y-Haplogroups I1 and R1b in European Countries, plus Ancient Migrations within Europe

Abstract

A concise summary of some of the ancient migrations of the people within Europe is given for general interest, and also to put into historical context TMRCA's and the various "Frequency by Region" plots.

This report compiles various sources of y-Haplogroup data for the common y-Haplogroups in Europe, and presents them in tables showing their frequency distribution for each region (or country) of Europe. In the case of **I1**, the proposed I1 STR clusters/clans¹ are presented in the same style of table.

Previously known east-to-west increasing frequency trends can be seen for y-Haplogroup **R1b**, which has been suggested as being related to the expansion of farming from the Near East into Europe during the Neolithic (from 7,000 BC to 4,000 BC) and the subsequent interaction with the indigenous hunter-gatherers of Europe at the time.

In the case of y-Haplogroup I1, whose MRCA may have existed after, or near the end of, the transition from hunting-gathering to farming in Europe, there is a south-to-north increasing frequency trend. The following more recent (as of say ~200 years ago) geographical frequency distributions are of note for the I1 STR clusters/clans:

- ★ I1-"BBA" is the dominant I1 STR cluster/clan in Norway and Sweden. Nearly 50% of Norway and Sweden's I1 population is I1-"BBA". It is also prevalent in other countries such as the Netherlands/Belgium and Denmark. Many I1-"BBA" and virtually all I1-"BBB" people have the L22 SNP mutation for I1d. This is consistent with the L22 mutation originally happening in a "BB" individual before the split into "BBA" and "BBB".
- * I1-"BAA" and I1-"BAB" (collectively called I1-"BA*") are the dominant I1 STR clusters/clans in Finland and account for about 75% of Finland's I1 population. Those two clusters/clans are associated with the L258 SNP mutation for I1d3a. Due to a back-mutation in DYS511 they could otherwise have been classified I1-"BBA". Within Finland, I1-"BAA" is more likely in western areas (and even across to Sweden/Norway), and I1-"BAB" is more likely in eastern areas. (Note that a DYS459a=7 cluster should be excluded from I1-"BA*".)
- ★ I1-"AABB" is relatively high in Ireland, Scotland, England, and the Netherlands/Belgium. That cluster/clan is associated with the L338 SNP mutation for I1f. So far, 16 out of 16 known L338+ people would be classified as I1-"AABB". Up to 10% of I1 people may have that mutation.
- ★ I1-"AABA" is particularly high in Wales but the sample size is low. So far, 8 out of 19 Welsh-ancestry I1 people, or about 40%, would be classified as I1-"AABA", whereas the European-wide rate is less than 10%.
- ★ I1-"AAA" is the dominant I1 STR cluster/clan in Europe as a whole. About 25% of European I1 people would be classified as I1-"AAA". There is a slight decrease in frequency from southern to northern Europe, due to the increasing frequency of I1-"BBA" in the north.

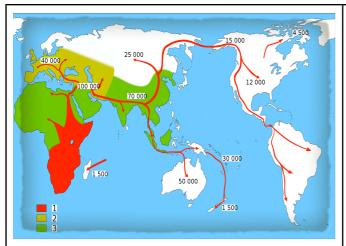
Some SNP mutations in y-Haplogroup II are found to be associated with a corresponding II STR cluster/clan.

Table of Contents

Ancient Migrations/Invasions within Europe	Page 2, 3
European y-Haplogroup Frequency by Region	Page 4
European R1b Sub-clade Frequency by Region	Page 5
European I Sub-clade Frequency by Region	Page 6
European I1 STR Cluster/Clan Frequency by Region	Page 7
European I1 STR Cluster/Clan Sample Size by Region	Page 8
European I1 STR Cluster/Clan MRCA	Page 9

¹ See http://www.goggo.com/terry/HaplogroupI1/ for the I1 STR Decision Tree showing the various I1 STR clusters/clans.

Ancient Migrations to and within Europe (page 1 of 2)



Modern Man to Europe (40,000 years ago)

The spreading ([Red] lines) of early modern man out of Africa.

