

Y-Chromosome Marker S28 / U152

Haplogroup R-U152

Resource Page

David K. Faux

To use this page as a resource tool, click on the blue highlighted words, which will take the reader to a relevant link.

How was this marker discovered? In 2005 Hinds et al. published a [paper](#) outlining the discovery of almost 1.6 million SNPs in 71 Americans by [Perlegen.com](#), and which were deposited in the online [dbSNP database](#). Gareth Henson noticed three SNPs that appeared to be associated with M269, what was then known as haplogroup R1b1c. Dr. James F. Wilson of [EthnoAncestry](#) developed primers for these Single Nucleotide Polymorphism (SNP) markers on the Y-chromosome, one of which was given the name of S28 (part of the S-series of SNPs developed by Dr. Wilson).

Who were the first to be identified with this SNP? In testing the DNA of a number of R-M269 males (customers or officers of EthnoAncestry), two were found to be positive for S28 (U152). These were [Charles Kerchner](#) (of German descent) and [David K. Faux](#), co-founder of EthnoAncestry (of English descent).

How is this marker classified? In 2006 the International Society of Genetic Genealogists (ISOGG) developed a [phylogenetic tree](#) since the academic grouping (the Y Chromosome Consortium – [YCC](#)) set up to do this task had lapsed in 2002. They determined, with the assistance of Dr. Wilson, that the proper placement would be R1b1c10, in other words downstream of M269 the defining marker for R1b1c. Karafet et al. (2008) (including Dr. Michael Hammer of the original YCC group) published a new [phylogenetic tree](#) in the journal *Genome Research*. The designation for S28 / U152 became R1b1b2h*. Recently a group of genetic genealogists, using findings from the deCODEme testing, discovered a SNP between M269 and S28 / U152, called S116 by EthnoAncestry, and P312 by the Hammer group (YCC), which caused ISOGG to revise their [phylogenetic tree](#). The most up to date classification for the YCC is Karafet et al. (2008) [Errata](#). Presently [Family Tree DNA](#) has termed U152+ as R1b1b2a1b4, if L2/S139+ add a “c” and if also L20/S144+ add a “1”. [ISOGG](#) (2009) has the category as R1b1b2a1a2d, and with L2/S139+ add a “3” and if L20/S144+ add an “a”. Perhaps the most sensible or parsimonious designation is R-U152 for most purposes.

What is the general nature of the S28 / U152 mutation? S28 is a neutral transition mutation (it does not adversely affect the functioning of any known gene) where an ancestral Guanine was replaced by the derived Adenine at a single location on the Y-

chromosome. The event presumably happened only once in the history of mankind. Thus some thousands of years ago a spot mutation occurred during meiosis in an R-P312* father. The mutation was passed to his son who in turn passed it on to every one of his descendants in the direct male line. Thus all R-U152 males have the same ancestor in the Y-line, although the age of the mutation (when S28 / U152 first appeared) has not been determined.

How old is the mutation? There is no molecular clock which will provide an absolute dating such as dendochronology (tree ring dating). The date depends on a host of assumptions such as generation age, whether to consider back mutations, accepting a mutation rate and so on. Robert McGregor used the Zhivotovsky et al. (2006) approach and obtained a date for the Most Recent Common Ancestor of 11,400 years (plus or minus 1000 years) of all those in the author's database. Vince Vizachero has estimated the age of S28 as 6600 years (Neolithic), and Dr. Ken Nordtvedt has obtained a date of about 3700 years (Bronze Age). While due consideration must be given these estimates, it is highly likely that what is being detected is an expansion of this haplogroup, whereas its origin lies in the Mesolithic, about 10,000 years ago. Thus the present author is inclined to accept the McGregor estimate since it accords best with all other data sources. Ultimately only ancient DNA samples are going to settle the matter to everyone's satisfaction.

What are some of the technical aspects of the S28 / U152 SNP? The essential facts can best be expressed in chart form:

Marker	Rs #	Forward PCR primer (5'-3')	Reverse PCR primer (5'-3')	Product Size (bp)	SNP position (bp)
U152	rs1236440: G>A	cttagctatacagcctcttttgg	aacattccacgcttgaggataa	172	127

What does the G to A mutation of a male who is S28 / U152+ look like along the sequence of base pairs in the vicinity of the marker? This can best be viewed by a [pherogram](#) / chromatogram printout from the program Sequencher to see the difference between an ancestral and derived Y-chromosome at this location.

