POLICY FORUM

GENETICS AND PRIVACY

Genealogy databases and the future of criminal investigation

The police can access your online family-tree research—and use it to investigate your relatives

By **Natalie Ram,[1] Christi J. Guerrini,[2] Amy L. McGuire[2]**

The 24 April 2018 arrest of Joseph James DeAngelo as the alleged Golden State Killer, suspected of more than a dozen murders and 50 rapes in California, has raised serious societal questions related to personal privacy. The break in the case came when investigators compared DNA recovered from victims and crime scenes to other DNA profiles searchable in a free genealogical database called GEDmatch. This presents a different situation from the analysis of DNA of individuals arrested or convicted of certain crimes, which has been collected in the U.S. National DNA Index System (NDIS) for forensic purposes since 1989. The search of a nonforensic database for law enforcement purposes has caught public attention, with many wondering how common such searches are, whether they are legal, and what consumers can do to protect themselves and their families from prying police eyes. Investigators are already rushing to make similar searches of GEDmatch in other cases, making ethical and legal inquiry into such use urgent.

In the United States, every state, as well as the federal government, has enacted laws enumerating which convicted or arrested persons are subject to compulsory DNA sampling and inclusion in the NDIS database. The NDIS contains more than 12 million profiles, and it is regularly used to match DNA from crime scenes to identify potential suspects. It is only helpful, however, if the suspect—or a family member of the suspect—has been arrested or committed a crime and their DNA has been collected and stored.

The case of the Golden State Killer is not the first instance of investigators turning to nonforensic DNA databases to generate leads. This was not even the first time investigators used genealogical DNA matches to develop and pursue a suspect in the Golden State Killer case itself. A year before investigators zeroed in on DeAngelo, they subpoenaed another genetic testing company for the name and payment information of one of its users and obtained a warrant for the man's DNA. He was not a match. Similarly, in 2014, Michael Usry found himself the target of a police investigation stemming from a partial genetic match between his father's DNA, stored in an Ancestry.com database, and DNA left at a 1996 murder scene. On the basis of the partial match, police were able to obtain a court order requiring Ancestry.com to disclose the identity of the database DNA match. After mapping out several generations of Usry's father's family, investigators zeroed in on Usry, eventually securing a warrant for his DNA. Ultimately, Usry was cleared as a suspect when his DNA proved not to match the crime scene DNA.

But there have also been reported successes. In 2015, for example, Arizona police arrested and charged Bryan Patrick Miller in the Canal Killer murders based in part on a tip drawn from a genealogical database search (*1*). Searches like these, drawing on genetic information unrelated to the criminal justice system, may offer substantial benefits. Allowing police to conduct similar database searches in other cases is likely to lead to more solved crimes. Moreover, expanding law enforcement investigations to encompass genealogical databases may help to remedy the racial and ethnic disparities that plague traditional forensic searches. In accordance with state laws, official forensic databases are typically limited to individuals arrested or convicted of certain crimes. Racial and ethnic disparities throughout the criminal justice system are therefore reproduced in the racial and ethnic makeup of these forensic databases. Genealogical databases, by contrast, are biased toward different demographics. The 23andMe database, for instance, consists disproportionately of individuals of European descent. Including genealogical databases in forensic searches might thus begin to redress, in at least one respect, disparities in the criminal justice system.

There are few legal roadblocks to police use of genetic databases intended to help individuals explore their health or identify genetic relatives. The Fourth Amendment's protection against warrantless searches and seizures generally does not apply to material or data voluntarily shared with a third party, like a direct-to-consumer genetics testing or interpretation company or a genetic matching platform like GEDmatch. Once an individual has voluntarily shared her data with a third party, she typically cannot claim any expectation of privacy in those data—and so the government need not secure a warrant before searching it.