[Red] - modern humans

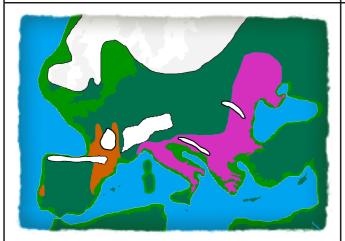
[Brown] - Neanderthals

[Green] - Homo Erectus

Modern humans shared Europe with the Neanderthals for 10,000 years or more, before the Neanderthals disappear from the fossil record around 30,000 years ago. World population remained less than 5 million people up until the expansion of farming in Europe from around 7,000 BC.

Image Source: http://en.wikipedia.org/wiki/

File:Spreading_homo_sapiens.svg



"Ice Age" Refugia (19,000 BC - 13,000 BC)

About 20,000 years ago, ice covered much of northern Europe. Sea-levels were lower too, and Britain and Ireland were joined by land to continental Europe. Refugia were situated in northern Iberia/south-west France, Italy, the Balkans, and Ukraine.

[White] - Ice sheet

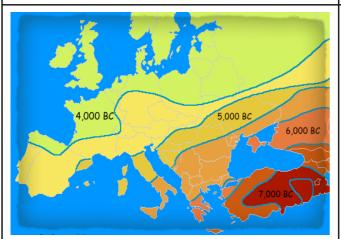
[Green] - Land at 20,000 years ago; [Dark Green] - Land today [Orange] - Solutrean Culture (came after Gravettian Culture)

[Magenta] - Gravettian Culture

After the ice retreats, these refugia people repopulate Europe.

Image Source: http://en.wikipedia.org/wiki/

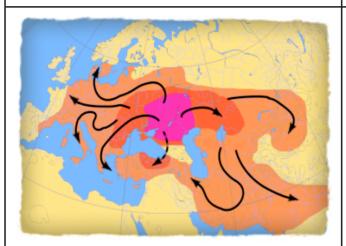
File:Europe20000ya.png



Expansion of Farming (7,000 BC - 4,000 BC)

The initial arrival of Neolithic farmers from the Near East to Greece occurred ~7,000 BC. The first wave (associated with impressed pottery) involved a maritime colonization of Crete ~7,000 BC, Southern Italy ~6,000 BC, and subsequently spread to coastal Mediterranean France and Spain. The second wave (associated with LBK) involved a migration to Central Europe, from Hungary to France, ~5,500 BC. Within a 3,000-year period, farming had replaced the indigenous hunter-gatherer culture of Europe, and reached Britain and Scandinavia ~4,000 BC.

Image Source: http://en.wikipedia.org/wiki/ File:European_Middle_Neolithic.gif



Indo-European Expansion (4,000 BC - 1,000 BC)

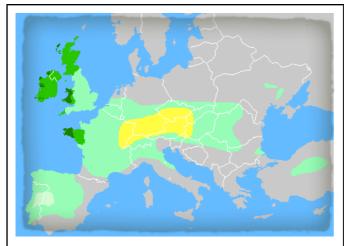
Indo-European migrations according to the Kurgan model. The Anatolian migration (dotted arrow) is across Caucasus or Balkans. [Magenta] - assumed homeland (Eastern Ukraine) at ~4,000 BC [Red] - settled by Indo-European-speaking people by ~2,500 BC [Orange] - settled by ~1,000 BC.

They met the pre-Celtic Urnfield culture (1,300 BC - 700 BC). Bronze Age cultures traded (grinding stones, flint, amber, salt) with each other along rivers (using canoes), and paths (using wagons drawn by oxen, reindeer, or later horses).

Image Source: http://en.wikipedia.org/wiki/

File:IE_expansion.png

Ancient Migrations/Invasions within Europe (page 2 of 2)



Migrations of "Celts" (1,000 BC - 250 BC)

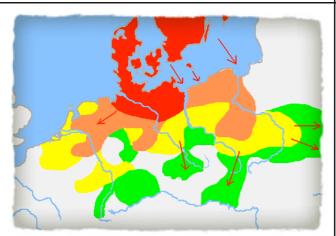
Distribution of Celtic people:

[Yellow] - Core Hallstatt territory, by 500 BC [Light Green] - Maximal Celtic expansion, by 250 BC [Green] - Areas with significant numbers of Celtic speakers [Dark Green] - Areas where Celtic languages spoken today

Gaulish-Celts went to France by 700 BC. Iberian-Celts went to Spain by 600 BC. Brythonic-Celts went to Britain and Ireland by 450 BC. Alpine-Celts remained around original core territory.