Where on the Y-chromosome is the S28 / U152 mutation? This marker can be viewed via the [Y-chromosome browser](#) created by Thomas Krahn.

Is there any way to relate S28 / U152 to the broad east – west classification system for R-M269? There appears to be an important division between the most common forms of R-M269 from Bulgaria eastward versus westward. An “old – fashioned” marker that is seldom used in studies today is the p49a,f Taq haplotype system. [Cinnioglu et al. \(2004\)](#) used this in describing their samples from locations such as Turkey and Armenia and those from Iberia and other locations in western Europe. The division is dramatic, Haplotype 35 is characteristic of most R-M269 east of the Balkans, whereas Haplotype 15 is most common in western Europe. Unfortunately no company at present will do commercial testing for this marker, but Vincent Vizachero has found a set of markers that can act as surrogates. The evidence is very clear, if someone has DYS393=12, DYS461=11, and DYF385=9,10 or 10,10 then it would appear that they are of the

“Eastern” Haplotype 35 variety. Similarly the motif characteristic of Haplotype 15 or “Western” variety is DYS393=13, DYS461=12, and DYF385=10,11. All R-U152 that have been tested for these markers fit into the latter pattern and hence are of the “Western” variety of R-M269 which likely expanded out of the Franco-Cantabrian refugium after the Last Glacial Maximum (whereas the “Eastern” type likely expanded from the West Asian area at this time).

How does the new SNP found via deCODEme testing, known as S116 / P312, relate to S28 / U152? Recent testing by [Thomas Krahn](#) at FTDNA has revealed that that R-U152, R-M153, R-M167, R-M222 and some R-M269* test positive for S116 / P312, meaning they have the derived A allele; whereas those who are R-U106 plus all Eastern R-M269* should have the ancestral C allele on this marker. This will perhaps link in a broad sweeping way those whose ancestors belonged to a Celtic culture (Insular or Continental) who will all share the S116 / P312+ polymorphism, dividing them from the S21 / R-U106 “Germanic” group as well as the large group of R-M269* “Eastern” found from Bulgaria to Kazakhstan and Anatolia.

Why is S28 also known by another name – U152? In 2007 Sims et al. published a [paper](#) in *Human Mutation* which identified the markers developed by Dr. Wilson, but using new names. Hence S28 is comparable to U152 in their paper.

Which companies test S28 / U152? From April of 2005 until February 2008 only EthnoAncestry offered this SNP. Presently [Family Tree DNA](#) tests the markers developed by Dr. Wilson, which they call the “U-series” or the “Garvey Panel”, which includes what they termed U152 (using the Sims et al. 2007 classification). This testing is only available to those who have tested at least 12 Y-STR markers (the standard Y-chromosome test in genetic genealogy) with FTDNA. Presently [deCODEme](#) and [23andME](#) also test this marker and the latter also tests for L2/S139 and L20/S144 (see later).

What do the certificates look like? Click for examples of [EthnoAncestry’s](#) and [FTDNA’s](#) certificates attesting to a positive finding for S28 and U152 respectively.

What databases are available for S28 / U152? Soon after S28 was identified [David K. Faux](#) began collecting haplotypes and genealogical information (particularly the place of birth of the earliest known ancestor in the Y-line). He also included research testing of his Shetland Islands samples, as well as some as yet unpublished data from academic research. Data is screened to ensure that, particularly in the case of those who are of Colonial American ancestry in the Y-line, that they have a solid paper trail lineage back to a European country. The data includes those tested at EthnoAncestry and FTDNA (and any other company who will offer this testing). In March 2008 it was announced that [Charles Kerchner](#) would be administering a FTDNA sponsored R-U152 project. It is only open to those who have SNP tested U152+ at FTDNA. [Adriano Squecco](#) is collecting all of the raw data from individuals who test with deCODEme and 23andME (Versions 1 and 2) and has presented in a spreadsheet database. In addition, [Ysearch](#) includes those who are identified as S28 or U152 and chose to include their data in this database. At

present the database appears to be in the process of shifting from the R1b1b2h to the current FTDNA designation, but much data seems to have vanished in the process (hopefully this will be corrected shortly). Another database is that of [John McEwan](#) who for many months collected data from those who tested positive on one of the subclade tests in relation to R-M269. There is a lot of other material here that would be of interest to those who are U152+. Unfortunately the database has not been updated since July 2007.

[Is there a website devoted to S28 / U152?](#) [Jacques Beaugrand](#) has established a website for this haplogroup, which includes links to many items of interest to those who fall into the R-U152 haplogroup.

