Title: Forensic DNA phenotyping: Developing a model
Privacy Impact Assessment

Authors: Nathan Scudder, Dennis McNevin, Sally F. Kelty,
Simon J. Walsh, James Robertson

PII: S1872-4973(18)30141-8
DOI: https://doi.org/10.1016/j.fsigen.2018.03.005
Reference: FSIGEN 1867

To appear in: *Forensic Science International: Genetics*

Received date: 8-6-2017
Revised date: 22-2-2018
Accepted date: 6-3-2018

Forensic DNA phenotyping: Developing a model Privacy Impact Assessment

Nathan Scudder [a, c] *

Dennis McNevin [a]

Sally F. Kelty [b]

Simon J. Walsh [c]

James Robertson [a]

[a] *National Centre for Forensic Studies, Faculty of Science and Technology, University of Canberra, ACT 2617, Australia*

[b] *Centre for Applied Psychology, Faculty of Health, University of Canberra, ACT 2617, Australia*

[c] *Australian Federal Police, GPO Box 401, Canberra ACT 2601, Australia*

*Corresponding author: Nathan.Scudder@canberra.edu.au

Correspondence Address: National Centre for Forensic Studies, University of Canberra ACT 2617

Declarations of Interests: None

Highlights

- A Privacy Impact Assessment is a useful policy response to step changes in technology, supporting a privacy by design approach.
- Applying this methodology to forensic DNA phenotyping requires careful consideration of issues of consent, public confidence and the right not to know.

- A careful and thorough application of this methodology will strengthen laboratory implementation and minimise privacy intrusion.

Forensic scientists around the world are adopting new technology platforms capable of efficiently analysing a larger proportion of the human genome. Undertaking this analysis could provide significant operational benefits, particularly in giving investigators more information about the donor of genetic material, a particularly useful investigative lead. Such information could include predicting externally visible characteristics such as eye and hair colour, as well as biogeographical ancestry. This article looks at the adoption of this new technology from a privacy perspective, using this to inform and critique the application of a Privacy Impact Assessment to this emerging technology. Noting the benefits and limitations, the article develops a number of themes that would influence a model Privacy Impact Assessment as a contextual framework for forensic laboratories and law enforcement agencies considering implementing forensic DNA phenotyping for operational use.

Keywords: forensic, DNA phenotyping, biogeographical ancestry, externally visible characteristics, privacy, privacy impact assessment

1. The privacy challenges of genetic information

With the advent of massively parallel sequencing (MPS), forensic laboratories can undertake more cost-effective analysis of informative parts of the human genome.[1] This is a major shift from current forensic DNA practice, which is focused on repetitive elements of DNA sufficient to estimate probabilities that different DNA samples have the same origin. In the context of forensic science, these capabilities would allow analysis of genetic material deposited by an individual at a crime scene and for probabilistic predictions to be made to inform law enforcement about possible attributes of the donor. This capability, referred to as *forensic DNA phenotyping*, is currently targeted at predicting biogeographical ancestry (BGA) and externally visible characteristics (EVCs), such as hair and eye colour [2,3].

Forensic DNA phenotyping could also soon be used to predict other donor traits, including male pattern baldness, biological age and fingerprints [4-6]. It could even be used to predict predisposition to certain diseases, for which they may be seeking medical treatment, or behavioural traits [7].

There are strong public policy grounds for using forensic science to assist law enforcement agencies to apprehend offenders. However, in addition to ensuring that new technology is scientifically sound, peer reviewed and quantifiable in terms of its error rate [8,9], the public interest must also be balanced against detriment to personal privacy. Genetic privacy is an important ethical issue, andaw enforcement agencies must also establish and maintain public confidence and trust in capabilities likely to intrude on the privacy rights of individuals or groups, or risk public criticism [10,11].

Historically, forensic DNA profiling has exploited medical testing capabilities, with the associated public policy discussions touching on medico-ethical issues of bodily integrity and privacy [12,13]. With forensic DNA phenotyping, it is necessary to further contextualise the operational capability within a more widely, ethically-informed privacy framework.

In broad terms, any exploitation of the human genome – particularly without the informed consent of the donor – presents a number of inherent risks. In the case of forensic DNA phenotyping, these can be categorised as:

1. Harm arising from the use or disclosure of predictive information generated to assist law enforcement. An example could be an individual becoming aware that their predicted BGA does not match their beliefs, based on their own cultural or familial self-identity.[14]
2. Ancillary information which could be derived from the predictive information described above. An example could be release of genetic marker information which, at a future time, is found to predict an individual’s health status.

Writing about whole genome sequencing, the Presidential Commission for the Study of Bioethical Issues [15] noted that:

> Strong baseline privacy protections require a spectrum of policies starting with data handling through the protection of persons from future disadvantage and discrimination…

The same report noted the balance between 'public beneficence' and 'responsible stewardship' by government, the need to respect the 'dignity and privacy' of individuals and avoid 'social stigma', concluding that whole genome sequencing 'substantially raises the stakes of medical information' [15].

Gostin [16] argues that '[g]enomic data are qualitatively different from other health data because they are inherently linked to one person', going on to describe them as 'unchanging and unchangeable' and a breach of genomic privacy as potentially giving rise to 'economic harms, such as loss of employment, insurance, or housing'. Similar themes can be found in the report of the Australian Law Reform Commission, which noted that genetics deals with 'possibilities rather than certainties', that 'genetic information has a familial dimension' and that there are cultural sensitivities around kinship and identity [17]. In addition, groups can be stigmatised in a genetic context and there are [inherent problems of confidentiality and informed consent when dealing with genomic information in the context of informational privacy [18,19].

Widespread use of DNA databases began in the late-1990s with agreement on specific genetic loci for DNA profiling [20]. These loci were, at the time, thought not to be informative, or to consist of 'junk DNA', and are used exclusively for identity, not phenotype prediction (with the exception of the amelogenin sex-determining locus) [21]. The notion that these loci are not informative for phenotypes has been challenged both in terms of functional genomics and in the context of the broader privacy debate over law enforcement use of DNA [12,22,23].

In addition, Curtis [24] described research undertaken in 2009 which showed most respondents to a survey held concerns about secondary use of their genetic information, if provided to law enforcement. Novas and Rose [25] explain that an individual's genetic identity must be placed in familial and societal contexts and a great deal of value from genetic information comes from its family links and an ability to compare genes between individuals [26]. The larger this body of genetic knowledge, the more accurate our probabilistic predictions and observations become and our ability to identify new phenotype-informative markers [20]. These familial and community links further complicate our assessment of forensic genomics from a privacy and ethical perspective, requiring us to approach the question of an adequate privacy framework holistically and with reference to broader community concerns. For example, the potential for phenotype-based intelligence to lead to prejudice against communities is a concern that is further exacerbated by the potential for particular populations to be overrepresented in BGA predictive modelling [27,28].