Image Source: http://en.wikipedia.org/wiki/

File:Celts_in_Europe.png



Migrations of "Germanics" (750 BC - 100 AD)

Expansion of Germanic tribes:

[Red] - Original settlements before 750 BC

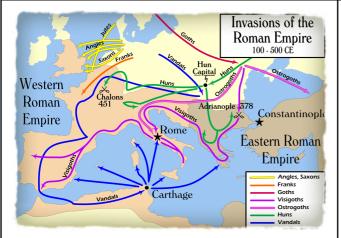
[Orange] - New settlements until 500 BC

[Yellow] - New settlements until 250 BC

[Green] - New settlements until 1 AD

The ancient Germanic and Celtic people split north of the Alps in ancient times, and intermixed with indigenous people as they migrated. **Germanic Tribes** were in Scandinavia by 1,500 BC.

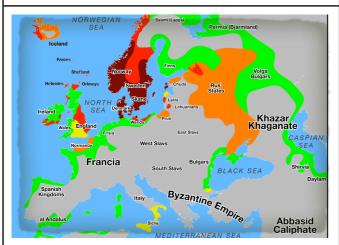
Image Source: http://en.wikipedia.org/wiki/ File:Germanic_tribes_(750BC-1AD).png



Migrations of "Barbarians" (100 AD - 500 AD)

Germanic tribes contributed to the collapse of Western Roman Empire. Tribes such as the Goths and Vandals split away. Huns (a non-Germanic people from Central Asia) attack the Ostrogoths in 376 AD and push other Germanic tribes westward. The Ostrogoths invade Italy. The Franks invade across the Rhine in 406 AD. The Vandals migrate to Iberia in 409 AD, then driven out by the Visigoths. The Visigoths invade Italy in 410 AD, later migrating to Iberia. The Angles, Saxons, and Jutes invade England in 450 AD. (Irish Gaelic Scoti spread to Scotland in ~500 AD.)

Image Source: http://en.wikipedia.org/wiki/ File:Invasions_of_the_Roman_Empire_1.png



Migrations of "Vikings" (800 AD - 1100 AD)

Scandinavian settlements:

[Brown] - 8th century AD [Coastal Norway/Sweden and Denmark] [Red] - 9th century AD [Norse/Dane settlers to Britain, Ireland]

[Orange] - 10th century AD [Norse to France; Swedes to East]

[Yellow] - 11th century AD [Normans invade England, Southern Italy]

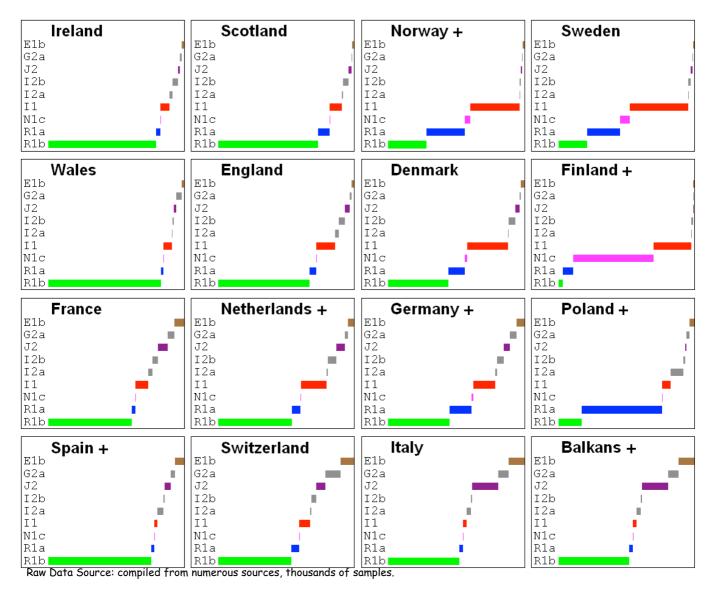
[Green] - Areas subjected to Viking raids, but little settlement

There were **Norse Vikings** (to Britain, Ireland), **Danish Vikings** (to Eastern England), and **Swedish Vikings** (to Eastern Europe).

