

European DNA pages:

Haplogroups; description Haplogroups; timeline Migration maps; Ancient DNA; Y-DNA by country; MtDNA by country

Origins, age, spread and ethnic association of European haplogroups and subclades

Version française

Last update November 2011 (Updated G2a, I1, J2, R1a and R1b)

Content

Introduction

Y-DNA Haplogroups

Chronological timeline

Haplogroup R1b

IE invasion vs acculturation

How did R1b become dominant ?

The upheavals ca 1200 BCE

Haplogroup R1a

Turkic R1a

Haplogroup I

I1

I2

Haplogroup G

Haplogroup J

Haplogroup E

Haplogroup T

Haplogroup N

Haplogroup Q

Haplogroup C

Haplogroup P

Haplogroup L

Haplogroup H

Haplogroup A

MtDNA Haplogroups

European mtDNA haplogroups & subclades

Haplogroups H & V

Haplogroups U & K

Haplogroups J & T

Haplogroup W

Haplogroup I

Haplogroup X

Haplogroup R

Finno-Uralic mtDNA

Berber mtDNA

Gypsy mtDNA

Sources

Useful links

Follow-up

Disclaimer

The information about the origin and ethnic association of haplogroups on this website should not be read as hard facts, but, as is often the case in science, as a model in constant evolution based on the present knowledge and understanding (of the authors). Whenever the advancement of genetics couldn't provide irrefutable answers, we have attempted to provide the most likely and logical hypothesis based on archeological, historical and linguistic evidence. This page is being updated regularly to keep up with recent studies giving additional insights or rectifying possibly erroneous theories. Feel free to add comments or share your opinion on the forum.

Introduction to genetic genealogy

DNA studies have permitted to categorise all humans on Earth in genealogical groups sharing one common ancestor at one given point in prehistory. They are called haplogroups. There are two kinds of haplogroups: the paternally inherited Y-chromosome DNA (Y-DNA) haplogroups, and the maternally inherited mitochondrial DNA (mtDNA) haplogroups. They respectively indicate the agnatic (or patrilineal) and cognatic (or matrilineal) ancestry.

Y-DNA haplogroups are useful to determine whether two apparently unrelated individuals sharing the same surname do indeed descend from a common ancestor in a not too distant past (3 to 20 generations). This is achieved by comparing the haplotypes through the STR markers. Deep SNP testing allows to go back much farther in time, and to identify the ancient ethnic group to which one's ancestors belonged (e.g. Celtic, Germanic, Slavic, Greco-Roman, Basque, Iberian, Phoenician, Jewish, etc.).

In Europe, mtDNA haplogroups are quite evenly spread over the continent, and therefore cannot be associated easily with ancient ethnicities. However, they can sometimes reveal some potential medical conditions (see diseases associated with mtDNA mutations). Some mtDNA subclades are associated with Jewish ancestry, notably K1a1b1a, K1a9,d K2a2a and N1b.

#### DNA Facts

Nucleotides are the alphabet of DNA. There are four of them : adenine (A), thymine (T), guanine (G) and cytosine (C). They always go by pairs, A with T, and G with C. Such pairs are called "base pairs".

The 46 chromosomes of human DNA are composed of a total of 3,000 million base pairs.

The Y-chromosome possess 60 million base pairs, against 153 million for the X chromosome.

Mitochondrial DNA is found outside the cell's nucleus, and therefore outside of the chromosomes. It consists only of 16,569 base pairs.

A SNP (single nucleotide polymorphism) is a mutation in a single base pair. At present, only a few hundreds SNP's define all the human haplogroups for mtDNA or Y-DNA.

#### Y-DNA Haplogroups

##### Chronological development of Y-DNA haplogroups

- K => 40,000 years ago (probably arose in northern Iran)
- T => 30,000 years ago (around the Red Sea or around the Persian Gulf)
- J => 30,000 years ago (in the Middle East)
- R => 28,000 years ago (in the Central Asia)
- E1b1b => 26,000 years ago (in Northeast Africa)
- I => 25,000 years ago (in the Balkans)
- R1a => 21,000 years ago (in southern Russia)
- R1b => 20,000 years ago (around the Caspian Sea or Central Asia)
- J1 => 20,000 years ago (in the Taurus/Zagros mountains)
- J2 => 19,000 years ago (in northern Mesopotamia)
- E-M78 => 18,000 years ago (in north-eastern Africa)
- G => 17,000 years ago (in the Middle East)
- I2 => 17,000 years ago (in the Balkans)
- E-V13 => 14,000 years ago (in the southern Levant or North Africa)

I2b => 13,000 years ago (in Central Europe)  
N1c1 => 12,000 years ago (in Siberia)  
I2a => 11,000 years ago (in the Balkans)  
G2a => 11,000 years ago (in the Levant or Anatolia)  
R1b1b2 => 10,000 years ago (north or south of the Caucasus)  
E-M81 => 9,500 years ago (in Northwest Africa)  
I2b1 => 9,000 years ago (in Germany)  
I2a1 => 8,000 years ago (in Southwest Europe)  
I2a2 => 7,500 years ago (in Southeast Europe)  
I1 => 5,000 years ago (in Scandinavia)  
R1b-L21 => 4,000 years ago (in Central or Eastern Europe)  
R1b-S28 => 3,500 years ago (around the Alps)  
R1b-S21 => 3,000 years ago (in Frisia or Central Europe)  
I2b1a => less than 3,000 years ago (in Britain)

Map of early Bronze Age cultures in Europe around 4,500 to 5,000 years ago

=> More maps of the Neolithic and Bronze Age expansion

## European haplogroups

### Haplogroup R1b (Y-DNA)

Main article : Haplogroup R1b

Distribution of haplogroup R1b in Europe

R1b is the most common haplogroup in Western Europe, reaching over 80% of the population in Ireland, the Scottish Highlands, western Wales, the Atlantic fringe of France and the Basque country. It is also common in Anatolia and around the Caucasus, in parts of Russia and in Central and South Asia. Besides the Atlantic and North Sea coast of Europe, hotspots include the Po valley in north-central Italy (over 70%), the Ossetians of the North Caucasus (over 40%) and nearby Armenia (35%), the Bashkirs of the Urals region of Russia (50%), Turkmenistan (over 35%), the Hazara people of Afghanistan (35%), the Uyghurs of North-West China (20%) and the Newars of Nepal (11%). R1b-V88, a subclade specific to sub-Saharan Africa, is found in 60 to 95% of men in northern Cameroon.

### Anatolian or Caucasian origins ?

The origins of R1b are not entirely clear to this day. Some of the oldest forms of R1b are found in the Near East and around the Caucasus. Haplogroup R1\* and R2\* might have originated in southern Central Asia (between the Caspian and the Hindu Kush). A branch of R1 would have developed into R1b\* then R1b1\* in the northern part of the Middle East during the Ice Age. It presumptively moved to northern Anatolia and across the Caucasus during the early Neolithic, where it became R1b1b. The Near Eastern leftovers evolved into R1b1a (M18), now found at low frequencies among the Lebanese and the Druze. The Phoenicians (who came from modern day Lebanon) spread this R1b1a and R1b1\* to their colonies, notably Sardinia and the Maghreb.

The subclades R1b1b1 and R1b1b2 (the most common form in Europe) are closely associated with the spread of Indo-European languages, as attested by its presence in all regions of the world where Indo-European languages were spoken in ancient times, from the Atlantic coast of Europe to the Indian subcontinent, including almost all Europe (except Finland and Bosnia-Herzegovina), Anatolia, Armenia, European Russia, southern Siberia, many pockets around Central Asia (notably Xinjiang, Turkmenistan, Tajikistan and

Afghanistan), without forgetting Iran, Pakistan, India and Nepal. The history of R1b and R1a are intricately connected to each others. Whereas R1b1 is found in such places as the Levant or Cameroon, R1b1b mostly likely originated in north-eastern Anatolia.

The North Caucasus and the Pontic-Caspian steppe : the Indo-European link

Modern linguists have placed the Proto-Indo-European homeland in the Pontic-Caspian steppe, a distinct geographic and archeological region extending from the Danube estuary to the Ural mountains to the east and North Caucasus to the south. The Neolithic, Eneolithic and early Bronze Age cultures in Pontic-Caspian steppe has been called the Kurgan culture (7000-2200 BCE) by Marija Gimbutas, due to the lasting practice of burying the deads under mounds ("kurgan") among the succession of cultures in that region. Horses were first domesticated around 4000 BCE in the steppe, perhaps somewhere around the Don or the lower Volga, and soon became a defining element of steppe culture. During the Bronze-age period, known as the Yamna horizon (3300-2500 BCE), the cattle and sheep herders adopted wagons to transport their food and tents, which allowed them to move deeper into the steppe, giving rise to a new mobile lifestyle that would eventually lead to the great Indo-European migrations.

The Pontic-Caspian steppe cultures can be divided in a western group, ranging from the Don River to the Dniester (and later Danube), and an eastern one, in the Volga-Ural region. The Pontic steppe was probably inhabited by men of mixed R1a and R1b lineages, with higher densities of R1b just north of the Caucasus, and more R1a in the the northern steppes and the forest-steppes.

R1b almost certainly crossed over from northern Anatolia to the Pontic-Caspian steppe. It is not clear whether this happened before, during or after the Neolithic. A regular flow of R1b across the Caucasus cannot be excluded either. The genetic diversity of R1b being greater around the Caucasus, it is hard to deny that R1b settled and evolved there before entering the steppe world. Does that mean that Indo-European languages originated in the steppes with R1a people, and that R1b immigrants blended into the established culture ? Or that Proto-Indo-European language appear in northern Anatolia or in the Caucasus, then spread to the steppes with R1b ? Or else did Proto-Indo-European first appear in the steppe as a hybrid language of Caucasian/Anatolian R1b and steppe R1a ? This question has no obvious answer, but based on the antiquity and archaic character of the Anatolian branch (Hittite, Palaic, Luwian, Lydian, and so on) an northern Anatolian origin of Proto-Indo-European is credible. Furthermore, there is documented evidence of loan words from Caucasian languages in Indo-European languages. This is much more likely to have happened if Proto-Indo-European developed near the Caucasus than in the distant steppes. R1b would consequently have been the spreading factor of PIE to the steppes, and from there to Europe, Central Asia and South Asia.

The Maykop culture, the R1b link to the steppe ?

The Maykop culture (3700-2500 BCE), in the North Caucasus, was culturally speaking a sort of southern extension of the Yamna horizon. Although not generally considered part of the Pontic-Caspian steppe culture due to its geography, the North Caucasus had close links with the steppe, as attested by numerous ceramics, gold, copper and bronze weapons and jewelry in the contemporaneous cultures of Mikhaylovka, Sredny Stog and Kemi Oba. The link between the North Pontic and North Caucasus is older than the Maykop period. Its predecessor, the Svobodnoe culture (4400-3700 BCE), already had links to the Suvorovo-Novodanilovka and early Sredny Stog cultures, and the even older Nalchik settlement (5000-4500 BCE) displayed a similar culture as Khvalynsk on the Volga. This may be the period when R1b started interacting and blending with the R1a population of the steppes.

The Yamna and Maykop people both used kurgan burials, with their deads in a supine position with raised knees and oriented in a north-east/south-west axis. Graves were sparkled with red ochre on the floor, and sacrificed domestic animal buried alongside humans. They also had in common horse riding, wagons, a cattle- and sheep-based economy, the use of copper/bronze battle-axes (both hammer-axes and sleeved axes) and tanged daggers. In fact, the oldest wagons and bronze artefacts are found in the North Caucasus, and spread from there to the steppes.

Maykop was an advanced Bronze Age culture, actually one of the very first to develop metalworking, and therefore metal weapons. The world's oldest sword was found at a late Maykop grave in Klady kurgan 31. Its style is reminiscent of the long Celtic swords, though less elaborated. Horse bones and depictions of horses already appear in early Maykop graves, suggesting that the Maykop culture might have been founded by steppe people or by people who had close link with them. However, the presence of cultural elements radically different from the steppe culture in some sites could mean that Maykop had a hybrid population. Without DNA testing it is impossible to say if these two populations were an Anatolian R1b group and a G2a Caucasian group, or whether R1a people had settled there too. The two or three ethnicities might even have cohabited side by side in different settlements. Typical Caucasian Y-DNA lineages (such as G2a) do not follow the pattern of Indo-European migrations, so intermarriages must have been limited, or at least restricted to Indo-European men taking Caucasian wives rather than the other way round.

Maykop people are the ones credited for the introduction of primitive wheeled vehicles (wagons) from Mesopotamia to the steppes. This would revolutionise the way of life in the steppe, and would later lead to the development of (horse-drawn) war chariots around 2000 BCE. Cavalry and chariots played a vital role in the subsequent Indo-European migrations, allowing them to move quickly and defeat easily anybody they encountered. Combined with advanced bronze weapons and their sea-based culture, the western branch (R1b) of the Indo-Europeans from the Black Sea shores are excellent candidates for being the mysterious Sea Peoples, who raided the eastern shores of the Mediterranean during the second millennium BCE.

The rise of the IE-speaking Hittites in Central Anatolia happened a few centuries after the disappearance of the Maykop culture. A back migration from the North Caucasus to northern Anatolia is very likely in this age of expansion. What is certain is that the Hittites used chariots, invented in the Volga-Ural steppes. R1a being found at low frequencies in Armenia and northern Anatolia, it is not unreasonable to imagine that a hybrid group of R1a-R1b from the Volga-Ural region migrated to this region sometime between 2000 BCE and 1650 BCE. The Maykop and Yamna cultures were succeeded by the Srubna culture (1600-1200 BCE), possibly representing an advance of R1a1a people from the northern and eastern steppes towards the Black Sea shores.

