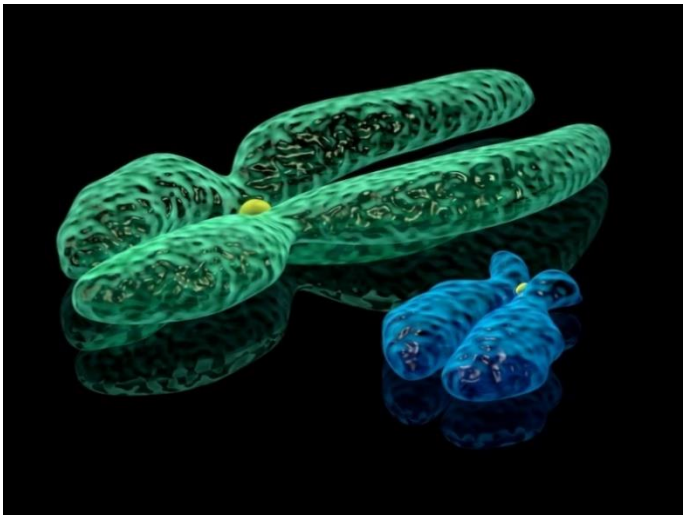
FAUX Y-DNA TESTING

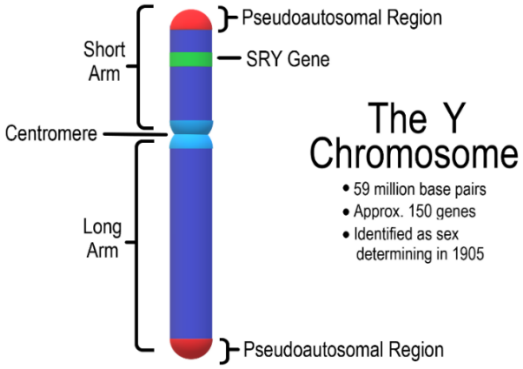
The Y Chromosome

In addition to 22 pairs of autosomes (which include most of the human DNA and coding genes), there are two sex chromosomes – the X and the Y. In general females of the species have two X chromosomes, and males have an X and a Y chromosome – the Y being inherited from the father and his father and so on in the direct paternal line back to the beginnings of the species.

The Y is among the smallest chromosome and includes about 63 genes along the 57,227,415 base pairs of the building blocks of DNA, adenine, guanine, cytosine, and thymine (AGCT) along the double helical shaped strands. One gene, the SRY, codes for "maleness". If born with this gene then generally one becomes a biological male (e.g., development of testis), otherwise an individual is born who is typically female (always a few exceptions – which are not germane to the discussion here).



X chromosome on upper left, Y chromosome on right



A good review paper on the structure and function of the Y can be found here.

Y-STRs and Haplotype Testing

In the early years of commercial genetic genealogy testing, when the goal was typically to determine whether two men were biologically related in the Y-chromosome paternal line, what was measured in Y-chromosome testing is a series of 12 to 67 markers (111 marker tests are available), called short tandem repeats (STRs), situated along the length of the chromosome providing a haplotype (genetic signature) for the person which should be almost identical to those who are related within a genealogical timeframe (e.g., past 500 to 1000 years). This DNA is essentially "junk DNA" in that it serves no known purpose, but is very useful for detecting similarities and differences between males. Basically the scores at all say 25 markers should be identical between a father and his biological son; and between individuals who are descended from a common ancestor in the last few hundred years (with occasionally up to three mutations in 25 markers). Most people are interested in the period since surnames were adopted - about 1200 AD. In that time frame of 800 or so years, if two individuals have the same surname but a very different patterns of scores (e.g., only 15 of the 25 STR markers match) they are not likely related through the male line. If, however, there are only four differences in the scores between two men with the same surname, it is probable that they had a common ancestor – IF it can be shown that they belong to the same haplogroup (more on this later).

1) Evidence from Various Y-STR Databases: There are many databases available to assist in interpreting haplotypes. Most, unfortunately, suffer from serious flaws that make them problematic for determining ancestry pre 1200 AD (approximately when surnames were adopted) - unless one has a very rare R1b haplotype – and databases using less than 25 markers are virtually useless due to the danger of matching coincidentally. Thus, one might have a 12/12 match that is not even the same haplogroup, meaning that the “match” is an illusion and you are only “identical by state” and not related within the past few thousand years.

The author’s father, his second cousin, a third cousin, and two 8th cousins have been tested using 37 markers. While there is a definite “Faux signature” on markers which rarely mutate, there are a number of false close matches at 25 markers, but within 5 mutations of 37 markers only other known Faux cousins have closely matching signatures.

The Faux Surname DNA Study i(somewhat out of date) is found here, and at the FTDNA Project site here. In comparing the results here to any public Y-STR database, the result tends to be much the same – the Faux signature is quite unique. There were a number of database options when the first version of this manuscript was written, but these latter now of historical interest only. Useful testing Y-STRs is currently available only through Family Tree DNA, With 100,000 or more haplotypes available, comparisons here may be promising, however at present there are no high resolution matches to the Faux haplotype in any of these databases. What emerges depends on the markers chosen, such that one could get the results desired simply by selecting a particular subset of the markers.

Using the databases above, upon entering the markers which seem most characteristic of the Faux "signature" (27), including a particular marker DYS464 which has 4 components, there were a

handful of matches at 23/27. Of those with attached genealogies and a known place of origin (other than being "brick walled" in the USA) the closest were one from Switzerland (Bern) but it dead ended in a non - paternity event; and one from Norway. The latter was very interesting in that the person's earliest ancestor was from Vestfold in Southeastern Norway – which takes on more significance in discussing haplogroup findings below.

Considering the academic research findings (noted below) this finding may be of some significance. The author then queried the database using the 6 most common R1b markers from academic papers and entered the values for the Faux haplotype as well as the rare 14, 15, 17, 17 motif for marker DYS464. There were 60 exact matches. While there are double the number of German haplotypes in this database relative to Denmark, there were only 3 German matches and 4 from Denmark. The latter were all from northern Jutland (e.g., Aalborg). Although sample sizes are small here these results may be of some significance.

An academic population genetics study by Dupuy et al. (2005) provides 1766 Norwegian haplotypes in an article entitled, "Geographical heterogeneity of Y—chromosomal lineages in Norway". The country is divided into Oslo, Bergen, South, West, Middle, North, and East. The author's 10 marker haplotype is found from 0 to 2 times in all regions except East where in is found in 6 samples. Even taking into account sample size, there are proportionately more of the haplotype in the East (but none in Oslo) relative to other areas of Norway. The East takes in the Vik (Vestfold) which was, as with Skane in Sweden, essentially a part of Denmark for most of its history.

Despite some tantalizing possibilities, the bottom line is that this line of inquiry holds little promise UNLESS there is a haplogroup connection.

Those Y-STR Tested Include David K. Faux (and his father Kenneth C. Faux); Robert Faux who is a second cousin to Kenneth; Christopher Faux who is a third cousin to Kenneth; as well as Ivor Faux and David A. Faux who are seventh cousins to Kenneth Faux. The matches in the Family Tree DNA database are shown on the following documents. Also shown are the individual repeat values at each marker on the certificates also attached.

All of the above are descendants of Thomas Faux born 1643, or Gregory Faux Jr. born 1642 (Ivor Faux and David A. Faux), the only sons of Gregory Faux Sr. of Croxton, Norfolk.

It was not necessary to conduct haplogroup testing on more than one person in the group since all will match. The testing was done using David K. Faux as proxy for the rest.

The limitations of what one can learn in relation to the early ancestral origins of the Faux Y-chromosome can be seen in considering the above Y-STR testing. However since the first edition of this manuscript was written there has been a "revolution" in the sense of huge advancements on the technological side of things such that much clearer answers are available with the newer testing of SNPs – and seen below.

