February 27, 2008

Race, Genes, and Justice: A Call to Reform the Presentation of Forensic DNA Evidence in Criminal Trials

Jonathan Kahn, *Hamline University School of Law*

Available at: http://works.bepress.com/jonathan_kahn/1/
Race, Genes, and Justice: A Call to Reform the Presentation of Forensic DNA Evidence in Criminal Trials

Jonathan Kahn, J.D., Ph.D.
Associate Professor
Hamline University School of Law
I. Introduction: Race, Genes, and Justice ................................................................. 3

II. Race and the Early Development of Forensic DNA ........................................... 11
   A. Origins of Forensic DNA Testing ................................................................. 11
   B. Early Questions and Challenges ................................................................. 14
   C. NRC II: Questions of Race Laid to Rest? .................................................. 23

III. Current Standard Practices Regarding Race and Forensic DNA Analysis ...... 27
   A. The Impact of NRC II ............................................................................... 27
   B. CODIS and the Move From VNTRs to STRs ........................................... 29

IV. Race, Technology, and “Care of the Data” ...................................................... 32
   A. Race v. Technology ................................................................................. 32
   B. Social v. Genetic “Race” ......................................................................... 35
   C. The Obvious Solution: A Non-Racial, General Population Data Base ....... 38
   D. The Inertial Power of Race ...................................................................... 40

V. Presenting Race and Forensic DNA in Court .................................................. 47
   A. Early Cases: DNA in Flux ....................................................................... 48
   B. From Pizarro to Wilson: Race in Flux .................................................... 53

VI. Race, Genes, and Relevance ......................................................................... 59
   A. When Race is not Relevant ..................................................................... 60
   B. When Race is not Reliable ....................................................................... 63
   C. Race, Genes, and Prejudice .................................................................... 67

VII. Conclusion .................................................................................................. 73
I. Introduction: Race, Genes, and Justice

How and when, if at all, is it appropriate to use race in presenting forensic DNA evidence in a court of law? In October, 2002, a California jury convicted William Curtis Wilson of first degree murder with use of a dangerous weapon during commission of an attempted rape and a lewd act upon a child. The court sentenced him to a term of life in prison without possibility of parole. DNA evidence played a central role in obtaining the conviction. This, in itself, is neither extraordinary nor unusual given the broad acceptance of the use of DNA evidence in courts across the country and, indeed, around the world. The case is noteworthy, however, for its discussion of the appropriate use of racially identified forensic DNA data bases in calculating the odds that DNA left at the scene of the crime by the perpetrator might be that of the defendant. The crime scene DNA was found to match the defendant Wilson’s DNA at nine distinct loci, or specific

1 I do not attempt to provide a set definition of “race” in this article. Rather I focus primarily on how actors in specific legal and scientific contexts have used the term. In the interests of economy and manageable syntax, in the remainder of this article I will often refer only to “race” when speaking generally of racial and ethnic categories. I am assuming both to be socially constructed categories that nonetheless have come to have biological implications as they play out in real world biomedical and forensic contexts. I will use the terms “race” and/or “ethnic” when referring to specifically marked groups. Thus, for example, the U.S. Census codes “White” or “Asian” as racial categories and “Hispanic” or “Latino” as ethnic categories. More information on the U.S. Census Bureau’s race and ethnic classifications is available at http://www.census.gov/population/www/socdemo/race/racefactcb.html (last visited: February 14, 2008). In the context of forensic practice, Hispanic is also sometimes referred to as a racial group. See Madeleine J. Hinkes, Race, Ethnicity and Forensic Anthropology: Realities of Racial Determinationin a Forensic Setting, 13 NAPA BULLETIN 49 (1993) (explaining that one’s “social race or ‘ethnicity’ may not be the same as [one’s] biological race. For example, self-proclaimed ‘Hispanic’ individuals may be Spanish (Caucasoid), Mexican (American Indian), Puerto Rican (Negroid admixture), or Filipino (Asian Mongoloid).”). Ethnic groups are often also discussed as sub-groups within races. For example, Italian or Irish might be understood as ethnic sub-groups within the racial category of Caucasian.

2 People v. Wilson, 21 Cal.Rptr.3d 102, 106 (2004).

points on the genome. The question then became what were the odds that someone else might share the defendant’s same genetic profile? Such odds are known as a Random Match Probability (RMP) – the probability of finding the same DNA profile identified in the crime scene sample in a randomly selected, unrelated individual.  

In calculating the RMPs, Nicola Shea, a criminalist with the California Department of Justice's Sacramento laboratory, found that Wilson’s genetic profile would be expected to occur in one of 96 billion Caucasians, one of 180 billion Hispanics, and one of 340 billion African-Americans.  

Shea noted that these profiles were extremely rare; after all, the world contains only about six and a half to seven billion human beings.  

On appeal, Wilson’s attorneys strenuously contested this use of race-specific DNA data bases to calculate odds to assist the trier of fact in reaching a verdict. Invoking the 2003 California Supreme Court decision in People v. Pizarro, they argued that the presentation of such race-specific odds was permissible only when the race of the perpetrator was known. Otherwise, they contended, the use of such evidence lacked sufficient evidentiary foundation because it was based on the improper assumption that the defendant was in fact the perpetrator.  

In July, 2006, the California Supreme Court rejected this argument, finding that the introduction of evidence of the odds of a DNA match calculating using race-specific data bases from major racial/ethnic groups represented in the local population was acceptable – thereby effectively overturning its recent ruling in Pizarro.

---

1. Butler, supra note 1, at 481, 486. Lewontin & Hartl, supra note 21, at 1746. That the individual be “unrelated” is significant because related individuals will have a higher likelihood of sharing a greater percentage of DNA, hence altering the probabilities of a random match.  
and focused site for the examination of debates relating to the use of racialized data bases in forensic DNA analysis.\(^9\) *Pizarro* involved an appeal from a case originally tried in 1990, when forensic DNA testing was still in its infancy. At the original trial, the forensic expert for the state testified that the odds of the defendant’s genetic profile occurring other population groups ranged from one in 250,000 in “Hispanics” up to one in 10,000,000 in “Caucasians”\(^10\) (frequencies in other racial populations were not presented to the jury).\(^11\) Here the difference produced by using racially marked data bases may be significant – not only statistically, but also as a legal matter. The lower bases may be significant – not only statistically, but also as a legal matter. The lower

\(^9\) At the outset, it is important to note that these questions are distinct from (yet ultimately related to) current heated debates concerning the use of DNA technologies to construct racial profiles of potential suspects from DNA samples left by an unknown perpetrator at the scene of a crime. See, e.g., Frederick Bieber, *Science and Technology of Forensic DNA Profiling: Current Use and Future Directions*, reprinted in DAVID LAZER, DNA AND THE CRIMINAL JUSTICE SYSTEM, 23-64 (David Lazer, ed., MIT Press 2004); Pilar N. Ossorio, *About Race, Forensic Genetic Testing for Race and Visible Traits*, 34 J. L., MED. & ETHICS 277-92 (2006). Such technology is largely prospective, being used in law enforcement contexts by investigating police authorities. It uses race to look forward to aid in the apprehension of a possible suspect. Bieber, supra, at 36-38. In contrast, the use of race-specific data bases at trial is largely retrospective. It is used by prosecutors to link an already apprehended suspect back to the crime by matching his or her DNA to a sample found at the crime scene. See, e.g., Edward Imwinkelreid, *The Relative Priority that Should Be Assigned to Trial Stage DNA Issues*, reprinted in DAVID LAZER, DNA AND THE CRIMINAL JUSTICE SYSTEM 91, 92-100 (David Lazer, ed., MIT Press 2004). The basic technology used is also often different. As discussed further below, courtroom uses of DNA involve constructing statistical probabilities for matches between two existing samples of DNA – the defendant’s and the crime scene sample. See infra notes 169-172, 186-87, 191 and accompanying text. Racial profiling for unsolved crimes generally involves examining a crime scene DNA sample for “Ancestry Informative Markers” (AIMS), which (though highly controversial) are believed by some scientists to provide indications of the likely ancestry, or mixture of ancestries of the human source of the sample. See, e.g., Mark Shriver, Tony Frudakis & Bruce Budowle, *Getting the Science and the Ethics Right in Forensic DNA Analysis*, 37 NATURE GENETICS 449-50 (2005). For a critique of this technology in the forensic context, see Mildred Cho & Pamela Sankar, *Forensic Genetics and Ethical, Legal and Social Implications Beyond the Clinic*, 36 NATURE GENETICS S8-12 (2004), and idem, *Reply to “Getting the Science and the Ethics Right in Forensic DNA Analysis,”* 37 NATURE GENETICS 450-51 (2005). These technologies and their forensic uses, though in many respects different, do share a common propensity to reify race as genetic. Such a propensity is both (a) socially dangerous, threatening to recreate or revitalize stigmatized conceptions of biologically superior and inferior races. See, e.g., Sandra Lee et al., *The Meaning of ‘Race’ in the New Genomics: Implications for Health Disparities Research*, 1 YALE J. HEALTH POL’Y L. & ETHICS 33, passim (2001); Troy Duster, *Race and Reification in Science*, 307 SCIENCE 1050, 1050-51 (2005); and b) such a propensity is also scientifically suspect, based on highly problematic assumptions about the correlations between race and biological ancestry and disregarding current broadly held understandings of the incoherence of race as a genetic concept. See infra notes 109, 116, 251, 257 and accompanying text.

\(^10\) *Pizarro* II, 3 Cal.Rptr.3d at 98.

\(^11\) *Id.* at 103.
denominators in this situation may be understood in large part as a function of the more rudimentary techniques for DNA analysis at the end of the 1980s. Race seemed relevant because it appeared to refine the results of a newly developing and still relatively crude technology. Hence, race became instantiated at the outset of forensic DNA analysis as a basic framework for presenting data. By the time of Wilson’s case, however, the technology had developed to such an extent that it was regularly capable of producing odds ratios on the order of one in the hundreds of billions. With such odds, the practical utility of distinguishing RMPs by race disappears. Nonetheless, race has remained ingrained in the framework of the production and interpretation of forensic DNA evidence.

This article considers how relations science and law must continually be reevaluated in light of changing social and technological developments. It questions the underlying assumption of the utility of race itself in forensic DNA analysis. Beginning with a review of the development of forensic DNA technology, it examines what race adds as a practical matter to the ability of a finder of fact to make fair and accurate decisions. Any such value must then be weighed against the potential dangers of bias created by introducing issues of race as genetic into the context of what is usually a violent crime. It argues that in most cases such evidence should be excluded as irrelevant, or if deemed relevant it should be held inadmissible because the dangers of infecting the proceedings with racial prejudice outweigh any possible benefit that introducing the race-based statistics could provide. This is not necessarily because the

12 See infra notes 13, 36, 39, 87, 151 and accompanying text.
mere mention of race will inevitably taint a jury’s deliberations, but rather because the benefits of using race in such context must be deemed so *de minimus* as to be incapable of outweighing even a remote danger of racial prejudice. Moreover, when examining the implications of connecting race and genetics in the context of violent crime, this article will consider how oftentimes such a danger may be far from remote. The article concludes that in most cases, given the current state of forensic DNA technology, there is no longer any justification (if there ever was any) for using race specific data bases in presenting DNA evidence to a jury. Ironically, given the power of current technology to generate powerful RMPs, adopting the recommendation to abandon the use of race in the presentation of forensic DNA evidence would not materially affect the ability of prosecutors to obtain convictions based on DNA evidence. It would, however, transform the way in which race, genes and violent crime are associated in the criminal justice system.

*   *   *   *

Part II of this article will provide some background to the history and technical aspects both of forensic DNA analysis and of debates regarding the relation between race and genetics. It will begin by introducing some of the details of early forensic DNA practices and procedures. It will also examine early debates over the appropriate use of racial categories in generating statistical match probabilities. In the early years of forensic DNA analysis, the late 1980s and early 1990s, these debates did not involve questions of *whether* to use race, but rather *how much* race to use. The concern of some early forensic DNA analysts was that using too broad a population group would generate odds of a match that were unfairly low. Advocates of using broad racial categories
argued that they provided a pragmatic and useful means to generate more accurate statistics than using an undifferentiated general population data base, while still providing odds that were fair to a defendant. Critics of this approach argued that since there is more genetic variation within racial groups than between them, data-bases should be characterized by smaller sub-groups within the larger racial designation in order to provide even better information about RMPs. This part of the article will also consider how these arguments relate to broader issues concerning the nature of race in relation to genetics. This section concludes with an examination of how the debates over using race to calculate RMPs were ultimately settled by the mid-1990s in favor of using broad racial categories which then became the norm for forensic DNA practice in the United States.

Part III will examine the current standards and protocols for conducting forensic DNA analysis and consider how, where, and when forensic experts inject race into their practices. This section begins with an examination DNA data bases, focusing on the FBI’s Combined DNA Index System (CODIS), which began as a pilot project in 1990 and has since evolved into a major data base of with over 5 million DNA profiles from convicted offenders nationwide. It reviews technological developments that have greatly increased the power and efficiency of forensic DNA analysis, and considers the selection of 13 genetic loci that have become the standard for generating DNA matches and calculating random match probabilities.

Part IV contrasts the elaboration and standardization of such technical protocols for DNA analysis with the protocols, (or lack thereof) for producing and using racial and ethnic categories in forensic DNA analysis. Here I argue that the use of racial categories

---

is woefully under-conceptualized and wholly inadequate – especially when contrasted with the great care taken to elaborate the technical protocols DNA analysis itself. I argue that similar care of the data should be given to methods for using racial categories in a genetic context. I put forward the use of general, non-racial reference population data bases as the obvious solution to this problem. This section concludes with an examination of what I call the “inertial power of race” to remain in a system of practice and analysis long after the initial reasons for using it have faded. It considers, as well, some of the dangers, beyond the court room, of allowing race to persist in a context that inappropriately reifies it as genetic.

