Musings on the Convergence of Deep and Recent Genetic Genealogy in the Briese Surname Group

David Briese (July 2015)

Summary

This paper examines how a consideration of the deep genealogy, based on single nucleotide polymorphisms (SNPs) of Y-DNA, of a surname lineage can inform the relationships inferred by a study of more recent genetic differentiation, as measured by short tandem repeats (STRs) of the Y-DNA. In the case of the Briese Surname DNA Project, SNP studies have identified a putative new haplogroup (R-A5588) which currently includes Briese and Klassen family lineages. This haplogroup, probably of north Germanic origin, is of the order of 1200-1500 years old and provides a "fixed point" from which to interpret more recent genetic relationships within one of the three Briese surname lineages. These latter relationships date back to a common ancestor who lived some 500 years ago, leaving a relatively small "black hole" of less than 1000 years in the genetic history of this lineage.

Introduction

The Briese Surname DNA Project [1] comprises individuals who are seeking an answer to the questions "Who am I?" and "Where do I come from?" As a starting point we have a common surname, but have generally exhausted the paper trail of traditional genealogy without getting too far down our ancestry path (3-6 generations) [2,3]. Only one pair of the Briese group could show a relationship from paper records and, while several originated from the same area in Prussia, a common ancestor could not be identified. Testing of Y-DNA samples using FTDNA [4], however, showed that six of the group members formed a related cluster with a common ancestor who lived some 500 years ago [5]. Three members were only very distantly related to this cluster, showing the complexity of origin of the Briese surname [6]. Genetic studies have clearly added value to traditional genealogy in this case. However, to date we have centred our studies on comparing relatively recent genetic differences between individuals and trying to find where they might merge to identify common ancestor. This raises the question of what might be gained by approaching the Briese lineage from the "deep end", i.e. by following our line of descent from much earlier ancestors and trying to find where and how it branches into the Briese lineage. This article looks at this convergence of deep and recent genealogy.

Measuring genetic difference: STRs and SNPs

Genetic genealogy has been advancing rapidly in recent years as more and more people provide DNA samples for testing and the methodology of extracting data and analysing it becomes more sophisticated. Traditional genealogy is largely based on researching the patrilineal inheritance of surnames and, consequently, is more aligned to the genetic study of Y-DNA, passed from father to son. Currently, Y-DNA studies are based on two types of gene mutation - short tandem repeats (STR) and single nucleotide polymorphisms (SNP).

Without going into too much detail, STRs are short sequences of DNA that are repeated a number of times. Many loci on the Y-chromosome contain such repeated regions and some of these have been selected as markers to create a genetic profile for an individual (haplotype). The numbers and pattern of repetitions within these haplotypes (based on 12, 37 or 67 markers) is characteristic of the lineage. Structural mutations can occur through the loss or addition of a repeated region over time, leading to differences between individual

