The genetic witness: forensic DNA phenotyping

Michele Van Laan¹

Abstract:

The underlying foundation of Forensic DNA Phenotyping (FDP) for Externally Visible Characteristics (EVCs) is that our DNA holds the key to our entire physical make-up, including our visible features. Forensic DNA testing has greatly impacted criminal investigations worldwide for the past three decades. As technology progresses, so too do the capabilities of DNA analytical techniques. FDP is a developing analytical process used to enhance the value of unknown DNA samples collected from crime scenes. The goal of FDP is to identify the perpetrator and eliminate innocent persons quicker than ever before. This article addresses the recent developments and current capabilities of FDP technology and its potential value in the context of criminal investigations. Finally, this article discusses the current technological capabilities of DNA analysis, future developments, legislative implications, and the most common criticisms of FDP and the potential solutions to those issues.

Keywords: forensic science, forensic DNA phenotyping, externally visible characteristics

¹ Forensic Investigations Student, Wayne State University, Eugene Applebaum College of Pharmacy and Health Sciences, 259 Mack Avenue, Detroit, Michigan, United States

Communicating Author Contact: Michele Van Laan, email: michelevanlaan@gmail.com
Discovery and Development of DNA Analysis

Prior to the discovery and development of DNA testing, forensic biological samples most often underwent serology processing, such as ABO blood typing and polymorphic isoenzyme analysis, in an attempt to reduce the number of potential donors of the sample. By doing so, investigators could eliminate all suspects who were not possible donors and instead focus on those who were. But while serology aided in reducing the donor pool, it could not individualize to a single donor. Moreover, the process of typing samples was based on statistics with little consideration for varying and unique population groups. Research has shown that many convictions based on testimony regarding serological findings in criminal cases proved to be inaccurate when DNA testing was performed on the same biological evidence years later. In fact, FBI research indicates nearly one-third of cases determined by serology would likely be excluded by DNA testing. The remarkably high error rate was most often due to the overstatement of serological findings and had devastating repercussions for innocent individuals convicted of offenses they did not commit, creating additional pressure for better testing methods.

Deoxyribonucleic acid, more commonly referred to as DNA, is the integral underlying structure of nearly all life forms and has been the star of the forensic investigation field for decades. In 1984, Sir Alec Jeffreys, a British geneticist studying inherited diseases, discovered DNA’s ability to individualize humans. Testing techniques have evolved over the last three decades, beginning with Restriction Fragment Length Polymorphism (RFLP). Using the RFLP method, a highly discriminatory and specific DNA analysis process, could successfully individualize at just four or five hyper-variable locations. Unfortunately, the sequence lengths for RFLP were so long that at least 250 nanograms of a good quality DNA sample was needed to produce accurate results, which was difficult to obtain from a forensic setting. Although the results were valuable, the testing process was extremely time-consuming, as a single analysis could take up to one month to complete.

The RFLP analysis method was eventually superseded by the Variable Number Tandem Repeat (VNTR) analysis method. VNTR studied only six loci within the DNA sequence, which reduced the size of the sample necessary to perform an analysis. The VNTR method was also more successful in processing degraded DNA samples because of the limited number of alleles necessary to conduct the analysis. However, although the VNTR method was an improvement from the original RFLP method, researchers still sought to reduce the sample size necessary for DNA analysis even further, and increase the individualizing capability of DNA testing.

As a result, Polymerase Chain Reaction of Short Tandem Repeat sequencing (PCR-STR) was developed. PCR-STR is the current and most effective DNA analysis method in use. The PCR process allows analysts to replicate DNA material, which reduced the size of the sample needed for processing while still providing accurate results. Specifically, the PCR-STR method only requires one-tenth of a nanogram- one ten-billionth of a gram- of DNA for testing. Once the PCR process has amplified the DNA sample, the STR process is then used to analyze that sample. Current STR analysis involves the examination of twenty specific loci, which are highly individualizing, both within and between individuals. The PCR-STR procedure is much quicker and more automatable than previous testing procedures, which has reduced the cost and resources necessary to perform the analyses. The STR method is much more sensitive and specific than the other methods, which, thereby increases the degree of individualization.
Although the PCR-STR method is the most commonly used method today, it is not the only DNA analysis method currently employed. Other techniques currently in use include Mitochondrial DNA (mtDNA), Y-chromosome Short Tandem Repeat (Y-STR), low copy number (Touch DNA) analyses, and familial searching based on the Y chromosome. As a result of technological advancements, most of these modern methods are sensitive enough to individualize a questioned sample even if there is not enough of that sample to identify what the substance is.

As DNA analysis progressed and became a reliable investigative tool for law enforcement, the FBI developed the Combined DNA Index System (CODIS). The CODIS database was originally developed as a software project in 1990, but was enhanced when the National DNA Index System (NDIS) was established in 1994. NDIS is the national component of CODIS, which contains the DNA profiles submitted by participating federal, state, and local forensic laboratories throughout the United States. The connectivity of CODIS allows all levels of law enforcement to exchange and compare information within NDIS, aiding in case linkages using DNA profiles. Resources are being invested to restructure the current software format to include the results from newer DNA analysis processes, including mtDNA, Y-STR, and mini-STR. As of December 2016, NDIS contained over 12.6 million offender profiles, 2.5 million arrestee profiles and 750,000 forensic profiles. To date, CODIS has produced more than 359,000 hits and assisted in over 344,000 investigations since the system began tracking its success. Empirical evidence such as this demonstrates the incredible value DNA evidence contributes to the criminal justice system.

**Forensic DNA Phenotyping (FDP)**

New DNA analyses are on the horizon and what was previously thought to be only scientific ideals are breaking through into reality. Forensic appearance prediction from DNA began early in the 21st Century and progressed slowly at first. However, DNA phenotyping is rapidly becoming a popular niche in DNA research, with specific intents for forensic applications.

