

Family Ties: The Use of DNA Offender Databases to Catch Offenders' Kin

Henry T. Greely, Daniel P. Riordan, Nanibaa' A. Garrison, and Joanna L. Mountain

“The sins of the fathers are to be laid upon the children.”¹

Just after midnight on March 21, 2003, a drunk stood on a footbridge over a motorway in a village in Surrey in southern England.² After eight pints of beer, he was drunk enough to decide to drop a brick from the overpass into traffic to see if he could hit something; unfortunately, he was not so drunk that he missed. The brick crashed through the windshield on the driver's side of a truck. It hit the driver, Michael Little, in the chest, triggering a fatal heart attack. He stayed conscious long enough to pull the truck safely to the side of the road, thereby perhaps saving other motorists; then he died. The crime was widely publicized, as was the driver's role in preventing any further accidents.

The police had no suspects, but they did have a clue – the brick had on it a mixture of DNA from the victim and someone else, presumably the perpetrator. The police also had blood from a nearby car that had been broken into that evening. The DNA from that blood matched the DNA on the brick. The police analyzed the DNA and compared it to their British DNA database, but found no match. Interviews began in the village, and voluntary DNA samples were taken from over 350 young men in the area, but without success. A £25,000 reward for information was offered, but nothing useful appeared. The police were eager to solve this crime, but, after six months, they had no suspect.

So the British police decided to check the DNA database for less than perfect matches in the hope that the perpetrator had a relative in the database. They set the search to pull up any offender in their database who matched at least eleven of the twenty DNA markers used by the British system. At first, they found too many matches to investigate. But after restricting the search to young white males from Surrey and Hampshire, two counties near the crime scene that are home to about 2.6 million people, they found about twenty-five partial matches, one of which matched on sixteen of the twenty DNA markers. The police interviewed

Henry T. Greely, J.D., is the Deane F. and Kate Edelman Johnson Professor of Law, and a professor, by courtesy, of genetics, at Stanford University, where he directs the Center for Law and the Biosciences and the Program on Stem Cells in Society. **Daniel P. Riordan** is a doctoral candidate in the Department of Genetics at the Stanford University School of Medicine. **Nanibaa' A. Garrison** is a doctoral candidate in the Department of Genetics at the Stanford University School of Medicine. **Joanna L. Mountain, Ph.D.**, is Assistant Professor in the Departments of Anthropological Sciences and Genetics at Stanford University, where she investigates the origins and implications of human genetic variation.

the person with the closest match and discovered he had a twenty-year-old brother living near the village where the crime had occurred. The brother, Craig Harman, denied involvement, but did agree to give a DNA sample.³ His DNA matched the DNA isolated from the blood on the brick; when confronted with the DNA match, Harmon confessed. In April 2004 he started serving a six year prison term for manslaughter.

Genes run in families. If you have information about one person's genome, you know something about the likely composition of the genomes of his or her biological relatives. This fact is now beginning to be used in criminal investigations. As of 2004, the British had used this method about twenty times, gaining valuable information about a quarter of the time.⁴ One newspaper article reported that familial searching had, through April 2005, solved nine cases in the United Kingdom.⁵ At least three of the successes, in addition to the Harman case, have been specifically discussed in the press. One case tied three 1973 murders to a suspect who had subsequently died; a second led to a confession in a 1988 murder.⁶ A third success led to a conviction in a 2002 rape-murder.⁷

In the United States, this method has been used successfully at least once. Willard Brown was convicted of a rape-murder from twenty years earlier after crime scene DNA provided a partial match to the DNA profile of his brother, Anthony Brown, which was in the North Carolina database. The same DNA test exonerated a man who had spent eighteen years in prison for the crime.⁸ Another variant of family forensic DNA, not using an Offender Index, seems to have helped Kansas police catch the BTK serial killer. After the police eventually developed a suspect, they analyzed DNA of his daughter, taken from a medical clinic, to confirm that the crime scene DNA could have been from one of her parents.⁹

The legal and policy implications of this kind of "family forensic DNA," which the British call "familial searching," have been discussed only rarely and briefly.¹⁰ This article provides both a fuller explanation of the science and technology behind such uses, and a different analysis of their policy implications. The article first sets out the background of forensic use of DNA, both how it works scientifically and the structure of existing DNA databases. It then discusses the scientific plausibility of using DNA to identify crime scene evidence as coming from a relative of a convict (or now, in the federal system and in some states, merely an arrested person) in a DNA database, both now and with some plausible "improvements." It ends by examining the implications of this technique, both generally and in light of the disproportionate number of African-Americans in U.S. forensic databases. We do not view

this article as providing definitive answers to the issues it raises, but we hope it will help start an informed discussion that can lead to useful policies concerning this technique.

Forensic DNA and DNA Databases

The Science of Forensic DNA Identification

The average human adult is made up of about fifty to one hundred trillion human cells, almost all of which contain, in their nuclei, forty-six chromosomes made up of deoxyribonucleic acid (DNA) wrapped around a protein backbone. Those chromosomes – pairs numbered one through twenty-two, plus either two X chromosomes (in most women) or one X and one Y (in most men) – comprise each human's "genome," the total of all that person's genetic information, apart from a very small amount of DNA in the cells' mitochondria. Each human's genome is made up of two copies of a variant of *the* human genome, one from his mother, one from his father. Each of these parental genomes, consisting of one each of the twenty-two pairs of chromosomes, plus the X or Y chromosome, is made up of about 3.2 billion "base pairs" of DNA. Each base pair makes up one "rung" of the double helix that is the DNA molecule, and comprises either an adenine connected to a thymine, a cytosine connected to a guanine, a guanine connected to a cytosine, or a thymine connected to an adenine: A-T, C-G, G-C, or T-A. The order of the A's, C's, G's, and T's is the genetic "sequence."

About three percent of the genome contains DNA instructions for making ("coding for") a particular version of ribonucleic acid (RNA), which in turn either provides the instructions for the cell to make a particular protein or has an independent function of its own. About four percent of the genome is not known to code for RNA but seems, from the fact that very similar genetic sequences are found in many species, to perform some important, but thus far unknown, function. The three percent that codes for RNA and probably much of the four percent with some other function make up the 20,000 to 25,000 "genes" that humans have. Some of the remaining ninety-three percent of our genomes appears to play structural roles in the physical activity of the chromosome; most of it, thus far at least, seems to have no function.

Nested within the vast majority of our genomes that have no known function are many different stretches of sequence called "microsatellites," or "short tandem repeats" ("STRs"). These are stretches of DNA where the DNA replicating mechanism appears to "stutter," resulting in different numbers of copies of repeated sequences. One common set of STRs involves repeated sequences of four bases – for example, the sequence ACGT. Some people inherited from one parent a stretch

of DNA with four repeats of ACGT; others inherited a stretch with six repeats, or one, or ten, or twenty-five. Each stretch with a different number of repeats is a different “allele.” The fact that these stretches of DNA have a different number of these repeats makes them useful as “markers.” Because their location on the chromosomes is known, they “mark” the location of genes that are nearby; because any individual will often have inherited a different length STR from his mother and his father, they can “mark” which chromosome came from which parent. Each of these STRs is found at one spot on one particular chromosome, a location known as a “locus.”

These tetramer repeats, as far as scientists know, have no function. They do not code for RNA, and they do not seem to be responsible for any difference in the structure or functioning of the people of who carry them. In other words, a person whose genome has two copies (one from each parent) of a marker, with twelve tetramer repeats, seems no different from a person with two copies of the marker with five repeats, or someone with one copy with seventeen repeats and another copy with three repeats.

These STRs can be used for identification. In the United States, crime laboratories typically use a set of thirteen STRs, known as the “CODIS markers,” named after the FBI’s Combined DNA Information System. These STRs are spread over twelve chromosomes.¹¹ Each individual has two copies of each of the thirteen STRs. On average, one of the CODIS markers has twelve different lengths, or alleles, found in significant numbers of the population, but the least variable CODIS marker has seven alleles and the most variable has twenty-three. One person might have two copies of the first marker that are four and eight repeats long, copies of the second that are eleven and twenty-three copies long, copies of the third that are three and ten copies long, and so on through all thirteen markers. That person – someone, possibly the perpetrator, who left DNA at a crime scene; someone who left DNA on some important evidence to a crime; or an unidentified person whose remains have been found – can thus be identified as thirteen pairs of numbers, one pair for each of the thirteen STRs. Those numbers constitute a “genotype” of the individual for those STRs (based on the alleles they have of those STRs).

