

Ancestry DNA Testing and Privacy

Haifa Kassis, MD¹ and Deborah A. Ferguson, PhD² / ¹Editor-at-Large, Crisp Writing, LLC, Boston, MA; ²Ferguson Biomedical, LLC, Stow, MA

The ancestry DNA testing market is booming. According to industry estimates, the number of people who had their DNA sequenced for ancestry analysis more than doubled during 2017 and now exceeds 12 million.¹ This rapid expansion has prompted many questions about the accuracy of these tests and their impact on consumers' privacy. Can DNA testing really pinpoint where someone's ancestors lived? Why do ancestry DNA results provided for the same individual vary between different companies? What happens to the genetic information and DNA samples after ancestry analysis is completed?

To address these questions and more, Dr Sheldon Krimsky, Professor of Urban and Environmental Policy and Planning at Tufts University, gave a presentation on "Ancestry DNA Testing and Privacy" to AMWA members and guests at the New England Chapter meeting on September 24, 2018. Dr Krimsky has published more than 200 essays and reviews on social and ethical aspects of science and technology. He is the author of 14 books, including *Genetic Justice: DNA Data Banks, Criminal Investigations, and Civil Liberties*. He also serves on the Board of Directors for the Council for Responsible Genetics and as a Fellow of the Hastings Center on Bioethics.²

Ancestry DNA Testing

As of October 2018, the International Society of Genetic Genealogy has listed 35 companies that provide ancestry DNA testing directly to consumers.³ Ancestry DNA testing relies on small variations within the human genome. Although 99.5% of the genome is identical from person to person, variations in single nucleotides—called single nucleotide polymorphisms (SNPs)—contribute to the remaining 0.5%. There are an estimated 10 million SNPs in the human genome, appearing in about 1 in every 300 nucleotides.⁴ A subset of SNPs, known as ancestry-informative markers,

have been detected at dramatically different frequencies among different geographic populations.⁵

Genetic testing companies are using this knowledge to develop proprietary databases of reference DNA samples obtained from people living in different geographic regions. By comparing a genetic sample provided by a consumer to those in their reference sample database, these companies determine what proportion of the consumer's ancestry-informative marker variations match those that are frequently found in a given geographic region. For example, if 60% of the consumer's DNA has variations that are found most frequently in Germany, there is a good chance that those regions of their DNA came from ancestors who lived in Germany. A given variation may frequently occur in one region; however, this does not mean that it never occurs in any other region. For example, a variation that is frequently found in reference samples from Germany could also be present in samples from Japan or Italy.

According to Dr Krimsky, tracing one's ancestry with genetic genealogy tests is largely done as a recreational activity, and there is a common misconception in the general public that ancestry information from genetic genealogy tests can be informative about one's race. He explained that social scientists consider race to be a self-identified variable that is influenced by physical traits such as skin color, eye color, and body form.⁶ Ancestry DNA testing can help consumers find out where in today's world other people share some of their DNA; however, this type of testing cannot determine where all of their ancestors lived in the past, what race or social groups those ancestors identified with, or how consumers should identify themselves today.

Dr Krimsky pointed out that consumers are often not aware of the science- and privacy-related issues that come into play once they voluntarily provide their unique genetic material and personal health information to for-profit

companies for genetic genealogy tests. Even if someone never takes such a test, his or her life might be impacted if a relative or a family member decides to do so.

Science-Related Issues

The results of ancestry DNA testing are heavily influenced by the quality and size of the reference database that is used for comparison. Due to limitations in technology and sample availability, it is not possible to sequence the DNA of our true ancestors who lived hundreds or thousands of years ago. To address this issue, genetic testing companies try to select reference samples from individuals whose families have lived in the region for several generations. However, other than statements that ancestry is validated through family documentation and that reference samples must be unrelated (to avoid bias), little is known about the requirements for inclusion or exclusion of reference samples in the databases.^{7,8}