Forensic DNA phenotyping (FDP) for Externally Visible Characteristics (EVCs) is of particular interest due to the great potential it holds for assisting in solving forensic cases, particularly those with no other promising leads. FDP is intended to be applied only to forensic DNA samples from unknown persons in cases where potential suspects cannot be eliminated by other investigative means. Once the focus is restricted to a smaller group of potential suspects with the EVCs predicted by FDP, conventional DNA testing would then be employed to confirm individual identification between the forensic sample and the suspect. The goal of FDP for EVCs is to enable analysts to develop a unique and distinguishable physical profile from an evidentiary DNA specimen, most commonly that left behind at a crime scene by the perpetrator.

Much of the scientific developments involving FDP have been achieved by a select group of researchers, namely Drs. Manfred Kayser, Tim Spector, and Susan Walsh. Dr. Kayser leads the forensic molecular biology department at Erasmus University Medical Center in Rotterdam, Netherlands. His research has laid the foundation for many advances in biological analysis, including the identification of male DNA using the Y chromosome. Dr. Spector is a professor of Genetic Epidemiology at Kings College in London, England, and the director of the TwinsUK Registry. Dr. Spector’s work has been instrumental in the study of the relationship between
genetic markers and EVCs. Dr. Walsh is a forensic geneticist at Indiana University-Purdue University in Indianapolis, Indiana, and she has been awarded grants from the U.S. Department of Justice to develop and improve DNA intelligence tools that can aid in the identifying unknown criminal suspects and unidentified human remains.

Other notable contributors include Drs. Mark D. Shriver, Peter Claes, and Wojciech Branicki. Dr. Shriver is a professor of anthropology and genetics at Pennsylvania State University and Dr. Claes is a research expert in morphometrics at the Medical Image Computing laboratory at KU Leuven in Belgium. Drs. Shriver and Claes “developed a complex mathematical method to represent faces, based on measuring the three-dimensional coordinates of more than 7,000 points on the face”, which is the basis for FDP imaging computer software. Dr. Branicki is a researcher in the Department of Genetics and Evolution at Jagiellonian University and a DNA expert at the Institute of Forensic Research, both of which are in Krakow, Poland. Dr. Branicki has completed many research studies related to the identification of genetic markers responsible for the prediction of EVCs.

The underlying foundation of FDP for EVCs is that our DNA holds the key to our entire physical make-up, including our visible features. Therefore, FDP is akin to a biological blueprint, allowing for predictions of physical feature developmental propensities. Tiny variations in the human genome, referred to as single nucleotide polymorphisms (SNPs), are responsible for the biologically-influenced development of physical features. There are millions of SNPs throughout the human genome; a single SNP is one nucleotide in the genome at a certain location. Phenotyping is based on the understanding that specific DNA segments have a high correlation with the expression of certain characteristics. Although genes do not actually code for specific characteristics, they play a large role in the expression of physical traits. Externally visible phenotypic characteristics are the unique and complex physical traits displayed in an individual, which are influenced by both genetic and environmental factors.

There are two forms of DNA phenotyping: indirect and direct. Indirect phenotyping traces the external physical features of an individual indirectly by first determining the individual’s geographic or ethnic origin. In contrast, direct phenotyping is the analysis of an individual’s genotype related to external physical characteristics. FDP for EVCs employs a combination of both approaches and “focuses exclusively on portions of the DNA molecule that control or contribute to physical characteristics.” DNA phenotyping uses PCR-STR to identify SNPs and focuses on areas that code for specific physical characteristics.

Most FDP advancements to-date are a product of genome-wide association (GWA) and population studies. GWA studies are utilized to locate genes involved in the expression of complex traits, and population studies provide the data for comparison and interpretation of common biologically-influenced physical characteristics. Sanger technology was used to complete the first sequencing of the full human genome, which took 13 years and cost $2.7 billion. In 2005, 454 Life Sciences, located in Branford, Connecticut, developed the Genome Sequencer 20 platform, which was the first commercial technology available to sequence the human genome, which takes just five months and costs $1.5 million, a mere fraction of the resources previously required. The speed at which the human genome can be sequenced subsequently increases the speed at which researchers can examine and identify individual SNPs related to the expression of EVCs.

FDP software programs, such as SNaPshot, analyze data from FDP profiles and create a computer-generated composite photograph using the most likely physical features of the sample donor. As with eyewitness-based composite sketches, FDP composites can help direct an
investigation toward individuals who possess the physical characteristics outlined in the DNA sample and eliminate suspects who do not possess those features. The added benefit of FDP, as opposed to information gleaned from eyewitnesses, is the improved accuracy and reliability of the information because FDP data is statistically supported and not influenced by preconceptions or circumstances of a crime. The ability to eliminate persons of interest in a case is just as important as focusing the investigation on the most likely suspects, as it clears extraneous information from the investigation and directs investigative resources towards relevant leads. Therefore, the use of FDP could accelerate investigations and ensure a more prompt release of innocent suspects. With this technology, DNA evidence can be used more effectively at the beginning of the investigation process rather than merely waiting until the end to identify an individual to a particular crime scene.

The EVCs most accurately predicted are eye, hair, and skin color as well as geographic ancestry. However, some accuracy can be achieved in the prediction of many other characteristics, including biological age, facial morphology, freckling, adult body height, hair morphology, male baldness and hair loss, and genetic diseases.

