

# THE GENEALOGY DETECTIVES: A CONSTITUTIONAL ANALYSIS OF “FAMILIAL SEARCHING”

David H. Kaye\*

## ABSTRACT

*“Familial searching” in law enforcement DNA databases has been pilloried as a step “towards eugenics and corruption of blood” and “lifelong genetic surveillance” that is “inconsistent with a basic pillar of American political thought.” Courts have yet to address the issue fully, but several commentators contend the practice is unwise, unjust, or unconstitutional. This Article examines the more significant constitutional claims. It concludes that although kinship matching should not be implemented simply because it is technologically seductive, neither should it be removed from the realm of permissible law enforcement information gathering on constitutional grounds. In reaching this conclusion, the Article describes the logic of kinship analysis; clarifies the nature of partial-match searching; shows how an advanced system of DNA databases could yield additional, accurate leads in the investigation of both routine and high profile crimes; and explains why this system, if properly implemented, is compatible with constitutionally protected interests of both convicted offenders and their close relatives.*

## CONTENTS

INTRODUCTION . . . . .	110
I. FROM KINSHIP ANALYSIS TO KINSHIP MATCHING . . . . .	113
A. <i>Kinship Analysis with a Suspect</i> . . . . .	115
B. <i>Kinship Analysis with a Database: The Two Types of Database Trawls</i> . . . . .	118
C. <i>The Efficacy of Kinship Matching</i> . . . . .	121
II. EQUAL PROTECTION . . . . .	125
A. <i>Racial Discrimination</i> . . . . .	125
B. <i>Arbitrary Discrimination</i> . . . . .	127

---

\* Distinguished Professor and Weiss Family Scholar, Dickinson School of Law, Graduate Program in Forensic Science, Penn State University. Versions of this Article were presented in 2011 at the 22d International Symposium on Human Identification, the 7th International Society for Applied Biological Sciences Conference in Forensic, Anthropologic and Medical Genetics, and at a faculty workshop at the University of Pittsburgh School of Law. Charles Brenner, Rockne Harmon, David Lazer, Kristin Lewis O’Connor, Jennifer Kristin Wagner, Bruce Weir, Kenneth Weiss, and David Witherspoon commented on part or all of an early draft as did Erin Murphy, who kindly offered especially valuable suggestions and corrections. © 2013, David H. Kaye.

III.	ACQUIRING DNA FOR TRAWLING AS A SEARCH . . . . .	129
A.	<i>The Definition of a Search</i> . . . . .	132
B.	<i>Physical Extraction of a DNA Sample as a Search</i> . . . . .	134
IV.	THE REASONABLENESS OF INNER-DIRECTED AND OUTER-DIRECTED TRAWLING . . . . .	138
A.	<i>The State's Interests</i> . . . . .	139
1.	<i>Authentication: True Identity</i> . . . . .	139
2.	<i>Intelligence: Investigative Leads and Exonerations</i> . . . . .	140
3.	<i>The Reduced Expectation of Privacy</i> . . . . .	142
B.	<i>Database Inhabitants' Interests in Avoiding Outer-directed Trawls</i> . . . . .	145
1.	<i>"A Blooming Family Tree"</i> . . . . .	145
a.	<i>The desire to keep quiet</i> . . . . .	147
b.	<i>The desire not to know</i> . . . . .	149
2.	<i>More Markers</i> . . . . .	151
C.	<i>Relatives' Interests in Avoiding Outer-directed Trawls</i> . . . . .	153
1.	<i>Sheltering Evidence of Guilt</i> . . . . .	154
2.	<i>Avoiding False Convictions</i> . . . . .	155
3.	<i>Being a Suspect</i> . . . . .	156
4.	<i>Maintaining Spatial Anonymity</i> . . . . .	159
D.	<i>The Unbearable Lightness of Shadow Databases</i> . . . . .	161
	CONCLUSION . . . . .	162

## INTRODUCTION

DNA databases are a darling of the detective's nursery. They began as a curiosity thought to be useful for solving only a few types of violent crimes.<sup>1</sup> Today, they are dazzling devices for enforcing criminal laws from car theft to murder.<sup>2</sup> Computerized matching of the DNA identification profiles from crime-scenes and victims<sup>3</sup> to profiles from known individuals has produced hundreds of thousands of "cold hits."<sup>4</sup>

But some people are never satisfied. A number of scientists have concluded the databases could produce many useful investigative leads if a technique known as

1. See COMM. ON DNA TECH. IN FORENSIC SCI., NAT'L RESEARCH COUNCIL, DNA TECHNOLOGY IN FORENSIC SCIENCE 120 (1992) ("[I]t is clear that crimes of most types will not afford the opportunity to recover relevant biological evidence that will allow the police to identify an unknown suspect . . .").

2. See Edwin Zedlewski & Mary B. Murphy, *DNA Analysis for "Minor" Crimes: A Major Benefit for Law Enforcement*, 2006 NIJ J. 2, 2, 5 (noting expansion of DNA databases to solve property crimes in the United States and the United Kingdom) (2006).

3. For brevity, this Article uses the phrase "crime-scene DNA" to refer to DNA that appears to have come from a criminal, whether found in a location associated with the crime or on or within the body or clothing of an apparent victim.

4. *CODIS-NDIS Statistics*, FBI, <http://www.fbi.gov/about-us/lab/codis/ndis-statistics> (last visited Nov. 23, 2012).

kinship analysis were routinely employed.<sup>5</sup> To take their idea to its logical extreme, we can envision a database system constructed to be especially useful for this kind of analysis. Like today's databases, this system would pick out any individuals in the database who are likely sources of crime-scene DNA samples. But the trawling would not stop there. Almost magically, it could lead to identifications of individuals outside the database who left their DNA at crime scenes or on their victims.

Unfortunately, there is a catch. These new leads would point only to very close relatives who are not themselves subject to inclusion in the federal and state databases because they have not been convicted of qualifying crimes. Kinship matching, therefore, has been pilloried as “function creep,”<sup>6</sup> “mission creep,”<sup>7</sup> “a major privacy intrusion in the life of families,”<sup>8</sup> “the worst kind of guilt by association,”<sup>9</sup> “genetic surveillance for all,”<sup>10</sup> and “lifelong genetic surveillance”<sup>11</sup> that is “inconsistent with a basic pillar of American political thought.”<sup>12</sup> And, as if

---

5. Frederick R. Bieber et al., *Finding Criminals Through DNA of Their Relatives*, 312 *SCIENCE* 1315 (2006). Dr. Bieber and his colleagues reason as follows:

Consider a hypothetical state in which the “cold-hit” rate—the chance of finding a match between a crime scene sample and someone in the offender database—is 10%. Suppose that among criminals who are not (yet) in the database themselves, even 5% of them have a close (parent/child or sibling) relative who is. From our projections that up to 80% (counting the 10 best leads) of those 5% could be indirectly identified, it follows that the kinship analyses we describe could increase a 10% cold-hit rate to 14%—that is, by 40%. There have been 30,000 cold hits in the United States up to now. Kinship searching has the potential for thousands more.

*Id.* at 1315–16.

6. Barbara Prainsack, *Key Issues in DNA Profiling and Databasing: Implications for Governance*, in *GENETIC SUSPECTS: GLOBAL GOVERNANCE OF FORENSIC DNA PROFILING AND DATABASING* 15, 28 (Richard Hindmarsh & Barbara Prainsack eds., 2010).

