FORENSIC DNA DATABASES IN HONG KONG AND CHINA: A BRICS COMPARATIVE PERSPECTIVE

ALE ROUX-KEMP

I. INTRODUCTION

Two developments at the turn of the 20th century changed criminal justice practice fundamentally: In 1984-1985, British geneticist Alec Jeffreys and colleagues developed “DNA fingerprinting,” the forerunner to the extraordinary DNA profiling techniques that today are regarded as “an indispensable tool for the identification of individuals” in criminal justice.1 And the second, which is often overshadowed by the first, is the advances in information and computing technology that have “rendered affordable, the compilation and maintenance of large databases as well as the rapid searching of these databases. . . .”2 This article will focus on an ensuing development, namely Forensic DNA Databases.

The first national criminal or Forensic DNA Database was the National DNA Database (DNAD) of the United Kingdom which became operational in April 1995.3 At the same time, the FBI developed the USA national (federal) DNA database (managed through the Combined DNA Index System software (CODIS)). By 2000, many European countries also established their own DNA databases, and as the application of DNA profiling continues to permeate the realm of criminal justice, more countries continue to extend the application of forensic DNA typing in criminal case work.4 Yet, with regard to the BRICS countries, very little information is generally available, and it seems as though these jurisdictions are either still in the process of establishing their national Forensic DNA Databases or lack the necessary regulatory and oversight legislation for managing their existing databases. BRICS – the acronym that is used for Brazil, Russia, India, China and South Africa – are a collective of middle income countries that are pivoting to form an alternative voice in a multipolar world still dominated by the United States and her western allies. BRICS is increasingly creating new platforms for development and cooperation, but vast developmental differences remain.

---

* Assistant Professor, School of Law, City University of Hong Kong (HKSAR).
An overview of the use and development of Forensic DNA technologies and databases in the BRICS countries reveal the following: South Africa has been conducting DNA profiling since 1991 and established a DNA database in 1997, but it only recently promulgated legislation - the South African Criminal Law Forensic Procedures Amendment Act 37 of 2013 - for the formal recognition and regulation of its DNA database. In 2008, Russia adopted the National Genetic Registration, the Russia Federal Act N242-FZ, enabling Russian law enforcement to create genome record databases of convicted offenders. This Act only came into operation in 2010. Brazil, in turn, promulgated legislation (Law No. 12654) for the establishment and regulation of a DNA database and DNA profiling on May 28th, 2012, while India is still in the planning stages of its national DNA framework and enabling legislation (Draft Human Profiling Bill 2012), despite India already being well-established for performing DNA profiling for other jurisdictions in the Southeast Asia region.5

The primary focus of this article is on China and one of its special administrative regions, Hong Kong. First, a general introduction to Forensic DNA Typing will restate the rudiments on which the application of DNA technology in the realm of law enforcement and criminal justice is based. This background is necessary for the discussion on database design in the final part of the article. The Forensic DNA Databases and related legislative frameworks of China and Hong Kong will then be considered and the relevant scientific and legislative features contextualized. Finally, the article will explore two distinct features with regard to the database design of the Forensic DNA Databases of China and Hong Kong, namely the population size of DNA Databases and the possibilities for cross-border collaboration. The discussion here will be supplemented with references to related developments in BRICS countries as well as selected references and comparisons with jurisdictions having more developed Forensic DNA Databases, particularly that of the UK and the USA.

II. FORENSIC DNA TYING: THE NATURE OF DNA AS AN IDENTITY MARKER

Deoxyribonucleic acid (DNA) is the chemical building blocks of a living organism’s genetic makeup. This genetic material is stored in chromosomes, which are located in the nuclei of the hundreds of billions of cells that make up our physical composition and, to a certain extent, also account for our behavioral characteristics.6 Although all humans share approximately 95% to 99% of their nucleotide “fingerprint,” a small percentage thereof is entirely unique (except for identical twins), and it is this highly individualizing feature that makes DNA such a powerful tool in forensic science.7

Human cells contain 23 pairs of chromosomes, one pair inherited from each

5. Id.
7. LIRIEKE MEINTJES-VAN DER WALT, DNA IN THE COURTROOM: PRINCIPLES AND PRACTICE 3 (Cape Town: JUTA, 2010).
parent. Each of these chromosomes consists of a singular linear molecule of DNA, known as a double helix. The double helix, in turn, consists of two coiled sugar-phosphate-based strands, which is formed by the hydrogen bonding of the chemical bases set on these strands. While the bonding of these chemical bases - Adenine (A), Thymine (T), Cytosine (C) and Guanine (G) - are fixed with T always pairing with A and G always with C, the order in which the bases appear on these single strands are unique. It is known as a DNA sequence. It is because of this unique order in which the bases appear on the single strands of a DNA molecule (in other words the uniqueness of the DNA sequence), that we are able to use DNA as an identity marker.

A. Non-coding and Coding DNA

In terms of DNA forensics, a DNA sample refers to the DNA material collected at a crime scene or from a suspect, an offender, or any other person. A DNA sample can therefore be in the form of blood, hair, skin, saliva, or other biological material found, for example, under a person’s fingerprints. A DNA profile, on the other hand, refers to ” . . . the numerical representation of specific loci that represents an individual’s unique sequence of DNA base pairs.” Loci are specific regions of a chromosome that analysts use in order to determine the likelihood whether a particular DNA profile belongs to a specified person. Loci, or a locus (singular), contain both coding and non-coding regions, and it is usually the non-coding regions that are used for the purpose of forensic DNA analysis.