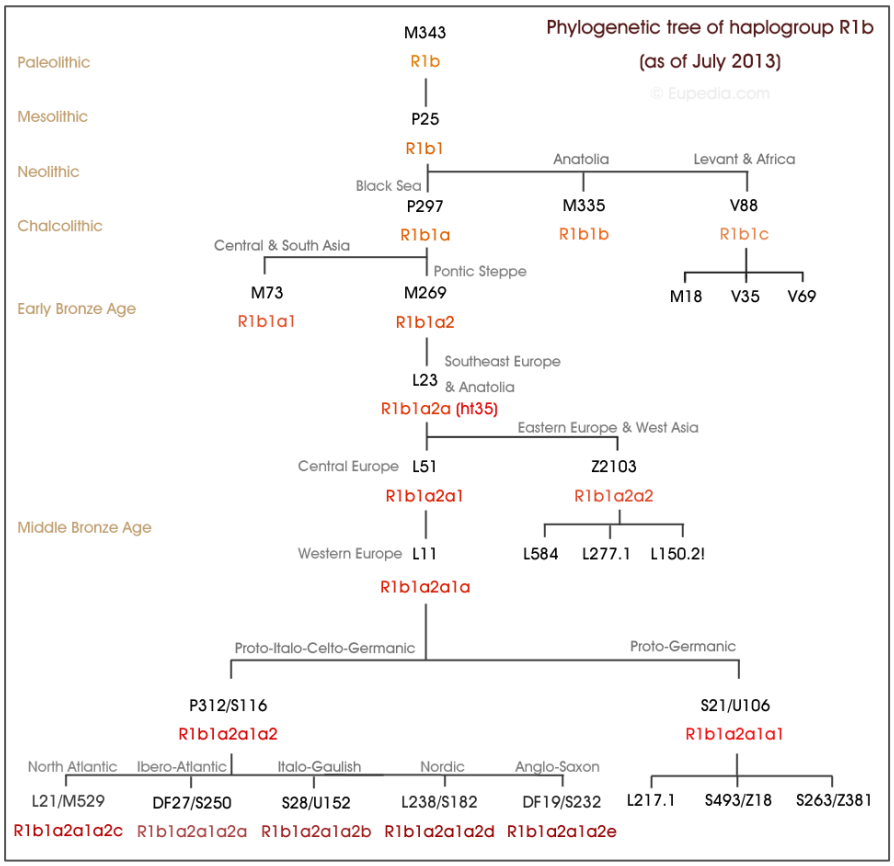

Fig 1. Deep genealogy – the tree of the R1b haplogroup (from Eupedia [8]). The Briese lineage is from M343>P25>P297>M269>L23>L51>L11>U106>Z381 and then continues on.

haplotypes and the analysis of these differences can inform us how closely related two or more individuals are. STRs are useful in defining paternal lineages up to 15 generations [7], which takes us back as far as the time when surnames were being adopted (500-800 years ago, depending on region). Some inferences can be made about more distant genealogy, but this becomes increasingly "fuzzy" due to convergence and reversal of STR mutations over time.

On the other hand, SNPs are the tool of choice for investigating deep genealogy, the study of the lineages of descent from the original common ancestor of modern man (aka "Adam") who apparently roamed the plains of Africa some 70,000 years ago. SNPs occur when a single nucleotide within a gene sequence is replaced (i.e. one of the four nucleotides that form the building blocks of a strand of DNA is replaced by another). As these replacement mutations occur in Y-DNA, different lineages develop from the founding male. Some become extinct, some proliferate and branch further as subsequent mutations occur, forming new haplogroups (groups that share a common ancestor with the same SNP mutation). Haplogroups are defined by that SNP and, from the knowledge of mutation rates and the pattern of SNPs, the pattern of formation of new haplogroups and the time scale from ancient to modern man can be determined (e.g. Fig. 1). Each new haplogroup identified narrows down the number of related people and brings the time when their most recent common ancestor lived closer to the present.

Thus, SNPs are generally used to track deep genealogy from the beginning to recent times, while STRs can be viewed as pushing back our genealogical knowledge from the present to more distant times. The interesting point is when and where they converge and, in the case of the Briese surname group, that moment has just arrived.

Deep genealogy of the Briese lineage

As mentioned before, members of the Briese Surname Group have had their y-DNA tested for STR markers. Haplotypes derived from these can predict to which of the 20 major haplogroups (clades) that define modern man we belong and, in some cases, can predict sub-clades of these. Thus, most genetic genealogy test results, such as FTDNA, provide not only the individuals haplotype, but a prediction of their haplogroup and a map of haplogroup migrations (see Fig. 2). For all but one of the Briese group, this was R-M269 (defined by the SNP labelled M269 [9]). This is not surprising, as it is the dominant lineage in Western Europe today. The R haplogroup (Indo-Europeans) had arisen in Central Asia some 24000 years ago and the R-M269 clade formed from it 5-7000 years ago. Archaeological evidence suggests that, some 4500-6000 years ago, people of this clade (including our Briese ancestors) swept out of western Asia and up the Danube River. Over time, they displaced much of the existing western European population by virtue of their technical superiority in weapons and materials [8].

