

G12: Painting the Past by Numbers

Calculating Who Was Where and When

WC 3787

“History”, they say, “is written by the victor” and nowhere is this more true than in genetic genealogy in which we are writing the history of our ancestors going back to y-Adam and Mitochondrial Eve. Those ancestors, each one of whom was as essential to our existence as our own parents, survived not only the rigors of climate, natural catastrophes, plagues, famines and wild beasts, but also the crimes against humanity our species has committed upon itself.

But, like other victorious historians — our old friend Tacitus comes to mind — we cannot escape two historiographical problems¹ which crop up time and again in any history: one is that we tend to see the past through the lens of our contemporary mind-set; the other is to confuse assumptions with facts. The first is a problem of interpreting what is often too few facts at our disposal. To give an example: it is almost customary in this Western post-feminist world to see polygamy among Moslems as exploitative of women, but if you were writing a history of Islam in the 7th Century you would probably agree with the Prophet that it was better for the few men who had survived the ferocious wars of the time to take more than one wife than to leave the thousands of surplus women to their fate — starvation, prostitution and slavery.

Confusing assumptions with facts is commonplace not only in history but in science as well. Often, when scientists publish their research in an academic journal they use a *hypothetical construct* to sum up their findings. So, for example, in my own field, in the early years of the 19th Century, researchers noted some people were quicker to solve problems, see relationships and so on and they summed this up as “Intelligence”. Despite the fact this is, strictly speaking, only a hypothetical construct, later generations of psychologists have gone on using it as though it were a real object or an organ in the human body.

Genetic genealogy suffers from all these problems, the more so because it is a multi-disciplinary study which incorporates the “findings” of several others, including most notably genetics, archaeology, linguistics, and statistics. How contributions from these — sometimes mutually hostile — disciplines can be combined to produce a credible account of our species’ history is still far from a congenial compromise!

¹ That is not to forget bias or selective presentation — propaganda springs to mind.

the other three SNPs we already know belong in this clade so I must now await new discoveries.

Presuming the Past:

With the exception of a few aDNA analyses of ancient skeletons found in pre-historic European graveyards², the only genetic test results we have to go on are those taken from people living in the 21st Century AD. It is all too easy to jump from these results to the wrong conclusions:

First, just because we have the most test results for a particular clade in one of other geographical locality does not mean that was where the initial mutation occurred many thousands of years ago. This is simply a matter of “testing bias”. The reverse is also true: for example, for legal reasons, fewer French people are tested than are Americans or British and this skews our understanding quite significantly of the spread of say R-L21, a matter which is being hotly debated at the present time.

Second, we cannot assume that haplogroups emerged in the locality in which they are most plentiful today. There is always the possibility — indeed, probability — that the mutation could have occurred in one locality but the bearer or his or her descendents moved elsewhere before numbers increased significantly — or “expanded” in genetic genealogy jargon. Barbara Arredi et al. made a neat point when they wrote³:

...lineage ages cannot be equated with dates of migrations: if humans colonize Mars, their mtDNAs may have a Paleolithic MRCA, but we might be unwise to infer from this a Paleolithic Martian colonization.

Apart from *Where?* the other big question we ask in genetic genealogy is *When?* Again, this requires caution — as well as considerable knowledge of statistics...

One of the main ways of estimating — and note, this can only ever be an estimate — the age of a particular clade is to consider the diversity of the haplotypes of the people within it. This can be so because it takes time for particular STR markers to mutate. Of course they do not all mutate at the same rate, some markers mutating more rapidly than others and much research time is being spent trying to determine the mutation rates of these STRs.

² Actually, there seems to have been more interest in testing the aDNA of Neanderthals!

³ Quoted in Arredi et al., *The Peopling of Europe*, p. 391 in Michael Crawford (Ed): *Anthropological Genetics, Theory, Methods and Applications*; Cambridge University Press, 2007 (I think after Barbujani et al., 1998 to whom they refer).

One of the problems of using haplotype diversity as an indicator of age of the haplogroup is that diversity will probably have been reduced somewhere along the line by bottlenecks. There is a clear history in Europe of bottlenecks of one kind or another: eg., the Black Death, countless wars, and perhaps the greatest of them all, the LGM.