#### The European branch

The Indo-Europeans' bronze weapons and horses would have given them a tremendous advantage over the autochthonous inhabitants of Europe, namely the native haplogroup I (descendant of Cro-Magnon), and the early Neolithic herders and farmers (G2a, J2, E-V13 and T). This allowed R1a and R1b to replace (=> see How did R1b come to replace most of the older lineages in Western Europe ? most of the native male lineages, although female lineages seem to have been less affected.

A comparison with the Indo-Iranian invasion of South Asia shows that 40% of the male lineages of northern India are R1a, but less than 10% of the female lineages could be of Indo-European origin. The impact of the Indo-Europeans was more severe in Europe because European society 4,000 years ago was less developed in terms of agriculture, technology (no bronze weapons) and population density than that of the Indus Valley civilization. This is particularly true of the native Western European cultures where farming arrived much later than in the Balkans or central Europe. Greece, the Balkans and the Carpathians were the most advanced of European societies at the time and were the least affected in terms of haplogroup replacement. Native European Y-DNA haplogroups (I1, I2a, I2b) also survived better in regions that were more difficult to reach or less hospitable, like Scandinavia, Brittany, Sardinia or the Dinaric Alps.

The first forays of steppe people into the Balkans happened between 4200 BCE and 3900 BCE, when horse riders crossed the Dniester and Danube and apparently destroyed the towns of the Gumelnita, Varna and Karanovo VI cultures in Eastern Romania and Bulgaria. A climatic change resulting in colder winters during this exact period probably pushed steppe herders to seek milder pastures for their stock, while failed crops would have led to famine and internal disturbance within the Danubian and Balkanic communities. The ensuing Cernavoda culture (4000-3200 BCE) and Ezero culture (3300-2700 BCE) seems to have had a mixed population of steppe immigrants and people from the old tell settlements. These steppe immigrants were

likely a mixture of both R1a and R1b lineages. Many Danubian farmers would also have migrated to the Cucuteni-Tripolye towns in the Eastern Carpathians, causing a population boom and a north-eastward expansion until the Dnieper valley, bringing Y-haplogroups E-V13, J2b and T in what is now central Ukraine. This precocious Indo-European advance westward was fairly limited, due to the absence of Bronze weapons and organised army at the time, and was indeed only possible thanks to climatic catastrophes. The Carpathian, Danubian, and Balkanic cultures were too densely populated and technologically advanced to allow for a massive migration.

The Bronze Age announces a very different development. R1a people appear to have been the first to successfully penetrate into the heart of Europe, with the Corded Ware (Battle Axe) culture (3200-1800 BCE) as a natural western expansion of the Yamna culture. They went as far west as Germany and Scandinavia. DNA analysis from the Corded Ware culture site of Eulau confirms the presence of R1a (but not R1b) in central Germany around 2600 BCE. The Corded Ware migrants might well have expanded from the forest-steppe, or the northern fringe of the Yamna culture, where R1a lineages were prevalent over R1b ones.

R1b1b2 is thought to have arrived in central and western Europe around 2500 BCE, by going up the Danube from the Black Sea coast. The archeological and genetic evidence (distribution of R1b subclades) point at several consecutive waves towards the Danube between 2800 BCE and 2300 BCE (beginning of the Unetice culture). It is interesting to note that this also corresponds to the end of the Maykop culture (2500 BCE) and Kemi Oba culture (2200 BCE) on the northern shores of the Black Sea, and their replacement by cultures descended from the northern steppes. It can therefore be envisaged that the (mostly) R1b population from the northern half of the Black Sea migrated westward due to pressure from other Indo-European people (R1a) from the north, like the burgeoning Proto-Indo-Iranian branch, linked to the contemporary Poltavka and Abashevo cultures.

It is doubtful that the Beaker culture (2800-1900 BCE) was already Indo-European (although they were influenced by the Corded Ware culture), because they were the continuity of the native Megalithic cultures. It is more likely that the beakers and horses found across western Europe during that period were the result of trade with neighbouring Indo-European cultures, including the first wave of R1b into central Europe. Nevertheless, it is undeniable that the following Unetice (2300-1600 BCE), Tumulus (1600-1200 BCE), Urnfield (1300-1200 BCE) and Hallstatt (1200-750) cultures were linked to the spread of R1b to Europe, as they abruptly introduce new technologies and a radically different lifestyle.

=> See also Metal-mining and stockbreeding explain R1b dominance in Atlantic fringe

Did the Indo-Europeans really invade Western Europe ?

Proponents of the Paleolithic or Neolithic continuity model argue that bronze technology and horses could have been imported by Western Europeans from their Eastern European neighbours, and that no actual Indo-European invasion need be involved. It is harder to see how Italic, Celtic and Germanic languages were adopted by Western and Northern Europeans without at least a small scale invasion. It has been suggested that Indo-European (IE) languages simply spread through contact, just like technologies, or because it was the language of a small elite and therefore its adoption conferred a certain perceived prestige. However people don't just change language like that because it sounds nicer or more prestigious. Even nowadays, with textbooks, dictionaries, compulsory language courses at school, private language schools for adults and multilingual TV programs, the majority of the people cannot become fluent in a completely foreign language, belonging to a different language family. The linguistic gap between pre-IE vernaculars and IE languages was about as big as between modern English and Chinese. English, Greek, Russian and Hindi are all related IE languages and therefore easier to learn for IE speakers than non-IE languages like Chinese, Arabic or Hungarian. From a linguistic point of view, only a wide-scale migration of IE speakers could explain the thorough adoption of IE languages in Western Europe - leaving only Basque as a remnant of the Neolithic languages.

One important archeological argument in favour of the replacement of Neolithic cultures by Indo-European culture in the Bronze Age comes from pottery styles. The sudden appearance of bronze technology in Western Europe coincides with ceramics suddenly becoming more simple and less decorated, just like in the Pontic steppes. Until then, pottery had constantly evolved towards greater complexity and details for over 3,000 years. People do not just decide like that to revert to a more primitive style. Perhaps one isolated tribe might experiment with something simpler at one point, but what are the chances that distant cultures from Iberia, Gaul, Italy and Britain all decide to undertake such an improbable shift around the same time? The best explanation is that this new style was imposed by foreign invaders. In this case it is not mere speculation; there is ample evidence that this simpler pottery is characteristic of the steppes associated with the emergence of Proto-Indo-European speakers.

Besides pottery, archeology provides ample evidence that the early Bronze Age in Central and Western Europe coincides with a radical shift in food production. Agriculture experiences an abrupt reduction in exchange for an increased emphasis on domesticates. This is also a period when horses become more common and cow milk is being consumed regularly. The overall change mimics the steppe way of life almost perfectly. Even after the introduction of agriculture around 5200 BCE, the Bug-Dniester culture and later steppe cultures were characterized by an economy dominated by herding, with only limited farming. This pattern expands into Europe exactly at the same time as bronze working.

Religious beliefs and arts undergo a complete reversal in Bronze Age Europe. Neolithic societies in the Near East and Europe had always worshipped female figurines as a form of fertility cult. The steppe cultures, on the contrary, did not manufacture female figurines. As bronze technology spreads from the Danube valley to Western Europe, symbols of fertility and fecundity progressively disappear and are replaced by sculptures of domesticated animals.

Another clue that Indo-European steppe people came in great number to Central and Western Europe is to be found in burial practices. Neolithic Europeans either cremated their dead (e.g. Cucuteni-Tripolye culture) or buried them in collective graves (this was the case of Megalithic cultures). In the steppe, each person was buried individually, and high-ranking graves were placed in a funeral chamber and topped by a circular mound. The body was typically accompanied by weapons (maces, axes, daggers), horse bones, and a dismantled wagon (or later chariot). These characteristic burial mounds are known as kurgans in the Pontic steppe. Men were given more sumptuous tombs than women, even among children, and differences in hierarchy are obvious between burials. The Indo-Europeans had a strongly hierarchical and patrilineal society, as opposed to the more egalitarian and matrilineal cultures of Old Europe. The proliferation of status-conscious male-dominant kurgans (or tumulus) in Central Europe during the Bronze Age is a clear sign that the ruling elite had now become Indo-European. The practice also spread to Central Asia and Southern Siberia, two regions where R1a and R1b lineages are found nowadays, just like in Central Europe. The ceremony of burial is one of the most emotionally charged and personal aspects of a culture. It is highly doubtful that people would change their ancestral practice "just to do like the neighbours". In fact, different funerary practices have co-existed side by side during the European Neolithic and Chalcolithic. The ascendancy of yet another constituent of the Pontic steppe culture in the rest of Europe, and in this case one that does not change easily through contact with neighbours, adds up to the likelihood of a strong Indo-European migration. The adoption of some elements of a foreign culture tends to happen when one civilization overawes the adjacent cultures by its superiority. This process is called 'acculturation'. However there is nothing that indicates that the steppe culture was so culturally superior as to motivate a whole continent, even Atlantic cultures over 2000 km away from the Pontic steppes, to abandon so many fundamental symbols of their own ancestral culture, and even their own language. In fact, Old Europe was far more refined in its pottery and jewellery than the rough steppe people. The Indo-European superiority was cultural but military, thanks to horses, bronze weapons and an ethic code valuing individual heroic feats in war (these ethic values are known from the old IE texts, like the Rig Veda, Avesta, or the Mycenaean and Hittite literature).

After linguistics and archeology, the third category of evidence comes from genetics itself. It had first been hypothesized that R1b was native to Western Europe, because this is where it was most prevalent. It has since

been proven that R1b haplotypes displayed higher microsatellite diversity in Anatolia and in the Caucasus than in Europe. European subclades are also more recent than Middle Eastern or Central Asian ones. The main European subclade, R-P312/S116, only dates back to approximately 3500 to 3000 BCE. It does not mean that the oldest common ancestor of this lineage arrived in Western Europe during this period, but that the first person who carried the mutation R-P312/S116 lived at least 5,000 years ago, assumably somewhere in the lower Danube valley or around the Black Sea. In any case this timeframe is far too recent for a Paleolithic origin or a Neolithic arrival of R1b. The discovery of what was thought to be "European lineages" in Central Asia, Pakistan and India hit the final nail on the coffin of a Paleolithic origin of R1b in Western Europe, and confirmed the Indo-European link.

All the elements concur in favour of a large scale migration of horse-riding Indo-European speakers to Western Europe between 2500 to 2100 BCE, contributing to the replacement of the Neolithic or Chalcolithic lifestyle by a inherently new Bronze Age culture, with simpler pottery, less farming, more herding, new rituals (single graves) and new values (patrilinear society, warrior heroes) that did not evolve from local predecessors.

These Proto-Italo-Celto-Germanic R1b people had settled around the Alps by 2300 BCE, and judging from the spread of bronze working, reached Iberia by 2250 BCE, Britain by 2100 BCE and Ireland by 2000 BCE. This first wave of R1b assumably carried R1b-L21 lineages in great number, as these are found everywhere in western, northern and central Europe. A second R1b expansion took place from the Urnfield/Hallstatt culture around 1200 BCE, pushing west to the Atlantic, north to Scandinavia, and as far east as Greece and Anatolia (=> see Dorian invasion below).

The new Bronze Age culture flourished around the Alps (Unetice to early Hallstatt) thanks to the abundance of metal in the region, and laid the foundation for the classical Celtic culture. The Celtic Iron Age (late Hallstatt, from 800 BCE) may have been brought through preserved contacts with the the steppes and the North Caucasus, notably the Koban culture (1100-400 BCE).

The Alpine Celts of the Hallstatt culture are associated with the S28 (a.k.a. U152) mutation, although not exclusively. The Italic branch (also S28/U152) is thought to have entered Italy by 1200 BCE, but there were certainly several successive waves, as attested by the later arrival of the Cisalpine Celts. The Belgae were another S28/U152 branch, an extension of the La Tène culture northward, following the Rhine, Moselle and Meuse rivers.

One common linguistic trait between Italic and Gaulish/Brythonic Celtic languages linked to the Hallstatt expansion is that they shifted the original IE \*kw sound into \*p. They are known to linguists as the P-Celtic branch. It is thought that this change occurred due to the inability to pronounce the \*kw sound by the pre-Indo-European population of central Europe, Gaul and Italy, who were speakers of Afro-Asiatic dialects that had evolved from a Near-Eastern language. The Etruscans, although later incomers from the Levant, also fit in this category. It has recently been acknowledged that Celtic languages borrowed part of their grammar from Afro-Asiatic languages. This shift could have happened when the Proto-Italo-Celtic speakers moved from the steppes to the Danube basin and mixed with the population of Near-Eastern farmers belonging to haplogroups E-V13, T, G2a and J2b. However, such an early shift would not explain why Q-Celtic languages developed in Ireland and Iberia. It is more plausible that the shift happened after the Italo-Celts had first expanded across all western Europe. The S28/U152 connection to P-Celtic suggests that the shift took place around the Alps and Italy after 1200 BCE.

R1b-S21 (a.k.a. U106) is found at high concentrations in the Netherlands and northern Germany. Its presence in other parts of Europe can be attributed to the 5th- and 6th-century Germanic migrations. The Frisians and Saxons spread this haplogroup to the British Isles, the Franks to Belgium and France, and the Lombards to Austria and northern Italy. The high concentration of S21/U106 around Austria hints that it could have originated there in the Hallstatt period, or originated around the Black Sea and moved there during the

Hallstatt period. In fact, southern Germany and Austria taken together have the highest diversity of R1b in Europe. Besides S21, the three major first level subclades of R1b1b2a1b (L21, S28, M167) are found in this area at reasonable frequencies to envisage a spread from the Unetice to Hallstatt homeland to the rest of western Europe.

