	Thangaraj et al. 2010
South Asia	Northeastern India	Rajbanshi	Indo-European	Indo-Aryan	45	29%	Borkar et al. 2011
South Asia	Pakistan	Baloch	Indo-European	Iranian	25	28%	Di Cristofaro et al. 2013
South Asia	South Pakistan	Baloch	Indo-European	Iranian	25	28%	Sengupta et al. 2006
South Asia	South India	Iyer	Dravidian		29	28%	Sengupta et al. 2006
South Asia	Maharashtra	Mahadeo Koli	Indo-European	Indo-Aryan	50	26%	Thangaraj et al. 2010
South Asia	Nepal	Newar	Sino-Tibetan	Tibeto-Burman	66	26%	Gayden et al. 2007
South Asia	South India	Pallan	Dravidian		29	24%	Sengupta et al. 2006
South Asia	South Pakistan	Brahui	Dravidian		25	24%	Sengupta et al. 2006
South Asia	Pakistan	Brahui	Dravidian		25	24%	Di Cristofaro et al. 2013
South Asia	Tamil Nadu, India	Mukkuvar	Dravidian		17	24%	Arunkumar et al. 2012
South Asia	Northeastern India	Paharia	Indo-European	Indo-Aryan	100	23%	Borkar et al. 2011
South Asia	West India	Nav Buddha	Indo-European	Indo-Aryan	14	21%	Sengupta et al. 2006
South Asia	Pakistan	Kalash	Indo-European	Indo-Aryan	19	21%	Sengupta et al. 2006

Region	Location	Population	Language Family	Branch	Sample Size	Percent R1a-Z93	Reference
South Asia	Tamil Nadu, India	Tamil Jains	Dravidian		100	20%	Arunkumar et al. 2012
South Asia	Pakistan	Kalash	Indo-European	Indo-Aryan	20	20%	Di Cristofaro et al. 2013
South Asia	Southwest India	Lingayat	Dravidian		101	20%	Chennakrishnaiah et al. 2013
South Asia	Tamil Nadu, India	Kota	Dravidian		62	19%	Arunkumar et al. 2012
South Asia	Andra Pradesh	Naikpod	Dravidian		68	19%	Thanseem et al. 2006
South Asia	Central India	Halba	Indo-European	Indo-Aryan	21	19%	Sengupta et al. 2006
South Asia	Gujarat	Dhodia	Indo-European	Indo-Aryan	63	19%	Khurana et al. 2014
South Asia	West Bengal	Rajbanshi	Indo-European	Indo-Aryan	51	18%	Debnath et al. 2011
South Asia	Tamil Nadu, India	Mullukurumba	Dravidian		29	17%	Arunkumar et al. 2012
South Asia	Tamil Nadu, India	Valayar	Dravidian		95	16%	Arunkumar et al. 2012
South Asia	Tamil Nadu, India	Maravar	Dravidian		80	15%	Arunkumar et al. 2012
South Asia	Tamil Nadu, India	Vanniyar	Dravidian		21	14%	Arunkumar et al. 2012
South Asia	South India	Ambalakarar	Dravidian		29	14%	Sengupta et al. 2006
South Asia	South India	Vellalar	Dravidian		31	13%	Sengupta et al. 2006
South Asia	Gujarat	Vasava	Indo-European	Indo-Aryan	24	13%	Khurana et al. 2014
South Asia	Andra Pradesh	Pardhan	Dravidian		128	13%	Thanseem et al. 2006
South Asia	Gujarat	Valvi Chaudhari	Indo-European	Indo-Aryan	32	13%	Khurana et al. 2014
South Asia	South India	Kota	Dravidian		16	13%	Sengupta et al. 2006
South Asia	Northeastern India	Oraon	Dravidian		110	12%	Borkar et al. 2011
South Asia	Tamil Nadu, India	Thoda	Dravidian		26	12%	Arunkumar et al. 2012

