Routledge
Taylor & Francis Group

Current Issues in Criminal Justice

ISSN: 1034-5329 (Print) 2206-9542 (Online) Journal homepage: https://www.tandfonline.com/loi/rcic20

Policy and regulatory implications of the new frontier of forensic genomics: direct-to-consumer genetic data and genealogy records

Nathan Scudder, Dennis McNevin, Sally F. Kelty, Christine Funk, Simon J. Walsh & James Robertson

To cite this article: Nathan Scudder, Dennis McNevin, Sally F. Kelty, Christine Funk, Simon J. Walsh & James Robertson (2019) Policy and regulatory implications of the new frontier of forensic genomics: direct-to-consumer genetic data and genealogy records, Current Issues in Criminal Justice, 31:2, 194-216, DOI: 10.1080/10345329.2018.1560588

To link to this article: https://doi.org/10.1080/10345329.2018.1560588

Published online: 01 Mar 2019.

Submit your article to this journal

Article views: 107

View related articles

View Crossmark data

CURRENT ISSUES IN CRIMINAL JUSTICE
2019, VOL. 31, NO. 2, 194–216
https://doi.org/10.1080/10345329.2018.1560588

Policy and regulatory implications of the new frontier of forensic genomics: direct-to-consumer genetic data and genealogy records

Nathan Scudder [a,b], Dennis McNevin [c], Sally F. Kelty [d], Christine Funk[e], Simon J. Walsh[b] and James Robertson [a]

[a]National Centre for Forensic Studies, Faculty of Science and Technology, University of Canberra; [b]Australian Federal Police, GPO Box 401, Canberra, Australia; [c]Centre for Forensic Science, School of Mathematical and Physical Sciences, Faculty of Science, University of Technology Sydney, Australia; [d]Centre for Applied Psychology, Faculty of Health, University of Canberra, Australia; [e]Attorney at Law, Saint Paul, Minnesota, United States

ABSTRACT

Law enforcement is moving from targeted forensic DNA analysis to more extensive use of genomics in support of criminal investigations and for related purposes, such as the identification of human remains. The field of forensic genomics is data-driven and will continue to evolve as new capabilities are developed and new datasets are made accessible. Intelligence capabilities using forensic genomics include the prediction of externally visible characteristics and biogeographical ancestry, and the relatively new field of forensic genetic genealogy. This technique expands these capabilities by accessing public genetic datasets to identify potential relatives of the donor of DNA relating to an investigation. This exploitation of public datasets poses a range of ethical, legal and privacy challenges. The extended reach of these techniques expands these issues to entire families, across multiple jurisdictions. These legal challenges increase as attention turns to much larger, but less accessible, genetic data held by direct-to-consumer genetic genealogy providers.

KEYWORDS

DNA databases; forensic DNA; forensic genetic genealogy; forensic genomics; forensic intelligence; genetic privacy.

Opportunities for law enforcement to exploit larger datasets are increasing (Ferguson, 2017). Law enforcement and security agencies now have at their disposal a range of new tools to help identify suspects, or to prevent, disrupt or deter crime, from automated trawling of open source content to the application of face or voice identification processes to digital recordings (Ferguson, 2017; Joh, 2014, pp. 61–63). They are also managing evidence and information holdings which, in terms of their volume, complexity and speed of acquisition, could be categorised as 'big data' (Moses & Chan, 2014).

Human genetic analysis has progressed significantly in recent years, and the availability of cost-effective technology continues to open up new forensic opportunities (Børsting & Morling, 2015; Smith, 2018). Large genomic datasets, incorporating a significant number of base variants, are of a magnitude which can be considered big data, and law

CONTACT Nathan Scudder ✉ Nathan.Scudder@canberra.edu.au National Centre for Forensic Studies, Faculty of Science and Technology, University of Canberra, ACT 2617, Australia

enforcement use of this information can be viewed as a logical – even inevitable – step from the current, targeted forensic DNA analysis (Drabiak, 2017, pp. 176–181).

In this article we examine the differences between law enforcement use of other intelligence sources, which may be rich in personal information, and genomic data. We will explain recent operational success in using forensic genomics to identify suspects and map how this technology may progress in future, and the policy, ethical and regulatory issues that will arise.

What is forensic genomics?

Forensic science has gradually shifted from targeted genotyping of so-called 'junk DNA' to a wider exploitation of the human genome (Smith, 2015, pp. 100–105; Stajano, Bianchi, Liò, & Korff, 2008). Initial establishment of DNA profiling for criminal investigations deliberately used repeating segments of DNA called 'satellites', not believed to encode information capable of predicting health or physical characteristics, to create a statistical-based model for identification. The term *forensic DNA analysis* has been widely used to refer to this capability, and has allowed law enforcement applications to differentiate themselves from more privacy-intrusive genomic applications (Butler, 2015; Murphy, 2015, pp. 3–18).

Initially forensic DNA analysis looked for exact matches with a person of interest or in a law enforcement database, making no attempt to use the inherited nature of DNA, a technique applied successfully in DNA parentage testing for decades. However, comparing DNA recovered from crime scenes with individuals in a police database and then hypothesising that the donor may be a close relative was a technique successfully applied in the United Kingdom as early as 2004 (Maguire, McCallum, Storey, & Whitaker, 2014, pp. 65–66; Smith & Urbas, 2012). This technique also led to the arrest of the 'Grim Sleeper' in California in 2010 (Mitchell, 2018; Murphy, 2010). The approach – termed *familial DNA searching* – relies on kinship likelihood ratios, an assessment of the probability of common genetic markers being shared between a questioned DNA sample and a putative family member of the donor given a proposed familial relationship. Familial DNA searching can also make use of paternal and maternal inheritance – for example, analysis of the Y chromosome – to further define a familial link (Liberty, 2015, pp. 478–479). These approaches, however, work only for higher-order family relationships (Ram, 2015, p. 902; Smith & Urbas, 2012, p. 65).

The wider exploitation of informative genetic markers, as a component of legal enquiries, can be broadly termed *forensic genomics* (Stajano et al., 2008). This new capability, which we will describe, increases not just the volume of data but inherent possibilities. More genetic points of analysis means distant relatives are within reach of familial searching (Murphy, 2018; Stajano et al., 2008). Further, genomic analysis of particular markers, known as informative markers, can allow prediction of a donor's biogeographical ancestry (BGA) – where their ancestors likely came from – and externally visible characteristics (EVCs), such as natural hair and eye colour variation (Kayser, 2015; Koops & Schellekens, 2008).

Just add data

Law enforcement in the 1980s and 1990s was arguably looking at the tip of the DNA iceberg. There were sound reasons for this. As a new, relatively expensive and resource-intensive technique, exploiting more than a handful of genetic markers was not technically feasible. The low sensitivity of initial techniques limited use to visible stains rich in DNA and required intrusive sampling of body fluid or tissue from individuals associated with an enquiry. These concerns were mitigated by using a relatively small number of genetic markers that were thought at the time to be uninformative for anything other than individualisation and gender (Williams & Wienroth, 2017).

It is only in recent years, with newer DNA genotyping platforms, that moving to a forensic genomics approach has become more cost effective. Increased sensitivity now allows analysis of even trace amounts of cells. As the application of this technology increases, moving from major investigations to cold cases and eventually to more routine casework, the arguments distinguishing forensic DNA analysis from more privacy-intrusive forms of genetic analysis cannot be maintained. Forensic genomics is data-driven. Even a profile generated using targeted forensic DNA analysis has the potential to match against, and thereby re-identify, other genetic data (Humbert, Huguenin, Hugonot, Ayday, & Hubaux, 2015; Kim, Edge, Algee-Hewitt, Li, & Rosenberg, 2018). Law enforcement and policy-makers will need to grapple with these issues now as these capabilities reach even limited operational use (Debus-Sherrill & Field, 2018; Smith, 2018).

The identification capabilities of forensic genomics will continue to increase as datasets – both genetic and non-genetic – are overlayed.

An application of forensic genomics: forensic genetic genealogy

Recent application of forensic genomic techniques to casework has demonstrated significant potential. This is particularly the case where a data integration approach has been adopted.

A leading example of this is *forensic genetic genealogy*, the use of expanded genetic profile data in conjunction with genealogy investigation (Fitzpatrick & Yeiser, 2013). This technique involves three interrelated data analysis concepts or enhancements: use of high-density genotyping; exploiting genealogy; and use of publicly available genetic datasets.

Use of high-density genotyping

The first concept is simply an expansion of the familial searching described above. The use of new, microarray-based DNA technology moves the analysis from a few dozen genetic markers to between half a million and a million markers, vastly improving data volumes and, consequently, utility (King & Jobling, 2009).