Part V comes back to People v. Wilson and related cases to examine the current state of how race is used in presenting forensic DNA data in courts of law. The focus here is primarily on developments in California case law as a case study that provides a particularly good site of analysis and raises issues that are relevant to understanding how racialized data is used in presenting DNA evidence in courts across the country. A close analysis of these cases provides a clear and factually specific context for understanding broader issues of how any why racialized genetic data is used in today’s criminal justice system.

Part VI develops a specific critique of current uses of racialized DNA evidence and elaborates the argument that in most cases race-specific genetic data that is used to generate RMPs should not be admitted in court. This section begins with a brief consideration of the court’s gate-keeping function in evaluating scientific evidence under Daubert. Using the Federal Rules of Evidence as a framework, it then considers the basic question of what does race add to RMPs generated through forensic DNA analysis?

More specifically, given that it is possible, with current technology, to generate astronomically low RMPs without using race, it questions whether any legally relevant information of practical use to the finder of fact added by introducing race into the calculation. The problem here is to consider whether such data passes the threshold for legal relevance in determining its admissibility. Here I assert that courts must be careful to distinguish between statistical relevance and legal relevance in evaluating such evidence. I argue that in many situations the different RMP’s produced by referencing distinct race-specific data-bases provide no useful additional information to the trier of fact and so should be excluded. I conclude that the only thing race adds to the proceedings is race itself. It associates race, genes, and violent crime in a manner wholly irrelevant to the determination of a particular defendant’s guilt or innocence.

I then move on to argue that even if such evidence is deemed relevant, it should nonetheless be excluded as prejudicial. Here I weigh any possible benefits to using race-specific RMPs in presenting evidence vs. the dangers of injecting the proceedings with unfounded associations of race, genes and crime which threaten to evoke attitudes among the jury, even if unconscious, that are improperly tainted with racism. In developing this section, I consider studies that examine the particular psychology involved in presenting DNA evidence to a jury and bring that work into dialogue with studies on the psychology of implicit prejudice – tacit attitudes towards race that are often held without an individual’s own conscious awareness of how they influence his or her perceptions of and responses to racialized subjects. Together, the persuasive authority of DNA evidence and the reality of implicit prejudice call into question the legitimacy of using racialized DNA evidence.
Concern over the prejudicial impact of such evidence gains added force when we situate those rather abstract psychological dynamics of implicit prejudice in the context of the highly racialized nature of the United States’ criminal justice system today. Here I argue that the pervasive racialization of violent crime in the United States takes such concerns out of the realm of mere speculation and gives concrete cause for concern. Given the de minimus nature of any possible practical utility to be gained from introducing race-specific RMP’s in most cases, I conclude that the prejudicial potential of such evidence clearly outweighs any possible benefits it might provide.

The article concludes with a brief synthesis of the arguments for ending the practice of using race frame the presentation of forensic DNA evidence. It notes this would not materially hinder the ability of prosecutors to obtain convictions using DNA evidence. Yet, by removing the gratuitous introduction of race into a context of genetics and violent crime, such reform would promote a positive and significant reorientation of the relation among race, genes, and justice.

II: Race and the Early Development of Forensic DNA

A. Origins of Forensic DNA Testing

DNA is made up sequences of four nucleotides: adenine, cytosine, guanine, and thymine – commonly represented as A, C, G, and T. Each nucleotide base is paired through a process known as hybridization: A is always paired with T; C is always paired with G. There are approximately three billion of these “base pairs” in the human genome. There are two major steps in using DNA for purposes of forensic identification. First, a sample left at the crime scene by the perpetrator is compared to a

16 Butler, supra note 11, at 18-20.
sample from a suspect. Second, if there is a “match” then statistics must be used to calculate to frequency of that DNA “profile” in an appropriate reference population.\(^\text{17}\) This latter step is required because, although every person’s DNA is unique, it is impractical to compare the full three billion nucleotide base pairs between two samples for forensic purposes. Therefore, two samples will be compared only at a limited set (usually between 4 and 13) of “loci,” or specific parts of the genome. For this practice to be effective, it is necessary to find loci that are highly variable between individuals and test only for them. Humans, however, are essentially identical in about 99.5% of their DNA.\(^\text{18}\) Finding the specific points of variation among individuals, therefore, can be difficult.

In 1985, English geneticist Alec Jeffreys first described a method for developing a DNA “profile” or a person in a manner that might be used for purposes of forensic identification.\(^\text{19}\) Jeffreys’ innovation consisted in observing that, in particular regions of the human genome, short segments of DNA – the ACGT nucleotide sequence – are repeated between 20 to 100 times.\(^\text{20}\) These repeat regions became known as “variable number of tandem repeats” or VNTRs. Different VNTR “alleles” – or variations – are composed of different numbers of repeats. In order to examine and visualize the VNTRs, Jeffreys employed a technique known as restriction fragment length polymorphism (RFLP) which uses a restriction enzyme to cut the regions of DNA surrounding the


\(^{18}\) Older analyses typically put the figure at 99.9% but a more recent study indicates that 99.5% may be a more accurate finding. See Rick Weiss, Mom’s Genes or Dad’s? Map Can Tell, WASH. POST, Sept. 4, 2007, at A01, available at http://www.washingtonpost.com/wp-dyn/content/article/2007/09/03/AR2007090301106.html (last visited Feb. 19, 2008).

\(^{19}\) Butler, supra note 11, at 2-3. See also NATIONAL INSTITUTES OF JUSTICE (NIJ), THE FUTURE OF FORENSIC DNA TESTING 14-15 (Lightning Source Inc. 2002) (hereinafter FUTURE DNA).

By looking at VNTRs from several distinct loci on the genome, it is possible to calculate the probability that a particular genetic profile comprised of distinct sets of VNTRs will appear in one or more individuals in a particular population. A standard way to estimate the frequency of a particular profile is to count occurrences in a random sample of an appropriate reference population and then use classical statistical formulas to place upper and lower confidence limits on the evidence. The resulting conclusion of identity or nonidentity between two samples is therefore necessarily probabilistic.

In conducting the comparison, investigators came to adopt the “product rule” for determining RMPs. Any given VNTR may be calculated to occur at a certain frequency in a random population. By the early 1990s the standard was to test for VNTRs at four independent loci on the genome. The product rule allows for multiplying each independent genotype frequency together to produce an overall probability of a match at all four loci.

Jeffreys’ innovation was first used in a forensic setting in England in 1986. Forensic DNA testing was first used in the U.S. in 1987. Soon thereafter some

---

21 Butler, supra note 11 at 2-3.
22 Lewontin & Hartl, supra note 19, at 1745-46; Mildred Cho & Pamela Sankar, Forensic Genetics and Ethical, Legal and Social Implications Beyond the Clinic, 36 NATURE GENETICS S8, S9 (2004).
23 Richard Lempert defines the product rule as follows:

According to the product rule, the probability of two independent events equals the probability of the first event times the probability of the second; with n independent events the separate probabilities of each of the n events are multiplied together to give the probability of their joint occurrence. Thus if the probability that a person had allele A = 1/10 and the probability that he had allele B = 1/10 and the probability that he had allele C = 1/10, and if the probability that the person had one of these alleles was not affected by whether or not he had either or both of the others, the probability that the person would have alleles A, B, and C would be 1/10 x 1/10 x 1/100, or 1/1000.

---

25 Butler supra note 11 at 481-486. Lewontin & Hartl, supra note 21, at 1746. That the individual be “unrelated” is significant because related individuals will have a higher likelihood of sharing a greater percentage of DNA, hence altering the probabilities of a random match.
26 Lewontin & Hartl, supra note 21, at 1746.
27 Butler supra note 11 at 3.
commercial laboratories made use of this “fingerprinting” procedure and in 1988 the U.S. Federal Bureau of Investigation implemented forensic DNA techniques. Critical to the acceptance of forensic DNA in courts was the development of standards of technical proficiency and accuracy in generating RMPs. The product rule was one such standard, requiring that each chosen loci be understood as being inherited independently of the others. Also important were basic crime scene management techniques for the identification and handling of DNA samples.

B. Early Questions and Challenges

Questions about the reliability of DNA evidence surfaced as early as 1989 in cases such as People v. Castro in New York and the Minnesota case of Schwartz v. State. Partially in response to these cases, several federal agencies called upon the National Research Council (NRC), an arm of the National Academies of Science (NAS), to study and recommend guidelines for the production and use of DNA evidence. The NRC created a Committee on DNA Technology in Forensic Science, which issued a report in 1992. It is in the context of the production of this report that race first enters the story front and center.

29 NIJ, supra note 18, at 14-15.
30 This latter area of concern was brought front and center in 1995 in the highly publicized murder trial of O. J. Simpson where defense lawyers undermined apparently airtight evidence connecting Simpson to the crime by calling into question the methods (or lack thereof) employed by the Los Angeles Police Department in collecting and handling of relevant DNA samples. See, e.g., DAVID LAZER, DNA AND THE CRIMINAL JUSTICE SYSTEM: THE TECHNOLOGY OF JUSTICE 4 (David Lazer ed., MIT Press 2004); and Sheila Jasanoff, DNA’s Identity Crisis, reprinted in DAVID LAZER, DNA AND THE CRIMINAL JUSTICE SYSTEM: THE TECHNOLOGY OF JUSTICE 340-45 (David Lazer ed., MIT Press 2004).
32 447 N.W.2d 422 (Minn. 1989).
33 Jasanoff, supra note 30, at 339-40.
34 See NRC I, supra note 21.
The Committee covered an array of issues relating to the forensic use of DNA technologies. Among its most controversial findings were those relating to reference populations and the appropriate methodology for calculating RMPs. In order to calculate the odds of any particular VNTR allele appearing at a given locus on the genome, one must have an appropriate reference population. The product rule depends on the assumption of statistical independence of the alleles tested – that is that they do not tend to occur in groups.  

Generally speaking, the more “related” a person is to a particular population group the higher are the odds of finding shared alleles -- or, alternatively stated, the less independence there is among alleles. Siblings would likely share more DNA than cousins; cousins more than others in the same isolated village; members of the same isolated village more than others in the same region; and so forth. Higher odds favor a suspect or defendant because they indicate a greater likelihood that some other person may have left the DNA sample found at a particular crime scene. The choice of reference population, therefore, can play a critical role in shaping the weight and authority of DNA evidence. The choice, however, is not always straightforward. Indeed, some of the earliest and most contentious controversies involving the use of DNA technology in forensic science involved choosing the appropriate population against which a suspect’s DNA should be compared and defining just how the suspect may be “related” to this population. Concepts of race played a central role in these debates and continue to frame the way forensic scientists, law enforcement, and the bar produce and interpret DNA evidence to this day.

35 See NRC I, supra note 21, at 12.
36 See generally, Butler, supra note 11, at 455-519.
The basic issue is whether or to what extent racial or ethnic categories should be used to characterize reference populations against which particular DNA samples could be compared to generate RMPs. The use of such categories may be particularly problematic in the arena of forensic DNA analysis because racial groups, especially those delineated in the U.S. Census, are fundamentally social not biological categories.\textsuperscript{37} Indeed, at least since the 1970s scientists have understood that race will statistically explain only a small portion of genetic variations.\textsuperscript{38} As a recent editorial in \textit{Nature Genetics} put it, “scientists have long been saying that at the genetic level there is more variation between two individuals in the same population than between populations and that there is no biological basis for ‘race.’”\textsuperscript{39} Nonetheless, to the extent that certain population geneticists understand particular racial groups as sharing a common genetic ancestry – usually by using race as a crude surrogate for geographic or continental ancestry -- members of those groups can be viewed as more “related” to each other (like an extended family) than to individuals from other groups. This problematic understanding of relatedness can then affect the calculation of RMPs. Generally speaking, the more fine-grained the characterization of a particular reference population the higher the odds of a random match – again, higher odds favoring the

\textsuperscript{37} For example, federally mandated racial and ethnic categories are not biomedical in origin. \textbf{United States Office of Management and Budget. “Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity,”} \textit{available at} http://www.whitehouse.gov/omb/fedreg/ombdir15.html (last visited: Feb. 19, 2008). Rather, they derive from the 1997 Office of Management and Budget’s (OMB) “Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity.” \textsuperscript{id} These standards set forth five minimum categories for data on race: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, and White. \textsuperscript{id} There are two categories for data on ethnicity: “Hispanic or Latino,” and “Not Hispanic or Latino.” \textsuperscript{id} These categories provide the basis for the classification of all federal data on race and ethnicity, most notably, the census. The OMB Standards, however, contain an important caveat: “The racial and ethnic categories set forth in the standards should not be interpreted as being primarily biological or genetic in reference.” \textsuperscript{id} These categories were developed to serve social, cultural and political purposes. \textsuperscript{id}

\textsuperscript{38} Richard C. Lewontin, \textit{The Apportionment of Human Diversity}, 6 \textit{EVOLUTIONARY BIOLOGY} 381 (1972).

\textsuperscript{39} \textbf{Editorial, Genes, Drugs and Race, 29 NATURE GENETICS} 239, 239 (2001).
suspect/defendant. In the early years of forensic DNA analysis (when typically only four VNTR loci were tested), there were concerns that using a general, undifferentiated population data base would produce inappropriately low RMPs. The decision to use race in constructing and categorizing reference populations was introduced into forensic DNA analysis in belief that it would improve the precision the calculations that generate RMPs.40

In early 1991 two pairs of eminent population geneticists squared off against each other in the pages of Science, a highly influential scientific journal, to debate the problem of using racial categories in forensic DNA analysis. On one side were Professors Richard Lewontin of Harvard University and Daniel Hartl of the University of Washington (described by an editorial accompanying the article as “two of the leading lights of population genetics.”)41 On the other side were Ranajit Chakraborty of the University of Texas and Kenneth Kidd of Yale University. Their dispute did not revolve around the question of whether to use race but rather how much race to use in constructing reference population data bases from which to calculate match probabilities.