Region	Location	Population	Language Family	Branch	Sample Size	Percent R1a-Z93	Reference
South Asia	Tamil Nadu, India	Paravar	Dravidian		27	11%	Arunkumar et al. 2012
South Asia	Tamil Nadu, India	Yadhava	Dravidian		107	10%	Arunkumar et al. 2012
South Asia	Tamil Nadu, India	Nadar TNV	Dravidian		59	10%	Arunkumar et al. 2012
South Asia	North Pakistan	Burusho	Isolate - Burushaski		20	10%	Sengupta et al. 2006
South Asia	East India	Agharia	Austro-Asiatic	Munda	10	10%	Sengupta et al. 2006
South Asia	West India	Maratha	Indo-European	Indo-Aryan	20	10%	Sengupta et al. 2006
South Asia	Pakistan	Burusho	Isolate - Burushaski		20	10%	Di Cristofaro et al. 2013
South Asia	Tamil Nadu, India	Pallar	Dravidian		51	10%	Arunkumar et al. 2012
South Asia	Tamil Nadu, India	Parayar NTN	Dravidian		52	10%	Arunkumar et al. 2012
South Asia	Gujarat	Dubla	Indo-European	Indo-Aryan	42	10%	Khurana et al. 2014
South Asia	Tamil Nadu, India	Vanniyar NTN	Dravidian		96	9%	Arunkumar et al. 2012
South Asia	Nepal	Tharu	Indo-European	Indo-Aryan	171	9%	Fornarino et al. 2009
South Asia	Tamil Nadu, India	Parayar	Dravidian		24	8%	Arunkumar et al. 2012
South Asia	West Bengal	Dhimal	Sino-Tibetan	Tibeto-Burman	36	8%	Debnath et al. 2011
South Asia	Nepal	Tharu	Indo-European	Indo-Aryan	170	8%	Fornarino et al. 2009; Underhill et al. 2015
South Asia	Gujarat	Nana Chaudhari	Indo-European	Indo-Aryan	25	8%	Khurana et al. 2014
South Asia	South India	Vanniyar	Dravidian		25	8%	Sengupta et al. 2006
South Asia	Southwest India	Vokkaliga	Dravidian		102	8%	Chennakrishnaiah et al. 2013
South Asia	East India	Mahishya	Indo-European	Indo-Aryan	13	8%	Sengupta et al. 2006

Region	Location	Population	Language Family	Branch	Sample Size	Percent R1a-Z93	Reference
South Asia	Gujarat	Mota Chaudhari	Indo-European	Indo-Aryan	27	7%	Khurana et al. 2014
South Asia	South India	Konda Reddy	Dravidian		30	7%	Sengupta et al. 2006
South Asia	Northeast India	Jamatia	Sino-Tibetan	Tibeto-Burman	30	7%	Sengupta et al. 2006
South Asia	Tamil Nadu, India	Betta Kurumba	Dravidian		17	6%	Arunkumar et al. 2012
South Asia	Tamil Nadu, India	Kanikaran	Dravidian		17	6%	Arunkumar et al. 2012
South Asia	Tamil Nadu, India	Kurumba	Dravidian		35	6%	Arunkumar et al. 2012
South Asia	Gujarat	Gamit	Indo-European	Indo-Aryan	18	6%	Khurana et al. 2014
South Asia	Tamil Nadu, India	Ezhava	Dravidian		95	5%	Arunkumar et al. 2012
South Asia	Tamil Nadu, India	Sourashtra	Indo-European	Indo-Aryan	40	5%	Arunkumar et al. 2012
South Asia	Northeastern India	Mech	Sino-Tibetan	Tibeto-Burman	20	5%	Borkar et al. 2011
South Asia	Northeastern India	Lachung	Sino-Tibetan	Tibeto-Burman	20	5%	Borkar et al. 2011
South Asia	Northeast India	Tripuri	Sino-Tibetan	Tibeto-Burman	21	5%	Sengupta et al. 2006
South Asia	Tamil Nadu, India	Brahacharanam	Indo-European	Indo-Aryan	21	5%	Arunkumar et al. 2012
South Asia	Nepal	Tamang	Sino-Tibetan	Tibeto-Burman	45	4%	Gayden et al. 2007
South Asia	Tamil Nadu, India	Kattunaickan	Dravidian		46	4%	Arunkumar et al. 2012
South Asia	Tamil Nadu, India	Nadar Cape	Dravidian		98	4%	Arunkumar et al. 2012