Non-coding regions contain polymorphic DNA markers that are highly individual and form the basis of existing methods of DNA analysis. It has also been described as ‘junk’ or benign DNA because it does not contain any important personal information about an individual. The coding regions of a chromosome, on the other hand, only make up about 5% of the total DNA strand and contain information needed to produce proteins responsible for dictating cell functions. These specific loci are generally referred to as genes. A gene occupies a specific region on a chromosome and is a portion of DNA of 1,000 to 100,000 base pairs in length. Genes are scattered throughout the DNA genome and are

9. Meintjes-Van der Walt, supra note 7, at 5; De Wet, Oosthuizen, & Visser, supra note 6, at 171.
11. Id.
12. De Wet, Oosthuizen & Visser, supra note 6, at 172.
13. Lewis, supra note 8, at 521.
14. De Wet, Oosthuizen & Visser, supra note 6, at 172.
15. Meintjes-Van der Walt, supra note 10, at 497; De Wet, Oosthuizen & Visser, supra note
separated by vast stretches of DNA that do not code anything.\textsuperscript{16} It is also important to note that the mere existence of a particular gene does not necessarily result in it producing a protein. It must first be activated, and a number of internal and external influences may be required for this to occur.\textsuperscript{17}

The exact location or site of either a specific gene or polymorphic marker on a chromosome is called a locus (plural: loci), and is important in the identification process since a copy of a gene or DNA marker always resides on the same locus on each of the chromosomes in a pair.\textsuperscript{18} Another concept and term that is important for the discussion on database design that follow in this article is “alleles,” which refers to the two copies of a gene on each of the two copies of a chromosome.\textsuperscript{19} A genotype represents the alleles at a specific locus and a combination of several genotypes from multiple loci forms the DNA profile of an individual.\textsuperscript{20} The entire DNA complement in the nucleus of a cell is known as a genome, and every cell in the human body, except for the red blood cells and sperm and female eggs, contains a complete copy of the human genome.\textsuperscript{21} The size of the genome will depend on the type of organism.\textsuperscript{22}

The remainder of this article will focus on non-coding DNA, the extraction of DNA profiles from such DNA, and the inclusion thereof on DNA databases. With regard to coding DNA, it can be noted that in 2003, a 13-year project, The Human Genome Project (HGP), was completed. The primary aim of this project was to identify all the approximately 20,000 – 25,000 genes in the human DNA and to determine the sequences of the 3 billion chemical base pairs that make up human DNA.\textsuperscript{23} Now that the genome sequence has been completed, research is shifting from the 99.9\% genetic similarity amongst humans to the 0.1\% unique genetic differences.\textsuperscript{24} This development also has important implications for behavioral genetics research and its increasing application in criminal justice.\textsuperscript{25}

\textsuperscript{6}, at 172.


18. De Wet, Oosthuizen \& Visser, \textit{supra} note 6, at 172.


20. De Wet, Oosthuizen \& Visser, \textit{supra} note 6, at 172.


22. Lewis, \textit{supra} note 8, at 522.


B. Current Methods of Forensic DNA Analysis

Forensic DNA analysis involves the comparison of a set number of loci of two (or more) DNA profile(s), and is based on the unique sequence of the base pairs identified in the DNA samples.26 Since the DNA of all organisms contains the same chemical components, this unique DNA sequence also accounts for the differences between species.27 The designated loci for DNA analysis are carefully selected based on its variability to ensure greater differentiability between samples during the forensic analysis.28 In fact, the higher the possible variation in DNA markers, the greater the discrimination between comparative DNA samples will be.29

Most DNA laboratories use the same methods and standards for DNA typing. First, a DNA amplification technique, called Polymerase Chain Reaction (PCR), is used to increase the amount of DNA at all the relevant loci to ensure a sufficient quantity for DNA typing. The PCR method targets and copies specific loci of DNA in order to make millions of exact copies of DNA from a biological sample.30 This process is necessary as there are usually only a small quantity of DNA found at crime scenes and such DNA samples may furthermore be compromised or severely degraded.31 Thereafter, Short Tandem Repeats (STRs) are used in the analyses to compare alleles at a particular locus that are repeated consecutively. The variability of the number of repeats between individuals makes this an extraordinary method of identification, and it has had a tremendous impact on criminal justice systems worldwide.32

The DNA profile - the set of genotypes possessed by a person at two or more loci - extracted from the DNA sample in terms of the methods and processes described above, can then be compared with other DNA profiles to determine whether the DNA samples belong to the same person.33 This comparison entails an analysis of all the alleles at the selected loci of the DNA profile.34 A “consistency” between two DNA profiles is found if all the alleles at all the same loci of the two profiles correspond.35 If the DNA profile obtained from two samples is indistinguishable, a criminal court would normally treat that as a “match.”36 Thus, the court will regard this consistency as [circumstantial]

27. Lewis, supra note 8, at 521.
29. De Wet, Oosthuizen & Visser, supra note 6, at 172.
30. Id. at 176.
31. Lewis, supra note 8, at 522; Jobling & Gill, supra note 26, at 741.
32. De Wet, Oosthuizen & Visser, supra note 6, at 178; Jobling & Gill, supra note 26.
33. De Wet, Oosthuizen & Visser, supra note 6, at 178; Jobling & Gill, supra note 26.
34. De Wet, Oosthuizen & Visser, supra note 6, at 176; Jobling & Gill, supra note 26, at 743.
35. Id.
36. Id.
evidence that the samples have a common source, for instance that semen found at a rape scene was contributed by the suspect with the matching DNA profile.\textsuperscript{37} While there are a few (older) cases where such consistencies in DNA profiles were viewed by the courts as direct evidence - see for instance \textit{State v. Mosely (1994/5)}\textsuperscript{38} - the more generally accepted view today is that consistencies in DNA samples treated as circumstantial evidence, albeit often highly reliable circumstantial evidence.\textsuperscript{39}

Once such a consistency is confirmed, the DNA profile is also compared to DNA profiles in a population database in order to determine the likelihood or probability of another individual, other than the accused or the suspect, being the depositor/donor of the DNA sample. Due to the inherent limitations of the DNA analysis process, absolute certainty of identification is not possible, and it is therefore necessary for the probative value of a consistency to be calculated statistically, in order to determine how likely it is for the DNA sample to have come from a third party with the same DNA profile as the suspect/accused. The statistical significance of a consistency – the frequency with which an entire pattern of the selected loci occurs in a target population – must therefore be determined by a series of calculations in terms of a reference population database. Population databases are compiled in terms of the principles of population genetics and take into account factors like race and ethnicity for a particular geographic area.\textsuperscript{40} Thus, by estimating the frequency with which that particular DNA profile would occur at random in that population, the significance of a consistency is determined. This is known as the random match probability and “describes the statistical probability of a randomly selected person having” the same DNA profile as the profile found at the crime scene or the DNA profile included in the database.\textsuperscript{41}