Ancestry DNA testing results may also be influenced by the number of samples that are included from a given region. Although some genetic testing companies disclose the number and distribution of reference samples in their databases, it is apparent that all geographic regions are not equally represented. For example, the current reference panel used by AncestryDNA® (Ancestry.com LLC, Lehi, Utah) contains more than 16,000 reference samples, many of which were obtained from Europe, including 2,072 samples from Germanic Europe, 1,959 from Eastern Europe and Russia, and 1,519 from England, Wales, and Northwestern Europe.⁹ By comparison, other regions are not well represented, with only 41 samples from Northern Africa, 31 from Senegal, and 30 from Sardinia.⁹ Moreover, as Dr Krimsky discussed, relatively little is known about how the reference samples were obtained—were they random samples or opportunity samples collected as a part of another study? Anything other than a random sampling of the population could introduce bias.

According to Dr Krimsky, ancestry predictions require complex algorithms and statistical analyses. Differences in analytical methods performed using the same reference database would likely result in different ancestry predictions. It is even possible that different results could be obtained from the same genetic testing company as databases and algorithms are updated over time.

At present, ancestry DNA testing has little oversight by scientific associations and is not regulated by the US Food and Drug Administration. In contrast, the Food and Drug Administration does regulate the marketing of direct-to-consumer tests for genetic predisposition to certain medical diseases or conditions.¹⁰ Although ancestry DNA tests do not make health-related claims, the SNPs they employ may be located within the coding or noncoding regions of genes, and

many studies have linked polymorphisms with susceptibility to disease, severity of illness, and response to drug treatment.¹¹ Certain SNPs may not be associated with a disease now, but new research may uncover a link in the future.

Privacy-Related Issues

Since the introduction of direct-to-consumer genetic tests, many privacy concerns and potential risks to consumers have been raised. Nonetheless, laws and regulations have not caught up with industry yet. According to Dr Krimsky, current laws are not broad enough to protect consumers' privacy. The Health Insurance Portability and Accountability Act protects privacy if DNA samples are collected by health care providers (such as a doctor or a hospital), but it does not apply when DNA samples are sent to private companies by consumers. The 2008 Genetic Information Nondiscrimination Act makes it illegal to use genetic information to discriminate against employees or applicants for *health care* insurance, but it does not protect consumers against discrimination in the context of *life* insurance or *long-term care* insurance.^{12,13}

It may not be clear to consumers that the contracts they sign to have their DNA sequenced for ancestry testing often include clauses that give away their rights to profits from any future products developed by using their genetic information.

The ancestry DNA testing industry stores a unique set of data about its consumers. Even though companies may take extensive precautions to protect consumers' data, the risk of hacking and database breaches always exists.¹² In the absence of oversight, it often falls to consumers to learn about and judge the reliability of the company's internal policies for data safety and disclosure of hacking.¹⁴ Consumers may not be aware that companies usually include a warranty clause waiving them from liability in the case of information theft.¹³

Millions of consumers have paid companies to have their DNA sequenced. The resulting large-scale genetic databases can be leveraged to discover mutations that cause diseases or confer response or resistance to certain drugs. As genetic data are usually collected along with individual health information, genetic databases may be used to calculate mutation rates in certain demographic or age groups. For these reasons, many

ancestry DNA testing companies sell their consumers' data to pharmaceutical companies for research and product development purposes.¹² For example, 23andMe (Mountain View, California) provides consumers with the choice of opting into research conducted by academic, nonprofit, and industry organizations.¹² Pharmaceutical companies that have used such data include Genentech for Parkinson disease research¹⁵ and Pfizer for lupus and inflammatory bowel disease research.¹² It may not be clear to consumers that the contracts they sign to have their DNA sequenced for ancestry testing often include clauses that give away their rights to profits from any future products developed by using their genetic information.^{12,13} Further, companies also often retain the right to change their terms and conditions as they find appropriate and without consumers' consent or notification.¹³ Finally, if a company goes bankrupt, is sold, or merges with another company, its privacy policies may no longer be valid or legally binding.¹²

