THE COLUMBIA
SCIENCE &
TECHNOLOGY
LAW REVIEW

VOL. XXI STLR.ORG FALL 2019

WHY WE FEAR GENETIC INFORMANTS: USING GENETIC GENEALOGY TO CATCH SERIAL KILLERS

Teneille R. Brown, JD*

Consumer genetics has exploded, driven by the second-most popular hobby in the United States: genealogy. This hobby has been co-opted by law enforcement to solve cold cases, by linking crime-scene DNA with the DNA of a suspect's relative, which is contained in a direct-to-consumer (DTC) genetic database. The relative's genetic data acts as a silent witness, or genetic informant, wordlessly guiding law enforcement to a handful of potential suspects. At least thirty murderers and rapists have been arrested in this way, a process which I describe in careful detail in this article. Legal scholars have sounded many alarms, and have called for immediate bans on this methodology, which is referred to as long-range familial searching (or "LRFS"). The opponents' concerns are many, but generally boil down to fears that LRFS will invade the privacy and autonomy of presumptively innocent individuals. These concerns, I argue, are considerably overblown. Indeed, many aspects of the methodology implicate nothing new, legally or ethically, and might even better protect privacy while exonerating the innocent. Law enforcement's use of LRFS to solve cold cases is a bogeyman. The real threat to genetic privacy comes from shoddy consumer consent procedures, poor data security standards, and user agreements that permit

* Teneille R. Brown J.D. is a Professor of Law at the University of Utah, S.J Quinney College of Law, and an adjunct Professor in the Department of General Internal Medicine. This research was made possible, in part, through generous support from the Albert and Elaine Borchard Fund for Faculty Excellence and funding from the Utah Center for Excellence in ELSI Research (UCEER). UCEER is supported by the National Human Genome Research Institute of the National Institutes of Health under Award Number P20HG007249. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health.

rampant secondary uses of data. So why do so many legal scholars fear a world where law enforcement uses this methodology? I surmise that our fear of so-called genetic informants stems from the sticky and long-standing traps of genetic essentialism and genetic determinism, where we incorrectly attribute intentional action to our genes and fear a world where humans are controlled by our biology. Rather than banning the use of genetic genealogy to catch serial killers and rapists, I call for improved direct-to-consumer consent processes, and more transparent privacy and security measures. This will better protect genetic privacy in line with consumer expectations, while still permitting the use of LRFS to deliver justice to victims and punish those who commit society's most heinous acts.

I. INTRODUCTION

Consumer genetics has exploded, driven by the second-most popular hobby in the United States: genealogy.[1] Kits from the market leaders—23andMe and Ancestry—are top-sellers on Black Friday.[2] As the Centers for Disease Control and Prevention reports, ancestry tests that have been bundled with information on genetic health risks are selling "like hotcakes," and direct-to-consumer (DTC) genetic tests have "continued to skyrocket."[3] These kits allow users to submit a saliva sample through the mail, without the involvement of a physician, and test for about 700,000 genetic mutations. The results are then returned to them online.

---

1. Scott Bowen & Muin J. Khoury, *Consumer Genetic Testing Is Booming: But What Are the Benefits and Harms to Individuals and Populations?*, CENTER FOR DISEASE CONTROL: GENOMICS AND PRECISION HEALTH (June 12, 2018), https://blogs.cdc.gov/genomics/2018/06/12/consumer-genetic-testing.

2. Shanna Mason, *Privacy of Information and DNA Testing Kits*, 27 CATH. U.J.L. & TECH. 161, 161 (2018) ("In 2017, AncestryDNA sold roughly 1.5 million kits from Black Friday through Cyber Monday, triple the amount of sales from 2016.").

3. Bowen & Khoury, *supra* note 1.

This "hobby" has been co-opted by law enforcement to solve cold cases. Officers can link crime-scene DNA with the DNA of one of the assailant's relatives, which had been previously uploaded to a non-forensic, DTC genetic database. A 2018 study predicted that within a couple of years roughly 90% of Americans of European descent will be genetically identifiable in this way, even those who have never submitted a saliva sample nor been tested themselves.[4] A relative's genetic data can act as a silent witness, or genetic informant, against the person who left the DNA at the crime scene. This "genetic informant" wordlessly guides law enforcement to a handful of potential suspects, by simply informing them that the suspect is very likely a third-cousin, nephew, or grandson of the person in the DTC database. Public records and newspaper clippings then provide the necessary details to put a name and location to the crime-scene DNA.

At least thirty murderers and rapists have been arrested after detectives identified them through a combination of genetic testing and genealogy research tools. This general method has been dubbed "genetic genealogy"—the use of DNA to infer relationships between individuals.[5] Given the hundreds of thousands of cold cases in the U.S., with some unknown number of case files containing DNA samples, experts predict that genetic genealogy will become a multimillion dollar forensic business.[6] One method in particular, called "forensic genetic genealogy," or FGG, has allowed law enforcement to significantly reduce the size of the suspect pool when no other good leads exist.[7]

The public generally seems to support the use of genetic genealogy to apprehend violent criminals in cold cases, even when

---

4. Yaniv Erlich et al., *Identity Inference of Genomic Data Using Long-Range Familial Searches*, 362 SCIENCE 690, 690 & 691 fig.1 (2018); *see also* Heather Murphy, *Most White Americans' DNA Can Be Identified Through Genealogy Databases*, N.Y. TIMES (Oct. 11, 2018), https://www.nytimes.com/2018/10/11/science/science-genetic-genealogy-study.html?module=inline.

5. Peter Aldous, *The Golden State Killer Case Has Spawned a New Forensic Science Industry*, BUZZFEED NEWS (Feb. 15, 2019), https://www.buzzfeednews.com/article/peteraldhous/genetic-genealogy-dna-business-parabon-bode.

6. *Id.*

7. Ellen Greytak et al., *Privacy and Genetic Genealogy Data*, 361 SCIENCE 857 (2018); *Interim Policy: Forensic Genetic Genealogical DNA Analysis and Searching*, DEP'T OF JUSTICE 6 (2019), www.justice.gov/olp/page/file/1204386/download; Heather Murphy, *Genealogists Turn to Cousins' DNA and Family Trees to Crack Five More Cold Cases*, N.Y. TIMES (June 27, 2018), https://www.nytimes.com/2018/06/27/science/dna-family-trees-cold-cases.html (Referring to FGG as "long range familial searches," or LRFS).

the individual contributing her DNA to a genealogical website had no notice her sample would be used in this way.[8] However, legal scholars have sounded many alarms. In op-eds in the New York Times[9] and Slate,[10] in scientific and legal scholarly articles,[11] and in lobbying efforts with their legislatures,[12] legal scholars have called for limitations or bans on these practices. Some propose requiring law enforcement to get warrants before police can access DTC genetic databases or immediately banning FGG and other genetic genealogy tools. In Maryland, legislators proposed a bill in 2019 to prohibit the searching of genealogical databases to find distant relatives of criminal suspects.[13]

The opponents' concerns are many, but they can generally be boiled down to fears that these new methods will invade the privacy and autonomy of presumptively innocent individuals by creating an involuntary and *de facto* forensic genetic database.[14]

---

8. The majority of Americans polled support police searches of genetic websites that identify genetic relatives and disclosure of DTC genetic testing customer information, as well as creation of fake profiles of individuals by police on genealogy websites. Respondents were much more supportive of these activities when the purpose was to identify perpetrators of violent crimes than when the purpose was to identify perpetrators of nonviolent crimes. However, the sample was more likely than the rest of the population to have been the victim of a crime. *See* Christi Guerrini et al., *Should Police Have Access to Genetic Genealogy Databases? Capturing the Golden State Killer and Other Criminals Using a Controversial New Forensic Technique*, 16 PLOS BIOLOGY e2006906, *3 (2018).

9. Elizabeth Joh, Opinion, *Want to See My Genes? Get a Warrant*, N.Y. TIMES (June 11, 2019), https://www.nytimes.com/2019/06/11/opinion/police-dna-warrant.html.

10. Natalie Ram, *The U.S. May Soon Have a De Facto National DNA Database,* SLATE (Mar. 19, 2019, 7:30 AM), https://slate.com/technology/2019/03/national-dna-database-law-enforcement-genetic-genealogy.html.

11. Natalie Ram et al., *Genealogy Databases and the Future of Criminal Investigation*, 360 SCIENCE 1078, 1079 (2018).

12. Natalie Ram, *Incidental Informants Police Can Use Genealogy Databases to Help Identify Criminal Relatives-but Should They*?, MD. B.J., July-Aug. 2018, at 8, 10.

13. H.B. 30, 440th Gen. Assemb., Reg. Sess., (Md. 2019) ("For the purpose of prohibiting a person from performing a search of a certain DNA or genealogical data base for the purpose of identification of an offender in connection with a crime for which the offender may be a biological relative of the individual from whom the DNA sample was acquired; and generally relating to DNA analysis.").

14. Privacy is a multifaceted concept that can include the right to be left alone, to be free from surveillance, to remain anonymous, to keep your information confidential, or entirely private, or to ensure that what is said about you is true. Each of these privacy concepts is potentially implicated here, as well as the additional idea that your privacy can be violated in ways that exploit you and violate your autonomy. *See* Daniel J. Solove, *A Taxonomy of Privacy*, 154 U. PA. L. REV. 477, 505-24 (2005); Ellen W. Clayton et al., *A Systematic Literature Review of Individuals' Perspectives on Privacy and Genetic Information in the United States*, 13 PLOS ONE e0204417, *14-18 (2018).

These concerns, I argue, are considerably overblown. Indeed, many aspects of FGG implicate nothing new, legally or ethically, and might even *better* protect the privacy of innocent individuals. That's right. This methodology might *reduce* the privacy violations that are rampant in ordinary police investigations. So why are so many legal scholars fascinated by genetic genealogy yet fear a world where law enforcement uses FGG? What is it about this methodology that triggers knee-jerk calls to ban the use of "genetic informants?"

