

NELCO
NELCO Legal Scholarship Repository

New York University Public Law and Legal Theory
Working Papers

New York University School of Law

7-1-2013

Legal and Ethical Issues in Forensic DNA Phenotyping

Erin Murphy
NYU School of Law, erin.murphy@nyu.edu

Follow this and additional works at: http://lsr.nellco.org/nyu_plltwp



Part of the [Criminal Procedure Commons](#)

Recommended Citation

Murphy, Erin, "Legal and Ethical Issues in Forensic DNA Phenotyping" (2013). *New York University Public Law and Legal Theory Working Papers*. Paper 415.
http://lsr.nellco.org/nyu_plltwp/415

This Article is brought to you for free and open access by the New York University School of Law at NELCO Legal Scholarship Repository. It has been accepted for inclusion in New York University Public Law and Legal Theory Working Papers by an authorized administrator of NELCO Legal Scholarship Repository. For more information, please contact tracy.thompson@nellco.org.

Legal and Ethical Issues in Forensic DNA Phenotyping

Erin Murphy*

NYU School of Law

In the recent Supreme Court case of *Maryland v. King*, which addresses the constitutionality of compulsory collection of DNA samples from felony arrestees, the state and federal government repeatedly underscored that forensic tests of biological samples look only at meaningless, non-sensitive information. These non-coding, non-expressive parts of the genome have even earned a nickname, “junk,” that alone does much to assure the public that the police are not scrutinizing confidential information. The moniker works in the legal community as well: the “it’s only junk DNA” refrain has been repeated by innumerable courts in their opinions upholding DNA collection statutes. Judges have consistently privileged the benefits to crime-solving against the minimal privacy intrusion posed by revelation of this otherwise meaningless string of numbers.

Yet despite the repeated invocation of the “junk” mantra in criminal courts, efforts at unlocking the deeper secrets of the human genome continue apace. Many of these projects are undertaken by private industry, or by scientists unaffiliated with the government in its law enforcement capacity. But the law enforcement arm of the federal government have also placed stakes in this race. For example, the National Institute of Justice, the research arm of the Department of Justice, has funded studies to isolate genetic polymorphisms keyed to age, ethnicity, skin tone, hair color, eye color, face shape, and other physical characteristics.¹ It has also funded “biosocial criminology” research, such as an examination of the link between genetic polymorphisms and antisocial behavior.² If money talks, then the government – including the law enforcement arm of the government – is telling us that it is very interested in genetic tests that go far beyond the “string of numbers.”

* Professor of Law, New York University School of Law. This research is generously supported by the The Filomen D'Agostino and Max E. Greenberg Research Fund. I have benefitted tremendously from comments from attendees at the NYU Law School faculty workshop, the participants at the 3rd National Conference on Genetics, as well as those at the 2013 PULSE Workshop at UCLA School of Law. Thanks to Nicole Geoglis and Taylor Stevens for their exceptional research assistance.

To some, the testing of DNA samples for anything beyond “junk” reads as sinister. But this chapter aims to identify with care precisely what issues are raised by what is commonly called phenotypic genetic testing, *i.e.*, the testing of biological samples for observable or expressed traits. Although the science is still very much in its infancy, recent breakthroughs suggest that legal actors and policy makers will soon have to confront these questions. Accordingly, this chapter proceeds in four parts. Part I briefly canvases the state of the technology today, projects where it might be headed tomorrow, and imagines the possible role that phenotypic testing may play in law enforcement. Part II focuses on the legal questions attending such testing in the United States, centering largely on the Fourth Amendment. Part III probes ethical and other concerns, largely grouped under the headings of privacy and discrimination, and Part IV sets forwards possible legislative or policy restrictions to address some of those concerns. The conclusion offers some thoughts about the path forward.

I. Introduction to Forensic DNA Phenotyping

A. The technology.

In the literature, DNA testing for observable characteristics is often referred to as “FDP” for Forensic DNA Phenotyping. Understanding the capacity of forensic FDP testing requires appreciation of some of the difficulties that impede efforts to assess expressed characteristics in biological terms. As a preliminary matter, it is worth noting that there is one former difficulty that will not, it now seems clear, pose an obstacle—namely, technological capacity. Advances in instrumentation and scientific knowledge have enabled rapid and affordable sequencing of genetic material, as well as other forms of analysis. Although there are skeptics, entrepreneurs today realistically race toward a \$100 whole-genome sequence. Thus, to the extent that technology once posed a realistic barrier to widespread analysis of complex characteristics, that seems less and less likely to be the case. Instead, the major impediments to FDP rests in three other areas.

First, contrary to the neat Punnett squares of Mendelian inheritance that schoolchildren routinely draw, it turns out that the actual biological mechanisms dictating even traits as simple as hair or eye color prove profoundly complex.³ For instance, there is no single gene that codes eye color as “blue” or “brown”; instead, there are an assortment

of regions that all seem to play a role, and yet even those known polymorphisms do not appear to operate as switches or toggles, but rather are subjected to as yet uncertain processes that determine ultimate outcome. Other characteristics may be subject to epigenetic forces which are presently little understood.⁴ Even sex, arguably among the easiest traits to discern given the relative ease of measuring the presence of the Y chromosome, can for complicated reasons at times be inaccurately reported. Thus, even if all of the relevant genetic regions can be identified, other processes may interfere with accurate prediction. And, note, analysis yields just that – predictions.

Second, testing of forensic samples may pose challenges unique to the criminal context. In the clinical setting, samples are collected and preserved in pristine conditions. But any effective forensic system must be able to contend with fieldwork challenges, like low quantities of template, and degradation due to light, heat, or moisture. Moreover, samples cannot always be retested if problems arise during analysis, and so the first shot may be the only one.

Third, although some traits manifest in largely objective terms with finite variations – sex, for instance – others sound in more subtle registers. Are his eyes blue, cobalt, cornflower, or cerulean? Is that blond hair or dirty blond hair or light brown hair? Is her skin chocolate, mocha, or caramel? The problem created by such nuance is three-fold: tests may be considered useful only inasmuch as they can discern these subtle variations (i.e., differentiate shades of blue); validation of testing methods requires reference to an objective arbiter of those characteristics (“The test predicted black hair, and the subject has black hair versus “The subject recorded her hair color as dark brown, not black!”); and consensus must exist as to the characterization of a trait so that validated tests can be translated by those in the field using them (in other words, police can distinguish between black and dark brown hair).

Despite these significant impediments, studies conducted now, during the earliest phases of this kind of research, show that phenotypic testing holds some promise for criminal investigation. Much of the current research is performed by an international cadre of scientists known as the VisiGen Consortium, or the International Visible Traits Genetics Consortium, which is led by Dr. Manfred Kayser of the Netherlands and Dr. Tim Spector of the United Kingdom. In October of 2012, VisiGen announced the results of a

study validating the Identitas (v1) Forensic Chip. The chip is the first commercially available tool that can simultaneously analyze genetic information from an array of sites on the genome and produce information related to biogeographic ancestry, eye and hair color, relatedness, and sex. Using 3,196 DNA samples of varying quality intended to emulate some forensic conditions, they found that 95% of samples produced results that were highly accurate for sex and first to third degree relatedness, averaged 94% accuracy for ancestry, 70-85% accurate as to eye color, and ranged from 48% to 85% accurate as to hair color.⁵

The Identitas Chip builds off of other systems already in use that predict eye color with 90%+ accuracy for varying hues of blue and brown eyes and hair color (including red, black, brown, and blonde) with 80-95% accuracy.⁶ A new multiplex, which can test both hair and eye color simultaneously from even low quality, degraded biological samples, showed similarly high rates of success in accurately predicting eye color, and variable success for hair color. When combined with bioancestry, that system (known as HIrisPlex) could predict with upwards of 86% accuracy whether a brown-eyed, black haired individual is of European versus non-European descent.

Oddly enough, age is one of the available characteristics; studies have also recently demonstrated the capacity to predict age with accuracy within 5 years.⁷ Still other studies have shown that facial morphology – such as the size and shape of the nose, lips, eyes, or overall face -- may be predicted from genetic variations, although this work is still developing.⁸ Research seems likely to uncover indicators connected to probabilities for adult body height, male baldness, freckling, and hair morphology.⁹

Of course, scientists are also seeking genetic clues to a number of other medical and psychological traits. In the rare case, some of these might prove indirectly helpful for forensic purposes, such as indicators of genetic disease (say albinism, dwarfism, or sickle-cell anemia), propensity for smoking, left-handedness, or stuttering.¹⁰ More pertinently, industry and academic research also continues to probe genetic predictors of traits like sexual deviance, chemical dependency, or propensity for violence – all of which have a clear and strong nexus to criminal justice. However, findings of this nature with meaningful practical application are likely to be far off. Even assuming genetics exerts a causal influence on this kind of behavior, these regulatory systems are likely to be so complex, and

so subject to environmental influences, that it seems unlikely in the short term that tools to discern such information will be available to law enforcement for an investigation. Nevertheless, it is not inconceivable that some genetic indicators will be uncovered, even if they initially hold low predictive power.

B. How FDP might be useful in criminal cases.

With that understanding of the state of the science, this subpart considers how forensic DNA phenotyping has been utilized thus far, and imagine applications in the immediate future. As an initial point of clarification, it is important to distinguish between the two general sources of biological material that might be tested: in other words, to separately address unknown forensic samples and known individual samples. Unknown forensic samples are those collected from crime scenes, and which typically belong either to a perpetrator or witness. Known samples are collected from identified persons, whether through voluntary or compulsory procedures. The range of potential uses in each category depends, of course, on the sophistication of the science. But for sake of argument, assume a wide array of capabilities.