Mitochondrial DNA analysis (mtDNA) and Y-chromosome analysis is sometimes also used to determine familial relationships between a DNA profile and a potential female or male relative respectively. MtDNA is especially effective when dealing with severely degraded DNA samples, as well as DNA samples that lack nucleated cellular material such as hair, bones and teeth. Such biological material cannot be analyzed by means of the conventional tests (RFLP, PCR and STR).\textsuperscript{42} (PCR and STR were discussed above. RFLP or Restriction Fragment Length Polymorphism was one of the first applications used for DNA analysis in the context of forensic investigations and is no longer widely used.)\textsuperscript{43}

In cases where the conventional tests are not effective in analyzing the biological


\textsuperscript{39} Kiely, \textit{supra} note 37, at 474.

\textsuperscript{40} De Wet, Oosthuizen & Visser, \textit{supra} note 6, at 175.

\textsuperscript{41} Id.

\textsuperscript{42} Id.

\textsuperscript{43} Id. at 176.
material, mitochondrial DNA that is passed from mothers to their offspring can be compared with the DNA profile of potential mothers in order to determine if there is a familial connection between the potential mother and the donor of the DNA sample. Y-chromosome analysis, on the other hand, analyses the genetic markers on the Y-chromosome in order to determine familial relationships among males, as the Y-chromosome is passed directly from father to son.\(^4\) (The concerns that have been raised about familial testing and constitutional guarantees like the right to privacy do not fall within the ambit of this article.)\(^5\)

**C. Forensic DNA Databases**

The expansive impact that forensic DNA profiling has had on criminal justice practice, would not have been possible had it not been for the concomitant advances in information and computing technology. Within a decade of the first successful case in which DNA profiling was used to clear one suspect and bring the real perpetrator to justice, authorities created (and subsequently expanded) databases of criminal offenders’ DNA profiles.\(^6\) While DNA profiling can therefore assist in identifying a person or excluding a person if the DNA profiles don’t “match,” the development of Forensic DNA Databases has further expanded the application of DNA in criminal justice practice by way of generating investigative leads.

For example, a Forensic DNA Database increases the capacity to solve and prevent crime by making it possible for a DNA profile to be evaluated and compared against a large data set of other DNA profiles. Such comparative or speculative searches are particularly effective and hold many advantages for law enforcement, including the potential to positively identify unknown DNA profiles, to link DNA profiles from different crime scenes to one another, to assist in the identification of missing persons and unidentified human remains, to eliminate individuals from investigations and exonerate the wrongly convicted, to combat transnational crime, and to ultimately establish a powerful and cost effective tool to combat crime generally.\(^7\)

A Forensic DNA Database typically contains the known or identifiable DNA profiles of offenders, as well as the unknown DNA profiles extracted from DNA


\(^{6}\) Houck, supra note 1, at 255-56.

samples collected at crime scenes. However, the DNA profiles of various categories of persons may be included in the database. Below is a summary of the main indexes or profiling categories that are usually included in Forensic DNA Databases:

<table>
<thead>
<tr>
<th>DNA PROFILE INDEXES</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Crime Scene / Forensic Index</td>
<td>Unidentified / unknown DNA profiles extracted from DNA samples collected at crime scenes, or from the victim or deceased’s body or from any object (like clothing) found at the crime scene or on a victim or deceased.</td>
</tr>
<tr>
<td>Convicted Offender Index</td>
<td>Most jurisdictions have promulgated legislation with regard to the categories of convicted offenders whose DNA profiles must be included on the database. It is usually only restricted to serious offenders / offences and these DNA profiles usually remain on the database indefinitely. Section 15J(1)(b) of the South African Criminal Law Forensic Procedures Amendment Act 37 of 2013 even provides for the retrospective inclusion of DNA profiles of offenders convicted before the Act came into operation.</td>
</tr>
<tr>
<td>Suspect / Accused / Arrestee Index</td>
<td>DNA profiles from persons suspected, reported, arrested, charged, and/or cautioned for specific categories of offences.</td>
</tr>
<tr>
<td>Victim / Volunteer / Investigative Index</td>
<td>Contains the DNA profiles of victims (especially with regard to sexual offences) as well as other persons who provide their DNA profiles voluntarily for personal reasons or reasons related to a particular case/investigation etc.</td>
</tr>
<tr>
<td>Reference / Elimination Index</td>
<td>DNA profiles of officials working in the collection and analysis of forensic samples, e.g. police officials and laboratory staff.</td>
</tr>
<tr>
<td>Missing Person Index</td>
<td>DNA profiles of missing or unidentified persons.</td>
</tr>
<tr>
<td>Unknown Deceased Index</td>
<td>DNA profiles of unknown deceased persons and unidentified human remains.</td>
</tr>
</tbody>
</table>

Pertinent aspects with regard to the Forensic DNA Databases of China and Hong Kong (HKSAR) will now be considered. Reference will also be made to the Forensic DNA Databases in other parts of the world, particularly the UK NDNAD, the USA CODIS System, and related developments in BRICS countries.

48. From the South African Criminal Law Forensic Procedures Amendment Act 37 of 2013. See the Act for more detailed definitions of each of these profiling indexes.
III. The Forensic DNA Databases of China and Hong Kong (HKSAR)