Lewontin and Hartl questioned the then current practice of calculating allele frequencies in the racial categories used in the Census such as “Caucasian,” “Black,” and “Hispanic” to provide the basis for calculating RMPs.42 They argued that such grouping were too broad and that substantial “genetic substructuring” occurred within the broad racial groupings that should be taken into account in calculating match probabilities. Using the broad racial groupings could produce RMPs with substantially lower odds than those that might be produced using more fine-grained ethnically identified sub-

40 Cho & Sankar, supra note 22, at S9.
42 Lewontin & Hartl, supra note 21, at 1747.
populations.\textsuperscript{43} These concerns grew logically out of Lewontin’s earlier path breaking work showing how genetic variation within socially identified racial groupings was actually greater than variation observed between such groups.\textsuperscript{44} This work laid the foundations for understanding that race was incoherent as a genetic concept, or at best, an overly-crude surrogate for genetic variation that improperly tended to reify race as genetic.\textsuperscript{45} Thus Lewontin and Hartl observed,

“Among genes that are polymorphic in European national or ethnic groups, the magnitude of the differences in allele frequency among subpopulations differs from one gene to the next. . . . For example, there are striking geographical clines of allele frequency across Europe for the ABO blood groups: the frequency of the B allele is 5 to 10\% in Britain and Ireland, increases across Eastern Europe, and reaches 25 to 30\% in the Soviet Union; the frequency of the O allele is 70 to 80\% in Sardinians, Irish, and Scottish populations but lower in Eastern European populations. These clines reflect the migrations and political history of Europe over the last few thousand years.”\textsuperscript{46}

Problems were even greater for the “heterogeneous assemblage” known as “Hispanic,” which was perhaps “the worst case for calculating reliable probabilities.”\textsuperscript{47} Consequently, they concluded that using reference databases organized by the broad racial groupings “Caucasian,” “Black” and “Hispanic” was “unjustified.”\textsuperscript{48}
Chakraborty and Kidd argued that Lewontin and Hartl exaggerated both the extent of ethnic substructuring in America and its significance for calculating match probabilities. While conceding that some substructuring existed, they argued that its effects upon frequency estimates generated by using the broader racial data bases was “trivial.” Chakraborty and Kidd did not deny that using finer-grained ethnic reference populations might produce more precise allele frequency estimates. Rather, their point was that such an approach was unnecessary – and unnecessarily burdensome. Current technology and understandings of population genetics, they asserted, justified the use of broad racial and ethnic categories, which were, additionally, far more practical and currently available. Race was at the center of this early debate. But, again, for these eminent scientists, it was not a question of whether to use race but how, or more specifically how much (i.e. how fine-grained) race to use.

This debate took place while the NRC Committee was conducting its study of DNA technology in forensic science. Its report, issued in 1992, discussed both sides of the issue without specifically taking sides. It did, however, choose “to assume for the sake of discussion that population substructure may exist and to provide a method for estimating population frequencies in a manner that may account for it.” The report recognized that “population genetic studies show some substructure within racial groups for genetic variants. . . . Thus, North American Caucasians, blacks, Hispanics, Asians, and Native Americans are not homogeneous groups.” In effect, this approach reflected
the concerns expressed by Lewontin and Hartl, recognizing that social categories of race did not map neatly onto discrete genetically definable population groups. The NRC’s 1992 report created problems for prosecutors. By taking cognizance of the difference of scientific opinion regarding the appropriate calculation of allele frequencies and RMPs, it seemed to assert that forensic DNA technologies lacked to sort of scientific consensus needed to support the introduction of such expert evidence. Thus, for example, in the 1992 case of People v. Barney the California Court of Appeal cited the NRC Report in concluding disagreement and uncertainty in the scientific community regarding the selection of appropriate reference populations precluded the admission of DNA evidence based on the product rule. By April 1993, the director of the FBI asked the National Academies of Science to conduct a rapid follow-up study to resolve these uncertainties. The NRC then appointed a second committee (NRC II) late in 1994 with a specific mandate to update and clarify discussions of population genetics and statistics as they applied to DNA evidence. Meanwhile, the debate continued in the scientific community. The position advocated by Chakraborty and Kidd received a major boost in 1994 when Eric Lander of M.I.T., previously a vigorous critic of the lack of adequate standards in DNA typing.

54 Kaye, supra note 16, at 102-03.
55 People v. Barney, 10 Cal.Rptr.2d 731, 743 (Cal. Ct. App. 1992). David Kaye notes that this case was followed in People v. Wallace, 17 Cal.Rptr.2d 121 (Cal. Ct. App. 1993); Commonwealth v. Lanigan, 396 N.E.2d 311 (Mass. 1979) (finding that product rule calculation method not prescribed by NRC panel for calculating frequency of DNA pattern is not generally accepted among population geneticists); State v. Vandeboeart, 616 A.2d 483 (N.H. 1992) (same); State v. Cauthron, 846 P.2d 502 (Wash. 1993) (finding error in allowing expert to testify that defendant was the source of the incriminating DNA and yet excluding testimony of frequency of the DNA pattern given that the NRC panel had proposed a generally accepted method of calculation); cf. State v. Bible, 858 P.2d 1152 (Ariz. 1993) (holding method as applied to 1988 data base not generally accepted); Springfield v. State, 860 P.2d 435 (Wyo. 1993) (holding frequency re-calculated with “the most conservative” NRC method admissible under relevance standard); People v. Atoigue, DCA No. CR 91-95A (Guam Dist. Ct. App. Div. 1992) (method not generally accepted among population geneticists). Kaye, supra note 16, at 103, n.20.
56 Kaye, supra note 16, at 397.
paired with Bruce Budowle, one of the principal architects of the FBI’s DNA typing program, to write an article in the journal *Nature*, declaring “DNA fingerprinting dispute laid to rest.”

The article argued that applying the product rule to the frequency estimates for four independent VNTRs generated odds of such magnitude that any technical statistical differences observed between the use of the broad racial data bases (as advocated by Chakraborty and Kidd) versus more fine grained ethnic sub-group data bases (as advocated by Lewontin and Hartl) were “of no practical consequence to the courts.”

As Lander and Budowle observed,

“In the vast majority of cases, jury needs to know only that a particular DNA pattern is very rare to weigh it in the context of a case: the distinction between frequencies of $10^{-4}$, $10^{-6}$ and $10^{-8}$ is irrelevant in the case of suspects identified by other means. . . . The most extreme positions range over a mere two orders of magnitude: whether the population frequency of a typical four-locus genotype should be stated, for example, at $10^{-5}$ or $10^{-7}$. The distinction is irrelevant for courtroom use.”

Lander and Budowle were not arguing that racial sub-groups themselves were not needed or desirable in calculating RMPs. The “distinction” they saw as “irrelevant” was the one between ethnic sub-groups, such as “Irish,” and larger racial groups, such as “Caucasian.” Thus, they were legitimating the then current standard FBI practice of using broad racial groups, such as “Black” and “Caucasian,” as reference data bases for generating allele frequencies for calculating RMPs. Significantly, Lander and Budowle did not argue for doing away with racial data bases altogether in favor of using an undifferentiated general

---

58 *Id.* at 735.
59 *Id.* at 738. (emphasis added).
population database. Given the current state of forensic technology which generated RMPs from examining VNTRs at only four loci, they deemed race relevant. They simply did not want too much of it – that is, they did not want law enforcement forced to undertake the burdensome task of developing more elaborate data bases that reflected the wide array of genetic population substructuring that actually occurs across the globe. Given the odds generated by testing at four VNTR loci, they deemed the broad racial categories of the Census more than adequate for forensic purposes.

Lander and Budowle made a critical distinction between statistical and legal relevance. Though hardly the first to do so, the distinction allowed Lander and Budowle to quiet both the scientific debates and the legal uncertainties swirling around this new and powerful forensic technology. This distinction continues to play a role throughout the continuing development and application of DNA technology up to the present day cases such as People v. Wilson. Another critic of the NRC’s 1992 report, David Kaye, made a similar distinction in a 1993 article in the Harvard Journal of Law & Technology. Kaye, who would sit on the second NRC Committee, wrote that in calculating RMPs, “the real issue . . . is not ‘statistical significance’ but rather practical or substantive significance.” The difference was critical for Kaye and others because it provided the basis validating then current law enforcement practices of using broad racial reference population data bases. By distinguishing between statistical vs. logical or practical significance, Kaye and others did not refute Lewontin and Hartl so much as bracket off their concerns as irrelevant to the legal applications of forensic DNA

60 See, e.g., Herman Trautman, Logical or Legal Relevancy – A Conflict in Theory, 5 VAND. L. REV. 385 (1952).
61 136 P.3d 864, 865-66 (Cal. 2006).
technology in courts. Of most immediate significance in terms of the unfolding story of the use of race in forensic DNA technology, is the fact that this distinction played a central role in the NRC’s second report, *The Evaluation of Forensic DNA Evidence* (NRCII), issued in 1996.⁶³

**C. NRC II: Questions of Race Laid to Rest?**

The NRC II report focused primarily on updating and clarifying issues related to population genetics and statistics as they applied to DNA evidence. It argued directly for “using separate databases for different racial groups”⁶⁴ even while it acknowledged Lewontin’s underlying argument that “the variability among individuals within a population is greater than that between populations.”⁶⁵ Recognizing the uncertainties inherent in calculating RMPs, the report noted that “the accuracy of the estimate will depend on the genetic model, the actual allele frequencies, and the size of the data base.”⁶⁶ It was confident, however, that “when several loci are used, the probability of a coincidental match is very small.”⁶⁷ Nonetheless, the report recommended incorporating a ten-fold margin of error in RMP calculation, stating, “if the calculated probability of a random match between the suspect and evidence DNA is 1/(100 million), we can say with confidence that the correct value is very likely between 1/(10 million) and 1/(billion).”⁶⁸

---

⁶⁴ Id. at 22.
⁶⁵ Id.
⁶⁶ Id. at 33.
⁶⁷ Id. at 34
⁶⁸ Id.
At first glance, such a range may strike the reader as rather large but the report legitimizes it by returning to the distinction between statistical and legal relevance. “The proper concern,” it asserted,

“is not whether the probability is large or small, but how accurate it is.

Probabilities are not untrustworthy simply because they are small. In most cases, given comparable non-DNA evidence, a judge or jury would probably reach the same conclusion if the probability of a random match were one in 100,000 or one in 100 million.” 69

In other words, the large range presented earlier in the report was of little practical or legal significance so long as it was good enough to guide a judge or jury in their deliberations. It was good enough for two reasons: first, because it was accurate – accuracy here was crucially distinguished from precision which the large range of probabilities certainly lacks; second, because the lower end of the range still presented odds so vanishingly small as to render it indistinguishable from the upper end of the range as a practical matter – that is, the difference was deemed to be insufficient to have any practical effect on the conclusion a judge or jury would reach in using the evidence.

And yet, even accepting this huge range of variance, the report persisted in using race as organizing categories in calculating RMPs. Thus, even while acknowledging that “some assert the word race is meaningless” in a genetic context, 70 the report adopted the categories “white (Caucasian), black (African American), Hispanic, east Asian (Oriental),

69 NRC II, supra note 61, at 56.
70 Id. at 57.
and American Indian (Native American)” as designated “racial groups” as a matter of
“convenience, uniformity, and clarity.”\footnote{Id.} It justified this choice by asserting that,
“there are reproducible differences among the races in the frequencies of DNA
profiles used in forensic settings, and these must be taken into account if errors
are to be minimized.”\footnote{NRC II, supra note 61, at 57.}

It is instructive to note here just where it is that “difference” made a difference in the
calculation of RMPs. Difference was deemed insignificant when it manifested as a
thousand-fold range for an “accurate” calculation using the product rule to compare a
single sample against a single reference population data base – that is, the “difference”
between one in 100,000 and one in 100 million in made no practical difference for use of
the data in a court of law. To be fair, as noted above, the NRC II report recommended
calculating RMPs with a margin of error limited to ten fold in either direction\footnote{Id. at 34.} – but this
still translates into a variation of one hundred fold between the lowest and highest
estimate. But when race was at issue in the NRC II report, the “difference” of
frequencies among racial reference populations became critical and had to be “taken into
account if errors are to be minimized.”\footnote{Id. at 57.}

Race enters into people’s consciousness in complex and often unanticipated ways.
The NRC II report clearly focused on issues of race in response to the questions raised by
the debate between Lewontin/Hartl and Chakraborty/Kidd. That debate involved the
relation between social groups of race and genetic variation. Both sides recognized that
racial categories were crude surrogates for capturing genetic variation across groups, but
Chakraborty and Kidd were, in effect, arguing that race was nonetheless not “too crude” – that is, it was good enough for practical use in law enforcement because of the ability to generate astronomically low RMPs even allowing for a substantial range of variation.

As a practical matter, the debate cast into doubt the admissibility of DNA forensic evidence in courts; hence the FBI’s urging that the issue be revisited by a second NRC Committee. The NRC II Report, therefore, aimed to quiet the dispute, rendering it irrelevant to the practical application of forensic DNA technologies in law enforcement.

Yet, it is unclear why the NRC II report characterized difference between racial reference populations as meaningful “error” while it deemed the hundred-(or even thousand-) fold range of variance within a single reference population to be of no practical significance. This seems largely to be an artifact of the report’s focus on addressing the issues raised by Lewontin and Hartl in manner that would allow forensic DNA testing to proceed unimpeded by concerns of the accuracy of using racial reference populations to calculate RMPs. The report needed to show that RMPs generated by using racial categories were good enough for practical use in courts of law. The utility and/or validity of using a general population data base without reference to either race or ethnic sub-groups was never really at issue.

In the end, the report issued to following formal recommendation for estimating RMPs:

“In general, the calculation of a profile frequency should be made with the product rule. If the race of the person who left the evidence-sample DNA is known, the database for the person’s race should be used; if the race is not known,
calculations for all the racial groups to which possible suspects belong should be made.” 75

The NRC II report thus legitimized the then-standard practice of using race to generate RMPs. In rejecting Lewontin and Hartl’s concerns about broad racial data bases, it seems also implicitly to have rejected – or at least failed fully to appreciate – Lewontin’s cognate concerns about the incoherence of race as a genetic category and the dangers of reifying race as genetic.

