'SATIABLE CURIOSITY

Hot or Not?

'Satiable Curiosity **is a column dedicated to the proposition that genetic genealogists are an untapped resource for resolving questions about DNA behavior--how DNA changes over the course of a few or many generations and how DNA patterns are distributed around the world. Some questions are so broad that it could take decades to arrive at a conclusion, yet others are narrow enough to answer in a shorter time frame, perhaps even within a semester or two for a student research project. The results may nonetheless be of considerable genealogical utility and scientific interest, worthy of publication in a technical journal.**

Some positions in the mitochondrial DNA control region have long been classified as "hotspots," due to the fact that mutations appear in multiple locations on a phylogenetic tree (Helgason, 2000).[1] Few of these studies, however, have conducted a detailed "property appraisal" of a region in the neighborhood of bases 514-524, which contains a two-base pattern (CA) repeated several times in a row.

Much like a Short Tandem Repeat (STR) on the Y chromosome, insertions and deletions of the CA repeat motif are common. One database lumps all variants together in a category called "gaps."[2] Table 1 divides those records into roughly equal parts, based on the status at the coding region mutation 12705. It is obvious that variations are present in substantial numbers throughout the global database.

Table 1. Number of Records According to Status at 514, Inside and Outside Haplogroup R

	514C	514 Gap	Ratio of C to Gap
12705C	838	221	0.21
12705T	736	348	0.32

[1] Sites 152, 16519, 16362,16093, 195, 16192, 16311, 150, 204, and 073 were especially variable.

[2] *mtDB - Human Mitochondrial Genome Database* http://www.genpat.uu.se/mtDB/, *viewed October 15, 2007*

The mutation 12705C is one of those that define Superhaplogroup R, comprising the major European haplogroups UK, JT and HV.

The region 514-524, technically part of HVR3, is now routinely sequenced for the genetic genealogy community, and much more data is currently available. This month's issue of JoGG contains an article by William R. Hurst entitled "Mitochondrial DNA Control-Region Mutations at Positions 514-524 in Haplogroup K and Beyond" (Hurst, 2007), in which he examines the behavior of this region in some detail.

Hurst is a keen observer of the Haplogroup K scene, and he noticed that mutations are not sprinkled here and there at random in all the subclades, as one would expect if forward and reverse mutations occur with such frequency and ease as to be labeled hotspots. Instead, there are cases where one version or another comes to predominate in certain subclades, almost to the point of predictability. Hurst proposes that this structure may instead derive from random fluctuations in pre-existing sources of the various alleles – to wit, heteroplasmy in the parent clades.

This immediately suggests an experimental approach, investigating the levels of heteroplasmy in subclades with varying percentages of a particular allele. Another haplogroup has a candidate for such a study, Haplogroup H. Many genetic genealogists have tested SNPs in the coding region to determine their subclade, and the haplogroup H project has enough members in

the H5 subclade and derivatives to form the nucleus of a study.

Table 2 Insertions (CA+) and Deletions (CA-) Found in Haplogroup H5[3]

Haplogroup	N	CA+	CA-	Percent CA-
H5	36	4	5	14%
H5a	27	0	19	70%
H5a1	14	0	11	79%

If the growing numbers of samples in H5a and H5a1 with CA deletions are not de novo mutations, the parent subclade H5 should exhibit a rather elevated level of heteroplasmy (compared to other haplogroups). The subclades, both with and without the deletion, might also retain substantial traces of the heteroplasmy, although not to the level customarily reported in conventional studies.

If this turns out to be the case, project members could also recruit close and distant cousins to participate in another phase of the study. This would provide insight into the heritability of heteroplasmy.

Demonstration of heteroplasmy in this particular population would not rule out the existence of de novo mutations in other cases. However, it would give some guidance in how to weight differences at 514-524, for phylogenetic studies as well as matches in the genealogical time frame.

Ann Turner

References

Helgason A, Sigurðardóttir S, Gulcher JR, Ward R, Stefánsson K (2000) mtDNA and the origin of the Icelanders: deciphering signals of recent population history. Am J Hum Genet, 66:999-1016.

Hurst WR (2007) Mitochondrial DNA Control-Region Mutations at Positions 514-524 in Haplogroup K and Beyond. J Genet Geneal, 3:47-62.

[3] Data derived from tables at *http://wiki.hmtdna.org/doku.php/results/basic* , viewed October 15, 2007