According to the 2009 Global DNA Profiling Survey conducted by Interpol, twenty countries (Australia, Bangladesh, Bhutan, Brunei, China and Hong Kong (HKSAR), Fiji, India, Indonesia, Japan, Republic of Korea, Malaysia, Maldives, New Zealand, Pakistan, Philippines, Singapore, Sri Lanka, Thailand, Uzbekistan, and Vietnam) in the Asia and South Pacific region are known to use DNA analysis in criminal investigations and seven of those countries (Australia, China and Hong Kong (HKSAR), Japan, Republic of Korea, Malaysia, New Zealand, Singapore) are known to have a national Forensic DNA Database.\(^{49}\) This can be compared with the situation in 2002, when only eleven countries used DNA profiling in criminal investigations and only one country had a national DNA Forensic Database.\(^{50}\) The survey also showed that countries in the region generally took a significant number of years from the date that they first started to use DNA profiling in criminal investigations before setting up a national Forensic DNA Database. For example, the national Forensic DNA Database of Australia was set up in 2007, eighteen years after DNA profiling was first introduced and approximately fifteen years after the first regional Forensic DNA Database of Australia was established in 1992.\(^{51}\) China, Japan and Malaysia took thirteen, fifteen and eleven years respectively before setting up their national DNA databases while the Republic of Korea and Singapore in contrast, set up their databases in the same year that they began to use DNA profiling in their criminal justice systems.\(^{52}\) The data collected for the 2009 Global DNA Profiling Survey therefore seem to suggest that the later a country begins with DNA profiling, the sooner that country will establish a DNA database.\(^{53}\) The 2009 Interpol survey also indicated that the most used marker system by countries in the region is Identifier\(^\circ\) and for those countries that operate national Forensic DNA Databases, the database software used is mostly of national design.\(^{54}\) The main features of the Chinese and Hong Kong Forensic DNA Databases will now be considered.

A. China

China, along with Australia and Japan, were the first in the Asia and South Pacific region to use DNA profiling in their criminal justice systems in 1989, with the Genetics Laboratory of the Institute of Forensic Sciences being the first DNA analysis unit in China.\(^{55}\) New Zealand started to use DNA profiling in its criminal

\(^{49}\) Interpol DNA Unit, supra note 4, at 25, 28.

\(^{50}\) Id. at 28.

\(^{51}\) Id.

\(^{52}\) Id. at 28, 29.

\(^{53}\) Id. at 29.

\(^{54}\) Id. at 25.

\(^{55}\) Id. at 29.
justice system shortly thereafter in 1990, and set up the first Forensic DNA Database of the region in 1997, with China and Japan following suit in 2004. While Japan also promulgated legislation targeted at the regulation and oversight of its database a year after the database was established (2005), China has – in terms of the 2009 Interpol report – not promulgated any legislation with regard to its Forensic DNA Database and also indicated in the 2009 Interpol survey that no legislation in this regard is planned for the future. New Zealand, in contrast, promulgated specific legislation prior to the establishment of its database in 1995. This absence of any regulatory guidelines and/or legislation contributes to the scant information that is available on China’s national Forensic DNA Database.

What is known about China’s national Forensic DNA Database today, is that it quickly grew to be the largest Forensic DNA database in the world. At the time of the 2009 Interpol survey the database included a total of 1 200 600 profiles from the profiling categories as indicated in Table 2 below, and was ranked third in terms of total DNA profiles after the databases of the United States of America and the United Kingdom.

<table>
<thead>
<tr>
<th>CHINA: DNA PROFILE CATEGORIES/INDEXES</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Crime Scene Index</td>
<td>126 000</td>
</tr>
<tr>
<td>Convicted Offender Index</td>
<td>918 000</td>
</tr>
<tr>
<td>Suspect / Accused / Arrestee Index</td>
<td>147 000</td>
</tr>
<tr>
<td>Unknown Deceased Index</td>
<td>9600</td>
</tr>
<tr>
<td>Total</td>
<td>1 200 600</td>
</tr>
</tbody>
</table>

Table 2

While China’s Forensic DNA Database did not include a ‘Missing Persons’ index in 2009, it indicated that such a database is underway, and that it also intended on including a ‘Victim/Volunteer’ index. More recent sources have reported that an anti-trafficking DNA database was indeed established separate to the Forensic DNA Database and that it includes the DNA profiles of missing and abducted children and their family members. By 31 March 2012, China’s database contained more than 13 million DNA profiles making it the largest Forensic DNA Database in the world.

---

56. Id.
57. Id.
58. Id.
59. Interpol DNA Unit, supra note 4, at 31.
60. Id.
B. Hong Kong (HKSAR)

Hong Kong, a Special Administrative Region of China, started using DNA profiling in its criminal justice system in 1992, and a regional Forensic DNA Database using CODIS software became operational in 2001. Similar to New Zealand, Hong Kong first introduced specific DNA database legislation prior to establishing its Forensic DNA database; the Dangerous Drugs, Independent Commission Against Corruption and Police Force (Amendment) Ordinance (Cap. 68) of 2000 - Short title: Independent Commission Against Corruption Ordinance - came into operation on 1 July 2001 and amended the Dangerous Drugs Ordinance (Cap. 134), the Independent Commission Against Corruption Ordinance (Cap. 204), and the Police Force Ordinance (Cap. 232) to provide for the taking of intimate and non-intimate samples from certain categories of persons and matters related thereto. This can be compared to Japan where legislation was only introduced after it had established and started using its Forensic DNA Database. Malaysia is also following this route while it seems as though India will first promulgate the necessary regulatory and oversight legislation before establishing its national Forensic DNA Database.

Section 59G of the Hong Kong Police Force Ordinance (Cap. 232) provides that a DNA database be established and managed under the authority of the Government Chemist and that DNA information derived from the forensic analysis of intimate and non-intimate samples taken pursuant to the provisions of the Ordinance be stored on the database. Intimate samples are defined in section 3 of the Ordinance as samples of blood, semen or any other tissue or fluid, including urine or hair (other than dead hair), as well as a dental impression and a swab taken from a private part of a person's body or from a person's body orifice other than the mouth. A non-intimate sample is defined in the same section as dead hair, biological material taken from a nail or under a nail, a swab taken from any part other than a private part of a person's body or from the mouth but not any other body orifice, saliva, or an impression of any part of the body other than an impression from a private part or the person's face or an impression from identifying particulars like a photograph, fingerprint, palm-print, sole-print, toe-print and the weight and height measurements of a person. DNA information obtained from such intimate and non-intimate samples may subsequently only be included and be retained indefinitely on the database where