Of course, there are technical constraints to the application of this technique, particularly for trace DNA analysis, as micro-arrays typically require at least an order of magnitude more DNA than testing using conventional forensic DNA analysis. While a 'forensic chip' for DNA intelligence was first reported in 2013 (Keating et al., 2013), it should not be regarded as a straight swap-out of existing DNA equipment. But, where it can be applied, it is akin to moving the output of DNA analysis from a postcard to a novel. When compared to reference data, that novel can tell us quite a lot. It can be used to predict BGA and EVCs as well as degrees of genetic relatedness (Zieger & Utz, 2015). For samples from male donors, it can also identify Y haplogroups which – apart from confirming male lineage – may give an indication of the donor's surname (Erlich & Narayanan, 2014; King, Ballereau, Schürer, & Jobling, 2006).

The analysis of high-density data to determine relatedness uses the concept of *identity by descent*: a technique comparing segments of DNA in high-density genetic data to determine a potential common ancestor (Erlich, Shor, Pe'er, & Carmi, 2018; Phillips, 2018). In addition to direct matching against individuals in a criminal DNA database and, through familial DNA searching, their immediate family members (as shown in Figures 1a and 1b), this technique has the potential to expand the matching capabilities well beyond immediate family, potentially to third or fourth cousins of the donor. This could comprise many hundreds of distant relatives (Court, 2018).

There is a problem here, however. When police collect DNA from a suspect or convicted offender and upload it to a criminal DNA database, they are uploading the postcard, not the novel. In Australia, there are detailed statutory requirements around both voluntary provision of a sample by a suspect, or another individual for elimination purposes, as well as the circumstances under which a suspect or convicted offenders can be compelled to provide a sample (Smith & Urbas, 2012, p. 77). While legislation is silent on the method of DNA analysis applied to those samples, the clear intent is to allow for matching of specific markers within a self-contained criminal DNA database.

Notwithstanding this, if DNA databases allowed for the recording of Y-chromosome, mitochondrial and other relevant autosomal DNA sequences for both crime scene and reference DNA samples, they would then be able to generate an investigative lead similar to the following:

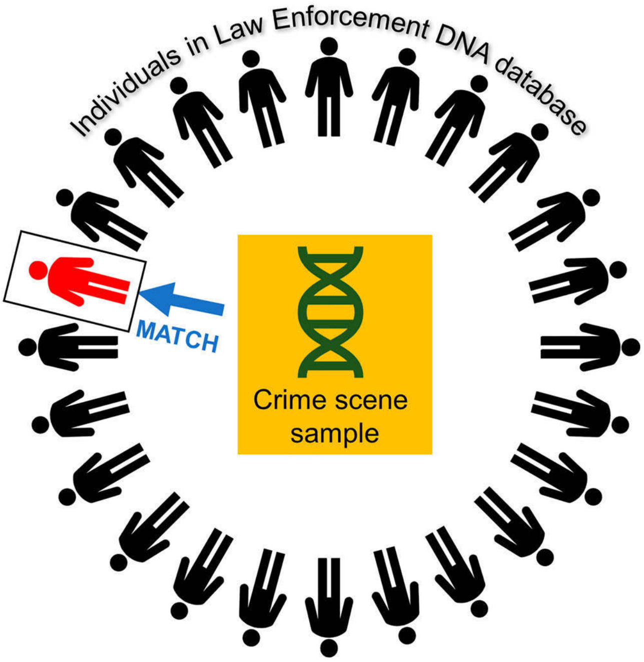

Figure 1a. Use of police DNA databases for direct comparison between reference and crime scene profiles

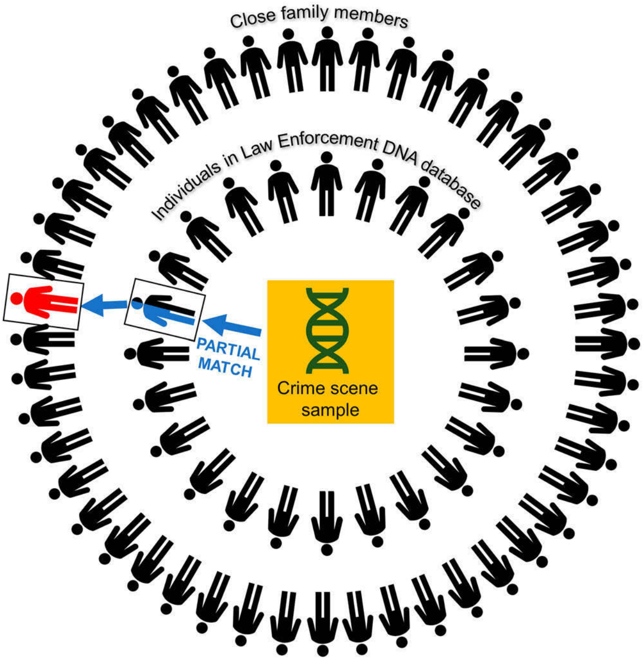

Figure 1b. Use of familial searching in combination with a close match in a police DNA database.

1. Crime scene sample ‘A’ did not match any individuals whose profiles are in the criminal DNA database.
2. Crime scene sample ‘A’ has segments of DNA in common with convicted offender ‘B’, whose DNA is in the database. Based on the amount of shared DNA, the donor of sample ‘A’ could be a second cousin of offender ‘B’ (or relative of equivalent genetic proximity).

Increasing the potential pool of suspects from individuals who have provided a DNA sample, either voluntarily or by direction of the state, to their extended families raises significant ethical and privacy issues. Identity by descent analysis has been demonstrated operationally for family relationships at least as far as third cousins, whose common ancestor would be one set of great-great-grandparents. It is feasible to extend the technique to even more distant relatives (Court, 2018; May, 2018; Pierce, 2018). Erlich, Shor, Pe’er, & Carmi (2018) analysed relationships within a dataset of 1.28 million individuals. Within donors of European descent, they found more than 60% revealed at least one potential third cousin or closer relative within the dataset. In this context, the Personal Genome Project – with its stated aim of publishing 100,000 human genomes – could in itself reveal thousands of distant familial relationships (Court, 2018).

The number of samples required by law enforcement to cover an entire population would be relatively small. Erlich, Shor, Carmi, & Pe’er (2018, p. 4) predict that comparing a suspect’s DNA to a database representing only 2% of total population size would, provided the suspect is from that general population group, almost certainly reveal a potential third cousin or closer relative. In fact, each suspect or offender could bring with them the potential to partially match against hundreds or possibly thousands of distant relatives: living or dead, local or abroad.

Law enforcement use of familial DNA searching, even to a limited extent using current markers, has attracted criticism (Murphy, 2010, p. 319; Ram, 2015) and calls to consider an appropriate regulatory regime (Smith & Urbas, 2012, pp. 78–79). Any proposal which would dramatically extend the reach into distant family members would attract a high degree of public scrutiny.

As Williams and Wienroth (2014) and Court (2018) note, the question of the rights of relatives has already attracted some judicial commentary in the European Union, in the case of *S and Marper*.[1] Murphy (2010) analysed likely constitutional considerations in the United States, both under the Equal Protection clause in the Fourteenth Amendment and protection against unreasonable searches in the Fourth Amendment. However, while familial DNA searching has attracted judicial scrutiny, there does not appear to be a firm basis for a constitutional or human rights challenge to existing practices.

It is possible that any shift to forensic genomics may lessen the likelihood of a successful constitutional challenge in the United States. It has been argued that familial DNA searching creates a sub-class of individuals, closely related to suspects or convicted offenders, and that this will have a greater impact on minority and disadvantaged groups in society who may be proportionally over-represented in criminal DNA databases (Murphy, 2010, pp. 321–325). However, this argument may logically diminish as forensic genomics extends the technique genetically further and further from suspects and offenders (Murphy, 2018).

The United Kingdom's National DNA Database contains just over 6 million DNA profiles in a population of 65 million people (Wiles, 2018). If these 6 million samples had been processed with high-density DNA technology – creating perhaps a few dozen terabytes of genetic data – then a very high percentage of the entire United Kingdom population would be within genealogical reach of at least one stored profile. Such a dataset would be five times the size that Erlich, Shor, Carmi, & Pe'er (2018) predict gives near-universal genealogical reach for a population group.

Rutherford (2018) discusses the complexities and interconnectedness of families through the centuries. The idea of a subset of the population being subject to law enforcement scrutiny through familial DNA searching while the remainder of the population are immune starts to diminish the closer we come to a near-to-whole population saturation point (Berkman, Miller, & Grady, 2018, p. 1078; Ram, Guerrini, & McGuire, 2018). Interestingly, Hazel, Clayton, Malin, & Slobogin (2018) recently argued for a community-wide genetic database, principally for criminal investigative purposes.

There remain significant privacy and ethical challenges to any increased scope of criminal DNA database, including risks of unintended exposure of health-predictive information. It could prove or disprove parentage or be used to establish or attempt to disprove BGA (Morrison, 2017).