Machado and Silva [29] compared ethical issues for medical biobanks and forensic DNA profiling, identifying commonalities between them that underpin their success. Key to this was transparency and accountability, the right to be informed and to provide consent where feasible. Forensic DNA phenotyping will, in many ways, further blur this distinction and require forensic scientists to learn from and adapt approaches used in medical research and diagnostic fields. Privacy considerations align with community concerns about the potential for DNA database capabilities to be used as a form of genetic surveillance [30].

While acknowledging that there has been a privacy debate around current forensic DNA practice, we do not seek to re-explore those issues except where they are particularly agitated by the adoption of forensic DNA phenotyping. The use of a small selection of DNA markers for forensic identity testing has delivered a robust, generally privacy-compliant system for more than twenty years. We instead seek to tease out the differences between this approach (the baseline) and changes arising from forensic DNA phenotyping.

2. Privacy Impact Assessment as a policy response

When a generational change occurs in forensic technology, a prudent service provider will seek legal advice as to any legislative impediments. Such advice would likely identify the provider's need to comply with relevant privacy legislation.

Purely considering statutory obligations would not necessarily ensure engagement with broader issues of privacy norms: concepts that may make a course of action lawful yet socially unacceptable [31,32]. In particular, Tene and Polonetsky noted the need to avoid privacy by 'regulatory fiat' and to embrace a 'a nuanced and sophisticated path' [31]. Clearly, a broader approach to privacy is required, with one example being the use of privacy impact assessments (PIA). A PIA is a technique used to examine new capabilities or projects through a privacy lens [33]. The requirement for, approach and content of a PIA varies between countries. This discussion, therefore, is limited to the broader concept with a view to informing a model PIA – a so-called 'straw' policy - that can be adapted to specific jurisdictional requirements. Such an approach can assist in identifying privacy risks and placing technology within a broader societal and technological context [34].

3. Approaches to a PIA

The use of a PIA is generally aligned to legal obligations on business and agencies to implement policies and procedures to ensure compliance with privacy laws. For example, the Australian Privacy Principles require an entity to 'implement practices, procedures and systems' to ensure privacy compliance. [35] The US *Privacy Act* requires agencies to publish details of systems of records in the Federal Register.[36] This approach, which encourages proactive consideration of potential privacy risks, is known as 'privacy by design' [37], and works alongside agencies' existing risk management frameworks [38].

A PIA is intended to minimise any adverse impact on personal privacy or risks around the handling of personal data 'while allowing the aims of the project to be met whenever possible', and considers both positive and negative privacy impacts [38,39].

The Office of the Australian Information Commissioner (OAIC) [38] describes a 'threshold assessment' for a PIA, explaining that the process may be unnecessary where:

> the project does not propose any changes to existing information handling practices, if the privacy implications of these practices have been assessed previously and controls are current and working well.

Embarking on a PIA requires a careful assessment of its scope, however. In the context of forensic DNA phenotyping, the PIA must consider the broader opportunities MPS presents to forensic science. It is arguably inappropriate to employ a PIA threshold assessment that merely addresses a technology upgrade and does not grapple with the broader issues of maintaining public trust in forensic DNA, the shift towards forensic genomics and concepts of 'big data' [40]. Equally important is to determine whether any PIA process should revisit broader questions around law enforcement use of DNA, or accept - as a baseline – some level of privacy intrusion as a necessary part of criminal justice processes.

The process of conducting a PIA requires a detailed mapping of how personal information will be collected, used and managed [38]. The United Kingdom Information Commissioner's Office [39] prescribes wide consultation as an integral part of a PIA, alongside the required analysis of local privacy laws and practice. A broad range of risk and mitigation factors should be considered and the OAIC [38] recommends areas of focus to include:

- Necessity (of collection);
- Proportionality (to broader realised benefits);
- Transparency and accountability;
- Implementation of privacy protections (such as staff training);
- Flexibility (considering differing community views and privacy expectations);
- Privacy by design; and
- Privacy enhancing technologies.

A further step in planning a PIA-based approach to forensic DNA phenotyping is to consider how it factors into a proposal to implement MPS capabilities. Given the timeframes usually involved in purchasing and validating a new MPS platform, there is argument for two separate PIA processes or, at the very least, a refresh of the PIA prior to operational use. A PIA conducted too early and with too narrow a focus could have limited relevance by the time the technology implementation is complete. A PIA has been described as an 'iterative process' [37] and there is considerable advantage in such a process running in parallel, and closely aligned to, an agency's implementation of forensic DNA phenotyping.

An important consideration in undertaking a PIA is who should lead its development. A PIA process can be much improved by engaging the right mix of personnel with scientific, legal, regulatory, governance, ethical and privacy expertise. Inclusion of these skillsets, whether as authors of a PIA or engaged stakeholders, would allow the final product to thoroughly consider proposed uses of forensic DNA phenotyping in the applicable legal and societal context. The exchange of ideas concerning privacy and forensic DNA, and the sharing of PIAs themselves, will assist.

3.1. Legal framework

Privacy laws vary between jurisdictions and, to provide a reasonably adaptive discussion of the privacy challenges of forensic DNA phenotyping, the model PIA approach in this paper incorporates some consideration of the privacy frameworks in Australia, the European Union, Japan and the United States.

In the subsequent sections, we use the hypothetical introduction of a forensic DNA phenotyping capability to work through the principle aspects of a PIA.

3.2. Assumptions and Conceptual Foundation

It is important to note that the PIA imagines privacy as an innate human right, but one which must be viewed through a prism of legal frameworks [41]. As such, a PIA must objectively balance the right to privacy with competing societal interests. Clarke [42] notes that

> [T]he PIA process is motivated by the need for public trust, and is framed in terms of risk management…The evolution of PIAs needs to be seen within the context of larger trends in advanced industrial societies to manage risk and to impose the burden of proof for the harmlessness of a new technology, process, service or product on its promoters.

As a tool that seeks both to reassure the community and manage inherent risk, particularly in technology projects, an important question is the level of specificity for a PIA [34]. Should it take a holistic or an incremental approach to privacy? For a step-change in technology, should it seek to make bold predictions about potential use or misuse? A PIA should be forward-thinking, anticipate and comment on foreseeable technological or procedural changes and make recommendations that may pre-empt privacy concerns before they arise. Wright [43] outlined sixteen steps in an optimised PIA approach. Important to this is identifying the information flows and areas of privacy impact. In doing so, the authors of a PIA must set parameters around how they see the technology may operate. It is reasonable to anticipate or, indeed, recommend a further PIA for future substantial shifts in the technology base, as well as to seek to embed privacy awareness as a cultural imperative [43].

