Comparison Between DNA STR Typing and Salivary Proteomic Profiling Study for Forensic Identification

By

HARRISON DOAN

THESIS

Submitted in partial satisfaction of the requirements for the degree of

MASTER OF SCIENCE

in

Forensic Science

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Cecilia Giulivi, Chair

James Angelastro

Ashley Hall

Joy Viray

Committee in Charge

2023

Abstract

Person-of-interest identification is crucial in forensic work, providing probative value for biological evidence obtained from crimes. The emergence of DNA typing as an identification tool has dominated the forensic biology field with DNA's potential for high statistical weight and ability to discern individuals [1], with current PCR amplification kits containing 20 or more STR loci. Despite the discriminatory power of DNA typing, many cases involve complex evidence due to the presence of mixtures, degraded DNA, environmental PCR inhibitors, and high temperatures affecting the resolving clarity of profiles [1-3, 29, 4, 12]. The emerging field of Forensic Proteomics offers potential due to the more robust chemical properties and abundance of proteins relative to DNA [19, 48]. This study compared the efficacy of individuals' identification between STR DNA typing and protein profiling using liquid, unstimulated female saliva samples collected from 15 subjects (Samples A1-A15) along with an additional "perpetrator"/reference sample (Sample A16). DNA short-tandem repeat (STR) typing results using real-time qPCR, (RT-qPCR) and capillary electrophoresis were compared to salivary protein profiling results acquired from tandem liquid chromatography-mass spectroscopy (LC-MS/MS) and analyzed with principal component analysis (PCA) was the bases of subject identification with testing occurring 1-3 years after initial sample collection.

Acknowledgements

Many thanks to the committee of this research, Dr. Cecilia Giulivi, Dr. James Angelastro, Dr. Ashley Hall, and Joy Viray for always being responsive in such short notice. Many thanks to Joy Viray for allowing me to use the Sacramento District Attorney's Office Biology Lab for DNA and providing the necessary reagents for DNA quantification, amplification, and typing, as well as guiding me through the process. My sincerest thank you to Dr. Cecilia Giulivi for providing the lab space at Veterinary Medicine Building 3B and the QIAGEN® DNA Blood and Mini Kit and the other reagents required for DNA extraction and being a supportive mentor. Lastly, thank you to my friends and family for all their help, with a special mention to my best friend Jonathan Ha for his undying emotional support during this research.

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1 Using Saliva in Forensic Casework

DNA has been the "Gold Standard" within the forensic sciences for circumstantial evidence assisting in criminal cases as a key identification method that includes a noninvasive method of collecting reference profiles [7, 25, 46]. Saliva evidence can be plentiful in cases ranging from buccal swabs, cigarette butts, used utensils, gum to more obscure items like half-eaten foods, vomit, and telephone receivers [1] to associate inclusion or exclusion of an individual to a crime scene [2]. In sexual assault cases, saliva on the neck, face, breast, and/or genitalia, in tandem with semen, can corroborate victim testimonies and link the offender(s) to the crime.

The identification of saliva is performed by presumptive tests. This group of tests includes amylase diffusion, RSID^{TM} saliva immunochromatographic field strips, and Phadebas test can screen for the presence of saliva. These techniques target the abundant amylases, comprising most of the salivary proteome [11, 2, 7]. The RSID^{TM} strips use two monoclonal antibodies to alpha-amylase and laminar flow to detect its presence, working on saliva samples collected from various surfaces [9-11]. The Phedabas test employs starch-bound blue dye complex, which is insoluble in water. In the presence of water and alpha-amylase, the complex, freeing the water-soluble blue dye, causes the region to become colorless, indicating the presence of the amylase [10]. Another popular test is radial diffusion, which employs blue starch-iodine gel test plate. Alpha-amylase hydrolyzes the starch from iodine, turning the area colorless [8]. The radius of the transparent circle determines the amount of amylase present. This is followed by DNA typing on saliva-positive samples, leaving saliva a crucial body fluid in forensic case work [1, 2, 7, 12].

