F18 Trouble With Y: Tribal Populations Cannot Be Pooled

Charles H. Brenner, PhD, 6801 Thornhill Drive, Oakland, CA 94611-1336; and Jami Johnson, LLD*, Federal Public Defender, 850 W Adams Street, Phoenix, AZ*

After attending this presentation, attendees will have learned that with current population data, a reasonable statistical evaluation of Y haplotype matching evidence in a Native American context is impossible. Therefore, its use at trial is precluded.

This presentation will impact the forensic science community by alerting attendees to appreciate: (1) the inadequacy of pooled Y haplotype population data in place of tribe-relevant data for tribal populations such Native Americans; and, (2) the confused and therefore confusing nature of a formula that Scientific Working Group on DNA Analysis Methods (SWGDAM) recommends.

Native Americans account for a large share of individuals prosecuted by the federal government for sexual assault. Y haplotype DNA is often vital evidence in sex cases; however, because Native Americans constitute a small segment of the United States population, too little attention has been given to the vital and distinguishing fact of tribal population structure, though the work of Hammer is important and helpful.[1]

In several recent Arizona sexual assaults, a Y haplotype match was featured as evidence. No population data was available for Y haplotypes of the relevant tribe. Instead, the prosecution offered a statistic based on pooled data from Native Americans in general. In one case, after a *Daubert* hearing, the court correctly understood that pooling Native Americans as a substitute for data from the relevant tribe "manufactures diversity" and therefore excluded the Y haplotype evidence.

Y haplotypes are patrilineally transmitted DNA data, copied as a unit from father to son, accurate except for the occasional (once per twenty generations) mutation. The Y picture is very different from the more typical autosomal forensic DNA (which comes from the 22 chromosome pairs excluding Y or X) primarily (though not only) because Y mutates much faster. (That's because the relevant sense of mutation here is any change in the transmitted unit, which is a single locus for autosomal but a block of about 17 loci for Y. Hence, the relevant mutation rate for Y is 17-fold faster.) Still, Y mutations are infrequent enough that a tribe arising from a small group 500 years ago will not have accumulated very much diversity. Hence, the chance for an innocent tribe member to be inculpated by a coincidental match is large. On the other hand, mutation is fast enough that over the 14,000 years Native Americans have inhabited the Americas, great diversity, mostly *between* tribes, has evolved. Therefore, a pooled database covering more than 100 tribes with common pairwise ancestors often 5,000 or more years in the past exhibits great diversity. The consequence, when it is used as it was in the Arizona cases, is to artificially exaggerate the strength of the evidence against the suspect. This is true even though the method of calculation of the matching statistic intends to be quite conservative by making statistical allowances.[2] But statistical allowances may not and do not compensate for the fundamental statistical sin of using population data that is not approximately representative of the relevant population.

The matching chance claimed from the pooled North American data was 1/35 — one chance in 35 that a random innocent man, if that's what the suspect was, would match. The above theoretical discussion explains why the procedure is wrong, but is it a genuine practical concern that the very modest-seeming prosecution 1/35 claim really overestimated the chance of a random match? Research from Finland says yes.[3] It reports a remarkable lack of diversity among men in some regions of Finland. For example, among the men in a region a few hours drive from Helsinki, two random men have one chance in ten to match — considerably larger than the supposedly generous estimate of one in 35. The Finns are not even an obviously tribal people. At a minimum, it is appropriate to regard evidence assessment based on the overall Native American database as merely random numbers.

Reference(s):

1. Hammer M., Redd A.J. (Jan 2006) *Forensic Applications of Y Chromosome STRs and SNPs*. Technical research report prepared US DOJ grant and submitted to the DOJ.
2. SWGDAM Y-chromosome Short Tandem Repeat (Y-STR) Interpretation Guidelines. http://www.fbi.gov/.../fsc/oct2009/standards/2009_01_standards01.htm/.
3. Palo *et al* (2007) High degree of Y-chromosomal divergence within Finland – forensic aspects. *Forensic Science International: Genetics*.(2007)1:120-124.

Y Haplotypes, Native American, Forensic Mathematics