Another potential privacy concern would be a temptation towards deliberate analysis of such a genetic dataset to, for example, identify potential offenders *a priori* by their predicted BGA, or to search for common genetic markers which may be linked to a predisposition to criminality (Kaye, 2006). Once such a dataset exists, it may be difficult to

[1] [2008] Eur Court HR 1581.

deflect attempts to gain access – at least to a de-identified version – to undertake such studies. Kaye (2006) notes that some legislation in the United States may preclude such analysis, and this situation likely applies in other jurisdictions, including under frameworks such as European Union General Data Protection Regulation.[2]

If it were attempted, such genetic analysis could have a significant and detrimental impact on individuals, families and communities by confusing socio-economic causes of criminality with genetic causes. Researchers in Germany in the decades before the Second World War published on race and eugenics, attempting to give a level of scientific support to government action against the Jewish population and other minorities (Ehrenreich, 2007).

Any genetic analysis of a dataset of offenders and suspects would need to be approached with great caution.

Exploiting genealogy

Identifying that sample 'A' may have been deposited by a second cousin of 'B' provides no actionable lead without overlaying other data – particularly genealogy information. While law enforcement would generally have access to government records of births, deaths and marriages – at least for their own jurisdiction – today's online genealogy market provides a ready opportunity to identify potential suspects through high-density genotyping, and then to narrow down that list based on other demographic information (Fitzpatrick & Yeiser, 2013).

If we suspect that the person who deposited the sample and 'B' are second cousins or share another relationship of similar genetic distance (such as first cousins twice removed) we then hypothesise that they share at least one common ancestor, most likely a set of great-grandparents. To identify suspects, it would therefore be necessary to build that family tree upwards three generations from 'B' (Figure 2a).

It is then necessary to build the four family trees of the great-grandparents. These family trees would then include a significant number of individuals, being biologically related grandparents, parents, great uncles and aunts, uncles and aunts, siblings, and first and second cousins of 'B' (Figure 2b).

Finally, using other investigative tools, it is then necessary to narrow in on potential suspects. Some individuals in those family trees may be deceased or living abroad. Some may be too close or too distant in their genetic relationship. Other available information or intelligence, such as eyewitness reports estimating the approximate age of the suspect, could narrow the pool even further. If successful, the technique will yield a small enough list of potential suspects to allow police to take further overt or covert action (Figure 2c). This could include obtaining a further DNA sample from the identified suspect for comparison with the original crime scene sample 'A'.

Use of publicly available genetic datasets

We have discussed some of the legal, ethical and technical challenges to any change to law enforcement databases, including increasing the number of autosomal, mitochondrial and

[2] (EU) 2016/679.

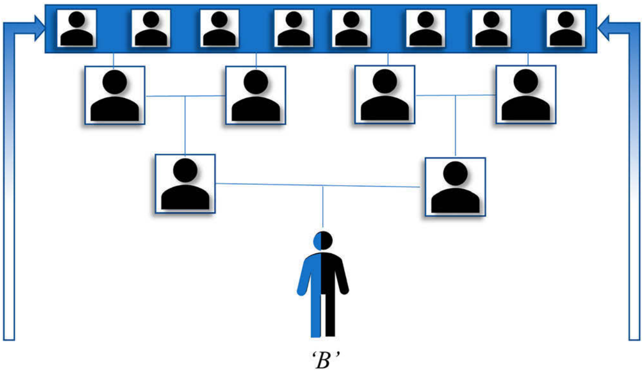

Figure 2a. Building the family tree for 'B' to identify a possible second cousin match. (Using technique described by Moore (2016)).

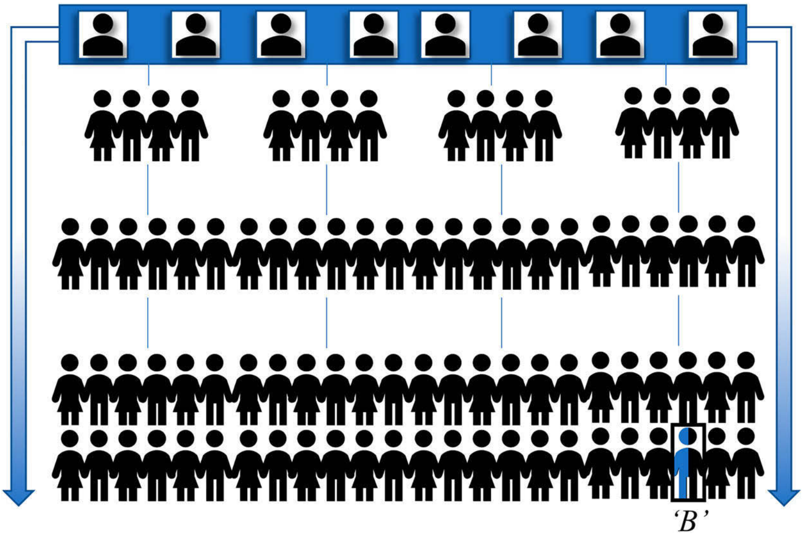

Figure 2b. Building the family tree for 'B' to identify a possible second cousin match.

Y chromosome markers for suspects and convicted offenders. Given the amount of genetic data stored, laws requiring an individual to submit to such genetic testing would challenge many of the long-held justifications and safeguards for law enforcement use of DNA, particularly that it is largely directed towards non-coding segments of the genome (Murphy, 2015). It could therefore be seen as an overreach or a disproportionate response.

Just as the European Court of Human Rights closely considered the use of suspect and offender DNA in the *Marper* case, discussed previously, in *Maryland v King*[3] the United States Supreme Court also considered the question of compulsory DNA samples from arrestees. In upholding the law in Maryland, the court noted [at 464] the use of 'noncoding parts of the DNA that do not reveal the genetic traits of the arrestee. While science can

[3] 569 U.S. 435 (2013).

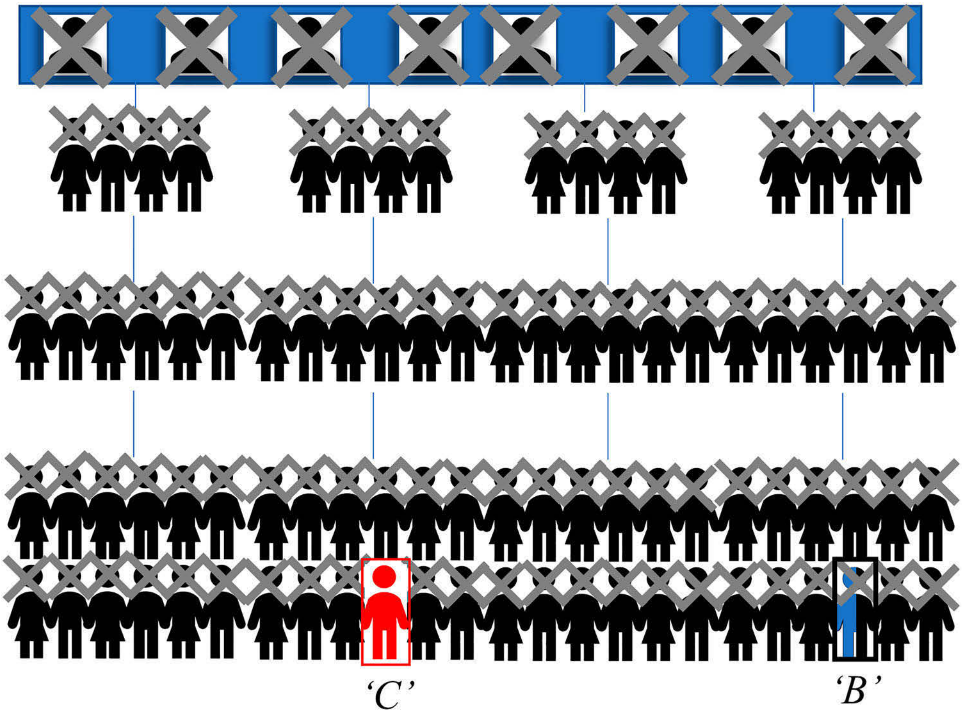

Figure 2c. Narrowing potential suspects based on other leads to ultimately identify a candidate for further DNA testing.

always progress further, and those progressions may have Fourth Amendment consequences, alleles … "are not at present revealing information beyond identification'".

Four Justices dissented and, in a decision written by the late Justice Scalia, [at 482] expressed doubt that 'the proud men who wrote the charter of our liberties would have been so eager to open their mouths for royal inspection'.

But what if the genetic data was already freely available to all? Genetic genealogy has progressed at a remarkable pace in recent years. Individuals are seeking to supplement searches for family tree documentary records with genetic testing, identifying likely relatives and common ancestors (King & Jobling, 2009).

