Operationalising forensic genetic genealogy in an Australian context

Nathan Scudder [a, d, e], Runa Daniel [b, e], Jennifer Raymond [c], Alison Sears [c]

a Australian Federal Police, GPO Box 401, Canberra, Australia
b Office of the Chief Forensic Scientist, Victoria Police Forensic Services Department, Macleod, Victoria 3085, Australia
c Forensic Evidence and Technical Services, NSW Police Force, NSW 2010, Australia
d Centre for Forensic Science, School of Mathematical and Physical Sciences, Faculty of Science, University of Technology Sydney, Australia
e National Centre for Forensic Studies, Faculty of Science and Technology, University of Canberra, Australia

I. Abstract

Forensic genetic genealogy, a technique leveraging new DNA capabilities and public genetic databases to identify suspects, raises specific considerations in a law enforcement context. Use of this technique requires consideration of its scientific and technical limitations, including the composition of current online datasets, and consideration of its scientific validity. Additionally, forensic genetic genealogy needs to be considered in the relevant legal context to determine the best way in which to make use of its potential to generate investigative leads while minimising its impact on individual privacy. This article presents these issues from an Australian perspective, with the observations and conclusions likely to be applicable to other jurisdictions.

II. Introduction

The East Area Rapist terrorised areas of California in the 1970s and 80s. The crimes were committed in communities where doors were once left unlocked (Castillo et al. 2018). The East Area Rapist, later referred to as the Golden State Killer, is reported to have committed at least 12 murders and 45 sexual assaults (Selk 2018). Never really a cold case investigation in the decades since, investigators continued to pursue investigative leads. But it would be a new forensic tool, the application of genetic genealogy, which would eventually lead to the arrest, and guilty plea, of Joseph de Angelo (Guerrini et al. 2018; McLaughlin 2020; Phillips 2018). The use of online family tree genealogy records allowed investigators to narrow their focus to a handful of suspects, ultimately making an arrest.

Scientific, ethical, privacy, security and legal considerations arise as the application of forensic genetic genealogy (FGG) continues to advance. This article will explore those considerations in an Australian legal and scientific context. In addition, the article will explore how these new capabilities may sit within existing forensic and investigative processes. In Section XII a checklist is provided outlining items for consideration by any agency seeking to implement this technique within casework.

Direct-to-consumer (DTC) DNA testing has been offered to consumers for several decades, but high cost meant it was reserved for the most determined of recreational genealogists. Early use sparked the interest of Bennett Greenspan who founded FamilyTreeDNA, the first mainstream genetic genealogy provider (International Society of Genetic Genealogy 2018).

Interest in genetic genealogy has increased substantially since its inception, with dramatic growth in the last few years. In 2017, the number of DTC tests held by commercial providers more than doubled, with estimates of 12 million kits analysed at that time (Regalado 2018).

There is no such thing as a guarantee of absolute privacy (Angrist 2013). Sociologist Gary Marx has described the concept of privacy as ‘a multi-dimensional concept with fluid and often ill-defined, contested and negotiated contours, depending on the context and culture’ (Wright & De Hert 2012). There is a high level of social contract and an assessment of proportionality as privacy laws and processes navigate competing interests of personal privacy and competing societal interests, such as the need for police to find and prosecute criminals (Skinner 2018).

There is increasing international interest, and operational success reporting in the US, in using FGG to identify suspects in cold cases (Aldhous 2019b). This article aims to explore some of the scientific, legal and practical considerations, culminating in a checklist to assist Australian forensic and law enforcement agencies considering operationalising this technique. Although this paper focuses on the Australian context, the general scientific and operational considerations are relevant for any forensic and law enforcement agencies.

III. Managing Stakeholder Expectations

FGG is not a technique that can be applied in every case. Several technical and practical considerations could make the technique unfeasible. In the first instance, most crimes occur between people who know each other (Murphy 2013). The specific circumstances of those cases will determine whether DNA evidence has any probative value. In cases where DNA evidence plays a central role, the use of existing police DNA databases may be sufficient to identify a suspect through routine Short Tandem Repeat (STR) profiling. It may also identify a close family member whose DNA is held by police, through forensic familial searching.

In developing a capability to deliver FGG it is important to manage the competing interests and expectations of victims, government and the broader community. As the processes underpinning forensic genetic genealogy are refined and enhanced, the argument may be put that – with enough scientific and investigative resources – it should theoretically be possible for the donor of any trace of unknown origin to be identified. In practice, this is unlikely to be the case. Clear guidelines to assist individuals to understand the opportunity afforded by FGG, and some of the limitations or obstacles that may arise, will assist in this regard.

IV. Feasibility for Australian Investigators

The two main methods currently used to generate whole genome Single Nucleotide Polymorphism (SNP) data for FGG are Whole Genome Arrays (WGA), also referred to as high density SNP arrays and microarrays, and Whole Genome Sequencing (WGS).

Both WGA and WGS (with the exception of whole mitochondrial DNA sequencing) are relatively new in their use in forensic DNA analysis. For this reason, utilisation of these methods is accompanied by scientific validation considerations.

Validation of forensic DNA analysis methods is guided by recommendations made by the United States Scientific Working Group on DNA Analysis Methods (SWGDAM). SWGDAM's recent statement on Investigative Genetic Genealogy (the alternate term for FGG), released in February 2020, provides an overview of the process (Scientific Working Group on DNA Analysis Methods 2020). However, to date, validation criteria or application considerations for the forensic community have not been developed.