For forensic DNA phenotyping, the MPS platform itself has wide application depending on the chemistry employed. It is therefore possible for an MPS instrument to be configured between runs to provide entirely different genetic information, to switch from identity markers to phenotype markers, or to run both simultaneously [44].

Table 1 outlines some current and future genomic applications to forensics, and acts as an example of how the scope of a PIA may be constrained by identifying different use cases which may range from a baseline or current state to future speculated potential.

4. Technology Description and Information Flows

Forensic DNA phenotyping commences with the acquisition of *genetic material*. This material is generally gathered by swabbing or sampling material at a crime scene. The genetic *sample* is then analysed, creating *genetic information* in analogue or digital form. This raw genetic information is then interpreted and distilled into information for comparison purposes (in the form of a genotype) or for forensic DNA phenotyping, interpreted and represented graphically, diagrammatically or as predictive statements.

At times, the sample will contain genetic information from more than one person. This could include an intimate swab from a victim, which includes genetic material from both the victim and the suspect. In the context of forensic DNA phenotyping, at least up until the point when a victim’s genetic information can be compartmentalised, they would maintain a personal privacy interest in genetic information, the exploitation of which is not intended to identify attributes about them.

4.1. How is the genetic information collected?

Informed consent in relation to access to an individual's genetic material can raise ethical, scientific and technical issues [16,40]. Genetic samples can come into law enforcement possession in several ways (see Table 2), including voluntarily but also from abandoned DNA or from material coercively obtained from suspects and offenders in accordance with legislative powers.

Donors consent to the collection of their DNA only in limited circumstances. In doing so, there is often a power imbalance between the state and the donor [12]. Questions concerning consent have seldom become key issues in criminal proceedings using current identity-focused DNA technology, as the use is generally quite limited to identification purposes. Given that Forensic DNA Phenotyping requires more intrusive analysis, if a court were to consider consent from a privacy perspective, it may well be construed narrowly. This is similar to diagnostic medicine, where any ambiguity in patient consent would likely be interpreted as extending only to analysis reasonably necessary to provide effective diagnosis and treatment [23].

4.2. What genetic information is being held?

MPS technology provides highly granular genetic information, at the base pair level, at selected targeted regions of an individual's DNA. The technique can analyse any fraction of the human genome and generates a significant amount of data for each sample processed, particularly compared to current forensic DNA analysis employing capillary electrophoresis. In Australia, genetic information is 'sensitive information', and is further restricted as to secondary use [45]. The European Union defines information about an individual's health status as 'data concerning health' [46]. The United States makes no such distinction, although the requirement on agencies to 'maintain in its records only such information about an individual as is relevant and necessary to accomplish a purpose of the agency' will be relevant in a genetic privacy context [36]. A similar requirement applies in Japan, stipulating that agencies 'may retain Personal Information only when the retention is necessary for performing the affairs under its jurisdiction provided by laws and regulations' [47]. In the selected sample jurisdictions, the information being collected would constitute information that attracts the protection of relevant privacy laws.

4.3. Analysis and access to genetic information

While a laboratory introducing MPS would need to map their internal information flows to determine an optimal balance between privacy concerns and efficient access to laboratory data, the actual laboratory workflows associated with forensic DNA phenotyping are unlikely to be significantly different to existing DNA analysis for identification purposes. The quantum of data would be greater and, from a privacy perspective, its potential to reveal personal or sensitive information about the donor is increased.

Privacy requirements concerning data security and disclosure therefore need to be more robust when dealing with forensic DNA phenotyping.

The resulting probabilistic phenotype predictions, in the form of a written, graphical or illustrative representation of BGA and/or EVC, also have privacy implications. The provision of information on likely appearance would appear consistent with similar collection of eyewitness evidence. However, where either BGA or EVC prediction, or a combination, allows investigators to infer possible community or cultural ties, and this intelligence is used as the basis to undertake further searches, these additional searches could raise Fourth Amendment constitutional issues in the United States [12,48]. This would particularly apply where the genetic information identifies a particular community group, or excludes other groups, but does not sufficiently individualise suspects within that population. The use of genetic markers to identify health information, such as that a donor may suffer a medical condition requiring a specific prescribed medication, will be discussed later.

4.4. Other information sharing

The specialised nature of forensic DNA phenotyping will likely encourage laboratories to offer testing to other organisations through a service-based approach. Whether as part of broader inter-laboratory cooperation or on a fee-for-service basis, it is therefore possible that samples or genetic information may be transferred between government laboratories, or between government and private sector laboratories.

While an outsourced forensic model does not, of itself, raise inherent privacy concerns if properly managed, there is the potential for genetic information to be diverted into research programs or be transferred insecurely between organisations. Medical and diagnostic testing has had to grapple with similar issues and have established relevant guidelines that could assist forensic laboratories considering an outsourced model [49].

Transnational movement of samples and genetic information can also be anticipated. There are requirements in most privacy laws with respect to international transfer of personal information or personal data [45,46]. The European Council adopted a resolution in 2001 concerning transnational DNA results which, in part, encourages only the transfer of 'chromosome zones containing no genetic expression' (i.e. identity markers) [50].

4.5. Withdrawal of consent

As previously noted, DNA phenotyping of samples of known origin adds no forensic value. However, at times, issues of withdrawal of consent can apply, particularly when samples are later identified through comparison with reference DNA.

Withdrawal of consent by the donor of a genetic sample would likely raise more privacy concerns and sensitivities in the context of the potential use of forensic DNA phenotyping than traditional DNA analysis. The Office of the Australian Information Commissioner [51] notes that voluntary consent can be withdrawn at any time. Japanese privacy law allows for suspension of use in certain circumstances [47]. While an individual volunteering genetic information may withdraw their consent at any time, the practical effect of that withdrawal may vary depending on other legislative, judicial or policy considerations.

There is potential for someone who volunteered a reference DNA sample, perhaps a victim or witness, to later seek to withdraw consent on the basis of concern that the reference sample could be used to re-identify their genetic information collected at a crime scene and possibly already subjected to MPS analysis. In considering the privacy requirements, laboratories must be mindful of public trust. While destruction of a voluntary sample may not always be feasible, or legally required, laboratories will need to be sensitive to this aspect in the context of broader public debate about genomics and genetic privacy.

4.6. Access to personal information

Privacy laws frequently provide a right of access to an individual's own personal information or records. [36,45-47] In the case of forensic DNA phenotyping, a right of access would be exercised when a donor seeks to obtain a copy of any genetic information or resulting predictive phenotype reports relating to them as donor of that genetic material.

Responding to a specific request from an individual for a copy of their genetic information would not ordinarily raise privacy concerns. The United States Armed Forces DNA Identification Laboratory, by way of example, provides a simple one-page form for family members to request a copy of their mitochondrial DNA sequence after analysis.[52] Laboratories have a corresponding duty to maintain security of personal information, however. Workflow analysis, in a PIA context, must balance making information accessible and securing sensitive information against unauthorised access or disclosure.