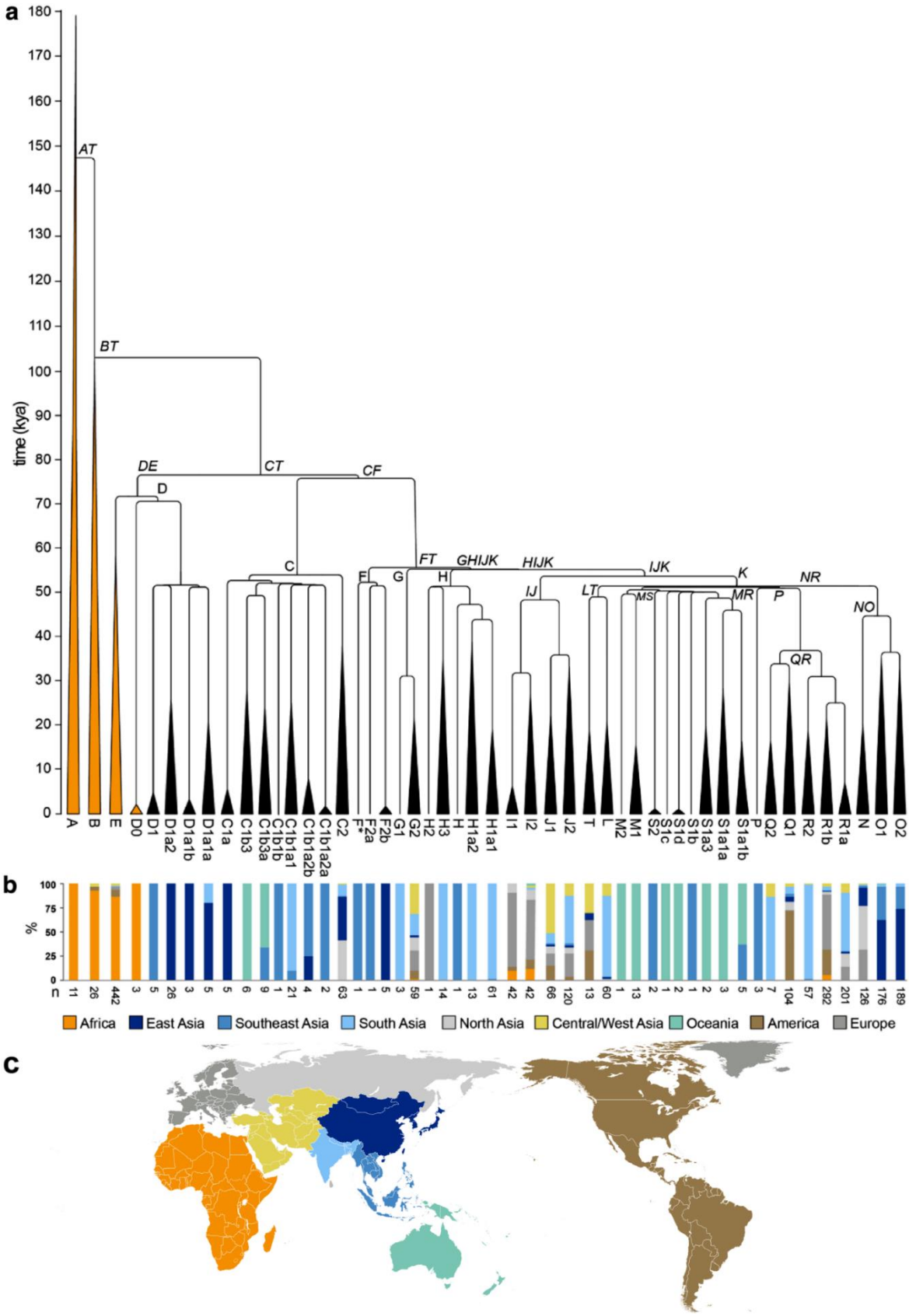
As noted above, and to repeat, in addition to its key role in human development, the Y also includes factors which can tell the story of a person's regional and ethnic ancestry, and their relationship to other males who may be related (e.g., whether they are likely say a third cousin, sharing the same great great grandparents). Commercial testing of STR (short tandem repeat) markers (e.g., a male has 25 ACCCTGGAT repeats at a marker named say DYS270) allow the comparison of the results for two males on for example 37 such markers. A 36/37 match would suggest a relatively close relationship, somewhere between first degree and say 8th degree relatives in the direct male line. While this testing can be informative within a genealogical time frame (e.g., since parish registers were first begun – typically in the 16th Century), it will falter in the attempt to explore distant ethnic ancestry (e.g., whether geographical origins are to be found in Switzerland or Ireland). This is where SNP (single nucleotide polymorphism) testing can offer potential answers.

Y-SNPs and Haplogroup Testing

All human males today descend from an African man who lived perhaps 200,000 years ago in East Africa and whose descendants spread to every corner of the globe. His haplogroup was named Haplogroup A (there are still Africans today who possess this haplogroup). At that point or soon thereafter humans began to migrate out of Africa and new haplogroups emerged when mutations occurred and that male had male descendants who carried the same mutation whose progeny are still here today.

Every 50 to 130 years (a wide estimate) a new mutation will appear on the Y chromosome of a male lineage. All descendants in the male line will carry that mutation (usually harmless). Thus, many thousands of years ago a male, likely living east of modern Turkey, developed a mutation that is today called M269 within the broader category of R1b (a distant descendant of the original group A). This mutation is currently found in over 33% of the males in Western Europe, and in locations such as Western Ireland it comprises over 90% of the male haplogroups.

The haplogroup distribution map below shows the world and regional distribution of haplogroup R1b (5th notation from right of the "a" part of the diagram). The percentage of R1b (the reason for using a series of letters and numbers to designate a haplogroup will be explained later) in Europe can be seen along the 5th bar from the right in "b". R1b is the basic grouping into which R-U152 and its "downstream" descendants such as R-L20 fall. After an extensive exploration of haplogroup R-L20, the author will describe the specific results of this form of testing which will apply to all who are Faux descendants in the male line from William Falke born circa 1385 in Worlingham, Suffolk, England – and the terminal SNP (which matches an ancient DNA sample).


a
time (kya)
AT
BT
DE
CT
CF
D
C
F
G
FT
GHIJK
HIJK
IJK
K
NR
H
IJ
LT
MS
MR
P
NO
QR
A B E D0 D1 D1a2 D1a1b D1a1a C1a C1b3 C1b3a C1b1b C1b1a1 C1b1a2b C1b1a2a C2 F* F2a F2b G1 G2 H2 H3 H H1a2 H1a1 I1 I2 J1 J2 T L M2 M1 S2 S1c S1d S1b S1a3 S1a1a S1a1b P Q2 Q1 R2 R1b R1a N O1 O2
b
%
n
11 26 442 3 5 26 3 5 5 6 9 1 21 4 2 63 1 1 1 5 3 59 1 14 1 13 61 42 42 66 120 13 60 1 13 2 1 2 1 2 3 5 3 7 104 57 282 201 126 176 189
Africa
East Asia
Southeast Asia
South Asia
North Asia
Central/West Asia
Oceania
America
Europe
c

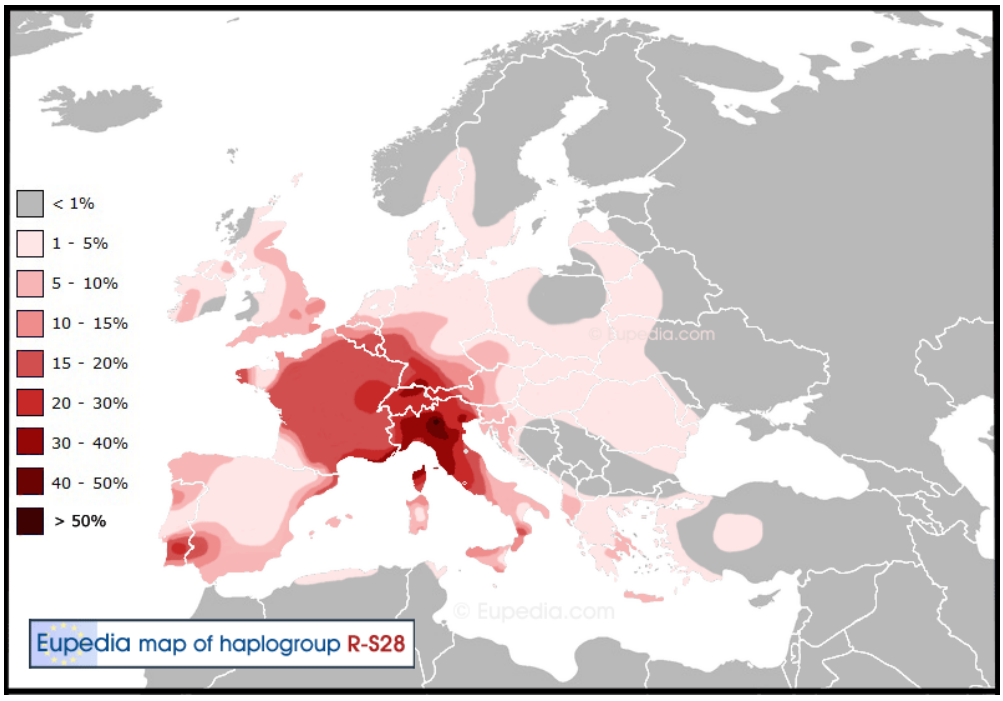
Haplogroup R-U152

Branchings of M269 occurred after about 3000 BCE/BC as a large wave of Indo-European speaking individuals made their way westward from the steepes between the northern Caspian and northern Black Seas and fanned out across Europe. Early in the Bronze Age (about 2000 BCE) we find descendant R-U106 concentrated in Northern Germany and Scandinavia; R-DF27 in Iberia; and R-U152 in the Alpine areas such as Switzerland and adjacent countries.

More detail about Y-DNA with a focus on ancestry testing can be found here.

First it is important to note that today many thousands of Y SNPs have been discovered, and each is placed into the "Y Tree" using a series of letters and numbers such as R1b. See here for the most up to date tree of the International Society of Genetic Genealogy.

Distribution of R-U152 Today: The focus in this study is on the descendants of R1b-M269 > U152, and how and when its descendant lines (subgroups) came to England and all the British Isles. As we will see later, only ancient DNA holds a clear answer to these questions.



An excellent overview of the history of R-U152 (or simply U152), which is also known as R-S28 (the designation used when the marker was first discovered) is found at Eupedia here, and in the diagram above which reflects the modern day distribution of the haplogroup.

As an aside, it should be noted that the author has chosen to use "U152" since it is the most common designation, however when first discovered it was known as S28 (see the "U152 and Subclades Project" for the history of the discovery, and the role of the current author). Some

authors, particularly in ancient DNA studies use "U152/S28" to designate the haplogroup. The "U152" designation will typically be used in the current study due to its prevalence in academic studies as well as the term used by FTDNA who currently lead the field in commercial Y-chromosome testing.