Testing of unknown crime scene samples would likely focus most on identification of a perpetrator, which means that the relevant genetic tests would be those that help distinguish individuals from one another. Such tests include testing for superficial or visibly apparent traits, and perhaps for distinguishing medical or other traits (such as albinism or a stutter). Once such tests helped isolate a suspect, then additional genetic testing (of the traditional forensic kind) would resolve whether the suspect's profile matched that of the evidence.

The second kind of testing – of samples associated with a known individual – has a broad range of application. Phenotypic DNA testing would be unlikely to be useful to solve past crime – because conventional DNA testing would be a more reliable indicator of guilt or innocence of another offense for which there was genetic evidence. But evidence of genetic predispositions could be viewed as either mitigating or aggravating conditions for decisions about pre-trial release or conditions, sentencing, and even preventative detention.

In light of the current state of the science, the most probable immediate use of phenotypic DNA testing is the former scenario – to test forensic unknown samples in order to identify a possible perpetrator. Indeed, that is the use to which FDP has already been put in a number of limited cases; one researcher has even claimed that it has already resulted in at least six arrests.¹¹ The most prominent example involved the Louisiana hunt for a serial killer linked to at least seven murders. Law enforcement had very little by way of leads; one of their few pieces of information identified the perpetrator as a white male in a white van. DNA linked the homicides to one another, but the perpetrator did not have a profile in any law enforcement database and so could not be identified.¹² Desperate, law enforcement enlisted the help of forensic psychologists and other “profilers,” and engaged in a 1000+ dragnet, but found no leads. They also turned to a new company, DNA Print, based in Sarasota, Florida. That company marketed a service called DNA Witness, which tested biological samples for ancestry, thereby allowing crude conjectures about skin tone. After testing the samples, the company reported that the perpetrator’s ancestry was 85% sub-Saharan African and 15% Native American. Based on their associative studies of ancestry and pigment, they projected that the “skin shade of the subject was most likely of average to darker than average tone relative to that of the African-American group.”¹³

Police refocused their investigation based on this new information. First, they made an association between the serial killer’s profile and another unsolved killing that they had failed to connect because the perpetrator in that case was identified as Black. Next, they compiled a new short list of suspects, which included a troubled man named Derek Todd Lee, who had earlier been brought to police attention, but was ultimately excluded when he did not match the description (including skin tone) offered by the sole witness. A sample from Lee was collected, which turned out to match. Lee fled the jurisdiction, but police used cell phone records to find and apprehend him in another state. In October of 2004, he was sentenced to death for one of the murders.

There are other examples of resort to forensic DNA phototyping in actual criminal investigations, but with much less success. The most commonly cited case occurred in 2004 in the United Kingdom.¹⁴ Investigators sought a burglar and rapist who had committed over 80 offenses, mostly against elderly women, since 1992. DNA evidence linked the cases to a single perpetrator. Again, DNA Witness was brought in, and

conducted a study that determined the overall make-up of the suspect's ancestry to be 82% sub-Saharan African, 6% European, and 12% Native American. Closer scrutiny discerned that one parent was dominantly sub-Saharan, while the other had greater admixture. Moreover, because Native American admixture is uncommon in Europeans, investigators concluded that this parent must have been a recent immigrant. Using U.K. immigration data, which revealed that the Caribbean immigrants from countries with high proportions of Native American influences were Guyana, Trinidad, Belize, Dominica, and St. Vincent. Some of those countries also tended to have European influence, however, which the suspect profile showed very little. Accordingly, investigators surmised that the suspect was most likely the child of an Afro-Caribbean from the Windward Islands area, rather than Jamaica. Police initiated a dragnet, but found no match. The case remains unsolved.

Although this example also suggests that phenotypic testing might ignite racial passions, it can also result in the defusing of such tension. Consider a case in the Netherlands involving the killing of a young Dutch girl. Suspicions focused on a nearby hostel popular with Middle Eastern and Northern African asylum seekers, but testing revealed the killer was likely of western European descent. Similarly, in investigating a killing in Mammoth Lakes, California that at first seemed related to an Asian gang dispute, testing revealed that the perpetrator was of Native American/Hispanic ancestry. That description jogged the memory of a park ranger, who recalled an earlier complaint from a woman about her abusive companion of similar ancestry. A killing in Concord, California was linked to a white male, which focused investigators on a homeless population that gathered in the area.

Each of these cases provides nice illustrations of some of the prior uses of forensic DNA phenotyping, and hint at some of its perils and limitations. They also point toward several additional points critical to an accurate understanding of the potential impact of FDP. First, as the number of unsolved investigations might suggest, DNA Print did not succeed as a business. It closed in 2009, suggesting that if the market is an indicator, then law enforcement did not find its services particularly beneficial. But DNA Print offered one flagship service – the ancestry and skin tone. It also made those predictions with more primitive technology than is currently available. Today, more sophisticated means of

phenotyping are available, capable of producing instantaneous (even on-scene) results across a greater array of physical characteristics with greater precision and certainty.

Nevertheless, even if testing on scene is rapid and readily available, it is unlikely that FDP will become a routine part of every crime scene investigation. The first reason is that such a tool would prove useful only in cases in which the perpetrator is unknown, but the vast majority of crimes are between people who know one another. To be clear, DNA testing is often still valuable in such cases – just not *phenotypic* DNA testing. That is, even in known offender cases, it is often useful to test biological samples to develop evidence supporting the guilt of a known perpetrator (e.g., to test an intimate swab to corroborate the victim’s account that her acquaintance sexually assaulted her). But that kind of testing employs the “junk” identifiers loci, not parts of the genome that express phenotypic traits. In known cases, there seems little benefit to testing swabs for physically identifying information.

Second, even in cases where identity is unknown, it will always be preferable to mine the biological sample for a straightforward forensic profile. In other words, police are more likely to test biological material for the identifying markers used in forensic DNA typing, and then search databases for an exact match to a known individual. Linking DNA to a perpetrator this way is simpler and more precise in every respect; it is only if there is no match to a profile in a DNA database that investigators would even consider turning to phenotypic testing.

This observation provides a nice bridge to the other conceivable use of forensic DNA phenotyping, which would be to test the samples of *known* individuals for reasons other than identity. There is generally no reason to test known samples for externally visible traits like hair or eye color, although there might be some limited useful applications -- for instance, to verify claims related to ancestry or heritage in immigration cases. But generally speaking, the only reason to test known forensic samples would be to discern highly sensitive behavioral or cognitive traits – things like mental illness, chemical dependency, sexual deviance, violence, sociopathy, and so on. The government would no doubt find these kinds of tests very appealing, whether for purposes of bail or detention hearings, sentencing, or civil commitment (such as that for sexual predators or for mental health). But of course they are precisely the kinds of tests that today seem most ambitious,

and most distant. Thus, the likelihood of known sample phenotyping would seem to depend largely on the capacity of scientists to make such associations.

With that foundation, the remainder of this chapter considers the legal and ethical concerns that might influence the actual deployment of forensic DNA phenotyping in criminal cases.

II. Legal frameworks

A. *Statutory*

In one of the few significant treatments of forensic DNA phenotyping in the legal literature, namely a 2008 article by Bert-Jaap Koops and Maurice Schellekens, the authors conducted a survey of positive law regulating forensic DNA phenotyping in eleven countries, using legal resources as well as qualitative assessments in the form of questionnaires sent to experts.¹⁵ Overall, they concluded that “[a]lmost all surveyed countries have some form of DNA legislation, but these laws are generally confined to traditional DNA forensics, that is, making DNA profiles or fingerprints from crime-scene material and comparing these to profiles stored in forensic DNA databases.”¹⁶

With respect to countries other than the United States, only some jurisdictions explicitly restrict the permissible kinds of testing or information-sharing in the genetic context. For instance, a European Council Resolution encourages the exchange of only non-coding genetic information, or loci “not known to provide information about hereditary characteristics” and suggests dropping any loci shown to provide information about hereditary characteristics.¹⁷ In most locations, the convention is to limit DNA testing to non-coding regions of the genome, but no explicit legal authority expressly requires as much.¹⁸ Two states expressly forbid phenotypic testing: Belgium, and Germany for traits other than sex.¹⁹ Only the Netherlands has enacted a statute expressly embracing it,²⁰ although both Japan and the United Kingdom appear to tacitly allow it.²¹ Generally speaking, the Dutch law limits testing to only externally perceptible traits from samples found at a crime scene and alleged to belong to an unknown suspect.²²

The legal status of forensic DNA phenotyping in the United States is largely dependent on state law. Only one U.S. state, Texas, has explicitly authorized phenotypic

testing. In its statute authorizing the creation of a DNA database, it provides that “information contained in the DNA database may not be collected, analyzed, or stored to obtain information about human physical traits or predisposition to disease *unless the purpose for obtaining the information is related to a purpose described in this section.*”²³ The purposes in the section include “investigation of an offense, the exclusion or identification of suspects or offenders, and the prosecution and defense of the case.” Thus, presumably, that statute permits testing of “physical traits or predisposition to disease” if it is for the purpose of locating witnesses or identifying potential suspects, or to otherwise aid in prosecution or defense.