In the absence of validation criteria or guidelines, forensic and law enforcement organisations that use WGA or WGS for FGG must employ a considered approach which should seek to assess, as a minimum, the accuracy and reliability of the methods for various forensic evidence or sample types.

WGA relies on high DNA input amounts and high quality DNA. In contrast to forensic STR profiling which requires half a nanogram of DNA to analyse approximately 20 STRs, WGAs often requires hundreds of nanograms to analyse more than half a million SNPs (Tillmar et al. 2020). Poor quality and degraded DNA samples, often retrieved from compromised biological evidence in cold cases, is expected to impact on the success of the WGA data and analysis. In addition, trace DNA may not meet the DNA input requirements for analysis using WGAs. WGS may be utilised to circumvent the high DNA input requirements of WGA. However, WGS may require prior preparation steps such as whole genome amplification to increase the likelihood of obtaining useable data which also has validation requirements and considerations. In addition, these methods may require specialist bioinformatics to impute missing SNP data.

The expertise and technology required for WGA and WGS may not exist in operational forensic laboratories; consequently, external expertise and service provision may need to be employed to ensure reliable application of these methods for FGG. This may occur through a DTC operator that

provides a full service from sample analysis, provision of whole genome SNP data, upload of SNP data to a database and/or genealogical services.

Alternatively, forensic laboratories may choose a 'semi-outsourcing' approach by using a genome facility such as the Australian Genome Research Facility (AGRF) to generate the whole genome SNP data. Using this approach, the forensic or law enforcement organisation would then be responsible for downstream analysis and reporting. This approach would require laboratories to consider further requirements such as accreditation of the external service provider, selection of method (WGA or WGS), chain of custody, selection of one or more genetic genealogy databases available for law enforcement searches and employing the services of a genealogist.

There are no international standards or certification for genealogists or genetic genealogists, although standards exist at a national level (Board for Certification of Genealogists 2014). Therefore, it may be challenging to determine the level of expertise of the genealogist or genetic genealogist.

In addition to external expertise, a forensic or law enforcement organisation may consider developing in-house expertise to reduce the issues/risks associated with external service provision.

A toolkit approach

Application of FGG to criminal investigations and unidentified human remains should be considered in the context of generating investigative leads. Therefore, the primary analysis of DNA retrieved from biological evidence will be STR profiling (autosomal and Y) followed by forensic familial searching (if available/permitted). In the absence of leads, investigative DNA methods such as the prediction of biogeographical ancestry (BGA) and externally visible characteristics, e.g., eye and hair colour, may be considered prior to, or at the same time as, FGG. Given the high proportion of individuals of European descent in the public genealogy databases (Thomson et al. 2020), knowing the BGA of the DNA donor may assist in the decision-making process as to the likely success of applying FGG.

In light of the technical considerations of WGA and WGS, application of FGG to biological evidence requires consideration of the sample type and the quantity of DNA remaining for analysis. In cases where evidence may be depleted using FGG, the prior application of all available forensically validated DNA analysis methods should be considered as a priority before applying WGA or WGS.

Cost of analysing samples

The cost versus benefit of any new forensic technique is an important consideration to ensure best value is being achieved. Pricing for genetic genealogy has two components:

- the DNA analysis cost; and
- the genealogical analysis.

1. DNA analysis costs

The cost to fully analyse a human genome has dramatically decreased in recent years, from more than $USD100 million in 2001 to approximately $USD1,000 in 2019 (Wetterstrand 2019). This price drop has greatly exceeded Moore's law which suggests the doubling of computer processing power whilst halving cost over time.

In addition, the cost to agencies to generate the high-density SNP data required for genetic genealogy is reasonably low. At the time of publication, the cost of a single WGA chip is in the order of $USD3,500, enabling the analysis of around 20-80 samples depending on the chip and the number of controls used. Generating WGS data (30x) through genome sequencing services, such as at the AGRF and KCCG Sequencing Laboratory at the Garvan Institute of Medical Research, ranges from approximately $USD1,200 to $USD1,350 per sample. WGS through a DTC ranges from approximately $USD560 to $USD1,000 per sample. However, the genealogy costs are likely to comprise most of the expenditure in the application of genetic genealogy to investigations.

2. Genealogy costs

Genealogy costs are entirely dependent on the complexity of the search required to identify an individual. Providers generally offer an initial triage service, whereby a profile is entered onto a database and the match results reviewed to determine if the person is likely to be identified, based on the shared centiMorgan (cM) value to determine the degree of relatedness between two individuals (Bettinger

2020). As outlined by Thomson et al. (2020), matches indicating that the closest potential relative is likely to be third to fourth cousins (<50 shared cM) would require significantly more investigative time and resources and have a reduced chance of successful identification.

It is difficult to determine an 'average' time for genealogical searches and therefore provide costing estimates. The genealogical searching in the study described by Thomson et al. (2020) demonstrated that initial triage of genetic information from de-identified UK volunteers required 2-3 hours work, with the quickest case solved in 3 hours and complex cases between 50 and 300 hours. Changes to GEDmatch terms and conditions in May 2019 could further impact on these estimates (Thomson et al. 2020).

Australian genealogists are likely to charge approximately USD$35 per hour, meaning that a complex case could run to USD$10,000 in analysis costs. This cost may be considered low in terms of the years of investigative time and resources invested in cold case homicides and may be outweighed by the substantial benefit of identification of a previously unknown offender. However significant investment in genealogy analysis gives rise to considerations as to the adoption of these skills in-house, and verification of the proficiency of genealogists employed. These issues are discussed further in sections IX and X.