We are right to be concerned about unleashing private, genetic information to the government or private actors. We are still unlocking the secrets of our genomes, and yet the promised value of using genetic data to guide health care treatment is enormous.[15] Because there is a great deal of money to be made developing health care products that are tailored to individuals based on their genomic information (the goal of so-called "precision medicine"), many private and public research institutions would love to get their hands on large datasets of genetic information, especially when that information is coupled with multi-generational pedigrees, traits, and lifestyle choices. Analysts estimate the global market for precision medicine initiatives will increase from an estimated $92.4 billion in 2017 to $194.4 billion by 2024.[16] The value of the data that 23andMe and Ancestry store is colossal. There is, therefore, a great risk of deliberate data breaches or weak user privacy protections.[17]

By concentrating on law enforcement's use of FGG, and failing to address the larger risks of genetic research and disclosure, these privacy scholars miss the mark. Law enforcement's use of genetic genealogy to solve cold cases is a bogeyman. The larger threat to genetic privacy comes from shoddy consumer consent procedures used by DTC genetic companies, poor data security standards, and user agreements that permit rampant secondary uses of the users' DNA and data. Unless police drastically expand the way

---

15. Geoffrey Ginsburg & Kathryn Phillips, *Precision Medicine: From Science to Value*, 37 HEALTH AFF. 694, 694 (2018) ("The assembly of genomic, environmental, digital health, and patient-reported data from a variety of sources serves as the foundation for a powerful precision medicine platform that, when coupled to other national and global data and clinical networks, will lead to the dissemination of knowledge that will enable other health care delivery systems to benefit.").

16. *Global Personalized Medicine Market 2017-2018 & 2024: Market is Projected to Reach US$194.4 Billion by 2024 from an Estimated US$92.4 Billion in 2017*, PR NEWSWIRE (Oct. 15, 2018, 7:45 AM), https://www.prnewswire.com/news-releases/global-personalized-medicine-market-2017-2018--2024-market-is-projected-to-reach-us194-4-billion-by-2024-from-an-estimated-us92-4-billion-in-2017--300730848.html.

17. Yaniv Erlich & Arvind Narayanan, *Routes for Breaching and Protecting Genetic Privacy*, 15 NATURE REV. GENETICS 409, 409 (2014).

they are conducting genetic genealogical searches, there is too much fear and fascination surrounding this methodology. This Article seeks to demystify this unfounded fear.

This Article proceeds in three parts. In the first part, I explain the rise of genetic genealogical testing and how it is employed by police for FGG. I also clarify how FGG is different from a traditional search of the federal Combined DNA Index System (CODIS) database.[18] In the second part, I challenge the many concerns that scholars have raised in response to FGG. Specifically, I counter the arguments that it violates the Fourth Amendment, invades the privacy of innocent individuals, renders people unintentional genetic informants, improperly relies on police deception and the involuntary participation of suspects, and creates a *de facto* federal database. These concerns reflect misunderstandings of ordinary criminal procedure, the legal might of online user "agreements," and the distinctions between clinical research and criminal law. In the third part, I provide a unique theory for why we seem to fear "genetic informants." I conclude with a call for more nuanced policy measures that will better protect genetic privacy consistent with consumer expectations, while still permitting the use of FGG to deliver justice to victims and help convict serial killers and rapists.

II. THE EXPLOSION OF CONSUMER GENETICS

In the last few years, the cost for genetic testing has dropped considerably, and the large genetic ancestry companies have also lowered their prices, resulting in a predictable spike in demand.[19] One market leader, Ancestry, boasts over 15 million customers while its primary competitor, 23andMe, has more than 10 million customers.[20] One consumer genetics businessman remarked that

---

18. The Combined DNA Index System of the Federal Bureau of Investigation is commonly referred to as CODIS and is the federal database that contains the short-tandem repeat satellite markers at 13 or 20 non-coding regions for the individuals who have been sampled. The sampled population consists mostly of criminal offenders, but has been expanded to include arrestees. States may contribute to the federally-maintained database. *See* 34 U.S.C. § 40702 (2012).

19. Jie Yuan et al., *DNA.Land is a Framework to Collect Genomes and Phenomes in the Era of Abundant Genetic Information*, 50 NATURE GENETICS 160, 160 (2018); Tim Caulfield & Amy L. McGuire, *Direct-to-Consumer Genetic Testing: Perceptions, Problems, and Policy Responses*, 63 ANN. REV. MED. 23, 23 (2012); *see also* Bowen & Khoury, *supra* note 1.

20. *See About Us*, 23ANDME, https://mediacenter.23andme.com/company/about-us/; *Ancestry.com Surpasses 15 Million DNA Customers*, ANCESTRY, https://blogs.ancestry.com/ancestry/2019/05/31/ancestry-surpasses-15-million-dna-customers.

"the inflection point [for DTC genetic test sales] started in the summer of 2016, and from there it's gone into the stratosphere."[21] Most of these sales have occurred in the United States, and roughly 1 in 25 Americans have had their samples analyzed online, without involving a physician or geneticist.[22] It is fair to say that the market for DTC genetic tests is booming.

However, the many DTC companies are considerably different from one another. Some offer tests that claim to match romantic partners based on their genetic soulmate, others create personalized travel holidays, and still others purport to identify favorite beverages, all based entirely on your genetic results.[23] Some even market themselves as being able to predict "how gay you are," by relying on a study that itself specifically dismissed the idea that their results could be used on individuals.[24] These kinds of tests lack clinical and analytic validity, and are a form of pseudoscience quackery.[25] Somewhere in the middle are health-related tests that claim to screen for food allergies or common drug side effects, with some following laboratory and privacy best practices, and others not. Then on the other end of the spectrum are the reputable DTC-companies such as 23andMe and Ancestry. These companies began by offering ancestry testing and have now branched out to offer health-related information.[26]

So far, it does not seem that finding out about an elevated risk for some disease changes behaviors, and in some cases it probably should not. The health-related risk information is often of very weak predictive value, particularly for complex diseases like cancer. Results often reflect small increases in overall lifetime risk, and cannot be interpreted without knowing someone's family history, personal risk, and environmental factors. Therefore, using DTC genetic tests to make health care decisions is often

---

21. Antonio Regalado, *2017 Was the Year DNA Consumer Testing Blew Up*, MIT TECH. REV. (Feb. 12, 2018), https://www.technologyreview.com/s/610233/2017-was-the-year-consumer-dna-testing-blew-up.

22. *Id.*

23. James W. Hazel & Christopher Slobogin, *Who Knows What, and When?: A Survey of the Privacy Policies Proffered by U.S. Direct-to-Consumer Genetic Testing Companies*, 28 CORNELL J.L. & PUB. POL'Y 35, 36–37 (2018).

24. Bizarrely, the app developers claimed it would be "absurd" for anyone to use the app in the way it was marketed, as its small-font disclosures technically warned that the app was not to be used to predict same-sex attraction. *See* Amy Maxmen, *'Gay Gene' App Provokes Fears of a Genetic Wild West*, 574 NATURE 609, 609 (2019).

25. Hazel & Slobogin, *supra* note 23.

26. *See* Megan A. Allyse et al., *Direct-to-Consumer Testing 2.0: Emerging Models of Direct-to-Consumer Genetic Testing*, 93 MAYO CLINIC PROC. 113, 116-117 (2018).

premature.[27] At present, it seems that most people pursue DTC genetic testing because they are tantalized by the idea of having as much information about themselves as possible, even information that has little predictive value.

A. Direct-to-Consumer Genetic Testing Companies are Under-Regulated and Provide Inadequate Consent and Privacy Protections

Because the DTC genetic testing industry is under-regulated, the quality of the privacy protections and clinical and analytic validity of the DTC genetic tests vary considerably.[28] This puts consumers at risk of data breaches with massive implications. Due to ineffective consent procedures at the initial stage when users submit a saliva sample and the failure of sites to require validation of user identities or to provide secure encryption, users may be unwittingly supplying the secrets of their genome to absolute strangers.[29] These strangers could then sell or share the valuable genetic profile data for legal (or illegal) insurance under-writing or pharmaceutical advertising schemes.

Importantly, there are no constitutional limitations on nefarious uses of data by private actors, and the federal Genetic Information Non-Discrimination Act (GINA) provides inadequate protection.[30] For example, GINA allows employers to use genetic risk information that they discover through "commercially available publications" such as newspapers, which might include websites like GEDMatch.[31] Further, GINA does not apply to life or disability insurance, and it requires that the discriminatory behavior be exposed. As with many forms of discrimination, detection is notoriously difficult, as is proving the intent behind the employer's or insurer's behavior. It is possible to argue that someone's insurance was priced the way it was or an adverse employment decision was made on some other pretextual, non-genetic basis. For example, in the context of racial discrimination, it is possible for an employer to successfully argue that the individual was fired due to very recent,

---

27. Jason Park et al., Question & Answer, *Privacy in Direct-to-Consumer Genetic Testing*, 65 CLINICAL CHEMISTRY 612, 613 (2019) (noting that "there is no solid evidence to support the contention that providing individuals with genetic information will . . . impact [their lifestyle choices].").

28. Hazel & Slobogin, *supra* note 23, at 40 ("DTC-GT remains largely unregulated in the majority of jurisdictions.").

29. Caitlin Curtis et al., *Protecting Trust in Medical Genetics in the New Era of Forensics*, 21 GENETICS MED. 1483, 1483-84 (2019) (proposing that "[i]t must not be possible for an individual to unwittingly sign an agreement that results in loss of control of their genetic data.").

30. *See* PROHIBITING EMPLOYMENT DISCRIMINATION ON THE BASIS OF GENETIC INFORMATION, 42 U.S.C. § 2000ff-1 (2019).

31. *Id.* at 42 U.S.C. § 2000ff-1(b) (2019).

and possibly fabricated, work performance issues, rather than due to the experience of racial bias. In the absence of data showing a disparate impact on a large number of employees, there is plausible deniability that the employer did not engage in intentionally unlawful employment discrimination. The same could occur with GINA. Finally, GINA is not written in stone; it is simply a Congressional statute that could be repealed. The privacy and security risks associated with DTC genetic testing led Senate Minority Leader Chuck Schumer (D-NY) to hold a press conference in 2017 to call for more regulatory oversight of DTC genetic testing.[32] Specifically, he asked the Federal Trade Commission to look closely at this industry and ensure that companies have fair privacy policies as well as adequate informed consent procedures. This has yet to occur.

This is big business and big data research—ancestry testing is just the gateway. Essentially, these companies are enormous biobanks. Because DTC genetic testing companies often also have pedigree and personal information, their genetic databases could be instrumental in assessing genetic risk for complex diseases. This makes their data incredibly valuable to pharmaceutical companies and clinical researchers, and Ancestry and 23andMe have publicized their relationships with these types of research bodies.