<table>
<thead>
<tr>
<th>TABLE 1: Prediction Accuracy Rates for Commonly Tested EVCs</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Gender &lt;sup&gt;20&lt;/sup&gt;</td>
</tr>
<tr>
<td>o 100%</td>
</tr>
<tr>
<td>• Race and skin coloring &lt;sup&gt;20&lt;/sup&gt;</td>
</tr>
<tr>
<td>o Caucasian: 90%</td>
</tr>
<tr>
<td>o Asian: 88%</td>
</tr>
<tr>
<td>o African: 71%</td>
</tr>
<tr>
<td>• Iris color (brown and blue) &lt;sup&gt;20&lt;/sup&gt;</td>
</tr>
<tr>
<td>o 76-99%</td>
</tr>
<tr>
<td>• Hair color (blonde, brown, and red) &lt;sup&gt;20&lt;/sup&gt;</td>
</tr>
<tr>
<td>o 90%</td>
</tr>
<tr>
<td>• Adult height &lt;sup&gt;20&lt;/sup&gt;</td>
</tr>
<tr>
<td>o 65%</td>
</tr>
<tr>
<td>• Age prediction within five years using methylation status approach &lt;sup&gt;24&lt;/sup&gt;</td>
</tr>
<tr>
<td>o Age group 2-19 years: 86.7%</td>
</tr>
<tr>
<td>o Age group 60-75 years: 50%</td>
</tr>
</tbody>
</table>

There are currently three major commercial companies specializing in FDP: Illumina, Identitas, and Parabon NanoLabs Inc. Illumina, the largest manufacturer of DNA sequencers, recently introduced a forensics product capable of predicting a small number of physical traits and performing conventional DNA profiling. Illumina products use Next-Generation Sequencing (NGS) technology, which utilizes automated, high-throughput equipment to analyze a sample, as well as interpret and review the results. Illumina’s NGS technology “uses clonal amplification and sequencing by synthesis (SBS) chemistry to enable rapid, accurate sequencing. The process simultaneously identifies DNA bases while incorporating them into a nucleic acid chain. Each base emits a unique fluorescent signal as it is added to the growing strand, which is used to determine the order of the DNA sequence.” This technology has traditional medical applications, including disease detection and identification, as well as forensic applications relevant to EVC predictions.
Identitas offers their IDentify software technology, which uses advanced genetics and an 800,000 SNP point database to predict a donor’s bio-geographic origin and physical characteristics with reliable accuracy.\textsuperscript{26} IDentify has the capability of predicting gender, hair and eye colors, bio-geographical ancestry and relatedness up to third degree relationships.\textsuperscript{26} In addition, Identitas is currently developing a program that can identify a single DNA profile from a mixed sample of up to ten contributors, which holds tremendous potential for forensic application.\textsuperscript{26}

Parabon NanoLabs in Reston, Virginia, utilizes computer software programming developed by Drs. Shriver and Claes in their aptly named program, SNaPshot, which has been in use for over a decade.\textsuperscript{14,21} SNaPshot is funded by the U.S. Department of Defense, uses Next Generation Mini-Sequencing technology to look at more than 20,000 DNA markers within a sample, and uses complex algorithms to compare the findings with over 10,000 donor profiles on file.\textsuperscript{15,27} The system then combines the raw data derived from the analysis to “reverse engineer an image of the suspect.”\textsuperscript{15} Parabon NanoLabs extensive database of trait connectivity is the result of research involving volunteer subjects who divulge their physical characteristics, submit a DNA sample that is examined at approximately one million locations, and undergo a 3-D facial scan to document the structure, composition, and dimensions of their facial features.\textsuperscript{16} The SNaPshot system then combines art and science to render a computerized visualization of an unknown individual, referred to as a forensic facial reconstruction, based on the phenotyping of the suspect’s DNA.\textsuperscript{28,29} Since FDP can only account for the biological component of physical characteristic expressions, a forensic artist is needed to incorporate any non-biological characteristics of the suspected perpetrator that were discovered through other investigation techniques, such as age, hair style, facial hair, scars, tattoos, or accessories like eyeglasses.\textsuperscript{28} SNaPshot analyses are limited to the prediction of EVCs, including genetic ancestry, eye color, hair color, skin color, freckling, and facial shape in individuals from any ethnic background, including those of mixed ancestry.\textsuperscript{29} SNaPshot technology holds promising potential for many forensic applications, including the analysis of unknown DNA specimens located at crime scenes, unidentified human remains, non-human samples (entomology, microbial, and wildlife), and organs recovered in human trafficking crimes.\textsuperscript{21} A single SNaPshot analysis costs approximately $5,000.\textsuperscript{30}

\textbf{SNPs Versus STRs}

FDP methods that focus on analyzing SNPs supersede STR technology in numerous ways, making it a viable option to replace the current STR approach in the future. SNP analysis uses very short fragments of DNA to allow for successful analysis of low amounts of highly degraded and fragmented DNA samples, something STRs processing cannot achieve given their highly polymorphic repetitive sequences.\textsuperscript{31} SNPs also have a significantly lower mutation rate and are shorter in length than STRs, thereby reducing the possibility of inter-sample contamination during the PCR process.\textsuperscript{21} Furthermore, there are far more SNPs than STRs within an individual cell, resulting in more SNPs being available for analysis in any given DNA sample, which improves the likelihood of producing a full profile with lower quality and quantity specimens than possible with the analysis of STRs.\textsuperscript{21}

However, there are some distinct drawbacks of SNP analysis, including lower discriminating power and limited success with mixed samples.\textsuperscript{21} Although there are many benefits to utilizing SNPs rather than STRs, it is unlikely SNPs will replace STRs in the near
future because this process would require the creation of a new database and the re-sampling of all individuals in the current databases, since SNP testing methods are incompatible with STR methods; thus, the results are not comparable to those currently in the STR-based systems.31