Time to the Most Recent Common Ancestor

Most genetic genealogists do not try to estimate the actual age *per se* of a haplogroup or its clade because it can probably never be known when the mutation actually occurred. Instead, what is called the *Time to the Most Recent Common Ancestor* (or TMRCA) is the preferred estimate of age.

Of course there is disagreement among statisticians about how best to calculate TMRCA. I am in the dark as to how any of these calculations are done, but I know one popular method is to count what the statisticians call the “nodes of the phylogenetic tree” and presumably factoring in mutation rates, to arrive at an age when two SNPs branched apart. Another which some argue is better, is by means of an analysis of variance, but while I remember this from “Stats” half a century ago, I don’t know how they do it these days.

There is also the question of what unit of time to use? While some of us like to know the TRMCA in generations, there is a problem with this rather than simply counting the years because no one can agree how long a generation is — some argue for 25 years, others for 30 and others for variations on the theme.

Jumping to Conclusions

In 2000 an article was published in *Science* by a team of very respected population geneticists led by Ornella Semino and her colleagues in which it was stated:

We estimated the age of M173 by using the variation of three microsatellites, namely DYS19, YCAIIa, and YCAIIb. Although an estimate of ~30,000 years for M173 must be interpreted cautiously, it is consistent with our hypothesis that M173 marks the Aurignac settlement in Europe or, at least, predates the LGM. The polymorphism M170 represents another putative Paleolithic mutation whose age has been estimated to be ~22,000 years⁴.

⁴ Ornella Semino et al.: The Genetic Legacy of Paleolithic *Homo sapiens sapiens* in Extant Europeans: A Y Chromosome Perspective , *SCIENCE*, 10 November 2000 Vol 290 - www.sciencemag.org. Footnotes in the original have been deleted here.

To translate, M173 is the SNP which defines yHaplogroup R1b and M170 defines yHaplogroup I. If not at the time of writing then shortly after Semino's paper was published it was realized that R1b did not leave Asia and that it was a descendent clade, R-M269 (at that time known as R1b1c) which was the first to enter Europe. Even so, Semino was here suggesting that the first *Homo sapiens* men to enter Europe were R1b and that this was the y-haplogroup of the Aurignacian people. She adds that yHaplogroup I also predated the LGM, consistent — as we would later realize — with the advent of the Gravettian culture.

It still seems to be widely accepted that Haplogroup I arrived with the Gravettian culture and sought refuge from the LGM in the Balkans⁵. It is also still commonly believed, as Semino *et al.* suggested — that the first women in Europe belonged to mtHaplogroup U (probably more accurately, U5) and that the women who accompanied the Hg I men into Europe belonged to mtHaplogroup HV. This later split into two separate haplogroups, H (which is now the largest European mtHaplogroup) and Hg V did not expand much until the LGM.

So we are fairly confident that the people in Europe before the LGM included the Gravettian yHg I men and mtHg HV. We are also fairly confident that the first women in Europe, the Aurignacian women, were mtHg U. However, as yet we really have no idea who were the first men who accompanied the Aurignacian women into the cold wastes of Europe ~40 kya.

It certainly does not appear to have been R1b men as Semino *et al* and many others since have contended. Nor was it their descendents who painted the wonderful images in the caves of southern France and Cantabria as Oxford Ancestors and several other labs have told their customers who tested R1b. Calculations released late in 2008 indicate that R-M269 came to Europe *after* the LGM, during the Neolithic. Summing up the controversy which erupted when the new calculations were announced, geneticist Vincent Vizachero⁶ wrote:

The "new theory" is that R-M269 entered Europe during the neolithic era (i.e. sometime after 8 kya ago). Different people use slightly different methodologies, but most TMRCA estimates for R-P311 (which accounts for over 90% of R1b in western Europe) are in the range of 4-8 kya.