V. What can we compare against?

Forensic genetic genealogy has a high dependency on reference datasets in genealogical databases. In theory, the technique might have limited internal application if policing agencies developed their own database to generate long-range familial links between unidentified crime scene profiles. But operational success requires data (Ford 2018).

Access to online genealogical databases can be achieved in several ways:

1. On a commercial basis

Law enforcement may be able to compare their crime scene SNP data against data uploaded by members of the public simply by creating an account, paying any fees (if applicable) and complying with terms and conditions.

This is the approach used by law enforcement agencies internationally to date. Conditions could include:

- A requirement to declare that the profile relates to an investigation.
- Restrictions on which profiles may be searched against, based on each user's individual privacy settings or consent.
- Restrictions on use depending on the seriousness of the alleged offence (Aldhous 2019a)
- Geographical restrictions on use of services (e.g. limited to United States agencies only).

2. Under warrant

There have now been instances where investigative agencies have sought to obtain access to data under warrant (Hill & Murphy 2019; Kaiser 2019). Such an approach would negate any conditions ordinarily imposed. But it could also allow a vendor to challenge the legal validity of a warrant.

There is an interesting divergence between the size of some online genealogy companies and the potential privacy impact. GEDmatch, until its recent sale to Verogen, was a small organisation. The ability of small online providers that do not permit law enforcement database searches to challenge judicial warrants could be limited. This raises broader public policy considerations and may ultimately result in consumers avoiding such providers in favour of larger companies such as AncestryDNA and 23andMe.

3. Covertly

The third approach raises significant legal and ethical risk. Some enforcement agencies have special powers to undertake covert or clandestine investigations. Whether such an agency could upload a crime scene profile without declaring the nature of the sample is a complex question beyond the scope of this article.

Are databases 'local' enough?

A key consideration in operationalising this capability is an assessment of its viability in the local context. Online genealogy providers do not routinely

release sufficient information about their users and profiles to enable an exact estimate of utility.

It has been estimated that GEDmatch.com, the website for the GEDmatch Genesis platform, was accessed 774,000 times in January 2020.[1] Nearly 59.6% of this traffic originated in the United States, where it was estimated in 2013 that there were 40 million people with an interest in genealogy (Rak 2017; Tallbear 2013).

It was estimated that 9.4% of the web traffic in January originated in Australia. Given the United States has fourteen times the population of Australia, theoretically, GEDmatch could have been accessed by Australians over 2,000 times a day: per capita even more frequently than by Americans.

These numbers do fluctuate month to month, and each website visit would not equate to the upload of a new genetic profile. But this analysis does support the anecdotal contention that Australians do have a high, per capita, interest in genealogy research. And, as such, suggests that it is plausible that FGG may assist with the identification of Australian suspects or unidentified remains in some cases.

This use, however, would not be expected to be homogenous across the whole Australian population. Marketing campaigns run by major online genealogy providers such as Ancestry.com promote interest in genealogy by, at times, focusing more on British, American or descendance from early European settlers to Australia, as opposed to tracing family roots in other countries.

Further study of the use of genetic genealogy tools by population groups in Australia is therefore warranted to ascertain the effectiveness of the technique for different ancestral populations, as well as to help quantify Australia's use of these tools overall.

VI. Is it legal?

Forensic procedures legislation

There is a longstanding question of whether crime scene samples could be subject to types of DNA analysis beyond what was anticipated for upload into Australia's National Criminal Investigation DNA Database (NCIDD). This question was briefly considered in a 2003 report into the Commonwealth legislation, with the Review preferring not to recommend limitations on which genetic markers police could analyse (Commonwealth of Australia 2003).

Legislation in effect in Australia is primarily focused on the safeguards for provision of DNA samples by suspects and volunteers.[2] While those samples are tightly regulated, it is reasonably clear that DNA of unknown origin at a crime scene can legally be subjected to other forms of genetic analysis, either sequentially or likely even in parallel with STR-based DNA analysis for upload to NCIDD.

However, there is one important caveat. To avoid some of the legislative requirements around privacy in Australia, discussed below, the source must be genuinely unknown. There would be other legal considerations if an investigator already had a reasonable suspicion as to the suspect's identity but nonetheless used FGG processes to gather additional evidence about a suspect's relatives.

The legal position in Australia does not entirely align with United States case law concerning discarded DNA (Joh 2006). As such, similar considerations would apply to the collection of covert samples from suspects, so as to confirm a genealogical hypothesis, or from relatives to covertly extend a genealogical record (Australian Law Reform Commission 2003, pp. 1049-53).

Health records legislation

In Australia, most states and territories have specific legislation dealing with health records. The legislation frequently deals with privacy and access to those records by a patient. Given the potential for secondary use of genetic information for health research purposes (Stoeklé et al. 2016) it is now even more difficult to exclude these records from what is frequently a broad definition under legislation.

In New South Wales, the most populous state in Australia, for example, health information includes 'personal information that is genetic information

[1] Estimates generated using Similarweb.com

[2] *Crimes Act 1914* (Cth), *Crimes (Forensic Procedures) Act 2000* (ACT), *Crimes (Forensic Procedures) Act 2000* (NSW), *Police Administration Act 1978 (NT)*, *Police Powers and Responsibilities Act 2000* (Qld), *Criminal Law (Forensic Procedures) Act* 1998 (SA), *Forensic Procedures Act 2000* (Tas), *Crimes Act 1958* (Vic), *Criminal Investigation (Identifying People) Act 2002* (WA).

about a person in a form that is, or could be, predictive about the person's health at any time'.[3] Whole genome analysis will produce information about specific DNA markers which clearly fall within this definition.