[Can the relationship between individuals who are identified as U152 positive be seen in a graphic form?](#) [Dale Bricker](#) has constructed a network diagram to graphically portray the relationships between U152+ individuals. What is interesting is that the closest neighbors of someone from say England may be Italy or Switzerland. However a definite clustering is seen for those in the DYS492=14 category (likely L2 ancestral, described later), which appears to include 1B07=8 also. The larger grouping with DYS492=12 (likely L2 derived) is clustered into 6 groupings with the important markers noted above, as well as DYS456 which appears to have structure in these classifications. [Beau Gunderson](#) has created a pdf contour map (preliminary, first draft) which clearly shows the apparent point of origin of R-U152 (shown as R1b), but which at this time shows the biases in the data with more from Britain and Ireland as well as southern Italy, and less from locations such as France, Spain, and Eastern Europe,

[Gary Felix](#) has established a set of slow moving Y-STR markers and has a table showing how far each haplotype diverges from the modal.

[Is there a map showing the geographical distribution pattern of S28 / U152?](#) [Jacques Beaugrand](#) has drawn a number of maps to illustrate where S28+ are found, which allows viewers to see patterns and overlap with historical and archaeological cultures. [Vince Vizachero](#) has placed pins on a Google map of Italy showing the location of each of the R-M269 SNP tested individuals in the Italy Geographic Project. With the assistance of John Laws and Robert Tarin the author has learned to use Google Maps to depict aspects of the distribution of R-U152 across Eurasia. This resulted in the following maps:

- 1) [Google Map](#) showing all who know their specific village of earliest origin in Continental Europe.
- 2) It has been necessary to construct a separate [Google Map](#) to show “Insular R-U152” (British Isles) since the original map has become so large that Google split it into two pages (all new entries going to page 2). This is less than acceptable so the division of Eurasia makes more sense. The British Isles are highly over represented in all Y-DNA databases, thus what appears to be a large number of R-U152 there is an illusion – it is a relatively rare haplogroup in that region.
- 3) A third [Google Map](#) shows the distribution of those who are L20+.

- 4) The next set of maps can be used as overlays by simultaneously opening one or more of the above and any of the following:
- a) A [Google Map](#) which shows the boundaries of the Continental Celtic world as depicted in Koch (2007). This haplogroup downstream of R-U152 is found in Central France, the Italian Lake District, Switzerland, Southern Germany, and Belgium. It is found in the British Isles but only in England to date, and only along the eastern coastal margins from Lincolnshire to Kent. A more sophisticated map in KML format can be seen [here](#).
 - b) A [Google Map](#) illustrating the regions where tribes from the Continent are known to have settled circa 100 AD, there being little evidence of any post – Neolithic migration of any significance before this date.
 - c) A [Google Map](#) showing the extent of the Angle Kingdoms and presumably early Angle settlement from Jutland; as well as the region where the Jutes also from that area are known to have migrated.
 - d) A [Google Map](#) portraying the extent of the Viking settlements in Britain and Ireland based on historical, archaeological, and linguistic (place name) evidence.

The reader will then be able to come to their own conclusions in relation to whether the distribution of R-U152 follows closely the known Celtic world; and which source or sources are responsible for the presence of R-U152 in the British Isles.

How common is S28 / U152? The percentage of R-U152 in Europe or the regions to which Europeans expanded (e.g., USA) is not clear at present. What seems apparent from the database noted above is that there are “hotspots” and so that for example in Switzerland S28 / U152 is going to be relatively common if not the predominant haplogroup. However, despite heavy sampling, to date no one with an aboriginal surname from Ireland (or the western regions of Scotland) has been assigned to the R-U152 category. Hence, its frequency will vary dramatically from country to country. In the McEwan study above (October 2006) he observed that at that time, of the 47 people from Ireland tested (R-M269 being higher in Ireland than anywhere else in Europe), none were R-U152; but 5% of the 21 from Scotland, 12% of 30 individuals from England, and 37% of 12 people from the Continent did test R-U152.

Are there any published studies using S28 / U152? In the Simms et al. (2007) study above using a European American sample from the USA (largely from the states of Virginia and South Dakota), 46% were some variety of R-M269, however only 3% were R-U152 (or 7% of the R-M269 population). The most recent study is that of [Niederstatter et al. \(2008\)](#) which examined the percentages of U106 and U152 in the M269 samples from Innsbruck, Austria. In that region the overall numbers of R-M269 were relatively low at 31% (N=42) of the sample of 135 males. About 62% of the M269 were S21 / U106, 21% were S28 / U152, and 19% were “unresolved” R-M269 (neither of the former two). Hence the fraction of both haplogroups was considerably higher in the Austrian (Central European) sample than the European American (probably Western European origin) sample of Simms et al. (2007).