63. Interpol DNA Unit, supra note 4, at 27, 32.
65. Interpol DNA Unit, supra note 4, at 29.
67. Id. § 3.
68. Id.
the persons from whom the samples were taken had been convicted of a serious
arrestable offence, or if the samples were given voluntarily for this purpose, i.e.
for inclusion on the DNA database.69 A serious arrestable offence is defined as
an offence for which a person may under, or by virtue of any law, be sentenced
to a term of imprisonment of no less than seven years, or any other offence as
listed in Schedule 2 of the Ordinance.70 In contrast to the retrospective action of
the comparable South African legislation, the Hong Kong provisions do not work
retrospectively and the inclusion of DNA profiles from offenders convicted prior
to the commencement of the Ordinance is therefore not possible.71

The management and oversight of the DNA database is shared by two
departments, the Hong Kong Police Force and the Government Laboratory
(Department of Health).72 The Hong Kong Commissioner of Police is specifically
tasked with the management and oversight of the DNA database and the
Commissioner must, for example, ensure that DNA samples and any related
information subsequent to the forensic analysis thereof be destroyed if the person
from whom the sample was taken is acquitted, discharged or if a conviction is
subsequently quashed.73 DNA samples must furthermore be destroyed as soon as
the forensic analysis has been completed and the extracted DNA profile – where
authorized - been included on the DNA database. This mandatory destruction of
DNA samples applies irrespective of the outcome of the criminal charges or
investigation against the person from whom the intimate or non-intimate sample
was taken.74 And, section 59D of the Police Force Ordinance (Cap. 232) places
strict limitations on what the DNA samples and results of the forensic analysis
may be used for in order to safeguard the integrity of the database and the genetic
privacy of the individuals from whom samples had been taken.75

At the time of the 2009 Interpol Survey the Forensic DNA Database of Hong
Kong included a total of 20 396 DNA profiles covering the various indexes as set
out below.76 It was also indicated that Hong Kong is working towards including
an index for missing persons and will consider conducting DNA post-conviction
testing in future.77

69. Id. § 59G(1).
70. Schedule 2 of the Ordinance includes the following offences: criminal intimidation,
assaults with the intent to cause certain acts to be done or omitted, procurement by false pretenses,
intercourse with a girl under the age of 16 years and procurement of a girl under the age of 21
years.
71. Ordinance, supra note 66, § 59E.
72. See id. §§ 59A-59I.
73. Id. § 59H.
74. Id. § 59H(6).
75. Id. § 59D.
76. Interpol DNA Unit, supra note 4, at 27.
77. Id.
To date, no ‘missing person’ index has been included in the Hong Kong Forensic DNA Database, nor has any post-conviction testing been conducted, but the database size has grown to an estimated 49,466 profiles in 2015.  

IV. DATABASE DESIGN: THE RELEVANCE OF SIZE AND COLLABORATION FOR THE FORENSIC DNA DATABASES OF HONG KONG AND CHINA

The value and benefits of DNA profiling and Forensic DNA Databases are indisputable and much research has already been conducted in this regard. The discussion and comparison between the Forensic DNA Databases of China and Hong Kong will now continue with reference to two pertinent aspects: First, the question about database size will be considered. The Forensic DNA Database of China is currently the largest known database of its kind, while the number of DNA profiles on the Hong Kong database reflected a mere 0.6% of the city’s total population in 2013.  

With the current trend towards database-driven investigations and the seemingly logical impetus that comparative searches on larger databases have a higher likelihood of producing a match, the question may rightly be asked whether Hong Kong should not also expand its Forensic DNA Database. This question will be considered with reference to the UK NDNAD, the USA CODIS databank, as well as two judgments by the European Court of Human Rights and the United States Supreme Court. Second, the potential and the obstacles for cross-border collaboration, not only between Hong Kong and China, but also with regard to other jurisdictions’ databases will be considered.

A. Does Size Really Matter?

Initially, when the UK NDNAD was first established in 1995, only the DNA profiles collected from offenders convicted of serious violent crimes such as murder, rape or serious assault were included on the database. This status quo

78. Id.
80. Id.
81. Liz Campbell, Non-Conviction DNA Databases in the United States and England:
was, however, altered with the incorporation of section 82 of the Criminal Justice and Police Act of 2001 which provided for the collection of DNA (and fingerprints) from any person suspected of having committed an offence. While section 64 of the Police and Criminal Evidence Act 1984 (PACE) previously required that fingerprints, samples and profiles be destroyed if the arrestee is ultimately acquitted of the offence, no comparable requirement was included in section 82(3) of the 2001 Act. 82 This ultimately led to the considerable expansion of the UK NDNAD and by end March 2006 almost 4 million DNA profiles had been loaded onto the database, accounting for 5.2% of the UK population. 83 By 31 March 2014 this number had increased to almost 6 million individual DNA profiles and more than 450 000 unknown DNA profiles collected from crime scenes. 84 Furthermore, it was estimated in the 2013/2014 National DNA Databank Strategy Board Annual Report that when a DNA profile found at a crime scene is searched against the NDNAD, there was a 61.9% chance that the database would produce a match. 85 The data in the report further reflected that the UK NDNAC produced 24,953 crime-person matches and 1410 crime-crime matches based on routine loading alone. 86

The more than 6 million DNA profiles included in the UK NDNAD fade in comparison to the more than 12 million DNA profiles of the world’s second largest DNA Database, the national (federal) CODIS databank of the United States of America. 87 The expansion of the federal and state DNA databases in the USA was also a gradual process. According to Rothstein and Talbott, state DNA databases in the USA began almost exclusively as collections of adult sexual offenders’ DNA profiles but soon expanded to include the DNA profiles of “all convicted felons, juvenile offenders, those convicted of certain misdemeanors, and even arrestees.” 88 With regard to the federal database, the DNA Fingerprinting Act of 2005 authorized the expansion of this database to include DNA profiles collected from “individuals arrested, and from non-United States persons who are detained under the authority of the United States.” 89 Similar to the situation in the United Kingdom, the primary impetus behind the expansion of the state and federal DNA databases in the USA is based on the prevailing

83. Graham, supra note 3, at 286.
85. Id. at 10.
86. Id.
87. Interpol DNA Unit, supra note 4, at 21.
89. Campbell, supra note 81, at 289-90; Rothstein & Talbott, supra note 88, at 153-54.
view that the greater the number of DNA profiles included in the database, the higher the likelihood of a match.\textsuperscript{90}