III. Current Standard Practices Regarding Race and Forensic DNA Analysis

A. The Impact of NRC II

The NRC II report became tremendously influential in shaping forensic DNA techniques and their acceptance in courts of law. It established new norms for calculating RMPs generally and for using race-specific data bases in particular. Following the NRC II recommendation, it has since become standard practice to present race-specific RMPs. 76 Thus, for example, in the 1999 case of People v. Soto, 77 the California Supreme Court noted that the dispute regarding population substructuring that had been at the heart of the 1992 case of People v. Barney had “been eclipsed by subsequent important scientific developments, most notably the publication of a completely new report by the NRC,” 78 The court concluded that use of the product rule as applied to broad racial data

75 NRC I, supra note 21, at 5.
76 See, e.g., Butler, supra note 11, at 474-517.
77 981 P.2d 958 (Cal. 1999).
78 Soto, 981 P.2d at 960. Further on in the case the court notes that significant developments since Barney included an FBI World Survey of genetic variation in 1993, the publication of the article in Science by Lander and Budowle, and “of greatest significance,” the NRC II report. Id. at 976.
bases “has gained general acceptance in the relevant scientific community.”79 Similarly, in People v. Wilson80 the criminologist from the California Department of Justice who presented the forensic DNA evidence to the court, testified that the “to help juries understand the significance of a DNA match, the Department followed the statistical approach recommended by [the NRC II report] . . . for presenting the frequency with which genetic profiles occur.”81

In cases such as Soto and Wilson we see that the use of racial data bases characterized by the broad terms of the U.S. Census categories had emerged as normative referents for the calculation of RMPs. Thus, for example, in justifying its calculation of RMPs in Wilson, the State of California argued that the lower court “correctly approve[d] the California Department of Justice’s generally accepted method for generating match probability statistics using reference data from major racial and ethnic groups. Typically, a range of statistics if provided using three major U.S. population data bases: African-American, Caucasian, and Hispanic. This method . . . is supported by NRC II.”82

As representative of current practice, People v. Wilson shows how fully integrated race has become in the conceptualization and practice of forensic DNA analysis. The use of race is understood as requiring no justification other than that it had become “generally accepted” and is “typical”.

79 Id. at 960. Further on in the case the court notes that significant developments since Barney included an FBI World Survey of genetic variation in 1993, the publication of the article in Science by Lander and Budowle, and “of greatest significance,” the NRC II report. Id. at 976.
80 136 P.3d 864 (Cal. 2006).
81 Id. at 866.
82 Answer Brief on the Merits, People v. Wilson, No. S130157 (Cal., Nov. 18, 2005) (emphasis added).
B. CODIS and the Move From VNTRs to STRs

The NRC II report itself was based largely on an assessment of the then current practice of testing samples at four VNTR loci. Ironically, by 1997, barely a year after the report had issued, a new technology had emerged to replace four loci VNTR analysis using restriction fragment polymorphism (RFLP) methods of analysis. In 1985 Kary Mullis and members of the Human Genetics group at the Cetus Corporation had discovered a technique known as Polymerase Chain Reaction (PCR) which enabled scientists to make millions of copies of a specific sequence of DNA in a matter of hours.\(^8^3\) The ability to amplify segments of DNA is critical to forensic analysis. PCR is sensitive, rapid and not limited by the quantity of DNA as are RFLP methods.\(^8^4\) PCR enabled a shift in focus from VNTRs to sections of DNA known as “Short Tandem Repeats” (STRs). VNTRs are typically 10-100 bases in length.\(^8^5\) STRs (also known as microsatellites) are regions of DNA only 2-6 base pairs in length.\(^8^6\) STRs highly variable across individuals and are easily amplified by PCR, thus making them very effective for purposes of human identification.\(^8^7\)

Beginning in 1996, the FBI commenced an effort to develop a set of core STR loci to be used as standard referents for the calculations of RMPs in forensic DNA analysis.\(^8^8\) In November 1997, the FBI settled on 13 core STR loci which were chosen to be the basis of the CODIS (Combined DNA Index System) national DNA Database which was launched in 1998.\(^8^9\) New technologies allowing for “multiplex” testing of

\(^8^3\) Butler, supra note 11, at 63.
\(^8^4\) Id. at 63.
\(^8^5\) Id. at 85.
\(^8^6\) Id.
\(^8^7\) Butler, supra note, at 94.
\(^8^9\) Id. at 13, 85.
multiple loci at once were soon capable of regularly generating RMPs rarer than one in one trillion. The Minnesota State Department of Public Safety has noted that “STRs are very discriminating for single-source samples. Typically, a complete DNA profile might be found in less than one in one hundred billion people. A typical DNA report would read ‘This profile would not be expected to occur more than once among unrelated individuals in the world population.’” By 2000 the FBI laboratory and many others stopped using RFLP analysis altogether in favor of PCR analysis of the 13 CODIS STRs. Because of their use in the FBI database, the 13 CODIS STRs have become a national (indeed international) standard and have come to “dominate the genetic information that has been collected to date on human beings.”

CODIS was initially authorized by the DNA Identification Act of 1994 and became operational in 1998. As described by the FBI, “CODIS is implemented as a distributed database with three hierarchical levels (or tiers) - local, state, and national. NDIS is the highest level in the CODIS hierarchy, and enables the laboratories participating in the CODIS Program to exchange and compare DNA profiles on a national level. All DNA profiles originate at the local level (LDIS), then flow to the state (SDIS) and national levels. SDIS allows laboratories within states to exchange DNA profiles. The

---

92 Butler, supra note 11, at 13.  
tiered approach allows state and local agencies to operate their databases according to their specific legislative or legal requirements.95 As of October 2007, there were over 5 million DNA profiles in CODIS.96 The profiles themselves are not classified by race. Rather they are primarily used, much like a database of fingerprints, to aid in the investigation of crimes by providing matches or “hits” to DNA evidence left at crime scenes.97 In the context of establishing an initial match using the CODIS database, race is therefore irrelevant.

Nonetheless, race has come to pervade the characterization of forensic DNA data generated using the standard 13 CODIS loci. This is because establishing a match is only the first step in applying forensic DNA technology. Once a match is found, whether using the CODIS database or not, law enforcement must still take the further step of calculating an RMP for any given DNA profile. It is at this stage that race enters CODIS – and in a more powerful way than ever before. In addition to the basic CODIS database, the FBI has generated a population file to estimate allele frequencies according to specifically identified racial or ethnic groups.98 This population file is based on a 2001 study led by Bruce Budowle which typed allele frequencies for the 13 CODIS loci from 41 population data sets.99 Budowle classified the results in terms of five “major population groups”: “African American, U.S. Caucasian, Hispanics, Far East Asians, and...

95 Federal Bureau of Investigation, CODIS: Mission Statement and Background is available at http://www.fbi.gov/hq/lab/codis/program.htm (last visited: Feb. 21, 2008); see also Maclin, supra note 27.
97 Butler, supra note 11, at 439. The database is also used to aid investigations in identifying human remains. See id.
98 Butler, supra note 11, at 439.
Native Americans." These allele frequencies have since become the standard reference data base for calculating racially identified RMPs. Thus, for example, in *People v. Wilson*, criminologist Nicola Shea referenced the Budowle study, when noting that the California Department of Justice “used databases that the Federal Bureau of Investigation published in the Journal of Forensic Sciences reflecting profile frequencies in the Caucasian, Hispanic and African-American populations.”

IV. Race, Technology, and “Care of the Data”

A. Race v. Technology

The casual and perfunctory assignment of social categories of race to biological samples in professional discussions of forensic DNA stands in marked contrast to the meticulous care taken concerning the more technical aspects of DNA extraction, amplification and analysis. The discussions of each in a 2005 article by Peter Vallone, Amy Decker and John Butler, of the National Institute of Standards and Technology’s (NIST) Human Identity Project team, are fairly typical. This particular article involved the characterization of allelic frequencies for 70 SNPs (single nucleotide polymorphisms)

100 Id.
102 People v. Wilson, 136 P.3d 864, 866 (Cal. 2006).
103 I am indebted to Mike Fortun for this concept.
in DNA samples taken from three racially marked groups: U.S. Caucasian, African-American, and Hispanic.\textsuperscript{106} The article presents its techniques for racially identifying the DNA samples as follows:

“Anonymous liquid blood samples with self-identified ethnicities were purchased from Interstate Blood Bank, Inc. (Memphis, TN) and Millennium Biotech, Inc. (Ft. Lauderdale, FL).” \textsuperscript{107}

“Self-identification” thus provides the sum total of all care and technique devoted by Vallone et al. to characterizing genetic samples by race. Contrast this with their discussion of the more apparently technical aspects of how they manipulated the samples once in the lab (this is quoted at length to heighten the contrast):

2. DNA extraction
   Blood samples were extracted using a modified salting out procedure.

3. Quantification
   Extracted DNA was quantified using UV spectrophotometry followed by a PicoGreen assay to adjust concentrations to approximately 1 ng/ml.

4. SNP markers
   The 70 autosomal SNP markers are listed in Table 1 (see also http://www.cstl.nist.gov/biotech/strbase/SNP.htm). The PCR primer sequences were obtained from Orchid Cellmark (personal communication, Jeanine Baisch, Orchid Cellmark Dallas). The exact chromosomal locations were ascertained using BLAT (http://genome.ucsc.edu/cgi-bin/hgBlat) and dbSNP (http://www.ncbi.nlm.nih.gov/SNP/) and are based on the July 2003 assembly of the human genome. All of the SNPs are C/T transitions.

5. PCR amplification
   For each sample, the 70 SNP markers were typed in 11 unique 6-plexes and a single 4-plex PCR. The final concentrations of the six (or 4) PCR primer pairs were present at 0.5 mM for all multiplex PCRs. Amplifications were performed in reaction volumes of 10 ml using a master mix containing 1X GeneAmp1 PCR Gold buffer (Applied Biosystems, Foster City, CA), 4.5 mmol/l MgCl2, 250 mmol/Idexynucleotide triphosphates (dNTPs; Promega Corporation, Madison,

\textsuperscript{106} Vallone et al., supra note 103, at 279.
\textsuperscript{107} Id.
Wt), 0.16 mg/ml bovine serum albumin (BSA) fraction V (Sigma, St. Louis, MO), and 0.5 unit of AmpliTaq Gold1 DNA polymerase (Applied Biosystems). The thermal cycling program was carried out on a GeneAmp 9700 (Applied Biosystems) using the following conditions in 9600-emulation mode (i.e., ramp speeds of 1°C/s):

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>95°C</td>
<td>10 min</td>
</tr>
<tr>
<td>Three cycles of {95°C for 30 s, 50°C for 55 s, 72°C for 30 s}</td>
<td></td>
</tr>
<tr>
<td>18 cycles of {95°C for 30 s, 50°C for 30 s +0.2°C per cycle, 72°C for 30 s}</td>
<td></td>
</tr>
<tr>
<td>11 cycles of {95°C for 30 s, 55°C for 30 s, 72°C for 30 s}</td>
<td></td>
</tr>
<tr>
<td>72°C for 7 min</td>
<td></td>
</tr>
<tr>
<td>25°C until removed from thermocycler</td>
<td></td>
</tr>
</tbody>
</table>

Following PCR amplification, unincorporated primers and dNTPs were removed by adding 4 ml of an Exo-SAP enzyme cocktail consisting of 1.4 ml Exonuclease I (10 U/ml) and 2.6 ml (1U/ml) of shrimp alkaline phosphatase (SAP; USB Corp., Cleveland, OH) to each 10 ml PCR reaction. Reactions were mixed briefly and incubated at 37°C for 90 min and then 80°C for 20 min to inactivate the enzymes.

6. Allele specific primer extension (ASPE)
ASPE reactions were also carried out in eleven 6-plexes and a single 4-plex. Multiplex primer extension reactions were conducted in a total volume of 10 ml using 2.5 ml of ABI Prism1 SNaPshotTM multiplex kit mix (Applied Biosystems), 0.5 ml of 10X AmpliTaq Gold1 PCR buffer, 3 ml of PCR template, 3 ml of water, and 1 ml of a stock solution of extension primers, which contained empirically balanced primers (approximately 1 mM each). Extension reactions were incubated as follows: 25 cycles of 96°C for 10 s, 50°C for 5 s, and 60°C for 30 s. Excess fluorescently labeled ddNTPs were inactivated by addition of 1 ml of SAP (1 U/ml). Reactions were mixed briefly and incubated at 37°C for 40 min then 90°C for 5 min.

7. Electrophoresis and typing
A 1.0 ml aliquot of each SAP-treated primer extension product was diluted in 14 ml Hi-Di™ formamide and 0.4 ml GS120-LIZ internal size standard (Applied Biosystems) and analyzed on the 16-capillary ABI Prism1 3100 Genetic Analyzer (Applied Biosystems) using filter set E5 without prior denaturation of samples. Samples were injected electrokinetically for 13 s at 1 kV. Separations were performed in approximately 30 min on a 36 cm array using POPTM-6 (Applied Biosystems). Automated allele calls were made in Genotyper1 3.7 using an in-house macro based on fragment size and dye color.