Unfortunately, the Common Rule, which provides protection for human subjects in research,[33] and the Health Insurance Portability and Accountability Act (HIPAA), which provides a bare minimum of protection for the security and privacy of identifiable health information,[34] do not apply to research that occurs outside of a health care setting and that is not federally funded.[35] Users may incorrectly expect certain health-related

---

32. *Schumer Reveals: Popular at Home DNA Test Kits Are Putting Consumer Privacy at Great Risk, as DNA Firms Could Sell Your Most Personal Info & Genetic Data to All-Comers; Senator Pushes Feds to Investigate & Ensure Fair Privacy Standards for All DNA Kits*, CHARLES E. SCHUMER: UNITED STATES SENATOR FOR NEW YORK (Nov. 26, 2017), https://www.schumer.senate.gov/newsroom/press-releases/schumer-reveals-popular-at-home-dna-test-kits-are-putting-consumer-privacy-at-great-risk-as-dna-firms-could-sell-your-most-personal-info-and-genetic-data-to-all-comers-senator-pushes-feds-to-investigate_ensure-fair-privacy-standards-for-all-dna-kits.

33. 45 C.F.R. § 46 (2019).

34. 45 C.F.R. § 160 (2019); 45 C.F.R. § 164 (2019).

35. *See* Clayton et al., *supra* note 14, at *14 ("The U.S. Common Rule also permits the use of de-identified data without consent and with limited to no IRB oversight and endorses an expansive role for broad consent of identified data.").

privacy regulations to apply because of the quasi-clinical nature of the information.[36]

Even if companies were to voluntarily comply with HIPAA, we cannot rely on existing laws to protect us because it is impossible to completely anonymize genomic information.[37] Every few months, new methods are developed to use genetic datasets and public records to re-identify anonymized samples. Given the enormous potential of genetic information to reveal health risks in the future, we need to restrict access to this information in ways that might at first seem paternalistic. More will be said below about how we might improve the consent procedures and limit secondary uses, but we should also reconsider whether health-related information should *ever* be relayed without a physician's or genetic counselor's interpretation.

B. DTC Genetic Tests That Rely on SNP Data Reveal More Than Just Ancestry

The genetics revolution, fueled in part by an explosion in DTC genetic testing, is upon us. In 2017 alone, about 7 million genetic testing kits were sold directly to individuals, and 20 million kits were expected to be sold in 2018.[38] Driven mostly by genealogical hobbyists, the majority of the DTC ancestry genetic testing services rely on single nucleotide polymorphisms (SNPs), which are mutations at the level of the individual nucleotides. While SNP data is not nearly as rich as data gathered from sequencing, it still provides a significant amount of information about future risk of disease.

SNP data can also reveal whether users share segments of their genome with other users, predicting relatedness through a common ancestor. This works by analyzing the percentage of overlapping bits of genetic code, so-called "identical by descent" sections, that one shares with relatives. Assuming no historical inbreeding, one likely

---

36. Park et al., *supra* note 27, at 614-15 (discussing the risks of cyber-attacks to DTC genetic testing databases, such as those waged on the 100000 Genomes Project data in the United Kingdom).

37. This will be discussed in more detail below. Nora von Thenen et al., *Re-Identification of Individuals in Genomic Data-Sharing Beacons Via Allele Inference*, 35 *BIOINFORMATICS* 365, 365 (2019). *See also* Bridget F.B. Algee-Hewitt et al., *Individual Identifiability Predicts Population Identifiability in Forensic Microsatellite Markers*, 26 CURRENT BIOLOGY 935, 937 (2016) (finding that forensic markers have nontrivial ancestry information); Michael D. Edge et al., *Linkage Disequilibrium Matches Forensic Genetic Records to Disjoint Genomic Marker Sets*, 114 PROC. NAT'L ACAD. SCI. 5671, 5671-76 (2017) (finding that genetic databases can be compared to identify individuals).

38. Erlich et al., *supra* note 4 at 690; Park et al., *supra* note 27, at 612.

shares roughly 12% of their genome with first cousins, about 3% with second cousins, and less than 1% with third cousins.39 Thus, by finding and quantifying overlapping genetic regions, DTC companies can predict genetic familial relationships. However, because parents do not contribute exactly half of their genome to their offspring and the reshuffling of DNA can be random, third cousins may share more DNA fragments than second cousins.40 Since genetic inheritance varies from child to child, as one moves beyond the level of third cousins, there is a decent chance that a known genealogical relationship will not be detectable genetically.41

In addition to predicting genealogical relationships, some DTC genetic tests now reveal SNPs linked to developing diseases and other traits.[42] While most complex traits cannot be reliably and accurately predicted through SNP data, there are thousands of individual mutations or "genotypes" that can increase the likelihood of developing a particular trait or "phenotype." Some companies require a physician to order these test kits, but the most popular ones do not.[43] Other niche tests focus on so-called "recreational" traits like detecting the smell of asparagus in urine or identifying nutritional needs and possible food allergies.

The leading consumer genetics companies, 23andMe and Ancestry, allow consumers to download their raw genetic data in plain-text format, which can then be uploaded to third-party websites.[44] These websites provide a range of additional services, including interpreting the clinical relevance of mutations and allowing individuals to expand the reach of their genealogical search. Up to 62% of DTC customers will upload their genetic data

---

39. *Average Percent DNA Shared Between Relatives*, 23ANDME, https://customercare.23andme.com/hc/en-us/articles/212170668-Average-percent-DNA-shared-between-relatives (last visited Nov. 18, 2019).

40. Catherine Rehder et al., *American College of Medical Genetics and Genomics: Standards and Guidelines for Documenting Suspected Consanguinity as an Incidental Finding of Genomic Testing*, 15 GENETICS IN MED. 150, 151 (2013).

41. Michael Edge & Graham Coop, *How Lucky Was the Genetic Investigation in the Golden State Killer Case?*, BIORXIV 5 (Jan. 29, 2019), https://www.biorxiv.org/content/biorxiv/early/2019/01/29/531384.full.pdf.

42. In 2013, the FDA sent cease and desist letters to 23andMe, ordering them to stop marketing and selling their health-related testing services until they received FDA approval for these devices. In 2017, the FDA approved 23andMe's carrier screening for hereditary Bloom syndrome, which created "DTC Testing 2.0." There was now precedent and a pathway for including disease-risk in the DTC panels. *See* Megan A. Allyse et al., *Direct-to-Consumer Testing 2.0: Emerging Models of Direct-to-Consumer Genetic Testing*, 93 MAYO CLINIC PROC. 113, 116-117 (2018).

43. Eline M. Bunnik et al., *Informed Consent in Direct-to-Consumer Personal Genome Testing: The Outline of a Model Between Specific and Generic Consent*, 28 BIOETHICS 343, 343-44 (2014).

44. Erlich et al., *supra* note 4, at 690.

to third-party websites for free or for a small fee.[45] One such third-party website is GEDMatch, an open-access service that is free for the most basic searches.[46]

C. Third-Party Sites Like GEDMatch Facilitate FGG

GEDMatch users can connect with even more distant relatives who used different testing services like FamilyTreeDNA or My Heritage. They do so by uploading their SNP profile, generated elsewhere, onto GEDMatch. The raw SNP data is analyzed using a simple algorithm, and the site then produces a list of likely relatives automatically, without the need to share any underlying genetic information with the putative relative. In just a few years, GEDMatch has cultivated a large community of hundreds of thousands of users.[47] While the user agreements of 23andMe and Ancestry state that they will not disclose users' genetic data without a legal subpoena or warrant and that users must not submit samples under false identities,[48] GEDMatch's agreement has never included such guarantees. In fact, GEDMatch allows users to use an alias rather than their real names to register.[49] In 2018, GEDMatch made explicit in their user agreement that law enforcement could submit profiles from crime scene DNA to find a suspect's distant relatives.[50] Even before then, however, the GEDMatch user agreement included the following warning: "DNA and Genealogical research, by its very nature, requires the sharing of information. Because of that, users participating in this site should expect that their information will be shared with other users."[51]

---

45. *See* Maxmen, *supra* note 24, at 610.

46. *GEDmatch.Com Terms of Service and Privacy Policy*, GEDMATCH, https://www.gedmatch.com/tos.htm (last updated May 18, 2019).

47. Yuan et al., *supra* note 19, at 160.

48. *See 23andMe Guide for Law Enforcement*, 23ANDME, www.23andme.com/law-enforcement-guide (last visited Nov. 19, 2019) and *Ancestry Guide for Law Enforcement*, ANCESTRY, www.ancestry.com/cs/legal/lawenforcement (last visited Nov. 19, 2019).

49. "Although you may provide a real name for registration and data upload, you have the option of providing an alias for either login or data." *Terms of Service and Privacy Policy*, GEDMATCH, www.gedmatch.com/tos.htm (last updated May 18, 2019).

50. *Id.* ("When you upload Raw Data to GEDmatch, you agree that the Raw Data is one of the following: Your DNA… DNA obtained and authorized by law enforcement to either: (1) identify a perpetrator of a violent crime against another individual… [or (2)] identify remains of a deceased individual.").

51. Cyrus Farivar, *GEDmatch, a Tiny DNA Analysis Firm, Was Key for Golden State Killer Case*, ARS TECHNICA (Apr. 27, 2018), https://arstechnica.com/tech-policy/2018/04/gedmatch-a-tiny-dna-analysis-firm-was-key-for-golden-state-killer-case.

In February 2019, FamilyTreeDNA announced that they too would allow law enforcement to submit crime scene DNA. Unlike GEDMatch, however, FamilyTreeDNA requires law enforcement to register all forensic samples and genetic files prior to uploading to the FamilyTreeDNA database. This is consistent with an interim policy issued by the Department of Justice (DOJ) in September 2019, which now requires law enforcement to submit the crime-scene derived profile explicitly on behalf of law enforcement, and not under false pretenses.[52] Permission to use FamilyTreeDNA's site for searching is only granted after the required documentation is submitted, reviewed, and approved. Permissible searches are limited to those identifying the remains of a deceased individual or a suspect in a homicide, sexual assault, or trafficking case.[53]

In part because of these lax standards, law enforcement agencies have been uploading genetic profiles increasingly quickly to GEDMatch and FamilyTreeDNA to link unidentified criminals with relatives. Profiles are constructed from samples of blood, semen, or tissue found at the crime scene. Often there is insufficient DNA available from the crime scene to develop a full SNP profile or run multiple genetic tests.[54] However, law enforcement's ability to extract whole-genome genotypes from degraded crime-scene samples is improving.[55] Further, if there is enough cellular material, which is usually the case with semen from sexual assaults, law enforcement can use SNP microarrays.[56] The microarrays generate dense genetic profiles indistinguishable from those developed by the major DTC genetic testing companies.[57]

Because the identity of the person from whom the crime-scene sample came is often unknown, law enforcement uses a false name—"John Doe," for example—and submits it to GEDMatch. Then, when their "John Doe" matches someone in the database, they use genealogical data to determine a common ancestor who

---

52. *Interim Policy: Forensic Genetic Genealogical DNA Analysis and Searching*, *supra* note 7.

