

Large-Scale Genetic Genealogy Privacy Concerns



I've been working on a presentation regarding the future of genetic genealogy, and one aspect of that future is the ability to trace DNA (SNPs, mutations, haplogroups, etc...) through recent history as the result of combining extensive genomic sequencing with massive family tree information. Although the ability to do this will have many uses (both for genealogy and for personalized medicine), it will also raise a number of privacy issues, as a recent paper suggests.

A New Privacy Study

In "[Inferential Genotyping of Y Chromosomes in Latter-Day Saints Founders and Comparison to Utah Samples in the HapMap Project](#)," author [Jane Gitschier](#) uses a combination of FamilySearch (<http://www.familysearch.org>) and Sorenson Molecular Genealogy Foundation (<http://www.smgf.org/>) to elucidate the Y-chromosome signature of two founders of the LDS Church. Gitschier then used that information to determine whether anyone who contributed DNA to the HapMap project was related to these individuals via the Y-chromosome (none appeared to be). However, Gitschier was able to predict the surname of many of the HapMap participants using these databases.

This research is related to two posts I wrote last October about using genetic genealogy results to determine surnames ("DNA Could Reveal Your Surname, Of Course," and "More On Revealing Surnames Using Genetic Genealogy"). I first mentioned this research last September when I highlighted some of the most interesting abstracts submitted for the American Society of Human Genetics' November meeting (see [here](#)).

The Abstract:

"One concern in human genetics research is maintaining the privacy of study participants. The growth in genealogical registries may contribute to loss of privacy, given that genotypic information is accessible online to facilitate discovery of genetic relationships. Through iterative use of two such web archives, FamilySearch and Sorenson Molecular Genealogy Foundation, I was able to discern the likely haplotypes for the Y chromosomes of two men, Joseph Smith and Brigham Young, who were instrumental in the founding of the Latter-Day Saints Church. I then determined whether any of the Utahns who contributed to the HapMap project (the "CEU" set) is related to either man, on the basis of haplotype analysis of the Y chromosome. Although none of the CEU contributors appear to be a male-line relative, I discovered that predictions could be made for the surnames of the CEU participants by a similar process. For 20 of the 30 unrelated CEU samples, at least one exact match was revealed, and for 17 of these, a potential ancestor from Utah or a neighboring state could be identified. For the remaining ten samples, a match was nearly perfect, typically deviating by only one marker repeat unit. The same query performed in two other large databases revealed fewer individual matches and helped to clarify which surname predictions are more likely to be correct. Because large data sets of genotypes from both consenting research subjects and individuals pursuing genetic genealogy will be accessible online, this type of triangulation between databases may compromise the privacy of research subjects."

My NEW eBook!

Download my FREE 28-page eBook:

"I Have The Results of My Genetic Genealogy Test, Now What?"

[Download Now](#)

Connect



[Email Me](#)