Donors could also become aware of genetic information in other ways in the course of the use of probabilistic phenotype prediction. This may result in the individual becoming aware of health information. If a forensic laboratory undertook phenotype prediction specifically looking for health information, so as to undertake particular investigative lines of enquiry, it would be reasonable to assume the donor could ultimately become aware of that information.

For example, consider an individual who provided a voluntary DNA sample (with informed consent) only to later learn police were looking for an individual matching their physical appearance who was likely taking a certain medication for a debilitating genetic condition? It may be that this represents the first time the individual becomes aware that they carry the gene sequence predisposing them to that condition. Does the ‘right to know’ or ‘right not to know’ for this person of interest differ if they were not involved in any crime, but had merely deposited genetic material at a certain location, where a crime later occurred; of if they were a suspect later exonerated; or if they were a suspect ultimately convicted? Should laboratories pro-actively notify individuals of health information, if they become aware of it as a result of Forensic DNA Phenotyping?

The ‘right not to know’ is ascribed in the Universal Declaration on the Human Genome and Human Rights [50]. In considering the privacy implications of the use of probabilistic phenotype prediction, the likelihood of an individual becoming aware of health information through a criminal investigative process must be considered and balanced against the law enforcement benefits of exploiting such technology.

A more vexed issue is that it is possible that predictive markers, or any markers for that matter, thought to be informative only for BGA or EVCs will later be found to be health informative. An individual who has either specifically requested, or otherwise obtained, genetic information arising out of forensic analysis, could inadvertently become aware of health informative associations years later.

Lunshof et al [18] challenges the absolutes with respect to consent or privacy in a genetic context, instead highlighting an approach based on risk. Informing individuals requesting copies of their genetic information that the document they are seeking may contain information that is health predictive, or will one day be health predictive, can assist in a more sophisticated access regime for personal information. Providing information as to the risk or likelihood of that occurring would further assist the individual to make their own decision as to whether they wished to exercise what is often a legal right of access to their own personal information.

Concern may also be raised about any obligation on a laboratory to self-initiate disclosure of personal information to an identified donor, should the laboratory form the view that it does contain information relevant to that individual’s health status or treatment [50,53]. A scheme that foreshadowed proactive release would need to consider the ‘right not to know’ and potentially any need for genetic counselling to ensure informed consent, potentially years before the information was known to be health informative.[14] An alternative approach is to appropriately explain to individuals, at the point of reference sample collection, that should the reference sample link to a crime scene sample which – now or in the future – is shown to contain predictive information, there would be no proactive disclosure of that information to the donor.

4.7. Anonymity/Pseudonymity

Privacy laws in Australia include a specific requirement for entities, where practical, to engage with individuals who wish to remain anonymous or to adopt a fictitious name or title. [45].

In the context of crime scene samples of unknown origin, the very aim of forensic analysis is to identify and attribute identity to the sample. A forensic laboratory's responsibilities within the criminal justice system would generally prevent it from knowingly allowing other individuals, such as victims of crime, to provide DNA anonymously or under a pseudonym, such that their real identity was not known.

5. Privacy Risks and Mitigation Strategies

One approach, yielding the highest privacy safeguards, is to restrict initial forensic DNA analysis to identity markers compatible with CODIS or similar DNA databases, effectively entrenching current DNA technology as a 'baseline' test. Predictive phenotyping would only be permitted:

1. When analysis of DNA identity markers, and subsequent comparison against DNA databases, yields no results; or
2. In a small number of cases where there is insufficient genetic material to ensure that a second sample could be derived and run. In such cases, identity marker analysis and

forensic DNA phenotyping could run concurrently (and, in many cases, on the same technology platform).

Given the high percentage of cases where the victim and offender are known to each other, and where phenotyping would present no obvious advantage over traditional DNA analysis, such an approach offers a high level of inherent privacy by design. [38,54]

In considering the privacy aspects of this approach, an evaluation of privacy concerns must also consider whether de-identified genetic information may one day be re-identified. This tends to put genetic information into an unusual, although not unique, category when dealing with de-identified personal information. Whether by investigative action, or other means such as data aggregation and re-identification, it may be possible to one day attribute unknown genetic information obtained from a crime scene to an individual [55].

A privacy compliant approach must therefore build in an assumption of re-identification, so that genetic information of unknown origin is accorded the same level of security and used solely for its primary purpose of collection: establishing the donor's identity as an investigative lead.

5.1. New genetic markers

As previously discussed, a PIA would be incomplete if it did not consider the strong likelihood that technology – and particularly MPS assays – will evolve over time. Individual laboratories may have little control over additional markers added by the instrument manufacturers in future.

From a privacy perspective, it is important to consider what processes should apply when new testing capabilities are made available, and how to ensure that privacy intrusion is considered against any perceived investigative benefit. There may be a temptation to provide as much information about an unknown crime scene sample to investigators as possible. However, this could result in undermining public confidence in the capability, raising concern amongst individuals and perhaps resulting in a decrease in their willingness to cooperate with police investigations [12]. This may well be the case if, as could reasonably occur, it is later found that the crime scene sample was deposited not by a suspect but by an innocent passer-by, then subject to intensive and intrusive genetic analysis. In the United States, analysis of discarded genetic material does not generally raise Fourth Amendment concerns [12]. However, the absence of appropriate safeguards and the likelihood of re-identification of samples could give rise to argument that it amounts, in some cases, to an unreasonable search. There are very real policy benefits, therefore, in ensuring that privacy safeguards are maintained [20].

Privacy considerations also extend to the processes around reporting information derived from forensic DNA phenotyping to police investigators. This issue becomes increasingly sensitive if that reporting extends beyond predicting BGA and EVC. It is possible to imagine a 'life and death' situation where investigators require full exploitation of genomic information from a crime scene sample to provide a time-critical, comprehensive intelligence briefing. In these instances, the best balance could be struck by forensic scientists working closely with investigators to ensure that genetic information is understood in its predictive context and as part of the totality of the evidence. Police and forensic scientists must be critically aware of the potential for forensic DNA phenotyping, like other forms of forensic evidence, to mislead if not considered in context [56].

A viable safeguard against privacy intrusion and scope creep, but equally valuable in guarding against the inadvertent misdirection of police resources away from the real offender, would be establishing a group of senior police officers and forensic scientists, similar to an ethics board, to authorise release of less reliable or privacy intrusive predictive information on a case by case basis. Such an approach would ensure that the likely investigative benefit is clearly weighed against the privacy implications, particularly as the technology matures. Such a group could ultimately make recommendations as to appropriate guidelines or standards, together with stakeholders with interest in law, governance, regulation, privacy and ethics.