Curiously, the legal trajectories of the DNA Databases of the United Kingdom and the United States of America have since developed in opposite directions.\textsuperscript{91} The extensive expansion of the UK NDNAD came under scrutiny as early as 2004,\textsuperscript{92} and came to head in 2008 when the European Court of Human Rights in \textit{S & Marper v. United Kingdom} found that such unfettered power to collect and retain forensic evidence, particularly DNA samples and profiles, was intrusive and failed to strike a fair balance between competing public and private interests.\textsuperscript{93} The UK parliament subsequently passed the Protection of Freedoms Act (PoFA) in 2012 and the Anti-Social Behaviour, Crime and Policing Act (ASBCPA) in 2014 to bring the regulations with regard to the collection and retention of DNA samples and DNA profiles on the UK NDNAC in line with the judgment by the European Court of Human Rights.\textsuperscript{94} A total of 7,753,000 DNA samples have since been destroyed and a total of 1,766,000 DNA profiles have subsequently been deleted from the database.\textsuperscript{95}

While the justices of the Grand Chamber of the European Court of Human Rights in \textit{S & Marper} emphasized privacy concerns in terms of the retention of certain categories of DNA profiles\textsuperscript{96} and were vigilant of how future technological advances could possibly affect the private-life interests of those individuals whose DNA profiles are included on databases,\textsuperscript{97} the U.S. Supreme Court in \textit{Maryland v. King} emphasized government’s legitimate interest in, and the extraordinary utility of DNA profiles and comparative searches, on DNA databases like CODIS.\textsuperscript{98} The facts of these two cases presented the respective courts with fundamentally different legal questions. The \textit{Marper} case essentially dealt with the retention of DNA profiles on the UK NDNAC where the donors of the profiles have been acquitted or the charges have been dropped, while the \textit{Maryland} case dealt with the legitimate government and public interest in taking a DNA sample in the form of a buccal swamp from certain categories of persons.

\begin{itemize}
  \item \textsuperscript{91} Campbell, \textit{supra} note 81, at 283.
  \item \textsuperscript{94} UK Home Office, \textit{supra} note 84, at 21.
  \item \textit{Id}.
  \item \textit{Id}.
  \item \textsuperscript{97} See Maryland v. King, 569 U.S. 435 (2013).
\end{itemize}
and for the DNA profiles so extracted to be used in speculative searches. However, the judgments of these two cases reflect the competing interests in the debate on the expansion of existing DNA databases. On the one hand, proponents of expansion emphasize the utility of DNA profiling and the potential of increased efficacy that an expanded DNA database holds, whereas the opponents show concern for individuals’ civil liberties and particularly their privacy interests with regard to their genetic material and information. Moreover, an important aspect that was largely ignored in the Maryland judgment is that in most states in the USA, the burden to destroy a DNA sample and expunge a DNA profile from a database is placed on the arrestee, and the majority of these states do not require that arrestees be informed of state expungement policies.99 Thus, “even if procedures exist to ensure that the profiles of some people should not be in a DNA database permanently, the practical hurdles set up by arrestee-initiated expungement likely mean that they are.”100

It is submitted that the weighing of these competing and equally important interests in the DNA Database expansion debate present a false choice as it overlooks the real issue: Is there really a direct correlation between the size of a DNA database and the likelihood of a match? If this premise is found to be false, the arbitrary expansion of a DNA database, and the tally of DNA profiles included in a database for the purpose of speculative searches will be of little real value in solving crimes. Mark Rothstein and Meghan Talbott unequivocally dismiss the suggestion that a larger Forensic DNA database will necessarily be more effective in producing “hits” and in curbing crime and criminality. They state:

[T]here is virtually no scientific, comprehensive, independent, peer-reviewed analysis quantifying the overall effectiveness of DNA databases in solving or preventing crimes. The only quantitative measure used to assess the value of DNA databases is the total number of “cold hits” or “investigations aided.” These totals make for good headlines and legislative testimony, but their use raises a number of serious methodological and policy concerns. To begin with, there is no clear definition of the terms “cold hits” or “investigations aided,” and thus the inclusion criteria vary widely among jurisdictions or even individual reports. There is also no comparative information available to estimate the likelihood that other forensic techniques or additional investigation would have identified the suspect. It is also not clear how many of the “investigations aided” actually result in conviction…[And] [a] match between a crime scene DNA and an individual’s database profile does not necessarily mean that the individual is guilty.101

100. Id.; Erin Murphy, License, registration, cheek swab: DNA testing and the divided court, 127 HARV. L. REV. 161, 172 (2013).
101. Rothstein & Talbott, supra note 88, at 154-55; see generally Carole I. McCartney, Tim
Rothstein and Talbott therefore caution against over-enthusiasm fed by media headlines: To this date, quantitative measures on the value of DNA databases remain sparse. Moreover, empirical evidence suggests that the utility and value of a DNA database do not depend on the total tally of DNA profiles included in the database, but rather on the type of DNA profile included. For example, an increase in Crime Scene Index profiles have been found to contribute to a higher likelihood of “hits” compared to an increase in the DNA profiles of offenders or arrestees.\(^{102}\) Ironically, however, and “despite popular assumptions, only a small percentage of crime scenes are checked for evidence that could be matched against a DNA database.”\(^{103}\)

The expansive UK NDNA\(^{D}\) is a case in point: While many countries keep a record of the number of times that a comparative search of a DNA profile from a crime scene “matches” a DNA profile stored on a database, only the United Kingdom also keep a record of DNA detections, which refer to cases where the database “match” ultimately led to a prosecution.\(^ {104}\) Recording so-called DNA detections in this manner can be particularly significant as not all “matches” necessarily identify the perpetrator of a crime and lead to a prosecution. Often, DNA “hits” in a comparative search can be due to a “match” with the victim’s DNA profile or with that of an innocent passer-by.\(^ {105}\) Thus, an analysis of the records kept with regard to the UK NDNA shows that despite, and throughout the period of the extensive expansion of the UK NDNA, the proportion of recorded crimes involving DNA detections remained roughly constant at 0.36% (ibid).\(^ {106}\) The wide-ranging expansion of the UK NDNA did not, therefore, increase the efficacy of the database, as many proponents for the expansion of Forensic DNA Databases argue.