Region	Location	Population	Language Family	Branch	Sample Size	Percent R1a-Z93	Reference
South Asia	Tamil Nadu, India	Piramalai Kallar	Dravidian		53	4%	Arunkumar et al. 2012
South Asia	Northeastern India	Ho	Austro-Asiatic	Munda	28	4%	Borkar et al. 2011
South Asia	Bangladesh	Marma	Sino-Tibetan	Tibeto-Burman	60	3%	Gazi et al. 2013
South Asia	West Bengal	Oraon	Dravidian		31	3%	Debnath et al. 2011
South Asia	West Bengal	Kol	Austro-Asiatic	Munda	62	3%	Debnath et al. 2011
South Asia	Northeastern India	Birhor	Austro-Asiatic	Munda	100	3%	Borkar et al. 2011
South Asia	West Bengal	Kharia	Austro-Asiatic	Munda	34	3%	Debnath et al. 2011
South Asia	Tamil Nadu, India	Irula	Dravidian		80	3%	Arunkumar et al. 2012
South Asia	Northeastern India	Munda	Austro-Asiatic	Munda	94	2%	Borkar et al. 2011
South Asia	West Bengal	Santhal	Austro-Asiatic	Munda	51	2%	Debnath et al. 2011
South Asia	Bangladesh	Chakma	Indo-European	Indo-Aryan	89	1%	Gazi et al. 2013
South Asia	Northeastern India	Santhal	Austro-Asiatic	Munda	90	1%	Borkar et al. 2011
South Asia	Tamil Nadu, India	Kadar	Dravidian		28	0%	Arunkumar et al. 2012
South Asia	Tamil Nadu, India	Vadama	Indo-European	Indo-Aryan	63	0%	Arunkumar et al. 2012
South Asia	Bangladesh	Tripuri	Sino-Tibetan	Tibeto-Burman	88	0%	Gazi et al. 2013
South Asia	East India	Bagdi	Indo-European	Indo-Aryan	11	0%	Sengupta et al. 2006
South Asia	West Bengal	Rabha	Sino-Tibetan	Tibeto-Burman	26	0%	Debnath et al. 2011

Region	Location	Population	Language Family	Branch	Sample Size	Percent R1a-Z93	Reference
South Asia	Tamil Nadu, India	Paliyan	Dravidian		95	0%	Arunkumar et al. 2012
South Asia	Tamil Nadu, India	Iyengar	Indo-European	Indo-Aryan	11	0%	Arunkumar et al. 2012
South Asia	Sikkim	Lachungpa	Sino-Tibetan	Tibeto-Burman	11	0%	Debnath et al. 2011
South Asia	South India	Kurumba	Dravidian		19	0%	Sengupta et al. 2006
South Asia	North Pakistan	Hazara	Indo-European	Iranian	25	0%	Sengupta et al. 2006
South Asia	Pakistan	Hazara	Indo-European	Iranian	25	0%	Di Cristofaro et al. 2013
South Asia	South India	Irula	Dravidian		30	0%	Sengupta et al. 2006
South Asia	East India	Lodha	Indo-European	Indo-Aryan	20	0%	Sengupta et al. 2006
South Asia	Tamil Nadu, India	Pulayar	Dravidian		63	0%	Arunkumar et al. 2012

Table 5.17.14. Contemporary Distribution of R1a1a1b1a1-M458 (Language View).

Language Family	Sub Family 1	Region	Location	Population	Sample Size	Percent R1a-M458	Reference
Indo-European	East Slavic	Eastern Europe	Belarus	Belarusians	364	15%	Underhill et al. 2010; Kushniarevich et al. 2013; Underhill et al. 2015
Indo-European	East Slavic	Eastern Europe	Belarus	Belarusians	565	15%	Kushniarevich et al. 2013
Indo-European	East Slavic	Eastern Europe	Lviv Region, Ukraine	Ukrainians	154	14%	Mielnik-Sikorska et al. 2013.
Indo-European	East Slavic	Eastern Europe	Ukraine	Ukrainians	498	13%	Underhill et al. 2010; Underhill et al. 2015
Indo-European	East Slavic	Eastern Europe	Russia	Russians	433	12%	Underhill et al. 2010; Underhill et al. 2015
Uralic	Finnic	Baltic	Russia	Veps	39	15%	Underhill et al. 2010; Underhill et al. 2015
Uralic	Finnic	Baltic	Russia	Karelians	140	14%	Underhill et al. 2010; Underhill et al. 2015
Indo-European	Italic	Eastern Europe	Moldavia	Moldavians	125	13%	Varzari et al. 2013
Indo-European	South Slavic	Eastern Europe	Croatia	Croatians	108	19%	Underhill et al. 2010; Underhill et al. 2015
Indo-European	West Slavic	Western Europe	Germany	Sorbs	123	57%	Rebala et al. 2013
Indo-European	West Slavic	Eastern Europe	Czech Republic	Czechs	53	30%	Underhill et al. 2010; Underhill et al. 2015
Indo-European	West Slavic	Eastern Europe	Poland	Poles	520	28%	Rebala et al. 2013
Indo-European	West Slavic	Eastern Europe	Poland	Poles	202	25%	Underhill et al. 2010; Underhill et al. 2015
Indo-European	West Slavic	Eastern Europe	Slovakia	Slovaks	257	21%	Underhill et al. 2010; Underhill et al. 2015
Indo-European	West Slavic	Eastern Europe	Slovakia	Slovaks	164	16%	Rebala et al. 2013

Table 5.17.15. Contemporary Distribution of R1a1a1b1a2b-M558 (Language View).