It appears that the most recent relevant academic study is that of Lucotte (2015). The general results are seen in the map below:

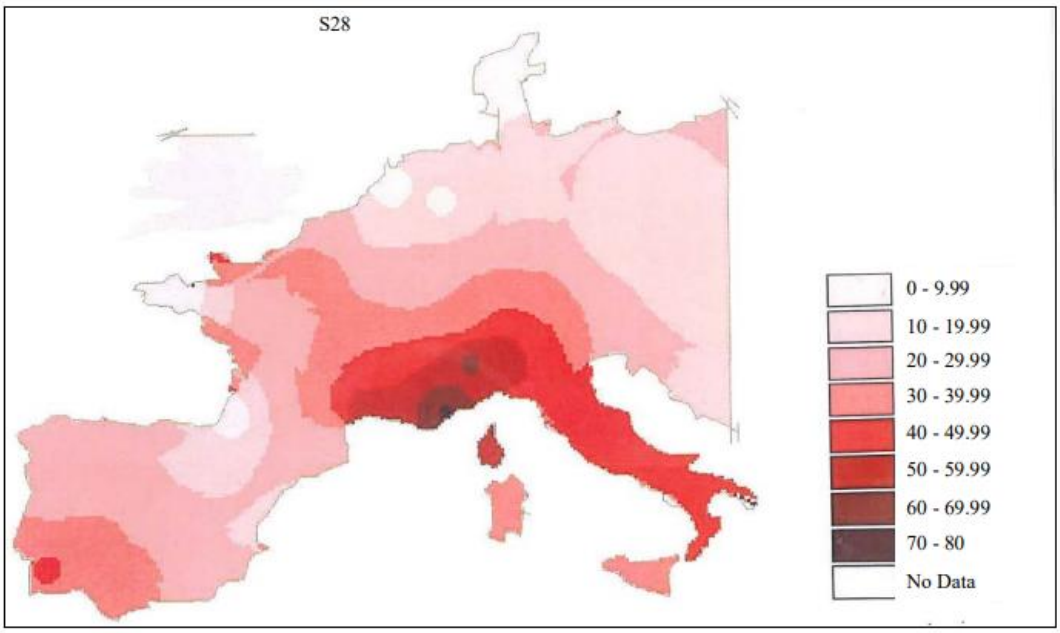


Figure 6. Isofrequency map of R-S28 in West-Europe; the various nuances of red depicts the structured pattern of variation for S28 frequencies.

Lucotte concluded that, "*the most elevated R-S28 values are concentrated at the interior of the Alps mountains in continental Italy*". Hotspots were southwest France and Northern Italy. Unfortunately, R-U158/S28 was not broken down into subclades – as in almost universally the case with all of the academic studies in the field. Published work can be found on the Family Tree DNA U152 and Subclades Project "Results" page here.

The most useful data has emerged from the Family Tree DNA (FTDNA) "Haplogroup U152 and Subclades" Project (see here). Thanks go out to Richard Rocca, Steve Gilbert and Tibor Feher who are administrators of this Project for the immense assistance that they have provided over the years.

Early in the study of this haplogroup the author kept a database, which included results from the testing of customers from companies including FTDNA and EthnoAncestry in a S28 / U152 database which was kept up to date until 2010. To that date the haplogroup has been seen in a band in the east from Greece through Poland, Ukraine, Turkey and Kazakhstan; southern and less so northern Germany, Italy, parts of France (mostly eastern), Benelux, Switzerland (the Alpine Region, ancestral home of the Celtic Helvetii, being a "hotspot") and west to the Bay of Biscay in

France and the Celtic speaking area of Iberia – all consistent with the range of the La Tene Celts of the Iron Age.

U152 in England: Soon after the discovery of S28/U152 by Dr. James F. Wilson, he came to term this an "eastern marker" in England. The primary locations where it was observed hugged the eastern, and to a lesser extent southern coastal areas of England. Reference to the above maps shows that the same general distribution pattern is seen today in the British Isles. Its presence in Ireland can most parsimoniously be explained as a function of the Middle Ages movement to that area by the Normans and later the English, as well as the Plantation Scottish in the north. Some contribution may also have occurred due to the earlier Viking settlements.

One fact is likely of some importance with respect to England, that being that a disproportionate number from here and their Colonial descendants have had commercial testing – it is a very popular hobby with this group. However we are less likely to have a clear picture of the pattern in for example countries in Eastern Europe, as well as France, at least at this time. This could be problematic when the goal is to be able to state with some certainty the origins of the haplogroup in the British Isles. As we shall see it did not originate there, nor is there any evidence that R-U152 males were residing in the British Isles prior to the Iron Age.

U152 in Scandinavia: In a small research sample of those identified as R-M269 from all regions of Norway, the subclade marker U152 was found immediately north of Jutland in Southeast Norway, but was not observed elsewhere in the country (whereas R-U106 predominates among those who are in the R-M269 category according to this, albeit small, sample), and only a single U152+ example was found in a much larger sample from Friesland, west of the Jutland Peninsula on the North Sea.

Despite very limited testing of Scandinavians for the subclades of R-M269, commercial testing by Family Tree DNA and 23andMe has found U152 (ancestral U152 as well as downstream subclades) in Norwegians from the mouth to the head of Oslofjord, as well as Swedes from the area east of Oslofjord, and Norwegians along the western aspect of this region known as the Vik or Vestfold. The Danish R-U152 is found primarily in the ancient Cimbri territory along the eastern aspect of northern Jutland (north and south of Limfjord), as well as Funen / Fyn. There is an immense gap between these Scandinavians and their Y genetic brethren in any direction. There have been few R-U152 men reported anywhere in the surrounds until central and particularly southern Germany where there is a strong concentration of this haplogroup.

A significant problem here is that population genetics studies of Denmark suffer from four shortcomings. First sample sizes are small. Second, the genotyping is too far upstream to be useful (generally not extending downstream to U152). Third, only one study has attempted to parse Denmark into regions. Four, the academic studies have tapered off such that there does not appear to be a relevant study since the above Lucotte 2015 study.

Earlier Myres et al. (2007) employed a Danish sample in their study of two haplogroups, J1 and R1b3 (R-M269). Of the 113 males, 34.5% were R-M269 (clearly10 – 15% lower than in other studies). In this study they examined two subclades, M405 (S21 / U106) and M467 (S29 / U198). Of the R-M269 group, 49% were U106+, 2% were U198+ (a variety of S21) and 49% were R-

M269 (except the previous two subclades). Of the latter (the approximately half not U106) it is expected that some will be R-U152 and the rest R-P312*.

The 2011 Myres et al. study finds hotspots of S28/U152 in the Alpine regions of Italy and across Switzerland (confirming the findings of most other studies). In terms of Denmark, they found zero of this haplogroup in their sample of Denmark East, Denmark Island East, and Denmark West. However, in terms of the area associated with the Cimbri there are numbers that are expected and could explain a sizeable proportion of the haplogroup in Britain via migrations during the time of the Angles and the Viking era. In Denmark North and Denmark South East the percentages are 4.8 and 4.5 respectively.

A more recent study which includes S28/U152 is that of Busby et al., 2011. They combine the above data of Myres et al. and others with their own dataset and conclude that this haplogroup is most highly concentrated in the Alps. They also show that of the Norway sample (from Hemsedalsvegen east of Bergen), 3.6% of the sample were S28/U152. Unfortunately they combined two regions of Denmark from the Myres et al. study and arrived at a South West Denmark percentage of 2% confounding the two and reducing the South East from the 4.5% seen in Myres et al. Here, however, the Busby study defined (by latitude and longitude) the region where the samples were obtained and the Northern Denmark sample showing 4.8% came from Store Habendal south of Skagen and west of Frederickshavn (Cimbri country). The South West sample was given the coordinates of Odense on Funen / Fyn (again within the tribal orbit of the Cimbri). The above figures accord well with the data seen in the S28/U152 database and associated Google map of the author showing these two locations as the only ones with U152 samples found in the databases accessible to the author (e.g., the Family Tree DNA U152 Haplogroup Project) – as noted above.

See the Resource pages for this haplogroup, created by the author (albeit somewhat dated), for more specifics.

If the hypothesis noted earlier, that the Cimbri began as a unique Celtic tribe, but over time became a mélange of Central and Eastern European Celtic tribes constituted circa 101 BC is valid, then there should be considerable variability in the genetic make up of the Cimbri of Denmark, Vestfold, and England. Preliminary data found in a database created by the author shows that the R-U152 genetic haplotypes of England are indeed what appears to be an amalgam of what is seen in those from other parts of the La Tene Celtic world (e.g., Switzerland, Northern Italy, Eastern Europe). The importance of these subgroups will be explored later. In addition, likely from an early date, the Cimbri merged with Germanic groups and thus haplogroups such as R-106 and I1a-M253 would potentially make up the majority of the Cimbri of the Iron and Roman (Imperial) Ages.