=> Trivia : Kings of many European countries have been confirmed to be R1b through genetic genealogy.

How did R1b come to replace most of the older lineages in Western Europe ?

Until recently it was believed that R1b originated in Western Europe due to its strong presence in the region today. The theory was that R1b represented the Paleolithic Europeans (Cro-Magnon) that had sought refuge in the Franco-Cantabrian region at the peak of the last Ice Age, then recolonised Central and Northern Europe once the ice sheet receded. The phylogeny of R1b proved that this scenario was not possible, because older R1b clades were consistently found in Central Asia and the Middle East, and the youngest in Western and Northern Europe. There was a clear gradient from East to West tracing the migration of R1b people (see map above). This age of the main migration from the shores of the Black Sea to Central Europe also happened to match the timeframe of the Indo-European invasion of Europe, which coincides with the introduction of the Bronze-Age culture in Western Europe, and the spread of Italo-Celtic and Germanic languages.

Historians and archeologists have long argued whether the Indo-European migration was a massive invasion, or rather a cultural diffusion of language and technology spread only by a small number of incomers. The answer could well be "neither". Proponents of the diffusion theory would have us think that R1b is native to Western Europe, and R1a alone represent the Indo-Europeans. The problem is that haplogroup R did arise in Central Asia, and R2 is still restricted to Central and South Asia, while R1a and the older subclades of R1b are also found in Central Asia. The age of R1b subclades in Europe coincide with the Bronze-Age. R1b must consequently have replaced most of the native Y-DNA lineages in Europe from the Bronze-Age onwards.

However, a massive migration and nearly complete annihilation of the Paleolithic population can hardly be envisaged. Western Europeans do look quite different in Ireland, Holland, Aquitaine or Portugal, despite being all regions where R1b is dominant. Autosomal DNA studies have confirmed that the Western European population is far from homogeneous. A lot of maternal lineages (mtDNA) also appear to be of Paleolithic origin (e.g. H1, H3, U5 or V) based on ancient DNA tests. What a lot of people forget is that there is also no need of a large-scale exodus for patrilineal lineages to be replaced fairly quickly. Here is why.

**Polygamy.** Unlike women, men are not limited in the number of children they can procreate. Men with power typically have more children. This was all the truer in primitive societies, where polygamy was often the norm for chieftains and kings.

**Status & Power.** Equipped with Bronze weapons and horses, the Indo-Europeans would have easily subjugated the Neolithic farmers and with even greater ease Europe's last hunter-gatherers. If they did not exterminate the indigenous men, the newcomers would have become the new ruling class, with a multitude of local kings, chieftains and noblemen (Bronze-Age Celts and Germans lived in small village communities with a chief, each part of a small tribe headed by a king) with higher reproductive opportunities than average.

**Gender imbalance.** Invading armies normally have far more men than women. Men must therefore find women in the conquered population. Wars are waged by men, and the losers suffer heavier casualties, leaving more women available to the winners.

**Aggressive warfare.** The Indo-Europeans were a warlike people with a strong heroic code emphasising courage and military prowess. Their superior technology (metal weapons, wheeled vehicles and warhorses) and attitude to life would have allowed them to slaughter any population that did not have organised armies with metal weapons (i.e. anybody except the Middle-Eastern civilizations).

**Genetic predisposition to conceive boys.** The main role of the Y-chromosome in man's body is to create sperm. Haplogroups are determined based on mutations differentiating Y-chromosomes. Each mutation is liable to affect sperm production and sperm motility. Preliminary research has already established a link between certain haplogroups and increased or reduced sperm motility. The higher the motility, the higher the

chances of conceiving a boy. It is absolutely possible that R1b could confer a bias toward more male offspring. Even a slightly higher percentage of male births would significantly contribute to the replacement of other lineages with the accumulation effect building up over a few millennia. Not all R1b subclades might have this boy bias. The bias only exist in relation to other haplogroups found in a same population. It is very possible that the fairly recent R1b subclades of Western Europe had a significant advantage compared to the older haplogroups in that region, notably haplogroup I2 and E-V13. Read more

Replacement of patrilineal lineages following this model quickly becomes exponential. Imagine 100 Indo-European men conquering a tribe of 1000 indigenous Europeans (a ratio of 1:10). War casualties have resulted in a higher proportion of women in the conquered population. Let's say that the surviving population is composed of 700 women and 300 men. Let's suppose that the victorious Indo-European men end up having twice as many children reaching adulthood as the men of the vanquished tribe. There is a number of reason for that. The winners would take more wives, or take concubines, or even rape women of the vanquished tribe. Their higher status would guarantee them greater wealth and therefore better nutrition for their offspring, increasing the chances of reaching adulthood and procreating themselves. An offspring ratio of 2 to 1 for men is actually a conservative estimate, as it is totally conceivable that Bronze-Age sensibilities would have resulted in killing most of the men on the losing side, and raping their women (as attested by the Old Testament). Even so, it would only take a few generations for the winning Y-DNA lineages to become the majority. For instance, if the first generation of Indo-Europeans had two surviving sons per man, against only one per indigenous man, the number of Indo-European paternal lineages would pass to 200 individuals at the second generation, 400 at the third, 800 at the fourth and 1600 at the fifth, and so on. During that time indigenous lineages would only stagnate at 300 individuals for each generation.

Based on such a scenario, the R1b lineages would have quickly overwhelmed the local lineages. Even if the Indo-European conquerors had only slightly more children than the local men, R1b lineages would become dominant within a few centuries. Celtic culture lasted for over 1000 years in Continental Europe before the Roman conquest putting an end to the privileges of the chieftains and nobility. This is more than enough time for R1b lineages to reach 50 to 80% of the population.

The present-day R1b frequency forms a gradient from the Atlantic fringe of Europe (highest percentage) to Central and Eastern Europe (lowest), the rises again in the Anatolian homeland. This is almost certainly because agriculture was better established in Eastern, then Central Europe, with higher densities of population, leaving R1b invaders more outnumbered than in the West. Besides, other Indo-Europeans of the Corded Ware culture (R1a) had already advanced from modern Russia and Ukraine as far west as Germany and Scandinavia. It would be difficult for R1b people to rival with their R1a cousins who shared similar technology and culture. The Pre-Celto-Germanic R1b would therefore have been forced to settled further west, first around the Alps, then overtaking the then sparsely populated Western Europe.

### The Greco-Anatolian branch

The Hittites (2000-1200 BCE) were the first Indo-Europeans to defy (and defeat) the mighty Mesopotamian and Egyptian empires. The Hittite ruling class was plausibly an offshoot of the late Maykop culture that conquered the Hattian kingdom. The northern Anatolians may also have been the original Indo-European speakers people who later founded the Maykop culture and spread their language and culture to the Pontic-Caspian steppes. Whichever way, northern Anatolian Bronze-Age Indo-European speakers would surely have belonged in great part to haplogroup R1b1b (and subclades). The Hattians might have had some older Middle-Eastern R1b mixed with the other haplogroups common in Anatolia nowadays (E-M78, G2a and J2).

Troy could well have been a Indo-European colony securing the trade routes between the Black Sea and the Aegean. The Trojans were Luwian speakers related to the Hittites (hence Indo-European), with proven cultural ties to the culture of the Pontic-Caspian steppe. The first city of Troy dates back to 3000 BCE, right in the middle of the Maykop period, and exactly at the time the first galleys were made. Considering the early

foundation of Troy, the most likely of the two Indo-European paternal haplogroups would be R1b1b, not R1a1a.

Greek R1b comes in many varieties: R1b1 from the Near-East, R1b1b from Anatolia, and the European R1b1b2, including the Proto-Celtic S116/P312 and Hallstatt Celtic S28/U152. The presence of R1b1b2 in Greece could be attributed to the Dorian invasion, thought to have happened in the 12th century BCE. The Dorians could have been related to the Trojans and the Hittites belonging to the oldest Indo-European linguistic branch, or to the Proto-Celts of central Europe and the Danube valley. One way of the other, their Y-DNA lineages would have been predominantly R1b1b or R1b1b2. The Dorians could be the descendants of the first (R1b) steppe nomads who settled in the Eastern Balkans (Cernavoda and Ezero cultures) and did not continue their migration up the Danube to central and western Europe.

Greek and Anatolian R1b-S28 lineages could be attributed to the Celtic invasions of the 3rd century BCE, but more probably to the Roman occupation. Older clades of R1b, such as R1b1 or R1b1a are only a small minority and would have come along E1b1b and J2 from the Middle East. The Mycenaean could have brought some R1b1b2 to Greece, but their origins can be traced back to the Seima-Turbino culture of the northern forest-steppe, which would make them primarily an R1a1a tribe.

### The Central Asian branch

An early group of R1b1b people is thought to have migrated from Caspian Sea region to Central Asia, where it evolved into the R1b1b1 (M73) branch. This variety of R1b occurs almost exclusively in very specific Central Asian populations. The highest percentages were observed among the Uyghurs (20%) of Xinjiang in north-west China, the Hazara people of Afghanistan (32%), and the Bashkirs (55%) of the Abzelilovsky district of Bashkortostan in Russia (border of Kazakhstan).

Central Asian R1b1b1 could correspond to the Tocharian branch of the Indo-Europeans. It is possible that the Tocharians split from the main R1b body as early as 7,000 BCE. Over the centuries some groups of these nomadic tribes ended up around the southern Urals, others in the Tarim Basin (Xinjiang) or in southern Central Asia. Another theory is that a group of early horse riders from the Repin culture (3700-3300 BCE) migrated from the Don-Volga region to the Altai mountain, founding the Afanasevo culture (c. 3600-2400 BCE), then moved south to the Tarim Basin.

Mummies of fair-haired Caucasian people were found in the Tarim Basin, the oldest of which date back to 1800 BCE. The modern inhabitants of the Tarim Basin, the Uyghurs, belong both to this R1b-M73 subclade (about 20%) and to R1a1 (about 30%). This could mean that they had become a hybrid R1b-R1a society by the time they reached the Tarim Basin. But R1a1 could also have arrived independently during the later Indo-Iranian migrations (approx. 2000 BCE), or much later through some nomadic Scytho-Iranian tribes (after 700 BCE).

### Haplogroup R1a (Y-DNA)

Main article : Haplogroup R1a

Distribution of haplogroup R1a in Europe

R1a is thought to have been the dominant haplogroup among the northern and eastern Indo-European speakers who evolved into the Indo-Iranian, Mycenaean Greek, Thracian, Baltic and Slavic branches. The Proto-Indo-Europeans originated in the Yamna culture (3300-2500 BCE), in the Pontic-Caspian steppe between modern Ukraine and south-west Russia. Their expansion is linked to the domestication of horses in the Eurasian steppes, and the invention of the chariot (see R1b above).

The eastern part of the Pontic-Caspian steppes is strongly associated with the Indo-Iranian and Balto-Slavic branches of Indo-European languages. Based on archeological, linguistic and genetic data, it is possible to say

that the pastoralist nomads who lived in the northern Russian steppes and forest-steppes 5,000 years ago carried predominantly R1a paternal lineages.

Nowadays, high frequencies of R1a are found in Poland (56% of the population), Ukraine (50 to 65%), European Russia (45 to 65%), Belarus (45%), Slovakia (40%), Latvia (40%), Lithuania (38%), the Czech Republic (34%), Hungary (32%), Croatia (29%), Norway (28%), Austria (26%), Sweden (24%), north-east Germany (23%) and Romania (22%).

### The Germanic branch

The first expansion of R1a took place with the westward propagation of the Corded Ware (or Battle Axe) culture (3200-1800 BCE) from the Yamna homeland. This was the first wave of R1a into Europe, one that is responsible for the presence of this haplogroup in Scandinavia, Germany, and a portion of the R1a in the Czech Republic, Slovakia, Hungary or Poland. The high prevalence of R1a in Balto-Slavic countries nowadays is not only due to the Corded Ware expansion, but also to a long succession of later migrations from Russia, the last of which took place from the 5th to the 1th century CE.

The Germanic branch of Indo-European languages probably evolved from a merger of Corded-Ware R1a (Proto-Slavic language) and the later arrival of Italo-Celtic R1b from Central Europe. This is supported by the fact that Germanic people are hybrid R1a-R1b, that these two haplogroups came via separate routes at different times, and also on the linguistics of Proto-Germanic language, which shares similarities with Italic, Celtic and Slavic languages. The Corded Ware R1a people would have mixed with the pre-Germanic I1 aborigines to create the Nordic Bronze Age (1800-500 BCE). R1b presumably reached Scandinavia later as a northward migration from the contemporary Hallstatt culture (1200-500 BCE). The first genuine Germanic tongue has been estimated by linguists to have come into existence around (or after) 500 BCE. This would confirm that it emerged as a blend of Hallstatt Proto-Celtic and the Corded-Ware Proto-Slavic. The uniqueness of some of the Germanic vocabulary points at borrowing from native pre-Indo-European languages. Celtic language itself is known to have borrowed from Afro-Asiatic languages spoken by Near-Eastern immigrants to Central Europe. The fact that present-day Scandinavia is composed of roughly 40% of I1, 20% of R1a and 40% of R1b reinforces the idea that Germanic ethnicity and language had acquired a tri-hybrid character by the Iron Age.