The odds that an unrelated person shares the same set of thirteen pairs are normally infinitesimal – at most one in several hundred billion, compared with a total of 6.3 billion living humans. Two random Americans will share, on average, about two or three alleles. On the other hand, identical twins will share all thirteen pairs – and first degree relatives (parent, sibling, or child) on

average will share at least half. This much higher rate of sharing among relatives is the reason for this paper.

Forensic DNA Databases

Many nations and every American state have established forensic DNA databases. The United Kingdom has one of the oldest and largest ones. Since 1995, it has collected and analyzed DNA samples from all those convicted of felonies and, indeed, often from those arrested but not convicted. The British analyze the DNA at STRs found at ten different markers and put the analyzed genotype (the twenty numbers that represent the length of the STRs for each of the ten pairs) into a database. Their database now contains genotypes for over 2.5 million Britons, about five percent of the United Kingdom’s population, and, as most of the samples are from adult males, it contains this genetic information on about one tenth of the country’s men.

Every jurisdiction in the United States has established its own “offender” database.¹² States set out different requirements about who must provide DNA for these databases. In some states only convicted sex offenders or violent criminals must provide DNA; increasingly, states are requiring DNA from all felons. More than thirty states have now gone to “all felon” databases and it seems likely that all states will adopt at least that broad a database soon. Several states, including California, Texas, Virginia, and Louisiana, now require DNA samples from some people merely arrested for or charged with felonies, and recent federal law amendments have allowed the Attorney General “to collect DNA samples from individuals who are arrested or from non-United States persons who are detained under the authority of the United States.”¹³ Many states also collect and analyze DNA from unsolved crime scenes, from missing persons, and from unidentified human remains.

Each state may analyze the DNA for any specific markers, and any number of markers it chooses, but general uniformity has been assured by CODIS. The FBI will accept entries from states for its database only if the states submit markers from its specified set of thirteen markers. federal law encourages states to join the CODIS system and to submit qualifying genotypes to the federal database from their own records.¹⁴ The genotypes are accompanied by identifiers for the specimen analyzed and for the laboratory and the personnel who performed the analysis. The federal database, called the National DNA Index System (NDIS), is made up of entries submitted by the states and federal law enforcement agencies. Every state except Mississippi has submitted genotypes, sending identifying information plus the twenty-six identifying numbers (thirteen pairs of STR lengths) from its own

DNA collections, based on whichever set of criminals or suspects it has decided to require to submit DNA samples – and whichever among those it has collected that it has actually analyzed – to the FBI.

The NDIS actually comprises two different indexes as part of CODIS. The “Forensic Index” contains geno-

An important question in the early period of adoption of DNA identification was whether different populations in the United States had sufficiently distinctive genetic structures as to make the chance of a random match within that population much greater than the overall chance. This population genetics question was explored in some detail as a concern in the 1992 report from the National Academy of Sciences on forensic use of DNA.¹⁸ In part to answer this question, the FBI collected information about the frequencies of these alleles in samples of different ethnic populations drawn from the database. These worries were resolved by a subsequent 1996 National Academy report, which found that, at least with respect to the

An important question in the early period of adoption of DNA identification was whether different populations in the United States had sufficiently distinctive genetic structures as to make the chance of a random match within that population much greater than the overall chance.

thirteen CODIS markers, the frequency of different length STRs did not vary enough from one group to another to have a substantial effect on the probabilities of a match.¹⁹ But little attention has been paid to another kind of population that clearly has distinctive genetic structures – families.

thirteen CODIS markers, the frequency of different length STRs did not vary enough from one group to another to have a substantial effect on the probabilities of a match.¹⁹ But little attention has been paid to another kind of population that clearly has distinctive genetic structures – families.

The Scientific Basis of Family Forensic DNA

DNA runs in families. Two people who are closely related genetically are likely to share more alleles than two people who are not closely related. The patterns of these similarities depend, however, on the type of familial relationship. A normal search through the CODIS database, or equivalent foreign databases, looks for a perfect, or near-perfect match.²⁰ If the search shows no perfect, or near-perfect, matches, it is considered a failure. But a search of crime scene DNA that shows a partial match to the Offender Index may actually indicate, to a high degree of probability, that the crime scene DNA came from someone who was closely related to the offender. The Harman case provides one example. Two random Britons would be expected to match on six or seven of the twenty alleles used in the United Kingdom’s system. The partial search turned up about twenty-five white males in the relevant geographical region who matched DNA from the brick and the car at eleven or more alleles. That DNA matched Harman’s brother’s DNA at sixteen out of twenty sites, far more than one would expect at random.

types from crime scene evidence; the “Offender Index” (originally the “Convicted Offender” index) contains the genotypes of those convicted or charged with crimes that state or federal law enforcement groups have submitted.¹⁵ A law enforcement agency can, via computer, check the genotype of an analyzed DNA sample against either (or both) of these two indexes. It can check crime scene DNA against DNA collected from previous offenders (the Offender Index) or from other crimes (the Forensic Index). Alternatively, it can check a particular suspect’s genotype against the Forensic Index to look for his possible involvement in other crimes.

As of November 2005, the CODIS Offender Index included information about 2.75 million people, while the Forensic Index had about 125,000 entries.¹⁶ Throughout the United States, roughly one million people are convicted of a felony each year. Some of those will already be in the CODIS database; nonetheless, the possibilities exist for rapid expansion of this database.

When law enforcement officials find DNA at a crime scene that they believe may be from the perpetrator, they generally analyze it for the thirteen CODIS markers. They can then check the genotype found in crime-scene DNA against the Offender Indexes in their state DNA database or in CODIS. The FBI claims that, through November 2005, the CODIS system has assisted in nearly 30,000 criminal investigations by providing a DNA match that would not otherwise have been developed.¹⁷ Its power seems certain to increase as more DNA samples from offenders are taken, analyzed, and submitted to CODIS. The rate-limiting factor has been the speed of states in spending the money needed to analyze the DNA samples they receive from convicts. (The federal database largely depends on submissions of DNA analyzed by the states; less than ten percent of felons in the United States are convicted of federal felonies.)

Determining, however, whether a high match is the result of a genetic family relationship between the offender in the database and whoever left the crime scene sample is not simple. It depends both on the nature of the postulated relationship and on the rarity of the genotype (set of alleles) involved.

First degree relatives share, on average, about fifty percent of each other’s DNA variants, including STR

lengths, by descent (as a result of their very recent shared ancestry). These are genetic parents, siblings, and children. Second degree relatives – uncles or aunts and nephews or nieces, grandparents and grandchildren, half-brothers and half-sisters – share one quarter of their DNA variations by descent; third degree relatives (first cousins or great-grandparents and great-grandchildren, among others) share one-eighth. First degree relationships are most likely to be useful for this investigative technique, but all first degree relationships are not the same.

While two unrelated people usually share only a few CODIS alleles, a genetic parent, say a father, and his child *must* match at no fewer than thirteen alleles, and are most likely to match at fourteen, fifteen, or sixteen

are a parent and child likely to have more total matching alleles than two unrelated people, but also that the way these genetic matches will occur between a parent and child is highly characteristic of that specific relationship – namely, every marker will have at least one of the two alleles in common, and relatively few markers will have more than one allele in common.

The unusual pattern of parent-child matches – that they *must* match at one allele at each marker – makes them a particularly useful kind of partial match. The FBI has published the frequency in the CODIS database of the different length variations for all thirteen of the markers it examines. This data allows one to calculate how likely it is that a genotype that matches a profile in the Offender Index at a certain number of alleles

comes from a relative of the person in the Offender Index, versus an unrelated person. On average, the chance that an unrelated person's genotype will match the genotype from crime scene DNA at thirteen or more of the twenty-six alleles, allowing for all possible ways of distributing the matches across markers, is around three percent. However, the chance that two unrelated people match at thirteen or more sites with every marker having at least one match (as will occur for

parent-child pairs) is about one in two thousand.²² Although these odds are low, with genotypes from 2.75 million people in the CODIS Offender Index, there should be many spurious matches at this level in the database. For an average genotype, around 2,000 to 3,000 people in the Offender Index are likely to have one or more matching alleles at all thirteen markers. For a rare genotype, the number will be much lower – perhaps none.