Case Studies: Practical Application of FDP

Identitas and Parabon NanoLabs are both working with law enforcement agencies analyzing DNA samples from cases with limited evidence or leads. The Indiana State Police, as well as law enforcement agencies in the Netherlands, Poland, and Australia, are participating in pilot studies related to their cold cases.3 Canadian police services in Ontario are also utilizing FDP technology to aid in the investigation of their active cases.30

FDP technology has been credited with the successful identification of Delroy Easton Grant, deemed the Minstead Rapist and Night Stalker, who was eventually convicted of numerous robberies and sexual assaults of victims in London, England, from 1992 to 2009.32 Ancestral prediction assisted in the identification of Derrick Todd Lee, better known as the Baton Rouge Serial Killer, who claimed the lives of seven women before his arrest in 2003, and Eric Copple for the brutal stabbing murders of Leslie Mazzara and Adriane Insogna in Napa, California in 2004.20 Bio-geographical ancestry testing of a DNA sample left on Meghan Landowski’s body helped narrow the investigation, which led to the identification, arrest, and conviction of Robert Barnes for her rape and murder in 2008 in Portsmouth, Virginia.33 Familial DNA and Y-STR testing along with a SNaPshot analysis assisted in identifying José Alvarez, Jr. as the person who murdered Troy and LaDonna French on February 4, 2012, in Reidsville, North Carolina.34

Several police departments in the United States- California, Colorado, Florida, Louisiana, Maryland, Massachusetts, North Carolina, Ohio, Rhode Island, South Carolina, Texas, Utah, Virginia, and Washington- have employed SNaPshot services.34 SNaPshot technology has been used to predict facial reconstructions of suspects in numerous unsolved murder cases, including the Bennett family in Aurora, Colorado in 1984;35 the skull of an unidentified male victim found in a trash can in Glen Burnie, Maryland in 1985;16 Lisa Ziegert in Springfield, Massachusetts in 1992;34 Sierra Bouzigard in Calcasieu Parish, Louisiana in 2009;16 and, Candra Alston and Malaysia Boykin in Columbia, South Carolina in 2011.14 The Toronto Police Service submitted over 30 DNA samples from cold cases dating from the early 1980s through 2014 to Identitas for analysis.14 Although no arrests or convictions have been made based on these SNaPshot and Identitas analyses, the profiles have resulted in fresh leads and renewed interest in the cases.

In 2012, nineteen-year-old Faith Hedgepeth was a sophomore at the University of North Carolina at Chapel Hill when she was bludgeoned to death with an empty rum bottle by an unidentified perpetrator. Her roommate reported that Hedgepeth was sleeping at their apartment when she left around 4:30am on September 7. The roommate returned at 11am and discovered Hedgepeth’s partially nude and bloody body. Investigators believe Hedgepeth was raped by her killer and a semen sample from an unknown male was found on her body. A DNA sample matching the semen sample was also collected from an ink pen used to write the phrase, ‘I’m not stupid, bitch, jealous’ on a discarded fast food bag located near Hedgepeth’s body. The DNA profile failed to match anyone in the current DNA databases and the investigation stalled due to the lack of evidence. After years of failing to solve the case and lacking any promising leads, investigators reached out to Parabon NanoLabs to have a SNaPshot analysis performed on the suspect’s DNA, which rendered a computer-generated photograph predicting the individual’s
likely appearance. The predictive composite photograph was released to the public and shown to
Hedgepeth’s family and friends but has yet to produce any additional leads in the case.
However, it has helped direct the investigators toward suspects of mixed ancestry, particularly
Native American and European or Latino, with brown-colored hair and eyes, which were
determined through the FDP analysis of the suspect’s DNA.36

**Future FDP Technological Developments**

In an attempt to speed up the developmental process and fully capitalize on the benefits
of FDP, many entities are investing in FDP technology and research. In 2015, the U.S.
Department of Defense granted Parabon NanoLabs $2 million and, a year later, awarded the lab
a two-year contract to further develop FDP analysis software, referred to as Keystone.30,37
“Keystone will be the first comprehensive forensics software platform able to analyze the gamut
of genomic datasets … and will integrate [next-generation DNA analytical] tools under a
common software infrastructure … The goal of this effort is to provide [the Department of
Defense] DNA laboratories with the latest forensic DNA analysis tools under a single, easy-to-
use platform.”37 Parabon NanoLabs also plans to offer the Keystone technology as a commercial
product to DNA laboratories around the world.37

Researchers are continually exploring the entire human genome, some with the specific
purpose of indentifying genetic interrelationships relevant to forensic applications. The ultimate
goal for FDP is to sequence every DNA molecule in a sample to create the most accurate and
unique profile of the suspect possible.22 The biggest challenge in achieving this goal is locating
relevant predictive DNA markers.11 To overcome this challenge, additional technologies must
be developed to permit parallel genotyping of large groups of SNPs, which may be possible with
NGS technologies.23,31 NGS essentially allows for the rapid analysis and sequencing of a large
volume of DNA markers simultaneously, significantly reducing the time and resources necessary
to conduct the analyses.

In addition, researchers are actively working to identify and strengthen the testing
reliability of additional physical characteristics, namely hair morphology, age determination,
facial structure, and adult height.12,38 They are also refining the interpretation of mixed samples
and tracking the effects the environment has on the human genome, in an effort to make
predictions about epigenetic changes.22,30 Other EVCs of interest include handedness, chin and
cheek dimpling, and earlobe attachment.20 However, limited resources have been invested in the
study of these features so far.

Hair morphology embodies the hair structure (i.e., thickness and texture) and the
potential for hair loss and baldness.10 However, the ease of changing the overall appearance of
one’s hair may negate the potential benefit of knowing the specific hair morphology and its
natural color from a suspect DNA sample.