This is still a long time in the past and, being a bit unsatisfied with this, in 2011 I had my DNA sample re-analysed by FTDNA, using the Deep Clade-R test [4]. This test looked for the presence of particular SNPs, which is the only accurate way of identifying your haplogroup. The results placed me (and the Briese cluster) in R-L48, a much more recent haplogroup (more specifically, this is R-M269>U106>L48). Around 4-4800 years ago, R-M269 branched into southern (P312) and northern (U106) haplogroups [10, 11]. By tracing paternal lineages back to their furthest known ancestor, our U106 lineage appears to have arisen in the north-western corner of Europe (Netherlands and north-west Germany) and is considered to be a proto-Germanic branch of the Indo-European family tree [8]. L48 is a more recent clade descended from U106 (but still over 3500 years old [10]) and seems to be centred more on northern Germany.

This is where things stood for a while, with a large black hole between the period from 3500 to 500 years ago. However, the technology of DNA testing is continually advancing and, a few years ago, FTDNA offered their Big Y test [12]. This takes a "whole of DNA" approach and can identify thousands of known SNPs and also novel variants (potentially new SNPs). I submitted my DNA

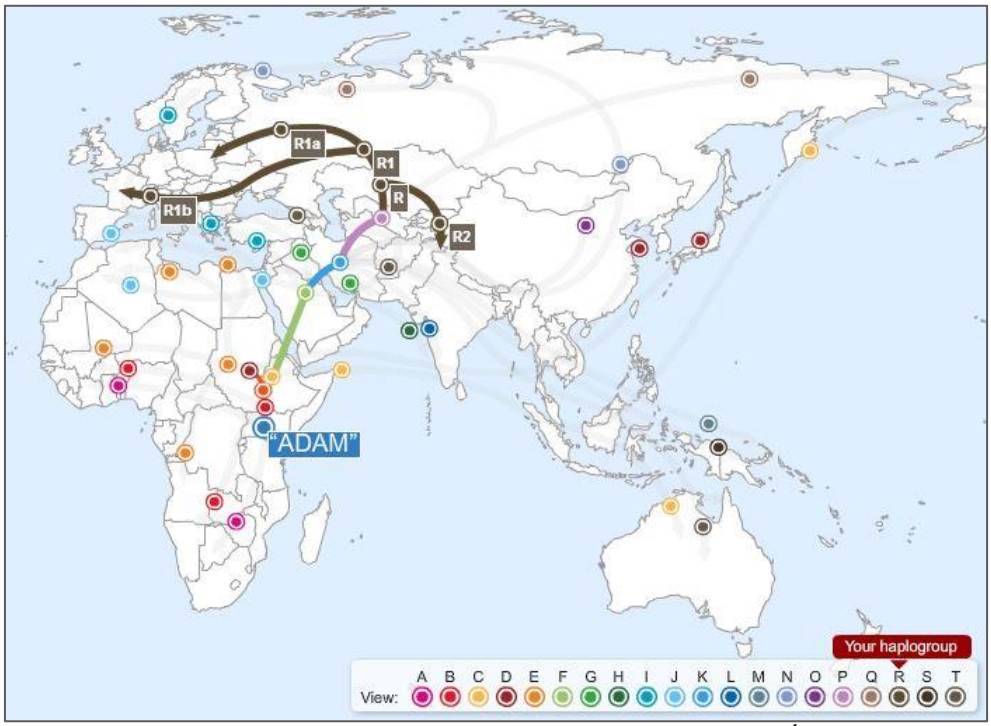

Fig. 2. Migration map of the R haplogroup (from FTDNA [4])

sample to the Big Y test in 2015, with very interesting results. Firstly, it refined the Briese haplogroup even further to Z326, whose origins are around 2200-2800 years ago [10]. Again this probably arose in the region of northern Germany, but also occurs amongst British men (probably as a result of the 4th-5th century AD invasions of England by Angles and Saxons and the 9th century invasions by the Danes).