7. Sonia M. Suter, *All in the Family: Privacy and DNA Familial Searching*, 23 *HARV. J.L. & TECH.* 309, 317 (2010).

8. Richard Willing, *DNA Database Can Flag Suspects Through Relatives: Critics Fear “Genetic Surveillance” of Innocent*, *USA TODAY*, Aug. 22, 2006, at 2A (quoting Tania Simoncelli).

9. *Id.*

10. Jeffrey Rosen, *Genetic Surveillance for All*, *SLATE* (Mar. 17, 2009), <http://www.slate.com/id/2213958>; cf. Frederick R. Bieber & David Lazer, *Guilt by Association: Should the Law Be Able to Use One Person’s DNA to Carry Out Surveillance on Their Family? Not Without a Public Debate Say Frederick Bieber & David Lazer*, *NEW SCI.*, Oct. 23, 2004, at 20 (“[Familial searching] brings with it the indirect lifelong surveillance of citizens simply because they are related to someone whose DNA profile is on record.”); Barbara Goldberg, *Virginia May Nab Serial Rapist Through Blood Relative DNA: Familial DNA Powerful Forensic Tool But Few States Use It Because of Privacy Concerns*, *ABC NEWS* (Aug. 9, 2010), <http://abcnews.go.com/News/TheLaw/familial-dna-expose-east-coast-rapist/story?id=11334748> (quoting Erin Murphy as saying “It’s guilt by association”).

11. SHELDON KRIMSKY & TANIA SIMONCELLI, *GENETIC JUSTICE: DNA DATA BANKS, CRIMINAL INVESTIGATIONS, AND CIVIL LIBERTIES* 175 (2011); Suter, *supra* note 7, at 311; Ellen Nakashima, *From DNA of Family, A Tool to Make Arrests: Privacy Advocates Say the Emerging Practice Turns Relatives Into Genetic Informants*, *WASH. POST*, Apr. 21, 2008, at A1 (quoting “Tania Simoncelli, science adviser to the American Civil Liberties Union”).

12. KRIMSKY & SIMONCELLI, *supra* note 11, at 83 (attributing this view to Professor Jeffrey Rosen).

all that were not enough, it has been tarred as “biological determinism”<sup>13</sup> and a step “towards eugenics and corruption of blood.”<sup>14</sup>

Although early commentators perceived no fundamental legal barriers to kinship matching<sup>15</sup> and most writing continues to focus on policy arguments,<sup>16</sup> some recent commentary displays more sympathy or support for constitutional objections.<sup>17</sup> The most prominent example is an essay by Professor Erin Murphy entitled *Relative Doubt: Familial Searches of DNA Databases*, which contends the technique is counterproductive for police practice, unfair, unjust, and of doubtful constitutionality.<sup>18</sup>

This Article provides a more complete examination of the two most significant constitutional issues—the Fourteenth Amendment’s guarantee of equal protection of the laws and the Fourth Amendment’s protection against unreasonable searches

---

13. Erin Murphy, *Relative Doubt: Familial Searches of DNA Databases*, 109 MICH. L. REV. 291, 304 (2010) (“[F]amilial searches . . . embody . . . biological determinism.”) [hereinafter Murphy, *Relative Doubt*].

14. Adam Schwartz, Senior Staff Counsel, ACLU of Ill., DNA Familial Testing: Civil Liberties and Civil Rights Concerns, Address at the Northwestern Law School Center on Wrongful Convictions Symposium on Familial DNA Searching (Mar. 3, 2011), at 4, <http://www.aclu-il.org/wp-content/uploads/2011/03/Speech-DNA-familial-testing.pdf>. The constitutional limitation on “corruption of blood” obviously has no application to criminal investigative practices. See Henry T. Greely et al., *Family Ties: The Use of DNA Offender Databases to Catch Offenders’ Kin*, 34 J.L., MED. & ETHICS 248, 257 (2006); David H. Kaye, *Associational Privacy, the Presumption of Innocence, and “Corruption of Blood” as Constitutional Metaphors in the Debate on “Familial Searching,”* AMER. CRIM. L. REV. (Nov. 26, 2012), <http://www.americancriminalawreview.com/Drupal/blogs/blog-entry/associational-privacy-presumption-innocence-and-“corruption-blood”-constitutional-m> [hereinafter Kaye, *Associational Privacy*].

15. Jules Epstein, “Genetic Surveillance”—The Bogeyman Response to Familial DNA Investigations, 2009 U. ILL. J.L. TECH. & POL’Y 141 (2009); Greely et al., *supra* note 14. Professor Greely recently testified in support, “with reservations,” of legislation to facilitate familial searching. *DNA Technology and Cold Cases: Hearing on H.R. 3361 Before the H. Subcomm. On Crime, Terrorism, and Homeland Security of the H. Comm. on the Judiciary*, 112th Cong. 25 (2012) (statement of Henry T. Greely, Professor of Law, Stanford Law School).

16. E.g., KRIMSKY & SIMONCELLI, *supra* note 11; Jessica D. Gabel, *Probable Cause from Probable Bonds: A Genetic Tattle Tale Based on Familial DNA*, 21 HASTINGS WOMEN’S L.J. 3, 57 (2010) (urging “vigilant and uniform” and “rigorous and practical guidelines” to cope with the “certainty” that kinship analysis in databases will become widespread); Joyce Kim et al., *Policy Implications for Familial Searching*, 2 INVESTIGATIVE GENETICS 22 (2011), available at <http://www.investigativegenetics.com/content/2/1/22>; Natalie Ram, *The Mismatch Between Probable Cause and Partial Matching*, 118 YALE L.J. POCKET PART 182 (2009); Natalie Ram & Michael Seringhaus, *O Brother, Where Art Thou?*, SLATE (June 14, 2010), <http://www.slate.com/id/2256966>.

17. Cf. Linda Bartusiak, Comment, *Plea Bargaining for DNA: Implications on the Right to Privacy*, 13 U. PA. J. CONST. L. 1115, 1126–28 (2011) (finding a violation of reasonable expectations of privacy from the involuntarily obtained DNA sample). *But see* Brett Mares, Note, *A Chip off the Old Block: Familial DNA Searches and the African American Community*, 29 LAW & INEQ. 395, 397 (2011) (concluding “familial DNA testing is constitutional under the Fourteenth and Fourth Amendments”); Mary McCarthy, Note, *Am I My Brother’s Keeper?: Familial DNA Searches in the Twenty-first Century*, 86 NOTRE DAME L. REV. 381, 382 (2011) (summarizing the literature and describing the legal arguments as “weak”).