Saliva has also been used in forensic drug testing for detecting illicit drugs such as Δ^9-tetrahydrocannobinol (THC), amphetamine (AMP), 3,4-methylenedioxymethamphetamine (MDMA), and many others using GC/MS. Gas chromatography volatilizes molecules into smaller functional groups, and mass spectroscopy uses electrons to fragment and ionize functional groups. The charged ions are then sorted based on their mass-to-charge (m/z) ratio and plotted. Saliva samples containing presumptive drugs can be confirmed with selective mass GC/MS. The sample spectra can be compared to a reference database of mass spectra of the target drugs as each compound fragments in very specific ratios [7, 12]. Thiocyanate, a metabolite product of cyanide, can also be screened for using gas spectrometry to identify cigarette smokers, if relevant to subject identification [7].

1.2 Forensic DNA STR Typing

Short Tandem Repeats (STRs) are noncoding microsatellite regions found throughout the DNA molecule. Across different loci, the number of repeats denotes the allele number. Partial repeats (incomplete nucleotide repeats) are given decimals along with the number of full repeats (Butler, 2010,). Forensic DNA profiling primarily revolves around analyzing inherited tetra-nucleotide and penta-nucleotide STRs [1, 14, 43].

DNA can be extracted from numerous types of cells including white blood cells, hair roots, spermatocytes, or the plentiful variations of epithelial cells. [1, 17]. There are multiple DNA extraction methods including Chelex and organic (phenol-chloroform), but silica bead extraction is the most common technique. This method uses sodium dodecyl sulfate (SDS) to weaken the cell membrane and isolate proteins and DNA, and chaotropic salts to disrupt hydrogen bonding networks, thus denaturing proteins [2]. Solution buffers more acidic than pH

7.5 cause DNA adhesion to the silica beads, allowing multiple washes to purify the DNA and elute under alkaline conditions or even with nuclease-free water [1, 2, 16].

After extracting the DNA from the sample, it can be quantified using qPCR. This allows for quantifying and amplifying DNA evidence samples by employing TaqMan oligonucleotide probes. The fluorescence generated by the probes is plotted on a Relative Fluorescent Unit (RFU) curve which can be used to quantify the DNA concentration at each cycle. This data, along with a standard curve and an internal positive control, is used to determine the quantity and quality of the sample, and the presence of degradation or inhibition and allows for optimal DNA concentrations to be used for subsequent analyses [2, 16]. After quantitation, a multiplex PCR kit amplifies the STR regions across multiple loci (dependent on the multiplex kit used).

The amplified STR products are resolved using capillary electrophoresis to count the number of repeats at each locus and translate the data into interpretable electropherogram peaks [2, 3, 43]. Capillary electrophoresis works by electrokinetic injection where the DNA molecules migrate towards a positive voltage through narrow injection zones within silica-surfaced capillaries. The amplified PCR products are separated by size due to the sieving properties of the complex polymer gel used as the medium, and the STRs are sized based on internal lane standards. An argon laser illuminates the amplified STR alleles and are detected using a photomultiplier tube or a Charge-Coupled Device (CCD), converting the detection to electrical signals that generate the electropherogram peaks [1]. The Promega Powerplex® Fusion 6C kit used for this study amplifies 27 loci.