The majority decision of the US Supreme Court in *Maryland* can, in light of this, rightly be criticized for the “almost giddy enthusiasm for scientific achievement” and the “rosy belief in the infallibility of the ‘identification sciences.’”\(^ {107}\) As was shown above, the arbitrary expansion of Forensic DNA

---


104. Wallace, Jackson, Gruber & Thibedeeu, *supra* note 103, at 58.

105. Kiely, *supra* note 37, at 474

106. Wallace, Jackson, Gruber & Thibedeeu, *supra* note 103, at 58.

107. Murphy, *supra* note 100, at 191.
Databases to include the DNA profiles of even more categories of persons under the rational that this will increase the efficacy of the database is indeed not scientifically sound and the weighing of this false assumption against competing public and private interests is therefore not only false, but also immaterial to an argument in favour of DNA database expansion. In addition to it being an unsound premise, the arbitrary expansion of a Forensic DNA database can also pose many practical difficulties. For example, the expansion of DNA Databases will place further strain on already burdened forensic laboratories and the limited resources available for law enforcement while inevitably also increasing the “potential for human error, mistaken identification, and wrongful conviction[s].”108 In this regard, the recently promulgated Criminal Law (Forensic Procedures) Amendment Act 37 of 2013 of South Africa has been criticized for the expansive provision that it makes for the collection of DNA samples upon arrest.109 This has not only raised human rights concerns, but it has also been questioned whether this is indeed the best way in which to use the already overburdened and limited available law enforcement resources in South Africa.110

In considering the relevance of these developments for the Forensic DNA Databases of China and Hong Kong, it is evident that the impetus behind the extended Forensic DNA Database of China, which is today regarded as the largest in the world, and the lack of any legislative framework or regulatory guidelines are questionable. While the more conservative Hong Kong database with its stricter regulatory framework that makes provision for the collection of forensic samples under specific circumstances and for the retention of extracted profiles on databases limited to specific circumstances, time frames, and/or crimes, is better aligned with the Marper judgment of the European Court of Human Rights, and can be said to better protect and serve both private and public interests.

B. Cross-border Collaboration

While the methods and technology used for DNA profiling today are generally the same in most jurisdictions, the loci selected by analysts for DNA typing, and the DNA profiles subsequently included on Forensic DNA Databases differ widely. For example, when the UK NDNAD was first established, it only included DNA profiles comprising of six STR markers (or loci).111 A further four loci were eventually added, and the UK NDNAD now requires DNA profiles consisting of ten loci for upload, analysis and comparison.112 The USA CODIS system requires a total of 13 loci, and with regard to the BRICS countries the following information is available:

108. Joh, supra note 99, at 4; Wallace, Jackson, Gruber & Thibedeau, supra note 103, at 58.
110. Id.
111. Jobling & Gill, supra note 26 at 743.
112. De Wet, Oosthuizen & Visser, supra note 6 at 172; Jobling & Gill, supra note 26 at 743.
<table>
<thead>
<tr>
<th>Country</th>
<th>Marker System</th>
<th>Number of loci</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>Russia</td>
<td>Pp16 or PowerPlex® System</td>
<td>15 loci plus Amelogenin (Amelogenin, also referred to as the AMEL locus or AMELX is used to determined gender.)</td>
</tr>
<tr>
<td>India</td>
<td>Ilil or Identifier® PCR amplification kit</td>
<td>15 loci plus Amelogenin</td>
</tr>
<tr>
<td>China</td>
<td>DNA Type 15, a multiplex kit developed by China (This kit has not been commercialized.)</td>
<td>14 loci plus Amelogenin</td>
</tr>
<tr>
<td>Hong Kong (HKSAR)</td>
<td>MSL, YNH24, PB30, TBQ7, Profiler (Pro or The AmpF/STR® Profiler® PCR Amplification kit amplifies nine loci), Profiler Plus (Pro® or AmpF/STR® Profiler Plus® PCR Amplification kit amplifies nine tetranucleotide STR loci), Identifier (the same as India), and CODIS DNA database software</td>
<td>More than 15 loci including Amelogenin</td>
</tr>
<tr>
<td>South Africa</td>
<td>Pro®</td>
<td>9 loci plus Amelogenin</td>
</tr>
</tbody>
</table>

Table 4

A further breakdown shows that in addition to the different number of loci used for DNA typing in these jurisdictions, the specific loci selected for the DNA analysis in each jurisdiction also differ. With regard to China and Hong Kong, it was reported in the 2009 Global DNA Profiling Survey conducted by Interpol that Hong Kong uses different markers compared to China and therefore does not contribute DNA profiles to China’s database.\textsuperscript{114} Compared to the USA CODIS database, China’s DNA profiles only include 11 of the 13 CODIS markers/loci.\textsuperscript{115}

\textsuperscript{113} Interpol DNA Unit, supra note 4.
\textsuperscript{114} Id. at 27.
\textsuperscript{115} Id. at 27, 32-33.
Given that Hong Kong indicated the use of various marker systems, it is unclear with how many of the CODIS markers the profiles stored on the Hong Kong database correspond. In contrast to this incongruence which hampers cross-border collaboration, South Korea has signed a bilateral agreement with the United States and New Zealand and is currently discussing the possibility of a bilateral agreement with the United States for cross-border collaboration in the sharing of DNA database data. The differences in the number and choice of loci of the DNA profiles stored on national Forensic DNA Databases therefore impact on the potential compatibility and sharing of information on a global scale. This is particularly important given the globalized nature of criminality and the potential role that Forensic DNA Databases can play in curbing transnational crime.