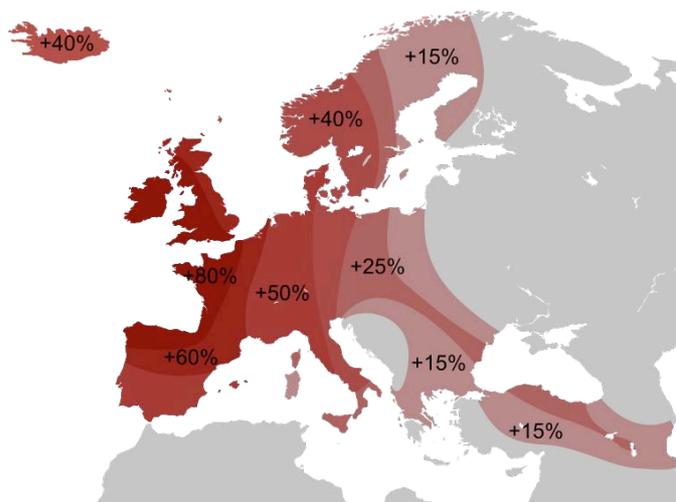
⁵ The last pre-Neolithic intrusion from the east into west-central Europe was the Badegoullian about 18,000BC. If these people carried Hg R, it would have been R1b which statisticians calculate would have been still in Asia and if they made it into a refuge during the LGM, it would have to have been the eastern one, not south of France.

⁶ Vincent Vizachero: TRMCA for R1b1, genealogy-dna@rootsweb.com, 28 Nov 2008

The SNP R-P311 (a.k.a. R1b1b2a1a) is “downstream” of R-M269 and the fact it accounts for such an overwhelming proportion of European men begs the question: *Where did they all come from?* The theory that they were the descendents of the first men in Europe, who found refuge in southern France and parts of Iberia during the LGM and then expanded north to re-populate Europe at the end of the LGM seemed nicely to account for this preponderance of sub-clades of Hg R in Western Europe. Books have been written assuming this scenario to have been true and many have created academic careers on the same presumptions.

Among those publications were two best-sellers, Bryan Sykes *Blood of the Isles* and the more erudite but still popular, *The Origins of the British* by Stephen Oppenheimer⁷. In this large work, Oppenheimer describes the origins of the many different peoples who have populated the British Isles, not only since the beginning of the Holocene but in the period between the end of the LGM and the onset of the Younger Dryas.

While this is a fascinating read, there are problems which mean that three years down the track, we must read it with caution. What follows is first, a brief explanation of Oppenheimer’s findings and second, the major problem which the re-calculation of the age of entry of “R1b” into Europe has created, not only for Oppenheimer but for all of us.



Distribution of R1b in Europe

The genetic landscape according to Oppenheimer

Oppenheimer’s book made the best-seller lists but among experts, he was criticised and his work called into question because he did not publish his findings in such a way that his work could be related to mainstream genetics or,

for that matter, replicated. A further criticism was that he employed only 6 y-STR markers, not SNPs, and failed to reveal which markers he had employed. These six markers, whatever they were, were sarcastically hailed as “Bikini haplotypes”, a term which seems destined to make it into the dictionary!

⁷ Oppenheimer, S: *The Origins of the British – A Genetic Detective Story: The Surprising Roots of the English, Irish, Scottish, and Welsh*; Carrol and Graf, NY. 2006.

Oppenheimer sorted these Bikini haplotypes into clusters according to whether their emigration from the refugia was during the late Upper Paleolithic or the Mesolithic. In explaining this, he relies heavily on a *pater familias* he calls *Rox*, a man who — like the dove from Noah’s arc — was the first to venture forth from the “Basque Refuge” as he calls it, and whose genetic signature on the “Bikini haplotypes” he labels the “**Basque Modal Haplotype**”. He also creates a fictional character called *Ruy*, son of *Rox*, who came out of the Basque Refuge after the Younger Dryas and whose DYS391 had meanwhile mutated from 10 to 11 so that his genetic signature now fitted the **Atlantic Modal Haplotype**. This, Oppenheimer says, accounts for 18% of British and 19% of Basques making it the single most common Y signature in Europe.