Image Source: http://en.wikipedia.org/wiki/

File: Viking_Expansion.svg

European y-Haplogroup Frequency by Region¹



R1b, shown as [Green], is the dominant y-Haplogroup in Europe. It is less prevalent in Scandinavian, and Poland. Over 100 million European men belong to R1b, with an increasing frequency from eastern to western Europe.

II, shown as [Red], is dominant in Norway and Sweden. Over 20 million European men belong to II.

N1c, shown as [Magenta], is dominant in Finland, with I1 following next. The R1b y-Haplogroup is rare in Finland.

R1a, shown as [Blue], is dominant in Poland. And I2a, shown as [Gray], is relatively high in Poland compared to other countries, but still at very low levels compared to other y-Haplogroups.

J2 and E1b1, shown as [Purple] and [Brown], are prominent in Italy and the Balkan states. Those two y-Haplogroups are rare in Northern Europe.

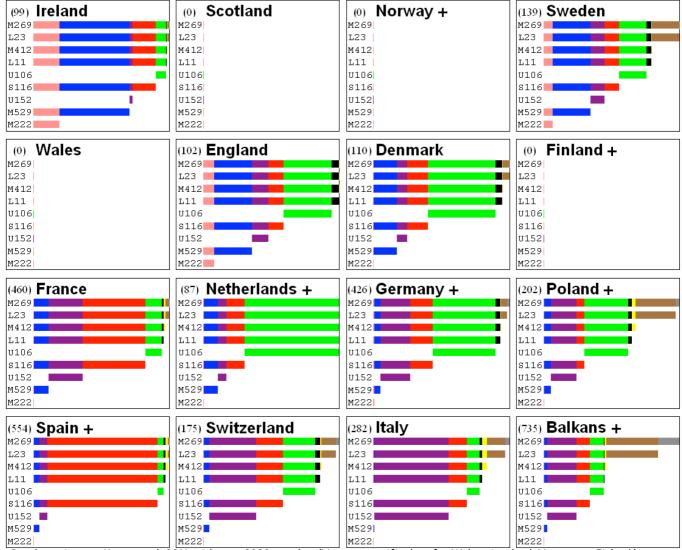
Norway+ = Norway & Iceland Finland+ = Finland & Estonia

Netherlands+ = Netherlands & Belgium Germany+ = Germany & Austria & Czech Republic

Poland+ = Poland & Lithuania Spain+ = Spain & Portugal

Balkans+ = (Albania, Bosnia-Herzegovina, Bulgaria, Croatia, Greece, Macedonia, Montenegro, Serbia, Slovenia) & Hungary, Moldova, Romania, Turkey

European R1b Sub-clade Frequency by Region¹



Raw Data Source: Myres et al, 2011, with over 2000 samples. (Note no specific data for Wales, Scotland, Norway, or Finland.)

R1b-M269, also known as R1b1a2 (ISOGG-2011), is the most common sub-clade of R1b in Europe as a whole. The graphic above shows R1b-M269 and its various deeper sub-clades using different colors.

R1b-M222, shown as [Pink], is prominent in Ireland and Scotland.

R1b-M529(xM222), shown as [Blue], is prominent in Ireland, and also Britain and Norway/Sweden.

R1b-U152, shown as [Purple], is prominent in Switzerland, and Italy.

R1b-S116(xU152,xM529), shown as [Red], is prominent in Spain/Portugal, and also France.

R1b-U106, shown as [Green], is prominent in the Netherlands/Belgium, Denmark, and Germany.

R1b-L11(xU106,xS116), shown as [Black], is present at low levels in England, and Denmark.

R1b-L23(xM412), shown as [Brown], is present in the Balkan states, Poland, Sweden, Italy, and Switzerland.