Beyond the Constitution, three federal laws protect some genetic data against certain disclosures, but these too are unlikely to provide an effective shield against law enforcement searches in nonforensic genetic databases. The Genetic Information Nondiscrimination Act (GINA) protects genetic data, but only against certain uses by employers and health insurers (*2*). GINA provides no protection against law enforcement searches. Similarly, most companies and websites offering DNA testing, interpretation, or matching services directly to individuals likely are not covered by the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule, which governs the use and disclosure of identifiable health information. These providers are usually careful to explain that they are not engaged in health care or the manipulation or provision of health data (*3*). Finally, although certificates of confidentiality protect scientific researchers from disclosing data to law enforcement—even against a warrant (*4*)—they do not extend to scenarios in which law enforcement is just another contributor to and user of online genetic resources, such as public databases and matching tools. Certificates of confidentiality have faced few challenges in court, and so it is also uncertain whether the protection they purport to provide will hold up against a challenge by law enforcement seeking access.

Consistent with this legal landscape, companies and websites that generate, interpret, or match genetic data directly for individuals often do not promise complete protection. In terms of law enforcement, for instance, 23andMe states in its privacy policy, "23andMe will preserve and disclose any and all information to law enforcement agencies or others if required to do so by law or in the good faith belief that such preservation or disclosure is reasonably necessary to…comply with legal or regulatory process (such as a judicial proceeding, court order, or government inquiry)…" (*5*). Ancestry.com similarly

[1]University of Baltimore School of Law, Baltimore, MD, USA. [2]Center for Medical Ethics and Health Policy, Baylor College of Medicine, Houston, TX, USA. Email: nram@ubalt.edu; amcguire@bcm.edu

discloses, "We may share your Personal Information if we believe it is reasonably necessary to: [c]omply with valid legal process (e.g., subpoenas, warrants)..." (*6*). And in the wake of the Golden State Killer arrest, GEDmatch has altered its terms of service to explicitly permit law enforcement use of its database to investigate homicides and sexual assault (*7*). Although these disclaimers are usually unambiguous, they are sometimes buried in terms of service or privacy policies that many individuals do not take care to read or fully understand.

Despite the lack of legal protection against law enforcement searches of nonforensic databases, such searches may run counter to core values of American law. The Fourth Amendment is a constitutional commitment to protect fundamental civil rights. Part of that is a commitment to protecting privacy or freedom from government surveillance. Police cannot search a house without suspecting a specific individual of particular acts—even if doing so would enable the police to solve many more crimes. Yet, database searches permit law enforcement to search the genetic data of each database member without any suspicion that a particular member is tied to a particular crime. Although the U.S. Supreme Court has approved suspicionless genetic searches for individuals with diminished expectations of privacy, like those arrested or convicted of crimes (*8*), ordinary members of the public are different. Familial searches, like those used in the Golden State Killer investigation, are an even further departure from the Supreme Court standard. Certainly, individuals who commit crimes and leave their DNA behind forfeit any expectation of privacy in that DNA. But a usable forensic identification requires two matching parts: a crime scene sample and a database profile that matches it. Suspects identified through familial searches cannot be said to have voluntarily shared their genetic profile in a database of known individuals, even if a genetic relative has.

The Supreme Court is poised to reconsider its broad rule that the voluntary sharing of data negates expectations of privacy—and thus negates Fourth Amendment protections against warrantless government searches. In *Carpenter* v. *United States*, the Supreme Court will determine whether police must obtain a warrant to justify access to historical cell phone records revealing the movements and location of a cell phone user over a long period of time (*9*). In the digital age, in which nearly all data are at least nominally shared with third parties like internet service providers, website hosts, and cell phone companies, the current rule means that the Fourth Amendment often does not apply. *Carpenter* may reshape this rule to account for the realities of a big-data world. A ruling in *Carpenter* that limits police use of historical cell phone data may substantially affect police practices surrounding genetic data as well, as merely sharing data with another might well be insufficient to permit its suspicionless search by the government for crime-detection purposes.