8. Analysis of data
The data were analyzed with PowerMarker v3.07. Allele frequencies, expected heterozygosity values and p-values (based on an exact test with 1000 resufflings) for each marker are provided in Tables 2–4 for the three U.S. sample groups.108

The point here is not to assess (or even understand) the intricacies of the technical analysis performed by Vallone et al. on their DNA samples. Rather it is to contrast the extreme care and detail devoted to elaborating the techniques performed in the lab with the casual and perfunctory discussion of how the samples came to be racially marked in the first place. As scientists, Vallone et al. understandably go into greatest detail with respect to techniques and practices in which they are professionally trained and proficient. This reflects their reasonable understanding that the extraction, amplification, and analysis of DNA take great care and expertise. The contrasting lack of care taken in characterizing the racial identity of the genetic samples indicates an implicit assumption that such characterizations are obvious, uncomplicated, and take no special expertise. This contrast may be understood more broadly as reflecting a conceptual separation of the world of the “social” from that of the “natural,” where the former is understood to contain transparent categories accessible to all while the latter requires specialized knowledge and expertise for proper analysis and interpretation. In other words, race is seen as easy and obvious; DNA is seen as difficult and complex.\footnote{This lack of comparable care is not restricted to the arena of forensics. For example, a recent survey of biomedical studies using race as a variable found that 72\% of 268 reports analyzed did not explain their methods of assigning race or ethnicity as independent variables. Hasan Shanawani \textit{et al.}, \textit{Non-reporting and Inconsistent Reporting of Race and Ethnicity in Articles that Claim Associations among Genotype, Outcome, and Race or Ethnicity}, 32 J. MED. ETHICS 724 (2006).} There is an utter failure to consider that social subjects such as race may demand similar rigor, expertise, and care in handling as scientific subjects such as DNA.

**B. Social v. Genetic “Race”**

Ironically, this separation of the social from the natural is enabled by the work of geneticists such as Lewontin who, together with a wide array of social scientists, have
worked diligently since World War II to reconfigure race from a biological into a social
construct.\textsuperscript{110} It is precisely because race is currently widely understood as a social
phenomenon that forensic scientists are able effectively to marginalize it from their
analysis of the biological construct of DNA. As a result, their care of the data extends
only to the analysis of DNA samples while wholly overlooking the complexities of using
racial categories in relation to genetics.

In effect, forensic scientists have simply adopted the broad categories of race and
ethnicity used in the U.S. Census to organize their genetic data. The Census, in turn, is
based on the Office of Management and Budget’s Directive 15 on “Standards for
Maintaining Collecting, and Presenting Federal Data on Race and Ethnicity-”\textsuperscript{111} which
provides the following categories as a minimum standard for maintaining, collecting, and
presenting data on race and ethnicity for all Federal reporting purposes: American Indian
or Alaska Native; Asian; Black or African American; Hispanic or Latino; Native
Hawaiian or Other Pacific Islander; and White.\textsuperscript{112}

excellent overview of the history of scientific and cultural understandings of race, see generally, Jonathan
Marks, \textit{Human Biodiversity: Genes, Race and History} (Aldine Transaction 1995). For some
influential statements on race by professional social science organizations, see, e.g., American
Anthropological Association (AAA), Response to OMB Directive 15: Race and Ethnic Standards for
http://www.aaanet.org/gvt/ombdraft.htm (last visited Feb. 21, 2008); American Anthropological
Association (AAA), Statement on Race, 1998, available at http://www.aaanet.org/stmts/racepp.htm (last
visited Feb. 21, 2008); American Sociological Association (ASA), \textit{The Importance of Collection Data and
Doing Social Scientific Research on Race}, 2003, available at http://asanet.org/galleries/default-

\textsuperscript{111} U.S. Office of Management and Budget (OMB), \textit{Revisions to the Standards for the Classification of
Federal Data on Race and Ethnicity}, 1997, available at
http://www.whitehouse.gov/omb/fedreg/ombdir15.html (last visited Feb. 21, 2008). Hereinafter OMB,
Revisions.

\textsuperscript{112} OMB, \textit{Revisions}, supra note \textsuperscript{110}, at 133; see also U.S. Census Bureau, \textit{Racial and Ethnic
Classifications Used in Census 2000 and Beyond}, available at
These federally mandated standards emerged as a consequence of major governmental programs and legal initiatives instituted since the 1960s. The OMB categories provide the basis both for census information and also for access to a variety of governmental goods and services that are contingent upon membership in a particular racial or ethnic group. For example, federal users of racial data provided by the census include: the Department of Education, Department of Justice, Department of Labor, Equal Employment Opportunity Commission, Federal Reserve, Department of Health and Human Services, Housing and Urban Development, Department of Agriculture, and the Veterans Administration. Alice Robbin notes, “groups must be counted in order to make credible claims for political representation, demonstrate discriminatory practices against them, seek and obtain legal remedies, receive governmental assistance for a host of social programs, and evaluate current, as well as develop new public policy.” Additionally, they provide the framework for evaluating school desegregation, electoral districting, and other civil rights initiatives.

Given the social and political uses which such standards were designed to serve, it should come as no surprise that Directive 15 explicitly acknowledges that the categories it provides are social in character, not biological or genetic. Using these same categories in the context of genetic research, however, presents issues of a different order. As Lee et al. note, “[r]esearch utilizing race serves to ‘naturalize’ the boundaries

---

113 Melinda Nobles, Shades of Citizenship: Race and the Census in Modern Politics 75-79 (Stanford University Press 2000).
117 OMB, Revisions, supra note 110, at 133.
dividing human populations, making it appear that the differences found reflect laws of nature. In fact, the use of race and ethnicity in biomedical research is problematic because it is caught in a tautology, both informed by, and reproducing, ‘racialized truths.’”

This dynamic reinforces what sociologist Michael Omi has characterized as an “interesting dilemma” facing scientists in the United States: “One the one hand,” Omi asserts, “scientists routinely use racial categories in their research . . . On the other hand, many scientists feel that racial classifications are meaningless and unscientific.”

C. The Obvious Solution: A Non-Racial, General Population Data Base.

Race was originally introduced into the calculation of RMPs in the early years of forensic DNA analysis in hopes of providing more refined statistical calculations. The rationale was grounded in the reasonable observation that there is a modicum of genetic variation across certain human populations. Capturing this variation might provide more accurate RMPs. Greater accuracy was important in the early years of forensic DNA analysis when generating RMPs using only four VNTR loci. With such limited data, the variation of RMPs generated using different reference populations could be of forensic significance.

Today the situation has changed significantly. With the advent of multiplex assays testing for the 13 standard CODIS loci, forensic scientists are now capable of regularly generating RMPs with denominators many times in excess of the entire world’s population.

---


As another article with the FBI's Bruce Budowle as lead author put it as early as 2000,

"By typing these [13] STR loci, the random match probability for a multiple locus profile will be exceedingly small. The average random match probability for unrelated individuals for the 13 STR loci is less than one in a trillion, even in populations with reduced genetic variability."

Under such circumstances the concerns originally expressed by Lewontin and Hartl that using broader racial categories will not produce accurate enough RMPs fade into irrelevance. When one is dealing with odds in the hundreds of billions or trillions, the more fine-grained characterizations of genetic variation among ethnic sub-groups called for in their original 1991 article in *Science* simply are not necessary as a practical matter.

The issue then shifts from *how much* race to use, to *whether* to use race at all. As is made evident by the range of odds generated in cases such as *People v. Wilson* (one of 96 billion Caucasians, one of 180 billion Hispanics, and one of 340 billion African-Americans), the use of a non-racially marked general reference population would still generate RMPs whose reciprocals would still exceed the world population many fold. Under such circumstances, any differences between RMPs generated by using race-specific reference populations vs. a general population are without forensic

---


122 Id.

123 Id.

significance. Thus, it is no longer necessary even to use the broad racial reference populations advocated by Kidd and Chakraborty back in 1991.\textsuperscript{125}

The possibility of abandoning racial reference populations in favor of a general population database, was broached in a 2000 report by the National Institute of Justice’s National Commission on the Future of DNA Evidence.\textsuperscript{126} In the context of discussing the rise of testing for STRs in contrast with the older method of VNTR analysis, the report noted that,

“It is already apparent that most of the STR variability is within groups. Although groups differ, the mean differences between groups are less than the individual differences within groups; profiles that are rare in one group tend to be rare in others. With enough loci it may be possible to have a single database for all the major groups in the United States.”\textsuperscript{127}

Given the ability to generate RMPs in the trillions, its seems obvious that we currently have enough loci to have a single non-racial reference population data base. The question remains, why do we continue to use race?

**D. The Inertial Power of Race.**

There is no easy answer to this question. I would like to suggest that there is an inertial power to race in American society that propels the continued use of race long after any original rationale for its introduction may have faded. In particular, I would like to consider three possible dynamics contributing to the persistent use of race in the


\textsuperscript{127} Id. at 27. (emphasis added).
presentation of forensic DNA evidence even after current technology has obviated the need for race-specific data bases: 1) the persistent conceptualization of race as genetic; 2) the confusion of statistical with forensic significance; and 3) the deep seated American identification of violent crime and race.

First, with respect to genetics: in spite of decades of efforts on the part of social and natural scientists to sever the ties between race and biology, large segments of American society continue to conceptualize race primarily in genetic terms. The rise of modern genomics was supposed to resolve the dispute. Upon the completion of the first draft of the human genome in 2000, President Clinton declared,

After all, I believe one of the great truths to emerge from this triumphant expedition inside the human genome is that in genetic terms all human beings, regardless of race, are more than 99.9% the same. At the same press conference, Craig Venter, president and CEO of Celera Genomics, reinforced Clinton’s message, asserting that “the concept of race has no genetic or scientific basis.” Yet, ironically, since this iconic press conference, genetic conceptualizations of race seem to have reemerged with a vengeance. As anthropologist Sandra Lee has noted,

the current trajectory of genomic research is increasingly focused on the 0.01 per cent genetic difference that is believed to separate one individual from another. The search for functional genetic variability is increasingly taken up in

129 *This was also quoted in Reading the Book of Life, White House Remarks on Decoding of Genome, NEW YORK TIMES, June 27, 2000, also available at http://query.nytimes.com/gst/fullpage.html?res=9502E1D81230F934A15755C0A9669C5B63 (last visited: Feb. 21, 2008).*
130 *Id.*
populations that are identified by conventional notions of ‘race’. This trajectory is the result of a confluence of factors, including a growing infrastructure of research materials that are racially categorized through the creation of biobanks. Such sorting practices reflect the ongoing conflict over the meaning of ‘race’ in science and medicine. In the emerging era of the new genetics, in which supercomputer technology has given way to an explosion of human genetic data, biobanks that utilize taxonomies of race in the classification, storage and distribution of DNA samples become ‘racializing technologies’ that promote notions of racial biology in research protocols designed to discover group difference.\textsuperscript{131}

Sociologist Troy Duster has further argued that “new claims that DNA analysis of crime scene data will assist criminal investigations” have led to a “molecular reinscription of race in the biological sciences.”\textsuperscript{132} The same technology underlying the creation of racialized forensic DNA data bases is also being used for drug development\textsuperscript{133} and to market new genetic ancestry tracing services.\textsuperscript{134} There have thus emerged both structural and commercial incentives to continue to use race in relation to genetics. This dynamic undergirds the inertial power of race in forensic DNA analysis by providing a broader context in which race is understood somehow to be naturally or logically connected to genetics. This dynamic is further reinforced by the tendency of forensic DNA experts, as

\begin{footnotes}
\item[131] Lee, supra note 128, at 448.
\item[134] Deborah Bolnick et al., The Science and Business of Genetic Ancestry Tracing, 318 SCIENCE 399 (2007).
\end{footnotes}
discussed above,\textsuperscript{135} to take race as an obvious, unproblematic category that does not require the same care and analysis as genetic data.

Second, the technical ability to generate statistically significant variation in RMPs across racial databases has led to the unquestioned assumption that such variation is also legally significant. Using the 13 CODIS loci, forensic experts around the world have characterized allele frequencies for numerous ethnically and racially marked populations.\textsuperscript{136} Modest frequency variation at each individual locus, when multiplied across loci by the product rule, can lead to apparently significant variations in RMPs across races. Thus, in cases such as \textit{People v. Wilson}, the variation in RMPs across race-specific data bases may appear, at first blush, to be important. In that case, RMPs varied from one in 96 billion Caucasians, to one in 180 billion Hispanics, and one in 340 billion African-Americans.\textsuperscript{137} According to the data bases, Wilson’s genetic profile was more the three times as likely to occur in a Caucasian as an African-American – an apparently significant difference. But in the forensic context, this statistically significant difference has no real practical importance. When the world’s population is under seven billion, the difference between an RMP of one in 96 billion and an RMP of one in 340 billion provides no meaningful distinction for a finder of fact. Both are astronomically low probabilities. Nonetheless, the experts’ ability to generate statistically significant differences across races seems to have propelled the continued use of racial data bases; even when these differences are of no practical legal significance.\textsuperscript{138}

\textsuperscript{135} See \textit{supra} notes 8, 35, 37 and accompanying text.


\textsuperscript{138} For further discussion of the difference between legal and statistical significance, see \textit{supra} notes 140, 41, 150 and accompanying text.
Ironically, the reverse logic is used by law enforcement to support the rising use of race to generate suspect profiles from DNA evidence left at the scene of a crime. This sort of genetic racial profiling uses allele frequencies to generate an estimate of the likely racial or ethnic background of an as yet unidentified perpetrator. In this context, Troy Duster notes that law enforcement officials themselves have made a distinction between theoretical and practical significance of racial difference in genetics. Thus, as Duster notes,

“When representative spokespersons from the biological sciences say that ‘there is no such thing as race’, they mean, correctly, that there are no discrete racial categories that come to a discrete beginning and end, that there is nothing mutually exclusive about our current (or past) categories of ‘race’, and that there is more genetic variation within categories of ‘race’ than between them. All this is true. However, when Scotland Yard or the Birmingham police force or the New York Police Department wants to narrow the list of suspects in a crime, they are not primarily concerned with tight taxonomic systems of classification with no overlapping categories. That is the stuff of theoretical physics and philosophical logic, not the practical stuff of crime-solving or the practical application of molecular genetics for health delivery via genetic screening, and all the messy overlapping categories that will inevitably be involved with such enterprises. That is, some African Americans have cystic fibrosis even though the likelihood of that is far greater among Americans of North European descent and, in a parallel if not symmetrical way, some American Whites have sickle cell anaemia even though the likelihood of that is far greater among Americans of West African descent.
But in the world of cost-effective decision-making, genetic screening for these disorders is routinely based on commonsense versions of the phenotype. The same is true with regard to the quite practical matter of naming suspects.139 Here the scientific understanding that race is not genetic is trumped in practice by the purported ability of some genetic tests to estimate the likelihood that a suspect belongs to one or another socially identifiable race. Law enforcement is using race because it is perceived as of practical significance – even if not scientific. Yet, in the courtroom context the reverse is the case - race is used because it is perceived as scientific – even if not of practical significance.