Numerous artifacts can show up on the electropherogram including spikes, dye blobs, pull-up, non-template peaks, and most commonly, stutter. Spikes are random voltage anomalies during the electrophoresis process. They appear as sharp peaks that can manifest in one or

multiple color channels at the same position. Dye blobs are caused by incomplete coupling of the fluorescent dye during the primer purification step that went undetected during manufacturing, and appear as a broad, low-level peaks. Pull-up artifacts are caused by tall peaks at loci above or below the pull-up position. This results in the spectral calibration failing to compensate for the observed signal on the CCD. Any of these artifact peaks requires further resolution from the analyst. Nonregistered allele microvariants or tri-allelic individuals can also cause an Off-Ladder Allele (OLA) label [1, 2]. Examples of artifact peaks are shown in **Figure 1.**

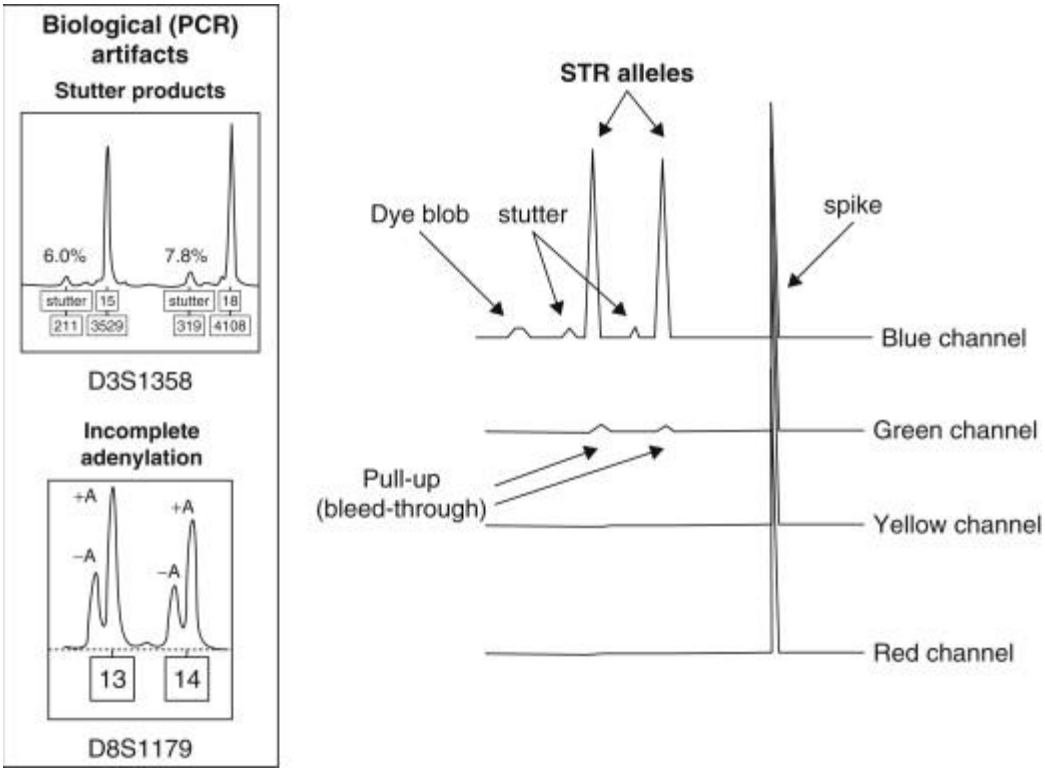

Figure 1. DNA STR Electropherogram Artifacts. Reproduced from (Butler, 2010) with permission from Elsevier Inc®.

Similar to stutter, non-template peaks can appear as partial adenylations one nucleotide before or after the true allele peak (N±1) referred to as incomplete adenylation caused by too much DNA template or thermal cycling optimizations during PCR, shown in **Figure 1**. Taq polymerase used for STR amplification usually favors the addition of an extra adenine on the 3'-end of PCR products (Butler, 2010).

Stutter is caused by primer or template strand slippage during PCR amplification resulting in either a deletion of a tetra-nucleotide repeat unit (N-4), an addition (N+4); or, for penta-nucleotides, an (N±5) from the true allele peak referred to as forward or reverse stutter, respectively [1, 3], and shown in **Figure 2**.