18. Murphy, *Relative Doubt*, *supra* note 13. Her arguments are summarized, with apparent approval, in Natalie Ram, *Fortuity and Forensic Familial Identification*, 63 STAN. L. REV. 751, 789–94 (2011). Similar concerns can be found in Suter, *supra* note 7. The first mention of “ethical and constitutional concerns” may be Michelle Hibbert, *DNA Databanks: Law Enforcement’s Greatest Surveillance Tool?*, 34 WAKE FOREST L. REV. 767, 786 (1999).

and seizures.<sup>19</sup> These constitutional provisions, I maintain, are not show-stoppers. Kinship matching should not be implemented simply because it is technologically seductive, but neither should it be taken off the legislative table on constitutional grounds. To reach this conclusion, Part I describes the logic of kinship analysis and how it can be applied to state-of-the-art forms of DNA databases that could yield accurate leads in the investigation of both routine and high profile crimes. It introduces a few standard terms from genetics, presents more neutral terminology than the slightly ominous phrase, “familial searching,” and explains how kinship matching differs from the partial matching the FBI allows in the national database (NDIS) that is part of the Combined Offender DNA Index System (CODIS). It also discusses the difficulty of measuring the efficacy of “familial searching.”

Parts II through IV analyze the two main constitutional objections to kinship matching. Part II argues the practice is clearly compatible with the established understanding of the Equal Protection Clause. Parts III and IV analyze the interests of all convicted offenders and their families to show why kinship matching in law enforcement databases can qualify as a reasonable search or seizure under the Fourth Amendment. Like every other investigative technique, it can adversely affect very close relatives, but the actual Fourth Amendment interests of the individuals in the database and their close relatives in keeping the state from finding investigative leads from crime-scene DNA are weak. The government interest in efficiently investigating crimes with a thorough and properly implemented system of kinship matching therefore outweighs these interests.

## I. FROM KINSHIP ANALYSIS TO KINSHIP MATCHING

Kinship analysis refers to comparing DNA from different individuals to see if those individuals might be related. It is done frequently in child support<sup>20</sup> and immigration<sup>21</sup> cases and in missing persons and human remains investigations.<sup>22</sup> It

---

19. Other constitutional provisions—involving associational privacy, the presumption of innocence, and corruption of blood—have been invoked, but they are makeweights. David H. Kaye, *Drawing Lines: Unrelated Probable Cause as a Prerequisite to Early DNA Collection*, 91 N.C. L. REV. ADDENDUM 1 (2012); Kaye, *Associational Privacy*, *supra* note 14. Consequently, they are not pursued here.

20. *See, e.g.*, 1 GEORGE E. DIX ET AL., MCCORMICK ON EVIDENCE § 211 (Kenneth S. Broun ed., 6th ed. 2006) [hereinafter MCCORMICK] (describing increased reliance on DNA testing in paternity suits); DAVID H. KAYE ET AL., THE NEW WIGMORE, A TREATISE ON EVIDENCE: EXPERT EVIDENCE § 14.3.2 (2d ed. 2011) (describing the application of DNA databases to paternity cases) [hereinafter KAYE ET AL., WIGMORE].

21. The first well-publicized application of forensic DNA typing occurred in a British immigration case. DAVID H. KAYE, THE DOUBLE HELIX AND THE LAW OF EVIDENCE 55 (2010) [hereinafter KAYE, DOUBLE HELIX].

22. *E.g.*, M. Prinz et al., *DNA Commission of the International Society for Forensic Genetics (ISFG): Recommendations Regarding the Role of Forensic Genetics for Disaster Victim Identification (DVI)*, 1 FORENSIC SCI. INT'L: GENETICS 3 (2007) (recommending further DNA sampling as a method to prepare for identifying victims of disasters).

is done in criminal cases when a rape victim has a child or an aborted fetus.<sup>23</sup> It is one reason to believe U.S. soldiers killed Osama Bin Laden rather than a man who merely resembled him.<sup>24</sup> These applications are uncontroversial.

To appreciate the more controversial use of the procedure in criminal database trawls, it is important to understand the scientific and statistical principles behind kinship analysis. These are straightforward, but a few technical details bear heavily on the procedure's efficacy and invasiveness. This Part, therefore, defines and briefly describes the types of chromosomes, alleles, and loci used in forensic DNA identification and how these generate likelihood ratios for specific genetic relationships between individuals that can be used to produce investigative leads.<sup>25</sup>

*Chromosomes.* In humans, DNA comes in packages known as chromosomes. Each cell nucleus normally contains twenty-three pairs of chromosomes.<sup>26</sup> In twenty-two of these pairs, called autosomes, the two chromosomes are about the same length.<sup>27</sup> The twenty-third pair differs between males and females.<sup>28</sup> Females have two copies of the X chromosome, while males have one X and one much smaller Y.<sup>29</sup>

Sex cells (eggs and sperm) are exceptional in that they have a reduced number of chromosomes.<sup>30</sup> Each sex cell contains only one chromosome from each homologous pair, chosen at random, giving them a total of twenty-three individual chromosomes.<sup>31</sup> When a sperm and egg cell combine, a new set of twenty-three pairs is formed.<sup>32</sup> The fertilized cell divides, as do its daughter cells, giving rise to trillions of cells in the offspring.<sup>33</sup> All the new cells (except for sex cells) have the same genome of twenty-three homologous chromosomes.<sup>34</sup> One member of each pair has been inherited at random from one parent, and the other member was

---

23. See KAYE ET AL., WIGMORE, *supra* note 20, § 14.3.2 (describing cases); Willing, *supra* note 8 (mentioning Florida's practice of searching its offender database for profiles indicative of the paternity of children born following a rape).

24. David H. Kaye, *Osama Bin Laden's DNA? 99.9% Accuracy and 0.1% Nonsense*, THE DOUBLE HELIX LAW BLOG (May 3, 2011, 12:23 PM), <http://www.personal.psu.edu/dhk3/blogs/DoubleHelixLaw/>; Donald G. McNeil Jr. & Pam Belluck, *Experts Say DNA Match Is Likely a Parent or Child*, N.Y. TIMES, May 3, 2011, at F2.

25. The discussion that follows oversimplifies some facts. For more detail and qualifications, see, for example, FORENSIC DNA EVIDENCE INTERPRETATION (John Buckleton et al. eds., 2005); JOHN M. BUTLER, ADVANCED TOPICS IN FORENSIC DNA TYPING: METHODOLOGY (2011); WILLIAM GOODWIN ET AL., AN INTRODUCTION TO FORENSIC GENETICS (2d ed. 2011); KAYE, DOUBLE HELIX, *supra* note 21; David H. Kaye & George Sensabaugh, *Reference Guide on DNA Evidence*, in NAT'L RESEARCH COUNCIL, FED. JUDICIAL CTR., REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 129 (3d ed. 2011).

26. Kaye & Sensabaugh, *supra* note 25, at 137.

27. *Id.* at 136, 200.

28. *Id.* at 137.