Race persists largely because it has become normative; an unquestioned, standardized practice that persists long after the rationale for it has faded. In Wilson, the State argued for the legitimacy of using race-specific RMPs primarily on the grounds that such was the “standard practice”140 and the “generally accepted method for generating match probability statistics”141 and that “typically” the state and federal labs used “three major U.S. population databases: African-American, Caucasian, and Hispanic.”142 General acceptance, typicality, and standardization – these all powerfully drive the inertial power of race.

Third, there is the unfortunate but well documented tendency in the United States to identify race and violent crime. In their book, Whitewashing Race, Michael Brown et al. discuss a cultural shift that began in the 1960s when the image of “the brave little girl walking up to the schoolhouse door in the face of jeering white crowds was replaced by

139 Duster, supra note 132, at 435-36
140 Wilson, 136 P.3d at 868.
141 Answer Brief on the Merits, People v. Wilson, No. S130157 (Cal. Nov. 18, 2005).
142 Id.
fearsome young black men coming down the street ready to take your wallet or your life." 143 In the context of the rising racialization of crime on the United States, Rothenberg and Wang observe that “[f]rom 1990 to 2004, blacks were five times more likely than whites to be incarcerated, and in 2000, blacks and Latinos comprised 63% of incarcerated adults, even though together they represented only 25% of the total population.” 144 Similarly, while examining the impact of DNA technology on the criminal justice system, Simon Cole concludes that,

“At the endpoint of this system is a carceral system that embodies gross race and class disparities, even if differential rates of offending are taken into account: two thirds of people in prison are racial and ethnic minorities, one in eight black males in their twenties are in prison or jail, three-quarters of persons in prison for drugs are people of color.” 145

Considering the dynamics that have produced such inequalities, Brown et al. review an array of historical, legal, and sociological data on race and crime in the United States. Discussing a classic observational study of police responses to juveniles in a Midwestern city in the 1960s, they note that police justified their different treatment of black youths on “epidemiological lines,” concentrating on “those youths whom they believed were more likely to commit delinquent acts.” 146 They argue, however, that the results of this ‘actuarial’ reasoning . . . is to exacerbate the very differences that are invoked to justify the racially targeted practices in the first place. This in

---

146 Brown, et al., supra note 143, at 149-50.
turn helps to cement the public’s image, and the police’s image, of the gun-toting
gangster or drug dealer as black or Latino. And this confirms the validity of the
police focus on youth in the same kind of vicious circle . . . described a generation
ago.”

The same sort of actuarial reasoning is at work in Duster’s identification of the use of
genetics in the “practical matter of naming suspects.” The association of crime and
race produces more racialized crime. As Dorothy Roberts has noted, the resulting mass
incarceration is “iatrogenic” – by damaging social networks, distorting social norms,
and destroying social citizenship, the disproportionate incarceration of minorities has
produced a vicious cycle of crime and repression that further reinforces the identification
of race and crime in the public mind.

Taken together, the persistent conceptualization of race as genetic, the confusion
of statistical with forensic significance, and the deep seated American identification of
violent crime and race may be understood to frame and facilitate the inertial power of
race to perpetuate itself as a salient category of forensic DNA analysis long after its
practical legal utility has passed.

V. Presenting Race and Forensic DNA in Court.

In discussing the possibility of moving beyond race in forensic DNA analysis, the
National Commission on the Future of DNA Evidence mentioned that a general

\[147 Id. at 151.\]
\[148 Duster, supra note 132, at 435-36.\]
\[150 See also Rose Brewer & Nancy Heitze, The Racialization of Crime and Punishment, 51 AMER. BEHAVIORAL SCIENTIST 625 (2008).\]
population data base “may appeal to those who would like to emphasize individual differences and ignore group differences.” The pertinent question presented by technological advances in forensic DNA analysis, however, is not “do we want to ignore group difference?” Rather, it is “what justification is there to continue to present DNA evidence in terms of race?” It is to this question that we now turn by looking at how race is currently used in presenting forensic DNA evidence in courts of law.

California provides an apt site to pursue this question. From People v. Barney in 1992 to People v. Wilson in 2006, a series of California cases have played important roles both in reflecting and shaping practices concerning the use of racial data in the presentation of forensic DNA evidence from its early years to the present day. These cases chart a trajectory from the initial rejection of RMPs calculated using broad racial data bases to the embrace of racial data bases in the aftermath of the NRC II report to renewed questioning of which data bases were appropriate to reference, to the full embrace of the standardized use of broad racial data bases in the calculation and presentation of RMPs to a jury in a criminal case.

A. Early Cases: DNA in Flux

As noted above the court in People v. Barney found that disagreement and uncertainty in the scientific community regarding the selection of appropriate reference

---

153 136 P.3d 864 (Cal. 2006).
154 Barney, 8 Cal.App.4th at 819.
155 People v. Soto, 981 P.2d 958 (Cal. 1999).
156 People v. Pizarro, 3 Cal.Rptr.3d 21 (2003).
158 See supra notes 53, 77, 152 and accompanying text.
populations precluded the admission of DNA evidence based on the product rule. The court discussed not only the findings of the first NRC report but also dealt at length with the 1991 dispute between Lewontin/Hartl and Chakraborty/Kidd in the pages of *Science* over the selection of appropriate reference populations for calculating RMPs. The court then went on to note how the first NRC report had taken note of this controversy and concluded that because the dispute among experts remained unresolved the court could not admit into evidence RMPs calculated using the broad racial data bases.

*People v. Barney* was followed by a flurry of discussion and general professional hand-wringing over the need to resolve the disputes so that DNA evidence could be used with impunity. The second NRC report issued in 1996 seemed to settle the argument in favor of Chakraborty and Kidd, opening the door to the free use of racial data bases in calculating RMPs. The 1999 California case of *People v. Soto* took note of this development when it effectively overruled *People v. Barney*, stating that “the use of the unmodified product rule in DNA forensic evidence has gained general acceptance in the scientific community.”

In *Soto* the defendant was charged with rape. The victim was unable to identify her assailant because he wore a mask, but she described him as a “white male... with light or blond hair and an olive complexion.” *Soto* was described as “Latino... with a dark complexion and black hair.” Nonetheless, DNA from semen left at the

---

159 *Barney*, 8 Cal.App 4th at 819
160 Id at 816
161 Id at 819
162 See supra notes 33, 77, 152 and accompanying text
163 See supra notes 54-55 and accompanying text
164 *Soto*, 981 P.2d at 966
165
166
167
crime scene matched Soto’s. Using the older RLFP analysis of four genetic loci, Robert Keister, a criminalist at the Orange County Sheriff’s Department (OCSD) testified at trial that “there was a probability of only 1 in 189 million of finding the same DNA pattern in individuals selected at random from the population represented by the OCSD’s Hispanic data base.” Interestingly, at the preliminary examination before the trial, Keister initially calculated a probability of 1 in 214 million. This difference of well over 10% resulted from “add[ing] some more samples to [the OCSD] data base and . . . [running] further tests on the augmented data base.” No explanation was given for why samples were added, where they came from, or how they were characterized as “Hispanic.” Further discussion revealed that Keister also used FBI Hispanic data bases to produce the following frequency estimates: (1) Southwest Hispanic (Texas): 1 in 55 million; (2) Southeast Hispanic (Florida): 1 in 2.3 billion; (3) U.S. Black: 1 in 2.4 billion; and (4) U.S. Caucasian: 1 in 3 billion. Asked to comment on the significance of the variation among frequency estimates, another DNA expert witness, Dr. Bruce Kovacs, professor of medicine at University of Southern California, testified, “They [the denominators] are astronomically large numbers. The significance of whether something is 1 in 55 million or 1 in 110 million versus 1 in 4 billion is something that I can’t really get my hands on in a real concrete way to distinguish that difference. It’s a very, very, very rare event.” In this context it seems clear that the experts are making an implicit distinction between statistical significance and forensic significance. Kenneth Kidd himself testified at the trial and made just such a distinction, offering the opinion that

---

168 *Id.* at 967.
169 *Id.* at 971.
171 *Id.* at 971.
172 Soto, 981 P.2d at 972.
“any difference in estimates over one in a million was pragmatically meaningless.”

From such detailed discussions of odds ratios using very specific numbers that carry the authority of scientific expertise, the witness ultimately characterized the odds of the match simply as “very, very, very, rare” – not particularly scientific terminology but apparently adequate to aid the relevant finder of fact in making a determination of guilt or innocence.

One might ask why it was acceptable in Soto to significantly change frequency estimates by the addition of more racialized data or reference to other databases but it was apparently not acceptable, or at least not considered, to change frequency estimates by removing race altogether? The Soto court’s further discussion of the rationale for using racialized databases provides some insight into this issue. Thus, the court noted that,

Major laboratories that do RFLP analysis... have developed their own separate population databases for each of several broad racial or ethnic categories such as Caucasian, Black, and Hispanic, the assumption being that mating among members of any one of those categories of the United States population is sufficiently random to justify using them in conjunction with the product rule to calculate the frequency of a DNA profile.

The assumption of random mating was central to overcoming Lewontin and Hartl’s concerns about ethnic substructuring and hence supporting the court’s decision to reject the reasoning of the Barney court and admit the race-specific frequency estimates into

---

173 Id. at 973.

174 People v. Soto, 21 Cal.4th at 526 (1999), (citations omitted).
evidence. The court goes on in a footnote, however, to consider and reject the possibility of using a non-racially marked data base for calculating RMPs, stating,

Conversely, the laboratories do not use a single interracial United States database, presumably because the incidence of random mating between members of the different racial categories is deemed low enough to preclude use of the product rule to calculate an overall frequency statistic for the United States population as a whole. 175

The characterization of a single general data base as “interracial” is in itself notable. It conflates the absence of racial markers assigned to DNA samples with the presence of interracial mating among the sources of such samples. Viewed in this light, the court’s characterization of an “interracial” genetic data base may be seen as reflecting a much deeper, unarticulated and misguided understanding of races as biologically distinct in meaningful ways that are perpetuated by mating patterns. Such presumptions reflect a Jim Crow era logic of anti-miscegenation where through either de facto or de jure historical patterns individuals were “deemed” to mate across racial lines in only small numbers. The failings of such an anti-miscegenation logic are evident in the historical reality of interracial mating throughout the country’s history176 and are especially pronounced with respect to the category “Hispanic,” which, as Lewontin and

175 Id. at 526, n.18

Hartl noted in the 1991 *Science* article “is a biological hodgepodge. It includes people of Mexican, Puerto Rican, Guatemalan, Cuban, Spanish and other ancestries.”\(^\text{177}\)

None of these points, however, were made at trial made because the expert witnesses for the defense were still looking backward to the Lewontin/Hartl article to argue that *more* ethnicity was needed not less to calculate RMPs. Thus, one defense expert argued for the need to develop separate data bases from the distinct “ancestral populations” that lived in places such as Cuba, Mexico, Spain, and Central America.”\(^\text{178}\)

This is perhaps still understandable in a case where the older RFLP technique was used to analyze only four genetic loci but soon this technique would be superseded by PCR analysis of up to 13 CODIS loci.

**B. From Pizarro to Wilson: Race in Flux**

The 2003 case of *People v. Pizarro*\(^\text{179}\) marked another major shift in the presentation and use of racially marked forensic DNA evidence in court. Michael Pizarro was convicted of murder and rape in 1990.\(^\text{180}\) He appealed the case in 1992 on the grounds that RFLP testing of DNA evidence was not at that time generally accepted in the scientific community.\(^\text{181}\) The case was remanded for a thorough evidentiary hearing which occurred in 1998, (by which time such issues were largely resolved in the

---

\(^{177}\) Lewontin & Hartl, [*supra* note 122], at 1749.

\(^{178}\) Soto, 21 Cal.4th at 529-30.

\(^{179}\) *People v. Pizarro*, 100 Cal.App.4th at 530 (2003).

\(^{180}\) *People v. Pizarro*, 100 Cal.App.4th at 540.

\(^{181}\) *Id.* at 540.
scientific community), and the court ruled that the evidence was admissible and reentered the judgment. 182

In this second appeal, Pizarro contended that there was a basic error in the presentation of the DNA evidence when the prosecution informed the jury that the DNA profile frequency was the probability of finding a matching profile in the Hispanic population. 183 Pizarro himself was identified as Hispanic but the ethnic identity of the perpetrator was not known independently. The court ruled in Pizarro’s favor, finding that the use of the Hispanic data base presumed that the perpetrator was in fact Hispanic when there was no sufficient evidentiary foundation to establish that fact. 184 It concluded that, “recurring thematically throughout the issues in this case are evidentiary violations founded on the improper assumption that defendant was in fact the perpetrator and that defendant’s traits therefore could be relied upon to provide or clarify those traits of the perpetrator forming the basis of the DNA evidence.” 185 The court argued that “in the absence of sufficient evidence of the perpetrator’s ethnicity, any particular ethnic frequency is irrelevant,” 186 and found that “the improper mention of ethnicity unfairly and unjustifiably encourages the jurors to focus on ethnicity and race-specifically the ethnicity and race of the defendant, the only suspect before them” 187

In a footnote, the court presented three options for presenting profile frequencies:

(1) establish that the perpetrator more likely than not belongs to a particular ethnic population, then present only the frequency in that particular ethnic population;

---

182 Pizarro, 100 Cal.App.4th at 540.
183 Id. at 622.
184 Id. at 622-23.
185 Id. at 540.
186 Id. at 631.
187 Id. at 540.

100 Cal.App.4th
(2) present only the most conservative frequency, without mention of ethnicity; or
(3) present the frequency in the general, nonethnic population. These options promote the goals of admitting only relevant evidence and eliminating unjustifiable and potentially prejudicial references to ethnicity and race.\textsuperscript{188}

The court here seemed to be acutely sensitive to the dangers of improperly injecting race into the presentation of DNA evidence. Significantly, it broached the possibility of moving beyond race in the presentation of frequencies. Nonetheless, it remained primarily concerned with the proper management of racial references and did not go on specifically to question the underlying utility (or lack thereof) of race itself as an analytic category in the presentation of DNA evidence.