In fact, if we consider the rarest possible genotype in the Caucasian population – corresponding to someone who has two copies of the least frequent allele at every single marker – then the chance of an unrelated individual randomly matching that genotype at least once at every marker is around one in ten trillion quadrillion (10^{28}). Therefore, in this best-case scenario, a partial match at this level is considerably less likely to be spurious than is the typical perfect genotype match (which is on average around one in ten quadrillion [10^{-16}] for the Caucasian population). On the other hand, in an extreme worst-case scenario (corresponding to someone who has one copy of each of the two most frequent alleles at every single marker), the chance of randomly matching at least once at every marker is just under one percent. In this worst-case scenario, it is unlikely that a true relative could be reliably identified by familial searching since there would be so many spurious matches.²³

The cost of looking to see how many partial matches exist for a given crime scene sample should be quite small – a tiny amount of computer processor time and a moment's glance at a computer screen.

alleles. One of the child's two alleles at each of the thirteen CODIS markers came from the father; except for the unusual event of a mutation in one of those alleles in the sperm that was part of the child's conception, those thirteen alleles *must* be the same. In addition, by chance, the father may share with the child's mother some of the thirteen alleles that he did not pass on to the child. For example, if two alleles of an average CODIS marker in genetically unrelated people, such as the child's mother and father, are likely to be identical by chance fifteen percent of the time, then the child will likely get two alleles from his mother that match the father's genotype. That child would match the father at one of two alleles at eleven markers (where the matching allele came only from the father), and at both alleles at two markers, where one came from the father and the other one, which came from the mother, happened by chance to be the same as the father's second allele.

The most likely number of alleles shared between parent and child will vary from population to population because the extent to which unrelated individuals share alleles is somewhat different in different populations. Using the Caucasian population (for which good published data exists) as an example, a father and his genetic child will share, on average, 15.7 of the twenty-six CODIS alleles, whereas two completely unrelated Caucasian individuals will share on average 8.7 alleles.²¹ However, it is also important to recognize that not only

It is possible for two siblings to share anywhere from zero to all twenty-six markers, but on average they share around 16.7 alleles.²⁴ Thirteen of the shared alleles are expected to occur, on average, due to common inheritance of the same alleles from their parents, whereas the additional matches can occur either when the two parents share some alleles with each other, or when either parent has two copies of the same allele. Again, not only the number of overall matching markers, but also the *pattern* of how the matches are distributed across the genetic markers are characteristic of a relationship between two siblings – a few markers are expected to have no shared alleles, most have one allele in common, and a fair number of markers have two shared alleles. For the same Caucasian population, an average pair of siblings has about one marker with no common alleles, about seven with exactly one shared allele, and about five with two alleles in common.

Although there is no simple pattern of partial genotype matching that can perfectly distinguish all sibling pairs from spurious partial matches, the difference in genotype matching patterns between siblings and unrelated individuals does provide considerable information that can be used to successfully identify pairs of siblings some of the time. If a false positive rate of one in two thousand were tolerated (comparable to the false positive rate at which one hundred percent of parent-child pairs can be detected), then about sixty percent of true sibling pairs could be reliably identified. However, about twenty percent of true siblings could be detected at a level that would be expected to yield only one in 100,000 matches by random chance, such that only around twenty to thirty spurious matches would result from searching the Offender Index. Thus, for a substantial fraction of cases, partial genotype match patterns can be used to reliably identify pairs of siblings with false positive rates comparable to or even much lower than observed for parent-child pairs (with one hundred percent detection).

Usually, the partial match by itself will not be overwhelming evidence that the person who left the crime scene DNA *is* a relative of the person in the Offender Index who provided a partial match. It will usually be the case that such a partial match could be made to many of the world's 6.3 billion people who do not have a relative in the CODIS Offender Index. The partial match is only a lead – a relatively weak one for a common genotype though possibly a very strong one for a rare genotype.

How strong or weak the lead is can be estimated. One should be able to estimate how many people in the overall population, or perhaps in defined subpopulations, match any given genotype at thirteen specified sites. Ultimately, though, the partial match would only

need to function as a lead and not as evidence in court. If a suspect were identified as a result of the partial match, his DNA could then be taken and analyzed (voluntarily, through a search warrant, or after arrest) and compared to the crime scene DNA, leading to a conclusive match or non-match.

Algorithms could be created easily to look for both parent-child and sibling-sibling matches. The first are more distinctive, because at least one allele at each site must match, but, at least in the early years of a database's existence, the second are likely to be more useful. Most of those convicted of felonies are relatively young. According to data collected by the Department of Justice, forty-eight percent of those convicted of felonies in 2002 in state courts were under thirty years old. Another twenty-nine percent were thirty to thirty-nine years old.²⁵ It will take some years before the primarily young men in the Offender Index have many children in the prime age for committing crimes, nor are their parents likely often to be suspects. The siblings of a person in the Offender Index are likely to be in or near the age groups responsible for most criminal activity.

The cost of looking to see how many partial matches exist for a given crime scene sample should be quite small – a tiny amount of computer processor time and a moment's glance at a computer screen.²⁶ The sample will have been genotyped and submitted to state databases and CODIS in any event, looking for a perfect match. It would cost little or nothing to have the program also show all people in the Offender Index who match the crime scene genotype in a way that is characteristic of either a sibling-sibling match or a parent-child match, and also to report statistical confidence levels for each partial match.

The cost of *following-up* the leads generated by family forensic DNA may be extensive, involving interviewing many offenders and then finding and interviewing any of their relatives who could be possible suspects. Sometimes, the computerized search will reveal hundreds of matches at that level. Sometimes, it will reveal only fifty such matches. Sometimes it might reveal a handful – or only one. Certainly, police are less likely to use these leads if the genotype is common and there are hundreds of partial matches, and more likely to use them if the genotype is rare and there are only a few leads. On the other hand, the police are more likely to chase even unlikely leads in a high profile offense – a political assassination, a terrorist attack, or a notorious murder.

In light of the low cost of looking for partial matches, it seems plausible that police, in the event they do not get a perfect match from their state DNA database or from CODIS, will increasingly request information on partial matches. Depending on the number of partial

matches and on the circumstances of the crime, police can then decide whether or not to talk with any or all of the offenders identified as partial matches about their relatives to see whether any of those relatives could be suspects in the crime.

Improving the Efficiency of Family Forensic DNA

Under the current CODIS system in the United States, the tactic of investigating partial matches may often be inefficient, not because of the cost of the familial search, but because of the cost of investigating the relatives of all the partial matches identified in the Offender Index. That efficiency could be substantially improved in two very different ways.

First, information could be collected and stored about the relatives of people in the Offender Index. If, instead of interviewing the people in the Offender Index who were partial matches, the police could check a database to find out information about their parents, siblings, and children, the cost of this kind of investigation would be greatly reduced. The police could quickly determine whether the partial match in the Offender Index had a relative (often the police will be only or primarily interested in young male relatives) of the right age and geographical location to be a suspect in the case. There would, of course, be costs in collecting this information and putting it in a database. There may be questions, which this article does not explore, about whether an offender could be compelled to provide that information, which, after all, is not relevant to the crime for which *he* has been convicted or charged. And the offender might lie, or be mistaken, about his relatives, thus lowering the efficiency of this approach. (Of course, he might also lie in a specific interview with police.) Information from the offender might be checked before entry, but that step would increase the initial costs.

A second, more powerful approach is more technical – increase the number of genetic markers analyzed for forensic DNA genotyping. Using additional genetic markers would greatly increase the probability of identifying true genetic family relationships with confidence and would thereby eliminate spurious leads. The more markers analyzed, the lower the probability that a partial match is the result of chance rather than family ties. If a thousand STRs similar to the CODIS markers were analyzed for both offenders and crime scene DNA, the probability that an unrelated person would match an offender at fifty percent of those markers would be very small, but most first degree relatives

(and all genetic parents and children) will match at half of the markers.

Of course, analyzing an additional thousand STRs would both impose substantial costs and increase the concern that some STRs provided information about features relevant to the individual's body, health, or behavior. But, in fact, if roughly twenty more markers were added, similar to the thirteen CODIS markers in the number of variants and their frequencies, an average crime scene genotype that matched a genotype in the Offender Index at at least one allele at each of the thirty-three markers would be likely to come from that offender's relative, as it would be expected to occur only once in every 200 million unrelated individuals from the same population. In that case the lead provided by such a match should be quite powerful; there would

The legal and policy implications of family forensic DNA have largely been overlooked in the voluminous literature on forensic DNA.

be few, if any, spurious matches to waste an investigator's time. It should also be possible to collect enough marker information to make it likely that a crime scene sample came from a second or even third degree relative of an offender; we have not yet calculated how many CODIS-like markers would be necessary to make that determination.