Language Family	Sub Family 1	Region	Location	Population	Sample Size	Percent R1a-M558	Reference
Indo-European	Armenian	Caucasus	Russia: Rostov and Myasnikovsky Regions	Armenians Don	92	13%	Balanovski et al. 2017
Indo-European	East Slavic	Eastern Europe	Russia	Russians	433	26%	Underhill et al. 2010; Underhill et al. 2015
Indo-European	East Slavic	Eastern Europe	Belarus	Belarusians	364	23%	Underhill et al. 2010; Kushniarevich et al. 2013; Underhill et al. 2015
Indo-European	East Slavic	Eastern Europe	Ukraine	Ukrainians	498	18%	Underhill et al. 2010; Underhill et al. 2015
Indo-European	South Slavic	Eastern Europe	Croatia	Croatians	720	19%	Sarac et al. 2016
Indo-European	South Slavic	Eastern Europe	Bosnia	Bosnians	78	12%	Underhill et al. 2010; Underhill et al. 2015
Indo-European	West Slavic	Eastern Europe	Slovakia	Slovaks	257	22%	Underhill et al. 2010; Underhill et al. 2015
Indo-European	West Slavic	Eastern Europe	Poland	Poles	202	19%	Underhill et al. 2010; Underhill et al. 2015
Indo-European	West Slavic	Eastern Europe	Slovenia	Slovenes	180	17%	Underhill et al. 2010; Underhill et al. 2015
Turkic		Eastern Europe	Tatarstan, Russia	Tatars	66	15%	Underhill et al. 2010; Underhill et al. 2015
Turkic		Eastern Europe	Russia, Bashkortostan	Tatars	39	13%	Underhill et al. 2010; Underhill et al. 2015
Turkic		Eastern Europe	Russia	Chuvash	117	11%	Underhill et al. 2010; Underhill et al. 2015
Uralic		Eastern Europe	Russia	Mari	51	24%	Underhill et al. 2010; Underhill et al. 2015

Language Family	Sub Family 1	Region	Location	Population	Sample Size	Percent R1a-M558	Reference
Uralic		Eastern Europe	Hungary	Hungarians	113	12%	Underhill et al. 2010; Underhill et al. 2015
Uralic	Finnic	Baltic	Estonia	Estonians	236	19%	Underhill et al. 2010; Underhill et al. 2015
Uralic	Finnic	Baltic	Russia	Veps	39	15%	Underhill et al. 2010; Underhill et al. 2015
Uralic	Finnic	Baltic	Russia	Karelians	140	10%	Underhill et al. 2010; Underhill et al. 2015
Uralic	Permian	Eastern Europe	Russia	Udmurt	54	11%	Underhill et al. 2010; Underhill et al. 2015

Table 5.17.16. Contemporary Distribution of R1a1a1b2- Z93 (Language View).

Language Family	Branch	Region	Population	Location	Percent R1a-Z93	Reference
Austro-Asiatic	Munda	South Asia	Agharia	East India	10%	Sengupta et al. 2006
Dravidian		South Asia	Iyengar	South India	30%	Sengupta et al. 2006
Dravidian		South Asia	Iyer	South India	28%	Sengupta et al. 2006
Dravidian		South Asia	Pallan	South India	24%	Sengupta et al. 2006
Dravidian		South Asia	Brahui	South Pakistan	24%	Sengupta et al. 2006
Dravidian		South Asia	Brahui	Pakistan	24%	Di Cristofaro et al. 2013
Dravidian		South Asia	Mukkuvar	Tamil Nadu, India	24%	Arunkumar et al. 2012
Dravidian		South Asia	Tamil Jains	Tamil Nadu, India	20%	Arunkumar et al. 2012
Dravidian		South Asia	Lingayat	Southwest India	20%	Chennakrishnaiah et al. 2013
Dravidian		South Asia	Kota	Tamil Nadu, India	19%	Arunkumar et al. 2012
Dravidian		South Asia	Naikpod	Andra Pradesh	19%	Thanseem et al. 2006
Dravidian		South Asia	Mullukurumba	Tamil Nadu, India	17%	Arunkumar et al. 2012
Dravidian		South Asia	Valayar	Tamil Nadu, India	16%	Arunkumar et al. 2012
Dravidian		South Asia	Maravar	Tamil Nadu, India	15%	Arunkumar et al. 2012
Dravidian		South Asia	Vanniyar	Tamil Nadu, India	14%	Arunkumar et al. 2012