Recent studies show that Z326, itself, comprises many more recent sub-clades [13], some yet to be formally identified. In fact, the test showed that the Briese lineage shared a number of novel variants with another member of haplogroup Z326 from the Klassen family. This degree of overlap between Briese and Klassen is sufficient to raise a putative new haplogroup (A5588) based on one of the novel SNPs (Fig. 3). The preliminary time estimate (from SNP mutation rates) for the origin of A5588 (and hence the time when Briese and Klassen shared a common ancestor) could be as little as 1100 years ago [10]. As Klassen has also had an STR test carried out, I was able to include that in the analysis of our Briese STR markers. This gave a slightly lower estimate of a common ancestor of between 800-1100 years ago. More importantly, though, as the Klassen-Briese relationship is certain (as shown by uniquely sharing a number of novel SNPs), it provides an ancestral and time "anchor" for all analyses of relationships within the Briese surname group. With only two members, it is not really possible to narrow the area of origin for A5588 from northern Germany, though the Klassen line can be traced back to similar West Prussian roots as the Briese line.

Where deep and recent genealogy meet

As mentioned, knowing with certainty the terminal SNP of a lineage provides an anchor for the study of more recent genealogy using STRs, i.e. it can help refine the predictions of genetic relatedness of individuals made using STR comparisons. As an example of this, the surname group contains one member that does not have the Briese name (Swanstrom), who joined on the basis of a close 37-marker STR match. While expanding this to a 67-marker test confirmed the closeness of other members of the Briese cluster, it showed Swanstrom to be genetically much more distant (see Fig. 4). Moreover, STRs infer a common ancestor, with varying degrees of probability - they do not prove it. As the Klassen-Briese SNP data was initially not available, it seemed that the apparent connection between Swanstrom and the Briese line at the 37-marker level may have been a "red herring" due to the small size of the first test.

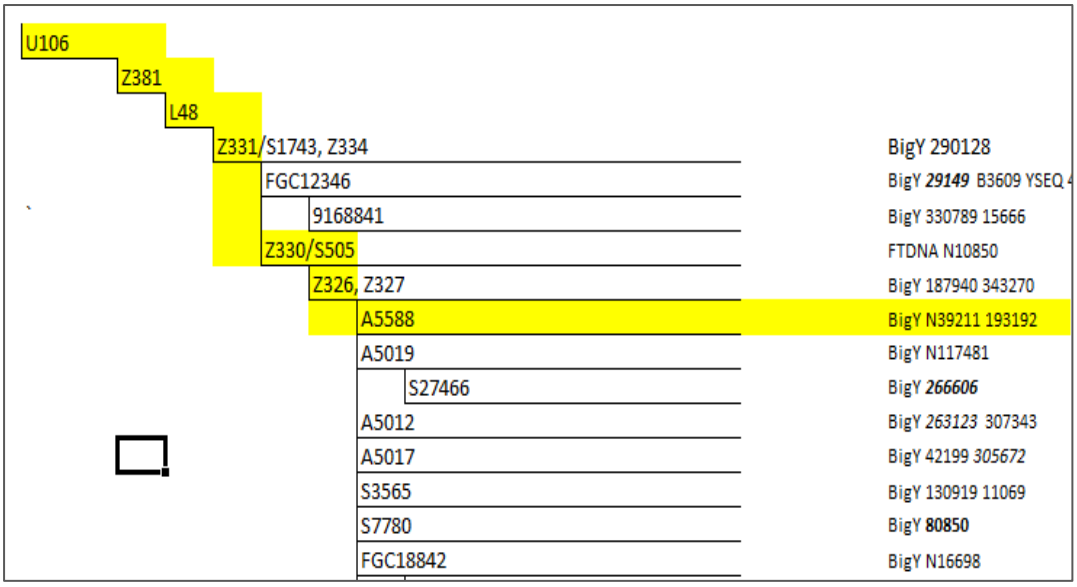

Fig. 3. The Briese lineage below R-U106, showing Z326 and the putative A5588 clade, where N39211 = Klassen and 193192 = Briese (from FTDNA Rb-U106 Project [13] - see text for details).

However, the new evidence from SNP testing "brings Swanstrom back into the fold". He is genetically closer to the Briese line than Klassen (whom we now know from the more accurate SNPs does share a common ancestor with Briese within haplogroup R-Z326). I recently did a network analysis of the STR markers of some members of the Z326 haplogroup, obtained from the FTDNA Rb-U106 Study Group [13] (Fig. 5). This showed that Swanstrom, and Briese and Klassen formed a grouping distinct from other Z326 members. This would seem to correspond to the putative A5588 haplogroup, though a Big Y test of Swanstrom's DNA would be needed to confirm it.