53. *Law Enforcement Guide*, FAMILYTREEDNA, https://www.familytreedna.com/legal/law-enforcement-guide (last visited Apr. 3, 2019).

54. New developments in massively parallel sequencing may be one way of getting more forensic data out of a limited amount of DNA in a sample. Denise S. Court, *Forensic Genealogy: Some Serious Concerns*, 36 FORENSIC SCI. INT'L: GENETICS 203, 203 (2018).

55. *See* Paul Ellenbogen & Arvind Narayanan, *Identification of Anonymous DNA Using Genealogical Triangulation* 5 (bioRxiv, Working Paper No. 531269, 2019).

56. Court, *supra* note 54.

57. Ellenbogen & Narayanan, *supra* note 55.

might be a great-great grandfather or grandmother. They then triangulate other data, such as birth, voting, and military records, to build out the pedigrees from that common ancestor, identifying all of the potential individuals who may be suspects. As we each have about 1,000 fourth cousins and 5,000 fifth cousins,[58] depending on the degree of relation, this process can be quite time-consuming. The methodology is known by different names. In the forensic genetics research community, it is referred to as "long-range familial searches" (LRFS). Law enforcement sometimes refers to this as Forensic Genetic Genealogy (FGG). This Article uses the term FGG.

The Golden State Killer, Joseph DeAngelo, was finally arrested using the FGG technique after eluding California police for decades. DeAngelo murdered at least 12 people and sexually assaulted at least 45 women. Although law enforcement had multiple samples of his DNA from crime scenes, his DNA did not match any samples contained in the federal CODIS "offender" DNA database.[59] Until the advent of FGG, it seemed like the identity of the Golden State Killer might never be known, and justice for his many victims might never be delivered.

Once DeAngelo was finally arrested, questions regarding the method of his identification began to surface. The police reluctantly acknowledged that they used FGG, following the steps described above.[60] The profile derived from the crime scene matched someone in GEDMatch—a distant cousin of the perpetrator. With the help of genealogists, law enforcement found a common Italian ancestor shared by the Golden State Killer and his distant relative.[61] They then built the family tree branch by branch to find people who were about the right age and sex at the time of the crimes.[62] They initially tailed the wrong person, following him until he left some trash behind that contained his DNA, which they tested against the crime scene samples. It was not a match. They eliminated that individual and kept looking for other possibilities. Eventually, they identified Joseph DeAngelo. After analyzing DNA he also left on a piece of trash, they had their suspect. It was a match. The former cop, now

---

58. Court, *supra* note 54.

59. More will be said about the database that the Federal Bureau of Investigations (FBI) maintains, CODIS, at p. 15. CODIS relies on a very different type of genetic profile, based on short-tandem repeats at 20 locations in the human genome that are not thought to code for traits. Because they are thought not to be coding regions, individual variation in these STRs is quite high, making them useful markers for differentiating individuals.

60. Edge & Coop, *supra* note 41.

61. Justin Jouvenal, *To Find Alleged Golden State Killer, Investigators First Found His Great-Great-Great-Grandparents*, WASH. POST (Apr. 30, 2018, 6:22 PM), https://wapo.st/2HCvivq.

62. *See* Guerrini, *supra* note 8, at *3.

in his early seventies, was finally arrested after evading law enforcement for decades.

The Golden State Killer was one of several suspects identified using FGG. Parabon® Nanolabs, Inc., a private company that has commercialized FGG for law enforcement, reports that a few dozen individuals have been arrested in this way.[63] Recently, another large forensic laboratory has entered the market.[64] Two decades after a man attacked ten women in their homes, investigators used FGG to identify Roy Charles Waller as the serial rapist. Within five minutes of viewing the GEDMatch data, they identified a close relative of the perpetrator. Because the individual in the database was a close relative, the Sacramento police had a suspect in under two hours, and Waller was quickly arrested.[65]

The head of Parabon's genealogy department, CeCe Moore, predicts that hundreds of crimes will be solved using FGG in the coming years, assuming it continues to be legal.[66] While the cost of genotyping crime-scene DNA has dropped considerably, the method remains quite costly due to the significant manpower required to sift through archives to complete family pedigree charts. Few GEDMatch users are close relatives of perpetrators, as was the case in the Waller arrest. But when no other leads are available, FGG may reopen a cold case.

D. How FGG Differs from CODIS

1. The Federal NDIS and CODIS Database Maintained by the FBI

The DNA Identification Act of 1994 established the National DNA Index System (NDIS), which stores the DNA profiles contributed by federal, state, and local forensic laboratories.[67] All 50 states, the District of Columbia, the federal government, the U.S.

---

63. Parabon's proprietary FDS methodology is called Snapshot Genetic Genealogy. *See Snapshot Genetic Genealogy*, SNAPSHOT, https://snapshot.parabon-nanolabs.com/genealogy (last visited Nov. 19, 2019).

64. Aldous, *supra* note 5.

65. Heather Murphy, *How an Unlikely Family History Website Transformed Cold Case Investigations*, N.Y. TIMES (Oct. 15 2018), https://www.nytimes.com/2018/10/15/science/gedmatch-genealogy-cold-cases.html.

66. Antonio Regolado, *"Hundreds" of Crimes Will Soon Be Solved Using DNA Databases, Genealogist Predicts*, MIT TECH. REV. (Sept. 13, 2018), www.technologyreview.com/s/612001/hundreds-of-crimes-will-soon-be-solved-using-dna-databases-genealogist-predicts.

67. Frequently Asked Questions on CODIS and NDIS, FED. BUREAU OF INVESTIGATION, www.fbi.gov/services/laboratory/biometric-analysis/codis/codis-and-ndis-fact-sheet (last visited Nov. 19, 2019).

Army Criminal Investigation Laboratory, and Puerto Rico contribute samples to the database.[68] The Act limits the categories of people whose profiles may be maintained in NDIS and details the quality assurance, privacy, and expungement requirements for participating laboratories. Once a match is identified by the CODIS system, the laboratories involved in the match share information to verify the match and identify the individual.[69] The only information contained in the CODIS database is an identifier of the contributing agency, a unique specimen identification number, the laboratory personnel associated with the analysis, and the "DNA profile."[70]

2. The Limited Value of CODIS STRs for Things Other Than Identification

The DNA profile that CODIS employs is very different from the profile used to apprehend the Golden State Killer. The CODIS database uses short tandem repeats (STRs) to identify individuals.[71] STRs, also known as microsatellites regions of DNA, are between two and six nucleotides in length.[72] For example, in one area, a string of nucleotides such as "gata" might be repeated three times in one person (gatagatagata), but thirteen times in another. These STRs were chosen because they are polymorphic, meaning that there is significant genomic diversity between individuals at these locations. This yields more accurate matches, as it is very unlikely that unrelated people would share the same number of repeats at these loci.[73] Forensic laboratory technicians create a genetic profile from complete STRs that is thought to have a "vanishingly small," but not zero, probability of being shared with another person.[74] The STRs are in non-coding regions of the genome, so they are not directly involved in coding for proteins.[75] However, despite their limited clinical usefulness, it is not accurate to label the STRs as "junk DNA."[76]

---

68. *Id.*

69. *Id.*

70. *Id.*

71. Daniel M. Bornman et al., *Short-Read, High-Throughput Sequencing Technology for STR Genotyping*, BIOTECHNIQUES 1, 1 (2012).

72. *Id.*

73. *Id.*

74. Susan Matheson, *DNA Phenotyping: Snapshot of a Criminal*, 166 CELL 1061, 1061 (2016).

75. The thirteen junk loci, or non-coding alleles, are "stretches of DNA that do not presently recognize traits and are not associated with any known physical or medical characteristics." Williamson v. State, 993 A.2d 626, 639 (2010).

76. Algee-Hewitt et al., *supra* note 37; Edge & Coop, *supra* note 41.

The 20 STRs are not known to contain health-related information, making them much less useful for clinical research than data from SNPs. However, ancestry information can be gleaned from the CODIS markers.[77] Because the locations of genes are not random as once thought (a phenomenon called "linkage disequilibrium"), CODIS markers can be used to predict some health risks and identify genetic profiles from biobanks that match a record in the CODIS database.[78]

Initially, the DNA Profile in CODIS only included STRs at thirteen loci, but as of January 2017, the government gathers data at 20 loci to achieve even higher confidence in matching.[79] These STR alleles are typically analyzed by amplifying the sample through multiplexed polymerase chain reaction, followed by capillary electrophoresis to separate segments.[80] This technique is time and cost-effective, but it does not allow for systematic genotyping of all STR loci.

Because the CODIS database relies on STRs at only 20 non-coding loci, there is not enough genetic information to provide matches to relatives beyond parents and siblings or to distinguish a fourth cousin from an unrelated person.[81] However, as most consumer genetics tests reveal hundreds of thousands of SNPs, identifying more distant relatives becomes possible—third cousins can usually be found, and many fourth cousins can be as well.[82] Because DOJ laboratories do not analyze SNPs during their forensic DNA casework, if they were to use this sort of analysis, it would need to be completed through an outside vendor laboratory.[83]
Even with SNP data, predicting an exact relationship based on shared DNA alone is not always possible, with the exception of identical-twin, parent-child, or full-sibling matches. There are certain relationships that produce similar patterns of shared DNA to each other. For example, a woman who shares 1750 centiMorgans (cMs) of DNA with you could be your half-sister, grandmother, granddaughter, or aunt. Likewise, a first cousin, grandchild, or a great-uncle/aunt/nephew/niece could all share roughly 950 cMs of

77. *See* Algee-Hewitt et al., *supra* note 37; Edge & Coop, *supra* note 41.
78. *See* Algee-Hewitt et al., *supra* note 37; Edge & Coop, *supra* note 41.
79. FED. BUREAU OF INVESTIGATION, *supra* note 67.
80. Bornman et al., *supra* note 71.
81. Edge & Coop, *supra* note 37, at 2-3.
82. *Id.*
83. DEP'T OF JUSTICE, *supra* note 52.

DNA.[84] To predict the type of relationship, other sources of data such as age and death records would need to be used.