### The Baltic branch

The Baltic branch is thought to have evolved from the Fatyanovo culture (3200-2300 BCE), the northeastern extension of the Corded Ware culture. Early Bronze Age R1a nomads from the northern steppes and forest-steppes would have mixed with the indigenous Uralic-speaking inhabitants (N1c1 lineages) of the region. This is supported by a strong presence of both R1a and N1c1 haplogroups from southern Finland to Lithuania and the adjacent part of Russia.

### The Slavic branch

The origins of the Slavs goes back to circa 3000 BCE. The Slavic branch differentiated itself when the Corded Ware culture (see Germanic branch above) absorbed the Cucuteni-Tripolye culture (5200-2600 BCE) of western Ukraine and north-eastern Romania, which appears to have been composed primarily of I2a2 lineages descended directly from Paleolithic Europeans, with a small admixture of Near-Eastern immigrants (notably E-V13 and T). Thus emerged the hybrid Globular Amphora culture (3400-2800 BCE) in what is now Ukraine, Belarus and Poland. It is surely during this period that I2a2, E-V13 and T spread (along with R1a) around Poland, Belarus and western Russia, explaining why eastern and northern Slavs (and Lithuanians) have a considerable incidence of haplogroups I2a2 with a bit of E and T. After just a few centuries, this hybridised culture faded away into the dominant Corded Ware culture.

The Corded Ware period was followed by the Trzciniec (1700-1200 BCE), Lusatian (1300-500 BCE), Chernoles (1025-700 BCE) and Milograd (600 BCE-100 CE) cultures in north-east Slavic countries. The last

important Slavic migration is thought to have happened in the 6th century CE, from Ukraine to Poland, the Czech Republic and Slovakia, filling the vacuum left by eastern Germanic tribes who invaded the Roman Empire.

Historically, no other part of Europe was invaded a higher number of times by steppe peoples than the Balkans. Chronologically, the first R1a invaders came with the westward expansion of the Corded Ware culture (from about 3200 BCE), then the Mycenaean invasion (1600 BCE), followed by the Thracians (1500 BCE), the Illyrians (around 1200 BCE), the Huns and the Alans (400 CE), the Avars, the Bulgars and the Serbs (all around 600 CE), and the Magyars (900 CE), among others. These peoples originated from different parts of the Eurasian steppes, anywhere between Eastern Europe and Central Asia, which is why such high STR diversity is found within Balkanic R1a nowadays. It is not yet possible to determine the ethnic origin for each variety of R1a, apart from the fact that about any R1a is associated with tribes from Eurasian steppe at one point in history.

## The Indo-Iranian branch

Proto-Indo-Iranian speakers, the people who later called themselves 'Aryans' in the Rig Veda and the Avesta, originated in the Sintashta-Petrovka culture (2100-1750 BCE), in the Tobol and Ishim valleys, east of the Ural Mountains. It was founded by pastoralist nomads from the Abashevo culture (2500-1900 BCE), ranging from the upper Don-Volga to the Ural Mountains, and the Poltavka culture (2700-2100 BCE), extending from the lower Don-Volga to the Caspian depression. The Sintashta-Petrovka culture was the first Bronze Age advance of the Indo-Europeans west of the Urals, opening the way to the vast plains and deserts of Central Asia to the metal-rich Altai mountains. The Aryans quickly expanded over all Central Asia, from the shores of the Caspian to southern Siberia and the Tian Shan, through trading, seasonal herd migrations, and looting raids.

Horse-drawn war chariots seem to have been invented by Sintashta people around 2100 BCE, and quickly spread to the mining region of Bactria-Margiana (modern border of Turkmenistan, Uzbekistan, Tajikistan and Afghanistan). Copper had been extracted intensively in the Urals, and the Proto-Indo-Iranians from Sintashta-Petrovka were exporting it in huge quantities to the Middle East. They appear to have been attracted by the natural resources of the Zeravshan valley for a Petrovka copper-mining colony was established in Tugai around 1900 BCE, and tin was extracted soon afterwards at Karnab and Mushiston. Tin was an especially valued resource in the late Bronze Age, when weapons were made of copper-tin alloy, stronger than the more primitive arsenical bronze. In the 1700's BCE, the Indo-Iranians expanded to the lower Amu Darya valley and settled in irrigation farming communities (Tazabagyab culture). By 1600 BCE, the old fortified towns of Margiana-Bactria were abandoned, submerged by the northern steppe migrants. The group of Central Asian cultures under Indo-Iranian influence is known as the Andronovo horizon, and lasted until 800 BCE.

The Indo-Iranian migrations progressed further south across the Hindu Kush. By 1700 BCE, horse-riding pastoralists had penetrated into Balochistan (south-west Pakistan). The Indus valley succumbed circa 1500 BCE, and the northern and central parts of the Indian subcontinent were taken over by 500 BCE. Westward migrations led Old Indic Sanskrit speakers riding war chariots to Assyria, where they became the Mitanni rulers from circa 1500 BCE. The Medes, Parthians and Persians, all Iranian speakers from the Andronovo culture, moved into the Iranian plateau from 800 BCE. Those that stayed in Central Asia are remembered by history as the Scythians, while the Yamna descendants who remained in the Pontic-Caspian steppe became known as the Sarmatians to the ancient Greeks and Romans.

The Indo-Iranian migrations have resulted in high R1a frequencies in southern Central Asia, Iran and the Indian subcontinent. The highest frequency of R1a (about 65%) is reached in a cluster around Kyrgyzstan, Tajikistan and northern Afghanistan. In India and Pakistan, R1a ranges from 15 to 50% of the population, depending on the region, ethnic group and caste. R1a is generally stronger in the North-West of the subcontinent, and weakest in the Dravidian-speaking South (Tamil Nadu, Kerala, Karnataka, Andhra Pradesh) and from Bengal eastward. Over 70% of the Brahmins (highest caste in Hinduism) belong to R1a1, due to a founder effect.

Maternal lineages in South Asia are, however, overwhelmingly pre-Indo-European. For instance, India has over 75% of "native" mtDNA M and R lineages and 10% of East Asian lineages. In the residual 15% of haplogroups, approximately half are of Middle Eastern origin. Only about 7 or 8% could be of "Russian" (Pontic-Caspian steppe) origin, mostly in the form of haplogroup U2 and W (although the origin of U2 is still debated). European mtDNA lineages are much more common in Central Asia though, and even in Afghanistan and northern Pakistan. This suggests that the Indo-European invasion of India was conducted mostly by men through war, and the first major settlement of women was in northern Pakistan, western India (Punjab to Gujarat) and northern India (Uttar Pradesh), where haplogroups U2 and W are the most common.

#### Turkic speakers and R1a

The present-day inhabitants of Central Asia, from Xinjiang to Turkey and from the Volga to the Hindu Kush, speak in overwhelming majority Turkic languages. This may be surprising as this corresponds to the region where the Indo-Iranian branch of Indo-European speakers expanded, the Bronze-Age Andronovo culture, and the Iron-Age Scythian territory. So why is it that Indo-European languages only survives in Slavic Russia or in the southern part of Central Asia, in places like Tajikistan, Afghanistan or some parts of Turkmenistan ? Why don't the Uyghurs, Uzbeks, Kazakhs and Kyrgyzs, or the modern Pontic-Caspian steppe people (Crimean Tatars, Nogais, Bashkirs and Chuvashes) speak Indo-European vernaculars ? Genetically these people do carry Indo-European R1a, and to a lesser extent also R1b, lineages. The explanation is that Turkic languages replaced the Iranian tongues of Central Asia between the 4th and 11th century CE.

Proto-Turkic originated in Mongolia and southern Siberia with such nomadic tribes as the Xiongnu. It belongs to the Altaic linguistic family, like Mongolian and Manchu (some also include Korean and Japanese, although they share very little vocabulary in common). It is unknown when Proto-Turkic first emerged, but its spread started with the Hunnic migrations westward through the Eurasian steppe and all the way to Europe, only stopped by the boundaries of the Roman Empire.

The Huns were the descendants of the Xiongnu. Ancient DNA tests have revealed that the Xiongnu were already a hybrid Eurasian people 2,000 years ago, with mixed European and North-East Asian Y-DNA and mtDNA. Modern inhabitants of the Xiongnu homeland have approximately 90% of Mongolian lineages against 10% of European ones. The oldest identified presence of European mtDNA around Mongolia and Lake Baikal dates back to over 6,000 years ago.

It appears that Turkic quickly replaced the Scythian and other Iranian dialects all over Central Asia. Other migratory waves brought more Turkic speakers to Eastern and Central Europe, like the Khazars, the Avars, the Bulgars and the Turks (=> see 5000 years of migrations from the Eurasian steppes to Europe). All of them were in fact Central Asian nomads who had adopted Turkic language, but had little if any Mongolian blood. Turkic invasions therefore contributed more to the diffusion of Indo-European lineages (especially R1a1) than East Asian ones.

Turkic languages have not survived in Europe outside the Pontic-Caspian steppe. Bulgarian language, despite being named after a Turkic tribe, is actually a Slavic tongue with a mild Turkic influence. Hungarian, sometimes mistaken for the heir of Hunnic because of its name, is in reality an Uralic language (Magyar). The dozens of Turkic languages spoken in the world today have a high degree of mutual intelligibility due to their fairly recent common origin and the nomadic nature of its speakers (until recently). Its two main branches Oghuz and Oghur could be seen as two languages about as distant as Spanish and Italian, and languages within each branch like regional dialects of Spanish and Italian.

#### The Greek branch

Little is known about the arrival of Proto-Greek speakers from the steppes. The Mycenaean culture commenced circa 1650 BCE and is clearly an imported steppe culture. The close relationship between Mycenaean and Proto-Indo-Iranian languages suggest that they split fairly late, some time between 2500 and 2000 BCE. Archeologically, Mycenaean chariots, spearheads, daggers and other bronze objects show striking similarities with the Seima-Turbino culture (c. 1900-1600 BCE) of the northern Russian forest-steppes, known for the great mobility of its nomadic warriors (Seima-Turbino sites were found as far away as Mongolia). It is therefore likely that the Mycenaean descended from Russia to Greece between 1900 and 1650 BCE, where they intermingled with the locals to create a new unique Greek culture.

### Haplogroup I (Y-DNA)

I is the oldest haplogroup in Europe and in all probability the only one that originated there (apart from deep subclades of other haplogroups). It is thought to have arrived from the Middle East as haplogroup IJ around 35,000 years ago, and developed into haplogroup I approximately 25,000 years ago. This means that Cro-Magnons most probably belonged (exclusively ?) to IJ or I. Nowadays haplogroup I accounts for 10 to 45% of the population in most of Europe. It is divided in four main subclades.

The earliest megalithic structures (5000-1200 BCE) of Europe were built by men belonging to haplogroup I, then were joined by Neolithic newcomers such as G2a and E1b1b.

### Haplogroup I1

Main article : Haplogroup I1

Haplogroup I1 (formerly I1a) is the most common I subclade. It is found mostly in Scandinavia and Northern Germany, where it can represent over 35% of the population. Associated with the Norse ethnicity, it is found in all places invaded by the ancient Germanic tribes and the Vikings.

During the Neolithic period, pre-I1 and I1 people were part of the successive Ertebølle culture (5300-3950 BCE) and Funnelbeaker culture (4000-2700 BCE). The Corded Ware period (3200-1800 BCE) marks the arrival of the Indo-European R1a people from the Ukrainian steppes.

I1 is identified by at least 15 unique mutations, which indicates that this lineage has been isolated for a long period of time, or experienced a serious population bottleneck. Although the first mutation splitting I1 away from I2 may have arisen as long as 20,000 years ago, people belonging to this haplogroup all descend from a single man who lived between 10,000 and 5,000 years ago. This corresponds to the arrival of the Indo-European, suggesting that a high percentage of the indigenous I1 men could possibly have been killed by the new immigrants.

Distribution of haplogroup I1 in Europe

### Haplogroup I2

Haplogroup I2 might have originated in southeastern Europe some 17,000 years ago and developed into four main subgroups : I2a1, I2a2, I2b1 and I2b2.

I2a1 (formerly I1b2) is found chiefly among the Sardinians and the Basques, and is rarely found outside Iberia, Western France, the West coast of Italy and the Mediterranean coast of the Maghreb. It accounts for approximately 40% of all Y-DNA haplogroups among the Sardinians. I2a1 is estimated to be 8,000 years old.

I2a2 (formerly I1b) is typical of the Dinaric Slavs (Croats, Serbs and Bosniaks). Its highest density is observed around ex-Yugoslavia and Moldova, but it is also common to a lower extent in Albania, Northern Greece, Bulgaria, Romania, Ukraine, Belarus, and southwestern Russia. The high concentration of I2a2 in north-east Romania, Moldova and central Ukraine reminds of the maximum spread of the Cucuteni-Tripolye

culture before it was swallowed by the Indo-European Corded Ware culture. This could mean that the Cucuteni-Tripolye culture was a native European group of hunter-gatherers who adopted farming after coming in contact (with perhaps some intermarriages) with the Levantine farmers who settled in the Balkans (haplogroups E-V13, J2b and T).

The modern territory of I2a1 and I2a2 (Illyria, Italy, Sardinia, Mediterranean coast of France and Spain) matches the extent of the Neolithic Printed-Cardium Pottery culture (5000-1500 BCE), that is believed to have started with the arrival of E-V13 and G2a farmers and herders from Thessaly (northern Greece). It was followed by the Terramare culture (1500-1000 BCE) in the Bronze Age. The R1b Celto-Italic people are thought to have crossed the Alps and invaded the Italian peninsula around 1,000 BCE, replacing most of the indigenous I2a, G2a and E-V13 people (especially in the northern half).