The Possible Conflation of Viking Dane and Angle from the Jutland Peninsula: It must be noted that there is another possible explanation for the distribution of R-U152 in England that relates to the Jutland Peninsula – the possible Y-Chromosome composition of the Angles from the southern neck region of Jutland. This area was entirely abandoned after about 475 AD and the Angles migrated en masse to England. Their historical and archaeological record show the geographical extent of their kingdoms from the 5th Century (at precisely the time the settlements

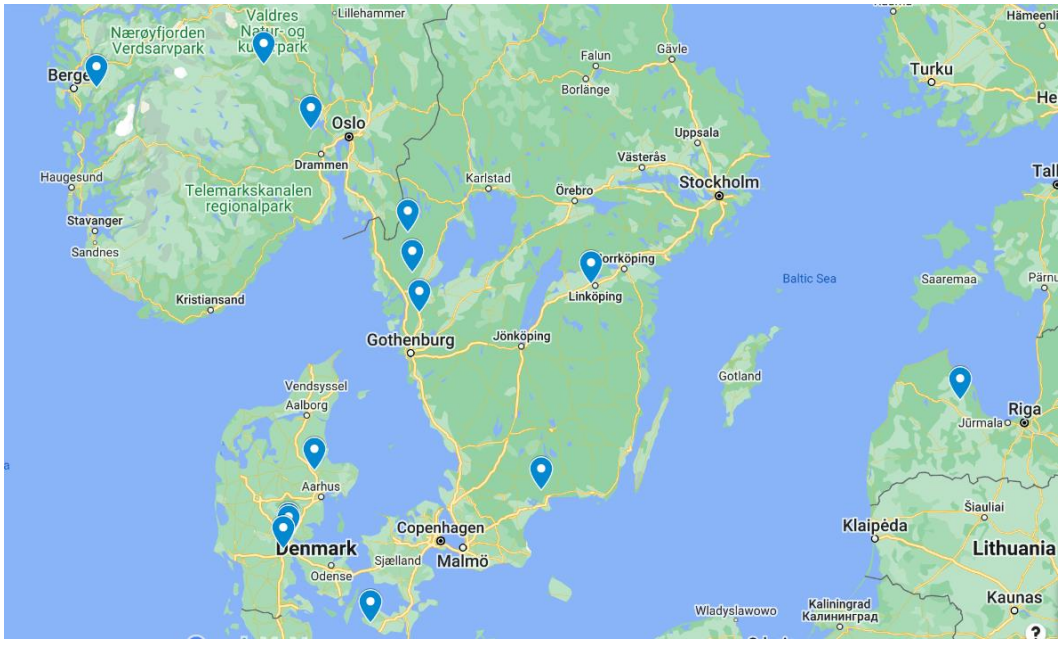
in the Angelin area are abandoned). The evidence of their migration from Jutland to England is seen via similar jewelry (e.g., cruciform brooches), shield styles, bracateate coin pendants, rune inscriptions found in both the Angle regions and contemporary artificats in England – but not in the Saxon parts of England or the Lower Saxony sites in Germany. Clearly close consideration must be given to the possibility that some of the R-U152 in England arrived from Jutland prior to the Viking era. A problem is that it will not be possible to assess the Y-Chromosome structure of circa 500 AD in the neck region of the Jutland Peninsula since the area was completely deserted for possibly two hundred years – leading to population replacement likely from Sealand and other areas of Denmark. It is also unknown what relationship the Angles had to the Cimbri, although they may be their direct descendants. This matter is taken up in some detail in another article by the author. Thus, until recently (e.g., 2015) all DNA statements must of necessity have been tentative inferences based on modern samples from living people in Scandinavia and England.

To date the academic community has largely failed to provide useful information on the array of U152 subgroup markers related to the current distribution of males across Europe. This problem has been addressed in the work of the current author (as noted above), and more importantly the FTDNA "U152 and Subclades Project". The most common of these subclades include R-U152 > Z36, R-U152 > L2, R-U152 > R-L2 > Z49, and R-U152 > L2 > L20/S144.

R-U152 > L2 > L20/S144

Subclade R-L20 Today: Various maps portray the distribution of this SNP in modern times. These include maps from Family Tree DNA's R-U152 and Subclades Project as seen here, and here, as well as from YSeq here. While the first two maps are a few years out of date, the picture here is of a SNP that is widespread, although not in Eastern Europe. It appears that the "hotspot" is Switzerland and the Alpine area, and along both sides of the Rhine River decreasing in numbers further north. In addition there are a significant number in Central and Western France, as well as England (although the latter is to a large extent influenced by the fact that this is an area where testing is "popular" among Brits and Colonials and so the numbers may be inflated in relation to say Sweden). Also, recently many more R-L20 have surfaced in locations such as Scandinavia. For example, there are three R-L20 individuals from the region in southwestern Sweden, a few miles north of Gothenburg; and two from eastern Sweden. As to Norway, there are R-L20 in Bergen, Hallingdal, and the eastern Oslofjord Vik. Similarly, in Denmark there are R-L20 males from Randers north of Aarhus; Agard west of Odense; one from north of Veile and one south of Veile west of Odense; and Tillitse, Lolland – none shown in the map below.

The distribution of R-L20 today among those who have tested as of 2016



The Distribution of R-L20 in Scandinavia Today (Genealogical Time)

A reasonable question relating to the current distribution of the haplogroup would be whether this pattern represents what would have been seen at the time of the Anglo-Saxon and Viking "migrations" to England. What is likely, but to an unknown degree, is that there was at one time many more R-L20 individuals living east of England but due to the above movements of people during the Middle Ages, there are fewer living today in the counties of origin of the migrants. A test of this hypothesis would be to examine the ancient DNA samples which, in the last 7 years, have gone from a handful of individuals to beyond 4,000 from all areas of the world, but particularly in Europe. What has driven this expansion is technological advances which have led to the ability to test skeletal samples many thousands of years old using upwards of one million SNP markers and high throughput equipment – as well as many very talented geneticists who have devoted their careers to this subject.

<u>Some Technical Details re the R-L20/S144 Designation</u>:

GeneticHomeland.com Mapping Technology for DNA, Surname & Genealogy Research

DNA Marker Index data for Marker: L20 on Chromosome: Y

Marker Name(s)	Notes	Identification	Ancestral[7]	Derived[8]	Chrom osome	Position (hg38)[1]	Position (hg19)[2]
R-L20 L20 S144 rs7067305	Found in haplogroup R1b under U152 on ISOGG, YFull, and FTDNA trees. hg38 Ref does not match ancestral allele value. Example is ancient DNA sample VK373 from 11th century Viking grave in Galgedil, Denmark. Note that an L20 lineage was one of the three samples used to compose the hg19 and hg38 human reference sequences. Phylogenetic Parent: Z367 / S255 View Pedigree View Map of descendants (Requires Registration Login) Phylogenetic Children: BY203086 BY204733 BY34065 BY5690 BY61198 BY69713 CTS9733 FGC54100 FT251781 FT340981 S23900 Z1909 Z291 NIH DBSNP b153 GRCh38p12 ClinVar? Position hg18: 12741292	Thomas Krahn, FTDNA	G	A	Y	12110586	14231292

The above chart indicates the position along the Y chromosome where R-L20 is situated, its "rs" code and other important details. It also lists "descendant lines" which must be very up to date since it includes BY61198 and the other recently discovered branches or terminal markers. There is one error in the above in that, as noted earlier the SNP was discovered by Dr. James F. Wilson of the University of Edinburgh who was the business partner of the current author in a company called "Ethnoancestry". The current author was the first individual to be tested for the parent marker of L2/S139, as well as L20/S144 (the "S" designation denotes that it was Dr. Wilson's discovery), and was found to be "positive" or "derived" for both markers.

For the purposes of genetic genealogy there are two descriptors that are most important.

1) Every Y SNP mutation (marker) is given a short hand name. In this case it is L20 or S144. One of the "children" of L20 is BY61198 which is a terminal marker for both the ancient DNA sample from Rookery Hill, and the present author (see here). It is located at Y

chromosome position 6901207 (Build HG38) where the ancestral C has been converted to a T. However, according to Family Tree DNA, I14538, the present author, and anyone else who tests positive for this SNP also **share 31 "private" named mutations** along the Y chromosome. It is estimated to have branched off a man who was R-L20* (the Asterix means that there are no other phylogenetically significant markers) about 50 CE (i.e., about 2000 years ago). Downstream of BY61198 is found a SNP known as BY55987, which has been detected in two individuals but they either don't know or have not reported their paternal line country of origin. It is estimated that their SNP branched off BY61198 about 200 years ago (see here). It should be noted that while the author has 12 "private SNPs", the BY55987 individuals have 21 mutations from BY61198 and three "private SNPs" (i.e., 24 private variants if only one individual had tested) – more on this later.

2) For research purposes all Y SNPs are given a unique letter and number designation reflecting its position on the Y phylogenetic tree. This work is updated every few years by the International Society of Genetic Genealogists (ISOGG), see here. The current version, used in the most recent studies, is Version 15.73 (2020) wherein L20/S28 is given the tag of R1b1a1b1a1a2b1a1 – or either R-L20 or R1b-L20 for clarity. Generally those who test to a deeper phylogenetic level will be assigned a "terminal SNP" below R-L20 (or assigned R-L20* if no other downstream markers are detected). The very uncommon R-BY61198 is the "terminal SNP" of the author.

Recent Techological Advances: It is recognized that the details above will need to be amended, as new information becomes available, and sources must be added, but the major facts and most of the conclusions appear to rest on solid ground. Fortunately, since earlier versions of this manuscript were written, the technology has changed dramatically (e.g., chip arrays of over one million markers replacing direct sequencing), and new results are much more revealing than anything published before.

Clearly all studies of European Y chromosomes, whether current populations or ancient DNA have until now been "inadequate" for one reason or another (e.g., genotyping to a level which is uninformative). Recent developments have changed the picture in a dramatic way, although few academic studies of modern populations have been reported. If the resources are available, population geneticists are now able to delve deeply into Y chromosomes and explore markers far "downstream" of for example M269 (the very "generic" marker found in 33% or more of European males) and the autosomal DNA of the 22 pairs of chromosomes – both for current groups and ancient DNA from samples thousands of years old.