Most states currently keep DNA samples indefinitely. Analyzing additional markers from a DNA sample that is already available is relatively inexpensive and is likely to become much cheaper in the future. Even if, for some reason, no samples from the 2.75 million people in the Offender Index can be found or used to analyze an additional fifteen markers, with roughly one million people each year being convicted of felonies, it would not take long to build up a very large database with the increased number of markers.

In addition to these two approaches, focusing on collecting family information or on analyzing more markers, other advances might improve the efficiency of family forensic DNA. More sophisticated scientific approaches that take into account both the identity of matching alleles and the identity of matching markers should improve the performance of familial searching strategies, compared to the simpler approach discussed here, which is based on the overall number of markers with zero, one, or two alleles shared between two genotypes.

Legal and Policy Implications of Family Forensic DNA

The legal and policy implications of family forensic DNA have largely been overlooked in the voluminous literature on forensic DNA. One set of discussions took place at a September 2004 workshop organized by the American Society of Law, Medicine & Ethics. The report from this workshop lists five fairly broad areas of consensus among the participants about the use of what they call “low stringency searches,” searches where an exact match is not required:

- Low stringency searches are an implicit database expansion that should be open to public debate.
- The investigations that follow low stringency database searches can involve asking non-suspects to reveal extremely intimate information.
- Low stringency searches and the subsequent investigations will amplify any racial or geographic bias already in the DNA database system.
- Clear standards and care are needed for choosing when and how to perform low stringency searches.
- The substantial costs to the criminal justice system and to the number of subjects of investigations must be weighed against the benefits of using low stringency searches.²⁷

There appear to be only two more detailed published discussions of these issues. One is from David Lazer and Michelle N. Meyer, in the concluding chapter of a book on the use of DNA in the criminal justice system.²⁸ In three pages, Lazer and Meyer set out the possibility of family forensic DNA and describe several early British uses of this technique. They raise several issues discussed in detail below, including the differential effect of such technology in the United States on African-Americans, and the consequences for the rights of the potentially implicated relatives of offenders.

The other discussion is from Robin Williams and Paul Johnson.²⁹ As one part of a wide-ranging article, Williams and Johnson discuss the United Kingdom’s experience with familial searching and point out several concerns about it. They do not, however, analyze what seems most troubling about this approach – the way it effectively increases police scrutiny and interest in people based on their relatives’ past involvement with the criminal justice system. This use raises concerns about fairness both in individual cases and, more broadly, in its differential effects on groups in American society. The first set of issues is disconcerting, but does not seem to disqualify the approach. The broader issues may make the use of this approach unacceptable.

This section of the paper first describes other objections Williams and Johnson raise. It then analyzes the legal and policy issues arising from this “suspicion by family association” aspect of family forensic DNA. It does not find any persuasive constitutional or policy arguments for prohibiting or greatly limiting the use of this technique as a general matter. It ends by examining the special problems in the United States raised by racial disparities in the CODIS Offender Index that lead to family forensic DNA having a disproportionate effect on African-Americans. This does raise substantial concerns about the technique.

Other Objections from Williams and Johnson

The other concerns raised by Williams and Johnson do not seem compelling. They first argue that familial searching may lead to knowledge of previously unknown genetic relationships, or to knowledge of the absence of what was previously believed to be a genetic relationship.³⁰ It may occur that someone an offender names as a parent or child will turn out, after genetic testing, to be demonstrably not genetically related. (Note that the possibility that two people are genetic siblings cannot be conclusively disproved by existing DNA analysis with CODIS markers without information from a third person who is related to at least one of the two; it is very unlikely, but not impossible, that genetic siblings have no CODIS alleles in common.)³¹ Unless that information is communicated to those people, the fact that someone involved in the criminal investigation knows this seems, at most, an invasion of privacy with no resulting harm. It is difficult to envision circumstances under which a relationship previously unknown to both parties would be discovered. If an offender names someone as a relative who does not suspect the relationship, DNA analysis could confirm that the relationship was possible or even likely, but, except for identical twins, it would not provide definite proof. And, again, unless the information is communicated to either party, there seems little to no harm.

Williams and Johnson then note that the use of familial searching with samples that were collected voluntarily may be inconsistent with promises of privacy or confidentiality given as part of the screening. Though true, this problem does not apply to samples, such as those in the FBI Offender Index, that are collected by force of law. They then express concern that family forensic DNA may exacerbate

pervasive problems associated with the confusion between “genetic” and “social” relatedness (“families” are not only constituted through genetic lines but through clusters of non-genetically related individuals) as well as the implicit assumption that

criminality is fostered because of such relatedness (either because of genetic or social reasons).³²

Although using familial DNA profiles to identify suspects has no logical connection with the idea that crime runs in either genetic or social families,³³ merely putting the words “crime,” “family,” and “DNA” in a sentence might conjure such an association for some readers. It is neither clear that such associations would have significant effects, nor that they could not be countered by appropriate disclaimers.

Next, Williams and Johnson posit that the police intervention with the “offender,” particularly an offender never convicted of a criminal offense, may invade that person’s privacy rights:

Such individuals will have to be approached by the police to name relatives whose own profiles are not on the NDNAD [National DNA Database]. It may well be claimed that this constitutes a disproportionate interference with their privacy rights under the European Convention. Even if the approach itself is licensed by their lessened right to privacy, which has resulted from their previous criminal arrest, a question is raised regarding their obligation to help the police with their inquiries in a case where their own DNA profile has already exonerated them from direct involvement.³⁴

Being compelled to assist the criminal justice system, in this or any other way, may be annoying, but is common. The argument that the police inappropriately invade a person’s privacy by asking, without any obligation, about his relatives, is very weak. Whether cooperation with the police in such circumstances should be considered truly “voluntary” might be doubted. (Of course, if one were concerned about the inconvenience or irritation to the “offender,” one could minimize the intrusions by collecting information about relatives only once and storing it for future use.)

The authors’ last objection does speak to this kind of suspicion based on family ties:

At the very least it seems likely that the use of this particular procedure will bring the police into contact with a number of individuals who have not been prosecuted for a recordable offence, who will have no criminal record, and *who are subject to interview only because they are genetically related to someone whose profile is on the NDNAD*³⁵ [emphasis added].

And to that concern we now turn.

Suspicion by Family Ties

Dan Krane, a DNA specialist at Wright State University in Dayton, Ohio, says familial searching “puts someone in jeopardy of investigation simply because his brother committed a crime...that’s the sins of the father being visited on the son...[it is] contrary to the whole idea of our criminal justice system.”³⁶

The “family suspicion” aspect of family forensic DNA clearly troubles people. It may well be responsible for the fact that this emerging technique is already being hedged round with limitations. In the United Kingdom, where it has received the most use, it can only be employed subject to apparently restrictive (but undisclosed) guidelines. Williams and Johnson report:

Discussions between ACPO [the Association of Chief Police Officers], the Home Office, the Information Commissioner, and representatives from the Human Genetics Commission, have resulted in an agreement about the circumstances under which such searches will be carried out and their results integrated into existing investigative procedures. However the agreement is operationally sensitive and has not been publicly disseminated.³⁷

Nothing similar seems to have emerged in the United States. A newspaper article reported in July 2005 that

The Scientific Working Group on DNA Analysis Methods, a group of 35 specialists who advise the federal government on DNA policy, began discussing the emerging legal and ethical questions last year. The group is likely to revisit the issue when the group meets this month at the FBI lab at Quantico, Va.³⁸

In the meantime, there is little legal guidance on family forensic DNA. Although the same article cites Thomas Callaghan, director of the FBI database program as claiming “Federal privacy law bars the FBI from performing familial searches within its own database,”³⁹ it goes on to say that Florida permits its state database operators to reveal information about offenders who match crime scene DNA at twenty-one of the twenty-six sites, and that Virginia permits disclosure when the match is “very, very close.”⁴⁰

The lack of legal guidance is not surprising. Although it is easy to argue that the American (and English) criminal justice system does not allow punishment of an offender’s family, it turns out to be hard to argue that it is opposed to *investigating* someone because of his family in spite of the uneasiness family forensic DNA causes. There seem to be no solid legal grounds, and only weak policy arguments, against such family-based investigation in general.