While some health records legislation excludes operational police services, those exemptions may not extend to specialist laboratories or professional genealogists preparing profiles for upload to online databases.

In cases where health records legislation applies, additional requirements may apply, particularly relating to security and safeguarding. Whether the legislation applies to genetic information about an unidentified person, or only at the point where there is a reasonable presumption as to identity, could determine whether health records legislation would pose any practical difficulties in using FGG.

Regardless, any agency considering use of FGG should, as a matter of good practice, review internal processes to ensure a high degree of compliance in relation to how genetic information is used, accessed and stored.

Human Rights Acts

Three Australian jurisdictions – the Australian Capital Territory, Victoria and Queensland – have enacted specific human rights legislation.[4] In addition to creating a framework to assess the impact of new legislation on human rights, certain obligations are placed on government entities in their decision-making and, in some jurisdictions, on the exercising of a 'function of a public nature'.

The laws acknowledge that human rights can be limited in certain circumstances and seeks to provide a means of balancing competing societal and individual interests.

Human rights legislation includes protection against arbitrary interference with privacy and protection of reputation in a way which is broader than in other laws, such as the *Privacy Act 1988* (Cth) which regulate personal information.

A consideration of FGG in the context of human rights laws would turn to both the privacy rights of the donor of genetic material at a crime scene, as well as the privacy rights of individuals who have uploaded their profiles into an online database. Those considerations would likely turn to questions of proportionality and consent.

Privacy laws

Forensic genetic genealogy raises issues of proportionality, or the balancing of competing social interests (Moran 2018). As Denise Syndercombe Court from King's College notes, the 'intrusion into privacy [from forensic genealogy] is obvious to all' (Syndercombe Court 2018).

Privacy in Australia is regulated under the *Privacy Act 1988* (Cth) and equivalent state and territory laws. While there are some differences, the privacy principles established by these laws are relatively consistent. Privacy laws have broad application to both the public and private sectors in Australia.

The Australian Privacy Principles, under the *Privacy Act*, restrict the way in which 'personal information' can be collected, stored, used and disclosed.[5] Genetic data is defined as 'sensitive information', which is further restricted. While it is not entirely clear on the face of the legislation, guidance material confirms that sensitive information is intended to operate as a subset of personal information rather than as a distinct class of information (Office of the Australian Information Commissioner 2019). As such, sensitive information must also meet the definitional requirements of personal information – that it is about 'an identified individual, or an individual who is reasonably identifiable'.[6]

This distinction is important. The intention of using this technique on a crime scene sample is to establish identity. At the time of upload to an online database, identity is unknown and therefore the requirements in the *Privacy Act* do not apply.

A possible exception is if investigators failed to take other reasonable steps to identify the donor. For example, failing to consider that genetic material came from a victim, from whom an elimination reference sample could readily be obtained. Or moving straight to FGG without first exploring whether the donor's profile is already on NCIDD. Either of these cases might conceivably give rise to

[3] *Health Records and Information Privacy Act 2002* (NSW), s 6.
[4] *Human Rights Act 2004* (ACT), *Charter of Human Rights and Responsibilities Act 2006* (Vic), *Human Rights Act 2019* (Qld)
[5] *Privacy Act 1988* (Cth), sch 1.
[6] *Privacy Act 1988* (Cth), s 6(1).

an argument that the information was 'reasonably identifiable' and that a privacy breach has occurred.

As noted, the intention of this technique is to establish identity. If the technique is successful, and law enforcement forms a view as to the donor's identity, all genetic data falls within the definition of 'sensitive information'.

To limit the potential for a privacy breach, genetic data should generally be removed from online databases once a list of potential relatives is obtained. Within the laboratory, it would be best practice – as in the case of health records – to handle all genetic data as though it were sensitive information from the time of analysis, and to apply rigorous protocols around security, access and use.

Forensic genetic genealogy could involve third-party providers, such as professional genealogists. This is explored further below. In addition to holding genetic data from crime scene samples, a genealogist may hold information of individuals obtained officially (for example, through births, deaths and marriages checks) along with other publicly available records.

This data could be quite sensitive. Some genetic data could be predictive of health status. The combination of genetic match data and official records may reveal anomalies in family trees around parentage. Even before identity of the suspect is established, many of these records could fall within the definition of personal information. It would therefore be prudent to put in place appropriate non-disclosure agreements and consider contractual requirements around use, storage and return of case-related information in the custody of the third party.

VII. Is it secure?

The question of security of genetic genealogy came to the fore in July 2020 when GEDmatch was subject to 'a security breach orchestrated through a sophisticated attack on one of our servers via an existing user account' (GEDmatch 2020; Kennett 2020; Whittaker 2020). This follows the hacking of MyHeritage in 2017, exposing user credentials but no genetic data (Syndercombe Court 2018).

Commentators have noted the potential for online genetic databases and, in particular, GEDmatch to become a target of so-called false relative attacks (Ney, Ceze & Kohno 2020). This approach uses the very matching capabilities at the heart of genetic genealogy services against them, to expose personal information.

The potential for a data breach involving genetic data appears to have been realised, and creates very real concerns around public trust and confidence in genetic genealogy capabilities. The addition of law enforcement genetic profiles creates an operational security (opsec) risk which must also be factored into future planning around how police engage with these capabilities.

The rise of 'genetic informants'

While, as noted above, FGG has a heavy reliance on data, it can be easy to forget that each of these data points searched in a database is an individual user, a person who has their own view on genetic privacy and family relationships.