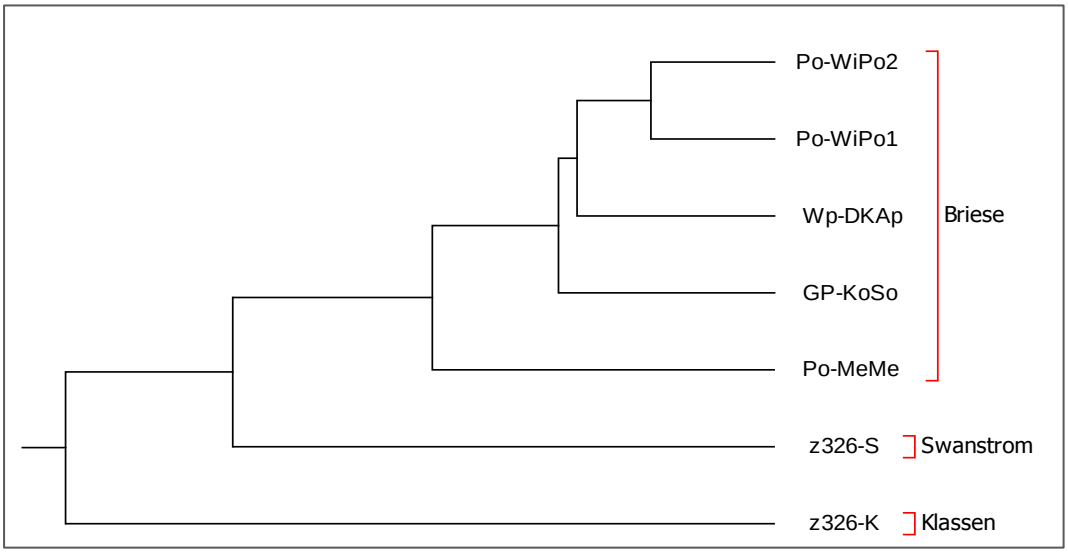

Fig 4. Phylogenetic relationships, based on genetic distances for 67 STR markers, between the known members of putative haplogroup R-A5588 (Briese and Klassen) and a possible member (Swanstrom).