The United States has broadly rejected one kind of familial punishment: “corruption of blood.” Corruption of blood, which the colonies inherited from English common law, stripped the descendants of anyone convicted of a felony of their right to inherit the felon’s estate, any noble title, or any other “hereditament.” The Constitution itself prohibits “corruption of blood,” at least for one crime, stating “The Congress shall have Power to declare the Punishment of Treason, but no Attainder of Treason shall work Corruption of Blood, or Forfeiture except during the Life of the Person attainted.”⁴¹

Today, twenty-eight state constitutions prohibit any “corruption of blood” as a result of any conviction or attainder,⁴² another six prohibit it by statute for all convictions,⁴³ as does the Uniform Act on Status of Convicted Persons.⁴⁴ In England, corruption of blood

– but there seems to be nothing unconstitutional about such an arrest.

The closest example, though, comes from a visual family resemblance. If a witness said that a person in a photograph was not the perpetrator but looked very much like him – “could have been his brother” – there seems no reason to think the police could not try to find out whether the pictured man had a brother. Family forensic DNA is a molecular version of this kind of family resemblance.

One might argue that the use of family forensic DNA is an unreasonable search of the offenders’ relatives in violation of the relatives’ rights under the Fourth Amendment.⁴⁶ The application of the Fourth Amendment to compulsory inclusion in forensic databases has been anything but clear, with courts almost always upholding the constitutionality of the statutes but

The United States has broadly rejected one kind of familial punishment: “corruption of blood.” Corruption of blood, which the colonies inherited from English common law, stripped the descendants of anyone convicted of a felony of their right to inherit the felon’s estate, any noble title, or any other “hereditament.”

was abolished in 1870.⁴⁵ These provisions, though, are all concerned with the consequences of conviction, not with suspicion or investigation.

We have found no general rules, or even any discussion, of whether family relationships may be considered in criminal investigation. It seems assumed that the fact of a family relationship may sometimes be a relevant, and useable, fact for police or other investigations. Some reflection shows why.

A family relationship link may lead to suspicion or arrest for several reasons. For some crimes, a family relationship may provide a motive. Family members of a person who was attacked might understandably be suspects if the attacker himself is then assaulted. Other kinds of crimes may be thought, for sociological reasons, to “run in the family.” Thus, relatives of members of organized crime groups or of gangs are probably more likely to be suspected by the police of involvement in similar crimes. Finally, an investigation of one family member may lead to evidence of crimes committed by another. If the police search a family’s home under a valid search warrant and, while doing so, observe the suspect’s child using illegal drugs, they are not prohibited from making an arrest. The relationship between the two, and the consequent fact that they lived in the same place, was a “but for cause” of the child’s arrest

through varying and murky invocations of the “totality of the circumstances” or “special needs” doctrines, both in the context where those compelled to give DNA have had some diminished expectation of privacy. The un-convicted relatives of offenders do not have that diminished expectation of privacy, but extending either of these unclear doctrines to family members has the problem that nothing has been seized from them and they have not been searched.

In the absence of clear law or precedent, one could argue that this new method of family forensic DNA violates Constitutional rights associated with the family. The Constitution has been held to provide some protection for family decisions, including the education of children⁴⁷ and other aspects of their upbringing.⁴⁸ (One could even argue that the contraception⁴⁹ and abortion⁵⁰ cases are, in a sense, “family rights” cases, though they are not generally so categorized.) Writing for the Court’s plurality in *Troxel v. Granville*, Justice O’Connor reviewed the past cases and concluded that, “In light of this extensive precedent, it cannot now be doubted that the Due Process Clause of the Fourteenth Amendment protects the fundamental right of parents to make decisions concerning the care, custody, and control of their children.”⁵¹ It seems entirely too long a jump, though, to go from constitutional protection from state interference for decisions about child-rear-

ing or child-bearing to constitutional protection from state use of family resemblances in DNA in criminal investigations. In short, there seems to be no general legal barrier to the use of family forensic DNA.

The more complicated question is whether family forensic DNA, as a matter of policy, should be prohibited. If a sample of DNA from a crime scene is a partial match for the DNA of someone in the CODIS Offender Index, the police may decide to investigate that offender's relatives. This investigation is likely to involve, at first, an interview with the offender or perhaps the review of information previously collected about his relatives. If the offender has a relative who is a plausible suspect – most often likely to be defined as a relatively young male – that relative may then be the subject of an investigation, including possibly an interview with the police and a request for a DNA sample.

It is not likely that this would lead to the incorrect arrest or conviction of a suspect. In addition to whatever other alibis the family-member suspect may have, the crime scene DNA either will or will not match his own DNA. We will leave to other analysis whether the partial match, by itself or with some other plausible evidence, could be sufficient to compel production of a DNA sample. A person who knows he was not at the crime scene is likely to be willing to give a voluntary sample and would then be exonerated, barring the very low chances of technical error, an identical twin who was at the crime scene, or planting or contamination of the crime scene DNA with this family member's DNA. Innocent people could be forced to go farther into the criminal process only if they refused to provide a DNA sample or if their DNA was innocently at the crime scene. These risks are not zero, but seem likely to be quite small.

Nonetheless, merely being a suspect is a non-trivial cost to the relative. It involves time, some risk, and what will likely be a less than enjoyable encounter with the police. Imposing these costs on people as a result of their family relationship to an offender, who in turn is a partial match to crime scene DNA, seems unfair. On the other hand, people are interviewed by the police in connection with investigations for all sorts of reasons, fair or unfair. Additionally, the successful use of family forensic DNA will also spare some people interviews with the police. On balance, it might reduce the total number of people interviewed as potential suspects or witnesses. Given the uncertain net effect of this technology and the relatively low likely costs to innocent people of becoming suspects as a result of a partial DNA match, the "unfairness" of the family connection does not seem compelling. And, barring constitutional issues, any unfairness to a guilty suspect, who other-

wise might not have been caught, seems unlikely to carry much weight.

The costs of those relatives interviewed are, of course, not the only relevant costs. As a matter of policy, one would want to be confident that this technique did not soak up more police time and money than its benefits justified. It certainly has that potential, as a partial sample of crime scene DNA might well trigger hundreds of partial matches, which, if pursued, could require interviewing thousands of relatives. One useful aspect of this technology is that the searching for partial matches to an already analyzed genotype is nearly costless. After seeing how many hits they get, the police could then make a decision of which, if any, to pursue. Undoubtedly, the early uses of this technique would require substantial learning about when it is and is not useful. It is certainly possible that family forensic DNA, as a new technological "toy," would be overused at first, potentially at the expense of more effective, but less exciting, investigative techniques. But this should be only a short-run concern. It takes little special faith in criminal investigators to assume that they know their business and will eventually learn when and how this tool is best used.

In summary, neither the legal nor the policy arguments against family forensic DNA seem strong in general. It may prove a useful and cost-effective tool, in some cases, with little risk of false convictions or other excessive interference with innocent people. We do not, however, live "in general," and the specific circumstances of the American criminal justice system in the early 21st century raise some special concerns, to which we now turn.

Family Forensic DNA, Race, and America

What is only slightly troubling in the context of individual families may be much more troubling in a broader setting. African-Americans constitute about thirteen percent of the U.S. population, or about thirty-eight million people. In an average year, over forty percent of people convicted of felonies in the United States are African-American. As a result, the set of individuals in the Offender Index is not racially neutral with regard to the American population. Although we have not been able to find confirmation of this, we assume, based on the felony conviction statistics, that African-Americans make up at least forty percent of the CODIS Offender Index, or roughly 1.1 million people out of 2.75 million. The problems of embedding this racial disproportion in the Offender Database have already been ably argued by David Kaye and Michael Smith.⁵²

Assume that, either using the current CODIS markers or an expansion to roughly twice as many markers, partial matches of crime scene DNA samples to

the CODIS Offender Index could generate useful leads from among an offender's first degree relatives – parents, siblings, and children – but not more distant relatives. Using some additional simplifying assumptions, the percentage of African-Americans who might be identified as suspects through this method would be roughly four to five times as high as the corresponding percentage of U.S. Caucasians (the term used, along with African American, without a hyphen, in the CODIS system). (Non-African-American Hispanics occupy a middle position between African-Americans and non-Hispanic U.S. Caucasians in terms of their proportionate representation in the CODIS Offender Index. For the sake of simplicity, they have not been treated separately in this analysis.)