There are millions of high-density genetic test results held by online genetic genealogy databases worldwide (Ball et al., 2018; Ney, Ceze & Kohno, 2018). Each of these profiles has the potential to identify familial relationships and therefore to lead law enforcement to the donor of a sample found at a crime scene or to the relatives of an unidentified deceased person.

Investigators with access to expertise in forensic genetic genealogy are now making use of publicly available genetic data, and there have been some notable investigative successes (Murphy, 2018; Skwarecki, 2018).

Case Example: traversing the East Area Rapist's family tree

Joseph James DeAngelo was arrested in California in April 2018 and is alleged to have committed at least 12 murders and 45 sexual assaults in the 1970s and 1980s

(Wamsley, 2018). The use of genealogy to arrest the suspected East Area Rapist, also known as the Golden State Killer, involved many months of work for police and genealogists (Selk, 2018). Investigators re-analysed a crime scene sample using microarray technology, generating data compatible with genealogy providers. They then uploaded the data file to a public genetic genealogy website called GEDmatch, identifying possible third cousins (Berkman, Miller, & Grady, 2018; Wamsley, 2018).

The use of publicly available genealogy information allowed police and genealogists to then build more than two dozen family trees, going back to DeAngelo's great-great-great grandparents (Arango, 2018). A suspect list was then generated, with investigators gradually narrowing their focus to DeAngelo. Before arresting the suspect, covert samples were obtained from a door handle on Mr DeAngelo's car and later from discarded rubbish to verify the match (Stanton & Smith, 2018).

Case Example: identifying the 'Buck Skin Girl'

In 1981, a young woman's body was recovered, wearing a buck skin jacket, on a roadside in Ohio. She was buried in a 'Jane Doe' grave and, despite police and public efforts over many decades, remained unidentified.

In 2018 the 'DNA Doe Project' began further work on the 'Buck Skin Girl' case. This charity had been set up through crowdfunding during 2017 to help identify unknown human remains. Using a similar process of forensic genetic genealogy, the team uploaded genetic data generated from a vial of the victim's blood to GEDmatch. The search identified a possible first cousin, once removed. This, in turn, led to a genealogy website where a user had entered the death details of a relative, of a similar age, as 'Unknown Missing-Presumed Dead' (Augenstein, 2018a).

In only a few hours, a team of genealogists had provided police with a solid investigative lead. Following DNA testing of immediate family, Marcia L. King, aged 22, was identified as the 'Buck Skin Girl'.

Law enforcement use of big data

Babuta (2017) argues that big data analytics is required 'when data is collected on such a large scale that it cannot be analysed with traditional data-management tools and methods'. Scale might include the volume of data, its complexity or data integration considerations. Forensic genomics, and particularly the application of forensic genetic genealogy, draws on some of these concepts, albeit a large part of the process is still performed through time-intensive manual review of records.

This is not to say that the process could not be largely automated in the future. Kaplanis et al. (2018) demonstrated a technique which, with minimal manual manipulation, generated family trees, the largest of which spanned 13 million individuals related through birth or marriage.

Moses and Chan (2014, pp. 645, 677–678) assessed the use of big data for investigative purposes from a technical, social and normative viewpoint, concluding that the inferences drawn are 'not neutral' and applications require critical analysis. Ferguson (2017, pp. 187–201) describes five possible questions to be asked around inputs, outputs and technology when considering applying a big data solution to policing. Both approaches highlight the need for accountability, transparency and data quality.

Current uses of forensic genetic genealogy arguably fail these tests, particularly through their reliance on commercial or privately managed online tools. GEDmatch changed its sign-up process in May 2018 to add a category for law enforcement profiles (Skwarecki, 2018), which would now allow the site's administrators to track law enforcement use more closely. However, earlier profiles uploaded by or on behalf of law enforcement were not as clearly categorised. In terms of accountability and data quality, the searches are wholly reliant on privately developed web-based tools, which are unlikely to have been externally validated.

Is this aspect unique to genetic datasets? Law enforcement is reliant elsewhere on data sourced from commercial and private entities, and social media companies frequently engage proactively with law enforcement (Facebook Inc., 2018). At least when approaching use for evidentiary purposes, arguably genetic data has an inbuilt safeguard – in the form of confirmatory DNA testing using conventional forensic DNA analysis – which does not exist for other data sourced from commercial providers.

There are few safeguards, for example, if a technology company erroneously provided law enforcement with the wrong user details in response to a request – e.g. for telephone records – implicating an innocent customer. Such errors are not limited to commercial providers. United Kingdom law enforcement in 2017 erroneously provided a passport photograph to an overseas law enforcement agency, ultimately resulting in an entirely different individual's photo being placed online as a fugitive with an Interpol Red Notice (Wheatstone, 2018).

Is it acceptable to rely on the taking of a later DNA sample to exclude a suspect who has been identified through forensic genetic genealogy? Are there differences between police analysing genetic records of individuals, compared to other personal information? It could be argued that all of these information-gathering techniques draw individuals into a criminal investigation. Genetic matching is less transparent, and errors or anomalies could be less obvious. There have been instances where familial DNA searching has resulted in a suspect being identified and later excluded (Selk, 2018). Even the East Area Rapist case had an initial lead, where DNA was obtained from an individual in a nursing home (Oremus, 2018). The individual volunteered his DNA and was excluded from the case.

It is important to note that the presence of biological trace material at a crime scene should never be immediately equated with guilt. As demonstrated in the inquiry in Victoria into the miscarriage of justice involving Arah Abdulkadir Jama, there could be a number of plausible scenarios, all of which must be carefully examined and excluded (Gill, 2014, pp. 27–30; Vincent, 2010). DNA evidence could be viewed by police as inculpatory but, in the context of forensic genetic genealogy, its probative value is more closely aligned to forensic intelligence as described by Ross (2015).

Law enforcement use of large genetic and genealogical datasets could also raise legal and ethical issues extending beyond that jurisdiction. Family members residing in other countries may be drawn into an investigation (Kennett, 2018). These family members might never have interacted with the online genealogy or genetic platform, with their name simply added to a family tree by a relative. Of course, if the individual does live abroad with no connection to the country in which the crime occurred, they should be quickly eliminated from enquiries. However, such an association could potentially remain in police indices, and this may even be required should there ever be a need for an investigator to justify any related background checks leading to the individual being eliminated as a suspect.

But such a record, like a bad credit history unknown to the borrower, could have a future impact. It could perhaps come to notice during a later application by that overseas family member for a visa to visit that country. As a result, there is an argument that any intelligence product of this type should, if not destroyed, be quarantined from wider police records.

Notwithstanding this and the considerable academic and media commentary on the policy and ethical considerations of this new forensic technique, Guerrini, Robinson, Petersen, & McGuire (2018, p. 3) conducted a survey in the United States of 1587 individuals in May 2018, finding 79% supported police searches of online genealogy websites ($p < 0.05$). The level of support was similar to police use of mobile telephone records or social media accounts (Guerrini et al., 2018, p. 4).

Mining public genealogy data: different policy models

Accessing publicly available genetic genealogy records from GEDmatch raises questions of third-party privacy. However, the issues are far broader. The use of Facebook data for political analysis by Cambridge Analytica shows the exponential increase in the value of data when an individual grants access to their networks and connections (Halpern, 2016; Moran, 2018). Gifting genetic data to the public (or, in the case of GEDmatch, allowing other users to compare their genetic data against your own) is, as Oremus (2018) explains, quite similar. It is truly a family donation (Krueger, 2018; May, 2018; Ram, 2015, pp. 929–939).

Users of GEDmatch and similar sites are free to withdraw their consent and, generally, sites will facilitate deletion of genetic data preventing any future searches. GEDmatch also allows the use of aliases or screen names when uploading genetic data (GEDmatch, 2018). It is likely that some individuals who have uploaded their genetic data to sites such as GEDmatch are now deceased, leaving what may be a permanent online record. This is particularly the case given the popularity of genealogy research amongst individuals in their retirement years, with some commentators noting an increase in DNA or ancestral tourism (Dickinson, 2018).

Current privacy laws, both in Australia and internationally, do not recognise these familial attributes. Genetic data itself is generally well protected in most jurisdictions, considered an individual’s personal information or, in some cases, a health record (Shoenbill, Fost, Tachinardi, & Mendonca, 2014). But those rights are not enforceable by relatives. While laws and policies around removing or memorialising online records for a deceased family member are catching up, there remain gaps in this area (Harbinja, 2017). The concept of intertwined personal information, where a document or record contains information about two or more people, is applied in a limited sense in the health sector, and broadening this approach was considered by the Australian Law Reform Commission (2003, pp. 238–240). While there have been instances where researchers have withheld historical genetic genealogy records to protect the genetic privacy of current generations (Larmuseau et al., 2016), such an approach would be problematic in the context of online genealogy records. It could only work by allowing someone to prevent upload of that portion of another’s genome which they share through family inheritance and would make online genetic genealogy unworkable.