29. *Id.*

30. *Id.*

31. *Id.*

32. *Id.*

33. *Id.*

34. *Id.* at 137–38.

inherited at random from the other parent.<sup>35</sup>

*Loci and alleles.* The DNA in a chromosome can be thought of abstractly as a string of four letters (chemical “base pairs”), designated A, T, C, and G.<sup>36</sup> The sequence of base pairs in a particular chromosome (number 16, for example) is mostly the same from one individual to another, but at some locations the sequences are different.<sup>37</sup> The sequence at each location, or “locus,” is called an “allele.”<sup>38</sup> The simplest alleles are a substitution, deletion, or insertion of a letter at a particular locus.<sup>39</sup> Such variations in the DNA sequence of the same chromosome in different people (or between the paired chromosomes in the same person) are “single-nucleotide polymorphisms,” or SNPs.<sup>40</sup> Another kind of polymorphism is a variation in the length of a region of DNA that arises from different numbers of several short, repeated letters.<sup>41</sup> For example, one chromosome number 16 in one individual might have the sequence GATA repeated eight times at a particular locus. Another chromosome 16 might have ten repeats of GATA. Short-tandem repeat, or STR loci, thus resemble trains with different numbers of boxcars. There are other kinds of sequence variations, but STRs currently are the most popular loci for identity and kinship testing, and SNPs are expected to come into widespread use in the near future.<sup>42</sup>

#### A. Kinship Analysis with a Suspect

With these terms defined, we can offer a simplified example of kinship analysis that will facilitate later discussion. The hypothetical example involves only two DNA samples and two STR loci, but it illustrates how kinship analysis could supply relevant information in a criminal investigation. A laboratory extracts from one sample, recovered in a rape case, the DNA from sperm cells. The laboratory also extracts DNA from cells obtained by swabbing the inside of Joe Suspect’s cheek pursuant to a court order.<sup>43</sup> Analyzing these samples, it finds the DNA profiles listed in Table 1, which also has figures on the frequencies of the alleles in the general U.S. Caucasian population.<sup>44</sup>

---

35. *Id.*

36. *Id.* at 136.

37. *Id.* at 137.

38. *Id.* at 139.

39. *Id.*

40. *Id.*

41. *Id.* at 139, 141–41.

42. See Manfred Kayser & Peter de Knijff, *Improving Human Forensics Through Advances in Genetics, Genomics and Molecular Biology*, 12 NATURE REV. GENETICS 179 (2011) (detailing current methods of DNA identification and anticipating future developments in the field).

43. Many jurisdictions provide for such orders on a showing of reasonable suspicion rather than the traditional standard of probable cause required for a search and seizure. Edward J. Imwinkelried & D. H. Kaye, *DNA Typing: Emerging or Neglected Issues*, 76 WASH. L. REV. 413, 420–22 (2001).

44. John M. Butler et al., *Allele Frequencies for 15 Autosomal STR Loci on U.S. Caucasian, African American, and Hispanic Populations*, 48 J. FORENSIC SCI. 1, 2 Table 1 (2003); John M. Butler et al., *Allele Frequencies for 27*

**Table 1. Hypothetical two-locus STR profiles and allele frequencies from a crime and a suspect. The D16S539 locus is on chromosome 16, and DYS444 is on the Y chromosome. The profile associated with the alleged rapist consists of eight repeats on one chromosome 16 and ten repeats on his other chromosome. Because a man has only one Y chromosome (having inherited it from his father and an X from his mother), there is only one allele at this locus. The percentages are frequencies for these alleles in samples of U.S. Caucasians.**

<i>Sample</i>	<i>Locus</i>		
	<b>D16S539</b>		<b>DYS444</b>
Crime (sperm cells)	8 (2%)	10 (6%)	15 (1%)
Suspect (Joe)	8 (2%)	12 (33%)	15 (1%)

The DNA testing apparently has absolved Joe because he does not match the crime sample completely. Joe cannot be the source of the sperm because his allele 12 at the D16S539 locus was not present in the sperm DNA. Suppose, however, that the victim picked Joe out of a lineup, and the police know that he has a brother, Jim, who resembles him. Does the near miss in the profiles of the sperm DNA from the unknown rapist and the cheek DNA from Joe support the surmise that Jim is the rapist?

Kinship analysis offers an answer. It contrasts the probability  $P_S$  of observing the STR types when Joe and the unknown source of the crime-scene sample are siblings to the probability  $P_U$  of observing these types when they are unrelated. The ratio of these probabilities,  $P_S/P_U$  is called the kinship index (or, in this case, the sibblingship index).<sup>45</sup> It is an example of a likelihood ratio. Likelihood ratios measure the probative value of evidence.<sup>46</sup> When the ratio is one, evidence is equally probable under the two hypotheses and thus has no probative value in deciding which hypothesis is true.<sup>47</sup> The extent to which the ratio exceeds one indicates how much the evidence supports the first hypothesis over the second.<sup>48</sup> For example, a positive result on a rapid test for the flu (known as the QuickVue

---

*Y-STR Loci with U.S. Caucasian, African American, and Hispanic Samples*, 156 FORENSIC SCI. INT'L 250, 253 tbl.2 (2006).

45. Jianye Ge et al., *Comparisons of Familial DNA Database Searching Strategies*, 56 J. FORENSIC SCI. 1448, 1448 (2011).

46. *E.g.*, KAYE ET AL., WIGMORE, *supra* note 20, at § 14.2.1; MCCORMICK, *supra* note 20, at § 185. Contrary to the description in Gabel, *supra* note 16, at 21, taken from NORAH RUDIN & KEITH INMAN, AN INTRODUCTION TO FORENSIC DNA ANALYSIS 149 (2d ed. 2002), the likelihood ratio does not give the relative probability of the hypothesis of relatedness. It gives the relative probability of the observed alleles under different hypotheses. Naively converting the likelihood ratio into the posterior odds is a form of the transposition fallacy. KAYE ET AL., WIGMORE, *supra* note 20, at § 14.2.1–2.

47. KAYE ET AL., WIGMORE, *supra* note 20, at § 14.2.1.

48. *Id.*

Influenza A+B Test) on a patient with flu-like symptoms supports a diagnosis of the disease because the test gives this result more often when the virus is present than when something else is responsible.<sup>49</sup> But the likelihood ratio is only about nine, indicating that the test, while rapid, is far from definitive.<sup>50</sup>

At the autosomal locus D16S539, the likelihood ratio for full siblingship is 6.5.<sup>51</sup> By itself, the one shared allele proves a little, but not much.<sup>52</sup> The matching Y-STR is more revealing. If Joe and the rapist are full siblings, both inherited their Y chromosome from the same man, so the probability of the matching DYS444 allele is  $P_S = 100\%$ . On the other hand, if Joe and the rapist are unrelated, then the probability of a match is roughly the proportion of Caucasian men with Y chromosomes who have this allele.<sup>53</sup> This frequency is  $P_U = 1\%$ . Thus, the likelihood ratio at this locus is  $100/1 = 100$ . It is 100 times more probable for the matching Y-chromosome evidence to arise when Joe and the rapist are full siblings (or otherwise in the same paternal lineage) than when they are two random men. Combining the two likelihood ratios gives the figure of  $6.5 \times 100 = 650$ .<sup>54</sup> The near miss in the STRs is 650 times as probable when the two profiles come from Joe and his (as yet untested) brother Jim than when the profiles come from Joe and an unrelated man.

This information would be quite useful to a magistrate asked to issue a warrant requiring the brother, Jim, to provide a DNA sample. Yet, Jim might object to informing the magistrate of the siblingship index on the ground that the police effectively searched his DNA when—without a new warrant—they reanalyzed Joe's profile to see whether the semen might have come from a possible sibling. He could argue that, up to that point, the police had no reason to suspect *him*—other than his being related to their suspect—and that it was unjust to place him under genetic surveillance because of his involuntary family ties.

However, the Fourth Amendment allows kinship analysis without a warrant. In

---

49. See Timothy M. Uyeki et al., *Low Sensitivity of Rapid Diagnostic Test for Influenza*, 48 CLINICAL INFECTIOUS DISEASES at E89, E90 (2009) (reporting the probability of a positive result when the virus is present is about 27% compared to a 3% probability of this evidence when something else is responsible).