R1b-M269(xL23), shown as [Gray], is present in the Balkan states.

```
ISOGG-2011 R1b1a2*
M269
     +- L23 +-
                                                          ISOGG-2011
ISOGG-2011
                                                                     R1b1a2a*
            +- M412 +--
                                                                                      (M412/T.51)
                                                          ISOGG-2011 R1b1a2a1a1*
                     +- L11 +----
                                                                                      (L11/P310)
                                                U106
                                                          ISOGG-2011
                                                                      R1b1a2a1a1a
                                                                                      (U106/S21
                            +- S116 +----
                                                          ISOGG-2011
                                                          ISOGG-2011 R1b1a2a1a1b3
                                                U152
                                                                                      (U152/S28)
                                    +- M529 +-
                                                          ISOGG-2011
                                                                     R1b1a2a1a1b4*
                                                                                      (M529/L21)
                                             +- M222
                                                        = ISOGG-2011 R1b1a2a1a1b4b
```

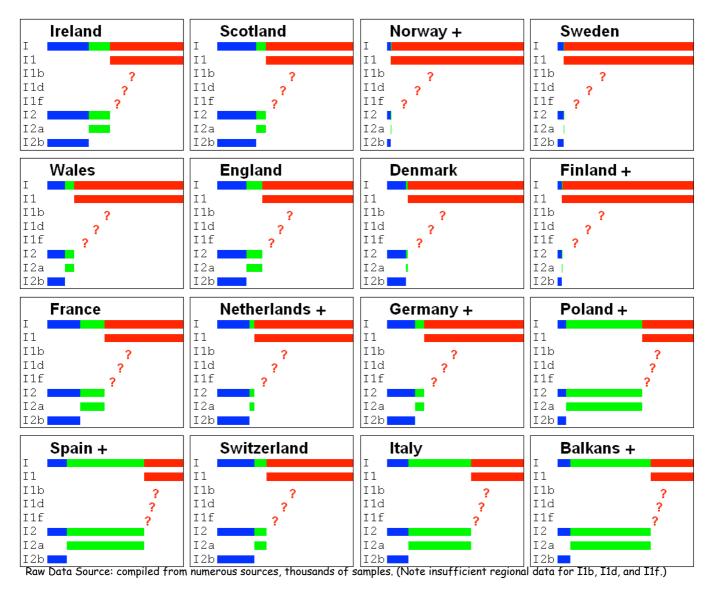
Norway+ = Norway & Iceland Finland+ = Finland & Estonia

Netherlands+ = Netherlands & Belgium Germany+ = Germany & Austria & Czech Republic

Poland+ = Poland & Lithuania Spain+ = Spain & Portugal

Balkans+ = (Albania, Bosnia-Herzegovina, Bulgaria, Croatia, Greece, Macedonia, Montenegro, Serbia, Slovenia) & Hungary, Moldova, Romania, Turkey

European I Sub-clade Frequency by Region¹



I-M253 (also known as I1), shown as [Red], is the dominant sub-clade of y-Haplogroup I in Scandinavia. That sub-clade is very prominent in Northern Europe, and there is an increasing frequency from south to north.

I-P37.2 (also known as I2a), shown as [Green], is the dominant sub-clade of y-Haplogroup I in Poland, the Balkan states, Italy, and Spain.

I-M436 (also known as I2b), shown as [Blue], is a somewhat prominent sub-clade of y-Haplogroup I in Northern Europe excluding Scandinavia. For I2b there is an increasing frequency from eastern to western Europe.

```
M170
                                                                           ISOGG-2011 I
                                                                           ISOGG-2011 I1*
ISOGG-2011 I1b
ISOGG-2011 I1d
       +- M253
                                                             M227
                                                             L22
                                                                                          I1d
                                                                           ISOGG-2011
                                                             L338
                                                                                          I1f
                                                                           ISOGG-2011
ISOGG-2011
      +- M438
                                                                                          12*
                                                              P37.2
                                                                                                  ---- M423 = I2a2
                                                                           ISOGG-2011 I2b
                                                                                                 ---- M223 =
```