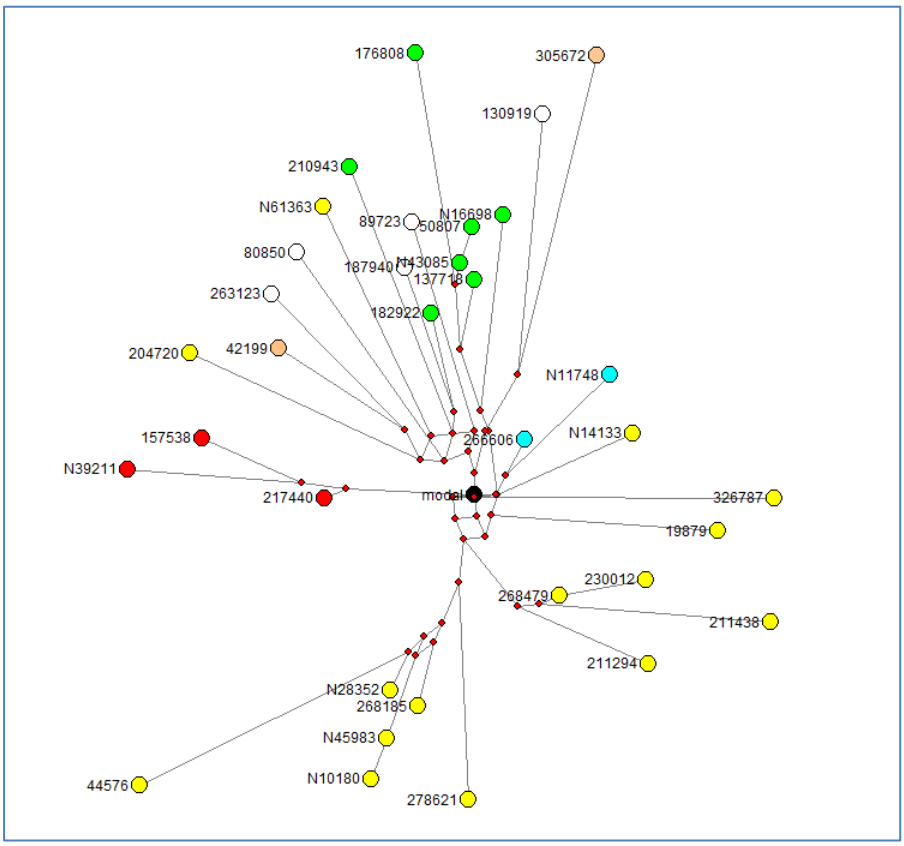

Fig 5. Network analysis of R-Z326 haplogroup (based on 67 STR markers). The red points indicate the putative new lineage A5588 (comprising Klassen, Briese and possibly Swanstrom). Other colours group other sub-clades of R-Z326.

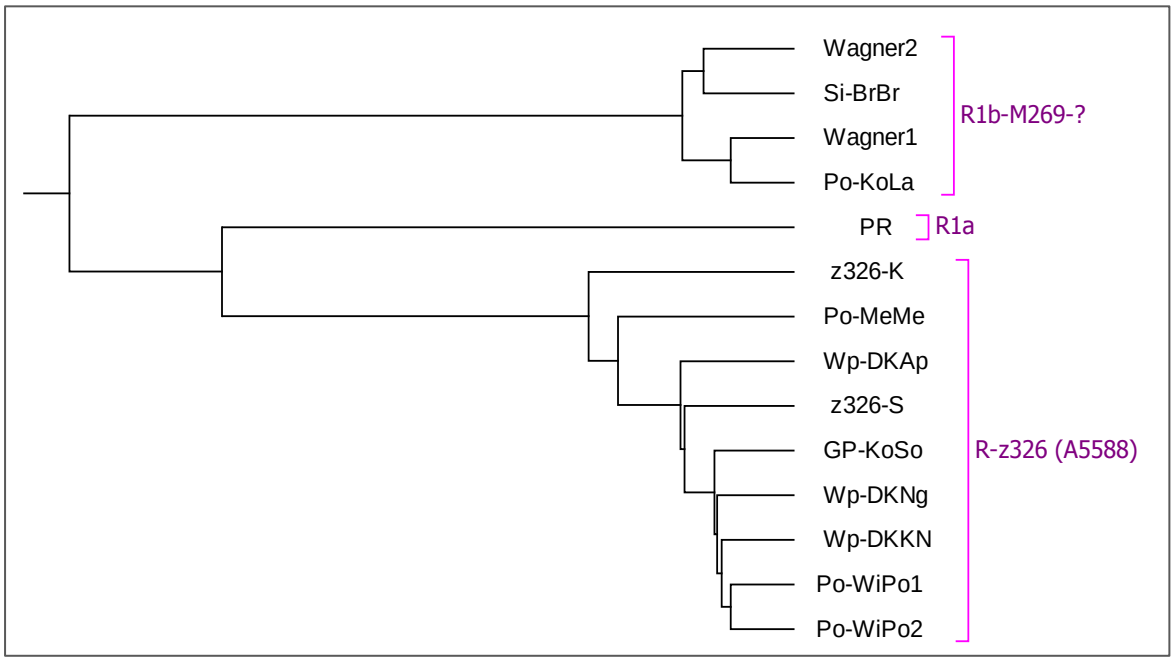

Fig. 6. Phylogenetic relationships (based on 37 STR markers) showing the three haplogroups within the Briese Surname Project (including other genetically-related surnames – Klassen (Z326-K), Swanstrom (Z326-S) and Wagner (see text for details).