Current legal considerations

Mining genetic databases gives rise to various ethical considerations. As Smith and Urbas (2012) discuss, forensic procedures legislation in Australia is focused on direct comparison between reference samples from individuals and DNA obtained from crime scenes. Exploiting crime scene DNA using analytical techniques falls outside that regulatory regime. Privacy laws in Australia likewise do not apply until a genetic profile is reasonably identifiable. But, when the donor's identity does become apparent, it will be necessary to consider relevant law enforcement exemptions around use of personal information without consent.

Ram, Guerrini, and McGuire (2018) discuss potential constitutional issues with use of public genetic databases such as GEDmatch in a United States context. Legal arguments precluding law enforcement use are hard to craft, although Ram et al. (2018) state that 'such searches may run counter to core values of American law' and, in particular, may ultimately be caught in a broader interpretation of Fourth Amendment protections.

In the meantime, even strict privacy regulations in Europe, which came into effect in May 2018, have law enforcement exceptions which would appear to make such use permissible in relation to European citizens (Massey, 2017). While Australia appears to be approaching this technology with some caution, law enforcement agencies elsewhere are increasingly assessing and adopting this technique (Graham, 2018; Skwarecki, 2018; The Local, 2018).

Accessing commercial provider data

Public genetic datasets are still a rarity, and some free online genealogy tools are disappearing (Estes, 2018). This trend has been attributed to the commencement of new European privacy laws, although a number of these services were already deprecated or legacy tools (Augenstein, 2018b; Estes, 2018). Current GEDmatch administrators are retirees (Zhang, 2018), and commentators have noted dominant genetic genealogy provider, Ancestry.com LLC, has a history of acquisitions and, for some smaller legacy systems, of sometimes discontinuing tools or databases (Leavenworth, 2018).

In addition to considering regulatory and policy considerations around use of public genetic datasets, it will soon be necessary to consider how to regulate law enforcement access and use information from commercial genetic databases. Accessing commercial databases would move the policy narrative from questioning whether it is ethical to make common ancestor predictions based on genetic information a relative has freely given away to the appropriateness of access to information where users have been given more explicit assurances around privacy and security of their genetic information.

Online providers, including social media and cloud storage companies, have been subject to requests from law enforcement for many years, developing varying policies on their engagement with police (Lynch, 2011). As most of these platforms are based in the United States, case law in this area has generally been litigated in American courts.

Where access is resisted, a warrant requiring production of genetic data would need to satisfy the threshold of probable cause in the United States. An international request would need to satisfy similar requirements but would generally be initiated through a mutual assistance request, for example under the Mutual Assistance in Criminal Matters Act 1987 (Cth). A purely speculative request, effectively trawling for evidence, could and would likely be challenged by online genealogy providers.

But, given the size of major commercial datasets and their extended reach into families, are we approaching a time when one or more family matches can be *expected* for any given search? If this is the case, would a search against records for any criminal investigation now meet the United States probable cause test? Effectively, law enforcement would argue that, for any given sample, there is almost certainly 'evidence' to be collected from any genetic database above a threshold population size.

For law enforcement to seek access to a commercial genetic genealogy provider's records to help identify the donor of a crime scene sample, the request would need to be framed in terms that required the provider to provide a list of users whose genetic data closely matched a genetic profile provided by law enforcement. There have been cases involving technology companies being compelled to assist law enforcement and provide user data. Apple Inc. was compelled by way of a writ issued by a United States Magistrate Judge to develop new software for law enforcement to install on a specific, locked mobile telephone to defeat standard security features (Farivar, 2018, pp. 26–37).[4] The writ also allowed Apple to recover its reasonable costs.

The case involving Apple never went to hearing, due to an alternative means of accessing the data, leaving open whether such a legal strategy could be successfully employed. Recent court action involving Facebook's instant messenger service could, however, reactivate this debate (Levine & Menn, 2018). There is an arguable difference in the level of privacy intrusion in accessing an individual's online communications and accessing their genetic information, or the genetic information of one or more of their family members. This distinction is, as yet, legally untested.

Can legal compulsion be resisted?

One genetic genealogy provider, 23andMe, uses genetic data of consenting customers for biomedical science. 23andMe asserts that genetic records are protected under a Certificate of Confidentiality issued by the United States National Institute of Health (23andMe, 2018). Certificates of Confidentiality protect research data where release to a third party could identify one or more individual subjects involved in the research. Wolf et al. (2015) argue that Certificates of Confidentiality have afforded a high degree of protection, including in criminal matters. However, applicability beyond biomedical research is untested. In addition, the consent arrangements and agreements to share identifiable information differ in relation to genetic genealogy information, when compared to analysis of health-informative genetic data. Courts may well accept a similar distinction in relation to law enforcement access to providers' records.

[4] *In the Matter of the Search of an Apple iPhone Seized During the Execution of a Search Warrant on a Black Lexus IS300, California License Plate 35KGD203* (CD Cal, Eastern Div No 5:16-CM-10, 19 February 2016).

A regulatory regime?

Should law enforcement access to online genetic information, either by directly accessing publicly available information, or by way of legal compulsion, be subject to a regulatory regime? Should it be outlawed?

A blanket ban on forensic genetic genealogy would maintain privacy of individuals who have generally provided no consent for the use of genetic information for the identification of suspects or unidentified deceased persons. However, this must be balanced against the policy benefits of identifying suspects. To date, use of the technique has mostly been limited to cold cases. But this has changed (McCarthy, 2019). Should a law prevent investigators from using public data, if doing so may help identify an active serial killer?

There are policy arguments for limiting the use of the technique. There would be genuine community concern, as it is possible that the scope of searches could progress from the most serious crimes, like serial homicide, to a variety of crime types. As Ram (2011) explained, while policymakers anticipated that familial DNA searching between immediate family members would generally be applied in the investigation of serious criminal offences, the technique was used in the state of Colorado to investigate a case involving theft of loose change from a motor vehicle.

Any regulatory approach, short of such a ban, needs to complement any existing forensic procedures legislation, ensuring the processes used by law enforcement for collecting, using and exploiting DNA capabilities are seamless. There are several options for such a framework:

1. Regulate access to all forms of genetic data, including prescriptive requirements for how law enforcement can use publicly available datasets. It could include processes allowing designated officials to compel online genetic genealogy companies to match data and provide user details.
2. It could adopt a limited approach, similar to the provisions in the Telecommunications (Interception and Access) Act 1979 around telecommunications metadata in Australia, allowing law enforcement to make a request and for genetic genealogy providers to confirm the existence of records, to a stated level of familial proximity. Law enforcement could then apply for a writ or search warrant to access those records.

A regulatory model could include prescribing:

1. Offence types, or a minimum prescribed penalty, for which forensic genetic genealogy could be applied.
2. How distant a familial relationship can be considered as an actionable lead by law enforcement. This would effectively balance the privacy cost to potential relatives drawn into an investigation with the benefit of identifying a suspect.
3. Obligations to maintain privacy of genetic information, which could also be extended to online genealogy providers who have possession of genetic data from crime scenes or unidentified deceased persons.

A regulatory scheme could also be used to put in place quality standards as well as implementing processes of transparency, reporting and oversight. While the privacy concerns around genetic information differ significantly from other forms of online communications, a similar oversight model does exist in Australia with respect to regulation of intercepted telecommunications, with the Commonwealth Ombudsman fulfilling an oversight and audit role (Commonwealth Ombudsman, 2017). Regular publication of audit information promotes public confidence in the transparency of law enforcement processes.

Industry standards

Another alternative would involve use of relevant industry standards. Genealogists, for example, can seek certification in some countries and agree to abide by a code of ethics (Australasian Association of Genealogists and Record Agents Inc., 2015; Board for Certification of Genealogists, 2017). Forensic laboratories have a number of international accreditation and certification options. Establishing an accreditation framework would likely increase community confidence in the use of the technology by ensuring practitioners are well trained and familiar with the risks and limitations of the technique. Processes could be aligned with best practice in forensic analysis and the development of intelligence products.

Industry is already seeking to address some community concerns. In July 2018, several major genetic genealogy providers have endorsed new guidelines requiring transparency and annual reporting of the number of law enforcement requests for information (Future of Privacy Forum, 2018; Romm & Harwell, 2018).

Training for law enforcement

Whether or not a regulatory or standards-based approach is considered, law enforcement agencies will need to carefully consider their approach to training and awareness. Investigators would need to understand the intelligence value and limitations of any leads obtained through genetic genealogy (Ney, Ceze, & Kohno, 2018). The widespread use of online genealogy means that family trees can contain inaccuracies or omissions. Awareness for judicial officers, who may ultimately need to assess information in the context of a warrant application, would also be beneficial.