Numerous approaches have been considered when attempting to predict the age of an
individual based on a DNA sample. The most promising approach is the study of cytosine
methylation, which focuses on the changes in gene expression throughout a lifespan.38,39
Methylation is the process of the genome activating and deactivating select genes over time.38
The age prediction error rate using methylation analysis appears to increase with age, likely as a
result of environmental effects on the genome; “the longer people live, the longer their
epigenome is influenced by the environment.”38 To date, there appears to be no difference in the
methylation process due to gender.39 Studies have focused mainly on blood and teeth samples,
highlighting the need for additional research utilizing other specimens such as saliva, semen, and bone. The potential to determine the biological age of an individual via a DNA sample could benefit not only the creation of a criminal profile, but also the examination of human remains from mass disasters and historical burial sites.

Researchers have also expressed interest in determining the chronological age of a DNA sample. The process involves the estimation of a sample’s age and deposition time, rather than the biological age of the individual from whom the sample originated. Researchers are studying the degradation rate of biological materials and evaluating the environmental effects on a deposited sample, resulting in speculation about the relationship between melatonin and cortisol concentration levels and chronological age. But, again, further research is necessary and essential to allow analysts to form a reliable analysis of this specific characteristic.

Facial features including the size and shape of the overall face and some individual features, such as the eyes, nose, and lips are believed to have a strong genetic component based on the striking resemblance between monozygotic twins; however, the complexity of the specific genes involved is not yet understood. The majority of current facial morphology prediction is based on sex and genetic ancestry. Genetic-based skin color analysis is still incomplete because studies have only covered a proportion of the total possible color variations. The complexity of facial composition is challenging because the features are considered modular, meaning the genes that affect one feature are more than likely also involved in the expression of other features.

Researchers are aggressively studying adult body height to enhance FDP profiles but the prediction accuracy rate for this trait is very low. Although familial studies have indicated that nearly 80 percent of adult body height is hereditary, this characteristic is greatly influenced by environmental factors, such as nutrition, and varies significantly within and between populations.

Recent studies regarding some genetic disease traits, including cleft lip, albinism, and dwarfism, may also help unravel the mystery behind normal appearance variation because of genetic influence. The resulting physical attributes of these diseases may be helpful in the development of a more unique and individualizing FDP profile image from a forensic DNA sample than is currently possible. In addition, rare genetic diseases that require regular treatment and medication, such as sickle-cell anemia, may allow investigators to locate a suspect via searches of local medical records.

Legislation

Legislation is slowly adapting to advances in scientific technology. Most countries have some form of general forensic DNA legislation but few have laws specifically regulating DNA phenotyping. In fact, the Netherlands is the only country currently regulating FDP, and it does so by restricting testing to visible EVCs present at birth. The Dutch Act on Determining Externally Perceptible Personal Characteristics from Cell Material (the “Act”) defines a “DNA investigation [as] the research of cell material which is only targeted at comparing DNA profiles or determining externally perceptible personal characteristics of the unknown suspect.” The Act further states that “if it is uncertain that the source knows about the trait, it may not be investigated … [and] any physical characteristic must be (1) externally perceptible; (2) visible; (3) present at the time of and since birth; and (4) publicly perceptible.” The Dutch Code of Criminal Procedure also allows public prosecutors and investigating judges to order testing, “but
only for offenses punishable by a maximum imprisonment of four years or more.”

The Netherlands approach to FDP has focused on the belief that solving a crime may trump individual rights.

Some countries, including Belgium and Germany, expressly forbid the use of FDP testing beyond determining sex. However, other forms of DNA analysis are permitted for specific purposes. Germany permits DNA testing only for determining parentage, gender, and whether DNA came from the suspect or victim. Canada, Belgium, Spain, and South Africa are restricted to typing non-coding regions of DNA, while Australia only allows for DNA genotyping.

Common law systems, like those in the United Kingdom and the United States, allow DNA testing “unless specifically prohibited by law.” The United Kingdom expressly permits DNA testing to provide “intelligence about the physical appearance of the offender” but limits phenotyping to ethnic inference and red-hair tests. The United States federal statute has no explicit restriction on phenotyping but limits DNA collection and testing of known samples for identification purposes only. However, “identification purposes” is not specifically defined and the statute only applies to samples in a DNA database; there are currently no U.S. federal laws to regulate the analysis of abandoned or crime scene samples.

The U.S. is likely lacking in regulation of FDP because its law enforcement agencies tend to lag in adopting new technologies as opposed to other countries operating at the forefront of technology. However, some states have instituted laws to regulate DNA processing in their specific jurisdictions. Texas is the only state to expressly authorize FDP testing, including physical characteristics and genetic diseases. Florida, Michigan, South Dakota, Vermont, Utah, and Washington all prohibit DNA testing for medical conditions or genetic disorders, but none of those states restrict FDP for EVCs. A Vermont statute is representative of the legislation in those other five states, and indicates that “[the] analysis of DNA samples is authorized … to type the genetic markers from DNA samples for law enforcement identification purposes … [but] analysis of DNA samples obtained pursuant to [the statute] is not authorized for identification of any medical or genetic disorder.”

Other states prohibit the use of DNA for determining physical traits (other than sex) or disease predispositions, including Indiana, New Mexico, Rhode Island, and Wyoming. Indiana’s statute specifically states that “[the] information contained in the Indiana DNA database may not be collected or stored to obtain information about human physical traits or predisposition for disease,” but it does not preclude the investigation of new samples “not yet in nor intended to be” put into the database. Rhode Island’s legislation states that “DNA samples and DNA records collected under [the statute] shall be used only for law enforcement identification purposes or to assist in the recovery of identification of human remains from disasters or for other humanitarian identification purposes, including identification of missing persons … DNA samples and DNA records … shall never be used under the provisions of this chapter for the purpose of obtaining information about physical characteristics, traits or predispositions for disease.” Finally, Wyoming law holds that “[the] information 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,” and only permits DNA genotyping.