#### Distribution of haplogroup I2a in Europe

I2b (formerly I1c) is associated with the pre Celto-Germanic people of North-Western Europe, such as the megaliths builders (5000-1200 BCE). The wide variety of STR markers within I2b could make it as much as 13,000 years old.

I2b is found in all Western Europe, but apparently survived better the Indo-European invasions (=> see R1b above) in northern Germany, and was reintroduced by the Germanic invasions during the late Roman period. Nowadays, I2b peaks in central and northern Germany (10-20%), the Benelux (10-15%) as well as in northern Sweden. It is also found in 3 to 10% of the inhabitants of Denmark, East England, and Northern France. It is rare in Norway, which concords with the fact that it hasn't been invaded by people from northern Germany.

There are two major subclades : I2b1 (M223+) and I2b2 (L38/S154+), further subdivided in at least 4 subclades each, although little is known about them yet. The subclade I2b1a (M284+) occurs almost exclusively in Britain, where it seemingly developed about 3,000 years ago.

#### Distribution of haplogroup I2b in Europe

#### Near-Eastern haplogroups

##### Haplogroup G (Y-DNA)

##### Main article : Haplogroup G2a

Haplogroup G is believed to have originated around the Middle East during the late Paleolithic, possibly as early as 30,000 years ago. At that time humans would all have been hunter-gatherers, and in most cases living in small nomadic or semi-nomadic tribes. Members of this haplogroup appear to have been closely linked to the development of early agriculture in the Levant part of the Fertile Crescent, starting 11,500 years before present. There has so far been ancient Y-DNA analysis from only four Neolithic cultures (LBK in Germany, Remedello in Italy and Cardium Pottery in Southwest France and Spain), and all sites yielded G2a individuals, which is the strongest evidence at present that farming originated with and was spread by members of haplogroup G.

G2a makes up 5 to 10% of the population of Mediterranean Europe, but is fairly rare in Northern Europe. The only places where haplogroup G2 exceeds 10% of the population in Europe are Cantabria, Switzerland, the Tyrol, south-central Italy (Molise, Central and Southern Apennine), Sardinia, northern Greece (Thessaly) and Crete - all mountainous and relatively isolated regions.

There are several theories regarding the origin of G2a in Europe. There are doubtlessly cumulative rather than exclusive.

#### Neolithic mountain herders

Chronologically, the first hypothesis is the advance of Neolithic farmers and herders from Anatolia to Europe between 9,000 and 6,000 years ago. In this scenario the Caucasian migrants would have brought with them sheep and goats, which were domesticated south of the Caucasus about 12,000 years ago. This would explain why haplogroup G is more common in mountainous areas, be it in Europe or in Asia.

The geographic continuity of G2a from Anatolia to Thessaly to the Italian peninsula, Sardinia, south-central France and Iberia already suggested that G2a could be connected to the Printed-Cardium Pottery culture (5000-1500 BCE). Ancient DNA tests conducted on skeletons from a LBK site in Germany as well as a Printed-Cardium Pottery site in southern France (Languedoc-Roussillon) confirmed that Neolithic farmers in Europe belonged primarily (even exclusively if assimilated local population are omitted) to haplogroup G2a.

Nowadays G2a is found mostly in mountainous regions of Europe, for example, Cantabria (over 10%) in northern Spain, Switzerland (10%), Austria (8%), Auvergne (8%) in central France, the mountainous parts of Bohemia (5 to 10%), and Wales (4%). It may be because Caucasian farmers sought hilly terrain similar to their original homeland, perhaps well suited to the raising of goats. But it is more likely that G2a farmers escaped from Bronze-Age invaders, such as the Indo-Europeans and found shelter into the mountains.

G2a3b1a, the Indo-European branch of G2a

The presumed homeland of R1b1b and Pre-Proto-Indo-European speakers is assumed to be in northern Anatolia and/or the North Caucasus. The Caucasus itself is a hotspot of haplogroup G. Therefore, it is entirely conceivable that a minority of Caucasian men belonging to haplogroup G (and perhaps also J2b) integrated the R1b community that crossed the Caucasus and established themselves on the northern and eastern shores of the Black Sea sometime between 7,000 and 5,000 BCE. Those Proto-Indo-European would have belonged evolved to R1b1b2a1 and G2a3b1a before their epic conquest of Europe starting timidly in the Balkans around 4000 BCE and completed when all the Atlantic fringe from Iberia to the British Isles was settled, around 2000 BCE. Contrarily to G2a\* and G2a3\*, which is more prevalent in mountainous areas, G2a3b1a is found uniformly throughout Europe, even in Scandinavia and Russia. More importantly, G2a3b1 is also found in India, especially among the upper castes. The combined presence of G2a3b1 across Europe and India is a very strong argument in favour of an Indo-European origin. The coalescence age of G2a3b1 also matches the time of the Indo-European expansion during the Bronze Age.

Distribution of haplogroup G in Europe, North Africa and the Middle East

Expansion of agriculture from the Middle East to Europe (9500-3800 BCE)

Roman redistribution

It is most likely that G2a arrived in Europe during the Neolithic or the Bronze Age and that the Romans helped spread it around, the whole of Italy being relatively rich in G2a. Migrations within the Roman Empire probably contributed to a moderate increase of G2a northward to Gaul and Britain. Indeed, the frequency of haplogroup G decreases with the distance from the boundaries of the Roman Empire. Haplogroup G is extremely rare in Nordic and Baltic countries nowadays, despite the fact that agriculture reached those regions around the same time as Britain or Ireland. This may just be a coincidence, because the forested lowlands of northern Germany, Poland and northern Europe happen to be poor in metals and would not have attracted Bronze-Age workers from the Caucasus. North-East Europe also has a fairly modest frequency of R1b, which further reinforces the correlation between the two haplogroups.

Alanic G2a1

The only ethnic group that has a majority of haplogroup G nowadays are the Ossetians in the Caucasus, in the modern Russian Republic of North Ossetia-Alania. They are thought to descend directly from the Alans, a Central Asian tribe related to the ancient Samartians. The medieval Kingdom of Alania was located in the northern Caucasus, in present-day Georgia and Ossetia.

G2a has been observed at a slightly higher frequencies in Picardy and Flanders than in surrounding regions. It has been hypothesised that G2a was brought to northern France and Belgium by the Alans, who traversed all continental Europe during the barbarian invasions in the 5th century and founded a short-lived kingdom in northern France.

Nonetheless, if there is Alanic G in Europe it must certainly belong to other subclades than those from the Neolithic or Bronze Age period (namely G2a3). G2a1 being the most common variety in the Caucasus nowadays, the fairly recent Alanic migration (from a genetic point of view) could have carried that particular subclade. In fact, G2a1 has been found all along the Alanic migration route (Hungary, France, Spain), as well as in Britain (Samartian element ?), but hardly anywhere else.

### Scythian G1

Romans were known to recruit Scythian or Sarmatian horsemen in their legions. According to C. Scott Littleton in his book *From Scythia to Camelot*, several Knights of the Round Table were of Scythian origin, and the the legend of Holy Grail itself originated in ancient Scythia. This hypothesis was also taken up in the 2004 movie *King Arthur*, which opens with the arrival of Scytho-Roman cavalry in Britain. However, Scythians were steppe people more likely to belong to haplogroup R1a. If any of them did belong to G, they presumably were G1, not G2a. This would explain the scattered cases of G1 in north-western Europe though. G2a2, which also been found in Britain and Anatolia, is another potential candidate.

### Haplogroup J (Y-DNA)

J is a Middle Eastern haplogroup, divided into the northern J2 and the southern J1. J2 is by far the most common variety in Europe.

### Haplogroup J2

#### Main article : Haplogroup J2

J2 originated in northern Mesopotamia, and spread westward to Anatolia and southern Europe, and eastward to Persia and India. J2 is related to the Ancient Etruscans, (Minoan) Greeks, southern Anatolians, Phoenicians, Assyrians and Babylonians.

In Europe, J2 reaches its highest frequency in Greece (especially in Crete, Peloponese and Thrace), southern and central Italy, southern France, and southern Spain. The ancient Greeks and Phoenicians were the main driving forces behind the spread J2 around the western and southern Mediterranean.

J2 is thought to have arrived in Greece from Anatolia sometime between the (late) Neolithic and the Bronze Age.

### Middle-Eastern and European J2a

The geographic distribution of J2a bears a strong correlation with the diffusion of agriculture from northern Mesopotamia (where it peaks) towards Anatolia, Greece, the whole Middle East, Iran, Afghanistan, Pakistan and western India. Its strong presence in Italy is owed to the migration of the Etruscans from the Near East to central and northern Italy, and to the Greek colonisation of southern Italy.

The Phoenicians, Jews, Greeks and Romans all contributed to the presence of J2a in Iberia. The particularly strong frequency of J2a and other Near Eastern haplogroups (J1, E1b1b, T) in the south of the Iberian

peninsula, suggest that the Phoenicians played a more decisive role than other peoples. This makes sense considering that the Phoenicians/Carthaginians were the first to arrive, founded the greatest number of cities (including Gadir/Cadiz, Iberia's oldest city), and their settlements match almost exactly the higher frequency zone of southern Analusia.

The Romans surely helped spread haplogroup J2 within their borders, judging from the distribution of J2 within Europe (frequency over 5%), which bears an uncanny resemblance to the borders of the Roman Empire.

The world's maximum concentrations of J2a is in Crete (32% of the population). The subclade J2a3d (M319) appears to be native to Crete. J2a also reaches high frequencies in Anatolia and the southern Caucasus.

#### Indian J2a

Within India, J2a is more common among the upper castes and decreases in frequency with the caste level. This can be explained by the assimilation of local J2a (and R2) people from Central Asia by the R1a Indo-European warriors who descended from modern Russia (Sintashta culture) and established themselves for a few centuries in southern Central Asia, immediately north of the Hindu Kush (including the Oxus civilization) before moving on to conquer the Indian subcontinent. J2a would have reached southern Central Asia with the expansion of Middle Eastern people during the Neolithic and mixed with the local hunter-gatherers belonging chiefly to R2 (and possibly some pre-Indo-European R1a).

#### J2b

The mutation founding the J2b subclade might have originated in Greece (or in Anatolia ?), like haplogroup E-V13 (see below) to which it is closely linked. The propagation of J2b and E-V13 corresponds roughly to the ancient Greek and Roman spheres of influence. Apart from Europe, J2b is also found all around India, but only at moderate frequencies in between Europe and India, meaning that, unlike for J2a, it was not a progressive and continuous diffusion, as is to be expected from the spread of agriculture. For this reason, and because it is generally found among the upper castes of India, it is thought that some J2b lineages might have been part of the Indo-Aryan invasions of South Asia (3,500 years ago) alongside R1a1a. It is conceivable that a minority of J2b, G2a3b1 and R1b1b from the Caucasus region migrated to the Volga-Ural region in the early Bronze Age (see R1b history), spreading with them the Proto-Indo-European language, bronze technology and domestic animals to the Caspian steppe before the expansion of this new culture to Central and South Asia (see R1a history).

#### Distribution of haplogroup J2 in Europe, the Middle East & North Africa

#### Haplogroup J1

##### Main article : Haplogroup J1

J1 is a Middle Eastern haplogroup, which probably originated in eastern Anatolia, near Lake Van in central Kurdistan. Eastern Anatolia being the region where goats, sheep and cattle were first domesticated in the Middle East, haplogroup J1 is almost certainly linked to the expansion of pastoralist lifestyle throughout the Middle East and Europe. J1 can be divided in two main groups: the J1c3 (P58) subclade, and the other forms of J1 (J1\*, J1a, J1b, J1c1 and J1c2).

J1c3 (J-P58) is by far the most widespread subclade of J1. It is a typically Semitic haplogroup, making up most of the population of the Arabian peninsula, where it accounts for approximately 40% t 75% of male lineages. J-P58 is also the Cohen Modal Haplotype. Roughly half of all Cohanim belong to the J-P58 subclade. In the Hebrew Bible the common ancestor of all Cohens is identified as Aaron, the brother of Moses.

J1c3 is thought to have expanded from eastern Anatolia to the Levant, Taurus and Zagros mountains and the Arabian peninsula at the end of the last Ice Age (12,000 years ago) with the seasonal migrations of pastoralists. Arabic speakers recolonised the Arabian peninsula in the Bronze Age from the north-west of the peninsula, close to modern Jordan. The rise of Islam in the 7th century CE played a major part in the re-expansion of J1c3 from Arabia throughout the Middle East, as well as to North Africa, and to a lower extent to Sicily and southern Spain.

Other subclades of J1 are less well studied due to their much lower frequencies. Most of the J1 in the Caucasus, Anatolia and Europe is of the non-J1c3 variety. Other types of J1 most probably spread to Europe during the Neolithic. J1 is particularly common in mountainous regions of Europe (with the notable exception of the Alps and the Carpathians), like Greece, Albania, Italy, central France, and the most rugged parts of Iberia (Asturias, Cantabria, Castile-La Mancha) as well as those with the highest density of Neolithic settlements (Portugal and Andalusia).

Like haplogroup G, J1 might have been of the principal lineages to bring domesticated animals to Europe. Both G and J1 reach their maximal frequencies in the Caucasus, some ethnic groups being almost exclusively J1 (Kubachi, Kaitak, Dargins), while others have extremely high levels of G (Shapsugs, North Ossetians). Most of the ethnic groups in the North Caucasus have between 20 and 40% of each haplogroup, which are by far their two dominant haplogroups. In the South Caucasus (Georgia, Armenia, Azerbaijan), haplogroup J2 comes into the admixture and is in fact slightly higher than either J1 or G.