Even if the Supreme Court decision in *Carpenter* does not revamp Fourth Amendment rules governing police access to shared data, the setting of that case suggests another way to resolve concerns about police access to nonforensic genetic databases. In the Stored Communications Act, Congress provided substantial statutory protection for email and other digital information maintained on the internet. Under the act, a court may order disclosure of electronic records if the government "offers specific and articulable facts showing that there are reasonable grounds to believe" that the records sought "are relevant and material to an ongoing criminal investigation" (*10*). This standard is less onerous than the Fourth Amendment's warrant requirement, but it is notably more demanding than any protections the law currently provides.

Enacting similar protection for genetic data stored in nonforensic databases would ensure that the government cannot subject ordinary individuals to suspicionless genetic searches, while allowing investigators to access genetic data where there is reason to believe a particular individual may be tied to a particular crime. A Stored Genetics Act would likely render law enforcement searches of nonforensic genetic databases unlawful for crime-detection purposes, as there can be no "specific and articulable" connection between particular database records and a particular crime when investigators seek to use such a search to generate leads, not investigate them. Thus, although such an approach would preserve freedom from perpetual genetic surveillance by the government, it may well result in fewer solved cases.

Legislatures may understandably be loath to enact a total prohibition of such searches. At a minimum, however, policy-makers should delineate under what circumstances such searches are acceptable. For example, several states, including California, Colorado, and Texas, have identified prerequisites to the use of familial searches of the state's own forensic database, including that the crime to be investigated is serious and that traditional investigative techniques have been exhausted without success (*11*). Similar constraints could be placed on law enforcement searches of nonforensic databases. The challenge of this approach is that limitations on the scope of use can erode quickly. Thus, although Colorado's policy governing familial searches of the state's forensic database limits such searches to crimes with "significant public safety concerns," police in that state used a familial search to solve a car break-in where the perpetrator "left a drop of blood on a passenger seat when he broke a car window and stole $1.40 in change" (*11*). The erosion of limits on crime-solving technology may well be inevitable, and it threatens our collective civil liberties and opens the door to socially and politically unacceptable genetic surveillance.

Whatever legislative solution is adopted, it must at least take into account public perspectives to clearly delineate acceptable uses and balance the social benefit of solving cases with individuals' interests in avoiding unwarranted government scrutiny. ■

ILLUSTRATION: LUSH/123RF

REFERENCES AND NOTES

1. M. Cassidy, "How forensic genealogy led to an arrest in the Phoenix 'Canal Killer' case," Arizona Republic, 10 November 2016; www.azcentral.com/story/news/local/phoenix/2016/11/30/how-forensic-genealogy-led-arrest-phoenix-canal-killer-case-bryan-patrick-miller-dna/94565410/.
2. 122 Statute 881.
3. J. Hsu, *I/S: A Journal of Law & Policy for the Information Society (Moritz College of Law)* **6**, 557 (2011).
4. 42 United States Code (U.S.C.) § 241 (d).
5. www.23andme.com/about/privacy/.
6. www.ancestry.com/cs/legal/privacystatement.
7. https://bit.ly/2IZKzGt.
8. *Maryland* v. *King*, 133 S. Ct. 1958 (2013).
9. *Carpenter* v. *United States*, No. 16-402 (argued 29 November 2017).
10. 18 U.S.C. § 2703 (d).
11. N. Ram, *Stanford Law Rev.* **63**, 751 (2011).

ACKNOWLEDGMENTS

Funding was provided by the National Institutes of Health, National Human Genome Research Institute (K01 HG009355 and R01 HG008918). We thank M. Majumder, D. Peterson, S. Pereira, R. Hsu, A. Gutierrez, and J. Robinson for contributions and assistance on this project.

10.1126/science.aau1083

Genealogy databases and the future of criminal investigation

Natalie Ram, Christi J. Guerrini and Amy L. McGuire

Science **360** (6393), 1078-1079.
DOI: 10.1126/science.aau1083