50. *Id.*

51. The likelihood ratio is given by the expression  $(1 + 2p)/8p$ , where  $p$  is the frequency of the one shared allele. Charles H. Brenner, *Likelihood Ratios for Sibship and Half-sibship*, FORENSIC MATHEMATICS, <http://dna-view.com/sibfmla.htm> (last visited Nov. 25, 2012). Here,  $p = 2\%$ , so the ratio is  $104/16 = 6.5$ .

52. The ratio for half-siblingship is  $(1 + 4p)/8p = 108/16 = 6.75$ . *Id.*

53. For refinements, see David Balding, *Short Tandem Repeats: Interpretation*, in 5 WILEY ENCYCLOPEDIA OF FORENSIC SCIENCE 2365 (Allan Jamieson & Andre Moenssens eds., 2009); J.S. Buckleton et al., *The Interpretation of Lineage Markers in Forensic DNA Testing*, 5 FORENSIC SCI. INT'L: GENETICS 78 (2011).

54. Multiplying the likelihood ratios is appropriate if the alleles on chromosome 16 are uncorrelated with the Y chromosome haplotype. In a large, randomly mating population, the Y haplotype will propagate over many generations (or be extinguished if men with that haplotype have no male offspring), whereas children will inherit one of every pair of autosomal chromosomes at random, and these chromosomes will change as a result of crossing over (the swapping of parts of the homologous chromosomes during meiosis). Kaye & Sensabaugh, *supra* note 25, at 137, 181–82. Over time, this recombination in autosomes will wash out any correlation between the Y haplotype and the autosomal genotype.

this example, the interference with Joe's interests, pursuant to court order, was fully justified, and Jim has no plausible Fourth Amendment claim to immunity from adverse inferences that flow from the legitimately acquired sample—even if it takes a little more investigative effort to draw these inferences. In this context, there is little reason to distinguish between (1) a court order that requires a suspect to surrender his DNA to ascertain whether the suspect matches the crime-scene sample and (2) the same court order issued with the understanding that the police can consider both whether the suspect's DNA matches and, if it does not, whether a close relative's DNA is likely to match. Parts IV and V present the more detailed analysis that underlies these conclusions in the context of a database trawl case. Before engaging in this analysis, however, we should consider how kinship analysis with a database works.

### *B. Kinship Analysis with a Database: The Two Types of Database Trawls*

Using the same principles of genetics, kinship searches can be conducted in a law enforcement database of DNA identification profiles. For brevity, we can call the convicted offenders, whose numerical profiles are recorded, “database inhabitants.”<sup>55</sup> A database is essentially an extended version of Table 1. Instead of a single row for Joe Suspect in our example, the table has rows for the profiles of all the database inhabitants. Moreover, each row is much longer, being based in the United States on thirteen autosomal STR loci (no X- or Y-STR loci).<sup>56</sup> Whichever loci are in the database, the profile of these loci derived from a crime scene can be compared with all the database profiles to see (1) if a database inhabitant's profile is a match (the usual “cold hit”), indicating that the crime-scene DNA may have come from the database inhabitant, or (2) if a database inhabitant's profile is a close enough partial match to establish that the crime-scene DNA may have come from a very close relative of the database inhabitant. I shall call searches for normal cold hits “inner-directed” trawls and searches for hits indicating a close relationship to someone outside the database “outer-directed” trawls.

---

55. More than half the states and the federal government now take DNA from mere arrestees. Julie Samuels et al., *Collecting DNA from Arrestees: Implementation Lessons*, NIJ J., June 2012, <http://www.nij.gov/nij/journals/270/arrestee-dna.htm>. The analysis in this Article would apply to these database inhabitants and their first-degree relatives as well, taking into account the lesser interest that the government has in acquiring DNA from individuals not convicted of an offense. See *infra* Part IV.A.3.

Professor Murphy uses the terms “offenders” and “database leads” for “the individuals who, by possessing a databased profile that partially matches a crime-scene sample, point toward suspects, one of whom may be the source.” Murphy, *Relative Doubt*, *supra* note 13, at 298. Suter, *supra* note 7, at 342, and KRIMSKY & SIMONCELLI, *supra* note 11, at 84, prefer the more colorful phrase “genetic informant.” Gabel, *supra* note 16, at 4, uses the term “genetic snitch.” Of course, by the same token, a patient who carries a genetic disease becomes a “genetic informant” or “genetic snitch” when she visits a physician who diagnoses the disease. See, e.g., *Pate v. Threlkel*, 661 So. 2d 278, 280 (Fla. 1995) (concluding a physician had a duty to warn relatives that a patient had a hereditary disease).

56. J. M. Butler & C. R. Hill, *Biology and Genetics of New Autosomal STR Loci Useful for Forensic DNA Analysis*, 24 FORENSIC SCI. REV. 15, 16 (2012).

For example, if the police trawl for a match to a crime-scene sample and Joe's profile is flagged as a full match, he becomes a suspect as a result of the inner-directed trawl. But if neither Joe nor any other database inhabitant matches in full, kinship ratios for all near-misses can be computed, in the hope of finding a database inhabitant (like Joe) with an unusually large kinship ratio to the individual outside the database who left the crime-scene sample (like Jim). With this outer-directed trawl, Jim—who was outside the database—might come to be a suspect.

As we have seen, the likelihood ratio for a particular partial match indicates how strongly the overlap in the profiles supports the inference of a particular relationship.<sup>57</sup> Even with the standard thirteen CODIS loci, which are not the best for kinship analysis, likelihood ratios for parent-child or sibling-sibling relationships can be quite large. Over ten years ago, the National Commission on the Future of DNA Evidence provided an example of a siblingship index for these loci of about one million—“that is, the match probability is a million times as great if the DNAs came from siblings [than] if they came from unrelated persons.”<sup>58</sup> If the samples had come from a crime scene and a database inhabitant (like Joe), this analysis would have strongly suggested that the crime-scene profile had come from a full brother or sister of the database inhabitant (like Jim).<sup>59</sup>

Outer-directed trawls popularly are known as “familial searching.”<sup>60</sup> Like that favorite phrase of opponents of estate taxation—“the death tax”<sup>61</sup>—“familial search” has a noxious connotation. It gives an impression of searching through an extended family—aunts, uncles, cousins, nephews, and so on. The book *Genetic Justice*, for instance, warns that “[a]nyone who has his or her DNA profiled in a state DNA data bank . . . brings his or her entire family under DNA surveillance.”<sup>62</sup> The current reality is somewhat less threatening. Although likelihood ratios can be produced for any desired relationship, with the limited number of

---

57. With a reasonable number of loci, the pattern of allele sharing across loci generally makes it easy to distinguish parent-child from sibling-sibling relationships. A child must possess at least one allele in common with each parent at every locus, while two full siblings have only a 25% chance of inheriting the same pair of alleles from their parents and a 50% chance of inheriting one allele in common. The probability of two siblings having at least one allele in common by descent at  $n$  loci in a row is therefore  $(3/4)^n$ . For 10 loci, about  $1 - (3/4)^{10} = 94\%$  of siblings will have at least one locus with no shared alleles. The figure will be lower when parents have alleles in common. For 13 loci, the figure is 98%; for 20 loci, 99.7%. In contrast (barring mutations), no parent-child pairs will fail to have at least one matching allele at every locus. Finding at least one allele in common at every locus is thus strong evidence of the parent-child relationship as compared to siblingship.