While it may not be feasible to include external experts routinely if decisions are being made on specific ongoing investigations, relevant agencies should involve such individuals as well as drawing from the medical science community, when formulating more enduring guidelines for the use of predictive phenotyping for law enforcement.

5.2. Wider use of genetic markers

A strictly limited approach to the use of genetic markers has several shortcomings. As Murphy [54] explains, most crimes occur between people who know each other. A sizeable proportion of DNA samples processed each year by crime laboratories are reference buccal swab samples from known individuals. As such, the limited approach assumes laboratories are willing to ignore potential economies of scale from using phenotyping for a broader range of samples, ultimately reducing the processing time and cost of DNA analysis, removing any requirement to re-analyse samples, and thereby fulfilling other public policy benefits.

Privacy concerns could arise with a broader adoption of forensic DNA phenotyping, including the potential for the MPS technology base to be used for both crime scene and reference samples. Table 3 outlines the various source of genetic samples, and how they may interact with various legislative privacy frameworks. If a laboratory elected to use forensic DNA phenotyping on a wider range of samples than crime scene samples of unknown origin, Table 3 also outlines how the status of the personal information or data may change.

Given these factors, it would likely be necessary to consider options to further mitigate privacy risks. This is consistent with the necessity and proportionality factors outlined previously.

Demonstrating a strong and robust mitigation strategy for privacy risks associated with genetic information held by law enforcement will help to counter arguments for legislation governing the use of phenotype markers [57]. Moreover, any privacy assessment must consider whether such expansive use is consistent with the consent given by donors, and with coercive arrangements for DNA collection from suspects and convicted offenders.

5.3. 'Masking' or encrypting personal information

If not strictly limiting DNA analysis to identity markers compatible with CODIS or similar DNA databases, one approach to protecting personal information gained from phenotype markers - not presently relevant to an investigation - would be to mask or encrypt the data, and to reveal it only as investigative priorities dictate. The concept of encrypting genetic information is being considered across a number of health care applications [58,59].

An ideal implementation of this safeguard would see an MPS platform deliver results that are partially obfuscated to the scientist, as illustrated in Table 4.

If analysis results in a full set of identity markers, uploading those markers to the unsolved crime scene index of a national DNA database, and returning no matches, could instead return an encryption key allowing the scientist to unlock all or a portion of the remaining genetic data, as illustrated in Table 5.

Such an approach, while providing a high degree of privacy protection and being consistent with accountability considerations, would require technical cooperation between laboratories, MPS manufacturers and administrators of DNA databases such as CODIS.

A hybrid model could be implemented within a laboratory by segmenting data and making the full genetic output available only to selected laboratory technicians. In a similar way, the technician could unmask the data, providing supplementary information to forensic scientists and, by extension, investigators. Tightly controlling and auditing the access to genetic information from an MPS platform is, in any event, highly desirable in any privacy-compliant implementation. Such a hybrid approach is also advantageous in that raw machine output is still available to laboratory staff responsible for the MPS platform, for quality assurance and related purposes.

This is not entirely dissimilar to the approach proposed by the United Kingdom Commissioner for the Retention and Use of Biometric Material to develop a ‘logical’ database separation of previously siloed data, under appropriate governance and privacy safeguards [60]. The common element of these approaches is the use of technology to provide an overlay or safeguard to help manage access to personal information.

5.4. Information security, de-identification, re-identification and deletion

A privacy assessment must carefully consider information security and security infrastructure requirements. With limited consent (and, in some possible cases, coercive collection of DNA), the implications of loss or misuse of genetic data would be severe [16]. Appropriate safeguards should be implemented to ensure access to genetic data is limited, and that the possibility of unauthorised access is minimised.

While commentators such as Kitchen [61] have suggested that police DNA databases will likely expand to include phenotype information, such an outcome appears unlikely, at least in the context of disaggregated genetic information. The sharing between policing agencies of BGA and EVC predictions, in the form of an intelligence report, could be beneficial in identifying repeat offenders across different cases and jurisdictions. For example, one case may involve DNA evidence (generating BGA and EVC information about a suspect who was likely to be of European background with green eyes and brown hair) and another case in a neighbouring jurisdiction may have no DNA evidence but an eyewitness who identified a person with those attributes. Coupled with other investigative information, such as a common *modus operandi*, it may be possible for investigators to hypothesise that the same offender may be linked to both cases.

However, such intelligence sharing does not require phenotype markers to be included in any criminal database. Intelligence in the form of phenotype markers is only useful in so far as it leads investigators to a suspect who could then obtain a conventional DNA profile (using identity markers) from the suspect for comparison with database records and/or crime scene biological evidence. Phenotype markers add nothing of investigative value to existing national DNA databases of identity markers. Identity markers are sufficient to link a crime scene to another crime scene, or a suspect to a crime scene.

Sharing of genetic information would only appear to be beneficial to aid in future research collaboration, such as to help identify new phenotype traits and thereby improve the capabilities of an MPS platform. It is already routinely used for this purpose where voluntary donation of DNA with informed consent for research is governed by the requirements of institutional ethics committees and scientific journals where the Helsinki Declaration is generally accepted as a minimum standard [62]. It has been suggested that uploading crime scene samples from cases with clearly identified suspects amounts to a ‘backdoor’ to coercive collection requirements. The erroneous upload of victim profiles has also been documented [12]. This would suggest that quite robust information security practices should be implemented to safeguard genetic information, irrespective of the platform used for analysis.

The separation of genetic information from research samples, necessary for validating and for enhancing MPS capabilities, from data derived from criminal investigations would be an appropriate privacy safeguard. The consent arrangements around the collection of genetic information for those purposes can be significantly distinct from those that apply to research samples (Table 6).

Privacy legislation often relaxes information security requirements when dealing with 'de-identified' personal information or data, in some cases relaxing the purposes for which the data can be used [35,36,46]. The *Federal DNA Identification Act of 1994*, as an example, makes it lawful for the FBI to allow very limited access to de-identified genetic information from their national DNA database for purposes such as population statistics or research [12].

However, there is a growing consensus that de-identification is a privacy enhancement and not a panacea [40]. As Anderson [63] explains, a surprisingly small amount of personal information is needed to re-identify supposedly anonymous personal data: Anderson's research showed that 87 per cent of Americans could be uniquely identified using just three pieces of information. Angrist [64] described work undertaken at the Massachusetts Institute of Technology to demonstrate re-identification of previously anonymised genetic profiles. Other commentators demonstrated similar results [65,66]. Law enforcement samples may be even more susceptible, given that – unlike an online genetic database – an adversary could make certain assumptions that an individual (such as a prominent suspect) are more likely to be within a given dataset.