The vast majority of states pattern their DNA collection and database laws after the federal statute, which does not contain any explicit restrictions on phenotypic testing. Rather, that statute (and its state analogs) outlines the permissible scope of collection and testing of *known* offender samples, and restricts testing of such samples to “identification purposes only.”²⁴ But this language falls short in two respects. First, “identification purposes” is easily be interpreted to include typing for more than just non-coding data. After all, hair color or skin tone is “identifying,” as would be “propensity to sexual deviance.” And, in fact, the Federal Bureau of Investigation has interpreted the federal statute as allowing testing for sex, as well as use of genetic profiles to draw inferences about ancestry and family relatedness. Second, the statute only applies to testing of forensic samples in the DNA database. Thus, it potentially leaves room for testing of forensic unknown samples that are not placed in the database.

Although most states follow the federal language, which leaves the interpretive gap just described, there are a handful of states that incorporate language explicitly intended to curtail some forms of phenotypic testing. For instance, ten statutes prohibit certain kinds of testing. New Mexico,²⁵ Rhode Island,²⁶ Wyoming,²⁷ and Indiana²⁸ phrase the language broadly, as covering “human physical traits or predisposition for disease.”²⁹ Six states -- Michigan,³⁰ Vermont³¹, South Dakota,³² Washington,³³ Utah,³⁴ and Florida³⁵ -- frame the restriction more narrowly to cover only “medical or genetic conditions.”³⁶ But again, many of these statutes, by their terms, apply only to samples collected or stored in a database under the authority of the statute, i.e., *known* or “databased” samples (as opposed to samples from “abandoned” items). As such, no statute clearly prohibits phenotyping in the

form in which it is more likely to occur -- collection and testing of crime scene *unknown* samples.

B. *Constitutional*

With respect to the United States, the strongest potential constraint on phenotypic testing for law enforcement purposes emanates from the Fourth Amendment to the Constitution.³⁷ Oddly enough, under current doctrine, the Constitution considers it far more important *what* sample is tested than the *kind* of genetic test that is performed. That is, Fourth Amendment protection is triggered by an initial finding that a person possesses an expectation of privacy in the place or things searched. For this reason, under current doctrine, testing of *known*, compelled DNA samples receives greater protection than testing of unknown crime scene samples and perhaps even known *voluntarily* submitted samples. That is because the Fourth Amendment clearly regulates the compulsory taking of a genetic sample from a person, whereas crime scene samples are conventionally considered “abandoned” material with no Fourth Amendment protection, as are known samples voluntarily relinquished (i.e., given by consent) to police.

But even though it may be clear that the Fourth Amendment will apply to phenotypic testing of known forensic samples that were collected under compulsion, it is not clear exactly *how* it applies. Courts have uniformly upheld testing of such known samples for non-coding genetic information – the “junk” that makes up the DNA database. Significantly, however, the specter of phenotypic testing was raised in most of the known offender cases, and courts either expressed skepticism that such a day would come, or intimated that further review would be undertaken if it did. Of course, because known samples were at issue, these discussions tended to reference highly intrusive forms of phenotypic testing (such as propensity for disease) rather than testing for superficial traits like hair or eye color. Indeed, although it was rarely raised in the cases, routine testing at the time also revealed sex and probable ancestry, and was capable of allowing assumptions about relatedness, but courts shied away from direct scrutiny of those issues.³⁸ Of course, revealing sex or bio-ancestry might not have been considered volatile in the context of known offender samples, where such information is typically visually apparent.

To the extent that law enforcement undertakes to test for sensitive traits some of the ten million plus *known* offender samples in its possession, it seems likely that some added constitutional scrutiny will occur. There might be hiccups in the constitutional doctrine that require resolution – for instance, is a known sample that was collected by compulsion different from a known sample offered voluntarily by the individual or collected surreptitiously by police? How may constitutional doctrine be invoked to invalidate later testing of a lawfully acquired sample, when typically it is acquisition of information, not analysis or use, that the Fourth Amendment regulates?³⁹ But such questions pose no inherently formidable doctrinal obstacle to review.

Indeed, in the argument before the Supreme Court in *Maryland v. King*, the United States was asked a version of just one of these questions. In *King*, the government defended its right to compel DNA from arrestees in part by claiming that the information it revealed was non-sensitive “junk.” Accordingly, the justices pressed counsel about whether the government could test for other, more sensitive genetic information using the (lawfully collected) sample that it retained. In response, the government attorney agreed that it was “probably correct, that the individual will retain a reasonable expectation of privacy in the genomic material that does not reveal identity,” and “additional Fourth Amendment scrutiny would be required before the government could make use of the rest of the genome.”⁴⁰ In other words, the government attorney seemed to accept that just because the Fourth Amendment has generally governed only the *acquisition* of material (here, the sample), and not its subsequent use, in the case of biological material, each test to which the sample is submitted – regardless of the lawfulness of its acquisition – might invite further constitutional scrutiny.

More perplexing, however, is the legal analysis pertinent to the more probable use of phenotypic testing – at least in the near term: testing of forensic *unknown* samples. The crux of the problem is that such samples are historically considered abandoned, and thus outside of the scope of Fourth Amendment protection. Yet while treating crime scene evidence as abandoned for purposes of run-of-the-mine genetic analysis and testing seems sound, surely greater protection might be warranted when it comes to more intrusive, state-of-the-art techniques.

By way of example: suppose police respond to the scene of a homicide and find a bloody wristwatch. They examine it closely: check the style and brand to see if it can be linked to a particular shop or buyer, and take samples to test for fingerprints or a DNA profile that might be associated with a known offender. None of this, under current doctrine, invokes any Fourth Amendment scrutiny, which seems right. It would be perverse to require that law enforcement obtain a warrant or have some suspicion before picking up and analyzing the watch, just as much as it would be odd to require a warrant to examine the weapon left at the scene or the scarf that fell off the suspect as she ran away.

But what about testing for more sensitive information? We can imagine a sliding scale, in which the item is examined first to determine that it has blood, then to type the blood, then to examine it for non-coding identifiers like the classic “DNA profile,” then on to external traits like hair or eye color, ancestry, likely skin tone, or age, then to physical morphologies like cleft palate or non-sensitive medical information like caffeine tolerance, and on to more sensitive characteristics like disease, including perhaps the most sensitive information of all, behavioral predispositions of various kinds. Is all of this testing equally allowed, and equally unregulated, because the sample is abandoned? Does it matter that the scope of what is currently considered “abandoned” goes beyond evidence left at crime scenes, and might include the cup that a person throws away at work or the flap of the licked envelope used to mail in one’s income taxes? Although resource constraints operate as a functional restraint on wide-scale law enforcement collection and typing of DNA samples, there is nothing in current doctrine that would prohibit a police program of collecting DNA samples at random, or without suspicion from targeted populations, and typing them at will. In other words, police now may only collect DNA from crime scenes, but it is departmental priorities that sets that restriction, not constitutional law. If the police were to decide that their time was well spent collecting DNA from the “abandoned” coffee cups or pizza crusts in certain neighborhoods or enclaves, then nothing in doctrine prevents it.

It seems outlandish to suggest that the Fourth Amendment has no stake in regulating law enforcement testing of DNA just because a sample was technically “abandoned,” and that seems true no matter who left it behind or why. Even samples collected at a crime scene and believed to belong to a putative perpetrator may turn out

instead to belong to an innocent bystander, and as such, even if criminal activity is considered a forfeit of genetic privacy, it will often not be certain *ex ante* that the sample came from the criminal.

Consider the analogy, equally vexing, of an iPhone found at a crime scene. Under conventional understandings of Fourth Amendment doctrine, such an item may be deemed “abandoned” and thus searched without restraint. The “crime scene” may be a too-loud party, a bar fight, a DUI, a burglary, or a rape – in all cases, the “abandoned” property becomes law enforcement fair game and police may search the phone and all its contents.⁴¹ But a power to search so untethered to the nature of suspicion is at odds with our constitutional framework. Instead, it points toward the need for greater nuance in Fourth Amendment doctrine in an era when technology allows us one act of absentmindedness (“where did I put my phone again?”) to expose a wealth of intimate information -- e-mails and text messages over the course of years, personal photos and videos, calendar, financial records, diary, reading and shopping habits, travel plans, weigh-loss goals, and the like.⁴²

By analogy, the DNA sample, even though technically “abandoned” and therefore lawfully acquired by police – likewise contains multiple layers of information. At the surface, it holds identifying but non-sensitive information – the “noncoding” loci. Testing for this basic information to ascertain identity seems fair, without constitutional scrutiny. This might be analogized to a look at the iPhone to ascertain who it belonged to, or perhaps even the last calls made. But more intrusive testing – for information that points toward sensitive characteristics like medical conditions or behavioral traits – should perhaps entail additional constitutional scrutiny. Otherwise law enforcement may legally go exploring through a person’s genome simply because the individual threw their coffee cup into the public trash. But if *some* scrutiny is necessary, then that simply begs the question of what kind.

First, precisely which kinds of tests would trigger scrutiny? At present, it does not seem that the answer is “any test that probes genes that express” or “any test that reveals a physical trait.” After all, abandoned samples today are routinely examined today for sex and basic ancestry, and no one has suggested that the Constitution forbids it. Even in those states that forbid testing for physical traits, it seems no challenge has been brought to the routine testing of sex.