Suggested restrictions
FDP is intended to assist in identifying the source of the unknown DNA samples and eliminating others who do not resemble the predicted characteristics. Therefore, FDP should never be conducted on DNA samples of known individuals already in a database, as any investigative benefit gained from doing so would be inappropriate. If a match for an unknown sample from a scene is located in a DNA database, there is no need for a FDP profile to be developed, as the identity of the suspect would already be known. Similarly, FDP should not be performed on any known sample in the context of an investigation for the same reason. Submitting a known sample for FDP, beyond the scope of informed empirical research, should be prohibited to prevent improper use of DNA materials. Further limiting FDP testing of an unknown sample to cases in which sufficient certainty exists to believe the sample belongs to the perpetrator and the nature of the offense warrants aggressive investigation (i.e., violent crimes like rape and murder) and apprehension of the perpetrator may help eliminate the testing of unrelated samples, such as those from witnesses and victims.\(^5\) However, consideration should be made to allow FDP processing for EVCs of DNA samples from unidentified victims to assist in the identification of these individuals as well, as the victim’s identity is an essential aspect of every criminal investigation. Similarly, this allowance should also be extended to FDP processing of unidentified human remains from mass disasters to provide clarification of those individuals’ identities and provide closure for the impacted families and communities.

Concerns and Remedies

The “Slippery Slope” Opposition

As a result of active research and technological advancement, forensic capabilities are rapidly evolving and the scope of possibilities is widening exponentially, inadvertently resulting in apprehension on the part of critics. The main argument against FDP is referred to as the slippery slope: the fear that allowing one form of FDP will eventually lead to the use of all forms of FDP, resulting in the misuse of information derived from the analysis of the human genome.\(^11\) Such fear has fed into the opposition against FDP advancement, resulting in suggestions of incredible and improbable repercussions, such as mass eugenics.\(^17\)

The ultimate question is whether it is acceptable to completely disregard the benefits and potential of FDP technology out of fear. The simple answer is: No. Various steps should be taken to regulate and define the acceptable use of FDP technology and analytical results to ensure FDP is used only for its intended and permissible purposes. Limiting and regulating the use of FDP would permit the proper use of FDP technology and reduce the opportunity to misuse or abuse the results. In an effort to reduce misinterpretation and profile inaccuracy, testing should be limited to EVCs with high predictive reliability and value.\(^11\) Emphasizing that inferences made from genetic markers are statistical likelihoods, rather than absolute certainties, and stressing that FDP results be evaluated within the context of other investigative information would assist in clarifying the capabilities and limitations of the technology.\(^11\) Similarly, all FDP-based composite images should be clearly labeled with a specific caveat of the predictive nature of the profile.\(^20\) For instance: “Please note these probabilistic estimates of [the person’s] appearance; there may be some variation and some characteristics may have been intentionally altered by the subject. [This is] an artist’s rendering of several possibilities….\(^20\) In addition,
limiting the extent to which the genome can be analyzed for forensic purposes may help alleviate concerns of opposing entities.11

Finally, limiting the circumstances under which FDP for EVCs can be used would prohibit the indiscriminate application of FDP on a large-scale basis. Requiring an analysis of the totality of circumstances on a case-by-case basis prior to performing FDP may significantly reduce the potential for abuse, including the utilization of FDP for political advantage. As previously suggested, use of FDP for EVCs should be limited to cases involving violent criminal activities, with a proposed guideline of crimes punishable by significant incarceration such as four years or more, as is the case in the Netherlands.

Right Not to Know

The “Right Not to Know” is a basic medical law principle and a founding code for the Universal Declaration of the Human Genome and Human Rights Declaration.20 The principle essentially grants the individual the right not to know information regarding their own medical status, including information identified through genetic testing. This right may be the main force in limiting FDP testing to EVCs, since external characteristics are visible and, therefore, the individual is highly likely to be aware of them prior to testing.

Proponents of FDP have pointed out that testing for EVCs is equivalent to information derived from an eyewitness statement, with the added benefit of scientific accuracy.11 EVC predictions are far more neutral and unbiased toward any population than that of individual memory and recollection. There is a high error rate in eyewitness statements but the reliability of FDP for EVCs can be statistically supported. There does not seem to be a difference between the public reaction to eyewitness statements regarding the physical appearance of a suspect and the EVCs determined via FDP.11

However, even FDP for EVCs has the potential to inadvertently reveal unknown facts, such as hidden parentage or disease-related characteristics, which may have detrimental effects on the individual and next of kin.19 Privacy right activists have also expressed concern regarding the eventual capability to reveal relationships between genetic markers and behavioral tendencies and the potential repercussions against individuals with propensities strongly related to criminal behavior, such as aggression, addiction, or pedophilia.19 Opponents worry that such information would be made admissible in court and viewed as aggravating conditions, effecting trial verdicts and parole eligibility, or result in forced treatment and preventative detention.5 Concern has also been expressed regarding FDP testing for behavioral tendencies and the possibility of dooming an individual to a type of self-fulfilling prophecy based on their genetic predispositions; similarly, individuals may use their predisposed tendencies as an excuse for their inappropriate behavior.5 Furthermore, the publication of sensitive traits or behavioral tendencies may result in emotional, mental, or social harm to individuals later determined to be innocent.