3. Familial Searching in CODIS Is of Limited Utility and Requires Additional Oversight

The FBI has discouraged law enforcement from using CODIS to identify partial STR matches. Partial matches occur when no one in the CODIS database matches the crime scene DNA at all 20 loci, but someone in the database matches at perhaps eight or so, indicating they probably are a sibling or parent of the person whose identity law enforcement is trying to determine. So-called "familial searching" in CODIS has been quite controversial, in part because this method produces a high rate of false positives. Also, like FGG, it identifies individuals by their association with people in the offender database, and not because they themselves chose to add their DNA.[85] Conducting these familial searches under CODIS requires greater regulatory oversight, and is limited to "the most serious cases."[86] Because familial searches can be unreliable, many states, such as Colorado, have created a committee that determines when a familial match is suggestive enough to disclose it to local investigators. Some have argued that the additional layers of oversight for CODIS-mediated familial matches should also be required for FGG. Erin Murphy, a Professor of Law at New York University, for example, supports the separation between the local police and the state committee overseeing familial searches to "ensure that incidental findings, such as adoption or non-paternity, are distanced from those in close contact with the family." [87] However, in practice, there is often no investigative reason to disclose such information to relatives, and policies discouraging such disclosure are reasonable.

Law enforcement's use of private databases to identify criminals is viewed by some as an "alarming end-run"[88] around

---

84. *The Limits of Predicting Relationships Using DNA*, THE DNA GEEK (Dec. 19, 2016), https://thednageek.com/the-limits-of-predicting-relationships-using-dna.

85. "[F]amilial searches should be forbidden because they embody the very presumptions that our constitutional and evidentiary rules have long endeavored to counteract: guilt by association, racial discrimination, propensity, and even biological determinism." Erin Murphy, *Relative Doubt: Familial Searches of DNA Databases*, 109 MICH. L. REV. 291, 304 (2010); *see also* Natalie Ram, *The Mismatch Between Probable Cause and Partial Matching*, 118 YALE L.J. POCKET PART 182, 185 (2009).

86. Erin Murphy, *Law and Policy Oversight of Familial Searches In Recreational Genealogy Databases*, 292 FORENSIC SCI. INT'L e5, e6 (2018).

87. *Id.*

88. *Id.* at e7.

forensic databases like CODIS, given that CODIS has many more technical requirements for registering samples, conducting searches, and returning the results to investigating agencies.[89] This argument is misleading. In practice, law enforcement turns to DTC genetic databases only when CODIS does not result in a match, consistent with the recent interim policy on FGG issued by the DOJ. More importantly, FGG differs from CODIS in ways that makes it *more* permissible and less intrusive, and thus not an "end-run" at all. FGG might be avoiding some of the limitations on CODIS, but the limitations are not required outside of CODIS because of the way the comparison samples are obtained. The procedure does not require that the government obtains samples involuntarily from individuals. With FGG, the government merely accesses a public genealogical database, albeit for forensic purposes.

a. The Data from DTC Genetic Databases Do Not Overlap with CODIS

The power of the FGG method lies in the sheer number of people who have contributed samples for DTC genetic tests. Data analysts project that a genetic database only needs to cover approximately 2% of the target population to "provide a third-cousin match to nearly any person."[90] Therefore, using population models that assume no inbreeding and random sampling of participants, researchers "predict that with a database size of ~3 million U.S. individuals of European descent…more than 99% of the people of [European] ancestry would have at least a single third-cousin match and more than 65% are expected to have at least one second-cousin match."[91] As the popularity of GEDMatch rises, and with FamilyTreeDNA announcing that they will allow forensic searching of their database, achieving this 2% target is within reach. With a warrant or subpoena, law enforcement could search 23andMe or Ancestry, which together already have tens of millions of users.[92]

---

89. "Thus, although corporations and individual citizens generate the largest storehouses of personal data today, the government—through its subpoena powers, contractual agreements, and public access to online data—can effectively bootstrap private information into its own domain without contending with the Constitution." Kimberly N. Brown, *Anonymity, Faceprints, and the Constitution*, 21 GEO. MASON L. REV. 409, 410 (2014).

90. Erlich et al., *supra* note 4. Others predict that with as little as 1% of the population genotyped with dense SNP data, accurate identification is possible in the "median" case. *See* Ellenbogen & Narayanan, *supra* note 55.

91. Erlich et al., *supra* note 4.

92. Ancestry.com's website states that AncestryDNA was "[l]aunched in May 2012, [and it] has more than 10 million people in its consumer DNA network, making it the largest in the world." *Ancestry Company Facts*, ANCESTRY, www.ancestry.com/corporate/about-ancestry/company-facts (last visited Nov.

b. FGG Uses Dense SNP Data, Which Contain Information about Disease Risk, Ancestry, and Physical Traits

Many of the panels that DTC companies employ are based on genome-wide association studies (GWAS) from thousands of unrelated individuals. Researchers look for point mutations on thousands of alleles and correlate them with disease risk. GWAS studies have transformed human genetics, with the discovery of thousands of mutations that are associated with increased (and in some cases decreased) risk of developing certain diseases. However, as members of the tested population likely have widely varying lifestyles and exposure to different environmental risks, the predictive effects for complex diseases are often small and in need of updating.[93] Even so, GWAS studies have the unusual scientific feature of being highly reproducible.[94] The ability to make disease-risk predictions from GWAS studies makes the data from the SNP microarrays, specifically those used by law enforcement in FGG, much more rich and sensitive. [95]

c. FGG Corrects the Racial Bias Inherent in the CODIS Database

There is another important way in which the genealogical databases differ from CODIS. The CODIS database is racially biased, due to its significant over-sampling of African Americans. [96] At least 40% of CODIS is comprised of African Americans, making it much more likely that they, and their family members, will be implicated in a crime through a profile in CODIS. Familial partial match searches would render about 17% of the African American population identifiable, as compared to just 4% of the Caucasian

---

18, 2019). From 23andMe's website, it states that "23andMe has more than 5,000,000 customers." *About Us*, 23ANDME, https://mediacenter.23andme.com/company/about-us (last visited Nov. 18, 2019).

93. The predictive ability of risk evaluation from GWAS studies "depends on the number and effect size of the loci associated with the probability of developing a given phenotype, and has to date been found to generally be modest for most multifactorial conditions." Joel Krier et al., *Reclassification of Genetic-Based Risk Predictions as GWAS Data Accumulate*, 8 GENOME MED. 1, 2 (2016).

94. Urko Merigorta, Replicability and Prediction: Lessons and Challenges from GWAS, 34 TRENDS IN GENETICS 504, 504 (2018).

95. Murphy, *supra* note 86, at e5.

96. Curtis, *supra* note 29, at 2; *see also* Henry T. Greely et al., *Family Ties: The Use of DNA Offender Databases to Catch Offenders' Kin*, 34 J.L. MED. & ETHICS 248, 258-259 (2006) (noting that African American suspects are four to five times more likely to be identified through CODIS searches than white Americans).

always sided with the authors.[233] The individual's First Amendment right to "tell her story" usually trumps the family member's expectation of privacy. This is true even when the subject whose private information is being revealed had no part in the construction of the memoir.[234] One court went even further, suggesting that whether or not the underlying material is newsworthy is irrelevant to whether the author has the right to reveal "her own identity."[235] It is presumed that she possesses this right, even if exercising it means revealing private information about those close to her.

These cases are instructive as applied to our genomic information. When I choose to obtain genetic tests to complete my genetic story, either to assist in precision genetics treatments or to connect with distant relatives, I am writing *my* story. I am gathering information and sharing it because it is my autonomous choice. This decision might indirectly implicate the privacy of others, and it might hurt them. Consider for example, a family who identifies as Native American only to discover through genetic ancestry testing, that there is no evidence, at least according to Ancestry's algorithm, that this family folklore is true. Do I have a privacy interest in my family's ancestry story? Perhaps. Will the law recognize this and protect it from disclosure by a family member? No. Just as with memoir, where "it is often difficult, if not impossible, to separate one's intimate and personal experiences from the people with whom those experiences are shared," it is likewise difficult to separate out which traits or SNPs run in your family, what those SNPs predict in isolation and in your particular pedigree, and what that shared genetic information reveals about any one person.[236] Put differently, Alex Wexler can certainly publish her story about her family's experience with Huntington's disease without being liable for an invasion of privacy claim. The case law is clear on this point. However, we are dealing with an even more amorphous kind of

---

233. "[T]here is an additional interest in this case: Kaysen's right to disclose her own intimate affairs. In this case, it is critical that Kaysen was not a disinterested third party telling Bonome's personal story in order to develop the themes in her book. Rather, she is telling *her own* personal story-which inextricably involves Bonome in an intimate way. In this regard, several courts have held that where an autobiographical account related to a matter of legitimate public interest reveals private information concerning a third party, the disclosure is protected so long as there is a sufficient nexus between those private details and the issue of public concern." Bonome v. Kaysen, No. 032767, 2004 WL 1194731, at *6 (Mass. Super. Ct. Mar. 3, 2004).

234. There is a First Amendment privilege to publish truthful information of legitimate public concern, and the privilege encompasses dissemination of information relating even to individuals who have not sought or who have attempted to avoid publicity. *See* Campbell v. Seabury Press, 614 F.2d 395, 397 (5th Cir. 1980).

235. *Anonsen*, 857 S.W.2d at 705.

236. *Bonome*, 2004 WL 1194731, at *6.

privacy threat. When my genetic story becomes even less concrete, and involves describing *potential* genetic risks rather than actual disease symptoms, the right to privacy for my relatives becomes even weaker.

Of course, one could argue that when police surreptitiously submit crime-scene DNA samples, they are forcing someone else to "tell their story" and connect with family members against their wishes. In one sense, this is true. But once again, given that the DNA is significant evidence of having committed a criminal *actus reus*, the privacy rights in that DNA are substantially diminished.[237] It hardly seems controversial ethically, and is certainly true legally, that a criminal has relinquished the right to tell his own story with his DNA. The law is utilitarian when it comes to privacy, not absolute. Because the purpose was to identify a murderer, the public certainly has an interest in the disclosure of the crime-scene DNA to GEDMatch to find a relative, if not in knowing its specific content.[238]

In conclusion, existing law would not require anything like familial consent before law enforcement could use FGG. Genomes are individually unique constructions, and the value of an entirely different genome, with a different combination of mutations and life experiences, cannot meaningfully be compared through the tenancy by the entirety metaphor. Assuming that our genes have independent property value, without any accompanying information about environmental risk or the interactive effects of our genes, relies on genetic determinism, wherein we assume we can know someone from their genes alone.[239]

Despite its allure, the property metaphor (which I assume is not meant to be literally implemented) just does not work well in the context of sensitive and highly idiosyncratic DNA information. Specifically, alienation, valuation, transfer, license, sale, etc., would be unwieldy. Not only does the case law make clear that family members can reveal photographs and private, scandalous family histories, but it really must be this way.