Williams referred to individuals drawn into investigations using familial DNA searches as 'genetic informants' (Haimes 2006; Williams & Johnson 2005). What are the potential implications for these individuals?

At one end of the spectrum, there is no impact. The individual whose online genetic profile helped triangulate the identity of a suspect, or whose family tree provided initial leads to a professional genealogist, may never become aware of that use.

They may become aware of what has occurred in several ways:

- They could be approached by investigators for more information about their family, or to help identify other family members who might agree to be tested.
- They may become aware because their details are led in evidence or otherwise included in court documents.

Being drawn into a police investigation can be stressful. Use in Australian investigations does not raise ethical issues concerning assisting in death penalty matters, which could arise in the United States. Notwithstanding, it could bring notoriety to a family member if shown to be related to an alleged serial killer or recidivist sex offender.

Citizens do not necessarily have a choice in their involvement in criminal investigations. A reluctant eyewitness can still be subpoenaed to give evidence for the prosecution. But the public policy

considerations can shift quickly if the personal safety of the 'genetic informant' is at risk.

Some affidavits in the United States have omitted names of distant relations but instead listed kit numbers. In most cases, this approach does not achieve full de-identification. The kit numbers may be publicly available on genealogy websites, allowing the police family tree to be reconstructed by the suspect, their lawyers or the media.

It is necessary to achieve a balance between the privacy of individuals whose online records have been used to identify suspects, and the rights of those suspects to disclosure of the prosecution case and to test or challenge techniques employed by police to identify them.

A judicial officer, prior to issuing a search warrant or similar instrument, needs to understand on what basis police have formed their hypothesis as to identity. Even if judicial officers were trained in these techniques, except where the affidavit reveals an obvious and significant deficiency in scientific or genealogical analysis, a judicial officer will be unlikely to be able to form an opinion as to whether the analysis is robust. Instead, judicial officers will likely need to satisfy themselves that a technique has been properly validated and that appropriate quality assurance safeguards exist.

Building privacy and security into analysis

A useful tool in assessing any new technology from a privacy perspective is the privacy impact assessment or PIA (Office of the Australian Information Commissioner 2014). This tool can be used to consider a capability through a privacy lens, and determine what potential problems could arise (Victorian Commissioner for Privacy and Data Protection 2017).

A PIA could be used to tease out issues with laboratory workflow, security and storage of genetic data, as well as long-term archiving or disposal. It can also consider questions as to whether genetic data is available to investigators or held only by laboratory personnel (Scudder et al. 2018).

Such a process can also assist with mapping out how the FGG process should work. SWGDAM's recently released overview of FGG/IGG recommended that profiles (whole genome data) be kept in online databases for the minimum time required to undertake genetic analysis and match comparison (Scientific Working Group on DNA Analysis Methods 2020). The PIA process can be used to consider the balance as to how long and how often a profile from an unknown crime scene sample remains in such a database, in the hope of obtaining additional matches with other relatives who may upload their profile after initial genealogy work is complete.

Such a decision is ultimately one of risk. If a case involved recidivism and an immediate threat to the community, and the source of the DNA sample gave a high likelihood that it came from the offender and nobody else, it may shift the balance in favour of more frequent uploading and comparison, or even support retaining the profile online for an extended period. In other cases, uploading the profile, receiving a list of potential relatives, and then deleting the profile would be low risk and most consistent with preserving individuals' privacy.

VIII. Maintaining Public Trust

In an article published in 2018, Mathias Wienroth from Newcastle University examined the governance of new or what he termed 'anticipatory technology' (Wienroth 2018). The author argued for three distinct approaches to self-governance, firstly aspirational regimes, secondly the setting of standards and thirdly application and training. While Wienroth's analysis extends beyond a purely regulatory approach, its central theme touches on the difficulties in establishing a governance arrangement around technology which is still evolving.

Forensic genomics cannot be entirely separated from wider developments in genomics. Some commentators have pointed to the fact that the groundwork for ethics in genomic screening dates as far back as 1975 (Shoenbill et al. 2014), While these underlying ethical principles for genetic analysis may be relatively constant, they must be viewed in the context of other developments, in data protection regulation and Big Data analysis.

Forensic familial searching, finding genetic relations of individuals of police DNA databases, has been in use in some countries for nearly 20 years. Forensic genetic genealogy raises a subset of the ethical issues associated with familial searching. Erica Haimes notes the ethical and policy challenges in prioritising identification of a suspect over personal identity of the suspect and their genetic relations.

The use of these techniques can prove or disprove genetic relationships, causing distress to not just a suspect but also members of their immediate, and possibly extended, families (Haimes 2006).

Police services have developed strategies to mitigate and manage ethical concerns around familial DNA. This has included the use of oversight boards, able to assess cases and provide advice (Maguire et al. 2014). While FGG utilises a different model for accessing genetic profiles, with consent, some of these strategies can also be used to oversee the use of the technique by law enforcement.

IX. Outsourcing

Procurement and governance

Operational Australian forensic laboratories do not currently possess instrumentation to conduct the WGA. Whilst a limited number of laboratories may possess equipment for WGS, WGS is not a capability which has been developed for FGG. In addition, no Australian forensic laboratory is accredited for this purpose. Therefore, the implementation of this technology may also entail a procurement and vendor assessment process. Issues such as chain of custody procedures, turnaround times, data security, reporting and logistics such as cost must be considered by any organisation seeking to outsource the analysis. Agencies may choose to divide the process amongst two or more service providers, as the DNA analysis and genealogical assessment may be conducted by separate vendors. The procurement assessment process should cover both aspects and identify acceptable output and success rates where possible.