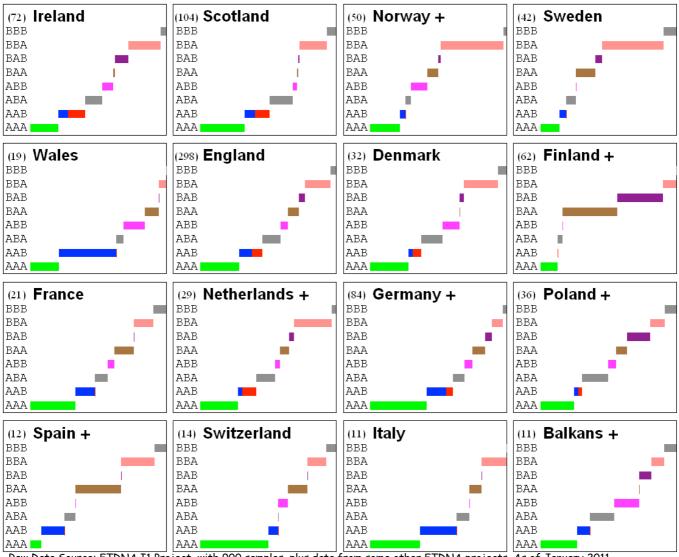
Norway+ = Norway & Iceland Finland+ = Finland & Estonia

Netherlands+ = Netherlands & Belgium Germany+ = Germany & Austria & Czech Republic

Poland+ = Poland & Lithuania Spain+ = Spain & Portugal

Balkans+ = (Albania, Bosnia-Herzegovina, Bulgaria, Croatia, Greece, Macedonia, Montenegro, Serbia, Slovenia) & Hungary, Moldova, Romania, Turkey

European I1 STR Cluster/Clan Frequency by Region¹



Raw Data Source: FTDNA II Project, with 900 samples, plus data from some other FTDNA projects. As of January 2011.

II-"BBA", shown as [Pink], is the dominant II STR cluster/clan in Norway and Sweden. Nearly 50% of Norway and Sweden's I1 population is I1-"BBA". It is also prevalent in other countries such as the Netherlands/Belgium and Denmark. Many I1-"BBA" and virtually all I1-"BBB" people have the L22 SNP mutation for I1d, which would be consistent with the L22 mutation originally happening in a "BB" individual before the split into "BBA" and "BBB".

II-"BAA" and II-"BAB", shown as [Brown] and [Purple], are the dominant II STR clusters/clans in Finland and account for about 75% of Finland's I1 population. Those two clusters/clans are associated with the L258 SNP mutation for I1d3a. Due to a back-mutation in DYS511 they could otherwise have been classified as I1-"BBA". Within Finland, I1-"BAA" is more likely in western areas, and I1-"BAB" is more likely in eastern areas.

II-"AABB", shown as [Red], is relatively high in Ireland, Scotland, England, and the Netherlands/Belgium. That cluster/clan is associated with the L338 SNP mutation for I1f. So far, 16 out of 16 known L338+ people would be classified as I1-"AABB". Up to 10% of I1 people may have that mutation.

II-"AABA", shown as [Blue], is particularly high in Wales - but the sample size is low. So far, 8 out of 19 Welshancestry I1 people, or ~40%, would be classified as I1-"AABA", whereas the European-wide rate is less than 10%.

II-"AAA", shown as [Green], is the dominant II STR cluster/clan in Europe as a whole. About 25% of European II people would be classified as II-"AAA". There is a slight decrease in frequency from southern to northern Europe.

See http://www.goggo.com/terry/HaplogroupI1/ for the I1 STR Decision Tree showing the various I1 STR clusters/clans.