58. NAT'L COMM'N ON THE FUTURE OF DNA EVIDENCE, THE FUTURE OF FORENSIC DNA TESTING: PREDICTIONS OF THE RESEARCH AND DEVELOPMENT WORKING GROUP 25 (2000).

59. The likelihood ratio for the partly matching profiles is essentially zero for a parent-child pair and about 1/500 for half siblings as compared full siblings. *Id.*

60. *E.g.*, Murphy, *Relative Doubt*, *supra* note 13, at 297–98; Suter, *supra* note 7, at 311.

61. See MICHAEL J. GRAETZ & IAN SHAPIRO, DEATH BY A THOUSAND CUTS: THE FIGHT OVER TAXING INHERITED WEALTH 4, 14 (2006) (finding the labeling of the estate tax as the “death tax” to be an effective strategy for its opponents).

62. KRIMSKY & SIMONCELLI, *supra* note 11, at 88.

STR loci used for database profiles, they are not large enough to generate useful leads to most relatives.<sup>63</sup> With databases of individuals of unknown pedigrees, kinship analysis with small numbers of loci generally is only effective for first-degree (parent-child and full sibling) relationships.<sup>64</sup>

With more and better loci, police could zero in on the correct first-degree database inhabitant, but even then, ascertaining and making use of more attenuated relationships rarely would be feasible. In a database that includes very specific lineage haplotypes (the genomes of the Y chromosome and the mitochondria), it would be possible to pick out a database inhabitant who is distantly related to a person who deposited DNA while committing a crime. However, these lineage markers only indicate the existence of a common ancestor at some time in history. They would not enable police to infer that a crime-scene sample comes from, say, a second cousin of a database inhabitant as opposed to someone else in a lineage stretching back an unknown number of generations.<sup>65</sup>

To avoid the connotation that kinship analysis in databases typically will go beyond the detection of first-degree relatives, I use the less tendentious terms “near-miss matching” or “kinship matching” for outer-directed trawls.<sup>66</sup> This terminology also clarifies the FBI’s slightly confusing effort to distinguish between “partial match” searching and “familial searching.”<sup>67</sup> One would think that partial-match searching simply denotes any method of identifying possible suspects in a criminal investigation short of a full match.<sup>68</sup> As such, “partial match searching” encompasses both inner- and outer-directed trawls (see Table 2). Existing CODIS software allows trawls for “low” or “moderate stringency” partial

---

63. See Stefan Wilkening et al., *STR Markers for Kinship Analysis*, 78 HUMAN BIOLOGY 1, 5–7 (2006) (finding current methods and STR markers in use are suitable only in detecting very closely related persons).

64. Steven P. Myers et al., *Searching for First-degree Familial Relationships in California’s Offender DNA Database: Validation of a Likelihood Ratio-based Approach*, 5 FORENSIC SCI. INT’L: GENETICS 493 (2011).

65. If the crime-scene sample was sufficient in quality and quantity, testing to produce high density SNP data would permit an inference as to the degree of the relationship (within one degree of the true relationship). See Chad Huff et al., *Maximum-likelihood Estimation of Recent Shared Ancestry (ERSA)*, 21 GENOME RESEARCH 768 (2011) (demonstrating that the ERSA’s statistical power approached the maximum theoretical limit imposed by the fact that distant relatives frequently share no DNA through a common ancestor). But beyond excluding many of the dataset inhabitant’s relatives as a likely source, how useful would it be to learn that the source probably is, say, a fifth, sixth, or seventh degree relative of the database inhabitant?

66. “Genetic proximity testing” is another phrase used in bioethics literature for outer-directed trawls. Prainsack, *supra* note 6, at 28–29.

67. See *Familial Searching*, FBI, <http://www.fbi.gov/about-us/lab/codis/familial-searching> (last visited Nov. 26, 2012) (defining that partial matching as “the spontaneous product of a regular database search where a candidate offender profile is identified as a possible close relative because of a similarity in the number of alleles shared between the two profiles?”).

68. A full match is a match to every allele at every locus in the crime-scene samples, though not necessarily in the database profiles. Some samples are too degraded to allow successful typing of all 13 core loci adopted for profiles in the U.S. databases. A seven-locus crime-scene profile, for instance, might produce one or more cold hits in a database trawl, making those database inhabitants suspects in the case. *Frequently Asked Questions (FAQs) on the CODIS Program and the National DNA Index System*, FBI, <http://www.fbi.gov/about-us/lab/codis/codis-and-ndis-fact-sheet> (last visited Nov. 26, 2012).

**Table 2. Types of trawls in a DNA database**

<b>Inner-directed</b> (focuses on database inhabitants)	<b>Outer-directed</b> (focuses on people outside the database)
<i>Full matching:</i> The database inhabitant's profile contains every allele detected in the crime-scene sample. As a result, the database inhabitant becomes a suspect.	
<i>Partial matching:</i> The database inhabitant's profile contains alleles not detected in the crime-scene sample, but these discrepancies might be due to ambiguities in mixed crime-scene samples or differences in the system used to profile the crime-scene sample. Database inhabitants are the only targets of the trawl, although inner-directed partial matches could lead investigators who exclude the database inhabitant as a suspect to consider relatives.	<i>Partial, near miss, kinship, genetic proximity, or familial matching:</i> The database inhabitant's profile differs from the one detected in the crime-scene sample in a way that excludes the inhabitant as a possible source but that is much more probable when a parent, child, brother, or sister outside the database is the source (as compared to an unrelated individual).

matches to deal with ambiguities resulting from mixed samples, technical problems, or variations in STR-typing kits from different manufacturers.<sup>69</sup> These partial matches are still the product of inner-directed trawls designed to pick out database inhabitants as possible sources. The CODIS partial matches are merely based on counting alleles without considering their frequencies in the population. Such matching does not use kinship analysis and is ill-suited for outer-directed searching for relatives.<sup>70</sup> For this latter purpose, kinship analysis is more effective (although it can be combined with allele counting rules of the kind used in CODIS partial matching).<sup>71</sup>

### C. *The Efficacy of Kinship Matching*

The distinction between near-miss searching for relatives through kinship matching, on the one hand, and partial-match searching for database inhabitants,

69. KRIMSKY & SIMONCELLI, *supra* note 11, at 68.

70. Ge et al., *supra* note 45, at 1448; Scientific Working Group on DNA Analysis Methods Ad Hoc Committee on Partial Matches, SWGDAM Recommendations to the FBI Director on the "Interim Plan for the Release of Information in the Event of a 'Partial Match' at NDIS," *Forensic Sci. Commun.*, Oct. 2009, [http://www.fbi.gov/hq/lab/fsc/current/standard\\_guidelines/swgdam.html](http://www.fbi.gov/hq/lab/fsc/current/standard_guidelines/swgdam.html).