Language Family	Branch	Region	Population	Location	Percent R1a-Z93	Reference
Dravidian		South Asia	Ambalakarar	South India	14%	Sengupta et al. 2006
Dravidian		South Asia	Vellalar	South India	13%	Sengupta et al. 2006
Dravidian		South Asia	Pardhan	Andra Pradesh	13%	Thanseem et al. 2006
Dravidian		South Asia	Kota	South India	13%	Sengupta et al. 2006
Dravidian		South Asia	Oraon	Northeastern India	12%	Borkar et al. 2011
Dravidian		South Asia	Thoda	Tamil Nadu, India	12%	Arunkumar et al. 2012
Dravidian		South Asia	Paravar	Tamil Nadu, India	11%	Arunkumar et al. 2012
Dravidian		South Asia	Yadhava	Tamil Nadu, India	10%	Arunkumar et al. 2012
Dravidian		South Asia	Nadar TNV	Tamil Nadu, India	10%	Arunkumar et al. 2012
Indo-European	Indo-Aryan	South Asia	West Bengal Brahmin	East India	72%	Sengupta et al. 2006
Indo-European	Indo-Aryan	South Asia	Pavagadhi Chaudhari	Gujarat	62%	Khurana et al. 2014
Indo-European	Indo-Aryan	South Asia	Sindhi	South Pakistan	52%	Sengupta et al. 2006
Indo-European	Indo-Aryan	South Asia	Sindhi	Pakistan	52%	Di Cristofaro et al. 2013
Indo-European	Indo-Aryan	South Asia	Koknasth Brahmin	West India	48%	Sengupta et al. 2006
Indo-European	Indo-Aryan	South Asia	Bengali	West Bengal	41%	Debnath et al. 2011
Indo-European	Indo-Aryan	South Asia	Chamar	Northern India	39%	Sengupta et al. 2006

Language Family	Branch	Region	Population	Location	Percent R1a-Z93	Reference
Indo-European	Indo-Aryan	South Asia	Uttar Pradesh Brahmin	Northern India	36%	Sengupta et al. 2006
Indo-European	Indo-Aryan	South Asia	Andh	Andra Pradesh	31%	Thanseem et al. 2006
Indo-European	Indo-Aryan	South Asia	Rajput	Northern India	31%	Sengupta et al. 2006
Indo-European	Indo-Aryan	South Asia	Makrani	Pakistan	30%	Di Cristofaro et al. 2013
Indo-European	Indo-Aryan	South Asia	Konkana	Gujarat	29%	Khurana et al. 2014
Indo-European	Indo-Aryan	South Asia	Thakur	Maharashtra	29%	Thangaraj et al. 2010
Indo-European	Indo-Aryan	South Asia	Rajbanshi	Northeastern India	29%	Borkar et al. 2011
Indo-European	Indo-Aryan	South Asia	Mahadeo Koli	Maharashtra	26%	Thangaraj et al. 2010
Indo-European	Indo-Aryan	South Asia	Paharia	Northeastern India	23%	Borkar et al. 2011
Indo-European	Indo-Aryan	South Asia	Nav Buddha	West India	21%	Sengupta et al. 2006
Indo-European	Indo-Aryan	South Asia	Kalash	Pakistan	21%	Sengupta et al. 2006
Indo-European	Indo-Aryan	South Asia	Kalash	Pakistan	20%	Di Cristofaro et al. 2013
Indo-European	Indo-Aryan	South Asia	Dhodia	Gujarat	19%	Khurana et al. 2014
Indo-European	Indo-Aryan	South Asia	Halba	Central India	19%	Sengupta et al. 2006
Indo-European	Indo-Aryan	South Asia	Rajbanshi	West Bengal	18%	Debnath et al. 2011
Indo-European	Indo-Aryan	South Asia	Valvi Chaudhari	Gujarat	13%	Khurana et al. 2014
Indo-European	Indo-Aryan	South Asia	Vasava	Gujarat	13%	Khurana et al. 2014
Indo-European	Indo-Aryan	South Asia	Maratha	West India	10%	Sengupta et al. 2006