Norway+ = Norway & Iceland Finland+ = Finland & Estonia

Netherlands+ = Netherlands & Belgium Germany+ = Germany & Austria & Czech Republic

Poland+ = Poland & Lithuania Spain+ = Spain & Portugal

Balkans+ = (Albania, Bosnia-Herzegovina, Bulgaria, Croatia, Greece, Macedonia, Montenegro, Serbia, Slovenia) & Hungary, Moldova, Romania, Turkey

Page 7 TDR, February 2011

European I1 STR Cluster/Clan Sample Size by Region¹

Region	AAA	AAB A,B	ABA	ABB	BAA	BAB	BBA	BBB	I1 all
Ireland	15	5, 9	9	6	1	7	17	3	72
Scotland	34	8, 11	18	3	1	1	21	7	104
Wales	4	8,0	1	3	2	0	1	0	19
England	85	23, 23	40	16	24	14	55	13	298
Norway+	11	2,0	2	6	4	1	23	1	50
Sweden	6	2,0	3	0	6	2	19	4	42
Denmark	9	1, 2	5	4	0	1	8	2	32
Finland+	8	0,0	2	0	25	21	6	0	62
France	7	3,0	2	1	3	0	3	2	21
Netherlands+	8	1, 3	4	1	2	1	8	1	29
Spain+	1	2,0	1	0	4	0	3	1	12
Switzerland	7	1, 0	0	1	2	0	2	1	14
Germany+	35	12, 4	7	5	8	4	7	2	84
Poland+	9	1, 1	7	2	3	6	4	3	36
Italy	4	3,0	1	0	1	0	2	0	11
Balkans+	3	1, 0	2	2	0	1	1	1	11
Europe	246	78, 53	104	50	86	59	180	41	897

The above sample-size counts for each of the I1 STR clusters/clans, is based on data, as of January 2011, collected for the FTDNA I1 Project. Only participants who had done the 67-marker test and who specified their male-line ancestral origin, were included in the above counts.

Alternatively, one could include data from ySearch.org. The sample-size counts (for each I1 STR cluster/clan) obtained from ySearch.org, as of January 2011, are as follows: 502 (AAA), 123,125 (AABA,AABB), 219 (ABA), 81 (ABB), 137 (BAA), 99 (BAB), 88 (BAB), 345 (BBA), 74 (BBB), for a total of 1793 (I1 all). About half of those samples have useable ancestral origin data.

Norway+ = Norway & Iceland Finland+ = Finland & Estonia

Netherlands+ = Netherlands & Belgium Germany+ = Germany & Austria & Czech Republic

Poland+ = Poland & Lithuania Spain+ = Spain & Portugal

Balkans+ = (Albania, Bosnia-Herzegovina, Bulgaria, Croatia, Greece, Macedonia, Montenegro, Serbia, Slovenia) & Hungary, Moldova, Romania, Turkey

European I1 STR Cluster/Clan MRCA1

Time Period	16,000 BC to 13,000 BC	13,000 BC to 10,000 BC	10,000 BC to 7,000 BC	7,000 BC to 4,000 BC	4,000 BC to 1,000 BC	1,000 BC to 2000 AD
Europe Events	Expansion out of Refugia after Last Glacial Maximum at 16,000 BC	End of Last Glacial Period by 10,000 BC	Sea Levels Rise 120m isolating Britain from Continent by 7,000 BC	Expansion of Farming from Near East to all of Europe by 4,000 BC	Expansion of Indo-Europeans out of Ukraine into Europe by 1,000 BC	Migration of "Celts", "Germanics", "Barbarians", "Vikings"
SNP Events	R1b-P25 E1b-M78 J2-M172	R1a-M17 I2b-M436 I2b-M223 N1c-P105	R1a-M458 R1b-M269 R1b-L23 G2a-P15	R1b-M412 R1b-L11 I1-M253 I2a-P37.2 I2a-M423	R1b-S116 R1b-U152 R1b-M529 I1-L22	R1b-U106 R1b-M222

The above timeline shows approximately when some SNP mutations occurred in the major European y-Haplogroups, and places them into historical context. Note, some authors have computed different SNP event dates to these.

Two methods for computing the TMRCA of an STR dataset are the "Walsh Method" and the "Variance Method". Both methods assume an idealized situation with up/down STR-mutation symmetry. The 67-marker mutation rates are taken from Chandler for the first 37 markers, and estimates used for the remaining markers; multi-copy markers are excluded. An inter-generation time of 30 years is used. Different assumptions will give different results, so the results below are for the specified assumptions and applied to the FTDNA I1 Project data.

