

Y-SNP rs34134567 Defines a Large Subgroup of Haplogroup G2a-P15

T. Whit Athey and James F. Wilson

Abstract

Evidence from case studies is presented demonstrating that Y chromosome SNP, rs34134567, also known as S126, defines a large subgroup of Haplogroup G2a-P15. S126/rs34134567 appears to be phylogenetically equivalent to the SNP, U8, which defines Haplogroup G2a3 in the current ISOGG phylogenetic tree (ISOGG, 2008).

Introduction

Y-chromosome Haplogroup G includes about 2% of all European men, about 10-15% of Middle Eastern men, and 30-70% in the Caucasus region. Its largest subgroup is Haplogroup G2a3, defined by the SNP U8 (ISOGG, 2008). Previously, Haplogroup G2a3 was thought to have three subgroups defined by the single nucleotide polymorphisms (SNPs), U16, U1, and M406 (ISOGG, 2008) as shown in Figure 1. The numbers in red in Figure 1 show the approximate relative size of each subgroup. The subgroup defined by the SNP U16 appeared on the first version of the 2008 ISOGG tree as in Figure 1, but in this study is shown to be a subgroup of the branch defined by M406.

The SNP rs34134567 is one of 858 Y-chromosome SNPs included on the Illumina 1M "gene chip." Testing of two people with this chip, one who was in Haplogroup G2a-P15, and the other who was in Haplogroup G2c-M377, revealed that they matched exactly on 857 of the 858 Y-SNPs. The single mismatch was on rs34134567. This discovery led to the present study to determine where rs34134567 fits into the G2a phylogeny. This discovery also led the company, Ethnoancestry, to develop a commercial test for the SNP, which has been named S126.

The current report presents the results of testing of S126 in several men who were known to be members of Haplogroup G2. The subjects were chosen to provide the maximum possible information on the phylogenetic placement of S126.

Methods

Six subjects were chosen for their membership in different branches of Haplogroup G2. Participants in the present study were designated as Subjects 101-106.

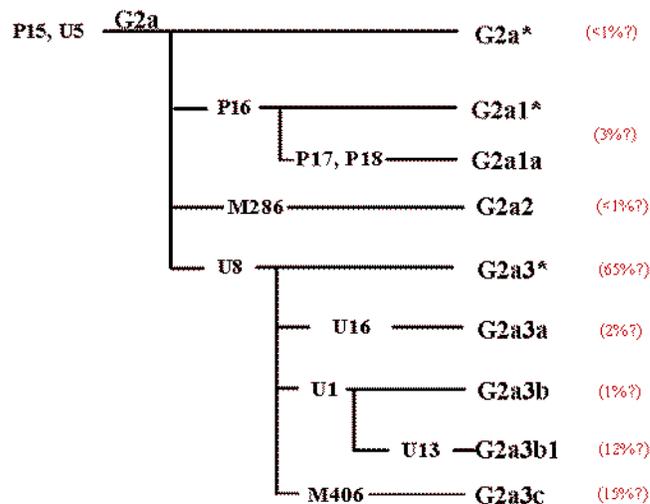


Figure 1 Phylogenetic chart for Haplogroup G2a according to ISOGG (June 2008 version).

For the present study, testing of S126 was carried out on Subjects 102-105 by Ethnoancestry (West Lothian, Scotland, UK). Y-chromosome markers were amplified by PCR with standard primers giving products from 200 to 500 bp in length. PCR products were then sequenced using dye terminator chemistry with electrophoresis on an ABI capillary sequencer. Alleles were called in the software package, Sequencher, by alignment with chromosomes of known allelic state (positive and negative controls).

Subjects 101 and 106 had their status at rs34134567 (same as S126) assessed using an Illumina 1M gene chip by DeCode Genetics (Reykjavik, Iceland).

Results

The status of all subjects on their most derived SNPs and their results from S126 testing are shown in Table 1.

Address for correspondence: T. Whit Athey, wathey@hprg.com

Received: September 15, 2008; accepted: October 12, 2008

Table 1 S126 Results for All Subjects

ID	Haplogroup according to Figure 1	S126 Status
101	G2c-M377	Ancestral
102	G2a1a-P18	Ancestral
103	G2a3a-U16	Derived
104	G2a3b1-U13	Derived
105	G2a3c-M406	Derived
106	G2a3-U8*	Derived

Discussion and Conclusions

Since subject 101 (Haplogroup G2c) was ancestral for S126, while Subject 106 (G2a) was derived, the branch defined by S126 must lie within G2a. Similarly, since Subject 102 (G2a1a-P18) was ancestral for S126 while subject 106 (G2a3-U8*) was derived, S126 must be located either within G2a3-U8, or in a position uniting G2a2 and G2a3. G2a2 is an extremely small group and no subject derived for M286 was available for testing, so we must necessarily ignore G2a2-M286 for now.

In order to resolve the location of S126 within Haplogroup G2a3-U8, results on S126 must be compared on subjects who represent all of the subgroups of G2a3. Subjects who were derived for U16, U1, and M406 were included in the present study. Prior to the present study, U16 was thought to define a subgroup of G2a3-U8 that was parallel to G2a3-U1 and G2a3-M406, but the one subject in the study by Ballantyne and Garvey (unpublished) who was derived for U16, in this study as Subject 103, was also found to be derived for M406 and S126. Therefore, U16 is actually downstream from

M406. Subject 104 is derived for U1/U13 and in this study was found to be derived for S126. Finally, Subject 105, also derived for M406, was found to be derived for S126. Subject 106 (as previously noted) is derived for U8, but not for any downstream SNPs (therefore in G2a3-U8*), and was derived for S126.

From these results and those discussed above, the SNP S126 is demonstrated to be approximately redundant with U8 in defining Haplogroup G2a3. At present we cannot rule out the possibility that some part of undifferentiated Haplogroup G2a3* will be ancestral for S126, in which case S126 would be slightly downstream from U8. We also cannot rule out at present the possibility that S126 would be sufficiently upstream from U8 that it would also be upstream from M286, which defines G2a2. In summary, S126 (rs34134567), within the presently available information, appears to be phylogenetically equivalent to U8 (see the latest version of the ISOGG 2008 tree).

The present study is a good example of the potential, pointed out by Turner (2008) in the Spring 2008 issue of this journal, for identifying new SNPs in the sets of Y-SNPs included on the gene chips that could be important in refining the structure of the Y phylogenetic tree. It will be very interesting to see the distributions of Haplogroup G2a3-S126 and its two subgroups in the Caucasus, Near East, Mediterranean, and Northern Europe, as more people are tested for this SNP.

Web Resources

<http://www.isogg.org/tree/>

ISOGG Y Phylogenetic Tree

References

ISOGG --International Society of Genetic Genealogy (2008) Y phylogenetic tree. See Web Resources.

Turner A (2008) SNPs on chips—A new source of data for Y-chromosome studies ('Satiabile Curiosity column'). *J Genet Geneal* 4:iii-iv.