In terms of investigative doctrine, careful consideration should be given to the advantages and disadvantages of obtaining DNA covertly before arrest, for example, by obtaining a discarded item from a suspect and undertaking forensic DNA analysis to compare to the original crime scene profile. This approach has been used in several cases involving familial DNA searching, including the Grim Sleeper and East Area Rapist investigations (McFerrin, 2012; Stanton & Smith, 2018). Such an approach can mitigate the risks of drawing an innocent person into an investigation but is resource intensive. The first covert sample in the East Area Rapist case was a mixture of DNA from three individuals (Stanton & Smith, 2018), necessitating the collection of a second sample.

What if the covert sample excludes the suspect? Is there a possibility of error in collection? Profiles that appear to exclude a suspect could be subjected to a similar form of forensic genetic genealogy as crime scene samples, effectively reversing the process. If law enforcement believes that an individual is likely a second cousin of ‘B’, but the covert sample reveals no relatedness, then either there is an anomaly in family tree records or – potentially – a chain of custody or contamination issue in the covert collection process.

There are also public policy considerations around the collection of covert samples, particularly when most jurisdictions require a warrant or some other authorisation to require a suspect to provide a DNA sample (Joh, 2006). Is it appropriate that an individual is never made aware that they were a suspect in an investigation and their genetic data was analysed? Does this differ from being drawn into an investigation through other covertly obtained intelligence? Again, concerns could be partially allayed through regulatory or doctrinal processes requiring destruction of genetic information after exclusion, or at least it being quarantined from other police intelligence holdings.

What if there are no leads?

Forensic genetic genealogy already shows a high degree of promise in providing some level of intelligence lead. Statistically, there would be a high chance of identifying a second or third cousin for some of the largest genetic datasets already in existence (Erlich, Shor, Carmi, & Pe'er, 2018). But there will be samples where identity by descent analysis reveals no family matches, or at least none that are close enough to be viable for an investigation. There will be amateur family tree records that are incomplete or erroneous, or even nonsensical. Can law enforcement take the next step and seek to expand online genetic holdings?

The DNA Doe Project, a charity set up to apply forensic genetic genealogy to unidentified human remains, has used social media to encourage individuals in some communities to upload their DNA to expand the number of available profiles. Utah's Cold Case Coalition called on individuals to voluntarily upload their DNA to GEDmatch to help solve crimes (Pierce, 2018).

If testing identified a relative but online genetic and genealogy records were scant, can police adopt a strategy of surreptitiously providing test kits to individuals listed in a particular family tree so as to expand the reach of the technique by inviting them to upload their DNA to a genealogy site? A relative could open their letter box to find a special invitation to claim a free DNA test kit.

Police can use government records of births, deaths and marriages. But might it be more efficient to target a family member who may have uploaded an incomplete family tree, sending them a free genealogy book to re-spark their interest? Such processes, while somewhat manipulative, would likely be legal in most jurisdictions. Individuals would be acting voluntarily if they decide to take up an offer actually being made by law enforcement.

Use of forensic genetic genealogy can also raise operational security implications. While GEDmatch, in particular, allows users to mark profiles as private, if a law enforcement agency did not do so, a suspect could receive an unexpected notification of a new 'twin', or a close relative advice of a new immediate family member. This could reveal to a suspect that law enforcement is likely already traversing their genetic family tree. These approaches would be alleviated by use of a regulatory or warrants-based scheme, which would ensure any searches are conducted directly by the provider.

Conclusion

Exploiting genetic datasets through forensic genomics, including the use of forensic genetic genealogy, opens up a new line of inquiry in many cold-case investigations. Whether individuals have concerns about the use of their genetic data and decide to delete their profiles is yet to be seen.

The approach is not yet mainstream and may well grow. It is timely to explore policy and regulatory responses around the use of these capabilities. Educating law enforcement, the public and the judiciary around this technology and its potential application is also very important in ensuring contemporary debate around the privacy and family implications.

While a technique still very much in its infancy, we have argued that there may come a time when access is not so readily available. Law enforcement needs to carefully consider its response to such a scenario, and regulators need to consider whether existing warrants and mutual assistance schemes sufficiently mitigate intrinsic genetic privacy risks.

It is likely that society is entering a time when, given enough investigative, genetic and genealogical resources, the source of virtually any sample can be ascertained. The privacy and social implications of this will play out over months and years.

Acknowledgements

The authors would like to thank the anonymous reviewers whose constructive criticism improved this article.

Research for this article was supported by an Australian Government Research Training Program Scholarship. This article also draws on research funded by the Endeavour Scholarships and Fellowships, a Department of Education and Training initiative.

Cases

Carpenter v United States 585 U.S. ___ (2018)
Maryland v King 569 U.S. 435 (2013)
Matter of 381 Search Warrants Directed to Facebook Inc., 132 AD3d 11 (NY 2015)
S and Marper v United Kingdom [2008] Eur Court HR 1581
United States v In the Matter of the Search of an Apple iPhone Seized During the Execution of a Search Warrant on a Black Lexus IS300, California License Plate 35KGD203 (CD Cal, Eastern Div No 5:16-CM-10, 19 February 2016)

Legislation

All Writs Act 28 USC § 1651 (1948)
General Data Protection Regulation (EU) 2016/679

ORCID

Nathan Scudder http://orcid.org/0000-0002-6011-9092
Dennis McNevin http://orcid.org/0000-0003-1665-3367
Sally F. Kelty http://orcid.org/0000-0001-9631-1740
James Robertson http://orcid.org/0000-0003-3634-4318

References

23andMe. (2018). *Research consent document*. Retrieved from https://www.23andme.com/about/consent/

Arango, T. (2018, May 3). The cold case that inspired the 'Golden State Killer' detective to try genealogy. *New York Times*. Retrieved from https://www.nytimes.com/2018/05/03/us/golden-state-killer-genealogy.html

Augenstein, S. (2018a). 'Buck Skin Girl' case break is success of new DNA Doe Project. *Forensic Magazine*. Retrieved from https://www.forensicmag.com/news/2018/04/buck-skin-girl-case-break-success-new-dna-doe-project

Augenstein, S. (2018b). Golden State Killer backlash? Public databases shutting down in wake of arrest. *Forensic Magazine*. Retrieved from https://www.forensicmag.com/news/2018/05/golden-state-killer-backlash-public-databases-shutting-down-wake-arrest

Australasian Association of Genealogists and Record Agents Inc. (2015). *Code of ethics*. Retrieved from http://www.aagra.asn.au/code-of-ethics/

Australian Law Reform Commission. (2003). *Essentially yours: The protection of human genetic information in Australia* (Vols. 1 and 2). Report 96.

Babuta, A. (2017). *Big data and policing: An assessment of law enforcement requirements, expectations and priorities*. Retrieved from https://rusi.org/sites/default/files/rusi-bigdata-press-2017.pdf

Ball, C., Battat, E., Byrnes, J. K., Carbonetto, P., Chahine, K. G., Curtis, R. E., … Wang, Y. (2018). *Genetic Communities™ White Paper: Predicting fine-scale ancestral origins from the genetic sharing patterns among millions of individuals*. Retrieved from https://www.ancestry.com/cs/dna-help/communities/whitepaper

Berkman, B. E., Miller, W. K., & Grady, C. (2018). Is it ethical to use genealogy data to solve crimes? *Annals of Internal Medicine*, *169*, 333–334. Retrieved from https://doi.org/10.7326/M18-1348

Board for Certification of Genealogists. (2017). *Genealogist's code of ethics*. Retrieved from https://bcgcertification.org/wp-content/uploads/2017/09/BCG-Code-of-Ethics.pdf

Børsting, C., & Morling, N. (2015). Next generation sequencing and its applications in forensic genetics. *Forensic Science International: Genetics*, *18*, 78–89, Retrieved from https://doi.org/10.1016/j.fsigen.2015.02.002

Butler, J. M. (2015). The future of forensic DNA analysis. *Phil. Trans. R. Soc. B*, *370*(1674), 20140252. Retrieved from https://doi.org/10.1098/rstb.2014.0252

Commonwealth Ombudsman. (2017). *A report on the Commonwealth Ombudsman's monitoring of agency access to stored communications and telecommunications data under Chapters 3 and 4 of the Telecommunications (Interception and Access) Act 1979 for the period 1 July 2015 to 30 June 2016*. Retrieved from http://www.ombudsman.gov.au/__data/assets/pdf_file/0018/45423/TIA-Act-Annual-Report-2015-16.pdf

Court, D. S. (2018). Forensic genealogy: Some serious concerns. *Forensic Science International: Genetics*, *36*, 203–204. Retrieved from https://doi.org/10.1016/j.fsigen.2018.07.011

Debus-Sherrill, S., & Field, M. B. (2018). Familial DNA searching: An emerging forensic investigative tool. *Science and Justice*, *59*: 1, 20–28. Retrieved from https://doi.org/10.1016/j.scijus.2018.07.006

Dickinson, G. (2018, May 31). Is 'DNA tourism' the next travel trend for millennials? *The Telegraph*. Retrieved from https://www.telegraph.co.uk/travel/comment/dna-heritage-ancestral-tourism/

Drabiak, K. (2017). Caveat emptor: How the intersection of big data and consumer genomics exponentially increases informational privacy risks. *Health Matrix*, *27*, 143–183.