71. BUTLER, *supra* note 25, at 605–06; Ge et al., *supra* note 45, at 1448; Myers et al., *supra* note 64, at 499. This is not to deny that partial-match searching as now practiced can be used, albeit inefficiently, to trawl for relatives. Many agencies are at sea when it comes to reporting out a reduced stringency search that also strikes an analyst as probative of kinship. For reviews of current, chaotic practices, see Ram, *supra* note 18, at 767; Suter, *supra* note 7, at 324–25.

on the other, is important because only the former holds the promise of generally productive trawls for suspects outside the database.<sup>72</sup> As noted in the introduction, a number of scientists have estimated kinship matching could generate thousands of useful investigative leads nationally.<sup>73</sup> That is a theoretical calculation, however, and direct data on the effectiveness of near-miss searching for relatives through kinship matching (outer-directed trawling) are extremely limited.<sup>74</sup> Because the technique is almost never used, proponents and opponents tend to rely on their choice of anecdotes about its value and accuracy.<sup>75</sup> *Relative Doubt* exemplifies this strategy. On the one hand, it acknowledges California's spectacular success in using kinship matching to locate a man alleged to be the Los Angeles "Grim Sleeper" murderer.<sup>76</sup> On the other hand, *Relative Doubt* emphasizes "one revealing fact: [Denver's District Attorney, Mitch Morrissey's] familial searches did not work. None of the three matches turned out to point toward a relative, much less the source, of the actual crime-scene sample . . . [I]t failed in three separate cases . . ."<sup>77</sup> However, these searches did not use a matching strategy designed and optimized to detect kinship. When Morrissey experimented with more appropriate software (known as DNA-VIEW) on Denver's local database, he obtained "the first [successful hit] ever [in] a deliberate familial search in the United States."<sup>78</sup>

---

72. Murphy, *Relative Doubt*, *supra* note 13, at 341–42, asserts that law enforcement personnel will engage in "strategic behavior" to perform inner-directed partial matching to generate leads to relatives; accord Ram, *supra* note 18, at 783–84. To prevent this, they would forbid a laboratory from advising police they should consider the possibility that a particular type of relative from outside the database is the source of a crime-scene sample. Murphy, *Relative Doubt*, *supra* note 13, at 345. Professor Murphy also suggests her proposed policy of silence about plausible leads is comparable to the policy geneticists doing family studies of disease inheritance follow when they discover misattributed paternity. The analogy is unpersuasive, since there is no large social cost to maintaining confidentiality in this research context. In the criminal context, the cost of withholding information exceeds the cost of investigating a potential suspect when kinship clearly is probable. A better argument against "accidental" kinship matching would be that, without formal kinship analysis, the suspicions of laboratory personnel might not be justified.

73. See Bieber et al., *supra* note 5, at 1316 (noting kinship searching has resulted in 30,000 cold hits with the potential for thousands more).

74. See Murphy, *Relative Doubt*, *supra* note 13, at 347 (suggesting procedures that could produce useful statistics on efficacy).

75. One incident that has largely escaped notice involved an ad hoc database, that is, a "DNA dragnet" in a small town in Alberta, Canada. There were no full matches, but two partial matches led police to the son of one of those men, whose DNA did match and who was convicted in 2005. Douglas Quan, *Familial DNA Searches Raise Tough Ethical Questions*, THE STARPHOENIX (Sasjatoon, Can.), May 21, 2011, available at <http://www2.canada.com/saskatoonstarphoenix/news/national/story.html?id=bf5cbe09-392e-4a76-9bdc-b931a18a4605&p=2>.

76. A more recent arrest in the Santa Cruz Coffee Shop rape case makes California's hit rate for its approach to kinship trawling two for thirteen. Stephen Baxter, *Rare, Familial DNA Used in Case of Rape and Robbery at Santa Cruz Coffee Shop*, SANTA CRUZ SENTINEL, Mar. 15, 2011, [http://www.santacruzsentinel.com/ci\\_17618326](http://www.santacruzsentinel.com/ci_17618326).

77. Murphy, *Relative Doubt*, *supra* note 13, at 293.

78. Raphaelle Moreau-Horwin, *States Using Familial Searches*, DNA FORENSICS LLC (Aug. 24, 2012), <http://www.dnaforensics.com/StatesAndFamilialSearches.aspx>. For more details on the case, see *Car Vandal Nabbed Thanks To Brother's DNA*, WUSA9 (Nov. 17, 2009), <http://www.wusa9.com/news/local/story.aspx?storyid=93841> (reporting Luis Jaimes-Tinajero's "brother was in the system because of a felony conviction for auto theft" and the "car was one of many at his Denver apartment complex broken into [that] morning").

So, is Morrissey's record 0 for 3 or 1 for 1? It is neither. Surely, there were many cases in which the Denver kinship searches drew blanks or false leads. But the usual, full-match searches do not always produce cold hits either. The latter hit rate depends on the fraction of perpetrators of the crimes who are database inhabitants. The former hit rate is more complicated. It depends on (1) sensitivity—that is, the probability that the kinship matching algorithm registers a hit when it processes the DNA profile of a true relative of the source of the crime-scene DNA; and (2) prevalence—the proportion of relatives of database inhabitants (a) who commit crimes with recovered DNA traces and (b) who are not themselves database inhabitants.<sup>79</sup> The sensitivity of various algorithms can be estimated,<sup>80</sup> but the prevalence is more uncertain. In a 1996 study, nearly half of the jail inmates surveyed stated they have close family members who have been incarcerated,<sup>81</sup> but we cannot know how many as yet untested close relatives of those incarcerated individuals have committed crimes for which DNA evidence is on file. Furthermore, with maturing databases, the chance of a criminal's relative being a database inhabitant will grow. When large databases have been in existence long enough, “[a] father's profile could lead to a son's apprehension as the younger man begins a life of crime.”<sup>82</sup> Nevertheless, until a state implements and tracks the results of a well designed form of kinship matching in a large number of cases, the real-life efficacy of the technique will not be known.<sup>83</sup>

Neither will its costs. These costs, clearly articulated in *Relative Doubt*, include indignities to innocent suspects and concomitant distress to loved ones, diverting

---

79. Category (b) reflects the fact that if two database inhabitants are close relatives with slightly different profiles and if one is the source of the crime-scene sample, the latter will emerge as an ordinary cold hit. Kinship matching will not be needed in this case, and the fact that the individual with the full match has a relative with a kinship hit will not be of interest. Erin E. Murphy, *Familial DNA Searches: The Opposing Viewpoint*, 27 CRIM. JUST. 1, 19, 21 (2012) [hereinafter Murphy, *Opposing Viewpoint*]. However, this is not to say that all the kin of database inhabitants who have a “shadow presence” in the database are “law-abiding relatives of offenders” (even as ascertained by the absence of a criminal conviction). *Id.* Some will have criminal records for lesser offenses that do not trigger DNA collection, and others will have completed their sentences.

80. Bieber et al., *supra* note 5; James M. Curran & John S. Buckleton, *Effectiveness of Familial Searches*, 48 SCIENCE & JUSTICE 164 (2008); Ge et al., *supra* note 45; Myers et al., *supra* note 64; Kristen Lewis O'Connor et al., *Evaluating the Effect of Additional Forensic Loci on Likelihood Ratio Values for Complex Kinship Analysis*, PROCEEDINGS OF THE 21ST INTERNATIONAL SYMPOSIUM ON HUMAN IDENTIFICATION (2010), available at <http://www.promega.com/products/pm/genetic-identity/ishi-conference-proceedings/proceedings-index-home/>; Thomas M. Reid et al., *Use of Sibling Pairs to Determine the Familial Searching Efficiency of Forensic Databases*, 2 FORENSIC SCI. INT'L: GENETICS 340 (2008).