Distribution of haplogroup J1 in Europe, the Middle East & North Africa

Haplogroup E1b1b (Y-DNA)

Main article : Haplogroup E1b1b Haplogroup E1b1b (formerly E3b) represents the last direct major migration from Africa into Europe. It is believed to have first appeared in the Horn of Africa or southern Africa approximately 26,000 years ago and dispersed to the Middle East during the Upper Paleolithic and Mesolithic periods.

On the European continent it has the highest concentration in north-west Greece, Albania and Kosovo, then fading around the Balkans, the rest of Greece and Western Turkey. Outside Europe, it is also found in most of the Middle East, northern and eastern Africa, especially in Morocco, Lybia, Egypt Yemen, Somalia, Ethiopia and South Africa.

Did E1b1b cross directly from North Africa to Europe due to climate change ?

It is still unclear when haplogroup E entered Europe. Recent DNA tests from Neolithic sites in southern Germany and southern France lacked all trace of E1b1b. This suggests a later arrival, either towards the end of the Neolithic/Chalcolithic or during the Bronze Age. In the absence of Y-DNA from Neolithic Greece, South Italy and Iberia, nothing rules out the possibility that E1b1b was present to these regions since the Neolithic, Mesolithic or even the late Paleolithic. North Africans carriers of E1b1b could have crossed the Mediterranean (probably in several independent waves) anytime between the Last Glacial Maximum (circa 20,000 years ago) and the last desertification of the Sahara that started when the monsoon retreated south 6,200 years ago.

At the Last Glacial Maximum, sea levels were 120 metres lower than today and the Strait of Gibraltar was just a few kilometres wide, permitting even the most primitive raft to cross it easily. Is it merely a coincidence that the last attested trace of Neanderthal in Iberia (actually in Gibraltar itself) dates from 24,000 years ago, a short time before the Last Glacial Maximum ? Could their disappearance be the result of an absorption by Homo Sapiens from North Africa ? The last Iberian Neanderthals did show some signs of hybridization with Homo Sapiens. Whereas Homo Sapiens indisputably colonised Europe from the Middle East, a counter-current

colonisation from Northwest Africa is plausible too. This would explain why there is so much Northwest African E-M81 in Portugal and Northwest Spain, which is not corroborated by any historical migration nor by any archaeologically demonstrable Neolithic migration from Northwest Africa.

The Sahara changed many times from a lush green place to a hot and arid desert in the last 20,000 years. It was as arid as today at the end of the last Ice Age 13,000 years ago, then the warming climate brought tropical monsoons again from 10,000 to 7,000 years before present. The desertification taking place today started around 4,200 BCE. These severe transformations of their environment surely had a tremendous effect on the indigenous (E1b1b) people, causing population booms during the green millennia of the Neolithic, and prompting migrations to milder climates once the rain had gone. It wouldn't be all that surprising that North Africans crossed the Mediterranean (again ?) in the late Neolithic. The region most affected by the desertification would have been around modern Libya. The northern Maghreb enjoys the protection of the mountains that stopped the advance of the desert. Egypt had the Nile and its delta. One hypothesis is that the Neolithic population of Libya migrated to what is now South Italy, Greece, Macedonia and Albania, bringing with them the E-V13 lineage, which is still found in Libya today, as well as in Iberia, Egypt and the Levant, but is far more common around Greece. Alternatively, instead of crossing directly the Mediterranean from Tunisia to Sicily, then to Italy and the southern Balkans, the migration could have taken place along the coast of the Mediterranean, through Egypt, the Levant and Anatolia, and eventually to Greece. Some migrants might have taken a westward route to Iberia, explaining why E-V13 is found in western Iberia, alongside the Maghreban E-M81, while Greeks never colonised that region.

## Distribution of haplogroup E1b1b in Europe, the Middle East and North Africa

### Subclades

E1b1b1a (or E-M78, formerly E3b1a) is the most common variety of haplogroup E among Europeans and Near Easterners. E-M78 is thought to have migrated out of Egypt in the Mesolithic or Neolithic to colonise the Middle East, where it mixed with the indigenous inhabitants belonging to haplogroups J and G.

The Phoenicians, from the Levant, also contributed to the spread of E1b1b1a (as well as J2, Q and T) to places such as Cyprus, Malta, Sardinia, Ibiza and southern Iberia. The lower incidence of E1b1b1a in Syria and Anatolia is almost certainly due to the competition from the other major Neolithic haplogroups : G2 and J2.

E-M78 is divided into 4 main branches : E1b1b1a1 (E-V12), E1b1b1a2 (E-V13), E1b1b1a3 (E-V22) and E1b1b1a4 (E-V65), each further subdivided in "a" and "b" subclades.

E-V13 is one of the major markers of the Neolithic diffusion of farming from the Levant. Like all the other subclades of E1b1b1a, E-V13 originated in North-East Africa around the end of the last Ice Age. Its frequency is now far higher in Greece, South Italy and the Balkans than anywhere else due to a founder effect, i.e. the migration of a small group of settlers carrying mostly this lineage (but also a small amount of other North-East African lineages, notably E-M123 and T). Archeological evidence shows that the region of Thessaly, in northern Greece, was the starting point (circa 6,000 BCE) for the diffusion of agriculture through the Balkans and the Danube basin, as far as northern France to the west and Russia to the east. The modern distribution of E-V13 hints at a strong correlation with the Neolithic and Chalcolithic cultures of Old Europe, such as the Vinča, Boian, Maritsa and Karanovo, cultures. However, the genetic testing of three male samples from the LBK culture only revealed the presence of haplogroups F and G2a. The sample is obviously too small to rule out that E1b1b also entered Europe during the Neolithic period though.

E-V13 is also associated with the ancient Greek expansion and colonisation. Outside of the Balkans and Central Europe, it is particularly common in Southern Italy, Cyprus and Southern France, all part of the Classical ancient Greek world.

E-V22 is the predominant subclade in the Levant and is therefore associated with the Phoenicians and Jews, in addition to the spread of agriculture. The Phoenicians could have spread E-V22 to Sicily, Sardinia, southern Spain and the Maghreb, and the Jews to Greece and mainland Italy and Spain. However, the Mediterranean route for the diffusion of agriculture (see map below) went through the exact same regions. It is therefore impossible to know at present which of the two periods (Neolithic or Classical Antiquity) played the stronger role in the spread of V22 around the Mediterranean.

E-V12 is the most common subclade of M78 in Egypt. Its low presence around Greece and Anatolia indicates that it probably already existed when E moved there in the early Neolithic.

E-V65 is found in North Africa, with a maximum frequency in Lybia, then Morocco. It is also likely to have originated in Egypt. In Europe it is found at low frequencies in Greece and Sicily, but interestingly makes up one fourth of Sardinian E. It could be due to immigration from the Phoenician colonies in the Maghreb to Sardinia (the Sardinian haplogroup I2a1 is also present at low frequencies along the coast of Algeria and Tunisia, confirming exchanges of population between the two regions, maybe when both were Phoenician colonies).

E1b1b1b (E-M81, formerly E3b1b) is characteristic of the Berbers of North-West Africa. In some parts of Morocco E1b1b1b can peak at 80% of the population. This sub-haplogroup is also found in Iberia, Italy and southern France, with the highest concentrations in southern Portugal (12%) and decreasing as we move north.

E1b1b1c (E-M123) and its main branch E1b1b1c1 (E-M34) is also associated with the diffusion of agriculture and ancient Middle-Eastern civilizations. This haplogroup peaks in the southern Levant (10-12% in Palestine and Lebanon), from where it expands in all directions over the Middle East, North Africa, South Asia and South-East Europe. The distribution of E-M123 matches almost exactly the expansion of farming (see map below) during the Neolithic period. E-M123 seems to go hand in hand with haplogroup G2a, with the difference that G2a reaches its maximum frequency around the Caucasus and Anatolia, where cattle, pigs and goats were first domesticated. Inside Europe, E-M123 follows more or less the distribution of E-V13, with the highest frequency (1 to 5%) observed in Greece, South Italy, the Balkans and the Danube basin, then fading towards Germany, Poland, Ukraine and Russia, where its frequency is under 1%.

#### Haplogroup T (Y-DNA)

T is a rare haplogroup in Europe. It makes up 1% of the population on most of the continent, except in Greece, Macedonia and Italy where it exceeds 4%, and in Iberia where it reaches 2.5%, peaking at 10% in Cadiz and over 15% in Ibiza. The maximal worldwide frequency for haplogroup T is observed in East Africa (Eritrea, Ethiopia, Somalia, Kenya, Tanzania) and in the Middle East (especially the Caucasus, South Iraq, Southwest Iran, Oman and South Egypt), where it accounts for approximately 5 to 15% of the male lineages. Besides these regions and Europe, T is found in isolated pockets as far as Central Asia, India, Cameroon, Zambia and South Africa. Its highest density is actually found among the Fulbe people of Cameroon (18% of the population).

Haplogroup T originated at least 30,000 years ago, making it one of the oldest haplogroups found in Eurasia, which may explain its vast dispersal around Africa and South Asia. It also makes its place of origin uncertain. The modern distribution T in Europe strongly correlates with the Neolithic colonisation of the continent by Middle Eastern farmers, who also included members of haplogroups E1b1b, G2a, J1 and J2. The hotspot in Estonia is very likely due to a founder effect in the Neolithic population.

Although haplogroup T is more common today in East Africa than anywhere else, its association with the rise of agriculture in the Middle East is a strong argument in favour of a Middle Eastern origin, and a colonisation of East Africa by Middle Eastern farmers. Another argument in that sense is that T is descended from haplogroup K, which is itself absent from Africa and spawned most of the Eurasian haplogroups (L, N, O, P,

Q, R and T), which are thought to have a common origin around Central Asia. The strong incidence of T from the Caucasus to central and southern Iran hint that early farmers might have descended from the Caucasus to southern Mesopotamia and southwest Iran. T might therefore be linked to the ancient Sumerians and Elamites.

The higher than average frequencies of haplogroup T in places like Cyprus, Sicily, Tunisia, Ibiza, Andalusia and the northern tip of Morocco suggest that some of the haplogroup T can be attributed to the Phoenicians colonisation (1200-800 BCE), and that ancient Phoenicia seemingly had a higher incidence of T than Lebanon does today.

The testing of Thomas Jefferson's DNA revealed that the US president belonged to haplogroup T.

Distribution of haplogroup T in Europe, the Middle East and North Africa

Other haplogroups found in Europe

Haplogroup N (Y-DNA)

N is found among Uralic speakers, from Finland to Siberia, and at minor frequencies as far as Korea and Japan. In Europe, haplogroup N is only found at high frequencies among modern Finns (58%), Lithuanians (42%), Latvians (38%), Estonians (34%) and northern Russians.

Haplogroup N is believed to have originated in Southeast Asia approximately 15,000 to 20,000 years ago, but the N1c1 subclade found in Europe likely arose in Southern Siberia circa 12,000 years ago, and spread to North-East Europe 10,000 years ago.

Haplogroup N1c1 is associated with the Kunda culture (8000-5000 BCE) and the Comb Ceramic culture (4200-2000 BCE), which evolved into Finnic and pre-Baltic people. The Indo-European Corded Ware culture (3200-1800 BCE) progressively took over the Baltic region and southern Finland from 2,500 BCE. The merger of the two gave rise to the hybrid Kiukainen culture (2300-1500 BCE). Modern Baltic people have a roughly equal proportion of haplogroup N1c1 and R1a, resulting from this merger of Uralic and Slavic cultures.

Distribution of haplogroup N1c1 in Europe

Haplogroup Q (Y-DNA)

Q is found predominantly in Central Siberia, Central Asia and among Native Americans. In the latter case it is the specific subclade Q1a3a.

One hypothesis is that Q came to Europe with the Huns in the 5th century. The Huns are thought to have originated from Central Siberia, where haplogroup Q is still common nowadays. Q is found in 2% of the people in Hungary and up to 5% in isolated pockets in the mountains of Slovakia, just north of Hungary. It is historically attested that Hungary was where most of the Hunnic invaders finally settled after wreaking havoc around Europe. The Nordic and Baltic states have the second highest frequency of Q in Europe. Based on the Hunnic hypothesis, it is possible that a group of Huns settled in Sweden and/or Norway along with their allies, the Goths. The Romans reported that the Huns consisted of a small ruling elite and their armies comprised mostly of Germanic warriors. An alternative scenario is that Nordic and Baltic Q came through the Uralic-speaking population of Siberia via Finland and Lapland, but this is unlikely because Q is not more common in Finland and does not correlate with the densities of the Uralic haplogroup N1c1.

Other Central Asian or Siberian migrations might have brought Q to Ukraine in the late Antiquity or Medieval period. For instance, the multi-ethnic Central Asian troops of Genghis Khan could very well have carried

some haplogroup Q (along with C, G, O and R1a) to Eastern Europe, but not to Central Europe or Scandinavia.

## Distribution of haplogroup Q in Europe

### Haplogroup C (Y-DNA)

Haplogroup C3 in Europe is most likely of Mongol origin. It is found everywhere at various concentrations in Genghis Khan's former empire, although only sporadically on the European continent. Other subclades of C come from ethnic groups too remote from Europe (Aboriginal Australians, Polynesians, South-East Asians) to be found among Europeans (apart from recent immigration).

### Haplogroup P (Y-DNA)

P is the parent group of Q and R (including R1a and R1b). It has almost disappeared nowadays, except around its place of origins in Central Asia. It is very rarely found in Europe. It may have been brought to Europe by Central Asian invaders, like the Huns or the Mongols.