Ehrenreich, E. (2007). *The Nazi ancestral proof: Genealogy, racial science, and the final solution*. Bloomington: Indiana University Press.

Erlich, Y., & Narayanan, A. (2014). Routes for breaching and protecting genetic privacy. *Nature Reviews Genetics*, *15*(6), 409–421. Retrieved from https://doi.org/10.1038/nrg3723

Erlich, Y., Shor, T., Carmi, S., & Pe'er, I. (2018). Re-identification of genomic data using long range familial searches. *bioRxiv*, 350231. Retrieved from https://doi.org/10.1101/350231

Erlich, Y., Shor, T., Pe'er, I., & Carmi, S. (2018). Identity inference of genomic data using long-range familial searches. *Science*, 9 Nov 2018, 690–694. Retrieved from https://doi.org/10.1126/science.aau4832

Estes, R. (2018). European Union's General Data Protection Regulation (GDPR). *Ogden family search library*. Retrieved from http://ogdenfsl.org/wp-content/uploads/2018/05/June-2018.pdf

Facebook Inc. (2018). *Facebook and law enforcement*. Retrieved from https://www.facebook.com/safety/groups/law

Farivar, C. (2018). *Habeas data: Privacy vs. the rise of surveillance tech*. Melville House.

Ferguson, A. G. (2017). *The rise of big data policing*. New York: New York University Press.

Fitzpatrick, C., & Yeiser, A. (2013). *Forensic genealogy* (2nd ed.). Fountain Valley: Rice Book Press.

Future of Privacy Forum. (2018). *Privacy best practices for consumer genetic testing services*. Washington DC. Retrieved from https://fpf.org/wp-content/uploads/2018/07/Privacy-Best-Practices-for-Consumer-Genetic-Testing-Services-FINAL.pdf

GEDmatch. (2018). *GEDmatch.Com terms of service and Privacy Policy*. Retrieved from https://www.gedmatch.com/tos.htm

Gill, P. (2014). *Misleading DNA evidence: Reasons for miscarriages of justice*. London: Elsevier Science.

Graham, J. (2018, May 6). Boston police explore using commercial DNA databases. *Boston Herald*. Retrieved from www.bostonherald.com/news/local_coverage/2018/05/boston_police_explore_using_commercial_dna_databases

Guerrini, C. J., Robinson, J. O., Petersen, D., & McGuire, A. L. (2018). Should police have access to genetic genealogy databases? Capturing the Golden State Killer and other criminals using a controversial new forensic technique. *PLoS Biology*, *16*(10), e2006906.

Halpern, S. (2016). They have, right now, another you. *New York Review of Books*. Retrieved from www.nybooks.com/articles/2016/12/22/they-have-right-now-another-you/

Harbinja, E. (2017). People are going to court over dead family members' Facebook pages – It's time for post-mortem privacy. Retrieved from https://theconversation.com/amp/people-are-going-to-court-over-dead-family-members-facebook-pages-its-time-for-post-mortem-privacy-78375

Hazel, J. W., Clayton, E. W., Malin, B. A., & Slobogin, C. (2018). Is it time for a universal genetic forensic database? *Science*, *362*(6417), 898–900. Retrieved from 10.1126/science.aav5475

Humbert, M., Huguenin, K., Hugonot, J., Ayday, E., & Hubaux, J.-P. (2015). De-anonymizing genomic databases using phenotypic traits. *Proceedings on Privacy Enhancing Technologies*, 2015(2), 99–114. Retrieved from https://doi.org/10.1515/popets-2015-0020

Joh, E. E. (2006). Reclaiming abandoned DNA: The Fourth Amendment and genetic privacy. *Nw. UL Rev.*, *100*, 857–884.

Joh, E. E. (2014). Policing by numbers: Big data and the Fourth Amendment. *Wash. L. Rev.*, *89*, 35–68.

Kaplanis, J., Gordon, A., Shor, T., Weissbrod, O., Geiger, D., Wahl, M., … Gymrek, M. (2018). Quantitative analysis of population-scale family trees with millions of relatives. *Science*, *360* (6385), 171–175.

Kaye, D. H. (2006). Behavioral genetics research and criminal DNA databases. *Law and Contemporary Problems*, *69*(1/2), 259–299.

Kayser, M. (2015). Forensic DNA phenotyping: Predicting human appearance from crime scene material for investigative purposes. *Forensic Science International: Genetics*, *18*, 33–48. Retrieved from https://doi.org/10.1016/j.fsigen.2015.02.003

Keating, B., Bansal, A. T., Walsh, S., Millman, J., Newman, J., Kidd, K., … Gasparini, P. (2013). First all-in-one diagnostic tool for DNA intelligence: Genome-wide inference of biogeographic ancestry, appearance, relatedness, and sex with the Identitas v1 Forensic Chip. *International Journal of Legal Medicine*, *127*(3), 559–572.

Kennett, D. (2018). The brave new world of genetic genealogy. *Technology Review*. Retrieved from https://www.technologyreview.com/s/611048/the-brave-new-world-of-genetic-genealogy/amp/

Kim, J., Edge, M. D., Algee-Hewitt, B. F., Li, J. Z., & Rosenberg, N. A. (2018). Statistical detection of relatives typed with disjoint forensic and biomedical loci. *Cell*, *175*(3), 848–858.

King, T. E., Ballereau, S. J., Schürer, K. E., & Jobling, M. A. (2006). Genetic signatures of coancestry within surnames. *Current Biology*, *16*(4), 384–388. Retrieved from https://doi.org/10.1016/j.cub.2005.12.048

King, T. E., & Jobling, M. A. (2009). What's in a name? Y chromosomes, surnames and the genetic genealogy revolution. *Trends in Genetics*, *25*(8), 351–360. Retrieved from https://doi.org/10.1016/j.tig.2009.06.003

Koops, B.-J., & Schellekens, M. (2008). Forensic DNA phenotyping: Regulatory issues. *Colum. Sci. & Tech. L. Rev.*, *9*, 158–202. Retrieved from https://www.doi.org/10.2139/ssrn.975032

Krueger, A. (2018, June 16). Are genetic testing sites the new social networks? *New York Times*. Retrieved from https://www.nytimes.com/2018/06/16/style/23-and-me-ancestry-dna.html

Larmuseau, M. H., Bekaert, B., Baumers, M., Wenseleers, T., Deforce, D., Borry, P., & Decorte, R. (2016). Biohistorical materials and contemporary privacy concerns: The forensic case of King Albert I. *Forensic Science International: Genetics*, *24*, 202–210, Retrieved from https://doi.org/10.1016/j.fsigen.2016.07.008

Leavenworth, S. (2018, May 30). Ancestry has a history of backtracking on promises to customers. *McClatchyDC*. Retrieved from http://amp.mcclatchydc.com/news/nation-world/article210969549.html

Levine, D., & Menn, J. (2018). Exclusive: U.S. Government seeks Facebook help to wiretap Messenger-sources. *Reuters*. Retrieved from https://www.reuters.com/article/us-facebook-encryption-exclusive/u-s-government-seeks-facebook-help-to-wiretap-messenger-sources-idUSKBN1L226D

Liberty, A. A. (2015). Defending the black sheep of the forensic DNA family: The case for implementing familial DNA searching in Minnesota. *Hamline L. Rev*, *38*, 467–518.

The Local. (2018, May 7). Swedish police mull using DNA family tree websites to catch killers. *The Local*. Retrieved from https://www.thelocal.se/20180507/swedish-police-mull-using-dna-family-tree-websites

Lynch, J. (2011). Social media and law enforcement: Who gets what data and when? *Electronic Frontier Foundation*. Retrieved from https://www.eff.org/deeplinks/2011/01/social-media-and-law-enforcement-who-gets-what

Maguire, C. N., McCallum, L. A., Storey, C., & Whitaker, J. (2014). Familial searching: A specialist forensic DNA profiling service utilising the National DNA Database® to identify unknown offenders via their relatives – The UK experience. *Forensic Science International: Genetics*, *8*(1), 1–9. Retrieved from https://doi.org/10.1016/j.fsigen.2013.07.004

Massey, S. R. (2017). *The ultimate GDPR practitioner guide: Demystifying privacy & data protection*. London: Fox Red Risk Publishing.