Another interesting aspect is that there are now three distinct lineages in the Briese surname group, and the A5588 cluster, which now contains genetically related people with different surnames, is internal to the main cluster (see Fig. 6). This means that the other two lineages of Brieses are genetically more distant to this cluster than the A5588 haplogroup. In fact, one has a predicted haplogroup of R1a, which shared a common ancestor with the Briese R1b haplogroup over 20,000 years ago. The third lineage is probably closer to A5588, even though the analysis showed it as more distant [14], as they both fall within the R1b haplogroup (Fig. 6). This cluster is, in itself, of interest as there appears to be a close genetic link to a family named Wagner. Could a similar story to the A5588 cluster eventually emerge or would an analysis with 67 markers instead of 37 separate the names in the cluster? Clearly, the mixture of people bearing a common surname and those having a close genetic relationship are not necessarily the same, with events such as multiple name origins, adoption and so-called "non-paternity" events having an influence over the years. This is the case, even in a relatively rare surname such as Briese, with a fairly definite geographical origin.

In conclusion, the deep genealogical studies have provided some important new information on the Briese surname and the group of people bearing it. As deep and recent genetic studies converge from either end, the "black hole" of ignorance of our genetic heritage has been reduced to a period of no more than 1000 years (from ca 700 AD to 1700AD). In addition, the Big Y results ensure that there is a repository of known SNPs and novel variants from the Briese lineage available for further analysis. As the number of people who test positive for Z326 grows and more individuals from this clade have their DNA tested at a deeper level, more matches may be found and the gap in line of descent reduced even further. While the purpose of resolving the more recent relationships between the different Briese families remains the primary purpose of this project, understanding its deeper line of descent will not only help that aim, but helps show how people with the Briese surname fit into the broader human family.

Acknowledgments

All Y-DNA tests were carried out by Family Tree DNA. I would like to thank the participants in the Briese Surname Project for providing the DNA samples. Identification of the new sub-clade of R-Z326 from the SNP results was made by members of the R1b-U106 Y-DNA Haplogroup Project.

I would like to thank Wayne Kauffmann for his comments on an earlier draft of this paper.

Notes and references

1. FTDNA Briese Surname DNA Project www.familytreedna.com/public/briese/
2. Briese Surname Project www.gang-gang.net/briesesurnameproject/index.php
3. Briese Family Genealogy Pages www.gang-gang.net/briesegenealogy/index.php
4. Family Tree DNA website www.familytreedna.com
5. Briese DT (2014) Preliminary Phylogenetic Analysis of Briese Family Relationships. *Surname DNA Journal* www.surnamedna.com/?articles=preliminary-phylogenetic-analysis-of-briese-family-relationships
6. Briese DT (2008) Why is the family tree a silver birch? The meaning of the Briese name. www.gang-gang.net/briesesurnameproject/documents/pdf_files/briesename.pdf
7. Family Tree DNA (2015) Expected Relationships with Y-DNA STR Matches. www.familytreedna.com/learn/y-dna-testing/y-str/expected-relationship-match/
8. Hay M (2014) Haplogroup R1b. *Eupedia* www.eupedia.com/europe/Haplogroup_R1b_Y-DNA.shtml
9. The nomenclature for haplogroups has recently changed and they are now given the name of the basal clade plus the terminal SNP, e.g. R-M269 as used here. This is equivalent to R1b1a2 in the old system, which is still often used. The problem with the old system is a new number or letter code must be added for each new sub-clade, which becomes increasingly unwieldy, e.g. the Briese haplogroup R-Z326 would have been R1b1a2b1a1a3b2b2a1.
10. Time estimates rounded from the data and analyses of Iain McDonald (July 2015) - for the exact calculations and standard errors see Yahoo Groups: R1b1c_U106-S21 Haplogroups / Age Analysis / snp-ages.xls.
11. ISOGG (2015) Haplogroup R1b (Y-DNA) http://www.isogg.org/wiki/Haplogroup_R1b_(Y-DNA)
12. Family Tree DNA (2014) Introduction to the Big Y. https://www.familytreedna.com/learn/wp-content/uploads/2014/08/BIG_Y_WhitePager.pdf
13. FTDNA R1b-U106 Y-DNA Haplogroup Project www.familytreedna.com/public/U106/
14. This shows the limitations of STR comparisons in determining relationships beyond more than 15 generations past.