Stanford University researchers recently demonstrated that it was possible, in a high proportion of cases, to re-identify genetic information associated with forensic DNA phenotyping by cross-referencing with identification markers (short tandem repeats) [67]. Of particular note, it was unnecessary to have an actual overlap of genetic information to use as a data linkage. Such a data linkage (for example, post code and age cross-referenced to name and date of birth) is the more common means of successfully re-identifying disparate data holdings [63].

Commentators such as Culnane et al [68] and Mittelstadt and Floridi [40] have discussed options to criminalise the re-identification of anonymised data, an approach recently proposed in Australia with respect to certain government datasets [69].

A forensic laboratory releasing de-identified genetic data could therefore quite easily find itself in a situation where a re-identification attempt was successful. The familial aspect of genetic information adds further privacy risks [70].

De-identification, however, can still be a useful mechanism for enhancing privacy. The separation of reference DNA profile information can usefully create a data silo, where an inadvertent or malicious data spill either results in the release of names, or genetic information, but not both. The siloing of genetic information has been proposed by Humbert et al [65] with potential application across a range of biomedical and public databases. While released information could still be re-identified, or could aid in re-identification of other publicly available genetic information, this would require both a data spill and a second deliberate, malicious step.

De-identification of genetic information may also be a legislative requirement for certain samples. For example, in Australia, forensic DNA profiles must be deleted or permanently de-identified in certain circumstances [71]. Given the risk of re-identification is demonstrably higher when dealing with predictive DNA, a laboratory may need to err on the side of actual deletion of genetic data, where possible.

The extent to which this is feasible depends on information flows through each laboratory system. It should be noted that the physical destruction of data is a particularly difficult task, given the way modern computer systems operate and the necessary backup regimes. A privacy assessment would need to consider whether reasonable efforts to delete data from a laboratory's operational systems are sufficient, when coupled with policies that would prevent inappropriate access to deleted data, and note that there is a residual privacy risk if, for example, a decision was made to restore a backup tape to access previously deleted genetic information.

6. Conclusion

We have attempted to examine the context underpinning a PIA for forensic DNA phenotyping. It explains the role that the PIA process can play, as a contributor to ensuring legal compliance but also engaging with the broader issues of public trust and confidence in forensic DNA. A thorough and thoughtful PIA will significantly strengthen any implementation of new technology in a forensic setting, particularly where there is a flow-on effect on individual privacy.

The article then goes on to draw out some of the key areas, within the construct of a model PIA or 'straw' policy. By ensuring careful consideration of information flows, consent and future applications, a robust privacy framework can be developed around this new technology.

Acknowledgements

See separate file

Acknowledgements

The authors would like to thank the two 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 Fellowships and Awards, a Department of Education and Training initiative.

References

1 Y. Yang, B. Xie, J. Yan, Application of next-generation sequencing technology in forensic science, Genomics, proteomics & bioinformatics, 2014;12(5):190-197, https://doi.org/10.1016/j.gpb.2014.09.001

2 C. Phillips, A. Salas, J. J. Sanchez, M. Fondevila, A. Gomez-Tato, J. Alvarez-Dios, M. Calaza et al., Inferring ancestral origin using a single multiplex assay of ancestry-informative marker SNPs, Forensic Science International: Genetics 1, no. 3 (2007):273-280, https://doi.org/10.1016/j.fsigen.2007.06.008

3 S. Walsh, F. Liu, A. Wollstein, L. Kovatsi, A. Ralf, A. Kosiniak-Kamysz, W. Branicki and M. Kayser, 2013, The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA, Forensic Science International: Genetics, 7(1):98-115, https://doi.org/10.1016/j.fsigen.2012.07.005

4 M. Marcińska,, E. Pośpiech, S. Abidi, J.D. Andersen, M. van den Berge, Á. Carracedo, M. Eduardoff, A. Marczakiewicz-Lustig, N. Morling, T. Sijen and M. Skowron, 2015, Evaluation of DNA variants associated with androgenetic alopecia and their potential to predict male pattern baldness. PloS one, 10(5), p.e0127852, https://doi.org/10.1371/journal.pone.0127852

5 R. Zbieć-Piekarska, M. Spólnicka, T. Kupiec, A. Parys-Proszek, Ż. Makowska, A. Pałeczka, K. Kucharczyk, R. Płoski and W. Branicki, 2015, Development of a forensically useful age prediction method based on DNA methylation analysis. Forensic Science International: Genetics, 17:173-179, https://doi.org/10.1016/j.fsigen.2015.05.001

6 S. E. Medland, D.Z. Loesch, B. Mdzewski, G. Zhu, G.W. Montgomery and N.G. Martin, 2007. Linkage analysis of a model quantitative trait in humans: finger ridge count shows significant multivariate linkage to 5q14. 1. PLoS genetics, 3(9), p.e165, https://doi.org/10.1371/journal.pgen.0030165

7 S. Rushton, Familial Searching and Predictive DNA Testing for Forensic Purposes: A Review of Laws and Practices, Victoria Law Foundation Legal Police Internship Program, 2010.

8 *Daubert v. Merrell Dow Pharmaceuticals, Inc.* 509 U.S. 579 (1993).

9 National Academy of Sciences (2009), Strengthening Forensic Science in the United States: A Pathway Forward.

10 R. Williams and M. Wienroth. Social and Ethical Aspects of Forensic Genetics: A Critical Review. Forensic Science Review 2017, 29(2), 145-169.

11 A. Etzioni, DNA Tests and Databases in Criminal Justice Individual Rights and the Common Good, in D. Lazer (ed.), DNA and the Criminal Justice System: The Technology of Justice, Cambridge 2004, 197-224.

12 E. Murphy, Inside the Cell: The Dark Side of Forensic DNA, New York NY: Nation Books, 2015.

13 R. Williams and W. Matthias, 'Ethical, social and policy aspects of forensic genetics: A systematic review' (2014).

14 J. Lunshof and R. Chadwick, Genomics, inconvenient truths and accountability, in: The Right to Know and the Right Not to Know: Genetic Privacy and Responsibility, 2014, 116-130.

15 Presidential Commission for the Study of Bioethical Issues, Privacy and Progress in Whole Genome Sequencing, 2012.

16 L.O. Gostin, Genetic Privacy, The Journal of Law, Medicine & Ethics. 1995;23(4):320-330, https://doi.org/10.1111/j.1748-720x.1995.tb01374.x

17 Australian Law Reform Commission, Report 96: Essentially Yours - The Protection of Human Genetic Information in Australia, 2003.

18 J.E. Lunshof, R. Chadwick, D.B.Vorhaus, G.M. Church, From genetic privacy to open consent, Nature Reviews Genetics, 9(5), 2008, 406-411, https://doi.org/10.1038/nrg2360.

19 Human Genetics Commission, Nothing to hide, nothing to fear: Balancing individual rights and public interest in the governance and use of the National DNA Database (2009).