### Haplogroup L (Y-DNA)

L is found mostly in the Indian subcontinent, but also at lower frequencies in Central Asia, Southwest Asia, and Southern Europe along the coast of the Mediterranean Sea (notably in Italy). L1 is typical of the Dravidian people of South India. Various subclades are found in Europe (L1, L2, L3) without any real geographic pattern. Europeans belonging to haplogroup L are likely to be descended from Indian (L1, L3) or Persian (L2, L3) merchants in ancient times, maybe at the time of the Roman Empire.

### Haplogroup H (Y-DNA)

Gypsies belong predominantly (about 50%) to haplogroup H1a. Haplogroup H is not otherwise found in Europe, but on the Indian subcontinent.

### Haplogroup A (Y-DNA)

A is the oldest of all Y-DNA haplogroups and the closest to the Y-chromosomal Adam. It originated in Africa over 70,000 years ago, most likely in the south-west, around modern Angola and Namibia. Modern populations with the highest percentages of haplogroup A are the Khoisan (such as the Bushmen) and the southern Sudanese. Isolated cases of individuals belonging to haplogroup A have been found in Western Europe (notably Ireland, Britain and Germany). It is believed that these people descend in direct paternal line from Nubians who probably came to Europe during the Roman period, probably as slaves (Nubian gladiators were popular in Rome). It is unlikely that they descend from slaves from the Atlantic slave trade (17th and 18th century) since these came from a part of Africa where A is very rare.

## MtDNA Haplogroups

All mtDNA haplogroups found in Europe descend from the N group, which is thought to represent one of the two initial migrations by modern humans out of Africa, some 60,000 to 80,000 years ago. Nowadays haplogroup N is only found at extremely low frequencies in various parts of Eurasia.

Unfortunately, the tiny size of mitochondrial DNA (approximately 16,500 base pairs as opposed to 60 million for Y-DNA) does not allow a very accurate tracing of ancestry. Mitochondrial haplogroups all arose during the Ice Age, a period when humans were nomadic hunter-gatherers, well before the establishment of cities and

civilizations. Mitochondrial haplogroups are only linked to ethnicities at a continental level. Those associated with European descent are H, I, J, K, T, U, V, W and X (except the branch X2a which found among Native Americans). Deep subclades can be associated with more specific regions, but do not necessarily match historical ethnic and linguistic groups. One likely reason is that women, through whom mtDNA is passed, tended to marry outside their ethnic group more often than men (e.g. to secure an alliance between two tribes or kingdoms).

#### Chronological development of mtDNA haplogroups

Note that the age of mitochondrial haplogroups is much more difficult to estimate than Y-DNA haplogroups, due to the tiny sequence of mtDNA and the few number of mutations available. The error margin for the dates below is typically of +/-5,000 years, but could even exceed that for older haplogroups.

N => 75,000 years ago (arose in North-East Africa)

R => 70,000 years ago (in South-West Asia)

U => 60,000 years ago (in North-East Africa or South-West Asia)

pre-JT => 55,000 years ago (in the Middle East)

JT => 50,000 years ago (in the Middle East)

U5 => 50,000 years ago (in Western Asia)

U6 => 50,000 years ago (in North Africa)

U8 => 50,000 years ago (in Western Asia)

pre-HV => 50,000 years ago (in the Near East)

J => 45,000 years ago (in the Near East or Caucasus)

HV => 40,000 years ago (in the Near East)

H => over 35,000 years ago (in the Near East or Southern Europe)

X => over 30,000 years ago (in north-east Europe)

U5a1 => 30,000 years ago (in Europe)

I => 30,000 years ago (Caucasus or north-east Europe)

J1a => 27,000 years ago (in the Near East)

W => 25,000 years ago (in north-east Europe or north-west Asia)

U4 => 25,000 years ago (in Central Asia)

J1b => 23,000 years ago (in the Near East)

T => 17,000 years ago (in Mesopotamia)

K => 16,000 years ago (in the Near East)

V => 15,000 years ago (arose in Iberia and moved to Scandinavia)

H1b => 13,000 years ago (in Europe)

K1 => 12,000 years ago (in the Near East)

H3 => 10,000 years ago (in Western Europe)

#### Mitochondrial DNA of prehistoric Europeans

The testing of ancient DNA helped understand how long each haplogroup has been in Europe. Only a few such tests have been successfully conducted so far. Mitochondrial DNA was extracted from the skeleton of a 28,000 year-old Cro-Magnon from southern Italy, and the haplogroup was determined as HV or pre-HV. Still preceding the Neolithic expansion from the Middle East, the 9,000 year-old Cheddar Man was found to belong to haplogroup U5a. (=> More examples of ancient mtDNA haplogroups).

Autochthonous (Cro-Magnoid) Europeans must have therefore belonged at least to haplogroups HV (and its offspring H and V) as well as U5a, which also happen to be the most common mitochondrial haplogroup everywhere in Europe. It has been speculated that over half of the matrilineal lineages in Europe descend directly from Paleolithic Europeans. Their male counterpart are Y-DNA haplogroup I.

#### European mtDNA haplogroups and their subclades

### Haplogroup H & V (mtDNA)

Haplogroup H is by far the most common all over Europe, amounting to about 40% of the European population. It is also found (though in lower frequencies) in North Africa, the Middle East, Central Asia, Northern Asia, as well as along the East coast of Africa as far as Madagascar.

H1, H3 and V are the most common subclades of HV in Western Europe. H1 peaks in Norway (30% of the population) and Iberia (18 to 25%), and is also high among the Sardinians, Finns and Estonians (16%), as well as Western and Central European in general (10 to 12%) and North-West Africans (10 to 20%). H3 is commonest in Portugal (12%), Sardinia (11%), Galicia (10%), the Basque country (10%), Ireland (6%), Norway (6%), Hungary (6%) and southwestern France (5%). Haplogroup V reaches its highest frequency in northern Scandinavia (40% of the Sami), northern Spain, the Netherlands (8%), Sardinia, the Croatian islands and the Maghreb. It is likely that H1, H3 and V, along with haplogroup U5, were the main haplogroups of Western European hunter-gatherers living in the Franco-Cantabrian refuge during the last Ice Age, and repopulated much of Central and Northern Europe from 15,000 years ago.

Haplogroup H13 is most common in Sardinia and around the Caucasus. Its distribution is reminiscent of Y-DNA haplogroup G2a. The same is true of H2 to a lower extent. This would suggest a Caucasian or Anatolian origin.

H5 and H7 are also common in the Caucasus, but their lower incidence around the Mediterranean, and higher frequency from Anatolia to the Alps via the Danube suggest a possible link with the spread of agriculture (YDNA E1b1b, J2 and T) or of the Indo-Europeans (R1b1b2).

### Haplogroup U & K (mtDNA)

Haplogroup U is extremely old. It originated some 60,000 years ago at the confine of North-East Africa and the Middle East, soon after the first Homo Sapiens ventured out of Africa. This is why each of its top-level subclade (U1, U2, U3...) can be seen as a haplogroup in its own right. The main European subclades are U3, U4, U5 and U8/K. U1 is mostly found in the Middle East, U6 in North Africa, U7 from the Near East to India, and the rare U9 from Ethiopia and the Arabian peninsula to Pakistan.

Haplogroup U2 is found primarily in South Asia, but probably is of Indo-European origin as it is found at low frequencies throughout the Pontic-Caspian steppe and has been identified in a 30,000 year-old Cro-Magnon from the middle Don valley in Russia. It might have been the dominant haplogroup of the northern forest-steppe foragers who later became the Proto-Indo-Iranian speakers (see R1a above) and moved massively to Central and South Asia.

Haplogroup U3 is centered around the Black Sea, with a particularly strong concentration in the north-eastern part. It could be related to the ancient Indo-Europeans, and probably more to R1b than R1a.

Haplogroup U4 are more common in Eastern Europe, Central Asia, northern South Asia (around Tajikistan for U4, and Pakistan for W), which also suggests an affiliation with the Indo-Europeans (correlated to Y-DNA haplogroup R1a). The same is true of haplogroups I, W, T2 and U2e to a lower extent.

Haplogroup U5 is the most common in Western and Northern Europe. DNA tests on ancient skeletons have shown that U5 was the principal mitochondrial haplogroup of Paleolithic and Mesolithic hunter-gatherers in Northern Europe. Ancient DNA tests conducted in Britain, Germany and Scandinavia indicate that the frequency of U5 has progressively declined over time through the Neolithic, Bronze Age, Iron Age and Middle Ages. Nowadays it remains most common in the far north of Europe, where the Mesolithic population has been least affected by subsequent migrations. For instance, 30 to 50% of the Sami people of northern Scandinavia belong to haplogroup U5b (and about 40% to haplogroup V, which is also pre-Neolithic European origin).

Haplogroup K is the main subclade of U8. It is found throughout Europe and Western Asia, as far away as India. Its highest concentration is in North-West and Central Europe, Anatolia and the southern Arabian peninsula. It is believed to have first arisen somewhere between Egypt and Anatolia approximately 16,000 years ago (estimates range from 22,000 years to as little as 10,000 years before present). It has the largest number of subclades of any haplogroup in spite of its fairly recent age. K1a is the largest subclade. The relatively important presence of K1a in the Near East suggest that it predates the Neolithic migration to Europe. This has been supported by the ancient mtDNA from Neolithic sites. Haplogroup K was never found in Europe prior to the Neolithic, then suddenly appears at a frequency (17%) much higher than in modern Europeans and similar to that of the present-day Levant. Most of the Neolithic K belongs to the K1a subclade.

Most K1a4, K1a10, K1b, K1c and K2 subclades are typically European. K1a4 is also common in Anatolia and Greece, and could indeed have spread to the rest of Europe from there during the Neolithic period, along with haplogroups J and T (and Y-DNA haplogroups E1b1b, J2 and T). The Indo-Europeans from Anatolia could also have contributed to the propagation of K. K1a1b1a and K1a9 are found primarily among Ashkenazi Jews.

#### Haplogroup J & T (mtDNA)

Haplogroup J originated in the Middle East 45,000 years, making it one of the oldest mitochondrial haplogroups in Europe and the Middle East. It is usually associated with the spread of agriculture. Haplogroup J being so common in Central Asia and around the Caspian and Black Sea, it is likely to have also a connection with the Indo-Europeans, especially the migration of Y-DNA haplogroup R1b (see R1b history above). J1 is common throughout the Middle East, as far as Central Asia and around Ukraine. In the rest of Europe it is mostly confined to Germanic countries (mimicking the distribution of Y-DNA haplogroup I1). J2 is much rarer than J1. J2a is found homogeneously across most of Europe. J2b is more frequent around Anatolia and in South-East Europe.

Haplogroup T is thought to have originated in the Middle East or North-East Africa at least 12,000 years ago. It is found throughout Europe, the northern half of Africa to Central Asia and Siberia, with pockets in India and North-West China (Xinjiang). The highest concentration of T1 has been observed in North-East Africa, Anatolia and Bulgaria, which suggests a Neolithic diffusion from Egypt to the Balkans. T2, the most subclade of T in Europe, is particularly common in North-East Europe and around the Aegean Sea. The overall distribution of haplogroup T points at an early Neolithic migration from North-East Africa to Eastern Europe, then a dispersal following the migration pattern of the Indo-Europeans (especially Y-DNA haplogroup R1a) to Europe and South Asia.

#### Haplogroup W (mtDNA)

Present at low frequencies in most of Europe, in Anatolia, around the Caspian Sea, and from the Indo-Pakistani border to Xinjiang, haplogroup W is one of the best maternal markers of Indo-European ancestry (mtDNA equivalent of R1a and R1b). Its highest frequency is in Ukraine, European Russia, Baltic countries and Finland (3 to 5% overall), as well as in northern Pakistan (15%), Punjab (9%) and Gujarat (12%). In Indian it is considerably more common among the upper castes and among Indo-European speakers (source).

#### Haplogroup I (mtDNA)

Haplogroup I has a similar distribution to haplogroup W, ranging from Europe to Pakistan and North-West India, with a characteristic presence in Pontic steppes and around the Caspian Sea. Its origin very probably lies in the Proto-Indo-European cultures (mtDNA mirror of R1a and R1b). Haplogroup I is nearly absent in parts of Europe from distant from the Pontic-Caspian steppes (Iberia, South-West France, Ireland) and strongest in Norway, southern Finland, Ukraine, Greece and western Anatolia.

#### Haplogroup X (mtDNA)

Haplogroup X is a very old and scattered haplogroup found all over Eurasia, North Africa as well as among Native North Americans. Its frequency rarely exceeds 5% of the population in any ethnic group, and is more often restricted to 1 or 2%. X1 is found almost exclusively in North Africa, while X2b is the only lineage present among Amerindians. X2a, X2c, X2d and X2e are found in Europe, Siberia and Central Asia. It is therefore possible that the latter be of Indo-European origin (R1b1b).

The strong presence of X2 around the Caucasus, progressively fading towards the Near East and Mediterranean, hints that it could be related to the spread of Y-DNA haplogroup G2a. R1b1b and G2a both having origins around the Caucasus it is unsurprising to find X2 alongside these two Y-DNA haplogroups.

See also : Identifying the original Indo-European mtDNA from isolated settlements

### Haplogroup R (mtDNA)

Haplogroup R is the main subclade of N, the one that was to generate the 6 most common European haplogroups (H, V, J, T, U, K). At the time of writing R subclades were numbered from R0 (a.k.a. pre-HV) to R31. Most of them are found in South Asia (R5, R6, R7, R8, R30, R31), Southeast Asia (R9, R21, R22, R24), East Asia (R9/F, R11/B), and even among Papuans (R14) and Australian aborigines (R12). R0a peaks in the southern Arabian peninsula is common among Arabs and Middle-Easterners. R1a (not to be confused with the homonymous Y-chromosome haplogroup) is found among the Adygei people from the North Caucasus (related to the Maykop culture => see R1b section), Brahmins from northern India, northwestern Russians and Poles - basically all people closely related with the Indo-European expansion. R2 is found from northwest India and Pakistan to Iran, Georgia and Turkey. It could be connected to the Indo-Iranians.