Conceivably, then, testing of crime scene samples for some expressed characteristics, namely those that are deemed low sensitivity, might proceed outside of constitutional coverage, leaving only the most sensitive kinds of tests to greater scrutiny. Of course, these are precisely the kinds of analyses that seem unlikely to be available until the distant future, if at all – such as determinations of mental health or behavioral traits.

Of course, just because testing of sensitive traits might invoke some Fourth Amendment supervision, does not mean that it would necessarily be prohibited. In fact, the opposite is often true – the Fourth Amendment normally dictates only that certain procedural standards be met, such as suspicion and a warrant, rather than forbidding outright the law enforcement action. So sensitive testing might be permissible, but only upon meeting certain procedural showings, such as a warrant or probable cause.

At this juncture it is beneficial to observe a key practical reality: there might be only a slim line of overlap between the kinds of tests that would be highly sensitive and those that would be broadly applicable for criminal investigations. In other words, examination of samples for sensitive traits like proclivities to violence, pedophilia, or the like, seems most useful in the *known* context -- such as when an individual is apprehended and the question is whether to grant bail or impose a particular sentence. At the margins, such information may prove of limited utility in a particular case – for instance, to support the evidence that the defendant committed the molestation because he is predisposed to sexual attraction to children. But admission of probabilistic evidence of this kind would be novel, to say the least.

In the context of investigation, however, testing crime scene samples for highly sensitive traits seems unlikely to serve any beneficial purpose. It does not help to learn that an unidentified serial killer is also genetically predisposed to violence -- not just because it is obvious, but because it does nothing to isolate or identify a particular individual as the offender. Such information is not even useful for confirmatory purposes, since of course the better way to ensure that the source of the sample is from the suspected individual is to conduct the more discriminating comparison of the non-coding loci. To be sure, there may be some outlier cases in which sensitive information gleaned from a crime scene sample might help solve a crime – the albino example, or a rare genetic disease that requires regular treatment and medication and so a perpetrator might be identified

through medical records. But for the most part, the kind of information needed to locate an unknown perpetrator is probably most likely to be precisely the kind of *non*-sensitive data that live witnesses ordinarily provide: height, hair, eye, skin, age, etc, rather than highly sensitive, unobservable traits. Such information is all but pointless with regard to testing of *known* samples, whereas the highly sensitive information is of utmost interest.

All of this is to suggest that to the extent that there is a uniform constitutional standard for what kind of samples receive protection, and what kinds of tests are permissible or not, then it is possible that the standards that exist now already map nicely onto a reasonable scope of coverage. In other words, even if crime scene samples are considered abandoned and thus outside of Fourth Amendment protection, the kind of tests that are useful in such situations may be precisely those that are least intrusive. In contrast, known offender samples are the logical candidates for testing of sensitive traits, but they are already subject to added constitutional (or statutory) protection.

One final observation along these lines: If analysis of known samples for sensitive traits were constitutionally forbidden, then it is possible that testing of crime scene samples for that information might be undertaken simply as an end-run around the prohibition. In other words, if law enforcement devises a test for pedophilia, but is prevented from using it on *known* offender samples, then it might seek to test crime scene samples if for no other reason than that such samples often are, eventually, connected to a known offender. Prevented from analyzing the felon's compulsorily collected biological material for a sensitive trait, the police might simply test the crime scene stain later matched to him and find out the same information. For this reason alone, it might be necessary to establish clear boundaries for sensitive testing of biological material from both crime scenes and known individuals, even if such testing is less obviously useful in the former situation.

III. Ethical and other concerns.

Having considered some of the legal frameworks that govern forensic DNA phenotyping, this chapter now turns to consider ethical and other concerns that such testing may raise. Undertaking this kind of inquiry with regard to an evolving science

inevitably entails some measure of speculation. However, for purposes of simplicity, this chapter assumes that FDP functions largely as it does today. In other words, this discussion assumes that FDP can be used to make probabilistic predictions, some very strong and others less so, about basic physical traits and characteristics, but that more complex characteristics, such as those pertaining to mental or behavior traits, are many decades from realization (although not inconceivable).

As the following discussion shows, the “genetic eyewitness” knows both more and less than the human eyewitness, which can be both an ethical advantage and an ethical concern. The advantages are simple: biological samples may provide clues to identity where other leads falter and they are arguably less susceptible (although not immune) to superficial biases.⁴³ Moreover, in a time in which the evidence offered by human eyewitnesses has come under great scrutiny, there is obvious benefit in being able to glean such information with high levels of probabilistic accuracy. So what about the disadvantages?

A. Breaking down the wall between testing of “junk” and expressive loci.

In some respects, the most formidable obstacle to phenotypic DNA testing may be psychological, in that to allow such testing would be to permit law enforcement to breach a perceived rigid boundary between analysis of biological material for “junk” and analysis for expressive purposes. An iron-clad rule that forbids all phenotypic testing and permits only analysis of non-coding loci avoids the difficulty of charting a principled course through the perilous waters demarcating various kinds of genetic information. There is much to be said for the certainty and simplicity of a clear line. So why not simply draw the line there?

As attractive as that solution might be, three reasons argue against it. First, the border between coding and non-coding loci is already, in current practice, more permeable than might seem at first glance. That is partly because we already engage in testing of coding loci, and partly because even non-coding loci can reveal sensitive information. For instance, present forms of testing routinely determine the sex of the individual, along with probable ancestry. In addition, non-coding DNA loci may be used to reveal familial associations, and recent studies have shown remarkably accuracy in predicting surnames,

particularly rare surnames, based on genetic profiles gleaned from the male chromosome.⁴⁴ If sensitive information can be gleaned not just from a particular type of intrusive test, but also from engaging in certain kinds of search techniques or by inference from otherwise non-intrusive testing, then it seems odd to focus singularly on how the information is obtained rather than *that* it is obtained. In other words, the fact that we *do* currently engage in some testing for sensitive information, and that even non-coding loci may reveal private information, suggests that the proverbial ship has sailed as regards any effort to draw a rigid line of either kind. Surely the ascertainment of natural hair color through genetics is less sensitive than sex at birth or biological parentage.

Second, relying simply on the coding/non-coding distinction is unwise because, as the next subpart explains more thoroughly, there are obviously some traits that are quite sensitive, but there may be others that are almost always *not* that sensitive. If law enforcement goals may be profitably served by engaging in some form of phenotypic testing in a manner that poses little to no collateral harm, then such testing is unquestionably worth considering.

Lastly, simply prohibiting testing for expressive information suggests that drawing such a line will alone suffice to rebuff the temptations to encroach on the boundary as technology develops. To be sure, these tests will materialize whether law enforcement directly supports their development or not. Although law enforcement may have a special interests in tests for traits like hair or eye color, a wide array of interests (academic, medical, corporate, etc.) would benefit from advances in genetic understandings of the mechanisms of addiction, mental illness, or other behavioral traits as well as characteristics like ancestry, facial morphology, and the like. And surely defendants will eventually seek to harness some of the power of genomics themselves, which may acclimate courts to hearing such evidence and therefore viewing it as less intrusive or undesirable. Given that forces other than law enforcement are major engines driving advances in the field, and in particular have an interest in developing the very techniques most ripe for ethical or legal challenge when used for law enforcement purposes, it behooves us to think carefully about what we will do when those tests become available. With ten million plus biological samples from known criminal offenders sitting in storage across the nation, it seems

irresponsible simply to assume that if such testing is not authorized then such testing will not be developed, and temptation will never come.

B. Setting new boundaries.

If we elide the distinction between “junk” and expressive DNA, however, then do any boundaries remain? Is any possible genetic test permissible, or should some tests be off-limits no matter how useful for solving crime? Several possibilities suggest themselves.

The most obvious suggestion, and the approach followed by a few of the jurisdictions to embrace phenotypic testing, is to limit such tests to externally visible characteristics, or “EVCs.” The Netherlands even adds a further limitation, and requires that the EVC be visible at birth. The rationale behind allowing only testing for EVCs is obvious; if a trait is visible to anyone who observes the person, and has been so throughout the person’s life, then testing is unlikely to reveal either information unknown to the person himself or herself or to any of that person’s associates.

As a matter of bright lines, then, the EVC standard seems a reasonable one. As a matter of principle, however, the EVC standard leaves much to be desired. The category of “externally visible” may be both under- and over-inclusive. It is under-inclusive because not all non-sensitive genetic information is externally visible. A person may have the gene that allows them to rapidly metabolize caffeine or be immune to bitter tastes; not visible, and yet hardly sensitive if revealed. And adding the “at birth” limitation also precludes typing for traits like age or adult height, which seems arbitrary.

The EVC standard may be over-inclusive because although it may generally speaking be true that an externally visible trait is unlikely to reveal sensitive information about a person, that need not always be the case. At the minor end of the spectrum, genetic testing might reveal whether a person has had plastic surgery, lies about their age, wears a hair piece, or dyes their hair. At the more intrusive end of the spectrum, it could reveal that someone has changed genders, call parentage into question, or cloud a person’s sense of self with conflicting information about ancestry. There is an entire television show devoted to precisely this kind of phenomenon -- taking a prominent member of one demographic community and revealing confirming or disconfirming facts about ancestry. Lastly,

externally visible traits may on occasion have genetic associations to sensitive information – say, a gene for detached earlobes that turns out also to be predictive of Parkinson’s.