20 D. Gusella, No Cilia Left Behind: Analyzing the Privacy Rights in Routinely Shed DNA Found at Crime Scenes, 54 B.C. L. Rev. 789 (2013), http://lawdigitalcommons.bc.edu/bclr/vol54/iss2/9

21 J.K. Wagner, Out with the "Junk DNA" Phrase. Journal of Forensic Sciences. 2013;58(1):292-294, https://doi.org/10.1111/j.1556-4029.2012.02252.x

22 M. Gymrek, T. Willems, A. Guilmatre, H. Zeng, B. Markus, S. Georgiev, M.J. Daly, A.L. Price, J.K. Pritchard, A.J. Sharp and Y. Erlich, 2016, Abundant contribution of short tandem repeats to gene expression variation in humans. Nature genetics, 48(1):22-29, https://doi.org/10.1038/ng.3461

23 H. Machado, S. Silva, Informed consent in forensic DNA databases: volunteering, constructions of risk and identity categorization, Biosocieties, 2009;4(4):335-348, https://doi.org/10.1017/s1745855209990329

24 C. Curtis, Public Understandings of the Forensic Use of DNA Positivity, Misunderstandings, and Cultural Concerns, Bulletin of Science, Technology & Society, 2014;34(1-2):21-32, https://doi.org/10.1177/0270467614549415

25 C. Novas, N. Rose, Genetic risk and the birth of the somatic individual, Economy and Society, 2000, 29(4), 485-513, https://doi.org/10.1080/03085140050174750

26 M. Turrini,B.Prainsack. "Beyond clinical utility: the multiple values of DTC genetics." Applied & translational genomics 8 (2016): 4-8, https://doi.org/10.1016/j.atg.2016.01.008

27 C.E. MacLean, Creating a wanted poster from a drop of blood: using DNA phenotyping to generate an artist's rendering of an offender based only on DNA shed at the crime scene. Hamline L Rev. 2014;36(3):1.

28 M.K. Cho, P. Sankar, 2004, Forensic genetics and ethical, legal and social implications beyond the clinic, Nature genetics, 36, https://doi.org/10.1038/ng1594

29 H. Machado, S. Silva, Public participation in genetic databases: crossing the boundaries between biobanks and forensic DNA databases through the principle of solidarity, Journal of Medical Ethics, 2015:medethics-2014-102126, https://doi.org/10.1136/medethics-2014-102126

30 E.R. Pike, Securing sequences: ensuring adequate protections for genetic samples in the age of Big Data. Cardozo L Rev 2015;37

31 O. Tene, J. Polonetsky, A theory of creepy: technology, privacy and shifting social norms, Yale JL & Tech, 2013;16:59.

32 A. Moore, Defining privacy, Journal of Social Philosophy, 2008 Sep 1;39(3):411-28, https://doi.org/10.1111/j.1467-9833.2008.00433.x

33 Victorian Commissioner for Privacy and Data Protection, Privacy Impact Assessment Template, 2017, https://www.cpdp.vic.gov.au/images/content/word/PIA_Template_March_2017.doc

34 D. Wright, P. De Hert, Introduction to privacy impact assessment, in: Privacy Impact Assessment, Springer, Dordrecht, 2012, 3-32.

35 Australian Privacy Principles, Sch 1, Privacy Act 1988 (Cth).

36 Privacy Act of 1974, 5 U.S.C. § 552a.

37 A. Johnston, Presentation: Privacy Impact Assessment Workshop. 2015, https://www.oaic.gov.au/resources/agencies-and-organisations/training-resources/privacy-impact-assessment-pia-workshop-slides.pdf

38 Office of the Australian Information Commissioner. Guide to undertaking privacy impact assessments. 2014

39 United Kingdom Information Commissioner's Office, Conducting Privacy Impact Assessments: Code of Conduct, 2014, https://ico.org.uk/media/for-organisations/documents/1595/pia-code-of-practice.pdf

40 B.D. Mittelstadt, L. Floridi, The ethics of big data: Current and foreseeable issues in biomedical contexts, in: The Ethics of Biomedical Big Data, Springer International Publishing, 2016, 445-480, https://doi.org/10.1007/978-3-319-33525-4_19

41 P. De Hert, A Human Rights Perspective on Privacy and Data Protection Impact Assessments, in: Privacy Impact Assessment. Springer, Dordrecht, 2012, 33-76.

42 R. Clarke, Privacy impact assessment: Its origins and development, Computer law & security review. 2009 Dec 31;25(2):123-35, https://doi.org/10.1016/j.clsr.2009.02.002

43 D. Wright, Making privacy impact assessment more effective, The Information Society 29, no. 5 (2013), 307-315, https://doi.org/10.1080/01972243.2013.825687

44 C. Børsting, N. Morling, Next generation sequencing and its applications in forensic genetics, Forensic Science International: Genetics. 2015;18(September 2015), 78-89, https://doi.org/10.1016/j.fsigen.2015.02.002.

45 Privacy Act 1988 (Cth)

46 General Data Protection Regulation, Regulation EU 2016/679

47 Amended Act on the Protection of Personal Information (Japan), 2016; Act on the Protection of Personal Information Held by Administrative Organs (Japan), 2003

48 M. Gloudemans, N. Shamaprasad, Current Issues in Forensic DNA Applications. http://pged.org/wp-content/uploads/2016/03/Current-Issues-in-Forensic-DNA-Applications.pdf, 2015 (accessed 17.05.16).

49 Forensic Genetics Policy Initiative, Establishing best practice for forensic DNA databases, 2017, http://dnapolicyinitiative.org/wp-content/uploads/2017/08/BestPractice-Report-plus-cover-final.pdf (accessed 16.12.17)

50 B. J. Koops, M. Schellekens, Forensic DNA phenotyping: regulatory issues. Colum Sci & Tech L Rev. 2008;9:158, https://doi.org/10.2139/ssrn.975032

51 Office of the Australia Information Commissioner. Australian Privacy Principles guidelines, 2015.

52 United States Armed Forces DNA Identification Laboratory, Request for Mitochondrial DNA Sequence Report, 2012, https://health.mil/Reference-Center/Forms/2012/02/01/Request-for-Mitochondrial-DNA-Sequence-Report

53 E. Kretowicz, Emerging DNA technology will impinge on privacy: Civil Liberties Australia, The Canberra Times, 17 Nov 2013.

54 E. Murphy, Legal and ethical issues in forensic DNA phenotyping. New York University Public Law and Legal Theory Working Papers Paper 415. 2013, pp 13-46, https://doi.org/10.2139/ssrn.2288204

55 B. Malin, L. Sweeney (editors), Re-identification of DNA through an automated linkage process, Proceedings of the AMIA Symposium, 2001, American Medical Informatics Association.