If family members were required to obtain the consent of their relatives before obtaining genetic test results or sharing them,

---

237. Hodge, *supra* note 146.

238. "[A]n involuntary loss of privacy is recognized in the modern formulations of this branch of the privacy tort, which require not only that the private facts publicized be such as would make a reasonable person deeply offended by such publicity but also that they be facts in which the public has no legitimate interest." Haynes v. Alfred A. Knopf, Inc., 8 F.3d 1222, 1232 (7th Cir. 1993).

239. This concept will be discussed in more detail in Section H.

would that extend only to immediate family members, or to anyone with shared DNA, including some holdout who could be a fourth cousin? Could a fourth cousin keep me from sharing my BRCA-1 results with my granddaughters? Given that most people do not consider their fourth cousins to be within the boundary of their "family," and given that without a service like Ancestry they would likely not even know who these fourth cousins were, it is a bit preposterous to suggest that their consent would be legally (as opposed to ethically) required *ex ante.*

Most importantly, in the context of precision medicine, it would be appalling to require an individual to receive her siblings' consent before undergoing genetic testing. Even without genetic testing, information about particular diseases (such as breast cancer, or depression) is already known to run in families. Family members often know their history; the genetic information may just confirm whether they are a carrier or at risk, but it's not a surprise that this risk is familial. Given the way this risk data is understood and communicated, an approach that treats it like information is far preferable.

G. Why Do Scholars Fear Genetic Informants?

1. Our Fear of Genetic Informants Reflects Moral Dumbfounding

As explained in the sections above, very few of the privacy concerns surrounding FGG stand up to scrutiny. In many cases, the arguments against FGG illustrate a form of "moral dumbfounding,"[240] where opponents find the method morally questionable but cannot locate a good argument as to why. Jonathan Haidt described this phenomenon as thinking like a "lawyer trying to build a case rather than a judge searching for the truth."[241] Moral dumbfounding leads to weak rationalizations for things that seem intuitively immoral (and may indeed *be* immoral, but for reasons we cannot articulate).

In some cases, critiques of FGG, based on an intuitive sense that it is immoral, have worked. Legislators in Maryland proposed bans on this useful prosecutorial tool because of inflated privacy concerns. These concerns reflect a misunderstanding of criminal

---

240. Jonathan Haidt et al., *Moral Dumbfounding: When Intuition Finds No Reason*, in 1 No. 2 LUND PSYCH. REP. 1, 1–29 (Dep't of Psychology, Lund Univ. ed., 2000).

241. Joshua May & Victor Kumar, *Moral Reasoning and Emotion*, in THE ROUTLEDGE HANDBOOK OF MORAL EPISTEMOLOGY 139, 142 (Karen Jones et al. eds., 2019).

procedure and of the *status quo ante*. The fact that such bright legal minds can be so confused about the actual privacy implications of FGG intimates that there is something deeper at work. This is more than just a failure to analyze case law precedent on the Fourth Amendment, witness testimony, the police use of deception, or the invasion of privacy.

2. Our Fear of Genetic Informants Reflects Genetic Essentialism

The reason, I suggest, that leads to these fears is subconscious endorsement of genetic essentialism. First, essentialist thinking leads people to believe that genetic explanations are truly exceptional. Therefore, relevant precedent is immediately and always distinguishable. Second, it leads people to believe that genes are more deterministic than they actually are. This in turn encourages an overemphasis of genetic causes of behavior over environmental factors and a need to protect these immutable blueprints from disclosure. Third, when people think about genes in essentialist ways, they may adopt genetic vitalism, where an almost mythic agency or intentionality is attributed to our genes. Vitalism may explain the view that FGG is "dystopian" or "creepy," as we consider our genes to be unintentional informants. This might explain why some fail to disambiguate the inanimate DNA from the individuals' actions, which allowed others to access the DNA.

Essentialist thinking has been demonstrated in many different cultures and contexts. We all engage in it to a degree. However, some people are more essentialist than others, and some topics, such as genetics, tribalism, or race, lead people to engage in more rigidly essentialist thinking. Essentialists find greater causal power in people's fixed characters than in their surroundings, and assume that much of one's behavior can be explained with reference to the essence that they have.[242] This sort of thinking is on display when we think that something represents "who we truly are," in a way that is "deep down and internal," "naturally determined," which "draws the boundaries" between social groups, and can be "transferred from individual to individual while preserving their original identity."[243] Natural things have an essence; synthetic things do not. In this way, essences are often conceived of as something like a soul, like *chi* in Chinese cultures or *prana* among Hindus.[244] Given these chief features, it is no wonder that genetics so handily lends itself to essentialism.

---

242. Steven Heine et al., *Essentially Biased: Why People are Fatalistic About Genes*, 55 ADVANCES IN EXPERIMENTAL SOC. PSYCHOL. 137, 142 (2017).

243. *Id.*

244. *Id.* at 148.

It seems some of the critiques of FGG sound in the type of genetic essentialism that treats genetics as completely exceptional. This might explain why scholars have failed to properly consider precedent regarding privacy and criminal procedure. They may assume that our genomes cannot be properly analogized to *anything that has come before*. In some ways, our genomes are special and present novel concerns about privacy. Namely, while we might disclose our genomes today, the predictive power of our genomes will continue to grow. Ten years from now, we will likely know even more about disease risk and penetrance than we do now, and that makes sharing genetic data related to disease risk very different. However, when SNP data are used to identify relatives, it is not taking on a form that is radically different from public archives, peer-to-peer sharing of pornographic files, or searching of social media posts. When it comes to the ways genetics is being used in the service of FGG, there are no sufficiently strong reasons to justify its exceptional treatment.

Genetic essentialism leads people to do a number of strange things. For one, they will "over-attribute a person's characteristics and behaviours, in all of their complexity, to their genetic makeup,"[245] and they will tend to view genetic risk factors as being more causal, or deterministic of outcomes, than they actually are.[246] This has the related effect of discounting other causes of disease, behavior, or identity, such as culture, diet, or the environment.[247] There are a handful of diseases wherein genes operate in hard, deterministic ways, and for which you could actually have "a gene for" the disorder that means you will definitely develop that condition. These Mendelian traits are actually pretty rare. Even so, "people overgeneralize from these to the far more common conditions where genes are not at all deterministic."[248]

Genetic determinism might move us to accord genes more heightened privacy protections than are warranted for their specific use. In the deterministic view, genes are perceived to so completely control our future that this justifies our need to keep them secret. I strongly suspect that some of the "moral dumbfounding" surrounding FGG has to do with precisely this. We worry about people having access to genetic information because we assume that

---

245. Ilan Dar-Nimrod et al., *Genetic Knowledge Within a National Australian Sample: Comparisons with Other Diverse Populations*, 21 PUB. HEALTH GENOMICS 133, 134 (2019).

246. Wren Gould & Steven Heine, *Implicit Essentialism: Genetic Concepts Are Implicitly Associated With Fate Concepts*, PLOS ONE, June 2012, at 1.

247. Ilan Dar-Nimrod & Steven J. Heine, *Genetic Essentialism: On the Deceptive Determinism of DNA*, 137 PSYCHOL. BULL. 800, 800 (2011).

248. Heine, *supra* note 242, at 150.

there is a one-to-one relationship between genotype and phenotype, and that if we have a mutation that increases the risk of depression or addiction, we will certainly develop depression and addiction.[249] This makes genetic information much more sensitive, as it is perceived to provide a crystal ball for our future.

In reality, the SNPs that are used to make disease predictions explain very little of the variance between individuals, so this data is currently of limited individual predictive value.[250] Most complex diseases are caused by many different mutations that interact with one another in as-yet-unknown ways. DTC results also fail to capture varying genetic expression, which is affected by lifestyle choices and environmental risks.[251] However, for some diseases, such as Alzheimer's, the identified variants are linked to a substantial increase in the risk of developing that disease, and they are thus more clinically useful. In the future, DTC companies might rely on sequencing the entire genome, which would provide vastly more information on rare diseases than that which is gleaned from just SNPs alone. The potential for using genetics in increasingly predictive ways means we should be mindful of the privacy and security of our genetic data. But this sort of prediction is not inherently at stake with FGG.

Long before we knew about the structure of the double helix, biologists debated whether biology could be explained solely by the principles of physics and chemistry.[252] Many eighteenth and nineteenth century biologists theorized that biology contained metaphysical and spiritual properties, which might never be knowable through the laws of the hard sciences. This came to be known as "vitalism," where biological processes were imbued with agency, desires, and ultimate goals.[253] Vitalism is now largely discredited and dismissed as superstitious.[254] Even so, vitalism is alive and well in our folk understandings of biology. Without being able to prove this directly, I suspect that some of our suspicions of the "genetic informant" or the "biowitness" have to do with our subtle and subconscious misattribution of agency and intent to our genes.

---

249. *Id.*; *see also* Mark Henderson, *'Fat' Gene Found by Scientists,* THE TIMES (U.K) Apr. 13, 2007, www.thetimes.co.uk/article/fat-gene-found-by-scientists-vbf7scwhhnn.

250. Park et al., *supra* note 27.

251. Bunnik et al., *supra* note 43, at 344.

252. Marc Kirschner et al., *Molecular Vitalism*, 100 CELL 79, 79 (2000).

253. Monica Greco, *On the Vitality of Vitalism*, 22 THEORY, CULTURE & SOC'Y 15, 16 (2005).

254. *Id.* at 15.

Anthropomorphism is a "false positive bias" where we over-attribute human-like characteristics such as agency, intention, purpose, or volition, to objects that possess no such capacities.[255] It can happen with inanimate objects, as well as with non-sentient organisms. Even if it is incorrect, it is the result of an over-active and fascinatingly adaptive cognitive process.[256] Because the anthropomorphism tendency is so pervasive, involuntary, and deeply-rooted, it is studied in many different disciplines. Within cognitive psychology, it is called "teleological obsession" or "overactive intentionality bias," and is employed to explain why people presume intentionality where there is no evidence of its presence.[257] Within psychiatry, it is known as "hyper-mentalizing" where it sheds light on why people with schizophrenia or other mental illnesses might imbue mailboxes and computers with secret intentions and surveillance capacities.[258] As applied to biological processes, this tendency is often referred to as vitalism.[259] Regardless of what you call it, there is consensus that anthropomorphism is expressed at such a young age, and is so powerful and automatic, that it feels (and may indeed be) innately hardwired.[260] We can therefore be forgiven for engaging in the particular form of anthropomorphism that implicitly recognizes intentional action in our genes.