[Are there any subgroupings of S28 / U152?](#) In the most recent revision of the Chromosome Consortium, R-U152 is being placed as R1b1b2a2g on the new [phylogenetic tree](#). What follows is a listing of all known SNPs downstream of U152.

1) Apparent Private or “Family” SNPs:

- a) Two subclades have been identified in the new tree, which were thought at one time to be phylogenetically equivalent to U152 (in other words sister clades to R-U152). These are M126-R1b1b2h1, and M160-R1b1b2h2. There have been a number of academic studies using these latter two markers, and they have been tested extensively in the commercial sphere but have not been observed beyond the original study in which they were published. Underhill et al. (2000) tested a sample of 60 “Europeans” and found 1 who was M126+, and 3 who were M160+. In closely examining the data, it appears that all of these individuals were Italians. It would, however, appear that these two markers may in fact be “private SNPs” and unlikely to be seen again except in the same area or with the same families tested in the original study.
- b) An individual from the Shetland Islands was found to have “private” or “family” SNP which EthnoAncestry termed S44. It is unknown at this time how widespread this mutation is since it has been tested in very few individuals.
- c) It has been reported that the individual whose ancestors are from Latvia (Jewish heritage) also has a private SNP as determined by Thomas Krahn at FTDNA who has labeled this SNP L4. Apparently at this time it is unknown whether this may be a marker unique to this man or this family, or found more widespread but appears to be associated with DYS492=14..
- d) Thomas Krahn at FTDNA also identified another SNP in the flanking regions of the amplicon (primer of circa 200 base pairs) used to detect the U152 SNP. This individual is from Lincolnshire, England. Considering the number of people tested for U152 at this point, it would appear that L3 may be a private SNP.
- e) In the process of testing four men with the deCODEme test, one, of Northern Italian ancestry (U152+, L2+, L20+) is M228+. This SNP has been reported in African Pygmies with haplogroup B, so this is some sort of parallel mutational event. It is probably private but only further testing can ascertain the frequency.
- f) Sims and Ballantyne (2008), while studying the Y-SNP M222, observed that one individual who was U152+ had a SNP situated close to the M222 marker. Sixteen other individuals who were U152+ were tested for this SNP, but were ancestral. Hence although the authors have placed the SNP in the phylogenetic category of R-U152, this may be premature since only 6.3% of their sample of R-U152 have this SNP – one person of a total of 16 – which suggests that the SNP may be in the “private” or “family” category.

2) Possible SNPs of Wider Phylogenetic Significance:

The first category of U152 males are those who test negative for all non private downstream SNPs and so are R-U152*. This group can be parsed into two groups by the STR marker DYS492. As a rule, those who are ancestral U152 will either have 12 or 14 repeats. The statistics are as follows for **U152 ancestral (U152*)**:

DYS492=14; 19

DYS492=12; 7

It is likely that as more individuals test for L2 (and don't assume that DYS492=12 will be L2+), the numbers with 12 and 14 repeats could emerge as almost identical.

Of the three newly discovered SNPs being offered by FTDNA for those who are U152+, one is **L2 / S139 or rs#2566671** which was discovered via 23ANDme testing of a U152 individual. Apparently FTDNA testing has found a mixture of individuals who are derived and ancestral on this new SNP, and hence it holds great promise in parsing the U152 haplogroup into meaningful subgroups – although the author believed that it would simply reflect the already established division between those who are DYS492=12 and 14 (see below), this is incorrect. To date three individuals who are DYS492=12 and one who is 13 have tested ancestral and thus R-U152*.

The most recent update from FTDNA is that of the 14 confirmed R-U152 tested:

L2 negative = 36%

L2 positive = 64%

More recent stats from the author's database give:

L2 negative = 28%

L2 positive = 72%

The first person to be identified as being L2 / S139+ (via 23andME testing) was Luc VanBraekel (of Belgian descent). Family Tree DNA then began testing this marker and a group of us (including DKF) were in the next "batch" to learn that they were positive on this marker.