56 P. Gill, Misleading DNA Evidence: Reasons for Miscarriages of Justice, International Commentary on Evidence 10, No. 1, 2012, https://doi.org/10.1515/ice-2014-0010

57 M. Smith, G. F. Urbas, Regulating new forms of forensic DNA profiling under Australian legislation: familial matching and DNA phenotyping, Australian Journal of Forensic Sciences, 2012;44(1):63-81. https://doi.org/10.1080/00450618.2011.581250

58 E. Vayena, U. Gasser, Between openness and privacy in genomics. PLoS Med. 2016;13(1):e1001937, https://doi.org/10.1371/journal.pmed.1001937

59 E. C. Hayden, Extreme cryptography paves way to personalized medicine, Nature 519, no. 7544 (23.03.2015), https://doi.org/10.1038/519400a

60 P. Wiles, Annual Report 2016: Commissioner for the Retention and Use of Biometric Material, 2017, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/644426/CCS207_Biometrics_Commissioner_ARA-print.pdf (accessed 16.12.2017)

61 A. N. Kitchen, Genetic privacy and latent crime scene DNA of non-suspects: how the law can protect an individual's right to genetic privacy while respecting the government’s important interest in combating crime, Crim Law Bull. 2015;52(2).

62 World Medical Association, 2001, World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects, Bulletin of the World Health Organization, 79(4), 373

63 N. Anderson, "Anonymized" data really isn't - and here's why not, Ars Technica, 2009.

64 M. Angrist, Genetic privacy needs a more nuanced approach. Nature. 2013;494:7, https://doi.org/10.1038/494007a

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

66 C. Lippert, R. Sabatini, M.C. Maher, E.Y. Kang, S. Lee, O. Arikan, A. Harley, A. Bernal, P. Garst, V. Lavrenko and K. Yocum, 2017, Identification of individuals by trait prediction using whole-genome sequencing data. Proceedings of the National Academy of Sciences, 114(38), pp.10166-10171,

67 M. D. Edge, B. F. B. Algee-Hewitt, T. J. Pemberton, J. Z. Li, N. A. Rosenberg, Linkage disequilibrium matches forensic genetic records to disjoint genomic marker sets, Proceedings of the National Academy of Sciences. 2017 May 15;114(22):5671–6, https://doi.org/10.1073/pnas.1619944114, https://doi.org/10.1073/pnas.1711125114

68 C. Culnane, B. Rubinstein, V. Teague, Can the government really protect your privacy when it 'de-identifies' public data?, The Canberra Times.,5 Dec 2016

69 Privacy Amendment (Re-identification Offence) Bill 2016

70 E. R. Pike, Securing sequences: ensuring adequate protections for genetic samples in the age of Big Data, Cardozo L Rev. 2015;37, 1977

71 Part 1D, Crimes Act 1914 (Cth)

Figures and table

Table 1. Possible forensic applications of genetic targets

Genetic targets	Law enforcement application	Included in this privacy impact analysis?
Identity markers (CODIS or other databases) in: • Crime scene samples • Reference samples	• Matching a questioned (evidential) biological sample to a suspect (reference) sample or database record • MPS may help to deconvolute mixtures because of enhanced capabilities over capillary electrophoresis • MPS may ultimately provide more efficient processing of samples than capillary electrophoresis	No (regarded as the baseline privacy position)
Markers for BGA and/or EVCs in: • Crime scene samples	• Can provide forensic intelligence to investigators.	Yes
Other phenotype markers relevant to identifying the donor of: • Crime scene samples	• Can provide forensic intelligence to investigators, such as the donor's likely health care or medicinal requirements.	Yes, as a possible future application
Mitochondrial DNA (mtDNA) including the full mitochondrial genome) in: • Crime scene samples • Reference samples	• Matching a questioned (evidential) biological sample to a suspect (reference) sample or database record • Can also be used for probabilistic phenotype prediction (especially metabolic phenotypes) • More sensitive when dealing with highly degraded samples	Yes
Whole genome in: • Crime scene samples	• May allow future 'cold case' analysis, retrospectively analysing other parts of the genome for new EVC or BGA markers not known to be predictive at the time of analysis.	Yes, as a possible future application.
Markers for DNA-based prediction of behavioural traits in: • Crime scene samples	• May provide information to the criminal justice system and predisposition to criminality, or behavioural profiling of offenders from	No. Requires a separate PIA.

- Reference samples	discarded DNA.	

Table 2. Possible sources of genetic information and their status under selected privacy laws

Source of sample from which genetic information derived	Donor identified or reasonably identifiable	Donor consent obtained	Derived genetic information is…			DNA phenotyping adds value to forensic investigation
			personal information (Australia)	personal data (European Union)	a privacy record (United States)	
Crime scene sample						
closely associated with an individual (e.g. blood stain believed to be from suspect in custody; victim of crime)	Yes	In some cases	Yes	Yes	No	No
unknown origin	No	No	Not until identified	Not until identified	Not until identified	Yes
Reference sample						
obtained voluntarily	Yes	Yes	Yes	Yes	Yes	No
obtained coercively	Yes	No	Yes	Yes	Yes	No

Table 3. Sources of DNA and EVC/BGA priority for analysis

Source of sample from which genetic information derived	Personal information/ data/ privacy record at time of first analysis	Result of database checks against reference DNA profiles	Phenotype relevant to case after database search	Personal information/ data / privacy record, if phenotyped	Personal information /data / privacy record, if subsequently identified
Crime scene sample					
closely associated	Yes	Probabilistic Match	Not relevant	Yes	Already identified

with an individual (e.g. blood stain believed to be from suspect in custody)		No Match	Not relevant	Yes	Already identified
unknown origin	No	Probabilistic Match	Not relevant	Yes	Already identified
		No Match	Relevant	No	Yes
Reference sample					
obtained voluntarily	Yes	Not Applicable	Not relevant	Yes	Already identified
obtained coercively	Yes	Not Applicable	Not relevant	Yes	Already identified

Table 4. Partially obfuscated analysis results

Identity markers	BGA	EVC	Health prediction
Available 🔓	Not available 🔒	Not available 🔒	Not available 🔒

Table 5. Unlocking BGA and EVC data after no criminal database match using identity markers

Identity markers	BGA and EVC	Health prediction	Other markers
No match	Available 🔓	Not available 🔒	Not available 🔒

Key ↥

Table 6. Types of sample donors and primary purpose of collection

Sample donor	Primary purpose for collection	Likely to be aware of genetic privacy implications	Would consent be valid for research purposes?
Victim of crime	Elimination	No	Possibly not
Bystander	Elimination	No	Possibly not
Suspect – volunteer	Comparison	No	Possibly not
Suspect – coercive	Comparison	No	No

Research participant	Research	Yes	Yes
Laboratory employee	Elimination or Research	Yes	Yes