May, T. (2018). Sociogenetic risks – Ancestry DNA testing, third-party identity, and protection of privacy. *New England Journal of Medicine*, *379*, 410–412. Retrieved from https://doi.org/10.1056/NEJMp1805870

McCarthy, K. (2019). I'm a crime-fighter, says FamilyTreeDNA boss after being caught giving folks' DNA data to FBI. The Register. Retrieved from https://www.theregister.co.uk/2019/02/01/familytreedna_fbi_link/

McFerrin, C. (2012). DNA, genetic material, and a look at property rights: Why you may be your brother's keeper. *Tex. Wesleyan L. Rev.*, *19*, 967–998.

Miller, A. (2018, February 19). College claims discovery of George Washington's hair, but fears DNA testing could destroy historic find. *ABC News*. Retrieved from https://abcnews.go.com/US/college-claims-discovery-george-washingtons-hair-fears-dna/story?id=53198413

Mitchell, M. (2018). As genealogy databases aid in crime-solving, are courts ready to tackle DNA privacy? *Legal Intelligencer*. Retrieved from https://www.law.com/thelegalintelligencer/2018/07/23/as-genealogy-databases-aid-in-crime-solving-are-courts-ready-to-tackle-dna-privacy/

Moore, C. (2016). History of genetic genealogy and unknown parentage research: An insider's view. *Journal of Genetic Genealogy*, *8*(1), 35–37.

Moran, K. S. (2018). Damned by DNA – Balancing personal privacy with public safety. *Forensic Science International*, *292*, e3–e4.

Morrison, P. (2017, March 8). Patt Morrison asks: Rachel Dolezal on racial fluidity and her changing identity. *Los Angeles Times*. Retrieved from http://www.latimes.com/opinion/op-ed/la-ol-patt-morrison-rachel-dolezal-20170308-story.html

Moses, L. B., & Chan, J. (2014). Using big data for legal and law enforcement decisions: Testing the new tools. *UNSWLJ*, *37*, 643–678. Retrieved from https://ssrn.com/abstract=2513564

Murphy, E. (2010). Relative doubt: Familial searches of DNA databases. *Michigan Law Review*, *109* (3), 291–348.

Murphy, E. (2015). *Inside the cell: The dark side of forensic DNA*. New York: Nation Books.

Murphy, E. (2018). Law and policy oversight of familial searches in recreational genealogy databases. *Forensic Science International*, *292*, e5–e9. Retrieved from https://doi.org/10.1016/j.forsciint.2018.08.027

Ney, P. M., Ceze, L., & Kohno, T. (2018). Computer security risks of distant relative matching in consumer genetic databases. *arXiv preprint arXiv:1810.02895*.

Oremus, W. (2018). How the Golden State Killer's DNA search is like the Cambridge analytica scandal. *Slate*. Retrieved from https://amp.slate.com/technology/2018/05/how-the-golden-state-killers-dna-search-is-like-the-cambridge-analytica-scandal.html

Phillips, C. (2018). The Golden State Killer investigation and the nascent field of forensic genealogy. *Forensic Science International: Genetics*, *36*, 186–188. Retrieved from https://doi.org/10.1016/j.fsigen.2018.07.010

Pierce, S. D. (2018, May 30). Utah's Cold Case Coalition wants you to share your DNA with cops to help solve murders. *Salt Lake Tribune*. Retrieved from https://www.sltrib.com/news/2018/05/30/utahs-cold-case-coalition-wants-you-to-share-your-dna-with-cops-to-help-solve-murders/

Ram, N. (2011). Fortuity and forensic familial identification. *Stanford Law Review*, *63*(4), 751–812.

Ram, N. (2015). DNA by the entirety. *Colum. L. Rev.*, *115*, 873.

Ram, N., Guerrini, C. J., & McGuire, A. L. (2018). Genealogy databases and the future of criminal investigation. *ScienceMag*, *360*(6393), 1078–1079. Retrieved from https://doi.org/10.1126/science.aau1083

Romm, T., & Harwell, D. (2018, July 31). Ancestry, 23andMe and others say they will follow these rules when giving DNA data to businesses or police. *Washington Post*. Retrieved from https://www.washingtonpost.com/amphtml/technology/2018/07/31/ancestry-andme-others-say-they-will-follow-these-rules-when-giving-dna-data-businesses-or-police/

Ross, A. (2015). Elements of a forensic intelligence model. *Australian Journal of Forensic Sciences*, *47*(1), 8–15. Retrieved from https://doi.org/10.1080/00450618.2014.916753

Rutherford, A. (2018). You're descended from Royalty and so is everybody else. *Nautilus*, *56*. Retrieved from http://nautil.us/issue/56/perspective/youre-descended-from-royalty-and-so-is-everybody-else

Selk, A. (2018, April 28). The ingenious and 'dystopian' DNA technique police used to hunt the 'Golden State Killer' suspect. *Washington Post*. Retrieved from https://www.washingtonpost.com/amphtml/news/true-crime/wp/2018/04/27/golden-state-killer-dna-website-gedmatch-was-used-to-identify-joseph-deangelo-as-suspect-police-say/

Shoenbill, K., Fost, N., Tachinardi, U., & Mendonca, E. A. (2014). Genetic data and electronic health records: A discussion of ethical, logistical and technological considerations. *Journal of the American Medical Informatics Association*, *21*(1), 171–180.

Skwarecki, B. (2018). Public DNA databases are now crawling with law enforcement and we better get used to it. *Lifehacker*. Retrieved from https://vitals.lifehacker.com/public-dna-databases-are-now-crawling-with-law-enforcem-1826196280/

Smith, M. (2015). *DNA evidence in the Australian legal system*, LexisNexis Butterworths.

Smith, M. (2018). Three stages in the development of DNA evidence in Australia. *Australian Journal of Forensic Sciences*. Retrieved from https://doi.org/10.1080/00450618.2018.1457719

Smith, M., & Urbas, G. F. (2012). Regulating new forms of forensic DNA profiling under Australian legislation: Familial matching and DNA phenotyping. *Australian Journal of Forensic Sciences*, *44* (1), 63–81. Retrieved from https://doi.org/10.1080/00450618.2011.581250

Stajano, F., Bianchi, L., Liò, P., & Korff, D. (2008). Forensic genomics: Kin privacy, driftnets and other open questions. In *Proceedings of the 7th ACM workshop on privacy in the electronic society* (pp. 15–22).

Stanton, S., & Smith, D. (2018, June 1). How detectives collected DNA samples from the East Area Rapist suspect. *Sacramento Bee*. Retrieved from http://amp.sacbee.com/latest-news/article212334279.html

Vincent, F. H. R. (2010). *Inquiry into the circumstances that led to the conviction of Mr Arah Abdulkadir Jama*. Melbourne: Victorian Government Printer.

Wade, N. (1999, July 7). Descendants of slave's son contend that his father was George Washington. *New York Times*. Retrieved from https://www.nytimes.com/1999/07/07/us/descendants-of-slave-s-son-contend-that-his-father-was-george-washington.html

Wamsley, L. (2018, April 27). In hunt for Golden State Killer, investigators uploaded his DNA to genealogy site. *National Public Radio*. Retrieved from https://www.npr.org/sections/thetwo-way/2018/04/27/606624218/in-hunt-for-golden-state-killer-investigators-uploaded-his-dna-to-genealogy-site

Wheatstone, R. (2018, June 10). Innocent Dorset mum horrified to find her 'mugshot' on world's most wanted list after Interpol blunder. *The Sun*. Retrieved from https://www.thesun.co.uk/news/6493985/adriana-barton-weymouth-housewife-interpol-most-wanted-mistake/

Wiles, P. (2018). *Annual report 2017: Commissioner for the retention and use of biometric material*. Retrieved from https://www.gov.uk/government/publications/biometrics-commissioner-annual-report-2017

Williams, R., & Wienroth, M. (2014). *Ethical, social and policy aspects of forensic genetics: A systematic review*. Northumbria Research Link. Retrieved from http://nrl.northumbria.ac.uk/16313/

Williams, R., & Wienroth, M. (2017). Social and ethical aspects of forensic genetics: A critical review. *Forensic Sci Rev*, *29*(2), 145–169.

Wolf, L. E., Patel, M. J., Tarver, B. A. W., Austin, J. L., Dame, L. A., & Beskow, L. M. (2015). Certificates of confidentiality: Protecting human subject research data in law and practice. *The Journal of Law, Medicine & Ethics*, *43*(3), 594–609.

Zhang, S. (2018). How a tiny website became the police's go-to genealogy database. *The Atlantic*. Retrieved from https://www.theatlantic.com/amp/article/561695/

Zieger, M., & Utz, S. (2015). About DNA databasing and investigative genetic analysis of externally visible characteristics: A public survey. *Forensic Science International: Genetics*, *17*, 163–172. Retrieved from https://doi.org/doi:10.1016/j.fsigen.2015.05.010