

Editor's Corner

A New YCC Phylogenetic Tree

The long-awaited update to the Y-Chromosome Consortium (YCC) phylogenetic tree has finally been published in *Genome Research* (Karafet 2008). The new tree was first previewed at the Family Tree DNA conference in October 2007 in a presentation by Michael Hammer, and a preliminary chart was provided to the attendees. However, the published version includes a few changes from that preliminary chart. For example, the published version includes new Haplogroups S and T, something that was only a proposal at the time of the presentation.

The Hammer group took the opportunity of the publication of the new tree to introduce dozens of new SNPs in the "P" series. Indeed, if these new SNPs had not been included, the "new" YCC tree would not even have been quite as up to date as the ISOGG (2007) tree, lacking as it is, some of the U SNPs and S SNPs that have not yet been published. With the new P SNPs, there is now considerably more resolution within several haplogroups than has been available previously. The Karafet article almost seems to be more about introducing new SNPs than it is about bringing the tree up to date, though obviously they go hand in hand. The Hammer lab has apparently been collecting new SNP discoveries for some time instead of publishing them as they were discovered.

Publishing dozens of new SNPs at once almost assures that the detailed testing information provided to support the proper placement of each SNP on the Y tree will mostly be lacking. One can only hope that the required testing has been done. However, past experience is not reassuring. This is the same group that for a couple of years insisted that Haplogroup I-M223 was not a subgroup of Haplogroup I-P38. Hopefully, this new batch of SNPs has been tested more thoroughly than was the case with P38 and M223.

Still, it is useful to have some of the leaders in the field get together every few years and share information and update the Y tree. The coauthorship of Peter Underhill indicates that the Stanford group provided many (but not all) of its latest discoveries. For example, the new SNP, M406, within Haplogroup G, announced in January by King et al. (2008), is missing from the new YCC tree.

Of course, discoveries are being made continuously, and the new YCC effort had to set a cutoff for inclusion at some point in time. But, this just points again to the need for an independent organization such as ISOGG to

maintain a Y tree with up-to-date information. It is unlikely that the YCC will be able to mount an effort similar to Karafet's every year, or even every couple of years, but the need for updates will only continue, especially with the data from the gene chips becoming available (see 'Satiabile Curiosity, this issue).

One very helpful inclusion in the Karafet article is the listing of all the new SNPs with the corresponding primers and location of the polymorphic sites. This will allow any commercial lab to develop tests for any of the new SNPs.

The new Karafet tree will undoubtedly help bring some order to the binary polymorphism offerings of the commercial labs. Up until publication of the new tree, a lab could choose its own tree to maximize the appearance of complete coverage by its own set of offerings. The new tree will undoubtedly force companies to provide many more SNPs in their commercial offerings, to the benefit of the community of genetic genealogists.

For ISOGG to succeed in maintaining a phylogenetic tree, it needs to have more cooperation from the main groups involved in population genetics and SNP discovery. At the beginning of the ISOGG effort, there seemed to be support. Liaisons were set up with the major labs and they were supposed to provide input and feedback to ISOGG. However, if one of the major groups saw ISOGG make a choice that was different from the choices they were in the process of making for themselves, by and large they kept it to themselves.

ISOGG is committed to a tree with the minimum of confusion for users, so naturally, with the publication of the new tree in Karafet (2008), ISOGG will be changing several haplogroup names to conform to the choices made by Karafet. Hopefully, in the future there will be more communication of discoveries to the ISOGG committee, so that when a major new effort such as Karafet's is published, there will be minimal differences with the then-current ISOGG tree.

If the ISOGG committee were fully supported by the professional community, it could better provide a valuable service as a clearing house for haplogroup phylogeny, SNP identification, and haplogroup names. Researchers who discovered a new binary polymorphism could see immediately at the ISOGG site the names that would be available for the new haplogroup.

Meanwhile, new SNPs are being announced or published almost every month. ISOGG's role will be to maintain a tree that is as up-to-date as possible, allowing us to see where each new SNP fits in. Some of these will cause additions or reorganizations of part of the tree, usually resulting in name changes for affected haplogroups. Slowly, the ISOGG tree will diverge from the benchmark tree provided in Karafet. The name changes will often be a little annoying, but that's just the way progress happens in the Y-chromosome field.

Whit Athey

Web Resources

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

ISOGG Y Tree

References

[Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF \(2008\) New binary polymorphisms reshape and increase the resolution of the human Y chromosomal haplogroup tree. *Genome Res*, 18:830-838.](#)

[King RJ, Özcan SS, Carter T, Kalfoglu E, Atasoy S, Triantaphyllidis C, Kouvatsi A, Lin AA, Chow CET, Zhivotovsky LA, M. Michalodimitrakis M, Underhill PA \(2008\) Differential Y-chromosome Anatolian influences on the Greek and Cretan Neolithic. *Ann Hum Genet*, 72:205-214.](#)