Despite these flaws, the external visibility of a trait may function as an effective measure of the permissibility of law enforcement testing by default. That is because it seems likely that, to the extent that it prohibits testing of non-sensitive internal traits, those traits are unlikely to be of much help to law enforcement (like caffeine metabolism). In contrast, most internal traits are precisely those sensitive characteristics that law enforcement would most likely seek (like predisposition to pedophilia or mental illness). Nevertheless, because it is conceivable to imagine that even some externally visible traits might reveal some sensitive information and some internal traits might not, a better taxonomy for considering the issue would key to the *sensitivity* of the trait, rather than its visibility.

Of course, there are additional possible lines of demarcation. For instance, the permissibility of testing might turn on the degree of certainty with which the trait can be discerned. Under this logic, only traits that may be predicted with a certain level of confidence should be tested. This approach appealingly minimizes the risk that law enforcement over-relies on genetic information, but it otherwise seems somewhat arbitrary. An eyewitness may have only minimal confidence in an observation, and yet law enforcement may pursue such leads. If anything, the numerical certainty of phenotypic prediction – the degree to which the investigator will be warned that the assessment is made with only a low amount of confidence – should reassure rather than exacerbate concerns about misuse. Witnesses say that they are certain, but these tests will provide concrete values and confidence intervals.

A second basis for distinction might be the immutability of the characteristic. Immutability seems reasonable both because it suggests that the information gleaned is likely to be of use, and also because it diminishes possible dissonance between the prediction and the characteristic of the individual later identified. In other words, if the trait tested is facial shape, and facial shape is deemed immutable, then it is unlikely that the person ultimately identified positively as the source will have to explain any incongruence between their genetic and actual appearances. But of course, virtually no phenotypic characteristic of immediate use to law enforcement is truly immutable: we can wear

colored contacts, dye hair, get plastic surgery, tan or bleach our skin. The genome may say the suspect is 50, but some 50 year olds look like Demi Moore.

If anything, the very mutability of even some of the most fundamental characteristics – like sex or hair shaft morphology – points toward one of the disturbing aspects of phenotypic testing. Namely, it gestures toward the possibility that efforts to evade law enforcement might necessitate radical reconstruction. Popular culture is rife with images of the fleeing felon dying her hair in the gas station bathroom; but there is also the dystopian image of the protagonist of *The Minority Report* buying a new set of eyes to avoid detection by the iris scans pervade society. Of course, many might find it more comical than disturbing that a fleeing felon now needs a nose job, but it does point to something more ominous about the nature of government omniscience in the 21st century. Still, such fears today remain mostly fanciful, in the sense that most of identifying characteristics are readily masked without much trouble or stigma (like the hair dye), and truly sensitive characteristics are ones unlikely to arise with regularity in the context of police pursuit (e.g., a transgender suspect).

A final basis for consideration relates back to the initial distinction between known samples and testing forensic unknown samples. It certainly seems defensible to forbid the testing of known samples for phenotypic characteristics, on the grounds that there is no identification purpose in such testing and that other values counsel against testing for purposes such as sentencing, preventative detention, or research. But the converse – allowing phenotypic testing of unknown samples – still requires finer tuning. It might be argued that limiting law enforcement to testing of unknown forensic samples works to limit the kinds of tests undertaken, because the truly sensitive tests offer benefits only if the sample source is known. But, of course, an unknown samples may eventually become a known sample. Indeed, that is the goal of testing the unknown sample – to identify the source so that it can be attributed to a person. Accordingly, allowing unfettered testing of unknown samples might simply encourage law enforcement to engage in intrusive tests in the hopes that they prove useful once the sample's source is identified.

C. Concerns that arise even if testing is only allowed for “non-sensitive” traits.

Assume that consensus can be reached that law enforcement should be allowed to test samples for low sensitivity traits, but forbidden from testing for high sensitivity traits, and that assignment of particular characteristics to each category is possible, if imperfect. Would limiting phenotypic testing to low sensitivity traits eliminate all grounds for concern? Certainly the obvious objections, such as fears about invasion of privacy or personal autonomy, would be less of an issue. Moreover, since such traits would effectively mirror the kinds of information supplied by actual eyewitnesses, it may seem that there is no special danger posed by the “genetic eyewitness.”

For the most part, that seems true. However, the fact that information is ascertained through biological means presents special problems worthy of deliberation. In particular, one stark issue arises, and it is the one that plagues all aspects of the criminal justice system: discrimination. At first glance, phenotypic DNA testing may seem race and ethnicity neutral – after all, how can a laboratory test discriminate? Unlike the flawed or biased identifications made by humans, the results of a DNA test would seem mercifully free from racial or ethnic prejudice. But closer consideration reveals several grounds for concern.

First, although a lab test may objectively report what it “sees,” the science itself need not necessarily see objectively. Certain characteristics are more readily discerned than others, which means that biological tests only identify persons with those characteristics. For instance, current tests cannot accurately predict green eyes, and struggle with black and blonde hair colors. In contrast, blue and brown eyes, and red and brown hair colors, may be accurately predicted at rates hovering around 95%.

That means that genetic tests, by dint or design, may prove especially effective at identifying suspects belonging to particular populations. For example, the now-defunct U.K. Forensic Science Service received sharp criticism when it announced that British Afro-Caribbean persons exhibit high rates of genetic variation, making it easier to differentiate members of that population.⁴⁵ It is likewise easy to imagine tests that either by accident or design concentrate only on the characteristics of particular groups; even medical tests may carry political connotations – consider a test for carrying sickle-cell anemia versus one for Tay-Sachs. Or consider that, in the United States, much attention was devoted to discerning

regional variations for different populations of Hispanics – in that case, in order to ensure the accuracy of the frequency tables used to assign meaning to a match.⁴⁶

As it stands, distinctions in the availability of certain tests today rests largely on the capacity of the science – we know how to find blue eyes because they are genetically distinct, not because law enforcement was desperate to develop a test for blue eyed perpetrators. But if phenotypic genetic testing is fully embraced, it might raise the specter of research or funding goals targeted toward developing tests keyed toward the distinct characteristics (whether sensitive or not) of certain sub-populations.

Although serious, these concerns might in some ways be overblown. Consider that the FSS also used to employ a red-hair predictor test. However, the FSS discontinued it due to limited utility – and this in a country concerned about Irish terrorism. It turned out that high accuracy for red headedness was of such limited utility in criminal investigation that to worry about its targeting effects is akin to worrying about the discriminatory effects of a test that could show whether the suspect has excessive body hair. And to the extent that genetic tests do not isolate with specificity *only* those traits belonging to historically targeted groups, then there seems to be little cause for concern.

That genetic tests skew toward the traits of certain vulnerable populations indubitably might generate actual or perceived injustices. But phenotypic testing raises a second, more damaging possible basis for discrimination. Namely, genetic eyes that only see in certain colors run the risk of replicating and inscribing racial and ethnic biases in the criminal justice system, rather than alleviating them. This effect could manifest in three pernicious ways.

If genetic tests are most effective at catching offenders of a certain demographic group, then it is safe to assume that members of other demographic groups will have greater success in committing crime without apprehension. That would mean that the consequences of illicit activity would fall disproportionately on certain groups. But the strongest argument against this claim is that phenotypic testing is always likely to be a last resort rather than a first choice, and it will also inevitably be capable of *some* degree of recognition beyond a specific group. In other words, we certainly would be concerned if phenotyping could only “see” brown skinned, brown eyed perpetrators of African-American ancestry, because that group would then be specially vulnerable to

apprehension. But it is already certain that tests will always sweep more broadly than that. And, in fact, there have already been examples of cases in which ancestry testing undercut assumptions about the possible racial identity of a perpetrator, both to the advantage and disadvantage of typically targeted groups.

But that raises the next concern, which is that phenotypic testing is of such limited utility that it may invite law enforcement dragnets, as has already happened in at least one of the cases in which phenotypic profiles were developed. However, there is nothing special about phenotypic testing that specially predisposes police to conduct a DNA dragnet; in fact, law enforcement has conducted DNA dragnets even without phenotypic characteristics. Indeed, phenotyping has in at least one case help narrow the pool of suspects in a way that might have prevented the dragnet undertaken before the ancestry information was revealed – in the Derek Lee Todd case. To the extent that DNA dragnets are a problem, it seems that phenotypic testing is not a fair tool to blame, and that the issue is more properly handled by forbidding dragnets, rather than forbidding the development of phenotypic profiling.

The most troubling discriminatory effect of phenotypic testing arises from articulating a relationship between genes and expressed traits in the criminal justice context. Allowing crime and criminals to be framed in biological terms opens a conversation that history suggests rarely ends fruitfully. Already the widespread use of racial categories to express the probabilistic significant of a DNA match has helped to inscribe inaccurate ideas about the biological nature of race. Conjectures about skin tone, facial morphology, or group identity drawn from ancestry may further erase the distinction between race as a social category and as a genetic classification. As scientists know, race has no genetic expression; instead, social categories of race at best loosely map onto ancestry patterns that show wide regional variation. In countries like the United States, moreover, the legacy of slavery and rising rates of inter-racial partnership combine to increasingly blur the genetic picture of any individual person even as a gross matter of ancestry. But widespread phenotypic testing that reports results from genetic analysis in terms that ring familiar racial bells – kinky hair, brown skin, wide nose, etc. – are likely to foster dangerous misapprehensions about a biological origin for race.