While the other shortcomings were sufficient to leave many genetic genealogists sceptical as to Oppenheimer’s “clusters”, the indication that R1b was a Neolithic intrusion, not a Palaeolithic continuation is a major blow to this otherwise respected scientist’s work because he, like many others, had accepted Semino’s assumptions that R1b preceded the LGM. In other words, whatever else their genes might have been, the men who re-populated north western Europe, including *Rox* and *Ruy*, could simply not have been any of those R1b clusters as Oppenheimer claims.

But if they were not from Haplogroup R, then **who were those men who endured the LGM in the Franco-Cantabrian refugia?** We know y-Haplogroup I was in Europe in time to experience the LGM but the evidence we have is that these people back-tracked down the Danube and Dnieper River system and found refuge in the Balkans and areas around the Black Sea — although there have been suggestions that some Hg I might have found refuge from the extremes of the Younger Dryas in that part of France and Spain.

There is a further very serious question which arises if we accept that the incursion into Europe by R-M269 was Neolithic: **if R-M269 did not arrive in Europe until the Holocene, how did R-M269 spread so rapidly and achieve such dominance**, and as a corollary, how did they achieve such dominance particularly on the Atlantic seaboard and in the British Isles? There is no archaeological record of a massive invasion which could have resulted in population replacement on such a grand scale. It seems more likely that some kind of mechanism of selection was responsible, that is, that a relatively small number of newcomers had one or more very powerful advantages which allowed their numbers to grow exponentially while the local, Mesolithic populations dwindled in comparison. Among those mechanisms which have been put forward to explain this are the introduction

of metals, Indo-European language, lactose tolerance, horses and wheeled wagons, some of which we will explore later... There are probably many more, all operating to create greatly favourable conditions for the newcomers, but whatever the combination, the mystery remains.



Two pioneering Neolithic cultures were dominant in Europe, the *Linearbandkeramic* or LBK, as it is more commonly known, and the *Impressed Cardial Ware*. The latter was a Mediterranean culture so we will leave it until the next Unit. However, for the moment we will deal with the LBK as a possible carrier of “R1b” (probably R-P311 or R-M269) into Europe.

This culture flourished from about 5500 to 4500 BC and fits the time frame described by the recent calculations for the TMRCA of those clades of R1b. At its maximum extent it ranged from the Paris basin in the west to the Vistula and upper Dniester rivers in the east, and stretched southward to the Danube and down into eastern Romania. The culture thus extended over 1,500 kms and took 360 years to encompass these limits with heaviest concentrations on the middle Danube and on the upper and middle sections of the Rivers Elbe and Rhine.⁸ Although probably not the first farmers in Europe, the LBK was certainly the most far-reaching and most influential.

These LBK people originated somewhere near the Carpathian Mountains, entering Europe up the valleys of the major rivers, including most importantly the Danube. Fairly recent research indicates that genetically they were characterised by the aDNA mtHaplogroup N1a⁹ but, as Wolfgang Haak et al.¹⁰ discovered, haplogroup N1a is found these days in European populations at very low frequency — only about 0.2% — suggesting that the

⁸ Another interesting feature of the LBK is that it did not extend to Denmark or the seaboard of Poland and Germany. These all remained Mesolithic for much longer, seeming to prefer to exploit the Atlantic salmon runs.

⁹ Ricaut, F-X et al., Genetic Analysis of a Scytho-Siberian Skeleton and Its Implications for Ancient Central Asian Migrations, *Human Biology* 76.1 (2004) 109-125

¹⁰ This is the study in which aDNA was analysed from skeletal remains in German cemeteries. See Haak, W. et al: "Ancient DNA from the First European Farmers in 7500-Year-Old Neolithic Sites". *Science*, 310:1016-1018. 2005.



number of immigrant farmers was relatively small compared to the majority Mesolithic populations of the time.