Most ancestry DNA testing companies emphasize that they anonymize their databases and share them with third parties only if the individual consumer opts in. Companies often let consumers opt out of data sharing and research whenever they want and allow them to ask for their entire data set to be permanently deleted at any time. Dr Krimsky pointed out, however, that anonymizing individual genetic data is a very challenging task. A highly skilled researcher with some knowledge about the family tree may theoretically be able to deanonymize the data.¹³ Moreover, if consumers opt out of data sharing and research long after their data had been sold to third parties, it could be very difficult to ensure that all shared data are destroyed.¹²

Recently, ancestry DNA testing companies have started to offer their customers the capability to search for lost or unknown family relatives. There is even a free, online, open-source amateur genealogy database (GEDmatch) that allows users to upload their unique DNA sequences and family trees to have them compared with other users' data on the site.¹⁶ These services may help locate estranged relatives who desire such a union; however, they may also uncover cases of infidelity as well as unknown or unwanted children, which may result in long-lasting family conflicts. Only a few companies warn their customers that their families may or may not want to know the information derived from the ancestry DNA testing and that their relationships with others may be harmed once the information is shared publicly.¹³

DNA sequencing data collected for ancestry analysis may also be requested by law enforcement agencies. Unless there is a court order, it is up to each company to decide whether it wishes to cooperate with law enforcement. Several of the leading ancestry DNA testing companies (eg, 23andMe) explicitly state that they resist law enforcement requests for individual

personal information as much as possible.¹⁷ In any case, it is unlikely that consumers will ever be asked to provide consent for access to their genetic information.¹³

Law enforcement officials, however, do not need a court order to access the publicly available GEDmatch database. Most recently, local police officers in Sacramento, California, used this database to identify the Golden State Killer, a criminal who terrorized California with a string of horrific rapes and homicides in the 1970s and 1980s.¹⁸ In this case, police had DNA evidence from the crime scenes, but investigators were unable to find a match until data from GEDmatch helped narrow down the possible suspects to a single family. As Dr Krimsky explained, officers then watched the house of a member of that family and collected a piece of his abandoned trash for DNA analysis. A match with the killer's DNA finally unlocked the identity of the elusive and unlikely suspect—a retired law enforcement officer.

Conclusions

Based on the popularity of ancestry DNA tests as holiday gifts last year,¹ it appears that many consider these tests to be a fun and harmless activity. However, the genetic information obtained from these tests is associated with serious medical and ethical issues that consumers may not be aware of. Better communication about the scientific accuracy and privacy implications of ancestry DNA testing is clearly needed.

Author contact: haifa.kassis@crispwriting.com

References

1. Regalado A. 2017 was the year consumer DNA testing blew up. *MIT Technology Review*. February 12, 2018. <https://www.technologyreview.com/s/610233/2017-was-the-year-consumer-dna-testing-blew-up>. Accessed October 25, 2018.
2. Sheldon Krimsky bio. Tufts University website. <https://sites.tufts.edu/sheldonkrimsky/biography>. Accessed October 25, 2018.
3. List of DNA testing companies. International Society of Genetic Genealogy Wiki https://isogg.org/wiki/List_of_DNA_testing_companies. Updated October 15, 2018. Accessed October 29, 2018.
4. What are single nucleotide polymorphisms (SNPs)? US National Library of Medicine website. <https://ghr.nlm.nih.gov/primer/genomicresearch/snp>. Published October 23, 2018. Accessed October 25, 2018.
5. Rosenberg NA, Li LM, Ward R, Pritchard JK. Informativeness of genetic markers for inference of ancestry. *Am J Hum Genet*. 2003;73(6):1402-1422.
6. Beckwith J, Bergman K, Carson M, et al. Using dialogues to explore genetics, ancestry, and race. *Am Biol Teach*. 2017;79(7):525-537.
7. AncestryDNA DNA reference panel. Ancestry.com website. <https://support.ancestry.com/s/article/DNA-Reference-Panel>. Accessed October 25, 2018.
8. Ancestry composition. 23andMe website. <https://www.23andme.com/ancestry-composition-guide>. Updated May 2018. Accessed October 25, 2018.
9. AncestryDNA Science Team. *Ethnicity Estimate 2018 White Paper*. Lehi, UT: Ancestry.com LLC; 2018. https://www.ancestrycdn.com/dna/static/images/ethnicity/help/WhitePaper_Final_091118dbs.pdf. Accessed October 25, 2018.