Viewing race as a biological category when it comes to superficial traits paves the way for the misapprehension that it is a viable basis for distinction for more sensitive characteristics. The public may begin to conflate the “biological” finding that the suspect meets a certain set of descriptors with a supposed “biological” finding that the class to which the suspect belongs is itself genetically distinguishable on other grounds. Put simply, the public may think that if genes can tell us that a Native American committed this crime, then that suggest a possible genetic explanation for Native American crime more generally. To be sure, such generalizations can occur even in the absence of genetic tests; for instance, disproportionate enforcement and policing practices can ingrain the same assumptions. But by injecting a scientific test into the equation, the possibility of a subtle shift to biological or scientific explanations increase, and such assumptions may be more difficult to dislodge given that they require greater sophistication to undercut them effectively.

Even apart from the lack of scientific basis for such a conclusion, this kind of biological speculation may distract from more important conversations about environmental or socio-cultural or political influences. In other words, given a high rate of arrests of young urban minority male for marijuana use, people might start asking after a biological explanations rather than probe police policies. And that is true even where a biological association is founded in legitimate science,⁴⁷ where conversations about observed inter-group differences might end rather than begin with biological distinctions.

This concern about mis- or over-interpretation of even non-sensitive scientific findings can occur even indirectly, simply due to the complexity of the material studied. For instance, one report of genetic sex typing showed a high rate of the male chromosome drop-out among males of Southern Asian ancestry when compared to Europeans. Although there is a clear scientific basis for the dropout, related to the primers used to splice the relevant section of the genome, repeated public misidentification of male perpetrators as female, or even reports of unknown gender due to uncertainty in the genetic tests, might be used as fuel for xenophobic or racist ideas. Similarly, studies suggest that dark skin tone is a common ancestral trait, a thesis consistent with the notion of a single origin of humans from Africa. Yet such findings could be misconstrued to feed racist and white supremacist claims that lighter skin tones signal “more advanced evolution.”

To be sure, the potential for abuse from phenotypic testing is high. We human beings have failed quite miserably when we have tried to use biology as a means of addressing social problems, whether with regard to sterilization programs of the first half of the twentieth-century, the eugenics theories of the mid-century, or the immigration debates that rage today. But before dismissing even non-sensitive phenotypic testing out of hand, it is worth observing that an entirely different set of reactions might unfold.

It is quite possible that phenotypic testing would in fact undermine and uproot, rather than affirm, racist and unfounded ideas. For instance, inevitably law enforcement will get it wrong – they will predict a blue eyed suspect and the actual perpetrator will turn up brown-eyed, or they will identify an individual of strong African ancestry and apprehend a person with porcelain skin. Forced to confront the reality that race and other supposed “natural” categories are in fact far more constructed and indeterminate than otherwise believed, society might begin to loosen its death-grip on those identifiers as biologically ordained features of social ordering. Consider this the Sally Hemmings effect – there is no more of a bigot’s denying that at core we are all the same when only the narrowest biological basis divides a Black family and White family that share Thomas Jefferson as an ancestor.

That is not to say that concerns about discrimination are not serious and real. But rather than reflexively dismiss phenotypic testing as inevitably opening a Pandora’s box of ills, it is worth considering both the extent to which we already engage in limited forms of such testing (by discerning ancestry and sex), as well as gauging the potential for such testing to dislodge, rather than inscribe, erroneous ideas about biology and physical characteristics.

D. Concerns raised by testing of “high sensitivity” traits.

Although it is conceivable that even non-sensitive phenotypic testing may infringe privacy interests or raise concerns about discrimination, the prior subsection suggests that those concerns may be less troubling than might at first glance appear. Rather, the greatest danger of starting down an avowed path of phenotypic testing is unquestionably not that it will reveal superficial characteristics of biological samples of unknown origin, but rather

that it will pave a trail toward testing of *known* samples for far more sensitive traits. Phenotypic testing of unknown samples is the more common point of discussion, but it is not the only important debate. Even the most sensitive medical or behavioral traits have little discriminatory power when it comes to finding a suspect – learning that the perpetrator is probably a male in his twenties with brown hair and brown eyes will not do much to narrow the pool, and only the rare case will arise in which it might help to learn that the suspect stutters or is balding, predisposed to addiction, or likely to one day suffer from a degenerative disease. Testing of unknown samples will simply have limited probative power unless and until facial morphology can be constructed with high degree of certainty – the genetic artist’s sketch. But it’s possible that such a day may never come.

All the talk about testing of unknown samples, however, in some ways distracts from the truly necessary conversation, which is what should be permitted with regard to *known* samples. Although the day may be far away that genetic testing can accurately predict traits like propensity to violence, addiction, pedophilia, or other asocial behaviors, it might not be that far away that genetic testing purports to come close to identifying such things. Already the system relies on crude predictive instruments to make a range of criminal justice decisions – ranging from informal measures like the pre- or post-trial interview that produces a report for bail or sentencing purposes to more formalized methods like exams for mental health or sexual predator commitment. The criminal justice relies on these imprecise tools because certainty is impossible – there is no way to know whether someone will reoffend or relapse, and so we administer hundred-page tests as a way to provide some sense of security.

Is it so hard to imagine, then, that the system would not latch onto even a primitive genetic test for something like violence or sociopathy? Perhaps not as a definitive basis upon which to incarcerate an individual, but certainly as one component of the ultimate determination. If such a conclusion sounds outlandish, then consider the 2011 report of the JASON advisory group to the Department of Defense, which was asked to study the relevance of advances in genetics to various aspects of national security.

The report explains that “[m]any phenotypes of relevance to the DoD are likely to have a strong genetic component,” such as “short- and long-term medical readiness, physical and mental performance, and response to drugs” as well as “phenotypic responses

to battlefield stress, including post-traumatic stress disorder” or tolerance for physically harsh conditions.⁴⁸ The group recommended that DoD “take a leading role in the personal genomics era, and become full partners with industry and academia in creating useful information from genotype and phenotype data.”⁴⁹ Given the “enormous reach-back potential” and comprehensiveness of the military’s collection of Veteran’s Administration health records, the DoD could conduct research on the 3.2 million plus genetic samples that it collects from military personnel aimed at identifying some of those “phenotypes of relevance.” In short, military advisors view phenotypic testing for sensitive traits as a fertile area for research, development, and subsequent decision-making for the benefit both of its soldiers and society at large. It is hard to imagine that criminal justice officials would disagree.

Yet although there might be some benefits, the testing of samples for sensitive traits for criminal justice purposes raises a number of ethical concerns. Of course, the fears outline above, namely those related to discriminatory testing or real or perceived biologization of characteristics, would be greatly exacerbated if the genetic information shifted from relatively benign externally visible traits to sensitive internal characteristics. This is true if for no other reason than that at least an incorrect prediction is immediately and decisively revealed as such when it comes to external traits – if the test says the suspect is a tall, blonde-haired woman and the actual perpetrator turns out to be a short, red haired male, then the genetic mistake is instantly revealed. But predictions for characteristics like violence or sexual predisposition might be harder to prove “untrue” even if they do not match up to a suspect’s history of behavior, and may doom an individual to a path of self-fulfillment by stripping individual agency and replacing it with genetic predestination.

That points to the three greatest concerns with the use of genetic tests to discern sensitive traits, beyond the concerns of discrimination outlined in the preceding Part. To be sure, each of these concerns also could arise incidentally with respect to non-sensitive traits, but they are most acute when it comes to more penetrating information. The first two concerns relate to the rights of the individual to whom the testing relates, whereas the third concerns the system itself.

First, and most obviously, such tests may invade personal privacy by revealing intimate facts. It is one thing to have it declared that a suspect is a redhead; it is another to declare that the person has tendencies toward schizophrenia. Revealing sensitive traits may undermine an individual's personal and professional relationships, or affect one's ability to obtain life or health insurance -- serious intrusions even if the person is a convicted criminal, and arguably indefensible with regard to one only suspected of criminal activity.

Second, and more significantly, typing for sensitive traits usurps individual autonomy in two related ways: it affixes a genetic label onto the individual regardless of the person's own intentions or desires with regard to the information, and it imposes actual or apparent biological determinism that may undermine the individual's exercise of free will. In the ethical literature, these autonomy concepts are often embodied in the notion of a "right not to know."⁵⁰ In the context of medical decisions, this right protects an individual from exposure to information that the person, for whatever reason, has decided not to learn. But even where overridden, the ultimate aim of information is to serve the patient's best medical interest. In contrast, in the criminal justice context, the individual would have no opportunity to exercise such a right, and the aim of such testing is to serve society's interest -- either by using the information to identify a perpetrator or to make more intelligent choices about punishment or detention.

This interest is even more problematic in the criminal justice context because revealing sensitive information of this kind may ultimately undermine some of the goals of the penal system. Labeling a person genetically, say as "violent" or "sexually deviant," may have the ancillary effect of supplying an excuse for asocial behavior. Not only could it potentially push individuals into behavior that they would have otherwise struggled to control, on the theory that it was always genetically predetermined, but it could also create unexpected tensions in their subsequent legal treatment. Phenotypic testing of this kind is inherently probabilistic, but would a genetic inclination provide grounds for mitigation or excuse for criminal behavior? Moreover, what message would it send to other members of the same family tree, who might in turn embrace criminality as a form of biological predestination.