Assume first that family structures are the same for African-Americans and for non-Hispanic U.S. Caucasians in the CODIS Offender Index. Assume further that the average person in the database has five living first degree relatives. (Data on this point has, thus far, proven impossible to find, certainly for offenders but even for adult males, either by race or in general.) Under these assumptions, the 1.1 million African-Americans in the Offender Index will have 5.5 million first degree relatives, leading to a total of 6.6 million African-Americans “findable” through the database – the offenders and their relatives. That constitutes about seventeen percent of all African-Americans. U.S. Caucasians (including non-African-American Hispanics) make up about sixty percent of the Offender Index or currently about 1.65 million people. They would have 8.25 million first degree relatives, for total coverage of 9.9 million people “findable” through the database. U.S. Caucasians, including non-African-American Hispanics, constitute about eighty-three percent of the American population or about 247 million people. The 9.9 million U.S. Caucasians who would be either in the Offender Index, or a first degree relative of someone in the Index would make up just four percent of the white population. Thus, more than four times as much of the African-American population as the U.S. Caucasian population would be “under surveillance” as a result of family forensic DNA and the vast majority of those people would be relatives of offenders, not offenders themselves. (If non-African-American Hispanics were analyzed separately from non-Hispanic U.S. Caucasians, the disproportion between African-Americans and U.S. Caucasians would be even greater.)

This analysis is simplified in several respects. For one thing, the actual “coverage” is lower because some people are double-counted. Some people are both offenders and first degree relatives of offenders; others are first degree relatives of more than one offender. If the family structures of African-American and U.S.

Caucasians offenders are the same, however, the relative proportion should remain the same. On the other hand, the expansion of the power of this technique to allow it to identify second degree relatives would, of course, increase the number of people covered in both groups while generally retaining the same disproportionate impact on the African-American community.

That disparate impact alone seems unlikely to lead a court to hold that family forensic DNA violates the Equal Protection Clause. It is not the result of any unstated racially discriminatory purpose or intent in the use of family forensic DNA, but a consequence of the vast disproportion, for whatever reasons, in felony convictions between African-Americans and U.S. Caucasians. And yet, like racial profiling, it does seem fundamentally unfair, in a way that has systemic implications broader than those affecting random families that include convicted felons.

Of course, African-Americans are disproportionately harmed by crime committed by other African-Americans. Some African-Americans may see the usefulness of family forensic DNA in solving crimes as a net benefit to their community. Similarly, those interviewed by police, as witnesses or as possible suspects, when crimes take place in African-American communities are likely to be African-American themselves. Successful uses of family forensic DNA might reduce the number of people so interviewed. Although there is no empirical evidence on this point, we suspect that such positive reactions would be overwhelmed by unhappiness at the number of times African-American relatives of offenders are interviewed by the police as a result of a partial match, particularly as the people who are spared either victimization by criminals or interviews by the police as the result of family forensic DNA are “statistical people” and not individuals who know that they have individually benefited. On balance, we suspect that African-Americans are likely to oppose family forensic DNA and to view its use as another racist action by the American criminal justice system. This unfairness, real and perceived, is important and deserves attention in the context of America's historical and current problems with race.

Conclusion

This paper has argued that family forensic DNA has substantial potential to extend the usefulness of DNA databases in generating investigational leads from crime scene DNA. Several plausible enhancements could make it even more useful. Using DNA from offenders to help catch their relatives is, at the least, unsettling. The legal arguments against it, however, are quite weak, as are the policy objections. Political arguments, based on its unsettling nature, may prove at-

tractive, but, in general, it is not clear that they should. On the other hand, the way that family forensic DNA puts African-Americans under much greater investigative scrutiny may not be unconstitutional, but seems unfair and quite possibly unwise.

Our goal in this paper is to explore and bring to light these possibilities and implications of family forensic DNA, not to propose a general "solution" to the issues it raises. We believe that our society needs to discuss these issues broadly and reach an open and politically legitimate resolution. We do note, however, that the racial implications of this technique, and, in fact, the technique itself, would disappear if a population-wide DNA identification database existed. Family searches would be unnecessary, except for people who had somehow slipped through the cracks of the universal system, and the database would not just be representative of the country's ethnic and racial diversity, but would fully embody that diversity. The legal, practical, and political obstacles to such a database are substantial, but, in light of the extension of existing databases through both expansive legislation and new techniques like family forensic DNA, it may be wise to consider an expressly population-wide database rather than ending up with something that, through the extension to family members, becomes a large but racially biased database.⁵³

Acknowledgements

Greely is primarily responsible for the legal and policy discussions as well as the overall structure of the paper; Riordan, Garrison, and Mountain are primarily responsible for the scientific discussion with Riordan playing the largest role. The authors wish to thank Greely's research assistant, Jason Tarricone, for his excellent help, and Professors Richard Banks, Mark Kelman, Allen Weiner, and Eugene Volokh and an anonymous reviewer for helpful suggestions. Riordan was supported by a Stanford Graduate Fellowship. Garrison was supported by the Stanford Genome Training Program (Grant Number T32 HG00044 from the National Human Genome Research Institute). Mountain was supported by National Institutes of Health grant GM28428. The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health (or anyone else).

References

1. W. Shakespeare, *The Merchant of Venice*, III verse 1. It is perhaps revealing of Shakespeare's view of this maxim that it is said by Launcelot, the fool, to Jessica, Shylock's daughter, of whose marriage and conversion to Christianity the play approves. Although the term sounds Biblical, the King James version of the Bible at least does not refer to the "sins" of the fathers, but instead talks of the "iniquity" of the fathers. "For I the Lord thy God am a jealous God, visiting the iniquity of the fathers upon the children unto the third and fourth generation of them that hate me." Exodus, 20:5. Similar sentiments are expressed in Exodus 32:7, Numbers 14:18, Deuteronomy 5:9, and Jeremiah 32:18. But see Ezekiel 18:20 ("The son shall not bear the iniquity of the father, neither shall the father bear the iniquity of the son: the righteousness of the righteous shall be upon him, and the wickedness of the wicked shall be upon him.")
2. This description of the case is taken largely from "The Sins of the Fathers," *The Economist* (April 24, 2004); S. Bird, "M3 Brick Killer

