Understanding SNPs and Haplogroups

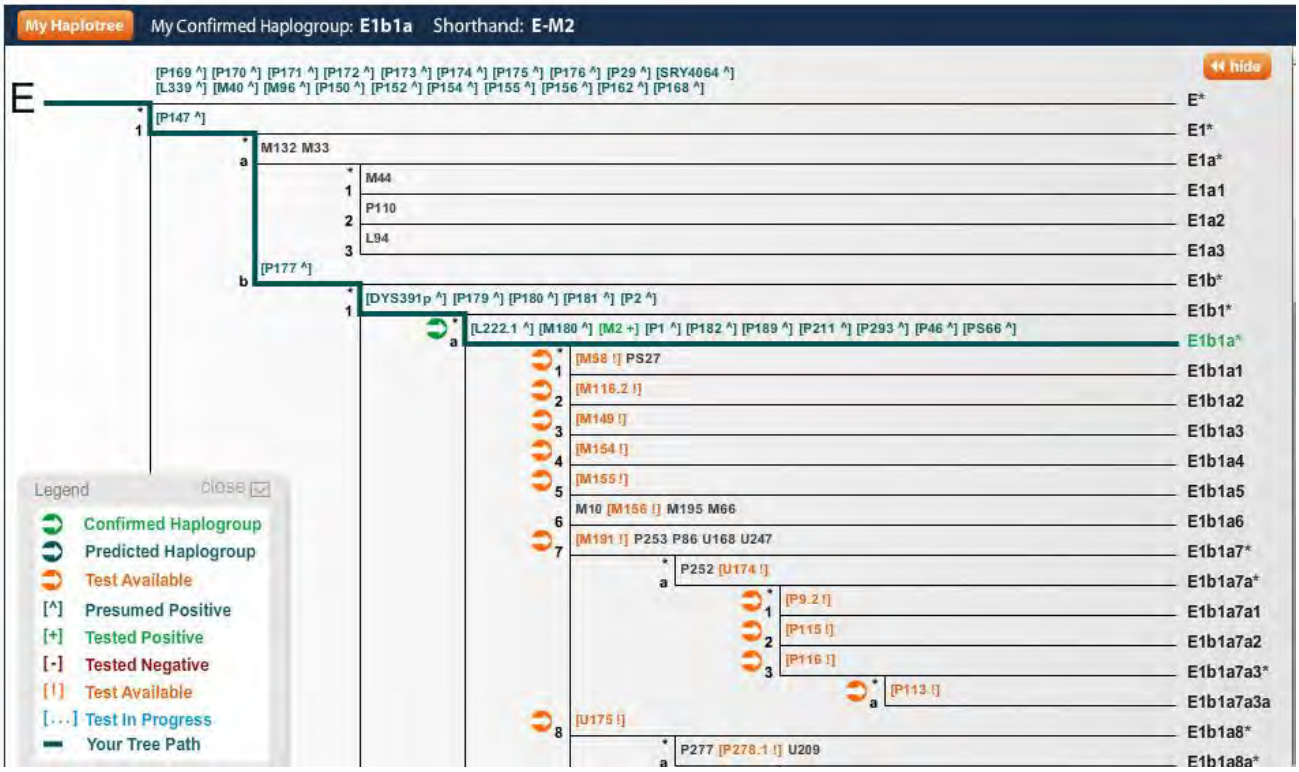
By Roberta Estes, copyright 2012

We all know by now that haplogroups define the different people of the world genetically. There are three or four main categories, European, African, Asian and Native American, which is a subset of Asian. Within these categories, there are different haplogroups. For example, in Africa, we find primarily paternal haplogroups E and A. For Native Americans, paternal haplogroups Q and C.

SNPs, or Single Nucleotide Polymorphisms, are what define haplogroups. These mutations are one "word" mutations in the long book of life on the Y chromosome. If you think of the Y chromosome as a street, with addresses, when we see something like SNP M3 or L222.1 as addresses on the street. Mutations at these addresses define a subgroup of haplogroup E1b1a, for example. These SNPs hold more information for us as well.

First, the letter tells us which lab discovered the SNP. The L means that the Family Tree DNA lab in Houston discovered that particular SNP. By way of trivia, the L was selected to honor a man named Leo Little, a genetic genealogist, who was instrumental in the discovery of the first SNP in that lab, and who died shortly thereafter. The number following the letter is the street address on the Y chromosome of where the mutation is located. In some cases, different SNPS are considered equivalent. For example, in haplogroup E1b1a, SNPs M2 and L222.1 are considered equivalent. If you are positive for any one of these, meaning you have the mutation, you will be positive for all of them. If anyone is discovered to NOT be positive for all of these mutations, then a new haplogroup subgroup has just been discovered.

The example shown below from haplogroup E1b1a shows several SNPS that are equivalent listed on each row. Not all SNPS have equivalent SNPS.



This seems all straightforward until you have a situation where two different haplogroups have defining mutations in the same location. This is seems confusing, but it really isn't!

Recently, someone who carried the L222.1 mutation in haplogroup E1b1a asked if they were related to someone in the J1c3d2 haplogroup who also carried a mutation at location L222.1. I was impressed with this man's observation!

The answer is no, they are not related since their common haplogroup separated, which, in the case of haplogroup J and E was about 50,000 years ago. Their common ancestor was haplogroup CR which was formed 45,000-55,000 years ago in the Ethiopia or Sudan region. Haplogroup F, the predecessor of J was formed about 45,000 years ago, and haplogroup DE, the predecessor of E was formed about 45,000-50,000 years ago in Africa. These divergent haplogroups are different streets, but just like every Main Street in every town, they all have the same addresses, because all haplogroups measure locations on the Y chromosome. Therefore, all haplogroups have the location of 222 - the only question is whether or not they incur a mutation at that address.

Locations such as 222.1 and 222.2 mean there are multiple insertions of DNA at that address (kind of like apartments). An insertion is one of three types of mutations. A deletion is another type and is often found in mitochondrial DNA sequences. The more typical mutation is when the "normal" or historical nucleotide is replaced by a different one. Think of everyone's house painted white on your street, but yours is painted red.

Mutations are the guideposts that provide our genetic history and differentiate us from each other. SNPS define our haplogroups and subhaplogroups which in turn tell us where our ancestors were and when. As we continue to discover new SNPS through the Walk Through the Y testing process at Family Tree DNA, what we are able to discern becomes more and more refined, in some cases now below the 5000 year threshold. One day, SNPs and STR testing, which are your individual marker results, will overlap, providing us with a smooth transition from our "family mutations" back in time indefinitely.

You can view the current haplotree of the Y chromosome and haplogroups at http://www.isogg.org/tree/index.html.