In the course of reaching this conclusion, the court’s opinion presents some revealing discussions of the meaning and significance of race in forensic DNA analysis. For example, the State argued that any reference to race was harmless in part because “frequencies do not vary greatly by ethnicity.”\textsuperscript{189} This, of course, raises the issue of why use ethnicity (or race) at all if the differences are so insignificant? Indeed, here it becomes clear that under such circumstances the only thing that race adds to the presentation of such DNA evidence is race itself – not simply as a marker of the suspect but as a conceptual framework for constructing a relationship among violent crime, genetics and race.

At the 1990 trial, the sole scientific witness testified that “the likelihood of finding another unrelated Hispanic individual with a similar profile as Mr. Pizarro is

\textsuperscript{189} Id. at 622, 629, 631.
approximately one in 250,000.”  Such odds fall well below the one in one million beyond which Kenneth Kidd saw no pragmatic significance; hence they may seem to justify the use of different racial data base in this case. It turns out, however, that Pizarro was actually identified as “half Hispanic and half Caucasian.” When asked how he could calculate RMPs in such a situation, the expert in the original trial stated that “there is nothing we can do other than to compare them to the two populations and we would use only the smaller one of the two in our report . . . [because it] is less detrimental to the defendant.” Pizarro’s “mixed” race presented a problem for the witness. (It certainly flies in the face of the anti-miscegenation logic of Soto). Analytically, the expert literally segregated Pizarro’s racial identities, producing separate RMPs with reference to distinct White and Hispanic data bases. His conceptual framework could not encompass the concept of mixed race – rather it was premised upon, and indeed demanded, a logic of racial purity.

In the aftermath of Pizarro, David Kaye, of the Arizona State University School of Law, wrote a powerful and influential critique of the court’s reasoning regarding the use of ethnic reference populations in calculating RMPs. Kaye essentially agreed with the Pizarro court that “if the perpetrator could have come from any of several racial groups, looking into only one racial group for a random match probability could be misleading.” He expressed grave concern, however, over the court’s conclusion that

---

190 Pizarro, 100 Cal.App.4th at 624.
191 See text accompanying notes ----, infra.
192 Pizarro, 100 Cal.App.4th at 624.
193 Id. at 624-25.
195 Id. at 214.
giving a range of frequencies for the major racial or ethnic groups in the United States was therefore unacceptable. Kaye noted that,

Since providing statistics from several racial groups is the standard way of assessing the significance of a match in cases in which the racial and ethnic status of the perpetrator of the crime initially is unknown, the opinion [in Pizarro] casts doubt on the outcome in innumerable cases. 196

Kaye disputed what he saw as the court’s presentation of an “unbridgeable gap between scientific and legal reasoning in this situation,” 197 asserting that

the scientific reasoning that the court questioned is nothing less than the kind of hypothesis testing – considering the principal alternatives and examining the probability of certain outcomes under each of these alternative hypotheses – that dominates modern statistical thinking. In this instance, the DNA expert simply testifies to how surprising the match would be if some major alternatives to the hypothesis that the defendant is the source of the biological samples were true. 198

Thus, he concluded, “When it comes to deciding what evidence is logically relevant, however, it is difficult to conceive of any substantive difference between legal and scientific reasoning.” 199 In the abstract, there is much merit to Kaye’s argument. But in declaring no difference between scientific and legal reasoning, he obscured the distinction between scientific and legal relevance.

---

196 Kaye, supra note 194, at 215.
197 Id.
198 Id.
199 Id.
When noting that a perpetrator may “share the defendant’s race or ethnicity” Kaye conflates race and genetics referring to a defendant’s “genetic heritage.” The concept of “sharing” is very peculiar and particular here. *Pizarro* involved someone who Kaye and the court define as “half-Hispanic” and “half Caucasian” – implicitly making these two categories mutually exclusive. Yet in social practice this makes no sense and reinforces the idea of genetically distinct and bounded “races” rather than continuums of variable “mixes”. Michael Pizarro could not be allowed to be a “Hispanic Caucasian” because the databases are not constructed that way. Two suspects may “share” the same race, but that “race” itself must be singular and unmixed – not “shared” with other races – but rather be capable of being broken down into parts “half x” and “half y.” In short, the entire model of using race to improve probability estimates depends on keeping genetic databases segregated by race – the segregation in turn produces RMPs.

Kaye correctly noted that there was no necessary presumption made about a perpetrator’s identity if the suspect’s DNA were compared to an array of racial data bases of populations to which the perpetrator might belong. He therefore adequately addressed and effectively undermined two of the *Pizarro* court’s three permissible options for the presentation of profile frequencies: “(1) establish that the perpetrator more likely than not belongs to a particular ethnic population, then present only the frequency in that particular ethnic population; (2) present only the most conservative frequency, without mention of ethnicity” Kaye’s logic, however, was based on a presumption that race

---

200 *Id.* at 212.

itself remained relevant in the calculation of RMPs. Thus, it failed to address the third option to “present the frequency in the general, nonethnic population.”

In the 2006 case of People v. Wilson, the California Supreme Court embraced Kaye’s arguments to disapprove of the reasoning in Pizarro and re-validate the calculation of RMPs using race-specific data bases even when the race of perpetrator is not otherwise known. In reaching its conclusions, the court asserted that,

The question here revolves around exactly what is the relevant population. The question is complicated by the fact that the odds vary with different racial and ethnic groups. Because of this variation, separate databases are maintained for different population groups, and the odds for each group are calculated separately.

The court then agreed with the lower court’s finding that

When the perpetrator’s race is unknown, the frequencies with which the matched profile occurs in various racial groups to which the perpetrator might belong are relevant for the purpose of ascertaining the rarity of the profile.

This effectively overturned Pizarro and reinstated the practice of using racially marked data bases to generate profile frequencies.

VI. Race, Genes, and Relevance.

In the 1993 case of Daubert v. Merrell Dow Pharmaceuticals, Inc., the Supreme Court notably articulated a “gatekeeping” role for the trial court judge in considering

\[\text{s}\]

\[\text{sd}\]

\[\text{d}\]

\[\text{e}\]

\[\text{f}\]

\[\text{g}\]

\[\text{h}\]

\[\text{i}\]

\[\text{j}\]

\[\text{k}\]

\[\text{l}\]

\[\text{m}\]

\[\text{n}\]

\[\text{o}\]

\[\text{p}\]

\[\text{q}\]

\[\text{r}\]

\[\text{s}\]

\[\text{t}\]

\[\text{u}\]

\[\text{v}\]

\[\text{w}\]

\[\text{x}\]

\[\text{y}\]

\[\text{z}\]
the admissibility of scientific evidence. Central to the holding in *Daubert* was the Court’s articulation of a requirement that the trial judge ensure that “an expert's testimony both rests on a reliable foundation and is relevant to the task at hand.” 208 Considered in light of the above discussion, it will now be made clear that the use of race in generating RMPs for forensic DNA matches should be deemed inadmissible by courts as neither relevant not reliable.

**A. When Race is not Relevant.**

Federal Rule of Evidence (FRE) 401 states that “Relevant evidence” means evidence having any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence.” 209 Clearly forensic DNA evidence is often relevant to a criminal proceeding. RMPs generated through reference to a population data base are therefore also often relevant. A central argument of this article, however, is that in the presentation of such RMPs, race is *not* relevant. Race does not add information that has “any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence.” 210

Taking the *Wilson* case as a paradigmatic example of how race is used in the presentation of forensic DNA evidence, we can see that central to the court’s decision were the assumptions it brought to bear regarding the relevance of race in producing DNA evidence. Following its assertion that data from racial groups to which the perpetrator might belong are relevant, the court made the relatively straightforward

---

207 *Id.* at 597.
208 *Id.*
209 Fed. R. Evid. 401.
210 *Id.*

**Deleted:** at 597
**Deleted:** RE
**Deleted:** FRE 401.
**Formatted:** Font: Italic
assertion that, “relevant evidence is evidence ‘having any tendency in reason to prove or
disprove any disputed fact that is of consequence to the determination of the action.’
(Evid.Code, § 210.)” The court went on to note that, “‘The test of relevance is
whether the evidence tends, “logically, naturally, and by reasonable inference” to
establish material facts such as identity, intent, or motive.’ ” These are basic rules of
evidence and consistent with the FRE 401 concept of relevance. The court, however,
framed the question of relevance in terms of “what is the relevant population” rather than
considering whether differentiation among populations itself provided any legally
relevant data. In Wilson’s case, and in most cases using current techniques of forensic
DNA analysis the answer to this latter question should simply be no.

The RMPs at issue in Wilson ranged from one of 96 billion Caucasians, one of
180 billion Hispanics, and one of 340 billion African-Americans. The court accepted
this racially marked data as relevant because forensic scientists had identified statistically
significant variation in frequencies when using different racial reference populations to
generate RMPs. The court went on to quote approvingly Professor Kaye’s assertion
that, “Contrary to the Pizarro court’s assertions, in a ‘general population case’-one in
which the investigation cannot be limited to a particular racial group-the statistics for a
range of groups surely are relevant.” Kaye made the apparently reasonable point that
having more data about RMPs for a range of populations would “surely” aid a jury in
establishing a material fact, such as identity. And indeed, the court concluded that, “it is

\footnotesize

\begin{footnotes}
\item[211] Wilson, 136 P.3d at 869.
\item[212] Id. (citation omitted).
\item[213] Id. at 867.
\item[214] Id. at 865-66.
\item[215] Id. at 869. (citations omitted).
\end{footnotes}
relevant for the jury to know that most persons of at least major portions of the general population could not have left the evidence samples.” 216

But forensic (and other) scientists have also repeatedly made the point that once a particular odds threshold is passed any difference among profile frequencies is of little or no practical significance. 217 As Yale geneticist Kenneth Kidd noted in the 1999 Soto case, “any difference in estimates over one in a million was pragmatically meaningless.” 218 Moreover, in Wilson itself, state criminologist Nicola Shea testified that when nine genetic markers are used (as in Wilson’s case), “the result would be a ‘pretty discriminating number’ no matter what population data base was used.” 219 Yet at no point did the court consider the logical implication of Shea’s statement – that under such circumstances using racially marked data bases to generate RMPs added nothing to the ability of the jury to make determination of guilt or innocence. The difference between one in 96 billion and one in 340 billion does simply does not “‘have[e] any tendency in reason to prove or disprove any disputed fact that is of consequence to the determination of the action.’” 220 Such information provides nothing of use to the finder of fact that would not already be available by using a general non-racially marked reference population which would have generated similarly powerful RMPs. In other words, where experts can generate such astronomically low RMPs, race simply is irrelevant and should not play a role in the presentation of DNA evidence.

Several other statements by prominent forensic DNA experts further highlight the glaring irrelevance of race to presenting DNA evidence given the power of current

216 People v. Wilson, 136 P.3d 864, 869 (Cal. 2006). (citations omitted).
217 See supra note 173 and infra note 248 and accompanying text.
219 Wilson, 136 P.3d at 866.
220 Id. at 869.
technology. Arguing in 1996 for the adequacy of using broad racial data bases to generate RMPs, the FBI’s own Bruce Budowle and Keith Monson noted that, “A profile would be considered rare whether is had an estimated frequency of 1/5,000,000 or 1/500,000,000. Obviously the difference in the rarity of such estimates would have little consequence in a forensic context.” More to the point, Budowle and Eric Lander, in their highly influential 1994 *Nature* article on forensic DNA technology argued that a distinction in population frequency between “$10^{-5}$ or $10^{-7}$” was “irrelevant for courtroom use.” The distinction in population frequencies across the diverse race specific RMPs generated in Wilson (roughly between $10^{-11}$ and $3.4 \times 10^{-11}$) was far smaller than that cited by Budowle and Lander as irrelevant. Given that current techniques regularly generate RMPs in the range of $10^{-11}$ (one in 100 billion) across diverse racial data bases any distinction among race-specific RMPs between must be understood as similarly “of little consequence in a forensic context” and hence “irrelevant for courtroom use.” In forensic contexts, the only thing that race adds to RMPs is race itself. It provides no additional information that is relevant to aiding the finder of fact to resolve any material issue at trial.

**B. When Race is not Reliable**

---


223 Id. at 738. (emphasis added)

224 See Budowle & Monson, supra note 221.

225 Id. at 738. (emphasis added)
The requirement that scientific evidence be “reliable” is typically discussed in terms of the following factors set forth in Daubert: (1) Whether the technique or theory underlying the evidence has been tested;\(^\text{226}\) (2) whether it has been subject to peer review and publication;\(^\text{227}\) (3) the known or potential rate of error of the technique or theory when applied;\(^\text{228}\) (4) the existence and maintenance of standards or controls;\(^\text{229}\) and (5) whether the technique or theory has been generally accepted in the scientific community.\(^\text{230}\)

When looking at these factors in relation to the generation and presentation of race-specific RMPs for DNA evidence, it is immediately clear that factors (1), (2), and (5) have been met. As discussed above,\(^\text{231}\) over the years numerous studies have been published in peer reviewed journals testing and evaluating the use of race-specific data bases to generate RMPs. Thus, since the inception of forensic DNA evidence, the use of race has been standard and generally accepted practice.\(^\text{232}\) Such general acceptance, however, is no longer the sole determining factors in assessing the reliability of scientific evidence. When scientific practices concerning the use of race in relation to forensic DNA are examined more closely, it becomes evident that they fail to meet factors (3) and (4): there has been little or no consideration of potential rates of error regarding the definition of race and its assignment to particular DNA samples; nor are there adequate standards or controls for the definition and assignment of racial categories to DNA


\(^{227}\) Id. at 593-94.

\(^{228}\) Id. at 594.