Ancient DNA

Since about 2015 advances in technology, including the ability to "dig deeply" into Y chromosome haplogroups beyond the standard generic M269, and even delve into the many newly discovered markers that characterize the U152/S28 phylogenetic tree, have revolutionized the understanding of the historical and pre-historical landscape of, for example, Europe. What is perhaps most astounding is that today we are able to use essentially the same technological advances and apply them to very early archaeological samples. Recently the DNA sequences of Neanderthal and Denisovian individuals who lived in Europe and Asia over 40,000 years ago have been published.

It is now known that individuals outside Africa have about 2% to 4% of their genomes inherited from these early hominids – via inter breeding with modern humans. Hence if we can locate skeletal samples from say the Medieval era (e.g., 800 AD), or the Hallstatt - Celtic era (e.g., 800 BC), if the samples are reasonably well preserved then we can provide data equivalent to that we can obtain from anyone donating a sample today.

Fortunately the authors of ancient DNA studies have recognized the importance of parsing the major Y haplogroups into subclades. Breaking down R-U152 into the varieties seen most commonly in ancient DNA and today, we note the following. R-U152 did not get quite as much "traction" in terms of descendant lines as one of its "children" (R-L2) although R-Z36 in particular has many descendants today. The R-U152 downstream marker R-L2 is heavily represented in ancient DNA samples as seen in the map here by Richard Rocca. R-L2 includes subgroupings with large descendant groups, such as R-Z49 as well as R-L20/S144. The focus of the current study will be on R-L20/S144 due to the importance of the ancient DNA findings relating to the Anglo-Saxons and the Vikings of Denmark, and the fact that this is the haplogroup of the current author.

Bronze and Iron Age Burials in Continental Europe and the UK:

In relation to the contents of the present paper, the published ancient DNA data (see the map of Ancient DNA finds by Richard Rocca at the Family Tree DNA U152 and Subclades Project) is extremely informative – and bodes well for future studies. Many of the finds have been attributed to the Bell Beaker cultural tradition circa 2800 BC to 1800 BC. The apparent earliest U152 find to date is recorded in Olalde et al., 2018 (see below) where a Bavarian sample from the Middle Bronze Age Tumulus Culture is dated to 2572 – 2512 BC.

The bulk of these finds have been published by Olalde et al., 2018 in an article entitled, "*The Beaker Phenomenon and the Genomic Transformation of Northwest Europe*" – see here. It is necessary to access the raw data via the Excel Supplementary Tables to view the specific haplogroups (beyond the more generic S116/P312 described in the text). Here we see that for example in the United Kingdom most of the samples are R-L21 (no U152). Even today L21 reaches its virtual saturation point in Ireland (most Irish are L21 or a subgroup of same) – and extends in a decreasing cline across Britain to the east coast where it is considerably less prevalent. This stands in stark contrast to the Continent where the samples in the Central and Western regions are largely U152 > L2 (37 of the total in this study). No L20 was detected in any sample in this study (despite the majority of samples being tested for this marker). These areas are also regions of high concentration of this haplogroup today – although much less so in the east (e.g., Hungary). The Bell Beaker culture was succeeded by the Unetice Culture, described earlier (with map) in this paper. Thus, these burials perhaps represent proto – Celtic individuals.

In a paper entitled, "*Genetic transition in the Swiss Late Neolithic and Early Bronze Age*" by Furtwangler et al., (Nature Communications, 2020), as seen here, they report ancient DNA data from 96 burials dated between 4500 to 2000 BC. All samples were recovered from cemeteries between Alsace, Southern Germany, and the Swiss Plateau. Here we find 10 samples that could be genotyped as U152 > L2. Patterson et al. reanalyzed this sample in their study and found it to

be R-L20 – which if correct, may mean that others who were originally labelled L2/S139 were actually R-L20. The current author is currently attempting to validate this finding via contacting the corresponding authors. The cemetery is situated in Singen, in the Black Forest area of Southern Germany close to the Swiss border west of Lake Constance.

The L20 sample is dated to plus or minus 1967 BC making it part of the Early Bronze Age. As noted above, today L20 is found primarily from the Lake District of Italy, through Switzerland north through Germany (particularly the Rhineland) and France, to the Baltic States, Sweden and Denmark over to England.

In the study by Patterson et al., entitled, "*Large-scale migration into Britain during the Middle to Late Bronze Age*" (Nature, 2021), as seen here, he provided data on samples from across Europe. He reported, among many R-L2, one R-L20 > Z291 > BY198578 > S1610 from Jinonice, Prague, Czechia dated at between 2200 and 1600 BCE. The same study also reported an Iron Age sample dated to circa 250 BC sample from Faux – Vesigneul, Marne, France which was genotyped as R-L20 > Z1909 (another assessment provided a R-L20* assignment).

Based on the findings of this study, it seems clear that none appear in the UK until the Iron Age. No **Bronze Age** samples of haplogroup R-U1152 have been located in the British Isles. However, the findings suggest that during the **Iron Age**, R-U152 individuals began to appear in the UK as follows - with individual, location of find, date, and subclade noted:

I16440 – Cornwall – 800 BCE to 43 CE – L2 > Z49
I19873 – Kent – 400 to 200 BCE – L2 > Z49 > Z142 > Z12222 > BY3616
I20589 – Oxford – 400 to 200 BCE – L2 > Z49 > Z142 > Z51 > L562 > Z57 > Z148 > Z52
I16422 – Southern Scotland – 364 to 121 BCE – L2
I20982 – Hampshire – 450 BCE to 1 CE – L20 > Z1909 > A1510
I13758 – East Yorkshire – 400 to 50 BCE – L2 > FGC13616 > FGC13633
I14097 – North Yorkshire – 162 BCE to 26 CE – L2 > DF110
I19208 – Oxford – 382 to 205 BCE – L2
I19047 – Cambridge – 1 to 50 CE – L2

Concerning the **Roman (Imperial) Era**, few samples from the UK have been assessed. However the findings from a noteworthy "Romano – Briton" burial site from near York, Yorkshire, England dating to between 200 and 400 AD has been published. In the study by Martiniano et al., 2016 entitled, "*Genomic signals of migration and continuity in Britain before the Anglo-Saxons"* as seen here, the authors explored the autosomal, mtDNA and Y chromosome DNA of 7 males, all of whom were decapitated. It is thought that this group may have been Roman soldiers. Five had "local" (Brithonic) Y chromosomal haplogroups (varieties of R-L21), whereas one had a variety of haplogroup J, and appeared by virtue of the autosomal DNA to have been from the Middle East. The other Y chromosome "outlier" was a S28/U152 > L2 individual – although whether local in origin or from the Continent cannot be discerned based on the autosomal DNA. Perhaps this individual was a Roman auxiliary soldier from the Continent, or a descendant of a probable Brythonic group from Yorkshire with cultural links to the Parisi or Belgae of the Continent – as described in the "Hypotheis C" article by the present author. Irrespective, the above two studies

suggest that U152, particularly the L2 subclade, was in the United Kingdom before the main Anglo–Saxon migration beginning circa 447 AD.

Anglo-Saxon Burials in England and Continental Europe:

In the most recent ancient DNA study published (22 September 2022), Gretzinger et al., "*The Anglo-Saxon migration and the formation of the early English gene pool*" (see here), the authors used 1,240,000 SNPs to explore the genomes of 460 ancient Northwestern Europeans to a sample of 278 individuals from England. They used the Y SNPs from the entire panel and compared them to the v.15.73 (2019-2020) version of the ISOGG Y haplogroup tree. The data was automatically assigned to a haplogroup, although they apparently did some degree of manual analysis. Unless the .bam full file of SNPs was consulted then the assignments might be off to some degree. Fortunately, at least 3 genetic genealogists did perform the complete analysis (S.S., Webb, Alex W.). Here the path to R-L20 was verified, and any downstream SNPs not "seen" by the authors were noted.

In this study, using 460 total samples, 182 were from Ireland, Denmark, the Netherlands, and Germany, with 278 individuals from England, where there was only one R-U152 individual reported.

The Continent: Individual number, location of cemetery, haplogroup -

IND003, Alt-Inden, North Rhine, Westphalia, Germany, R-L2 > Z258
IND017, Alt-Inden, North Rhine, Westphalia, Germany, R-L2 > BY13147
ADN007, Hanover-Anderten, Lower Saxony, Germany, R-L2 > Z49 > FGC22963
ADN012, Hanover-Anderten, Lower Saxony, Germany, R-L2 > Z258
GRO005, Groningen, Groningen, Netherlands, R-U152 > BY3654 > BY41129

England: Individual number, location of cemetery, haplogroup –

I14538, Rookery Hill, Bishopstone, Sussex, England, R-L20 > BY61198

The English sample was of particular interest to the present author. I14538 was from an early (early 5th Century to late 6th Century) Anglo-Saxon cemetery from Rookery Hill, Bishopstone, Sussex, on the south coast of England. There is little archaeological information available (although it exists in various locations), but with a Quoit Broach being noted and an array of other diagnostic information, it would appear that the individual in Grave 43 was likely a mercenary associated with the British king Vortigen's interaction with the Jutish Princes, "Hengest and Horsa" and the Jutish people were given territory in a specific treaty area between the Ouse and Cuckmere Rivers (Jolliffe, 1933; Cunliffe, 1997). The date, place and time suggesting that this was a settlement of people who would be known as Jutes.