Language Family	Branch	Region	Population	Location	Percent R1a-Z93	Reference
Indo-European	Iranian	Central Asia	Pashtuns	Afghanistan	52%	Di Cristofaro et al. 2013; Underhill et al. 2015
Indo-European	Iranian	South Asia	Pathans	Pakistan	49%	Lee et al. 2014
Indo-European	Iranian	South Asia	Pathans	Pakistan	40%	Di Cristofaro et al. 2013
Indo-European	Iranian	South Asia	Pathans	North Pakistan	40%	Sengupta et al. 2006
Indo-European	Iranian	South Asia	Baloch	Pakistan	28%	Di Cristofaro et al. 2013
Indo-European	Iranian	South Asia	Baloch	South Pakistan	28%	Sengupta et al. 2006
Indo-European	Iranian	Caucasus	Tats	Daghestan	25%	Karafet et al. 2016
Indo-European	Iranian	Central Asia	Tajik	Afghanistan	24%	Di Cristofaro et al. 2013; Underhill et al. 2015
Isolate - Burushaski		South Asia	Burusho	Pakistan	10%	Di Cristofaro et al. 2013
Isolate - Burushaski		South Asia	Burusho	North Pakistan	10%	Sengupta et al. 2006
Sino-Tibetan	Tibeto-Burman	South Asia	Newar	Nepal	26%	Gayden et al. 2007
Turkic		Central Asia	Kyrgyz	Kyrgyzstan	47%	Underhill et al. 2010; Underhill et al. 2015
Turkic		Northern Eurasia	Altai (Altaians, Altay)	Russia	38%	Underhill et al. 2010; Underhill et al. 2015
Turkic		Northern Eurasia	Khakas	Russia	33%	Underhill et al. 2010; Underhill et al. 2015
Turkic		Caucasus	Balkars	Caucasus	25%	Underhill et al. 2010; Underhill et al. 2015
Turkic		Caucasus	Karachays	Caucasus	25%	Underhill et al. 2010; Underhill et al. 2015

Language Family	Branch	Region	Population	Location	Percent R1a-Z93	Reference
Turkic		Central Asia	Uzbeks	Afghanistan	23%	Di Cristofaro et al. 2013; Underhill et al. 2015

Table 5.17.17. Contemporary distribution of R2a-M124 (Regional View).

Region	Location	Population	Language Family	Sub Family 1	Sample Size	Percent R2a-M124	Reference
Central Asia	Kazakhstan	Kazakh (Zhanakorgan)	Turkic		94	12%	Zhabagin et al. 2017
East Asia	Northern Thailand	Mon	Mon-Khmer		18	17%	Brunelli et al. 2017
Middle East	Turkey	Armenians (Sasun)	Indo-European	Armenian	104	17%	Hovhannisyan et al. 2014
Northern Eurasia	Siberia, Russia	Kalmyks	Mongolic		60	10%	Malyarchuk et al. 2011
South Asia	Tamil Nadu, India	Paravar	Dravidian		27	37%	Arunkumar et al. 2012
South Asia	East India	Lodha	Indo-European	Indo-Aryan	20	35%	Sengupta et al. 2006
South Asia	East India	Agharia	Austro-Asiatic	Munda	10	30%	Sengupta et al. 2006
South Asia	Tamil Nadu, India	Kadar	Dravidian		28	29%	Arunkumar et al. 2012
South Asia	Tamil Nadu, India	Iyengar	Indo-European	Indo-Aryan	11	27%	Arunkumar et al. 2012
South Asia	Nepal	Newar	Sino-Tibetan	Tibeto-Burman	66	26%	Gayden et al. 2007
South Asia	South India	Kota	Dravidian		16	25%	Sengupta et al. 2006
South Asia	West Bengal	Oraon	Dravidian		31	23%	Debnath et al. 2011
South Asia	East India	West Bengal Brahmin	Indo-European	Indo-Aryan	18	22%	Sengupta et al. 2006
South Asia	Gujarat	Konkana	Indo-European	Indo-Aryan	24	21%	Khurana et al. 2014
South Asia	West India	Koknasth Brahmin	Indo-European	Indo-Aryan	25	20%	Sengupta et al. 2006
South Asia	West India	Maratha	Indo-European	Indo-Aryan	20	20%	Sengupta et al. 2006
South Asia	Gujarat	Valvi Chaudhari	Indo-European	Indo-Aryan	32	19%	Khurana et al. 2014
South Asia	Tamil Nadu, India	Mukkuvar	Dravidian		17	18%	Arunkumar et al. 2012
South Asia	Tamil Nadu, India	Parayar NTN	Dravidian		52	17%	Arunkumar et al. 2012
South Asia	Tamil Nadu, India	Yadhava	Dravidian		107	17%	Arunkumar et al. 2012