Thanks to 23andMe testing another SNP within (downstream of) L2 / S139 has been found. It is **L20 / S144 or rs#7063305**. To date of about 15 individuals who tested with 23andME, there are nine who are negative, two thirds L2/S139+, and one third are U152 ancestral. Of those who have tested this marker via 23andme, FTDNA or EA, the following have been found to be L20/S144+, individuals whose ancestry is from: six of East Anglia, England ancestry (including the present author); one southern Germany; one from Switzerland; one Italian whose direct paternal line ancestors are from Civenna near

Switzerland, Lombardy, Italy; one from near Bourges in Central France; plus an individual whose ancestry cannot be proven to any European country are the only two who have tested positive. Of 27 who have tested both L2 and L20, 7 are L20+ or about 25% of the L2+ total.

David K. Faux was the first to be designated as having a mutation at this location and Giuseppe Belgieri (of Italian Alpine descent) was the second – both via 23andME testing.

EA is offering both of these markers via their “Custom Select” (S139 and S144); and FTDNA are now offering both within their “Advanced Menu”.

A [chart](#) showing the relationship of all of the known SNPs below M269 has been created by Thomas Krahn. It illustrates how L2 and L20 as well as the private SNPs fit within the phylogenetic categorization.

Dale Bricker has made a [prediction](#), largely hinging on whether the value of DYS456 is 15 (L20 / S144 positive) or 16,17 (L20 / S144 negative); and in relation of the marker 1BOF7 (8 repeats for U152* and 10 repeats for L2+). It will be a while before we are able to assess the validity of this interesting hypothesis, largely due to few having ordered 1BOF7.

When the results of more individuals who are R-U152 are available via the deCODEme and 23ANDme testing, it is possible that further subgroups will become apparent when individuals are compared on SNPs which fall within the broad R-M269 category.

There is at present almost no difference in the modal scores of 67 marker Y-STR (short tandem repeat – the typical markers used in genetic genealogy) haplotypes between R-M269* (“Western Variety”), R-P312*, and the non-Frisian grouping of R-U106. However, it appears that within R-U152 there is a bimodal distribution for one marker, DYS492 with 12 predominating (and being the ancestral variety in most other R-M269 clades), but with about 10% having a score of 14 on this allele, but one third among those with ancestry in Germany. It is anticipated that at some point the 12 / 14 split will show a more specific meaningful geographic distribution, but at present nothing can be offered beyond limited speculation.

Has anything been written about what the data appears to tell in the way of geographic distribution patterns or a connection with any cultural groups?

In 2007 David K. Faux first wrote a two part manuscript (continually updated) purporting to tell the “[story of U152 as a La Tene marker](#)”. This study compares the archaeological, historical, and linguistic data with the genetic Y-chromosome data. There is a remarkable overlap between the regions known to have been aboriginal to the Hallstatt and La Tene Central European Celts, or areas where there is documented expansion (largely in the early 6th and 4th Centuries BC) which took peoples from the Alpine regions such as Switzerland and Austria (plus eastern Gaul, Southern Germany) as far west as Central Spain (possibly circa 7th Century), and north to and outpost in Jutland and nearby areas in Norway and Sweden, scattered as far east as Anatolia and Western Ukraine (by the 3rd Century BC),

and Britain (between the 3rd and 1st Centuries BC). A recent finding of R-U152,L2+ in a Kipchak tribesman from Kazakhstan may relate to the incursions of this group into Ukraine and Hungary in the Middle Ages and eastward migration when the empire collapsed. [A recent publication](#) by Martin Ballauf is even more extensive, being 417 pages in length, in German (supplemented with many maps and charts), focusing on the link between the Alpine cultures and S28 / U152. Three studies by David K. Faux relate specifically to Britain. The first explores the possible connection between the [Cimbri tribes of Denmark, the Danelaw, and R-U152](#). A second focuses on the [Angle tribe of Schleswig, the Angle Kingdoms of England, and R-U152](#). Another referring specifically to the possibility of a connection between the [Belgae tribes of the Continent, the southern part of England, and R-U152](#). These could for example be the Belgae who settled in England before the time of Caesar, the Normans, or the Flemings.

David K. Faux, Ph.D.

R-U152, S139 / L2+, S144 / L20+ (S28 as tested by EthnoAncestry in 2005; and U152 as tested by FTDNA and deCODEme, as well as L2 and L20+ by FTDNA in 2008, and all three tested via 23andMe in 2008).

Seal Beach, California and Caledonia, Ontario
fauxdk [at] yahoo.com

Version: 12 December 2009
Copyright David K. Faux, 2008-2009