The reason why this individual is of particular interest is that in addition to obtaining a haplogroup of R-L20, both S.S. and Family Tree DNA, using the .bam files, determined that the individual in Grave 43 was also positive for the downstream SNP BY61198. **This SNP is precisely the same as that of the author, and after hundreds of L-20 having been tested, they are among the only**

three lineages who have tested positive for this (clearly rare) SNP. It should be noted that two individuals have BY61198 but also another SNP that Family Tree DNA reports split from the parent circa 1800 CE. Nothing is known of their ancestral country of origin. The ancient DNA finding demonstrates that I14538 shares a recent ancestor in common with the author (estimated at 50 AD by Family Tree DNA), and furthermore that his point of origin on the Continent is close to the "ancestral home" of the author's patriline lineage. Since R-U152, R-L2, and R-L20 are considered to be "Central European Alpine Celtic" it is important to show that these haplogroups were / are to be found in Scandinavia – particularly what is today Denmark and other areas which may have spawned the Anglo-Saxons who migrated to England beginning in the early 5th Century, Danish Vikings post 793 CE. The author has shown above that the Jutes were were from Jutland, and likely descendants of the Celtic Cimbri people of the Jutland Peninsula (called by Roman writers the Cimbric Cheronesus).

It is unclear why so few R-U152 individuals are found in the Gretzinger et al. study, however one problem is the fact that in Jutland from the late Bronze Age through Anglo-Saxon times, the overwhelming number of burials there are cremation. This is true of many, especially Anglo-Saxon cemeteries such as Sprong Hill, Norfolk where there were 2259 cremation burials and 57 (later) inhumation burials. Thus if the overwhelming number of Angle burials are cremation – our chance of being able to sample their Y chromosome haplogroups becomes problematic. There were no samples in this study from for example northern or central Jutland, which compounds the problem. Only with much larger and broader sampling will it be possible to make more definitive statements.

However, in supporting the hypothesis under consideration in the current study, what is most pertinent to the present work is the recent study of Viking era genomes from across the "Viking world".

Viking Era Genomes:

A landmark study of 2020 entitled "*Population genomics of the Viking world*" (Margaryan et al.) (seen here) explored the genomic structure of those who peopled Scandinavia and the diaspora during the Viking Age (c. 750 to 1050 AD). They used an a 26,083 biallelic SNP (single nucleotide polymorphism) array and were able to place each male Y chromosome into a category from the ISOGG (International Society of Genetic Genealogy) phylogenetic tree. This study is one of only a few to use the unprecedented depth of analysis that has become available by 2015 – even for modern samples, let alone samples 1000 or more years old. As with the Olalde et al., 2018 study noted above, in order to view the "deep" haplogroup levels of each sample it is necessary to refer to the Excel Supplementary materials (file 04). It is also necessary to convert the letter – number sequences such as R1b1a1b1a1a2b to the more manageable short hand versions, in this case R-L2 or R-U152 > L2. R1b1a1b1a1a2b1a1 translates to R-L20 or R-U152 > L2 > L20.

As expected by the present author, although not most others in the genetic genealogy community, there were individuals with the "Continental Alpine Celtic" Y Haplogroups of R-U152 and derivatives R-L2 and R-L20 among the Viking Age "ancient DNA" samples from Denmark and southwestern Sweden – as well as a Dane from Oxford, UK. Some of the most relevant findings as they relate to the Cimbri and Y Haplogroup R-U152 include the observation that in Viking Age

(VA) Scandinavia, "Many individuals from southwestern Sweden (e.g., Skara) cluster with Danish present-day individuals from the eastern islands (Funen, Zealand), skewing towards the 'Swedish' cluster with respect to early and more western Danish VA individuals (Jutland)." Skara is inland but just east of the tip of the Jutland Peninsula. Furthermore, "We also observe several individuals with large amounts of South European ancestry in Denmark and southwest Sweden during the Viking period." An inspection of the raw data shows that it was common for those VA individuals from Funen and associated areas to have much higher amounts of Southern European ancestry (e.g., 20%) than those from for example western Norway. This finding of Southern European – Like ancestry is entirely consistent with the hypothesis being advanced in the present paper since the Cimbri appear to have originated in southcentral Europe (e.g., as the Boii from the modern Czech Republic and Northern Italy). Also, after the defeat of 101 BC at the hands of the Romans, the historical record shows that the remnants, originally from eastern Jutland, Funen and proximal islands, were accompanied back to the Jutland region by Swiss Celtic tribes such as the Helveti. It is predicted that no matter what the Y chromosome Haplogroup, the people residing in the Funen and Jutland area would have, via Cimbri ancestry, a larger Southern European contribution to their genomes than those from for example western Norway.

The specific U152 and downstream findings for each sample number are as follows (Sample Number; Haplogroup; Location of Find; Major Autosomal Ancestry included). There were originally (when study first published) only 5 U152 samples, but now with a more detailed analysis and re-testing, there are 10. HT is Haplotree SNPs, and YF is YFull - identified SNPs. More in depth haplogroup testing of the R-L20 samples is included below. These samples are still being analyzed so the SNPs below are likely to be added to or amended over time:

VK177 = R-U152; HT=R-FT31867; YF-R-B83: Oxford, UK - 28% "UK-Like"; 36% "Danish-Like"; 33% "Southern European-Like". Sample size 22 (all male). Date c.1000 AD.

VK40 = R-U152 > L2 > Z49 > Z142 > Z150; HT=RBY166438; YF=R-Y3140: Skara, Varnhem, Sweden - 21% "UK-Like", 67% "Danish-Like". Sample size 31. Date c.1050 AD.

VK138 (AQQ) = R-U152 > L2 > Z49 > Z142 > Z150; HT=R-S149; YF=R-S1491: Galgedil, Funen, Denmark - No data. Sample size 16 (10 males). Date c.950 AD.

VK273 = R-U152 > L2 > Z49; HT=BY61747; YF=R-FT186424: Gnezdovo, Russia – 36% "Danish-Like"; 50% "Polish-Like"; 1% "Southern European-Like". Date c.1000 AD.

VK335 = R-U152 > L2 > Z49; HT=R-FT304388; YF=R-Y56490: Oland, Sweden – 1% "UK-Like"; 61% "Danish-Like"; 24% Southern European-Like". Date c.950 AD.

VK159 = R-152 > Z36; HT=R-A7982; YF=R-A12417: Pskov, Russia. Date c.1000 AD. Date c.1000 AD.

VK286 = R-U152 > L2 > Z367 > L20 > Z1910; HT=R-S10708; YF=R-Y6789: Bogovej, Langeland, Denmark - 23% "Danish-Like"; 63% "Southern European-Like". Sample size 13 (8 males). Date c.950 AD.

VK326 (K1578) = R-U152 > L2 > L20 > CTS9733 > S15057 > Y52895; HT=R-Y52895; YF=R-Y58552: Ribe, Jutland, Denmark – 12% "UK-Like"; 56% "Danish-Like"; 31% "Southern European-Like". Sample size 9 (5 males). Date c.950 AD.

VK373 (BER) = R-U152 > L2 > Z367 > L20 > FGC56105; HT=L20; YF=R-Y82494: Galgedil, Funen, Denmark - 69% "Danish-Like"; 23% "Southern European-Like". Sample size 16 (10 males). Date c.950 AD.

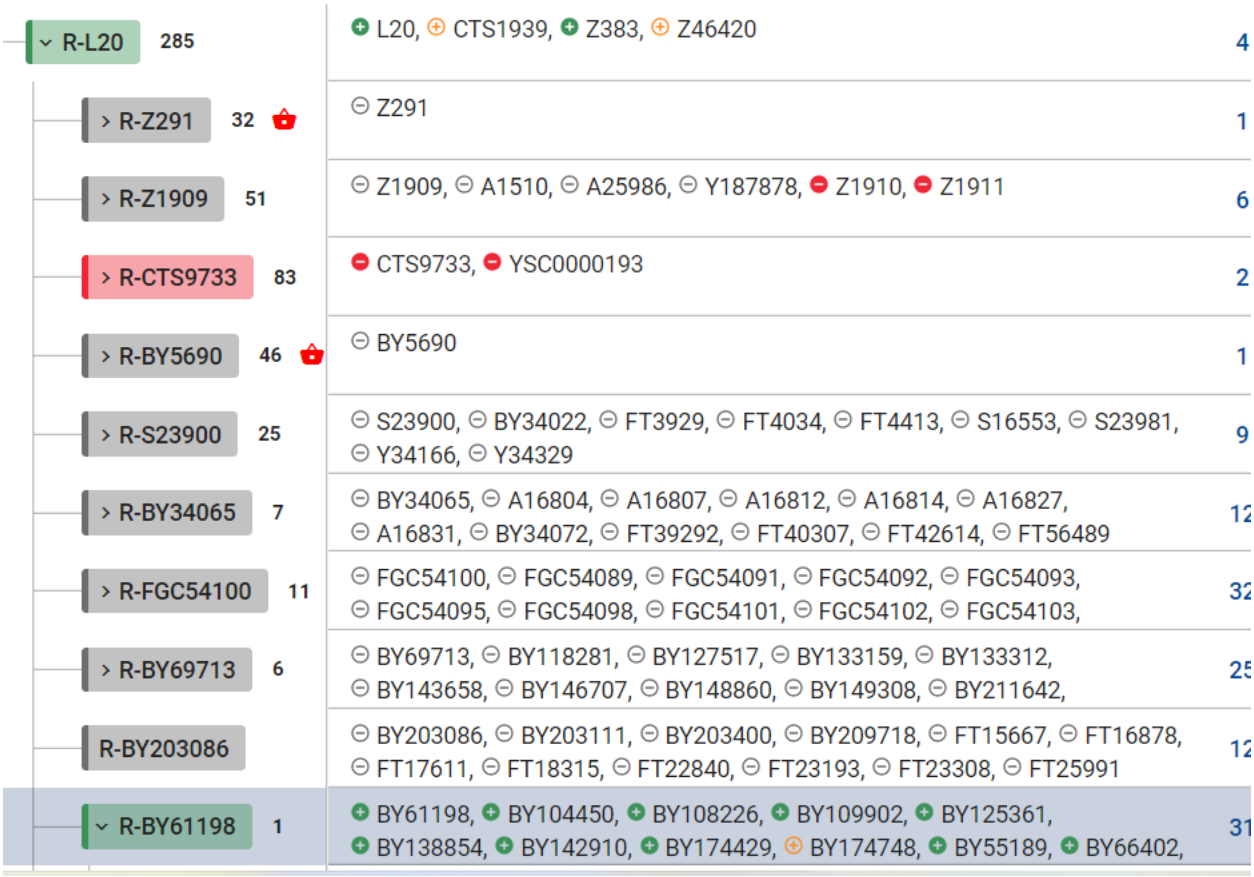
What is evident from the wealth of inter - disciplinary data is that all of the Viking Age U152 burials have ancestry from Denmark. Even the L2 individual from Skara, Sweden likely had origins (at some unknown time) in Denmark. The DNA evidence supports the observation (found in Margaryan) by the archaeological investigator Maria Vretemark that, the entire Vastergotland area (including Varnhem – Skara), is "a region that had long been part of the Danish sphere of influence." Galgedil on Funen Island (L2 and L20 samples) and Bogovej on Langeland Island (L20 sample) as well as Ribe (L20 sample), on the west coast of Jutland were within the Cimbri original territory. No graves on the east side of Jutland (Funen is an extension of this land mass) were explored for the purposes of the Margaryan study. Since the Venerable Bede, an Angle, recorded that the ancestral land of his people was still abandoned in his day (791 AD), it is possible that most of the U152 Cimbri descendants from the Angle region left Jutland (as Jutes and Angles) for England during the Anglo – Saxon Era. Hence it is likely that inhumation cemeteries in Jutland dating to say 500 AD would include a mixture of Y chromosome Haplogroups, but potentially a significant number of U152 individuals whose male line lineage could be traced to Central Europe in the Iron Age. It is noteworthy that the DNA of the U152 samples although strongly "Danish-Like", included a Southern European-Like component ranging from 23% to 61%. This suggests an isolated population whose origins in the Iron Age or earlier may have been the Alpine region among the Celtic tribes such as the Heleveti.