Rather than implementing a blanket restriction on all FDP, testing should be performed on a trait-by-trait basis and publicity of the results limited to the EVCs most likely to aid in the identification of the suspect, allowing investigators to tap into FDP’s value but reducing the threat to individual rights.19 Limiting FDP testing to EVCs greatly reduces the risk of “embarrassment or shock upon disclosure,” as the individual is, in all likelihood, already aware of these characteristics.5 Furthermore, non-visible characteristics are not very useful when pursuing, identifying, and eliminating suspects because the features are not visible to the public. Similarly, most genetically-based disorders have a low penetrance rate, so even if they are
identified, an individual will rarely express the disorder, greatly decreasing the predictive value of such traits. Further, the value of behavioral tendency predispositions is limited, as it would come as no surprise that a suspected murderer has a tendency toward aggression. Therefore, restricting FDP for genetic diseases and behavioral tendencies may be appropriate, as the information derived from such testing is unlikely to assist in identifying a suspect.

**Individual Freedoms and Privacy**

Some entities believe FDP is an invasion of privacy and personal autonomy. In order to address this concern, it is essential to analyze the underlying premise of samples used for FDP testing. The Fourth Amendment of the U.S. Constitution grants citizens the right to privacy and protection against illegal searches and seizures. However, crime scene samples are legally considered “abandoned” material, which have no constitutional protection, since the sample is no longer a unified component within the individual, there is no coercion used or physical integrity violated in the process of obtaining the DNA. Upon abandonment, the DNA material transitions from private to public property and is, therefore, susceptible to testing, just like all other items of physical evidence. Criminal behavior subsequently results in a suspect surrendering many of their rights, causing the public’s interest to outweigh the suspect’s individual liberties.

**Data Protection**

Prevention of the release of excess personal information and data protection are also of great concern. Thus, limiting the amount of information released to individuals outside of the laboratory environment may be helpful. Law enforcement should only be provided with the predictive results, as there is no practical need for authorities to possess the raw DNA data. The extent of information released to the public should be even further limited to only that which is pertinent to the public’s safety and ability to assist in visually identifying a suspect. It is unnecessary to advertise the source of the profile information to the public, especially when it would be possible to derive the same information from other investigative means. If additional information is derived from FDP testing, such as genetic disease or behavioral propensities, it should remain confidential, and the donor individual should be given the choice of requesting the full profile.

Disposing of the FDP data after a suspect has been identified would also reduce the likelihood and opportunity for misuse. Considering the intended purposes of FDP for EVCs, the resulting profile information becomes irrelevant once the suspect is identified and confirmed via traditional DNA testing methods. FDP for EVCs has investigative value but is not intended for court usage, as a direct comparison of the scene DNA sample and the suspect’s known sample would be performed to confirm the match. Therefore, there is no need to retain or store the FDP profile beyond the investigative stage since it is intended to be used for tactical rather than evidentiary purposes.

**Racial Profiling**

Some critics have discouraged making EVC predictions based on DNA geographic ancestry because appearance traits are not restricted to a particular geographic region.
Although ancestry and appearance overlap, they are not the same thing, which is important to remember when reviewing a FDP profile.\textsuperscript{12} Furthermore, the extent to which genetic admixture within an individual is reflected in their appearance remains unclear and the interpretation of admixed samples is currently underdeveloped.\textsuperscript{31} In order to overcome these challenges, researchers have suggested that appearance estimations be made using genetic markers known to be influential on those specific EVCs instead of making inferences based on bio-geographic ancestry.\textsuperscript{31}

In light of the current tensions between law enforcement and minority populations, some fear that “law enforcement might misconstrue the probabilistic, context-dependent nature of bio-geographical ancestry information… to justify targeting racially identified populations” and result in investigative “tunnel vision.”\textsuperscript{19,31} This type of discrimination is referred to as racial profiling, “[which] is the misuse of patterns or the misuse of somebody’s racial information… to focus unfairly on a certain population.”\textsuperscript{3} Activists have argued FDP profiles may reinforce prejudices against specific populations and encourage stigmatization of those groups, leading to increased vulnerability of minorities and subjecting them to higher rates of undue suspicion and greater investigative intensity.\textsuperscript{17,20} Furthermore, FDP profiles of minority suspects could reinforce stereotypical beliefs that some races have a greater propensity to commit crimes and result in more discrimination.\textsuperscript{20}

The flip side of this argument is that the use of FDP at the beginning of an investigation may help “reverse inaccurate or racially motivated accusations” by identifying the suspect’s likely bio-geographic origin.\textsuperscript{3} In addition, genetic testing indiscriminately seeks traits and is, therefore, not inherently biased toward or against any particular population.\textsuperscript{5} DNA material does not change to fit certain stereotypes, or unjustly discriminate against an individual, whether or not that person is part of a minority population. Considering the unknown DNA sample was left behind at a crime scene, the analysis of the DNA is warranted and absolutely necessary for a thorough investigation. The individual who left behind the sample is the party responsible for any connotations derived from their presence at a crime scene; analysts and investigators are not responsible for the DNA sample having been placed at the scene and have no control over who commits crimes. Individuals need to be held personally accountable for their actions and accept the consequences of their decisions; the individuals who commit crimes are responsible for the resulting statistics that lead to stereotypes. Blaming scientific technology, statistical data, and the public’s resulting beliefs is misplacing the responsibility. An FDP profile has an equal chance of contradicting racial stereotypes as it does of reinforcing them. The chance of reinforcing stereotypes should not supersede the opportunity to seek the truth and potentially alleviate biases.

\textit{FDP Technological Limitations}

Although many entities sing the praises of DNA technologies, they often fail to acknowledge that DNA typing is not absolutely infallible.\textsuperscript{42} FDP predictions come with a significant level of uncertainty, which may result in the misinterpretation of the results and the overstating of the evidence.\textsuperscript{12} Leading developers of FDP technologies expressed their concerns regarding the contrast between the current capabilities of FDP technology and that portrayed by the media: “the work so far provides an ‘analytical framework’ [for FDP] but considerable further research is needed… one must pay attention to the current lack of knowledge,
inaccuracies and issues linked with DNA phenotyping, unfortunately often ignored by the media.”