Cross-border collaboration and exchange of DNA information have been an ongoing endeavor amongst European Union member states. The Prüm Treaty (7 July 2005) and its implementing provisions (23 June 2008) provides for the automated exchange of DNA, fingerprints, and vehicle registration data, as well as other forms of police cooperation between the 27 EU States and specifically requires of member states to search for and share DNA matches automatically. By mid-2014, twenty of the 27 states had already implemented the agreement: Austria, Germany, Slovenia, Luxembourg, Spain, Finland, France, Bulgaria, Slovakia, Romania, Latvia, Lithuania, Hungary, Poland, Cyprus, Estonia, Sweden, Czech Republic, Malta, and Belgium. Another example of such cross-border collaboration is the envisaged integrated European DNA Database proposed by the European Network of Forensic Science Institutes (ENFSI) and the European DNA Profiling Group (EDNAP). While the scientists involved in this latter project recognize the importance of maintaining the various current national databases across Europe and the loci of choice of each country, they are also calling for expanded market sets to allow for greater statistical match probabilities to be generated through international searches. This, they argue, will only be possible if more common loci – they are advocating for 13 common loci in addition to the existing markers – are included in each country’s newly generated DNA profiles.

But there are also ethical concerns with regard to the cross-border collaboration and sharing of DNA profile matches. The United Kingdom, for example, specifically opted out of the Prüm Treaty due to “concerns about likely

116. Id.
117. Interpol DNA Unit, supra note 4, at 33; Wallace, Jackson, Gruber & Thibedeau, supra note 103, at 61.
118. Graham, supra note 3, at 286.
120. Graham, supra note 3, at 287.
121. Id.
122. Id.
large numbers of adventitious matches between individuals” DNA profiles held on the unusually large UK National DNA Database and crime scene DNA profiles stored in other countries.\textsuperscript{123} McCartney and her colleagues describe this concern in terms of the varying levels of national investment and organizational efficiency, influencing countries’ capacity to exchange, and the quality of the DNA information to be exchanged.\textsuperscript{124} These imbalances are compounded in the sphere of international cooperation and while this variability on the quantity and quality of DNA information available for cross-border sharing may at first seem to be only a technical or economical issue, it ultimately gives rise to questions of legitimacy wherein the citizens of one jurisdiction may be subjected “to a far greater extent to the full surveillance capabilities of DNA databases and their dis/empowering effects.”\textsuperscript{125} The validity of such shared DNA information can also come into question when the same quality controls with regard to the collection, analysis, retention, and management of DNA profiles are not complied with.\textsuperscript{126} Finally, the differences in domestic legislation with regard to data protection and the manner in which DNA evidence is dealt with in criminal justice systems of different jurisdictions may also hamper cross-border collaboration. It is important to note that with regard to the Prüm Treaty referred to above, no personal data is initially transferred between countries after a DNA match has been found. This is only done later in the process by way of “mutual legal assistance channels.”\textsuperscript{127}

While cross-border collaboration and sharing of DNA information for criminal justice purposes can therefore be a powerful tool in a globalized era of transnational criminality, the technical variabilities in database design, including the variations in national forensic bioinformation collection and retention policies, may pose an insurmountable obstacle for true and effective collaboration. This is particularly true of the Forensic DNA Databases of Hong Kong and China, where the differences in database design, criminal justice processes, and regulatory and guiding legislation seem too disparate in fostering true and effective collaboration.

V. CONCLUSION

Forensic DNA Databases are important investigative tools in criminal justice systems. It not only offers the convenience of centralized and computerized storage of DNA profiles, but it also enables the systematic comparison and automated matching of DNA profiles obtained from a wide variety of sources.\textsuperscript{128}

\textsuperscript{123} Wallace, Jackson, Gruber & Thibedeeau, supra note 103, at 61.
\textsuperscript{124} McCartney, Wilson & Williams, supra note 101, at 307.
\textsuperscript{125} Id.
\textsuperscript{126} Id. at 310.
\textsuperscript{127} Id. at 316.
\textsuperscript{128} Filipe Santos, Forensic DNA databases in European countries: Is size linked to performance?, 9:12 LIFE SCI., SOC’Y & POL’Y 1, 2 (2013).
Forensic DNA Databases have, for example, been used to link crime scenes with one another, link potential suspects to particular crimes, positively identify missing or unknown (deceased) persons, and to clear suspects of and exonerate convicted persons from crimes that they did not commit.

Yet, while the advantages in the use of Forensic DNA Databases for criminal justice purposes are undeniable, it is also true that scientific (forensic) advances are often adopted in law faster than social scientists and policy-makers can consider the consequences.129 Valueable comparative lessons can therefore be learnt; comparative lessons that go beyond the mere technicalities of DNA typing and database design, but reflect on the more fundamental approaches to evidence and guarantees underlying criminal justice practice. In this article, the standard rudiments underpinning the application of DNA technology in the realm of law enforcement and criminal justice was considered. But, the great variability in the design and utility of Forensic DNA Databases were also emphasized by way of a comparative analysis of a number of well-established and developing databases from different jurisdictions. The Forensic DNA Databases of China and Hong Kong were furthermore considered in relation to these other databases in terms of two pivotal database design features, namely database size and possible cross-border collaboration. It was evident from this analysis and discussion that much can be learned from such a comparative perspective, and that we should be sensitive to related developments in other parts of the world in calibrating the present functioning and future development of our own Forensic DNA Databases. While such mutual gain in terms of knowledge, know-how, and best practice hold the potential of a more standardized and universal approach for closer collaboration, it also lays bare the stark technical, procedural, and ideological national differences that exist with regard to the role that DNA typing and particularly Forensic DNA Databases (should) play in criminal justice systems.

With regard to the Forensic DNA Databases of Hong Kong and China, it was evident that stark differences exist, including with regard to the DNA profile indexes and marker systems used, the number of loci selected for DNA analysis, as well as the size of the two databases, which also implies that differences exist with regard to the procedures and regulations governing the collection and retention of DNA samples and profiles. These stark differences do not only pose practical obstacles for possible future collaboration, but also reflect more fundamental differences in criminal justice practice. While it is possible for technical harmonization to be achieved in terms of database design between the Hong Kong and China databases and even between the databases of the BRICS countries considered in this article, national differences with regard to the collection and retention of DNA samples and DNA profiles, that reflect the fundamental approaches to evidence and guarantees underlying the various criminal justice systems, will continue to challenge absolute standardization and effective collaboration.