The most informative of the above samples, in terms of the hypothesis being examined in the current study, are the three R-L20 individuals, plus the Jutish Anglo-Saxon individual all of whom had their origins in what is today in Denmark, in the region where the Cimbri once resided, meaning Jutland and the eastern offshore islands such as Funen.

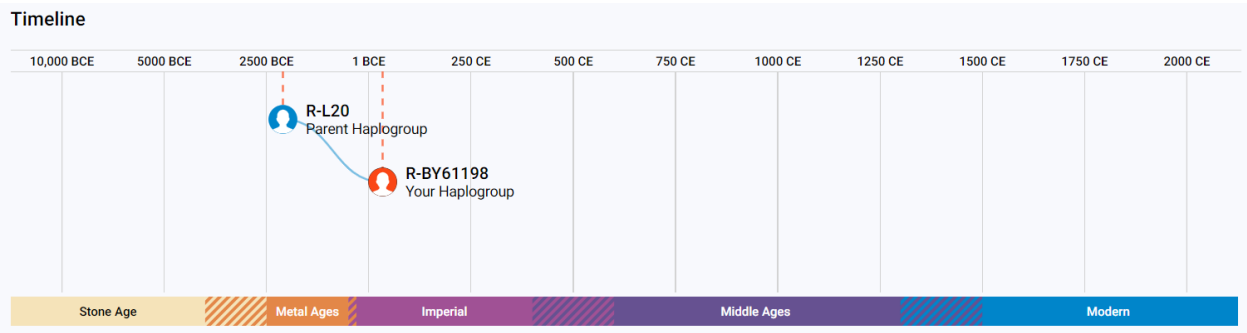
***Map showing the location of the ancient R-L20 samples from the Margaryan study*.**

SNP Testing of Faux Y-Chromosome

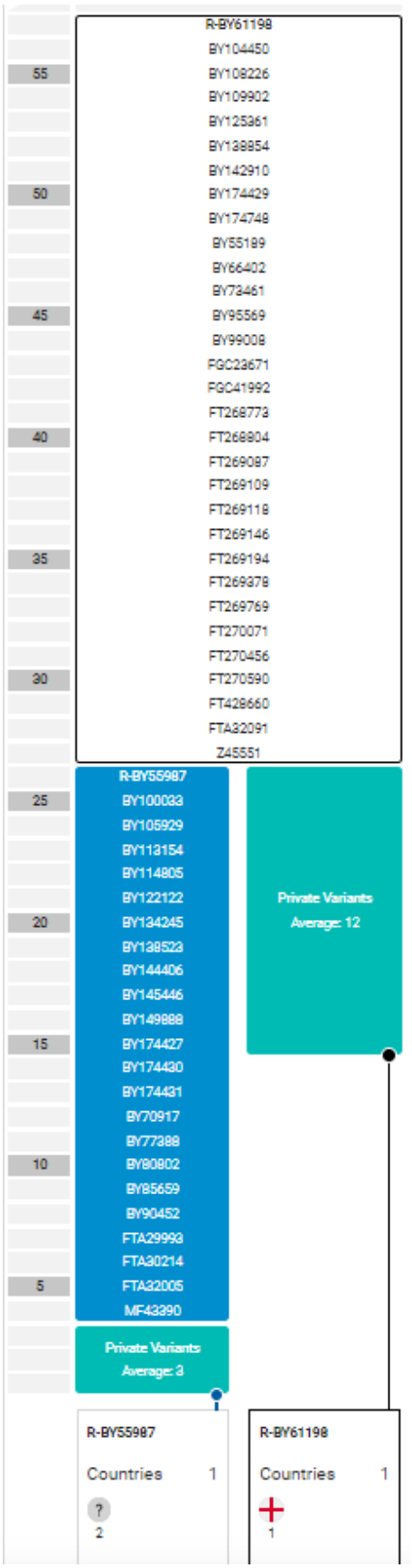
The author, serving as the proxy for the Falke – Faux family who are descendants of Willliam Falke born circa 1385 in Worlingham, Suffolk, England took the Family Tree DNA Big Y 700 test to probe deeper than R-L20/S144 which was determined in 2004. The results showed that he was R-BY61198 seen in the context of a phylogenetic tree below:



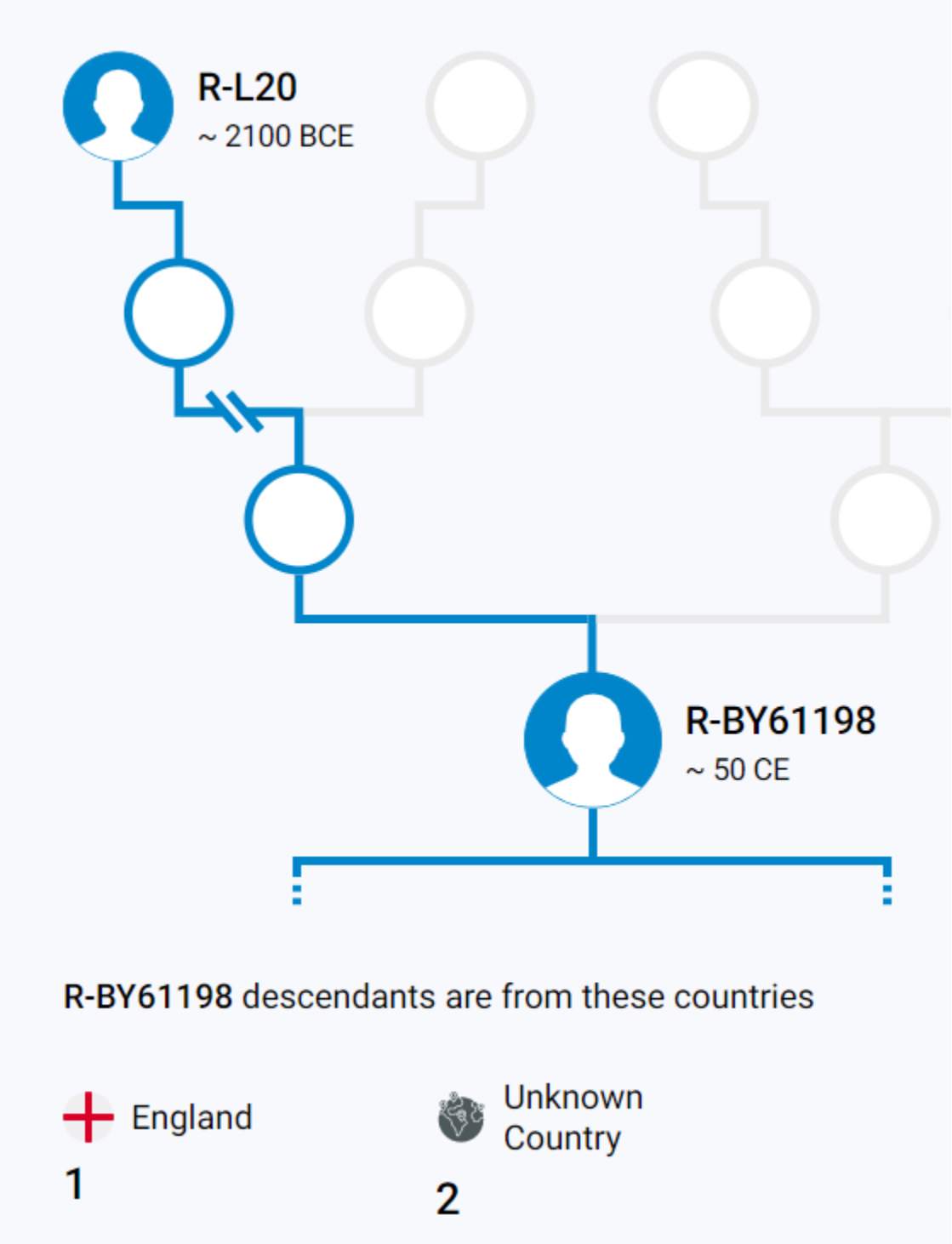
The list on the right are private variants. Thus I and all other R-BY61198 have each of the 31 associated private variants.
The timeline is as follows:



I have 12 private variants that are to date unique to me. As seen in the block tree below:

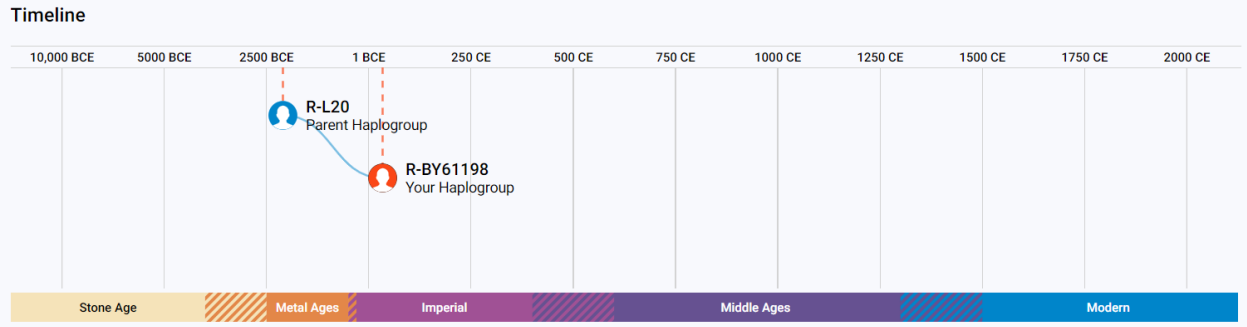

R-BY61198
BY104450
55
BY108226
BY109902
BY125361
BY138854
BY142910
50
BY174429
BY174748
BY55189
BY66402
BY73461
45
BY95569
BY99008
FGC23671
FGC41992
FT268773
40
FT268804
FT269087
FT269109
FT269118
FT269146
35
FT269194
FT269378
FT269769
FT270071
FT270456
30
FT270590
FT428660
FTA32091
Z45551
R-BY55987
25
BY100033
BY105929
BY113154
BY114805
BY122122
20
BY134245
BY138523
BY144406
BY145446
BY149888
15
BY174427
BY174430
BY174431
BY70917
BY77388
10
BY80902
BY85659
BY90452
FTA29993
FTA30214
5
FTA32005
MF43390
Private Variants
Average: 12
Private Variants
Average: 3
R-BY55987
Countries 1
?
2
R-BY61198
Countries 1
1

The diagram below from the “Discover More” option shows that my haplogroup branched off from its parent R-L20 about the year 50 CE.



Further information on ancient DNA matches include the above Rookery Hill, Bishopstone, Sussex Anglo-Saxon individual, as well as the above R-L20 Viking Age individuals from Denmark.

The timeline is as follows:



Further information can be found here.

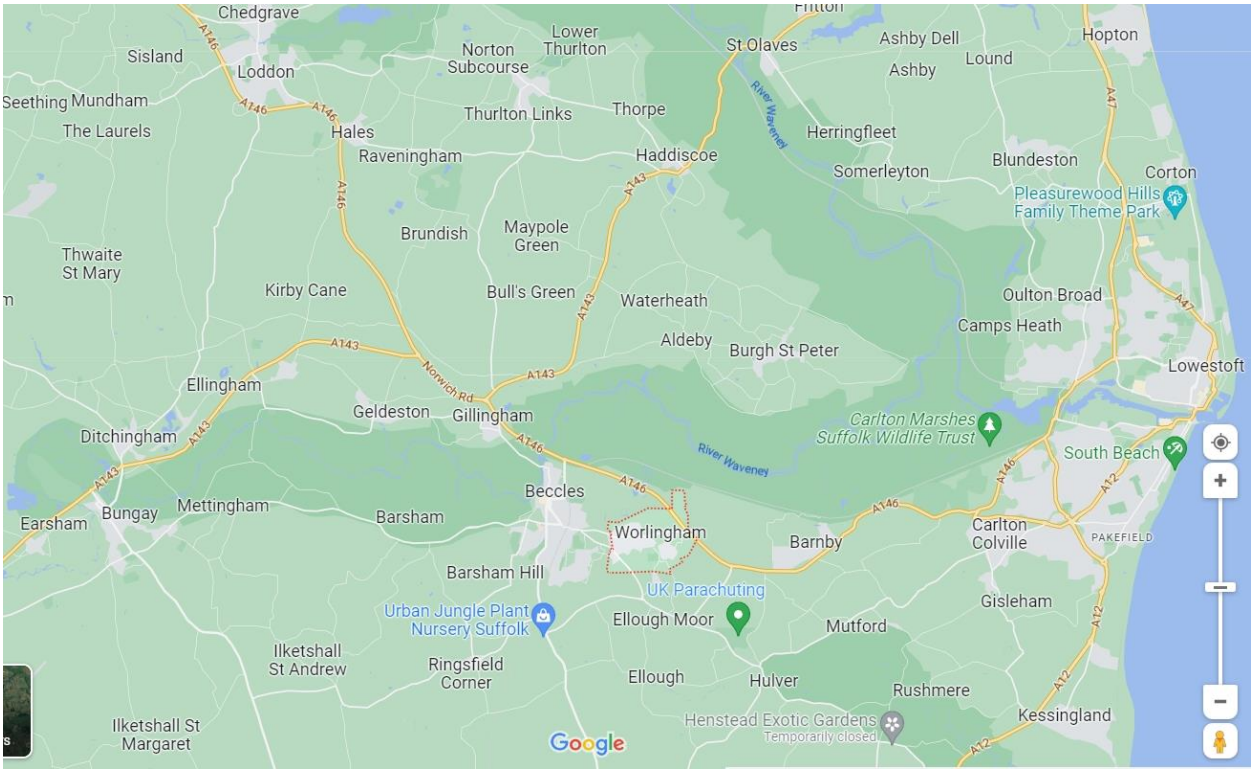
Conclusion

Most of the above material is taken from a study by the current author of the Celtic Cimbri of Jutland, Denmark, and articles on their probably successors the Jutes and the Angles (see here).

It was of particular interest to the present author that three of the U152 samples from Viking Era Denmark were L20, showing that this downstream haplogroup of U152 has been present in Denmark from at least Viking times, and likely much earlier. The Anglo-Saxon Era Jute from a grave in Sussex, England with the same R-L20 downstream haplogroup as the author points to the origin of their most recent ancestor in common being from the north and east - Denmark.

Family Tree DNA has predicted that the BY61198 mutation emerged on the Y chromosome of a R-L20* ancestor about 50 CE (an estimate with a very large confidence interval). Therefore, in order to extend the pedigree further back in time, it will be necessary to find an individual who tests as R-L20* and includes most or some of the 31 private SNP variants which all BY61198 possess. In all likelihood the individual would be from Jutland, or perhaps from an earlier geographical location such as Switzerland. Hopefully this scenario is just around the corner. The way the field has progressed since 2015, a find of this nature enters the realm of "a very distinct possibility".

In England, it would be helpful to find an ancient DNA BY61198 individual, with some of the author's private DNA variants, from Viking sites such as Repton, or in East Anglia where the author's Y line ancestors have resided since at least 1385 CE. The family had a Scandinavian surname (Falke), lived in a location (Worlingham and North Cove) surrounded by villages with Scandinavian place names, and were wealthy ship builders.



Map of villages in the vicinity of Worlingham, Suffolk

The villages that can be clearly identified with the Vikings in Suffolk above include, Lowestoft, Carlton Colville, Ilketshall St. Andrew, Ilketshall St. Margaret, Barnby, and Hulver. In Norfolk these include Thorpe, Ashby, Lound, Thwaite St. Mary, Kirby Cane, and Alderby. Some of the others may be "Grimston hybrid names" (amalgam of Anglo-Saxon and Viking name components).

Therefore, perhaps the most probable scenarios include the Y – line ancestor of the author coming to East Anglia with the "Great Army" of Danish Vikings who settled there under their leader Guthrum, who became King of East Anglia in the 9th Century; or with one of the subsequent waves of Danish people up to the time of Thorkell the Tall who ruled East Anglia in the 11th Century. It is also possible that my East Anglian Grandfather was correct in maintaining that our Faux ancestor was Norman – in other words a Danish Viking who took a more circuitous route, with a lengthy stop - over in France, before arriving in England to put down roots. At any rate, in the early Middle Ages (Anglo-Saxon times) the ancestor in Denmark would likely have been known as a Jute, and prior to that, during the Iron Age, a Cimbri.

Dr. David K. Faux
Cypress, California and Caledonia, Ontario

Version: 28 January 2015; 21 November 2022; 15 April 2023