Forensic genetic genealogy comprises functions that sit across traditional investigative and scientific roles, creating a challenge as to the long-term oversight. It is possible over time that both the DNA analysis and genealogy components may be adopted in-house, with the employment of genealogists and/or upskilling of staff in this methodology. The question arises as to whether the genealogical component should fall within the remit of the forensic or investigative arms of a policing agency. Some merit may be given to the alignment of genealogical searching within a traditional forensic quality framework, discussed further below.

International transfer of samples and data

Most commercial providers of FGG and related services are based in the United States. As such, there may be a need to ship forensic samples or to send data derived from those samples outside of Australia.

The *Privacy Act* provides a framework for cross-border disclosure of personal information. Generally, an entity subject to that Act is required to take reasonable steps to ensure that the recipient of the data will not breach the Australian Privacy Principles in the handling of that data.

There is an exception where reasonably necessary for ‘enforcement-related activities’, but only where the recipient has a similar enforcement-related function. Arguments will arise as to whether genetic data from an unidentified crime scene sample fits within the definition of personal information at the time of the cross-border disclosure.

Engaging a forensic genetic genealogist or other specialist from overseas to undertake genealogy work, where this involves providing information collected by an enforcement agency subject to the *Privacy Act* in Australia, would likely not fit within the exclusion for enforcement-related activities. Where the *Privacy Act* requirements apply, it would be necessary to consider whether the receiving party is contractually bound to meet, or is otherwise substantially subject under their own laws, to the requirements in the Australian Privacy Principles.

Future-proofing and vendor lock-in

Currently the access to databases for the purposes of genetic genealogical support of law enforcement investigations is limited to three service providers (Scientific Working Group on DNA Analysis Methods, 2020): GEDmatch, FamilyTreeDNA and now Othram’s DNASolves™ (Othram Inc. 2019). All of these providers are now private companies following the acquisition of GEDmatch by Verogen. This leaves the law enforcement use of genetic genealogy in the hands of commercial providers, in contrast to the strictly controlled, government-owned DNA data present in NCIDD. The future ownership and permitted use of this new genetic data may change depending on the success, viability and

business model of the commercial vendors. The risk of loss of access by law enforcement would be alleviated through the generation of equivalent whole genome SNP data in-house, with the added benefit of greater relevancy to Australian investigations given the database composition of local donors. However, this would require extensive time, resourcing, and legislative support to implement. A national conversation around the projected future control and use of these new genetic databases is warranted.

X. Establishing a Framework

Quality standards

Forensic science within Australia has a well-established path and process for quality standards and accreditation, through the National Association of Testing Authorities (NATA). A common refrain around the use of novel DNA testing such as genetic genealogy and phenotyping is that it is 'just intelligence', and that once a suspect has been identified through these methods, routine STR analysis will be utilised for confirmation and use in trial proceedings. Nonetheless, genetic genealogy evidence has already been tested in court hearings in the United States (Molteni 2019) lending weight to the requirement for a quality management framework to ensure its validity and ongoing application in criminal investigations.

Despite concerns that quality management has the potential to stifle forensic innovation and agility (Crispino & Roux 2016; Roux, Ribaux & Crispino 2018) the process can be adapted to be fit-for-purpose. The reactive field of military forensics, which traditionally relies on rapid intelligence rather than pristine evidence, has recognised the need for a quality framework (Wilson et al. 2018). It is therefore conceivable for standards to be developed around the provision of intelligence resulting from genetic genealogy, particularly to ensure consistency of information provided by different vendors and agencies. Whilst the information provided is considered intelligence, the consequences of incorrect information could be severe, including impacts on privacy, wrongful arrests, and lost time and resources in a misled investigation.

Proficiency Testing

A quality framework around FGG should encompass both components of the method. Validation of novel DNA analysis methods is well understood and should be applied here. Equally, the validity and proficiency of genetic genealogists is a critical component of the success of the method. Currently the education and accreditation of genealogists, and more specifically genetic genealogists, is largely unregulated and inconsistent.

Centre of Specialisation Model

The development of an FGG skillset within an organisation would require considerable resourcing, and likewise, procurement, validation and implementation of the instrumentation required to generate whole genome data is a substantial investment. Given that the number of applicable cases within Australia is likely to be small for reasons described previously, it would be inefficient for every jurisdiction to introduce this capability. Best value would see expertise in this field centralised to a small number of jurisdictions. Alternatively, given the dual nature of method (DNA analysis and genealogy), casework could be divided between two agencies along these lines to reduce the resourcing impost to develop expertise in both components.

XI. Conclusions

Forensic genetic genealogy presents many opportunities for law enforcement, particularly to assist with generating investigative leads that may solve cold cases or identify human remains. By carefully examining both technical and legal considerations in an Australian context, the authors have sought to provide a framework for Australian law enforcement considering operationalising this technique. By balancing privacy interests, considering appropriate oversight and understanding the potential limitations, including with database composition, the technique will potentially prove useful in Australia and withstand legal and public scrutiny.