- Tracked Through Relative's DNA," *The Times* (London) April 20, 2004, at 11; B. Taylor, "Killer Who was Trapped by his Relative's DNA," *Daily Mail*, April 20, 2004, at 25; *Forensic Science Service Annual Report, 2004-05*, at 9, available at <http://www.forensic.gov.uk/forensic_t/inside/about/annual.htm> (last visited February 22, 2006). One law review article describes the Harman case, though it does not analyze this kind of forensic search. See S. E. Henderson, "Nothing New Under the Sun? A Technologically Rational Doctrine of Fourth Amendment Search," *Mercer Law Review* 56 (2005): 507-563, at 554-55. Other overlapping news coverage of the Harman case includes J. Brookfield, "Trail-Blazing Use of DNA Evidence is Helping Police Beat the Criminals," *The Express*, April 21, 2004, at 17; S. Clough, "World First for Police as Relative's DNA Traps Lorry Driver's Killer," *The Daily Telegraph*, April 20, 2004, at 6; J. Twomey, "World First as New Technique Helps Track Down Yob," *The Express*, April 20, 2004, at 15; S. Raif, "DNA Link Helped to Trace Killer," *The Journal* (Newcastle), April 20, 2004, at 16; A. Williams, "Brick Killer Trapped by his Relative's DNA," *The Mirror*, April 20, 2004, at 13; P. Cheston and R. Mowling, "M3 Brick Killer Trapped by DNA Link in Family," *Evening Standard*, April 19, 2004, at C5; N. Allen, "Pioneering Technique Used to Track Killer," *Press Association*, April 19, 2004.
3. Almost all of the sources describe the person in the database only as a "close relative," apparently because of legal protections for the relative's privacy. Taylor, *supra* note 2. *The Economist* calls the two men brothers, which, given that Harman was twenty at the time and the "close relative" in the database was under thirty-five years old, seems most likely.
 4. *Forensic Science Service*, *supra* note 2.
 5. R. Willing, "Suspects Get Snared by a Relative's DNA," *USA Today*, June 7, 2005, at 1A.
 6. Bird, *supra* note 2. See also R. Williams and P. Johnson, "Inclusiveness, Effectiveness, and Intrusiveness: Issues in the Developing Uses of DNA Profiling in Support of Criminal Investigations," *Journal of Law, Medicine & Ethics* 33 (2005): 545-557, at 553-556. Reprinted in this issue, *Journal of Law, Medicine & Ethics* 34 (2006):234-247.
 7. "He May Never Have Been Caught," *Nottingham Evening Post*, May 28, 2004, at 2.
 8. Willing, *supra* note 5.
 9. Kansas authorities have not spoken much about their use of DNA in the BTK case, but several press reports indicate that they used a search warrant to a medical clinic to get tissue samples of the suspect's daughter. C. Moon, "Sides Trade Charges Over Abortion Records," *Topeka Capital-Journal*, September 2, 2005, at C1; S. Simon and P. J. Huffstutter, "Clues Were Clear but Slow to Add Up," *Los Angeles Times*, March 6, 2005, at A1. At least one reputable newspaper reported, however, that the daughter voluntarily provided the DNA sample in order to clear her father. M. Davey, "Computer Disk Led to Arrest in Killings, Pastor Says," *New York Times*, March 2, 2005, at A12.
 10. This use of DNA databases is discussed in the second workshop held by the American Society of Law, Medicine & Ethics under a grant to study issues in forensic DNA. DNA Fingerprinting and Civil Liberties Project, *Report of Workshop 2*, available at <http://www.aslme.org/dna_04/work2/report.php> (last visited February 22, 2006). Although one of the authors of this paper is part of the grant and participated in the third workshop, neither he nor any of the other authors were involved in the second workshop. Discussion also appears in several published sources. The most extensive written discussion of the issues appears to be four pages in R. Williams and P. Johnson, *supra* note 6, at 553-556. See the discussion in note 29, et. seq. See also, R. Williams, P. Johnson, P. Martin, *Genetic Information & Crime Investigation: Social, Ethical and Public Policy Aspects of the Establishment, Expansion and Police Use of the National DNA Database* (2004) at 108-110, available at <<http://www.dur.ac.uk/p.j.johnson/>> (last visited March 10, 2006). The issue also receives several pages of discussion in D. Lazer and M. N. Meyer, "DNA and the Criminal Justice System: Consensus and Debate," in D. Lazer, ed., *DNA and the Criminal Justice System: The Technology of Justice* (Cambridge, MA: MIT Press, 2004): at 371, 374-76. See discus-

- sion in note 28. Another chapter in the same book discusses in two paragraphs the technology of sibling-sibling partial matches. F. R. Bieber, "Science and Technology of Forensic DNA Profiling: Current Use and Future Directions," in D. Lazer, ed., *DNA and the Criminal Justice System: The Technology of Justice* (Cambridge, MA: MIT Press, 2004): at 47-48. A few other sources mention the possibility of such searches without any substantial analysis. T. Simoncelli and B. Steinhardt, "California's Proposition 69: A Dangerous Precedent for Criminal DNA Databases," *Journal of Law, Medicine & Ethics* 33 (2005): 279-291, at 284 (one paragraph discussing the possibility); E. Juengst, "I-DNA-Fication, Personal Privacy, and Social Justice," *Chicago Kent Law Review* 75 (1999): 61-82, at 79 (one paragraph mention of the possibility); and Willing, *supra* note 5. Cf. J. Buckleton and C. M. Triggs, "Relatedness and DNA: Are we taking it Seriously Enough?" *Forensic Science International* 152 (2005): 115-119 (suggesting that when reporting match probabilities, the probability that a sample matches a sibling should routinely be included).
11. Crime scene DNA is also usually analyzed for another marker, called amelogenin, which has two different lengths in men and only one length in women, thus allowing the DNA sample to be sexed.
 12. For ease of reference only, this article will refer to people in the Offender Index as "offenders," even though at least some of them will be people charged with, but not convicted of, a crime.
 13. Cal. Penal Code § 296 (2006) (murder and rape initially, all felonies in 2009); Tex. Government Code Ann. § 411.1471 (2006) (specified felonies, plus some former offenders when re-arrested); Va. Code Ann. § 19.2-310.2:1 (2006) (arrested for violent felonies); La. Rev. Stat. Ann. § 15:609 (2006) (all felonies); DNA Fingerprint Act of 2005, Title X of the Violence Against Women and Department of Justice Reauthorization Act of 2005, Section 1004(a)(1)(A), Pub. Law No. 109-162, 119 Stat. 2960 (2006).
 14. Since the passage of the Justice For All Act, in late 2004, federal law has authorized the states to submit DNA profiles to CODIS from a) Persons convicted of crimes; b) Persons who have been charged in an indictment or information with a crime; and c) Other persons whose DNA samples are collected under applicable legal authorities, provided that DNA profiles from arrestees who have not been charged in an indictment or information with crime, and DNA samples that are voluntarily submitted solely for elimination purposes shall not be included in the National DNA Index System. Pub. L. 109-405, § 203(a)(1). The FBI can also put into the index records that are "(2) analyses of DNA samples recovered from crime scenes; (3) analyses of DNA samples recovered from unidentified human remains; and (4) analyses of DNA samples voluntarily contributed from relatives of missing persons." 42 U.S.C. § 14132(a) (2006).
 15. The FBI also maintains a "population file" of anonymous genotypes used in determining the significance of any match.
 16. NDIS Statistics, available at <<http://www.fbi.gov/hq/lab/codis/clickmap.htm>> (last visited February 22, 2006).
 17. FBI CODIS webpage, *Investigations Aided*, available at <<http://www.fbi.gov/hq/lab/codis/aidedmap.htm>> (last visited February 22, 2006). This number may well be overstated; it is not clear how seriously the DNA analysis must assist the police in order to be counted as an aid to an investigation. For a fuller discussion of this topic, see in this issue, F. Bieber, "A Home Run or Just a Base Hit? Measuring and Improving the Efficacy of Forensic DNA Databank Programs," *Journal of Law, Medicine & Ethics* 34 (2006): 222-233.
 18. National Research Council, *DNA Technology in Forensic Science* (National Academies Press: 1992).
 19. National Research Council, *The Evaluation of Forensic DNA Evidence* (National Academies Press: 1996).
 20. There is some uncertainty in determining the length of an STR; a person who matches perfectly on twenty-five out of twenty-six markers and just misses matching on the last allele may really be a perfect match with the length of one allele misread on either the crime scene sample or in the analysis entered into the database.
 21. These and subsequent calculations concerning the number of matching alleles were performed by co-author Riordan with assistance from co-authors Mountain and Garrison. A scientific article setting out these calculations in more detail is in preparation.
 22. These odds vary depending on ethnicity. For most ethnicities in the United States, the odds are about 1 in 2,000; for Navajo, the odds are about 1 in 300.
 23. It is also possible that crime scene DNA might show partial matches to genotypes of two offenders who are known to be related to each other. This could make the probability much higher that the crime scene DNA came from someone in their family. We have not calculated the extent of the increased probability, but it seems likely, in some cases, to be substantial.
 24. This excludes identical (or "monozygotic") twins, who will necessarily share all twenty-six alleles.
 25. Bureau of Justice Statistics, U.S. Dept. of Justice, *State Court Sentencing of Convicted Felons, 2002 - Statistical Tables*, Table 2.1, available at <<http://www.ojp.usdoj.gov/bjs/abstract/scscfst.htm>> (last visited February 22, 2006).
 26. The British materials refer to a cost of £5,000 for a familial search, but this is not the cost but the *price* - the price charged to the police by the Forensic Science Service to run the search.
 27. DNA Fingerprinting and Civil Liberties Project, Report of Workshop 2, *supra* note 10.
 28. Lazer and Meyer, *supra* note 10.
 29. Williams and Johnson, *supra* note 6, at 553-56; see also Williams, Johnson, and Martin, *supra* note 10, at 108-110.
 30. Williams and Johnson, *supra* note 6, at 555.
 31. It would, however, be possible to prove that they did not share the same mother, using an analysis of mitochondrial DNA, which descends only from the mother. If both putative siblings were male, it could also be possible to prove that they did not have the same father, using an analysis of their Y chromosomes, which come only from fathers and go only to sons.
 32. Williams and Johnson, *supra* note 6, at 555. For another perspective on this issue of familial searching and families, see in this issue, E. Haimes, "Social and Ethical Issues in the Use of Familial Searching in Forensic Investigations: Insights from Family and Kinship Studies," *Journal of Law, Medicine & Ethics* 34 (2006): 263-276.
 33. Nevertheless, they are able to quote the custodian of the British DNA database as saying, "in a public meeting of the Human Genetics Commission held in February 2004 that '[Familial searching] is based on some very important assumptions that *criminality can run in families*, that a relative could be on the database, the families tend to live in the same area, and that offenders tend to offend close to their homes or in areas that they frequently visit.'" *Id.* (emphasis added). That statement can only be true in the very weak sense that such testing requires an assumption that it is possible that two members of the same family could both commit crimes - hardly a troubling position.
 34. *Id.*
 35. *Id.*
 36. Willing, *supra* note 5.
 37. Williams and Johnson, *supra* note 6, at 556.
 38. Willing, *supra* note 5.
 39. Willing, *supra* note 5. It is not clear on what privacy law Callaghan based that assertion. The statute authorizing the creation of the Offender Index allows the use only of information maintained by federal, state, or local agencies "pursuant to rules that allow disclosure of stored DNA samples and DNA analyses only (A) to criminal justice agencies for law enforcement identification purposes," among other purposes. 42 U.S.C. § 14132(b)(3)(A) (2006). The statute directly regulating CODIS and the Offender Index, in the provisions entitled "privacy protection standards," expressly allow "the results of DNA performed for a federal law enforcement agency" to criminal justice agencies for law enforcement identification purposes." 42 U.S.C. § 14133(b)(1) (2006). Furthermore, nothing in the statutes authorizing mandatory collection of DNA from federal offenders, 42 U.S.C. § 14135a (2006), or District of Columbia offenders, 42 U.S.C. § 14135b