Region	Location	Population	Language Family	Sub Family 1	Sample Size	Percent R2a-M124	Reference
South Asia	Tamil Nadu, India	Pulayar	Dravidian		63	16%	Arunkumar et al. 2012
South Asia	Southwest India	Lingayat	Dravidian		101	16%	Chennakrishnaiah et al. 2013
South Asia	West Bengal	Santhal	Austro-Asiatic	Munda	51	16%	Debnath et al. 2011
South Asia	East India	Mahishya	Indo-European	Indo-Aryan	13	15%	Sengupta et al. 2006
South Asia	North Pakistan	Burusho	Isolate - Burushaski		20	15%	Sengupta et al. 2006
South Asia	Pakistan	Burusho	Isolate - Burushaski		20	15%	Di Cristofaro et al. 2013
South Asia	Gujarat	Dhodia	Indo-European	Indo-Aryan	63	14%	Khurana et al. 2014
South Asia	Tamil Nadu, India	Brahacharanam	Indo-European	Indo-Aryan	21	14%	Arunkumar et al. 2012
South Asia	Tamil Nadu, India	Vanniyar	Dravidian		21	14%	Arunkumar et al. 2012
South Asia	Tamil Nadu, India	Maravar	Dravidian		80	14%	Arunkumar et al. 2012
South Asia	Southwest India	Vokkaliga	Dravidian		102	14%	Chennakrishnaiah et al. 2013
South Asia	Tamil Nadu, India	Pallar	Dravidian		51	14%	Arunkumar et al. 2012
South Asia	Tamil Nadu, India	Ezhava	Dravidian		95	13%	Arunkumar et al. 2012
South Asia	Pakistan	Baloch	Indo-European	Iranian	25	12%	Di Cristofaro et al. 2013
South Asia	South Pakistan	Brahui	Dravidian		25	12%	Sengupta et al. 2006
South Asia	South India	Vanniyar	Dravidian		25	12%	Sengupta et al. 2006
South Asia	Maharashtra	Mahadeo Koli	Indo-European	Indo-Aryan	50	12%	Thangaraj et al. 2010
South Asia	Pakistan	Brahui	Dravidian		25	12%	Di Cristofaro et al. 2013
South Asia	Andra Pradesh	Naikpod	Dravidian		68	12%	Thanseem et al. 2006
South Asia	Northern India	Chamar	Indo-European	Indo-Aryan	18	11%	Sengupta et al. 2006
South Asia	Tamil Nadu, India	Tamil Jains	Dravidian		100	11%	Arunkumar et al. 2012
South Asia	Andra Pradesh	Pardhan	Dravidian		128	11%	Thanseem et al. 2006
South Asia	Northeastern India	Ho	Austro-Asiatic	Munda	28	11%	Borkar et al. 2011

Region	Location	Population	Language Family	Sub Family 1	Sample Size	Percent R2a-M124	Reference
South Asia	South India	Kurumba	Dravidian		19	11%	Sengupta et al. 2006
South Asia	South India	Iyer	Dravidian		29	10%	Sengupta et al. 2006
South Asia	South India	Iyengar	Dravidian		30	10%	Sengupta et al. 2006
South Asia	Pakistan	Makrani	Indo-European	Indo-Aryan	20	10%	Di Cristofaro et al. 2013
South Asia	South India	Irula	Dravidian		30	10%	Sengupta et al. 2006

Table 5.17.18. R2a-M124 aDNA.

St. Clair Reference Number	aDNA Reference from Literature	Region	Location	Death (years ago)	Cultural Tradition	Best Resolved Informative Mutation	Reference
99	I4085	Central Asia	Turkestan, Tepe Anau, Turkmenistan	5450	Neolithic	R2a-M124	Narasimhan et al. 2019
112	I7492	Central Asia	Sappali Tepe, Uzbekistan	3827	Bronze Age	R2a-M124	Narasimhan et al. 2019
98	I1954	Middle East	Ganj Dareh, Iran	10093	Neolithic	R2a-M124	Narasimhan et al. 2019

Table 5.17.19. Contemporary Distribution of R2a-M124 (Language View).

Language Family	Sub Family 1	Region	Population	Location	Sample Size	Percent R2a-M124	Reference
Austro-Asiatic	Munda	South Asia	Agharia	East India	10	30%	Sengupta et al. 2006
Austro-Asiatic	Munda	South Asia	Santhal	West Bengal	51	16%	Debnath et al. 2011
Austro-Asiatic	Munda	South Asia	Ho	Northeastern India	28	11%	Borkar et al. 2011
Dravidian		South Asia	Paravar	Tamil Nadu, India	27	37%	Arunkumar et al. 2012
Dravidian		South Asia	Kadar	Tamil Nadu, India	28	29%	Arunkumar et al. 2012
Dravidian		South Asia	Kota	South India	16	25%	Sengupta et al. 2006
Dravidian		South Asia	Oraon	West Bengal	31	23%	Debnath et al. 2011
Dravidian		South Asia	Mukkuvar	Tamil Nadu, India	17	18%	Arunkumar et al. 2012
Dravidian		South Asia	Parayar NTN	Tamil Nadu, India	52	17%	Arunkumar et al. 2012
Dravidian		South Asia	Yadhava	Tamil Nadu, India	107	17%	Arunkumar et al. 2012
Dravidian		South Asia	Pulayar	Tamil Nadu, India	63	16%	Arunkumar et al. 2012
Dravidian		South Asia	Lingayat	Southwest India	101	16%	Chennakrishnaiah et al. 2013
Dravidian		South Asia	Vanniyar	Tamil Nadu, India	21	14%	Arunkumar et al. 2012
Dravidian		South Asia	Maravar	Tamil Nadu, India	80	14%	Arunkumar et al. 2012
Dravidian		South Asia	Vokkaliga	Southwest India	102	14%	Chennakrishnaiah et al. 2013