\(^{229}\) Id.

\(^{230}\) Id.

\(^{231}\) See supra notes 140-42, 175, 186, 206, 223-25, and accompanying text.

\(^{232}\) See supra notes 81, 141, 180, 230, and accompanying text.
samples. Such lack of basic scientific rigor calls into question the reliability of RMPs generated using racial categories.

Specifically, in the context of forensic DNA research, we see that the scientists who have developed racialized databases have in effect let the concept of “self-identification” supplant the need for any scientifically rigorous or coherent rationale for classifying genetic data by race. This is apparent in the highly influential article written by Budowle et al. in 2001 which has become a primary reference in calculating race-specific RMPs.233 Titled, “CODIS STR Loci Data from 41 Sample Populations,” the Budowle article purports to “present [] STR allele distribution data on 12 or 13 of the CODIS core STR loci in several sampled populations from each of the following major population groups: African American, U.S. Caucasian, Hispanics, Far East Asians, and Native Americans.” 235 This distribution data was derived samples provided by 20 laboratories distributed widely across the U.S., Canada, the Caribbean, and Mexico.236

Like the Vallone article discussed above,237 the Budowle article takes care to specify the technical laboratory instruments and practices used to analyze the samples. But with regard to race, Budowle et al. fall below even the Vallone article’s meager reference to self-identification,238 and provide absolutely no information on how or by whom racial identity was ascribed to these samples. Most of the samples came from law enforcement agencies.239 Self-identification may have been used, but it is also quite

233 Personal E-mail from John Butler, leader of the NIST Human Identity Project Team, to author, (April 26, 2007, time xx:xx:xx) (on file with author).
235 Id.
236 Id.
237 See supra Part IV, Section A.
238 Vallone, et al., supra note 103, at 229.
239 Budowle, et al., supra note 234, at 453.
likely that law enforcement authorities themselves ascribed racial identities to the samples.

In an article on the ethical, legal and social implication of forensic DNA analysis, Cho and Sankar discuss at length a British study showing that external ascriptions of racial identity by law enforcement authorities correspond very poorly with underlying patterns of genetic variation. The article notes that in the British study, “classifications into the five ‘ethnic’ groups [Caucasian, Afro-Caribbean, Indian sub-continental, Southeast Asian, and Middle Eastern] were assigned by police officers by visual characteristics,” based on perceptions of outward appearance rather than on knowledge of individual ancestry. The actual correspondence of these external ascriptions to the “true” ancestry of the individuals ranged from 30% for the Middle Eastern category up to 67% for Afro-Caribbean, with Caucasian falling in between at 56% percent. In other words, if the samples providing the basis for the Budowle article classified based on external ascriptions of race by law enforcement authorities, it would not be unreasonable to suppose that somewhere around 50% of the classifications were inaccurate in terms of their relation to genetically based ancestral origins. If this is the case, it calls into question the legitimacy of any RMPs derived from these reference populations.

Perhaps the samples provided to Budowle were classified by self-identification. This, however, would not solve the problem. Self-identification is a social practice, not genetic. Moreover, as Cho and Sankar note, “individual self-classification is not stable. [F]or example, one U.S. study found that one-third of people change their own self-

240 Budowle, et al., supra note 234, at 453.
identified race or ethnicity in two consecutive years.” Complicating matters still further, a recent study by Condit et al. found people often have very incomplete knowledge of the biological ancestry. Of a sample of 224 subjects interviewed for a study on attitudes toward race-based pharmacogenomics (the tailoring of drugs to genetic profiles), Condit et al. found that 39.6% did not know all four of their biological grandparents. In such situations, self-declared race may fail to capture significant variation in biological ancestry.

This lack of care given to the meaning and attribution of race in a genetic context contrasts markedly with the obvious scientific rigor applied to the elaboration of the more technical aspects concerning the extraction, amplification, and analysis of forensic DNA samples. Clearly, the general practice of using forensic DNA to help identify criminal suspects meets all the Daubert standards of reliability. It is only with respect to the handling of race that the reliability of particular RMPs should be called into question. The use of a general reference population to generate RMPs without regard to race would directly overcome this lack of reliability.

C. Race, Genes, and Prejudice

Even if race-specific RMPs were deemed somehow relevant and reliable they should still be excluded as prejudicial. Federal Rule of Evidence 403 states that,

---

244 See supra notes ----, and accompanying text.
Although relevant, evidence may be excluded if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury, or by considerations of undue delay, waste of time, or needless presentation of cumulative evidence.\footnote{Fed. R. Evid. 403.}

The probative value of race-specific RMPs must be evaluated in relation to the alternative probative value of non-race-specific RMPs. As noted above,\footnote{See supra notes 173, 176, 190, 218, and accompanying text.} Yale geneticist Kenneth Kidd testified in \textit{People v. Soto} that once the threshold odds ratio of one in one million was crossed, any further differentiation among RMPs was pragmatically meaningless.\footnote{People v. Soto, 21 Cal.4th 512, 534 (1999).} Similarly, the FBI’s Budowle and Monson discussed the difference between odds of $1/5,000,000$ and $1/500,000,000$ as of little or no forensic significance.\footnote{See Budowle \& Monson, supra note 221.} Since current technology can use non-racial general reference populations to generate regularly RMP’s far in excess of $1/5,000,000$ (and even in excess of $1/500,000,000$), then the probative value of any refinement of the odds provided by the addition of race-specific RMPs, even if relevant, should be deemed \textit{de minimus}.\footnote{Jeffrey M. Prottas \& Alice A. Noble, \textit{Use of Forensic DNA Evidence in Prosecutors’ Offices}, 35 J.L. MED. ETHICS 310 (2007) (“DNA evidence is most often used in sexual assault cases It appears to play role in a significant minority of murder cases and is rarely employed otherwise.” \textit{Id.} at 312.)}

What concerns for prejudice should then be balanced in the scales against this \textit{de minimus} relevance? The dangers of racial bias tainting the evaluation of forensic evidence are of paramount concern in this context. DNA evidence is usually presented in cases of violent crimes, often of the most heinous variety.\footnote{See text accompanying notes \textit{--- supra}} Where race is gratuitously injected into a the context of violent crime and genetics is added to the mix, the danger of
creating stigmatizing racial stereotypes by conflating race, violence, and genes\textsuperscript{251} should be deemed to outweigh any \textit{de minimus} probative value provided by race-specific RMPs.

Concern to ensure that racial prejudice does not infect the justice system must be primary in any evaluation of the admissibility of forensic DNA evidence. As the U.S. Supreme Court noted in \textit{McCleskey v. Kemp}, “[b]ecause of the risk that the factor of race may enter the criminal justice process, we have engaged in ‘unceasing efforts’ to eradicate racial prejudice from our criminal justice system.”\textsuperscript{252} Thus, for example, the prosecution may not challenge a juror on the basis of race;\textsuperscript{253} a change of venue may be constitutionally required as a result of widespread racial bias in a community;\textsuperscript{254} and the prosecution is barred from appealing to racial prejudice in its argument to the jury.\textsuperscript{255}

In their treatise on Federal Practice and Procedure, Wright and Graham note that “any reference to race by prosecutor must be justified by compelling state interest.”\textsuperscript{256} They caution, in particular, that “While many jurors would reject crude appeals to prejudice, more sophisticated forms of this technique may not be recognized as such. \textit{Today the appeal to prejudice is apt to be disguised as some form of science.}”\textsuperscript{257} With specific reference to FRE 403, they conclude that, fairness in adjudication does not consist entirely in the accuracy of the factual determinations but may require some sacrifice of accuracy to avoid the suspicion that the decision rests on prejudice disguised as science. Therefore, the party who


\textsuperscript{255} \textit{People v. Cudjo}, 6 Cal.4th 585, 625 (1993).

\textsuperscript{256} \textit{Id.} (emphasis added).

\textsuperscript{257} \textit{Id.} (emphasis added).
asserts a major premise based on one of the suspect classifications must expect that his premise will be more rigorously scrutinized than is typical in rulings on relevance.

Thus far, the use of race the presentation of forensic DNA evidence has received virtually no scrutiny from courts in terms of the value or lack thereof that race adds to the accuracy of the RMPs thus generated. Wright and Graham allow that some measure of accuracy may need to be sacrificed to avoid the suspicion of racial prejudice. In the case of presenting RMPs without regard to race any such a sacrifice would be negligible.

Wright and Graham’s reference to the distinctive power of science to disguise appeals to prejudice is especially apt in the context of forensic DNA evidence. In a study of mock jurors, Jonathan Koehler found that, “the way in which DNA match statistics are framed and presented to legal fact finders may affect how they think about and use the DNA evidence.” Koehler’s study looked only at different probabilistic frames for presenting the same statistic, but it is important to consider that similar subtle psychological dynamics may be at work in the framing RMPs in terms of race.

Sheri Lynn Johnson has argued that the use of negative racial stereotype pervades the presentation of criminal cases to juries. She argues that “If the entire body of relevant data is surveyed, the inference that race influences many white jurors’ determinations of guilt is unavoidable.” In their recent article, “Implicit Bias: Scientific Foundations,” Greenwald and Krieger discuss the science of implicit

---

261 Id., at 1804.
cognition, which “suggests that actors do not always have conscious, intentional control over the processes of social perception, impression formation, and judgment that motivate their actions.” They define “implicit biases” as “discriminatory biases based on implicit attitudes or implicit stereotypes.” Being implicit, such biases are not conscious – yet they are significant. They note that,

Implicit biases are especially intriguing, and also especially problematic, because they can produce behavior that diverges from a person’s avowed or endorsed beliefs or principles. The very existence of implicit bias poses a challenge to legal theory and practice, because discrimination doctrine is premised on the assumption that, barring insanity or mental incompetence, human actors are guided by their avowed (explicit) beliefs, attitudes, and intentions.

Greenwald and Krieger go on to review data from the “Implicit Attitude Test” (IAT) which is widely used to assess implicit attitudes toward African Americans relative to European Americans. They observe that researchers have consistently found what they describe as “implicit attitudinal preference” for European Americans over African Americans. They conclude that “a substantial and actively accumulating body of research establishes that implicit race bias is pervasive and is associated with discrimination against African Americans.” To the extent that such implicit race bias might already be present among average jurors, injecting race into the presentation of

---

263 Greenwald & Krieger, supra note 262, at 946.
264 Id.
265 Id.
266 Id.
267 Id., at 952.
268 Id., at 966.
269 Id., at 951.
forensic DNA evidence presents a significant danger of tainting the proceedings with unfair prejudice.

This danger is heightened by the pervasive association of race and violent crime in the public mind. For example, Hurwitz and Peffley argue that since the infamous “Willie Horton” ad run by the National Security Political Action Committee (NSPAC) against Democrat Michael Dukakis during the 1988 presidential campaign, subtly associating race and crime has been a staple of modern politics. In that spot, “the narrator notes that Willie Horton, a convicted murderer, received multiple weekend furlough passes from prison, during the last of which, the narrator informs us, he ‘fled, kidnapping a young couple, stabbing the man and repeatedly raping his girlfriend.’ While the ad could have conveyed exactly the same information without graphics, NSPAC elected to superimpose the most menacing possible picture of Horton, an African American, over the narrative.”

Hurwitz and Peffley go on to note that ad was particularly effective because of its “implicitness” which allows White Americans to internalize the association of African-Americans and violent crime, without directly challenging their conscious commitments to norms of racial equality.

Professor of Theology Ted Peters further cautions that “if we identify crime with genes and then genes


271 \textit{Id.} at 100.

272 \textit{Id.} at 100-01.
with race, then we may inadvertently provide a biological support for prejudice and discrimination.\footnote{TED PETERS, \textit{PLAYING GOD? GENETIC DETERMINISM AND HUMAN FREEDOM}, 73 (2d ed., Routledge 2003).}

Taken together, the presence of racial imagery in criminal trials, the psychological dynamics of implicit prejudice, and the prominent association of race and violent crime in the public mind, counsel strongly against the unnecessary introduction of race into the presentation of forensic DNA evidence. More specifically, the dangers they present of infecting criminal proceedings with racial bias clearly outweigh the minimal probative value provided by the use of race-specific RMPs. Thus, even if court deem race-specific RMPs to be relevant, they should nonetheless exclude such evidence as prejudicial.

\textbf{VII. Conclusion}

Race has been present in forensic DNA evidence since its inception. Over the past twenty years, the use of race-specific RMPs has become a normative, routine, and largely unquestioned practice. Whatever justifications may have originally been proffered for this practice have long since been superseded by basic technological developments which allow for the calculation of extremely powerful RMPs without reference to race. In relation to the presentation of forensic DNA evidence to juries, race is simply a concept whose time has passed. Race-specific RMPs provide little or no relevant information to finders of fact. They present a significant danger of unfairly prejudicing deliberations through the gratuitous association of race with genetics and violent crime. Ending the practice of generating race-specific RMPs will not materially
impede the ability of law enforcement to obtain convictions using DNA evidence. Forensic experts will still be able regularly to generate astronomically low RMPs (often with denominators far in excess of the world’s population) using a non-differentiated general reference population. There is no legal or practical justification for the continued presentation of forensic DNA evidence in terms of race. The practice can and should be ended. It should be replaced with the use of non-racial general population data bases to generate RMPs.

Given current technical ability to generate minuscule RMPs, even using a general population base, these recommendations may not change the specific outcomes of individual cases. They will, however, affect larger issues of how the criminal justice system is implicated in constructing, perpetuating, or deepening broader racialized understandings of the relations among race, genetics and violent crime. By eliminating at least one powerful site for the improper use genetics as a prism through which to view race and crime, these recommendations aim to take a step toward developing a more appropriate understanding of the complex relations among race, genes, and justice.
FORENSIC SCI. COMMUNICATIONS

Accepted Practices by the Forensic DNA Community Supported by NRC II Report. GENETIC IDENTITY
CONFERENCE PROCEEDINGS
SEVENTH INTERNATIONAL SYMPOSIUM ON HUMAN IDENTIFICATION – 1996. Available at
omitted)