- (2006), appears to limit the use of those samples for partial match searches. The portion of that statute entitled “privacy protection standards” merely restricts their uses to those authorized by the collection statutes and expressly permits their use under circumstances where CODIS samples may be used, cross-referencing 42 U.S.C. § 14132(b)(3). It is possible that Callaghan is construing “law enforcement identification purposes” to be limited to cases where there is an exact, and hence (almost) perfectly identifying match, but that seems strained. He may be referring to other federal privacy statutes, although the Federal Privacy Act, for example, contains a general authorization for disclosure for law enforcement purposes. 5 U.S.C. § 552a(b)(7) (2006). It is impossible to analyze Callaghan’s claim fully without more information about it. In any event, it does not sound as though Callaghan is asserting that state law enforcement officials cannot perform family searches of the Offender Database.
40. Willing claims that regulations in New York and Massachusetts allow familial searching. This assertion is based on an analysis of Massachusetts and New York regulations done for the American Society of Law, Medicine & Ethics. S. Axelrad, *States Regulations on Low Stringency/Familial Searches of DNA Databases* (2004), available at <http://www.aslme.org/dna_04/reports/index.php> (last visited February 22, 2006). The actual regulations, as they appear in Axelrad’s report, appear to be more concerned with the submission of partial genotypes – when, for example, for technical reasons fewer than all twenty-six alleles could be obtained from crime scene DNA – for search to the database, not the reporting from the database of partial genotype matches. Mass. Regs. Code 515 § 2.14; N.Y. Comp. Codes R. & Regs. 9 § 6192.3.
 41. U.S. Const. art. III, § 3, cl. 2. An almost identical prohibition was included, more than seventy years later, in the Confederacy’s constitution. Constitution of the Confederate States of America, art. III, § 3, cl. 2, available at <<http://www.yale.edu/lawweb/avalon/csa/csa.htm>> (last visited February 22, 2006). The Confederate constitution used exactly the same words as the Federal constitution, with minor variations in capitalization and punctuation.
 42. Ala. Const. Art. I, § 19; Alaska Const. Art. I, § 15; Ariz. Const. Art. II, § 16; Ark. Const. Art. II, § 17; Colo. Const. Art. II, § 9; Conn. Const. Art. IX, § 4; Del. Const. Art. I, § 15; Ga. Const. Art. I, § 1, para. XX; Idaho Const. Art. V, § 5; Ill. Const. Art. I, § 11; Ind. Const. Art. I, § 30; Ky. Const. § 20; Me. Const. Art. I, § 11; Md. Const. Art. 27; Minn. Const. Art. I, § 11; Mo. Const. Art. I, § 30; Neb. Const. Art. I, § 15; N.C. Const. Art. I, § 29; Ohio Const. Art. I, § 12; Okla. Const. Art. II, § 15; Or. Const. Art. I, § 25; Pa. Const. Art. I, § 19; S.C. Const. Art. I, § 4; Tenn. Const. Art. I, § 12; Tex. Const. Art. I, § 21; Wash. Const. Art. I, § 15; W. Va. Const. Art. III, § 18; and Wis. Const. Art. I, § 12.
 43. Haw. Rev. Stat. Ann. § 831-3 (2005); N.H. Rev. Stat. Ann. § 607-A:3 (2005); N.J. Stat. Ann. § 2A:152-2 (2005); N.D. Cent. Code § 12.1-33-02 (2005); R.I. Gen. Laws § 12-19-4 (2005); and Va. Code Ann. § 55-4 (2005). Several of the states with constitutional prohibitions also have statutory bans: Ark. Code Ann. § 14-14-805 (2005); Md. Code Ann., Est. & Trusts § 11-109 (2005); Tex. Prob. Code Ann. § 41 (2005), Tex. Crim. Proc. CODE ANN. § 1.19 (2005); and W. VA. CODE ANN. § 61-11-4 (2005).
 44. This uniform act is somewhat mysterious. It was recommended to the states for passage in 1964 by the National Conference of Commissioners on Uniform State Laws, but “was withdrawn from recommendation for enactment by the NCCUSL in 1966 due to being obsolete.” 11 Uniform Laws Annot. 235 (2001). At some point two states, Hawaii, Haw. Rev. Stat. §§831-1 *et seq.*, and New Hampshire, N.H. Rev. Stat. §§607-A:1, *et seq.*, adopted it. Neither the three mentions of the Act in law review articles that a search of Lexis uncovered nor the seven mentions in judicial opinions shed any light on its rapid obsolescence, or anything else about it. No mention of it can be found on the NCCUSL web site, at <<http://www.nccusl.org/>> (last visited February 22, 2006).
 45. It appears that corruption of blood for any treason or felony was abolished by the Forfeiture Act of 1870. The Act did not, however, abolish corruption of blood for outlawry (which seems to mean that if any people were declared outlaws, with no civil rights or status under the law, after 1870, they continued to experience corruption of blood). J. P. Bishop, *Bishop On Criminal Law*, J. M. Zane and C. Zollman, eds., vol. 1, 9th ed. (Chicago: T. H. Flood and Company, 1923): at 717; J. F. Stephen, *A History of the Criminal Law of England*, vol. 1 (originally published 1883) (Buffalo, NY: William S. Hein & Co., 1980): at 488.
 46. We owe the suggestion that this be discussed to an anonymous reviewer.
 47. See, e.g., *Pierce v. Society of the Sisters*, 268 U.S. 510 (1925) (state ban on private schools held unconstitutional); *Meyer v. Nebraska*, 262 U.S. 390 (1923) (state ban on teaching children German held unconstitutional); *Yoder v. Wisconsin*, 406 U.S. 205 (1972) (state law requiring high school education for fourteen- and fifteen-year-olds held unconstitutional, largely under the Free Exercise clause of the First Amendment, but also on the parents’ rights to determine the children’s education, or lack thereof).
 48. *Troxel v. Granville*, 530 U.S. 57 (2000) (state law giving grandparents visitation rights violates parents’ constitutional rights to control children’s rearing). *Moore v. City of East Cleveland*, 431 U.S. 494 (1977) (zoning ordinance with narrow definition of family violates the Due Process clause, plurality opinion).
 49. *Griswold v. Connecticut*, 381 U.S. 479 (1965).
 50. *Roe v. Wade*, 410 U.S. 113 (1973); *Planned Parenthood v. Casey*, 505 U.S. 833 (1992).
 51. *Troxel*, 530 U.S. at 66 (Justice O’Connor, writing for the plurality).
 52. D. H. Kaye and M. E. Smith, “DNA Identification Databases: Legality, Legitimacy, and the Case for Population-Wide Coverage,” *Wisconsin Law Review* (2003): 413-459.
 53. See Kaye and Smith, *supra* note 50.

Copyright of *Journal of Law, Medicine & Ethics* is the property of Blackwell Publishing Limited and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.