10. FDA allows marketing of first direct-to-consumer tests that provide genetic risk information for certain conditions. US Food and Drug Administration website. <https://www.fda.gov/newsevents/newsroom/pressannouncements/ucm551185.htm>. Published April 6, 2017. Accessed October 25, 2018.
11. Smart A, Bolnick DA, Tutton R. Health and genetic ancestry testing: time to bridge the gap. *BMC Med Genomics*. 2017;10(1):3. doi:10.1186/s12920-016-0240-3
12. Rosenbaum E. 5 biggest risks of sharing your DNA with consumer genetic-testing companies. CNBC website. <https://www.cnbc.com/2018/06/16/5-biggest-risks-of-sharing-dna-with-consumer-genetic-testing-companies.html>. Published June 16, 2018. Accessed October 25, 2018.
13. Krinsky S, Johnston DC; Council for Responsible Genetics. *Ancestry DNA and Privacy: A Consumer Guide*. Washington, DC: International Center for Technology Assessment; 2017. <https://sites.tufts.edu/sheldonkrinsky/files/2018/05/pub2017AncestryDNAPrivacy.pdf>. Accessed October 25, 2018.
14. Mooy MD. Commentary: your genetic data could be a hacker's jackpot. *Orlando Sentinel*. June 6, 2018. <https://www.orlandosentinel.com/opinion/os-ed-genetic-data-hackers-jackpot-beware-testing-20180606-story.html>. Accessed October 25, 2018.
15. New genetic associations for Parkinson's disease identified. 23andMe website. <https://blog.23andme.com/23andme-research/new-genetic-associations-parkinsons-disease-identified>. Published September 11, 2017. Accessed October 25, 2018.
16. GEDmatch.com terms of service and privacy policy. GEDmatch website. <https://www.gedmatch.com/tos.htm>. Revised May 20, 2018. Accessed October 25, 2018.
17. Transparency report. 23andMe website. <https://www.23andme.com/transparency-report>. Published May 15, 2018. Accessed October 25, 2018.
18. Jouvenal J. To find alleged Golden State Killer, investigators first found his great-great-great-grandparents. *Washington Post*. April 30, 2018. https://www.washingtonpost.com/local/public-safety/to-find-alleged-golden-state-killer-investigators-first-found-his-great-great-great-grandparents/2018/04/30/3c865fe7-dfcc-4a0e-b6b2-0bec548d501f_story.html?utm_term=.3408919f2583. Accessed October 25, 2018.

Introduction from Members Matters Section Editor: Melory Johnson

As one of the Local Networking Coordinators (LNCs) for my chapter, I have attended more than 20 networking events over the past 2 years. At every gathering, I am astounded not only by the eagerness of AMWA members to openly and honestly share their knowledge and experiences but also by the quality of our discussions, so much so that I began to wonder how we could share some of our local exchanges on a national level. As fate would have it, so did Jim Cozzarin, the *AMWA Journal* Editor! He too recognized the value of sharing information from the local level with the AMWA community as a whole—and what better platform than the *AMWA Journal*? So, with that, the Members Matters section was born.



I am very much looking forward to serving as editor of the new Members Matters section and to hearing from both AMWA newcomers and seasoned members about the journal-worthy topics shared at your local gatherings and events. If you have discussed a topic locally that you would like to share on a national level, please email me at melory.johnson@mjmedcom.com.

Melory Johnson, VN, President and Principal Medical Writer at MJ Medcom, LLC, is a freelance medical and health writer specializing in medical affairs and medical devices with more than a decade of experience in management and education.

Jump-start your career in regulatory writing.



AMWA EDUCATION
Write better. Write now.

www.amwa.org/regulatory123