If one thinks of FGG as employing DNA as "genetic informants" or "biological witnesses,"[261] it calls to mind something out of a dystopian science fiction novel. The use of these anthropomorphizing labels suggests that it is in fact the *DNA itself* that is testifying against you. This is of course deeply unnerving, in part because of humanity's inherent desire to control its environments. If an independent and intentional genetic informant can speak for us in this way, this reveals something profoundly scary—that we humans are not fully in control of our actions or their consequences. This is particularly troubling when the consequences seem so great. Rather than being able to decide whether to implicate our relatives in a crime, our genes are seemingly making this decision for us, without our consent. I suspect that this phenomenon is also at the root of our response to FGG.

Viewing FGG as employing "genetic informants," as many of its opponents do, exploits the psychological fear of losing control

---

255. Marco Antonio Correa Varella, *The Biology and Evolution of the Three Psychological Tendencies to Anthropomorphize Biology and Evolution*, FRONTIERS IN PSYCHOL., Oct. 2018, at 1.

256. *Id.* at 10.

257. Varella, *supra* note 255, at 1–4.

258. *Id.* at 1, 2.

259. Kirschner et al., *supra* note 252.

260. Varella, *supra* note 255.

261. Kayser, *supra* note 165, at 45.

of our environment and our free will.[262] Humans have a strong desire to "master their environments by increasing the environment's predictability" through its "apparent controllability."[263] Put simply, if we are not controlling our genes, we fear that they are controlling us. The need to feel in control of our environments is referred to as "effectance motivation," and is a big driver of anthropomorphism.[264] Attributing human-like properties to biological entities enables a sense of familiarity and, therefore, control over them.[265] This in turn provides a feeling of comfort when interacting with these agents.[266] Ironically, we employ anthropomorphisms to make us less afraid of losing control. However, because our brains evolved to over-apply it in novel contexts, it can actually make us more afraid of unintentional agents, like robots or genes.[267]

Of course, our genes are not intentionally testifying against us. But given the pervasiveness of anthropomorphizing, especially toward a biological phenomenon like genes, it would not be at all surprising if this causes our strong negative reaction to FGG. The language of the "genetic informant" or "biowitness" unnecessarily imbues the DNA with properties that it does not possess. For our purposes, these terms are misleading because they obfuscate the intentional actions of ordinary people.

III. CONCLUSION: REFORMS THAT BETTER ADDRESS THE REAL PRIVACY CONCERNS OF FGG, WITHOUT HAMPERING THE PROSECUTION OF SERIOUS CRIMES

Secondary use of genetic information by private actors without consent should certainly give us pause. And the unlimited potential uses by the government should also generate considerable worry. But by focusing on the Fourth Amendment concerns of FGG when it is used exclusively by law enforcement to identify the perpetrators of crime, privacy advocates emphasize the wrong privacy boogeymen and thus propose the wrong remedy. We should not ban FGG when used to identify suspects in cold cases. We might, however, decide that DTC genetic companies, while technically consumer enterprises, might share enough features with clinical care

262. Nicholas Epley et al., *On Seeing Human: A Three-Factor Theory of Anthropomorphism*, 114 PSYCHOL. REV. 864, 872 (2007).

263. *Id.*

264. Adam Waytz et al., Making Sense by Making Sentient: Effectance Motivation Increases Anthropomorphism, 99 J. PERSONALITY & SOC. PSYCHOL. 410, 410 (2010).

265. *Id.*

266. *Id.* at 879.

267. *Id.* at 872.

to legally require additional consent measures before people participate. We might also limit the secondary uses of DTC genetic information in some meaningful ways. I briefly sketch out the justifications for these very different reforms below.

a. Statutory Secondary Use Restrictions

I suspect it would be different if, rather than using regions of homozygosity and ancestry information, law enforcement connected the family through a rare and serious disease mutation that ran through the family, and was evidenced in health records or a SNP profile. If law enforcement identified and publicized that the GEDMatch user was affected, this would likely run afoul of legally recognized privacy rights. But even if this *had* been the case, there is no indication that law enforcement would ever need to reveal the rare mutation discovery to the GEDMatch user, just as they never identified Joseph DeAngelo's distant relative who provided the genetic connection that led to his arrest.

The possibility that law enforcement, or other commercial entities, could glean clinically relevant information from GEDMatch profiles is troubling. And yet, because users voluntarily uploaded their SNP data to an open-access, amateur genealogical website, they have no recourse under existing law, at least against secondary data users.[268] In order to better address this sort of harm, we must educate consumers more about the vast amounts of data they are sharing with sites like GEDMatch, or even 23andMe and Ancestry. There also should be significant secondary use limitations that delineate how law enforcement may and may not use the genetic profiles that they, and other private actors, have obtained. Perhaps we need a federal statute that allows law enforcement or private entities to use these SNP profiles only to identify perpetrators after they have committed crimes, to identify the remains of bodies,

---

268. The relatives of the family member (proband) who uploaded her data to GEDMatch may theoretically have a right of action in tort law against the proband. For example, if I upload my SNP profile to GEDMatch, or some other third-party site with insufficient privacy protection by design, my daughter might be able to sue me for public disclosure of private facts, if the state court were willing to find that this sort of disclosure was embarrassing, unreasonable, and at least reckless. However, this might not be recognized, depending on how strong the public interest is in the disclosed information. *See* Cox Broadcasting Corp. v. Cohn, 420 U.S. 469, 487–91 (1975); Time, Inc. v. Hill, 385 U.S. 374, 383 n.7 (1967); *see also* Virgil v. Time, Inc., 527 F.2d 1122, 1127 (9th Cir.1975). The purpose of the tort remedy is to protect the individual against unwarranted publication of private facts. The individual's right to privacy must be balanced with the privilege of the press to publicize matters of public interest that arise out of the desire and "the right of the public to know what is going on in the world and the freedom of the press and other information agencies to report it." *See generally* RESTATEMENT (SECOND) OF TORTS § 652 (1979).

or to reunite displaced persons or victims of human trafficking with their families. There could conceivably be limitations on the types of crimes for which it could be used, based on political sentiment.

As far as we know, law enforcement has not yet used SNP data to do anything other than identify suspects of serious crimes. If we fear that law enforcement could use this data to predict health traits or conduct internal research, then the solution is not to ban FGG. It is not just the police who have access to this rich SNP data. *Anyone* who has access to SNP profiles could upload genetic SNP data to a public site and search for the sample's relatives. This is because GEDMatch and FamilyTreeDNA do not require any identity authentication.

Rather than banning FGG, the solution to this sort of problem is to craft thoughtful legislation that limits certain secondary uses of DTC genetic information. Obvious candidates would be prohibiting the use of SNP data to price health, life or disability insurance, or increasing the fines associated with detection of these uses. Unfortunately, the federal statute prohibiting the use of genetic information to make health insurance coverage or employment decisions might not account for low detection rates, and the scope of GINA's broad exceptions have been insufficiently fleshed out.[269] Another candidate would be requiring specific consent to use samples in research. A large part of 23andMe's business model rests on the secondary use of genetic samples for health research. By sharing SNP data with outside researchers, promising research is being pursued that will help us treat cancer or heart disease more effectively. Not all secondary uses are sinister. But we might reconsider the model where simple disclosure of many terms and acceptance by consumers is considered legally sufficient.

b. Mandatory Updates of the Consent Process for DTC Genetic Tests

Multiple reviews of the consent process for DTC genetic testing reveal that the user agreements and consent procedures are wholly inadequate, especially given that no physician is made available to explain the risks and benefits.[270] There is a reason that countries like France, Germany, Luxembourg, Poland, and Portugal demand that physicians be involved in the ordering of any genetic

---

269. Bradley A. Areheart & Jessica L. Roberts, *GINA, Big Data, and the Future of Employee Privacy*, 128 YALE L.J. 710, 728 (2019). For example, employers can use genetic information that they obtain inadvertently or via commercially available documents, like newspapers that contain obituaries.

270. Clayton et al., *supra* note 14.

test.[271] Regulators in those countries appreciate the enormous risks posed by the inappropriate delivery and interpretation of this sort of information. Consumers of DTC genetic tests may not understand how to interpret the various findings. There are mutations associated with tiny relative risk increases, mutations with unknown expression or population penetrance (because many DTC tests are run in asymptomatic individuals, the likelihood of someone developing the disease associated with a mutation is unknown), or mutations of unknown significance. Further, DTC tests have varying levels of laboratory and clinical quality. Some are not reliable, meaning that many runs of the test will produce varying results, and some are not valid, meaning that the results do not mean what we think they do. We also might worry about this sort of sensitive information being relayed without the possibility of counseling to put the findings in context. For these reasons and others, many countries require that genetic susceptibility tests only be performed as part of clinical care. These countries therefore require physician involvement and patient-informed consent.[272]

A recent review of European regulations found that fourteen countries had specific requirements for informed consent for DTC genetic testing. Some of these countries require meaningful disclosure of all relevant risks and benefits, a right "not to know" the information should they change their mind before the results are delivered, and appropriate time be given before the test and after disclosure for the individual to consider consenting.[273] There is considerable variation among countries regarding which DTC genetic tests require informed consent, what must be present in the consenting process, and the consequences of non-compliance. In Germany, for example, failing to provide advance, express written consent is punishable by imprisonment or a fine. [274]

Regrettably, the United States has not taken this approach. Instead, in the U.S., DTC genetic testing operates largely outside of the realm of "health care," despite interacting with it in many ways. Consumers who undergo testing may become patients who seek physicians' help confirming or interpreting the results. They also may use the testing to decide whether to have a preventative procedure like a hysterectomy or whether to become involved in clinical research. Another large number of them may never use the data in an officially clinical way, but it might subtly change the way

---

271. Louiza Kalokairinou et al., *Legislation of Direct-to-Consumer Genetic Testing in Europe: A Fragmented Regulatory Landscape*, 9 J. COMMUNITY GENETICS 117, 123 (2018).