	MR<i>CA</i> ("Walsh Method")	MRCA ("Variance Method")
I1-"BBB"	2,500 BC ± 1,200 years	?,?00 BC
I1-"BBA"	3,400 BC ± 1,400 years	?,?00 BC
I1-"BAB"	3,500 BC ± 1,400 years	?,?00 BC
I1-"BAA"	3,200 BC ± 1,400 years	?,?00 BC
I1-"ABB"	3,400 BC ± 1,400 years	?,?00 BC
I1-"ABA"	3,400 BC ± 1,400 years	?,?00 BC
I1-"AABB"	1,900 BC ± 1,000 years	?,?00 BC
I1-"AABA"	3,700 BC ± 1,400 years	?,?00 BC
I1-"AAA"	4,400 BC ± 1,600 years	?,?00 BC
T1 M2E2	E E00 DC - 2 000	2 200 P.C
I1-M253	5,500 BC ± 2,000 years	?,?00 B <i>C</i>

The "Walsh Method" returns a probability distribution for the TMRCA, and that distribution has been simplified in the table to just a central date with an error in years (corresponding roughly to one standard deviation). Model assumptions are critical to the computed date of the MRCA. The "Variance Method" gives much younger dates.

 $^{^{1}}$ MRCA is the most-recent-common-ancestor, and TMRCA is the time-to-most-recent-common-ancestor.

² The "Walsh Method" (Walsh, 2001), assumes the Stepwise Mutation Model, and computes the TMRCA between a pair of haplotypes. FTDNA uses this method in their FTDNATiP™ calculator. The MRCA of a large population, requires finding the TMRCA of pairs of haplotypes from different branches.

³ The "Variance Method" (Slatkin, 1995; Stumpf, 2001) assumes that the variance (average-squared-distance from ancestral value) of each STR marker in a large population, is proportional to the TMRCA of that population. To get the proportionality relationship, one needs the mutation rates either an average rate for all STR markers, or separate rates for each marker. Combining all the marker results will give the TMRCA for this method.

References

Kinder et al (1974) - "The Penguin Atlas of World History: Volume 1", revised edition 2004

Slatkin (1995) - "A Measure of Population Subdivision Based on Microsatellite Allele Frequencies", Genetics (1995) 139:457-462

Stumpf et al (2001) - "Genealogical and Evolutionary Inference with the Human Y Chromosome", Science (2001) 291:1738-1742

Walsh (2001) - "Estimating the time to the most recent common ancestor for the Y chromosome or mitochondrial DNA for a pair of individuals", Genetics (2001) 158 (2): 897-912

Rootsi et al (2004) - "Phylogeography of Y-Chromosome Haplogroup I Reveals Distinct Domains of Prehistoric Gene Flow in Europe", American Journal of Human Genetics (2004) 75:000-000

Chandler (2006) - "Estimating Per-Locus Mutation Rates", Journal of Genetic Genealogy (2006) 2:27-33

Underhill (2007) - "New Phylogenetic Relationships for Y-chromosome Haplogroup I: Reappraising its Phylogeography and Prehistory", in Rethinking the Human Evolution, pp. 33-42

Lappalainen et al (2008) - "Migration Waves to the Baltic Sea Region", Annals of Human Genetics (2008) 72:337-348

Karafet et al (2008) - "New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree", Genome Research, published online April 2, 2008

Wiik (2008) - "Where Did European Men Come From?", Journal of Genetic Genealogy (2008) 4:35-85

Chiaroni et al (2009) - "Y chromosome diversity, human expansion, drift, and cultural evolution", PNAS (2009) vol. 106 no. 48 20174-20179

Balaresque et al (2010) - "A Predominantly Neolithic Origin for European Paternal Lineages", PLoS Biology (2010) 8: e1000285

Underhill et al (2010) - "Separating the post-Glacial co-ancestry of European and Asian Y chromosomes within haplogroup R1a", European Journal of Human Genetics (2010) 18:479-484

Burgarella et al (2011) - "Mutation rate estimates for 110 Y-chromosome STRs combining population and father-son pair data", European Journal of Human Genetics (2011) 19:70-75

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