XII. Checklist for Law Enforcement

- ❑ Policy development
- ❑ Privacy Impact Assessment
- ❑ Operational security assessment
- ❑ Legal advice
- ❑ Consideration of oversight/advisory committee
- ❑ Scientific and forensic validation
  - ❑ Sensitivity
  - ❑ Specificity
  - ❑ Mixtures
  - ❑ Sample type and analysis
- ❑ Vendor Assessment – DNA analysis
  - ❑ Chain of custody
  - ❑ Cost
  - ❑ Logistics (sample transport and recovery)
  - ❑ Type of analysis
  - ❑ Physical and data security
- ❑ Vendor Assessment - Genealogy
  - ❑ Applicability to different biogeographical ancestries
  - ❑ Data security
  - ❑ Confidentiality agreement
  - ❑ Access to appropriate records and databases)
- ❑ Pilot trial of casework
- ❑ Stakeholder education and awareness (e.g. homicide and cold case detectives, missing persons, sexual assault teams)
- ❑ Development of briefing notes/press releases

Acknowledgements

Dr Scudder would like to thank his doctoral supervisors at the University of Canberra, with whom he has previously published research. This article in part builds on this earlier research in an Australian context.

Disclosure Statement

No potential conflicts of interest are reported by the authors.

ORCID

Nathan Scudder http://orcid.org/0000-0002-6011-9092

Runa Daniel http://orcid.org/0000-0002-7262-6972

Jennifer Raymond http://orcid.org/0000-0003-0691-4105

Alison Sears

References

Aldhous, P 2019a, 'The Arrest Of A Teen On An Assault Charge Has Sparked New Privacy Fears About DNA Sleuthing', *Buzzfeed*, 14 May 2019, <https://www.buzzfeednews.com/article/peteraldhous/genetic-genealogy-parabon-gedmatch-assault>.

Aldhous, P 2019b, 'The Golden State Killer Case Has Spawned A New Forensic Science Industry', *Buzzfeed*, 15 Feb 2019, <https://www.buzzfeednews.com/amphtml/peteraldhous/genetic-genealogy-dna-business-parabon-bode>.

Angrist, M 2013, 'Genetic privacy needs a more nuanced approach', *Nature*, vol. 494, p. 7, <http://www.nature.com/news/genetic-privacy-needs-a-more-nuanced-approach-1.12363>.

Australian Law Reform Commission 2003, *Essentially Yours--The Protection of Human Genetic Information in Australia, Volume 1 and Volume 2. Report 96*, Canberra.

Bettinger, B 2020, *Version 4.0! March 2020 Update to the Shared cM Project!*, viewed 16 July 2020, <https://thegeneticgenealogist.com/2020/03/27/version-4-0-march-2020-update-to-the-shared-cm-project/>.

Board for Certification of Genealogists 2014, *Genealogy Standards, fiftieth anniversary edn*, Ancestry, Nashville TN.

Castillo, A, Tchekmedyian, A, Serna, J & Bermudez, E 2018, 'For victims of Golden State Killer, the horror never ended', *Los Angeles Times*.

Commonwealth of Australia 2003, *Report of Independent Review of Part 1D of the Crimes Act 1914 - Forensic Procedures*.

Crispino, F & Roux, C 2016, *Forensic-led regulation strategies*, Taylor & Francis Group Didcot (UK) Abingdon (UK).

Ford, M 2018, 'How the Supreme Court Could Rewrite the Rules for DNA Searches', *New Republic*, 30 April 2018, <https://newrepublic.com/article/148170/supreme-court-rewrite-rules-dna-searches>.

GEDmatch 2020, *Facebook announcement - 21 July 2020*, <https://www.facebook.com/officialGEDmatch/posts/>.

Guerrini, CJ, Robinson, JO, Petersen, D & McGuire, AL 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*, vol. 16, no. 10, p. e2006906, <https://doi.org/10.1371/journal.pbio.2006906>.

Haimes, E 2006, 'Social and ethical issues in the use of familial searching in forensic investigations: insights from family and kinship studies', *The Journal of Law, Medicine & Ethics*, vol. 34, no. 2, pp. 263-76.

Hill, K & Murphy, H 2019, ''Game-Changer' Warrant Let Detective Search Genetic Database', *New York Times*, <https://www.nytimes.com/2019/11/05/business/dna-database-search-warrant.html>.

International Society of Genetic Genealogy 2018, *Family Tree DNA*, viewed 19 Feb 2019 2019, <https://isogg.org/wiki/Family_Tree_DNA>.

Joh, EE 2006, 'Reclaiming Abandoned DNA: The Fourth Amendment and Genetic Privacy', *Nw. UL Rev.*, vol. 100, pp. 857-84.

Kaiser, J 2019, 'A judge said police can search the DNA of 1 million Americans without their consent. What's next?', *Science Magazine*, 7 Nov 2019, <https://www.sciencemag.org/news/2019/11/judge-said-police-can-search-dna-millions-americans-without-their-consent-what-s-next>.

Kennett, D 2020, 'Major privacy breach at GEDmatch', *Cruwys News*, 19 July 2020, <https://cruwys.blogspot.com/2020/07/major-privacy-breach-at-gedmatch.html>.

Maguire, CN, McCallum, LA, 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*, vol. 8, no. 1, pp. 1-9, <https://doi.org/10.1016/j.fsigen.2013.07.004>.

McLaughlin, EC 2020, 'Hearing details ghastly crimes of Golden State Killer as he pleads guilty to killings', *CNN*, 30 June 2020, <https://edition.cnn.com/2020/06/29/us/golden-state-killer-plea-expected/index.html>.

Molteni, M 2019, 'A Murder Trial Will Allow DNA Evidence From a Genealogy Site', *Wired*, 11 June 2019.

Moran, KS 2018, 'Damned by DNA—balancing personal privacy with public safety', *Forensic science international*, vol. 292, pp. e3-e4, <https://doi.org/10.1016/j.forsciint.2018.09.011>.