81. Bieber et al., *supra* note 5, at 1316 (“46% of jail inmates indicated that they had at least one close relative who had been incarcerated.”).

82. BUTLER, *supra* note 25, at 608.

83. Various figures have been quoted. The most recent for the United Kingdom—which uses only 10 loci—is 27% as of 2010. *Id.* at 606–07. See also KRIMSKY & SIMONCELLI, *supra* note 11, at 81 (“Proponents . . . in the United Kingdom boast a 90 percent success rate for those cases where it has been employed.”); *id.* at 175 (“By the beginning of 2008, the United Kingdom analyzed 148 cases using familial searching techniques; only 15 of them had been resolved with 9 convictions.”); Murphy, *Relative Doubt*, *supra* note 13, at 301 n. 49 (quoting statistics from familial searching in the United Kingdom and New Zealand which resulted in eighteen matches and thirteen convictions out of seventy searches since 2004).

police resources to unprofitable leads, and adding more DNA tests to produce better profiles for outside-the-database searches. Deciding whether the game, especially with current technology, is worth the candle is beyond the scope of this article. My primary focus is on the mildly futuristic system of kinship analysis with profiles constructed to be useful in identifying crime-scene sources outside the databases—that is, in outer-directed as well as inner-directed trawls. In this system, there would be more loci in the database. In fact, this enhancement may be on its way, as seven new autosomal STR loci have been proposed for inclusion in the CODIS “core loci.”<sup>84</sup> Moreover, loci on the sex chromosomes and mitochondria could be part of the profiles.<sup>85</sup> The additional, digitally searchable data would allow kinship analysis programs (that analyze the pattern of matching alleles within partially matching profiles and the population frequencies of these alleles as in Table 1) to filter out efficiently most database inhabitants who are unrelated to the source of the crime-scene DNA.<sup>86</sup>

This system changes the perspective of current thinking about kinship matching slightly. Because existing databases are confined to autosomal STRs at only 13 loci, near-miss matches require supplemental testing if the DNA itself is to weed out false leads. For example, California’s restrictive trawling rules require laboratory technicians to retain and reanalyze samples in the repositories to verify that the paternally inherited Y-STRs are the same in the crime-scene and the data-banked samples of males.<sup>87</sup> Recording such data at the outset avoids this extra step and, with it, the incentive to retain the physical samples that contain a wealth of personally sensitive information.<sup>88</sup> Adding mitochondrial sequence data at the same time would enhance detection of maternally related sources. And using carefully chosen SNPs in place of or in addition to existing STRs would improve the resolution of possible kinship.<sup>89</sup>

---

84. Douglas R. Hares, *Addendum to Expanding the CODIS Core Loci in the United States*, 6 FORENSIC SCI. INT’L: GENETICS e135 (2012); Douglas R. Hares, Letter to the Editor, *Expanding the CODIS Core Loci in the United States*, 6 FORENSIC SCI. INT’L: GENETICS, at e52 (2012).

85. Mitochondria are organelles outside the nucleus that are maternally inherited. Just as Y-STRs (or other Y-chromosome loci) help distinguish paternal lineages, mitochondrial DNA is a maternal lineage marker. KAYE, DOUBLE HELIX, *supra* note 21.

86. See Greely et al., *supra* note 14, at 254 (stating with “roughly twenty more markers . . . similar to the thirteen CODIS markers . . . there would be few, if any, spurious matches to waste an investigator’s time”); O’Connor et al., *supra* note 80 (modeling the effect of using 20 and 40 STR loci in outer-directed trawling). Professor Ram is pessimistic about the current specificity of outer-directed searches. However, she discusses almost none of the relevant research reported. Ram, *supra* note 18, at 764–65.

87. Information Bulletin from Edmund G. Brown, Jr., Attorney General, DNA Partial Match (Crime Scene DNA Profile to Offender) Policy No. 2008-BFS-01 (2008).

88. See David H. Kaye, *Behavioral Genetics Research and Criminal DNA Databases*, 68 LAW & CONTEMP. PROBS. 259 (2006) (outlining arguments for and against sample retention) [hereinafter Kaye, *Behavioral Genetics*]. Professor Murphy maintains, without explanation, that sample retention lacks justification and serves no purpose. Erin Murphy, *Paradigms of Restraint*, 57 DUKE L.J. 1321, 1409 (2008).

89. C. Phillips et al., *Resolving Relationship Tests That Show Ambiguous STR Results Using Autosomal SNPs as Supplementary Markers*, 2 FORENSIC SCI. INT’L: GENETICS 198 (2008).

But even if one concludes these enhancements would produce significant benefits in terms of more accurate leads and reduced supplemental search costs, the constitutional and other questions of legality and political morality of using databases to target outside individuals remain to be addressed. I turn now to the principal constitutional arguments.

## II. EQUAL PROTECTION

The Fourteenth Amendment mandates every person receive the equal protection of the law. Arbitrary classifications, which treat similarly situated people differently, or differently situated people equally, are impermissible. But almost all laws rest on imperfect generalizations about people and the problems the laws are designed to address. Even substantial under- or overinclusiveness in a legislative classification is ordinarily constitutionally tolerable, and a mere “rational basis” to draw the line where the legislature chooses will justify the law. Last year, for example, New York expanded the scope of its DNA database law to encompass virtually all convicted criminals.<sup>90</sup> Arguably, the earlier law was underinclusive; or perhaps the newer one is overinclusive. But neither possibility makes the legislation irrational. A line must be drawn somewhere, and legislatures ordinarily have great latitude in deciding where to draw it and when to move it.

But not always. For example, racial classifications, which have long reflected animosity toward certain groups, are especially likely to constitute unjustifiable disparate treatment. Accordingly, laws establishing overt racial classifications, and laws that are neutral on their face but are “administered . . . with an evil eye and an unequal hand”<sup>91</sup> to benefit one race at the expense of another, are constitutionally suspect. Such state action demands the most compelling justification.

Applying these general principles<sup>92</sup> to outer-directed trawling demonstrates that the procedure satisfies the Equal Protection Clause. As explained below, uniformly pursuing investigative leads from the genetic data at hand is neither a racial classification nor an irrational one.

### A. Racial Discrimination

DNA databases have a greater effect within some racial groups than others. Both full-match and near-miss searching affect a higher proportion of racial minorities (relative to their representation in the general population) simply because these groups are disproportionately arrested and incarcerated, and hence included in law

---

90. Press Release, Governor Andrew M. Cuomo, Governor Cuomo Signs Law to Expand New York’s DNA Databank, Mar. 19, 2012, <http://www.governor.ny.gov/press/03192012-Dna-Databank>.

91. *Yick Wo v. Hopkins*, 118 U.S. 356, 373–74 (1886).

92. The thumbnail sketch here is incomplete. For additional complications, categories, and nuances, see, for example, Kenji Yoshino, *The New Equal Protection*, 124 HARV. L. REV. 747 (2011).