272. *Id.*

273. *Id.*

274. *Id.* at 124.

they think about their health, and for-profit entities might exploit their data in ways that affect the individual's access to care in the future. Even so, DTC testing in the U.S. is considered mostly recreational. Thus, testing generally does not require physician involvement or clinical informed consent. This is the root of the problem. Despite perhaps being ethically required, outside the context of clinical care and research, the requirement for informed consent is not widely recognized.[275]

As a result, a few different research teams have documented the consent processes for DTC consumer tests and found them troublingly flawed. Indeed, the use of the phrase "consent processes" is misleading, as most DTC genetic testing companies just employ user agreements that consumers must click on in order to use the services or website. Third-party sites like GEDMatch have been found to have particularly poor procedures for documenting agreement with the terms of service.[276] For most DTC genetic tests, users find privacy protections important and desire control over the dissemination of their genetic information. However, they mistakenly assume that the DTC genetic companies are fully protecting their privacy rather than sharing their data with third-party researchers.[277] A study of DTC genetic testing companies targeting Canadian consumers found that "67% provided information insufficient for consumers to determine how their data and sample would be treated."[278] This is alarming, and certainly indicates that a significant part of our trouble with FGG might stem from concerns over meaningless consent at initial stages of testing.

Although more companies now meet guidelines relating to transparency regarding data security protocols, few companies disclose which secondary uses of users' data would be permitted, with whom they would contract, how long the information would be stored, and what might happen in the event that the company was sold or went bankrupt. This finding was supported by smaller, more in-depth studies, which revealed failures to convey the risks of re-

---

275. Bunnik et al., *supra* note 43, at 343.

276. Lauren Badalato et al., *Third Party Interpretation of Raw Genetic Data: An Ethical Exploration*, 25 EUR. J. HUMAN GENETICS 1189, 1190 (2017).

277. Clayton et al., *supra* note 14; Juli Murphy Bollinger et al., *Attitudes About Regulation Among Direct-to-Consumer Genetic Testing Customers*, 17 GENETIC TESTING AND MOLECULAR BIOMARKERS 424, 424-28 (2013); Saskia Sanderson et al., *Public Attitudes Toward Consent and Data Sharing in Biobank Research: A Large Multi-site Experimental Survey in the US*, 100 AM. J. HUMAN GENETICS 414, 414–27 (2017); Sara Chandros Hull et al., *Patients' Views on Identifiability of Samples and Informed Consent for Genetic Research*, 8 AM. J. BIOETHICS 62, 66 (2008).

278. Amanda Singleton et al., *Informed Choice in Direct-to-Consumer Genetic Testing (DTCGT) Websites: A Content Analysis of Benefits, Risks, and Limitations*, 21 J. GENETIC COUNSELING 433, 433 (2012).

identification and to obtain proper consent for the secondary use of data.[279] The privacy and security risks are critically important, given the market value for this sort of data and the temptation for unsavory companies to hack into databases to glean genetic and phenotypic information for research.[280] Unfortunately, these concerns have only been amplified with the increasing use of more modern and thorough genetic sequencing technologies.[281]

It is not uncommon for commercial entities to exploit the appearance of being "medical" to confuse consumers into thinking the experience has the legitimacy and scientific backing of the healthcare system. Based on a review of their advertising practices, many DTC genetic testing companies seem to do exactly this.[282] Several were found to exploit the professional legitimacy of a clinical encounter to establish trust between the company and the consumer, while in their small print user agreement they disavowed many of the regulatory protections that come with clinical experiences such as informed consent, HIPAA compliance, and notification and opportunity to withdraw data if the user no longer wishes to participate in the research database.[283] Clinical care obviously differs considerably from consumer advertising, both in its ethical and legal requirements. It is important for consumers to know what they are buying and sharing.

Medical ethics requires that clinicians respect the autonomy of their patients by disclosing all relevant risks and benefits of a procedure or test *before* it is performed. This has become so standard that the law of every state now requires informed consent in order for physicians to be free from either negligence or battery liability. More modern takes on informed consent stress that disclosure, wherein the clinician provides information in written form and obtains a signature by the patient, should not be a mere formality. Rather, to affect meaningful informed consent, patients ought to be able to understand enough of the procedure to ask questions and have time to reflect before making a decision. They also ought to be given the opportunity to change their consent status, when possible.

---

279. Linnea I. Laestadius et al., *All Your Data (Effectively) Belong to Us: Data Practices Among Direct-to-Consumer Genetic Testing Firms*, 19 GENETICS IN MED. 513, 513 (2017); Emilia Niemiec & Heidi Carmen Howard, *Ethical Issues in Consumer Genome Sequencing: Use of Consumers' Samples and Data*, 8 APPLIED TRANSLATIONAL GENOMICS 23 (2016).

280. Laestadius et al., *supra* note 279; Niemiec & Howard, *supra* note 279.

281. Niemiec & Howard, *supra* note 279.

282. Manuel Schaper & Silke Schicktanz, *Medicine, Market and Communication: Ethical Considerations in Regard to Persuasive Communication in Direct-to-Consumer Genetic Testing Services*, 19 BMC MED. ETHICS 56 (2018).

283. *Id.*

However, none of these legal or ethical standards have been found to apply to DTC companies because they are regulated as if they provide a purely commercial service. This is a mistake. It may be that part of the reaction to FGG lies in consumer discomfort with law enforcement having access to our genetic information, or a misunderstanding of who has access to this complex information in the first place. This suggests our consent process is not working, and for this and other reasons, it might be unwise to allow consumers to be tested "recreationally" for genetic mutations that carry potential for significant future non-recreational uses.

Because genetic information is so complicated, it is not possible to provide all the relevant risk and benefit information in one setting, particularly if done online. Informed consent for DTC testing ought to include an initial layer of basic information on how the saliva sample will be obtained, the laboratory procedures, and what tests may be run on that sample.[284] Customers should also be told about the risks of false positives, false negatives, and laboratory errors. If consumers affirmatively click to agree to these conditions, then a second layer of information should convey background knowledge on genetic expression and penetrance. This layer should include videos and images to simplify the complex information and identify and use some traits with high and low penetrance and expression as examples. Consumers should then be given notice on what others with this data can do with this information. This layer should include information about GINA, and what it protects. It can also describe how our genetic information does not change over the course of our lifetime and how this information could be used to identify us or others through methods like FGG. Again, the use of interactive videos might be useful to help illustrate the data behind these processes.

Next, if someone is interested in pursuing genetic information about health risks, there ought to be *another* layer of consent where consumers are briefed on the different kinds of health risk data. Specifically, consumers should have to opt-in to receive carrier or susceptibility information regarding penetrant health traits, particularly for which there are no known treatments or for which there is a significant and negative prognosis for either themselves or their offspring. Examples of this would include disclosing risks of Huntington's disease or Alzheimer's disease.[285] It is shocking that companies are not required to run these tests through a medical provider, given the potential psychological impacts of the results.

---

284. Bunnik et al., *supra* note 43.

285. Bunnik et al., *supra* note 43, at 345.

Some consumers may altruistically wish to contribute their genetic information to a database used for research purposes. If consumers agree to participate in a commercial research database managed by the DTC company, and have their samples or personal information used by research teams, then another layer must explain this process in detail. Consumers must be told what types of research might be conducted, what personal information of theirs will be used, how they might be later identified using this data, and with whom the data might be shared. Because genetic information—even when stripped of our names, birthdates, addresses, and other readily identifiable information—still has the potential to identify us and is immutable, consumers should have the opportunity to stop participation in this research at any time. Of course, this cannot undo the research and data sharing that has already occurred, but it is important in the event that a research company is sold to a less reputable company that one may then change the user agreement and protections.

Third-party sites like GEDMatch should also require a more robust consent procedure. If the sites are not conducting the tests themselves, they would not need to discuss laboratory errors and the risks of false negatives and positives. However, they should still be required to explain with whom the data might be shared, how it will be stored, and what data security and privacy measures will be taken to encrypt the data. Even where data can be uploaded anonymously with a pseudonym, users should be informed of the risks of re-identification by triangulating data from birth and death records and other accessible genetic databases.

Regardless of which mutations the consumers agree to have tested or shared, they should be given the DTC company's contact information (a phone number or email), which must be maintained so long as the company is marketing their services, where they can direct questions and be told how and with whom their information will be stored and shared in the future. The fact that U.S. law does not require any sort of informed consent for DTC genetic testing means that the quality of the consent processes vary considerably. Consumers are therefore only protected through consumer protection statutes, which are notoriously weak at preventing misrepresentation.[286]

While the specifics of what would be necessary for meaningful informed consent will need to be worked out in

286. Samuel Issacharoff, *Group Litigation of Consumer Claims: Lessons from the U.S. Experience*, 34 TEX. INT'L L.J. 135, 140 (1999).

particular detail by others, we should at least agree that consumers need better information before they, perhaps naively, upload their genetic information online. We have mistakenly assumed that genetic information could be given directly to consumers, without the involvement of health care professionals. Unlike the measures taken in Europe and elsewhere, we have also failed to require specific privacy and security protections for data shared online.

Millions of Americans have already submitted their SNP profiles to public sites like GEDMatch, exposing colossal amounts of personal data to strangers. The review of the DTC companies' consent processes reveals that many fall short of providing adequate information. That means that countless consumers are submitting their saliva or cheek cells, being tested, and then sharing their genetic information online, without appreciating the full risks of doing so. This is the true problem with FGG. The many privacy concerns advanced in opposition to FGG would largely subside if we made sure that consumers had meaningfully consented to the initial test, and that they had appreciated the volume of data contained in our genomes and how it was being shared. If consumers appreciated the privacy, security, and re-identification risks of genetic testing when they *first* submitted their saliva or blood sample, as well as when they then uploaded their data to a third-party site, we could feel more confident in law enforcement's use of this data in their investigation of crime. But our initial suspicions about the quality of the consumers' consent to DTC testing is spilling over into a misplaced concern about FGG. Coupled with our tendency to attribute intentionality to inanimate objects like DNA, to think that genetics plays a larger role in predicting our futures than the environment does, and to feel discomfort with not having control over who (or what) "speaks" for us, we have allowed far too much fear of genetic informants to take hold. Privacy advocates have misdiagnosed the problems with FGG and have thus proposed the wrong remedy—banning FGG. Rather than feeding into this unwarranted fear and banning the methodology, we must directly address the risks at their roots. We can do this by shoring up consent to DTC genetic testing and passing legislation aimed at prohibiting certain secondary uses of genetic data.