

Study: DNA websites cast broad net for identifying people

11 October 2018, by Malcolm Ritter



In this Friday, April 27, 2018 file photo, Joseph James DeAngelo, 72, who authorities suspect is the "Golden State Killer" responsible for at least a dozen murders and 50 rapes in the 1970s and 80s, is accompanied by Sacramento County Public Defender Diane Howard, right, during his arraignment in Sacramento County Superior Court in Sacramento, Calif. Authorities said they used a genetic genealogy website to connect some crime-scene DNA to DeAngelo. (AP Photo/Rich Pedroncelli)

About 60 percent of the U.S. population with European heritage may be identifiable from their DNA by searching consumer websites, even if they've never made their own genetic information available, a study estimates.

And that number will grow as more and more people upload their DNA profiles to websites that use genetic analysis to find relatives, said the authors of the study released Thursday by the journal *Science*.

The use of such databases for criminal investigations made headlines in April, when authorities announced they'd used a genetic

genealogy website to connect some crime-scene DNA to a man they then accused of being the so-called Golden State Killer, a serial rapist and murderer.

In general, such searches begin on a site by finding a relative linked to a DNA sample. Then sleuths can use other information like published family trees, public records and lists of survivors in obituaries, plus whatever they know about the person whose DNA began the process. They can build their own speculative family trees. Eventually, that can point to someone whose DNA is then found to match the original sample.

With DNA databases "you need just a minute fraction of the population to really identify many more people," said Yaniv Erlich of Columbia University, an author of the study.

Each person in a DNA [database](#) acts "as a beacon that illuminates hundreds of distant relatives," said Erlich, who is also chief scientific officer of the MyHeritage website.

His paper focused on Americans of European descent because such people are over-represented in DNA databases, which makes it easier to find relatives.

The researchers started with the 1.28 million participants on the MyHeritage site at the time they did the work. Most had a northern European genetic background. For each, they looked for relatives more distant than first cousins elsewhere in the database.

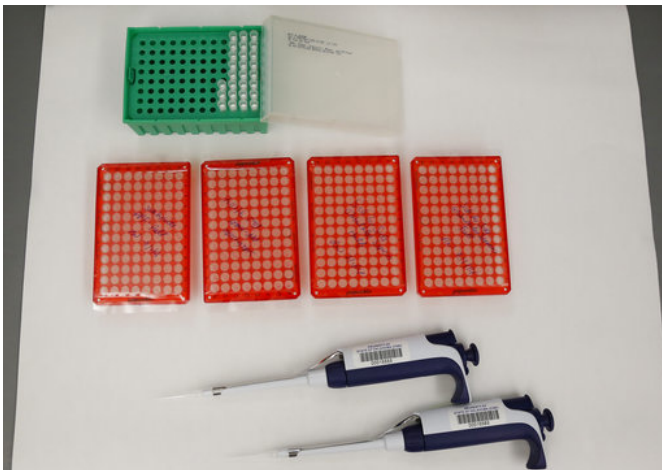
About 60 percent of the time, they found someone whose genetic similarity was at least equal to that of a third cousin, similar to the degree of relatedness that led to the Golden State Killer suspect. Third cousins share great-great-grandparents.

With some basic assumptions about what kind of data would be available for a criminal suspect, the researchers calculated they could pare down the possible identity of the initial person to just 16 or 17 people. That's limited enough that police could zero in with further investigation, Erlich said.

Erlich and his co-authors suggested that such searches could cast a broader net in the near future. A database with DNA profiles of just 2 percent of a population is enough to match nearly everybody with somebody who's as closely related as a third cousin, researchers said. From that, they calculated that the genetic profiles of about 3 million Americans of European descent could deliver the equivalent of a third cousin for more than 90 percent of that ethnic grouping.

Websites are getting very close to that, said Erlich, noting that MyHeritage now has more than 1.75 million participants. He said the [website](#) does not allow forensic searches.

Two DNA experts unconnected to the study said third and fourth cousins can both lead to identifications.



In this Aug. 1, 2014 file photo, tools used for DNA testing are shown in a DNA lab at the forensic science center of the Oklahoma State Bureau of Investigation in Edmond, Okla. About 60 percent of the U.S. population with European heritage may be identifiable from their DNA by searching consumer websites, even if they've never made their own genetic information available, a study estimates. That number will grow as more and more

people upload their DNA profiles to websites that use genetic analysis to find relatives, researchers said. (AP Photo/Sue Ogrocki, File)

"Because the average person has so many of these distant cousins, it becomes reasonably probable that one or more of them is in a publicly searchable database, even if only a small fraction of the U.S. population is included," Graham Coop and Michael Edge of the University of California, Davis, wrote in a statement to The Associated Press.

"The fact that most suspects could be identified in this way is predictable" from mathematical calculations, and the new paper provides a convincing demonstration, they said.

However, the work raises important policy questions, they said. Should anyone other than law enforcement be allowed to conduct such searches? And under what circumstances should they be permitted?

"How should we react to the fact that the decisions of our fourth cousins, whom one may never have met, affect one's privacy?" they asked.

In an interview, Edge noted that when people add their DNA profiles to a publicly searchable genealogy site, "they're not necessarily thinking about the genetic privacy of their [distant relatives](#)."

Amy McGuire, a professor of biomedical ethics at the Baylor College of Medicine in Houston, said that police searches using DNA and genealogy websites have sometimes pointed to an incorrect person.

"You would hope ... the victim of the false lead can be easily cleared" by providing DNA, she said. "But you still have some invasion into that person's personal life by being investigated."

Some people would say that's worth it to aid the cause of justice, but others "would find that very distressing," she added.

McGuire said there's an active legal debate about

whether police should be able to "go on a fishing expedition" using DNA genealogy websites without a warrant.

She recently published a survey that suggests most people support letting police search genetic genealogy databases. But support was much higher for investigations involving violent crimes and crimes against children than for nonviolent crimes.

More information: Y. Erlich et al., "Identity inference of genomic data using long-range familial searches," *Science* (2018).
[science.sciencemag.org/lookup/...
1126/science.aau4832](https://science.sciencemag.org/lookup/doi/10.1126/science.aau4832)

© 2018 The Associated Press. All rights reserved.

APA citation: Study: DNA websites cast broad net for identifying people (2018, October 11) retrieved 5 May 2021 from <https://medicalxpress.com/news/2018-10-dna-websites-broad-net-people.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.