Language Family	Sub Family 1	Region	Population	Location	Sample Size	Percent R2a-M124	Reference
Dravidian		South Asia	Pallar	Tamil Nadu, India	51	14%	Arunkumar et al. 2012
Dravidian		South Asia	Ezhava	Tamil Nadu, India	95	13%	Arunkumar et al. 2012
Dravidian		South Asia	Vanniyar	South India	25	12%	Sengupta et al. 2006
Dravidian		South Asia	Brahui	Pakistan	25	12%	Di Cristofaro et al. 2013
Dravidian		South Asia	Brahui	South Pakistan	25	12%	Sengupta et al. 2006
Dravidian		South Asia	Naikpod	Andra Pradesh	68	12%	Thanseem et al. 2006
Dravidian		South Asia	Tamil Jains	Tamil Nadu, India	100	11%	Arunkumar et al. 2012
Dravidian		South Asia	Pardhan	Andra Pradesh	128	11%	Thanseem et al. 2006
Dravidian		South Asia	Kurumba	South India	19	11%	Sengupta et al. 2006
Dravidian		South Asia	Iyer	South India	29	10%	Sengupta et al. 2006
Dravidian		South Asia	Irula	South India	30	10%	Sengupta et al. 2006
Dravidian		South Asia	Iyengar	South India	30	10%	Sengupta et al. 2006
Indo-European	Armenian	Middle East	Armenians (Sasun)	Turkey	104	17%	Hovhannisyan et al. 2014
Indo-European	Indo-Aryan	South Asia	Lodha	East India	20	35%	Sengupta et al. 2006
Indo-European	Indo-Aryan	South Asia	Iyengar	Tamil Nadu, India	11	27%	Arunkumar et al. 2012
Indo-European	Indo-Aryan	South Asia	West Bengal Brahmin	East India	18	22%	Sengupta et al. 2006
Indo-European	Indo-Aryan	South Asia	Konkana	Gujarat	24	21%	Khurana et al. 2014
Indo-European	Indo-Aryan	South Asia	Maratha	West India	20	20%	Sengupta et al. 2006
Indo-European	Indo-Aryan	South Asia	Koknasth Brahmin	West India	25	20%	Sengupta et al. 2006

Language Family	Sub Family 1	Region	Population	Location	Sample Size	Percent R2a-M124	Reference
Indo-European	Indo-Aryan	South Asia	Valvi Chaudhari	Gujarat	32	19%	Khurana et al. 2014
Indo-European	Indo-Aryan	South Asia	Mahishya	East India	13	15%	Sengupta et al. 2006
Indo-European	Indo-Aryan	South Asia	Dhodia	Gujarat	63	14%	Khurana et al. 2014
Indo-European	Indo-Aryan	South Asia	Brahacharanam	Tamil Nadu, India	21	14%	Arunkumar et al. 2012
Indo-European	Indo-Aryan	South Asia	Mahadeo Koli	Maharashtra	50	12%	Thangaraj et al. 2010
Indo-European	Indo-Aryan	South Asia	Chamar	Northern India	18	11%	Sengupta et al. 2006
Indo-European	Indo-Aryan	South Asia	Makrani	Pakistan	20	10%	Di Cristofaro et al. 2013
Indo-European	Iranian	South Asia	Baloch	Pakistan	25	12%	Di Cristofaro et al. 2013
Isolate - Burushaski		South Asia	Burusho	North Pakistan	20	15%	Sengupta et al. 2006
Isolate - Burushaski		South Asia	Burusho	Pakistan	20	15%	Di Cristofaro et al. 2013
Mongolic		Northern Eurasia	Kalmyks	Siberia, Russia	60	10%	Malyarchuk et al. 2011
Mon-Khmer		East Asia	Mon	Northern Thailand	18	17%	Brunelli et al. 2017
Sino-Tibetan	Tibeto-Burman	South Asia	Newar	Nepal	66	26%	Gayden et al. 2007
Turkic		Central Asia	Kazakh (Zhanakorgan)	Kazakhstan	94	12%	Zhabagin et al. 2017