Example of LBK pot (Wikipedia GNU)

Characteristics of the LBK

Agriculture came to Europe at the beginning of the “Atlantic climatic phase”, when temperatures were about 2°C warmer, humidity and precipitation were higher

(ranging from 8-22%) and winters milder than today. The “package” the earliest LBK people introduced included cattle, sheep and goats; pigs; dogs; emmer, einkorn and spelt wheat; peas and vetch; and flax. Of these, sheep and goats were predominant in more southerly regions, cattle in the north while pigs were more common in the west. While the diet was based principally on these domesticated species, some wild food resources — from 5 -20% — were exploited, including aurochs, wild pigs, red deer, horse, fish, apples and pears, and of course berries (mostly blackberries and raspberries).

The LBK people lived in long houses built in cleared areas usually within 500m or so of a river or stream. Commonly, these long houses were located close to each other, strung out along the river valley, suggesting a high degree of local cooperation and social integration. There is little or no evidence that these communities were organised in any particular power hierarchy — if



other, modern long-house societies can be any guide, someone would have been the “*primus inter pares*”, qualified more by age and experience than by birth or other feature.

A reconstruction of a LBK house (Photo by Hans Splinter, Flickr).

The LBK people lived in *long houses*, that is, communal dwellings,

occupied by several families who were probably inter-related¹¹. These longhouses were very large and massively built, rectangular in shape, oriented northwest-southeast, and most commonly divided into three sections. The centre part seems to have been the living space while the north-western section was reserved for animals and the south-eastern end, for grain and other storage. A village therefore, consisted in a number of long-houses strung out along the river, with the gardens behind them furthest from the river bank.

Archaeologists originally believed the LBK people employed “slash and burn” farming methods but recent evidence shows that far from farming a garden for a season or two and then moving elsewhere, these early farmers continued to grow their crops in the same gardens for decades and sometimes centuries, possibly without manuring. This is testimony to the richness of the deep loess soils deposited there in previous glacials.

Although for a long time the advent of farming to Europe was believed to have been a peaceful process, recent evidence points to this being a time of considerable violence, particularly in the more western regions and in later rather than earlier phases¹². Excavations have shown that many long houses were fortified by the construction of U or V-shaped ditches around them. Whether this was to defend the inhabitants from other LBK communities or from marauding bands of Mesolithic hunter-gathers whose territories they had invaded, archaeologists have not been able to tell. What they have been able to tell, however, is that there were murders and even massacres — for example, at Tallheim on the Middle Rhine 34 skeletons were found in a mass grave. They all had wounds on the back of the head caused by LBK stone axes or adzes or arrows lodged in their backs suggesting they were victims fleeing from their attackers. In one site near Stuttgart, the skeletal remains are more robust than those usually found in LBK graves, possibly indicating these were also the bodies of indigenous Mesolithic people. On an even darker note, there is occasional evidence of cannibalism and of human sacrifice, especially of women and children....

From whatever cause — violence, cult sacrifices, disease or even climate change — the LBK population declined drastically and the trade connections almost disappeared. By about 4,800BC new “daughter” cultures were emerging, including the Rössen in Western Germany and the Netherlands, the

¹¹ Long-houses occupied by present-day pre-literate peoples — eg, in Kalimantan — are usually roughly divided into compartments each of which serves as the home of an individual family. LBK longhouses seem to have been communal.

¹² This is just a speculation, but if LBK were the carriers of clades of R1b, this violence might help explain why R is so much more plentiful on the Atlantic seaboard.

Villeneuve or Saint Germain in France, and in the eastern-most reaches of the old LBK, the Lengyel. Even more important were the Corded Ware, Funnel Beaker and Bell-Beaker cultures which emerged and, loosely speaking, filled the void left by the old LBK.

The next cultures which might have seen the expansion of Haplogroup R-M269 were first, the Corded Ware people and following them, the Urnfield — although this latter is more a complex of localized cultures which had in common only the method of burial¹³ from which archaeologists derived the name. Following these, the next major and most serious contender as the culture responsible for the massive expansion of the R-M269 haplogroup in Europe was of course the Celts. But before we look at any of these, we must go back and consider the Neolithic from several other points of view. However, no matter how you view it, the Neolithic is a complicated and even today, little understood period in human history.

¹³ Cremation followed by burial of the ashes in urns. Some archaeologists prefer to call this the *Urnfield Complex*.