Lastly, testing for sensitive traits is beneficial only inasmuch as that information is used in criminal justice decision-making, and yet relying heavily on probabilistic assessments may challenge notions of procedural due process and fundamental fairness. But of course, the system currently relies upon an array of largely subjective risk assessment instruments with minimal scientific basis. On the one hand, basing decision-making on probabilistic estimates rather than softer forms of analysis, regardless of accuracy, seems anathema to dignitary values that resist reducing individuals to actuarial numbers. Thus turning from psycho-social to biological assessment structures could be viewed as a step up, even in the light of significant error bars.

On the other hand, the current use of bad information to make important decisions should not justify resort to different sorts of equally problematic predictive tools. It may be that interviews and assessments completed by court personnel are poor bases upon which to make critical decisions about incarceration and release, but at least some measure of those tools relies on factors over which the individual exercises free will. To be sure, risk assessment tools often rely on static factors such as “raised in a single-parent household,” but many also employ potentially dynamic inquiries such as “highest level of education” or “current employment.” Moreover, many such instruments rely on acts or offenses committed by the person, which may have stronger predictive power inasmuch as they indicate not just a vague genetic disposition, but a demonstrated willingness to behave in a criminally punishable way. It is in deference to the supremacy of free will that criminal law has typically, as any first year law student can recite, embodied an “act” requirement.

IV. Tentative recommendations.

It is difficult to know where exactly this analysis leads us. Only two conclusions seem clear: phenotypic tests to discern both sensitive and non-sensitive traits will be developed regardless of whether law enforcement specifically goes looking for them, and law enforcement will almost certainly be tempted to exploit those tests once they are available.

Testing unknown samples for non-sensitive phenotypic traits akin to that which might be externally observed by an eyewitness, or even non-sensitive traits that cannot be

observed but pose little risk of embarrassment or shock upon disclosure – the caffeine metabolism example – seems unobjectionable standing alone. We can imagine cases in which such observations will intrude upon the privacy or autonomy of a suspect, but these instances seem manageable. The bigger question is whether it is possible to draw meaningful lines to accept routine phenotypic testing of some kinds, while precluding other such forms. As it stands, there is a strong legal and cultural barrier – albeit it imperfect – between testing of “coding” and “junk” loci. If that line is breached, then can another rise to take its place at the better point of demarcation?

At a minimum, there might be some ways to craft legislation aimed at creating some such barriers, and minimizing the adverse impact of any phenotypic testing that occurs. As a preliminary matter, it might be presumed that phenotypic testing of any DNA sample is forbidden unless a legislature expressly authorizes otherwise. Beyond that, legislative bodies might consider several approaches to circumscribe phenotypic testing.

One comprehensive formulation would allow only phenotypic testing of unknown samples, for non-sensitive traits, based on a multi-factor inquiry into 1) the certainty that the donor of the crime scene sample is the perpetrator of the offense; 2) the nature of the offense and the perceived culpability of the targeted perpetrator; and 3) the perceived utility of the sought information to the investigation. Thus, testing would occur only when the offense involved was serious, the unknown perpetrator played a central role in it, and no other good leads existed. This would ensure that less testing would occur, given that it would not be allowed in minor cases, or to find a witness rather than a perpetrator, or as a first rather than last resort. Alternatively, a more discrete limitations might focus on, for instance, a restriction that only allowed testing of unknown samples. But of course, such a restriction does not address the concern that unknown samples often become *known* samples once the offense is linked to a proven perpetrator.

Another line of regulation might center on the rules surrounding disclosure and retention of test results. Public report of results might be restricted unless established thresholds of reliability (confidence in the asserted phenotype) and utility (capacity of the phenotype to meaningfully narrow the suspect population) were met. Rules might limit the information gleaned from phenotypic testing to investigative, rather than evidentiary, use. Or rules might prohibit phenotypic testing of crime scene samples that would consume the

entire sample, requiring that testing for better discriminating non-coding loci be completed first.

¹ See generally National Institutes of Justice, *Forensic DNA: Alternative Genetic Markers*, Research Projects List, available at <http://www.nij.gov/topics/forensics/evidence/dna/research/alternative-markers.htm> (“Alternative genetic markers and assays can potentially provide further information about biological samples under investigation, such as an estimation of ethnic origin, physical characteristics, and skin, hair, or eye color.”); see also Grant to GWU, *Ancestry/Phenotype SNP Analysis & Integration with Established Forensic Markers* (2011), available at <http://grants.ojp.usdoj.gov:85/selector/awardDetail?awardNumber=2011-CD-BX-0123&fiscalYear=2011&applicationNumber=2011-90385-DC-IJ&programOffice=NIJ>; Grant to Akonni Biosystems Inc., *A Microfluidic Microarray Instrument to Type SNPs for Physical Appearance* (2010), available at <http://grants.ojp.usdoj.gov:85/selector/awardDetail?awardNumber=2010-DN-BX-K209&fiscalYear=2010&applicationNumber=2010-91205-MD-DN&programOffice=NIJ>.

² Kevin M. Beaver, *Intersection of Genes, the Environment, and Crime and Delinquency: A Longitudinal Study of Offending*, National Criminal Justice Reference Series (Aug. 2010), available at <https://www.ncjrs.gov/app/publications/abstract.aspx?id=253671> (studying five genes for interactions).

³ Of course, not every genetic trait is this complex. For example, sickle-cell anemia is highly predictable because it depends on the variants of single gene; one mutant copy and one normal copy manifests as resistance to malaria, whereas two mutant copies manifests as sickle-cell anemia. However, many phenotypic traits involve much more complex articulation. In light of this, two recent research trends have developed. First, the falling costs of whole-genome sequencing has led to the emergence of genome-wide association studies (GWAS), which are undertaken as a form of backwards reasoning from a trait to possible involved genetic markers. Second, as researchers have learned that gene expression is not just a function of genetic sequence, but also to other aspects of genetic function, a new field of epigenomics has emerged that studies such influences. See Adrian Bird, *Perceptions of epigenetics*, 447 Nature 396 (May 24, 2007). One article describes epigenetics as “a potential antidote to genetic determinism” that has “a deliciously Lamarckian flavor.” *Id.*

⁴ See, e.g., Mario F. Fraga, et al., *Epigenetic differences arise during the lifetime of monozygotic twins*, 102(30) PNAS 10604 (July 26, 2005) (observing that “epigenetic differences may be an important part” of understanding “twin discordance for common diseases and traits”); Albert H.C. Wong, et al., *Phenotypic differences in genetically identical organisms: the epigenetic perspective*, 14 Human Molecular Genetics R-11 (2005) (“Epigenetic mechanisms may explain paradoxical findings in twin and inbred animal studies when phenotypic differences occur in the absence of observable environmental differences....”).

⁵ Brendan Keating *et al.*, *First all-in-one diagnostic tool for DNA intelligence: genome-wide inference of biogeographic ancestry, appearance, relatedness, and sex with the Identitas v1 Forensic Chip*, *Int. J. Legal Med.* (Nov. 13, 2012). The study was complicated by several factors, including the use of self-reporting rather than objective reporting. Thus, for instance, accuracy rates were likely lowered due to counted “misses” that were in fact the product of self-reports that might not match objective observation.

⁶ Wojciech Branicki, *et al.*, *Model-based prediction of human hair color using DNA variants*, 129 *Human Genetics* 443 (2011); Susan Walsh, *et al.*, *IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information*, 5(3) *Forensic Sci. Int’l Genetics* 170 (2011); Susan Walsh, *et al.*, *DNA-based eye colour prediction across Europe with the IrisPlex system*, 6(3) *Forensic Sci. Int’l Genetics* 330 (2012); Susan Walsh, *et al.*, *The HIRISplex System for simultaneous prediction of hair and eye colour from DNA*, 7(1) *Forensic Sci. Int’l Genetics* 98 (2013).

⁷ Sven Bocklandt, *Epigenetic Predictor of Age*, 6(6) *PLOS ONE* e14821 (June 22, 2011); D. D. Zubakov *et al.*, *Estimating human age from T-cell rearrangements*, 20(22) *Current Biology* (2010). Interestingly, one of the genes used in the Bocklandt study has been shown to have a relationship to pancreatic cancer and Parkinson’s disease.

⁸ Lavinia Paternoster *et al.*, *Genome-wide association study of three-dimensional facial morphology identifies a variant in PAX3 associated with nasion position*, 90(3) *Am. J. Hum. Genetics* 478 (2013); Fan Liu, *et al.*, *A genome-wide association study identifies five loci influencing facial morphology in Europeans*, 8(9) *PLoS Genet* e1002932 (2012).

⁹ Walsh, *HIRISplex*, *supra* note 5, at 99 (citing studies). Studies suggest that heritability of some common traits ranges from 100% for features like hair and eye color, bitter taste capacity, and tongue rolling, to around 50% for moles, migraine susceptibility or finger ratios. Brendan Keating, *Promega*, *supra* note X, at 7.

¹⁰ Bert-Jaap Koops & Maurice Schellekens, *Forensic DNA Phenotyping: Regulatory Issues*, 9 *Colum. Sci. & Tech. L. Rev.* 158, 165 (2008).

¹¹ Tony Frudakis, *Molecular Photofitting: Predicting Ancestry and Phenotype Using DNA* 614 (2007).

¹² *Id.* at 599-605.

¹³ *Id.* at 602.

¹⁴ *Id.* at 603.