A combination of infinite genetic components and environmental factors are ultimately responsible for the overall physical appearance of an individual. Because of the intertwining of biology and environmental influences, there is no guarantee predicted characteristics will come to fruition. Therefore, analysis requires some level of speculation, which opens FDP to additional criticism for the use of subjectivity in a field built on scientific objectivity. As seen with other pattern recognition-based disciplines, such as fingerprint examination and bullet striation comparison, reliance on methods requiring the interjection of subjective opinion leads to the demand for proof of method validation and testing integrity. The subjective opinion of an analyst is only as reliable as the training and skill of that particular analyst.

The most notable limitation of FDP for EVCs is the number and range of EVCs that can be predicted with a high rate of accuracy at this time. The challenge with analyzing and interpreting the genome specific to the needs of FDP is that most physical appearance markers are in non-coding areas. In addition, many markers serve more than one purpose and associate with other markers to express more than one trait. This is evident in the case of hair, eye, and skin coloring via FDP, as there is a tremendous amount of overlap in the markers that influence these features. Although FDP can differentiate between major population groups, it is not as accurate in differentiating between closely related population groups or admixed ancestries. Further studies must be performed to better understand the effects of mixed geographic samples and the influence of geographic origin on genetic-based EVCs. According to Dr. Kayser, “all DNA-predictable EVCs are [currently] group-specific, [and although] group-specific traits are expected to be useful to reduce a large group of potential suspects, it is the individual-specific appearance prediction that is sought.” Additional research is also essential to locate other SNPs influential in the expression of EVCs to increase the degree of predictability of these and other characteristics that may be beneficial in creating a more complete profile from an unknown sample.

The ability to combine multiple defining physical characteristics further strengthens the physical profile developed through FDP. For instance, the combination of highly-predictable EVCs with additional refining of facial structure and individual age prediction could result in the creation of an individual-specific profile, increasing the likelihood of identifying the source of the sample. Current FDP imaging software is limited to the production of generic profile images based on predictive likelihoods, not absolute certainties. Therefore, the results could be misleading to both law enforcement and the public, resulting in the elimination of persons of interests that fail to display the predicted features but, in reality, fall into the margin of error. This highlights the fact that FDP technology must remain only one of many tools available and utilized during an active investigation.

Finally, as with all DNA analyses, there is concern regarding the potential for sample contamination and the resulting effects of that contamination. Proper specimen collection and handling is essential to DNA processes, FDP analysis included. Ensuring proper handling and processing requires FDP analysts receive intensive training and gain experience before being permitted to perform FDP. In addition, all laboratories performing FDP should be required to undergo frequent auditing and proficiency testing to encourage ethical and objective handling standards. Laboratories and individual analysts must be held accountable for their actions when handling samples, performing their analyses, and reporting the results.
Conclusions

Imagine a day when a miniscule amount of DNA left behind by a suspect will lead to the production of a true-life photographic composite of that individual; a time when a suspect profile will no longer be generic but so specific the photo could be broadcast and the suspect rapidly identified and arrested. If suspects fear inevitable identification based on traditional DNA processing, imagine the dread when the full potential of FDP for EVCs is realized. The looming certainty of being identified may deter potential criminals or result in a perpetrator turning himself/herself in to relieve that immense pressure of impending identification and apprehension. Or consider the possibility of deriving the unique details of a suspect’s fingerprints from an abandoned DNA sample, resulting in the additional benefit of utilizing alternative databases, such as AFIS, to search for the suspect instead of being limited to a DNA search in CODIS. Although science has not caught up with these ideals yet, researchers are aggressively working towards the capabilities through the advancement of FDP technologies.

FDP for EVCs appears to hold the most promising potential for forensic applications in the near future and has already been employed by many law enforcement agencies around the world to aid in their investigations. However, FDP technology is still in its infancy and requires a tremendous amount of development before its full potential will be available. The science of FDP requires further studies and validation of testing methods to increase the current level of reliability in the process and capabilities. Areas of specific interest include the examination of environmental epigenome change, admixed sample interpretation, and the identification of additional EVCs with high rates of predictive accuracy.

Fear and misconceptions should not prevent further development of promising new technologies. However, countries intending to utilize FDP should first establish laws to regulate its use, restricting activities that may lead to misuse or abuse. Limiting FDP testing to unknown samples believed to belong to the perpetrators and unidentified victims of violent crimes, only testing EVCs with a high rate of predictive reliability, and disposing of FDP information after its investigative purposes have been fulfilled may be the essential regulations necessary to alleviate many of the concerns expressed by critics. Science is but one tool available to an investigator and should never be used exclusively nor out-of-context of other investigative information. Ultimately, FDP technology must only be used with the intent of objectively seeking the truth and pursuing justice.

Acknowledgements

I appreciate the encouragement provided during the development of this article, especially that from my Wayne State University academic mentors: Peter D. Frade, Ph.D. (Associate Dean of the College of Pharmacy & Health Science), Michael Kusluski (Criminalistics Professor), and Mark T. Evely, J.D. (Associate Department Chair of the College of Pharmacy & Health Science). In additional, I am extremely thankful for my family who made sacrifices to support the advancement of my education, namely my husband, Joshua, daughter, Emmalynn, and parents, Michael and Tina Dinnan. Above all, I owe my deepest gratitude to my Lord and Savior, Jesus Christ, who has made all of this possible.
References


13. News Reporter- Staff News Editor. (2015, Feb. 8). DNA Research; IUPUI awarded $1.1 million grant to develop tools to predict physical appearance from DNA. Vertical News, Retrieved from


