COMING CLEAN ABOUT "JUNK DNA"

Simon A. Cole [*]

It is a challenge to reply to a response when its very title pleads that we put the issue of whether forensic DNA profiles contain predictive medical information to rest.[1] I agree that the recent exchange between Professors Joh, Kaye, and myself has probably beaten the "junk DNA" horse past the point of expiration. One thing we all agree upon is that the potential privacy violations engendered by the storage of forensic DNA *profiles* in law enforcement databases is a "distraction,"[2] as Professor Kaye puts it, from the potential privacy issues posed by the storage of DNA *samples* in law enforcement and other government repositories.

Nonetheless, this exchange has not been a useless exercise. It began when I discovered Professors Joh and Kaye's contributions during my effort to better understand—and, therefore, more clearly convey in my own writing—the state of scientific knowledge concerning the claim that the information held in law enforcement genetic databases is innocuous from a privacy standpoint. Professor Joh asserted that the claim of innocuousness was not true,[3] and Professor Kaye countered that Professor Joh's claim was flatly "false."[4] Under such circumstances, I was at a loss as to what to tell my own readers. Therefore, I traced back Professor Kaye's key source, and offered my own contribution to the debate, suggesting that both authors had engaged in a certain degree of oversimplification.[5]

Professor Kaye's most recent contribution to this exchange brings further clarity to the issue. As his meticulous exposition of the precise mechanisms behind contemporary genetic screening demonstrates, when he and

---

[*] Associate Professor of Criminology, Law & Society, University of California, Irvine; Ph.D. (science & technology studies), Cornell University; A.B., Princeton University.

[1] D. H. Kaye, *Please, Let's Bury the Junk: The CODIS Loci and the Revelation of Private Information*, 102 Nw. U. L. Rev. Colloquy 70 (2007), http://www.law.northwestern.edu/lawreview/colloquy/2007/25/ (link) [hereinafter *Bury the Junk*].

[2] *Id.* at 81.

[3] Elizabeth E. Joh, *Reclaiming 'Abandoned' DNA: The Fourth Amendment and Genetic Privacy*, 100 Nw. U. L. Rev. 857 (2006) (link).

[4] D. H. Kaye, *Science Fiction and Shed DNA*, 101 Nw. U. L. Rev. Colloquy 62, 63 (2006), http://www.law.northwestern.edu/lawreview/colloquy/2006/7/ (link) [hereinafter *Science Fiction*].

[5] Simon A. Cole, *Is the 'Junk' Designation Bunk?*, 102 Nw. U. L. Rev. Colloquy 54, 54 (2007), http://www.law.northwestern.edu/lawreview/colloquy/2007/23/ (link) (arguing that if Joh had oversimplified by treating all non-coding DNA as equivalent, Kaye had oversimplified by treating the absence of a causal connection between non-coding DNA and disease as a lack of any predictive value).

other proponents of forensic DNA databases say that forensic DNA profiles have "no predictive value," they actually mean that the profiles have predictive value, but that it is so small as to be practically useless. Likewise, when Professor Joh and other opponents of such databases say that forensic DNA profiles "contain predictive medical information,"[6] they also mean that forensic STRs have only a very small amount of predictive value, at least currently.

Professor Kaye's response clarifies his declarations that law enforcement DNA profiles "have no meaning,"[7] "reveal nothing about propensities for disease, behavioral traits, or the like,"[8] "can tell nothing about a person,"[9] and are "as meaningless as fingerprints,"[10] and explains how his claims that "no forensic STR locus has been found to be predictive"[11] and that "any claim that the DNA profiles currently used for identification constitute 'predictive medical information' is false,"[12] over the course of his substantial body of work on the subject were shorthand for the more complex explanation contained in his response. His response makes clear that forensic STRs contain predictive information, but that he cannot envision feasible exploitations of this information given the current state of genetic knowledge.[13]

I do not have the genetic knowledge to challenge Professor Kaye's claims.[14] However, many readers may understand Professor Kaye's body

---

[6] Joh, *supra* note 3, at 870.

[7] D. H. Kaye & Michael E. Smith, *DNA Databases for Law Enforcement: The Coverage Question and the Case for a Population-Wide Database*, *in* DNA AND THE CRIMINAL JUSTICE SYSTEM: THE TECHNOLOGY OF JUSTICE 247, 256 (David Lazer ed., 2004).

[8] David H. Kaye, *Who Needs Special Needs? On the Constitutionality of Collecting DNA and Other Biometric Data from Arrestees*, 34 J.L. MED. & ETHICS 188, 194 (2006).

[9] D. H. Kaye & Michael E. Smith, *DNA Identification Databases: Legality, Legitimacy, and the Case for Population-Wide Coverage*, 2003 WIS. L. REV. 413, 431 (link).

[10] David H. Kaye, *Two Fallacies about DNA Data Banks for Law Enforcement,* 67 BROOK. L. REV. 179, 188 (2001) (link).

[11] *Science Fiction*, *supra* note 4, at 64.

[12] *Id.* at 62–63.

[13] As my colleague William C. Thompson argues, if this is indeed the case then there should be no objection to making the entire database of DNA profiles publicly available for scientific research in de-identified form. *See* William C. Thompson, Statement to the California Commission on the Fair Administration of Justice (Jan. 10, 2007), *available at* http://www.ccfaj.org/documents/reports/problems/expert/Thompson%20Testimony.pdf (link).

[14] I still find Professor Kaye's account excessively presentist. I understand Professor Kaye's aversion to what he calls the "speculation" about how it may be possible to exploit genetic databases as scientific knowledge progresses. *Bury the Junk*, *supra* note 1, at 73, However, it seems incorrect to label any extrapolation of future knowledge as "science fiction," as he has done twice in this exchange. *See id.* at 81; *Science Fiction*, *supra* note 4, at 62. The term "science fiction" implies untruth, whereas in fact, given enough science fiction scenarios, one future scenario must turn out to be correct. In other words, while any particular prediction of the future state of genetic knowledge may be unlikely to be correct, we *do* know that genetic knowledge is likely to advance *in some way as yet unforeseeable*. Therefore, a minimal assumption that genetic knowledge will advance seems appropriate. Insisting, as

of work to be saying something more dismissive than what he describes in his most recent response. Professor Kaye's arguments may or may not convince other readers that his shorthand description of the admittedly very complicated and technical state of scientific knowledge is appropriate. We can perhaps all agree, however, that this admirably meticulous fuller explanation only benefits the public discourse.

Professor Kaye does, that any potential uses of genetic profiles must plausibly proceed from our current understanding of genetics knowledge (or effectively doing the same by refusing to "speculate" about such advances) is no less "science fiction" than assuming any particular scenario. Professor Kaye's prediction that "the information coded in the databases is and will remain, with . . . limited exceptions . . . useful only for identification," *Bury the Junk*, *supra* note 1, at 71, is itself only one of many possible extrapolations of the future, a science fiction scenario. Professor Kaye's insistence on labeling all extrapolations of the future state of genetic knowledge that cannot be supported by reference to current theory as "science fiction" puts opponents of DNA databases in an unfair bind because it essentially demands solid evidence of the state of future knowledge, something no one can produce.

Presumably, Professor Kaye would respond that his extrapolation of the future is more defensible than others because it is "based on current knowledge and practice." *Id.* It may be more defensible, but that does not mean it is any more likely to be correct. Would the current capability of genetics have been predictable from the state of knowledge and practice in 1960? If not, there is no reason to assume that the capabilities of genetics in 2050—when the law enforcement DNA databases we are building today will likely still be in place and encompass a large portion of the population—must be wholly predictable from the current state of theory and knowledge.