### Finno-Uralic mtDNA

Finno-Uralic people have an overall mtDNA admixture similar to other Europeans, with a higher percentage of W and U5b, and a small percentage of Siberian haplogroups such as N or A. The Sami are characterised by a high percentage of haplogroups U5b1 and V.

### Berber mtDNA

The Berbers are the indigenous population of north-west Africa. Although their Y-DNA is almost perfectly homogenous, belonging to haplogroup E-M81, Berber maternal lineages show a much greater diversity, as well as regional disparity. At least half (and up to 90% in some regions) of the Berbers belong to some Eurasian lineages, such as H, HV, R0, J, T, U, K, N1, N2, and X2, mostly of Middle or Near Eastern origin. 5 to 45% of the Berbers will have sub-Saharan mtDNA (L0, L1, L2, L3, L4, L5). There are only three native North African lineages, U6, X1 and M1, representing 0 to 35% of the people depending on the region.

Haplogroup U6 has been observed from the Iberia and the Canary Islands to Senegal in the West, and from Syria to Ethiopia and Kenya in the East. It is also found at low density in Europe, though mostly limited to Iberia. Approximately 10% of all North Africans belong to this lineage.

### Gypsy mtDNA

The Gypsies (Romani people) originated in the Indian subcontinent and mixed with local population in the Middle East and Eastern Europe over the centuries. About half of the Gypsy population belong to haplogroup M, and more specifically M5 (reflected by Y-haplogroup H1a), which is otherwise exclusive to South Asia. The other mtDNA haplogroups found among the Gypsy community are mostly of Eastern European, Caucasian or Middle Eastern origin, such as H (H1, H2, H5, H9, H11, H20, among others), J (J1b, J1d, J2b), T, U3, U5b, I, W et X (X1b1, X2a1, X2f) (sources). The same diversity exist on the Y-DNA side (45% of H1a, followed by I1, I2a, J2a4b, E1b1b, R1b1b, R1a1a).

## Sources

The list below is non-exhaustive and include many of the numerous references linked on these websites. Some studies and databases not published on the Web were also used.

### Haplogroup-specific studies (Y-DNA)

Phylogeography of Y-Chromosome Haplogroup I Reveals Distinct Domains of Prehistoric Gene Flow in Europe, Rootsi et al. (PDF)

Y chromosomal haplogroup J as a signature of the post-neolithic colonization of Europe, Di Giacomo et al. (PDF)

The emergence of Y-chromosome haplogroup J1e among Arabic-speaking populations, Chiaroni et al. (PDF)

J1-M267 Y lineage marks climate-driven pre-historical human displacements, Tofanelli et al. (PDF)

Tracing Past Human Male Movements in Northern/Eastern Africa and Western Eurasia: New Clues from Y-Chromosomal Haplogroups E-M78 and J-M12, Cruciani et al. (PDF)

Origin, Diffusion, and Differentiation of Y-Chromosome Haplogroups E and J: Inferences on the Neolithization of Europe and Later Migratory Events in the Mediterranean Area, Semino et al. (PDF)

Phylogeographic Analysis of Haplogroup E3b (E-M215) Y Chromosomes Reveals Multiple Migratory Events Within and Out Of Africa, Cruciani et al. (PDF)

Separating the post-Glacial coancestry of European and Asian Y chromosomes within haplogroup R1a, Underhill et al. (PDF)

A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe, Myres et al. (PDF)

### Iberian peninsula

Reduced genetic structure of the Iberian peninsula revealed by Y-chromosome analysis: implications for population demography, Flores et al. (PDF)

The Genetic Legacy of Religious Diversity and Intolerance: Paternal Lineages of Christians, Jews, and Muslims in the Iberian Peninsula, Adams et al. (PDF)

Micro-phylogeographic and demographic history of Portuguese male lineages., Belezza et al.

Y-chromosome Lineages from Portugal, Madeira and Acores Record Elements of Sephardim and Berber Ancestry, Gonçalves et al. (PDF)

Y Chromosome and Mitochondrial DNA Characterization of Pasiegos, a Human Isolate from Cantabria (Spain), Maca-Meyer et al. (PDF)

The Basques in the Genetic Landscape of Europe, by Kristin Leigh Young (PDF)

### Italy & Greece

Y chromosome genetic variation in the Italian peninsula is clinal and supports an admixture model for the Mesolithic-Neolithic encounter, Capelli et al. (PDF)

Clinal patterns of human Y chromosomal diversity in continental Italy and Greece are dominated by drift and founder effects, Di Giacomo et al. (PDF)

Population Structure in the Mediterranean Basin: A Y Chromosome Perspective, Capelli et al. (PDF)

Peopling of three Mediterranean islands (Corsica, Sardinia, and Sicily) inferred by Y-chromosome biallelic variability, Francalacci, et al. (PDF)

Differential Greek and northern African migrations to Sicily are supported by genetic evidence from the Y chromosome, Di Gaetano et al.

Y-chromosome genetic structure in sub-Apennine populations of Central Italy by SNP and STR analysis, Onofri et al.

Slow and fast evolving markers typing in Modena males (North Italy), Ferri et al. (PDF)

Male haplotypes and haplogroups differences between urban (Rimini) and rural area (Valmarecchia) in Romagna region (North Italy), Ferri et al.

Y-Chromosome Based Evidence for Pre-Neolithic Origin of the Genetically Homogeneous but Diverse Sardinian Population: Inference for Association Scans, Contu et al. (PDF)

Peopling of Three Mediterranean Islands (Corsica, Sardinia, and Sicily) Inferred by Y-Chromosome Biallelic Variability, Francalacci et al. (PDF)

Paleolithic Y-haplogroup heritage predominates in a Cretan highland plateau, Martinez et al. (PDF)

Y-chromosomal evidence for a limited Greek contribution to the Pathan population of Pakistan, Firasat et al. (PDF)

The coming of the Greeks to Provence and Corsica: Y-chromosome models of archaic Greek colonization of the western Mediterranean, King et al. (PDF)

France & Benelux

Micro-geographic distribution of Y-chromosomal variation in the central-western European region Brabant, Larmuseau et al.

Phylogeography of French male lineages, Ramos-Luis et al. (PDF)

Central Europe & Balkans

High-Resolution Phylogenetic Analysis of Southeastern Europe Traces Major Episodes of Paternal Gene Flow Among Slavic Populations, Pericic et al. (PDF)

Significant genetic differentiation between Poland and Germany follows present-day political borders, as revealed by Y-chromosome analysis, Kayser et al. (PDF)

MtDNA and Y chromosome polymorphisms in Hungary: inferences from the palaeolithic, neolithic and Uralic influences on the modern Hungarian gene pool, Semino et al. (PDF)

Y-Chromosomal Variation in the Czech Republic, Luca et al. (PDF)

The genetic structure of the Slovak population revealed by Y-chromosome polymorphisms, Petrejčíková et al. (PDF)

The Peopling of Modern Bosnia-Herzegovina : Y-chromosome haplogroups of the three main ethnic groups, Marjanovic et al. (PDF)

Y chromosomal heritage of Croatian population and its island isolates, Barac et al. (PDF)

Paternal and maternal lineages in the Balkans show a homogeneous landscape over linguistic barriers, except for the isolated Aromuns, Bosch et al. (PDF)

Y-chromosomal evidence of the cultural diffusion of agriculture in southeast Europe, Battaglia et al. (PDF)

Northern Europe

Y-chromosome diversity in Sweden - A long-timeperspective, Karlsson et al. (PDF)

Geographical heterogeneity of Y-chromosomal lineages in Norway, Dupuy et al. (PDF)

Different genetic components in the Norwegian population revealed by the analysis of mtDNA and Y chromosome polymorphisms, Passarino et al. (PDF)

Estimating Scandinavian and Gaelic Ancestry in the Male Settlers of Iceland, Helgason et al. (PDF)

Excavating Past Population Structures by Surname-Based Sampling: The Genetic Legacy of the Vikings in Northwest England, Bowden et al. (PDF)

A Y Chromosome Census of the British Isles, Capelli et al. (PDF)

People of the British Isles: preliminary analysis of genotypes and surnames in a UK-control population, Winney et al. (PDF)

Y-chromosomal diversity suggests that Baltic males share common Finno-Ugric-speaking forefathers, Laitinen et al.

Y Chromosome and Mitochondrial DNA Variation in Lithuanians, Kasperaviciute et al. (PDF)

The Western and Eastern Roots of the Saami—the Story of Genetic “Outliers” Told by Mitochondrial DNA and Y Chromosomes, Tambets et al. (PDF)

Eastern Europe & Russia

Two Sources of the Russian Patrilineal Heritage in Their Eurasian Context, Balanovsky et al.

Y-Chromosome distribution within the geo-linguistic landscape of northwestern Russia, Mirabal et al. (PDF)

Gene Pool Structure of Eastern Ukrainians as Inferred from the Y-Chromosome Haplogroups, Kharkov et al.

Structure of the Gene Pool of Bashkir Subpopulations, Lobov, et al. (PDF)

Genetic Evidence for the Mongolian Ancestry of Kalmyks, Nasidze et al. (PDF)

Near East & Caucasus

Excavating Y-chromosome haplotype strata in Anatolia, Cengiz Cinnioglu et al. (PDF)

Testing hypotheses of language replacement in the Caucasus: evidence from the Y-chromosome, Nasidze et al.

Mitochondrial DNA and Y-Chromosome Variation in the Caucasus, Nasidze et al. (PDF)

Culture creates genetic structure in the Caucasus: Autosomal, mitochondrial, and Y-chromosomal variation in Daghestan, Marchani et al. (PDF)

MtDNA and Y-chromosome Variation in Kurdish Groups, Naside et al. (PDF)

Geographical Structure of the Y-chromosomal Genetic Landscape of the Levant: A coastal-inland contrast, El-Sibai et al.

## Y-DNA Projects & Databases

FTDNA : Ireland Y-DNA Project

FTDNA : Scottish DNA Project

FTDNA : Wales Cymru DNA Project

FTDNA : German Language Area DNA Research Project

FTDNA : French Heritage Project

FTDNA : House of Spain DNA Project

FTDNA : Italy DNA Project

FTDNA : Haplogroup G (Y-DNA) Project

FTDNA : Haplogroup I2\* Project

FTDNA : Haplogroup I2b1 Project

FTDNA : Haplogroup I2b2 Project

FTDNA : Haplogroup Q Project

R1b-S28/U152 database, David Faux

Iberian Roots : commercial Y-DNA database of Spanish and Portuguese regions

Multiple regions (Y-DNA)

Y-Chromosomal Diversity in Europe Is Clinal and Influenced Primarily by Geography, Rather than by Language, Rosser et al. (PDF)

The Genetic Legacy of Paleolithic Homo sapiens sapiens in Extant Europeans: A Y Chromosome Perspective, Semino et al.

Where Did European Men Come From, Kalevi Wiik (PDF)

Human Y-Chromosome Variation in the Western Mediterranean Area: Implications for the Peopling of the Region, Scozzari et al.

Multiple regions (mtDNA)

Geographic Patterns of mtDNA Diversity in Europe, Lucia Simoni et al.

Phylogenetic Network for European mtDNA, Saara Finnila et al.

High-resolution mtDNA evidence for the late-glacial resettlement of Europe from an Iberian refugium, Luísa Pereira et al.

Y-Chromosomal Diversity in Europe Is Clinal and Influenced Primarily by Geography, Rather than by Language, Rosser et al.

Ancient DNA

Ancient DNA provides new insights into the history of south Siberian Kurgan people, Keyser et al.

Ancient DNA Reveals Lack of Continuity between Neolithic Hunter-Gatherers and Contemporary Scandinavians, Malmström et al.

Ancient DNA from European Early Neolithic Farmers Reveals Their Near Eastern Affinities, Haak et al.

Molekulargenetische Verwandtschaftsanalysen am prähistorischen Skelettkollektiv der Lichtensteinhöhle, Schilz et al. (PDF)

News from the west: Ancient DNA from a French megalithic burial chamber, Deguilloux et al. (PDF)

Ancient DNA reveals male diffusion through the Neolithic Mediterranean route, Lacan et al.

Palaeogenetic evidence supports a dual model of Neolithic spreading into Europe, Sampietro et al.

Books on European prehistory

The Horse, The Wheel, and Language: How Bronze-Age Riders from the Eurasian Steppes Shaped the Modern World, by David Anthony

The Lost World of Old Europe, by David Anthony

Encyclopedia of Prehistory: Europe, Volume 4, By Peter Neal Peregrine and Melvin Ember

Useful links

ISOGG (International Society of Genetic Genealogy)

The Journey of Mankind : interactive map of the spread of humans on Earth

Y-search : public Y-DNA database

Mitosearch : public mtDNA database

SNPedia : human DNA mutations database

Mitomap : mtDNA tree with mutations (PDF)

---

Copyright © 2004-2011 Eupedia All Rights Reserved

About Eupedia - Privacy Statement - Terms of Use - Contact - Advertising - Sitemap  
[http://www.eupedia.com/europe/origins\\_haplogroups\\_europe.shtml](http://www.eupedia.com/europe/origins_haplogroups_europe.shtml)