¹⁵ Koops & Schellekens, *supra* note 9, at 166. The survey, while global in nature, did focus attention on those countries that have demonstrated an interest and capacity to engage in routine forensic testing. They received written replies from experts in Belgium, France, Spain, Australia, Israel, South Africa, and Brazil.

¹⁶ *Id.*

¹⁷ *Id.* at 166 (citing Council Resolution, art. III(1), 2001 O.J. C 187/1) and art. III(2), 2001 O.J. C 187/1). *See also* Prüm Convention, Convention on the Stepping Up of Cross-Border Cooperation, Particularly in Combating Terrorism, Cross-Border Crime and Illegal Migration art. 2(2), May 27, 2005, Council Secretariat, Brussels, 7 July 2005, Document no. 10900/05.

¹⁸ Koops & Schellekens, *supra* note 9, at 167 (citing Australia, South Africa, and Spain).

¹⁹ *See id.* Koops & Schellekens cite to Art 44ter § 1 Code of Criminal Procedure (Belgium) (“Comparative DNA investigation under this code may have only the purpose of comparing

DNA profiles of bodily cell material found or taken in order to identify directly or indirectly persons involved in a crime. This comparative investigation can only involve non-coding segments.”) and Section 81e StPO [German Code of Criminal Procedure] (“Material obtained ... may also be subjected to molecular-genetic examinations insofar as such measures are necessary to establish descent or the criminal act, to ascertain whether traces found originate from the accused or the victim; in doing so, the gender of the person may also be determined...”).

²⁰ *Id.* at 168-69 (citing Wet van 8 mei 2003 tot wijziging van de regeling van het DNA-onderzoek in strafzaken in verband met het vaststellen van uiterlijk waarneembare persoonskenmerken uit celmateriaal [Act of 8 May 2003 To Adapt the Regulation Of Forensic DNA Investigation In Relation To Determining Externally Perceptible Personal Characteristics From Cell Material], Staatsblad van het Koninkrijk der Nederlanden 201 (2003) (Neth.)). The statute limits the cases in which such analysis may occur to those in which pre-trial detention is allowed. It further requires the characteristics be externally perceptible as visible from birth, and forbids disease-related characteristics unless visible at birth.

²¹ *Id.*

²² *Id.*

²³ Tex. Gov’t Code Ann. § 411.143.

²⁴ 42 U.S.C. § 14132(b)(3)(A) (2006).

²⁵ New Mexico Admin Code § 10.14.200.11 (“The information contained in the DNA identification system database shall not be collected, stored, or released for the purpose of obtaining information about physical characteristics, traits, or predisposition for a disease or mental illness or behavior and shall not serve any purpose other than those specifically allowed by the DNA Identification Act.”).

²⁶

R.I. Gen. Laws § 12-1.5-10(5) (2007) (“DNA samples and DNA records collected under this chapter shall never be used under the provisions of this chapter for the purpose of obtaining information about physical characteristics, traits or predispositions for disease.”);

²⁷ Wyo. Stat. Ann. § 7-19-404(c) (2007) (“[I]nformation contained in the state DNA database shall not be collected or stored for the purpose of obtaining information about physical characteristics, traits or predisposition for disease.”).

²⁸ Ind. Code Ann. § 10-13-6-16 (West 2004) (“[I]nformation contained in the Indiana DNA data base may not be collected or stored to obtain information about human physical traits or predisposition for disease.”).

²⁹ *Id.*

³⁰ M.C.L.A. § 28.175a (“(2) DNA samples provided under this act shall not be analyzed for identification of any medical or genetic disorder.”).

³¹ Vt. Stat. Ann. tit. 20, § 1937(b) (“(b) Analysis of DNA samples obtained pursuant to this subchapter is not authorized for identification of any medical or genetic disorder.”)

³² S.D.C.L. § 23-5A-17 (“Analyses of DNA samples obtained pursuant to this chapter are not authorized for identification of any medical or genetic disorder.”).

³³ West’s RCWA § 43.43.753. Washington’s statute provides a policy statement that declares that “The legislature further finds...Washington state patrol has no ability to

predict genetic disease or predisposal to illness. Nonetheless, the legislature intends that biological samples collected ...be used only for purposes related to criminal investigation, identification of human remains or missing persons, or improving the operation of the system....” This language at least implicitly suggests that the “purposes” clause is limited to testing that does not reveal disease or illness.

³⁴ U.C.A. § 53-10-406(1)(f) (imposing duty to “ensure that the DNA identification system does not provide information allowing prediction of genetic disease or predisposition to illness”).

³⁵ F.S.A. § 943.325(13)(b) (“The analyses of DNA samples collected under this section shall be used only for law enforcement identification purposes or to assist in the recovery or identification of human remains or missing persons and may not be used for identification of any medical or genetic condition.”).

³⁶ *Id.*

³⁷ Other constitutional provisions might conceivably come into play, such as the Equal Protection Clause, the Fifth Amendment (although that typically applies only to testimonial evidence), or even some idea of substantive Due Process. But given that the Fourth Amendment most directly restricts the policing power, it provides the likeliest source of restraint.

³⁸ Erin Murphy, *Relative Doubt: Familial Searches of DNA Databases*, 109 Mich. L. Rev. 291, 292-93 (2010).

³⁹ *But see Ferguson v. City of Charleston*, 532 U.S. 67 (2001) (striking down program of testing of urine samples of pregnant women despite argument that samples were abandoned or lawfully obtained).

⁴⁰ *Maryland v. King*, No. 12-207, Oral Arg. Tr. at 16 (Feb. 26, 2013).

⁴¹ There is a limited amount of support for the proposition that property accidentally or inadvertently abandoned still receives some Fourth Amendment protection. *See, e.g.*, *State v. May*, 608 A.2d 772 (Me. 1992). *But cf. People v. Daggs*, 133 Cal. App. 4th 361 (2005); *United States v. Oswald*, 783 F.2d 663 (6th Cir. 1986); *State v. Huerta*, 223 Ariz. 424 (Ariz. App. 2010). However, even assuming such a principle, it nonetheless has uneven application with regard to DNA. Is the forgotten hat now covered by the Fourth Amendment? Is the hat “abandoned,” but not the skin cells on the brim? Are skin cells on a hat “abandoned,” but not those left on the light switch that was turned off, on the theory that the hat was a visible vector of biological material, whereas ordinary “shedding” is uncontrollable?

⁴² Orin Kerr recently proposed one such limit. Drawing from the *Gant* doctrine for automobile searches, which requires that officers limit their search for contraband related to the crime of arrest, Orin S. Kerr, *Foreword: Accounting for Technological Change*, 36 Harv. J. L. Pub. Pol’y 403 (2013).

⁴³ *See, e.g.*, Dov Fox, *The Second Generation of Racial Profiling*, 38 Am. J. Crim. L. 49, 77-78 (2010) (“The prospect of reliable DNA phenotyping gives occasion to rethink the prominent role that race plays in the investigation of crime.”).

⁴⁴ Koops & Schellekens, *supra* note 9, at 163. In a small-scale investigation in Britain, the correct surname could be “predicted” from DNA in 19% of the cases. If a surname is less common, this percentage is higher. The 80 least occurring surnames involved in the investigation were correctly predicted in 34% of the cases. *Id.* ((citing Bryan Sykes &

Catherine Irven, Surnames and the Y Chromosome, 66 *Am. J. Hum. Genetics* 1417 (2000)); see also Uta-Dorothee Immel et al., *Y-chromosomal STR haplotype analysis reveals surname-associated strata in the East-German population*, 14 *Eur. J. Hum. Genetics* 577, 580 (2006).

⁴⁵ Koops & Schellekens, *supra* note 9, at 173, 190.

⁴⁶ One study of African-American, European-American, and Hispanic haplotypes in the United States concluded that there were no significant variation among the frequency of observed haplotypes according to geography. In other words, East Coast populations genetically resemble West coast, Midwest, and Southern populations. Manfred Kayser, et al., *Y Chromosome STR Haplotypes and the Genetic Structure of U.S. Populations of African, European, and Hispanic Ancestry*, 13 *Genome Research* 624 (2003). There is one exception – samples from Texas exhibit heterogeneity to the haplotypes of Whites and Hispanics.

⁴⁷ For example, consider a study in the United Kingdom that investigated a genetic basis, among others, for observed higher rates of schizophrenia in a particular immigrant population. Although researchers have concluded that there was no genetic link, even preliminary findings to the contrary might be viewed as legitimating racist or biased policies of exclusion, and result in the stigmatization of certain groups with regard to housing, healthcare, employment, and the like. Rebecca Pinto et al., *Schizophrenia in black Caribbeans living in the UK: an exploration of underlying causes of the high incidence rate*, 58(551) *British J. Gen. Pract.* 429 (2008).

⁴⁸ JASON, The Mitre Corporation, *The \$100 Genome: Implications for the DoD*, JSR-10-100 (Dec. 2010), at 43.

⁴⁹ *Id.* at 5-6.

⁵⁰ Unlike the right of informed consent, there is a debate in the literature as to the defensibility of a right not to know. Compare David E. Ost, *The “Right” Not to Know*, 9 *J. Medicine & Philosophy* 301 (1984) (arguing against the right, on grounds that autonomy requires rational self-determination, and likening to Mill’s justified paternalism), with Mark Strasser, *Mill and the Right to Remain Uninformed*, 11 *J. Medicine & Philosophy* 265 (1986) (agreeing with rationality requirement, but using Mill and harm principle to defend right to remain ignorant).