Murphy, E 2013, 'Legal and ethical issues in forensic DNA phenotyping', *New York University Public Law and Legal Theory Working Papers. Paper 415*, no. 13-46, <https://doi.org/10.2139/ssrn.2288204>.

Ney, P, Ceze, L & Kohno, T 2020, 'Genotype extraction and false relative attacks: security risks to third-party genetic genealogy services beyond identity inference', in *Network and Distributed System Security Symposium (NDSS)*.

Office of the Australian Information Commissioner 2014, *Guide to undertaking privacy impact assessments*, <https://www.oaic.gov.au/resources/agencies-and-organisations/guides/guide-to-undertaking-privacy-impact-assessments.pdf>.

Office of the Australian Information Commissioner 2019, *Australian Privacy Principles guidelines*.

Othram Inc. 2019, *DNA Solves*, <https://dnasolves.com/>.

Phillips, C 2018, 'The Golden State Killer investigation and the nascent field of forensic genealogy', *Forensic Science International: Genetics*, vol. 36, pp. 186-8, <https://doi.org/10.1016/j.fsigen.2018.07.010>.

Rak, J 2017, 'Radical connections: genealogy, small lives, big data', *a/b: Auto/Biography Studies*, vol. 32, no. 3, pp. 479-97.

Regalado, A 2018, '2017 was the year consumer DNA testing blew up', *Technology Review*, 12 Feb 2018, <https://www.technologyreview.com/s/610233/2017-was-the-year-consumer-dna-testing-blew-up/>.

Roux, C, Ribaux, O & Crispino, F 2018, 'Forensic science 2020–the end of the crossroads?', *Australian Journal of Forensic Sciences*, vol. 50, no. 6, pp. 607-18.

Scientific Working Group on DNA Analysis Methods 2020, *Overview of Investigative Genetic Genealogy*.

Scudder, N, McNevin, D, Kelty, SF, Walsh, SJ & Robertson, J 2018, 'Forensic DNA phenotyping: Developing a model Privacy Impact Assessment', *Forensic Science International: Genetics*, vol. 34, pp. 222-30, <https://doi.org/10.1016/j.fsigen.2018.03.005>.

Selk, A 2018, 'The ingenious and 'dystopian' DNA technique police used to hunt the 'Golden State Killer' suspect', *Washington Post*, 28 April 2018, viewed 12 May 2018, <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, EA 2014, 'Genetic data and electronic health records: a discussion of ethical, logistical and technological considerations', *Journal of the American Medical Informatics Association*, vol. 21, no. 1, pp. 171-80, <https://doi.org/10.1136/amiajnl-2013-001694>.

Skinner, D 2018, 'Forensic genetics and the prediction of race: What is the problem?', *BioSocieties*, pp. 1-21, <https://doi.org/10.1057/s41292-018-0141-0>.

Stoeklé, H-C, Mamzer-Bruneel, M-F, Vogt, G & Hervé, C 2016, '23andMe: a new two-sided data-banking market model', *BMC Medical Ethics*, vol. 17, no. 19, <https://doi.org/10.1186/s12910-016-0101-9>.

Syndercombe Court, D 2018, 'Forensic genealogy: Some serious concerns', *Forensic Science International: Genetics*, vol. 36, pp. 203-4, <https://doi.org/10.1016/j.fsigen.2018.07.011>.

Tallbear, K 2013, 'Native American DNA', *Minneapolis: U of Minnesota P*.

Thomson, J, Clayton, T, Cleary, J, Gleeson, M, Kennett, D, Leonard, M & Rutherford, D 2020, 'An empirical investigation into the effectiveness of genetic genealogy to identify individuals in the UK', *Forensic Science International: Genetics*, p. 102263.

Tillmar, A, Sjölund, P, Lundqvist, B, Klippmark, T, Älgenäs, C & Green, H 2020, 'Whole-genome sequencing of human remains to enable genealogy DNA database searches–A case report', *Forensic Science International: Genetics*, vol. 46, p. 102233.

Victorian Commissioner for Privacy and Data Protection 2017, *Privacy Impact Assessment Template*, <https://www.cpdp.vic.gov.au/menu-resources/resources-privacy/resources-privacy-checklists-and-tools>.

National Institutes of Health 2019, *The cost of sequencing a human genome*, by Wetterstrand, KA, <https://www.genome.gov/about-genomics/fact-sheets/Sequencing-Human-Genome-cost>.

Whittaker, Z 2020, 'Gedmatch investigating after users' DNA profile data made available to police', *Techcrunch*, 20 July 2020, <https://techcrunch.com/2020/07/19/gedmatch-investigating-dna-profile-law-enforcement/>.

Wienroth, M 2018, 'Governing anticipatory technology practices. Forensic DNA phenotyping and the forensic genetics community in Europe', *New Genetics and Society*, vol. 37, no. 2, pp. 137-52, <https://doi.org/10.1080/14636778.2018.1469975>.

Williams, R & Johnson, P 2005, 'Inclusiveness, effectiveness and intrusiveness: issues in the developing uses of DNA profiling in support of criminal investigations', *The Journal of Law, Medicine & Ethics*, vol. 33, no. 3, pp. 545-58.

Wilson, LE, Gahan, ME, Robertson, J & Lennard, C 2018, 'Fit for purpose quality management system for military forensic exploitation', *Forensic science international*.

Wright, D & De Hert, P 2012, *Privacy Impact Assessment*, Springer, Dordrecht.