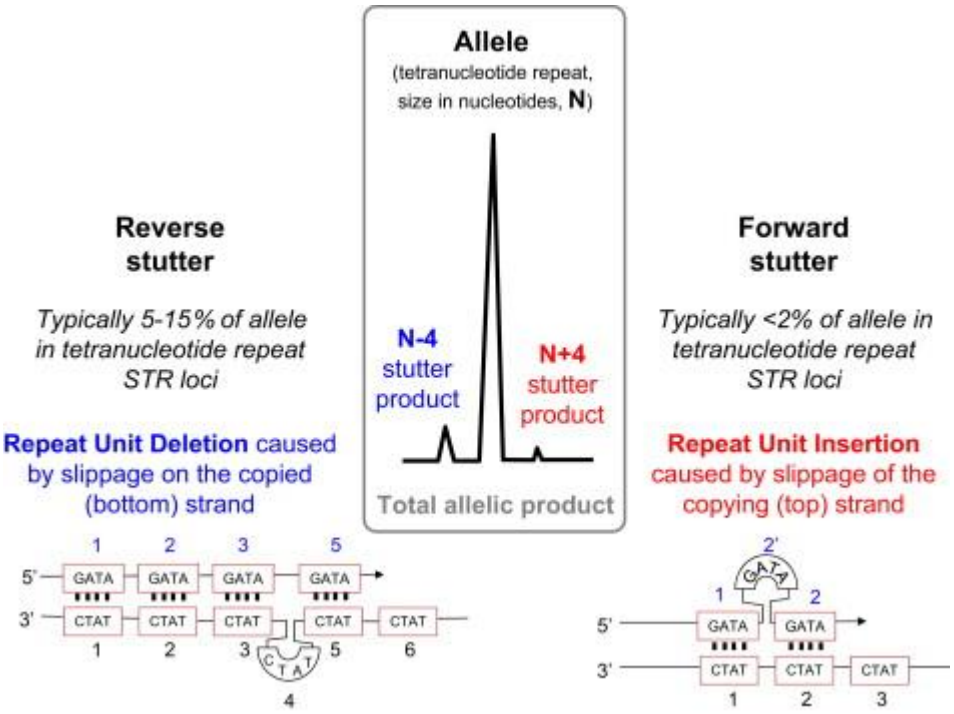

Figure 2. DNA STR Electropherogram depiction and possible mechanism of forward and reverse stutter peaks. Reproduced (Butler, 2015) with permission from Elsevier Inc®.

The NRCII endorsed statistical calculations to like Random Match Probability (RMP) or likelihood ratios to add weight to DNA evidence in court using frequency data generated from the FBI's population studies [33]. These calculations are possible because STRs are found in noncoding regions of genes and follow Hardy-Weinberg Equilibrium. Presenting the value as a 1/RMP Likelihood Ratio (LR) is interpreted as the probability of finding a random individual matching the exact same profile within a certain population [33, 43].

1.3 Limitations of DNA STR Profiling

The "Crime Scene Investigation Effect" (or CSI Effect) has glorified DNA typing to be consistent and absolute, ignorant of the many caveats that come with DNA analysis (Butler, 2015). Biological samples can be exposed to harsh environmental conditions before testing. These include high temperature, UV damage, high humidity, and bacterial nucleases that degrade template DNA within the samples [1-3, 5, 14]. The C=C double bond in pyrimidines and purines are susceptible to oxidative attack leading to ring fragmentations and base modifications, blocking replication and negatively impacting amplification with Taq polymerases during PCR [13]. This results in inaccurate sizing of STR repeats [4]. Other ions or molecules can also inhibit PCR reactions, most notably Ca^{2+} (divalent metal ions), melanin, and hemoglobin, by binding to necessary Mg^{2+} ions or collagen can bind to Taq polymerase [2, 3]. The degradation of template DNA leads to insufficient due to either primers' inability to bind to loci properly or inadequate efficacy of Taq polymerase during the amplification process. What results is a *"ski slope"* pattern in the electropherogram data [4] along with allele dropouts, affecting bigger loci the most [2, 13].

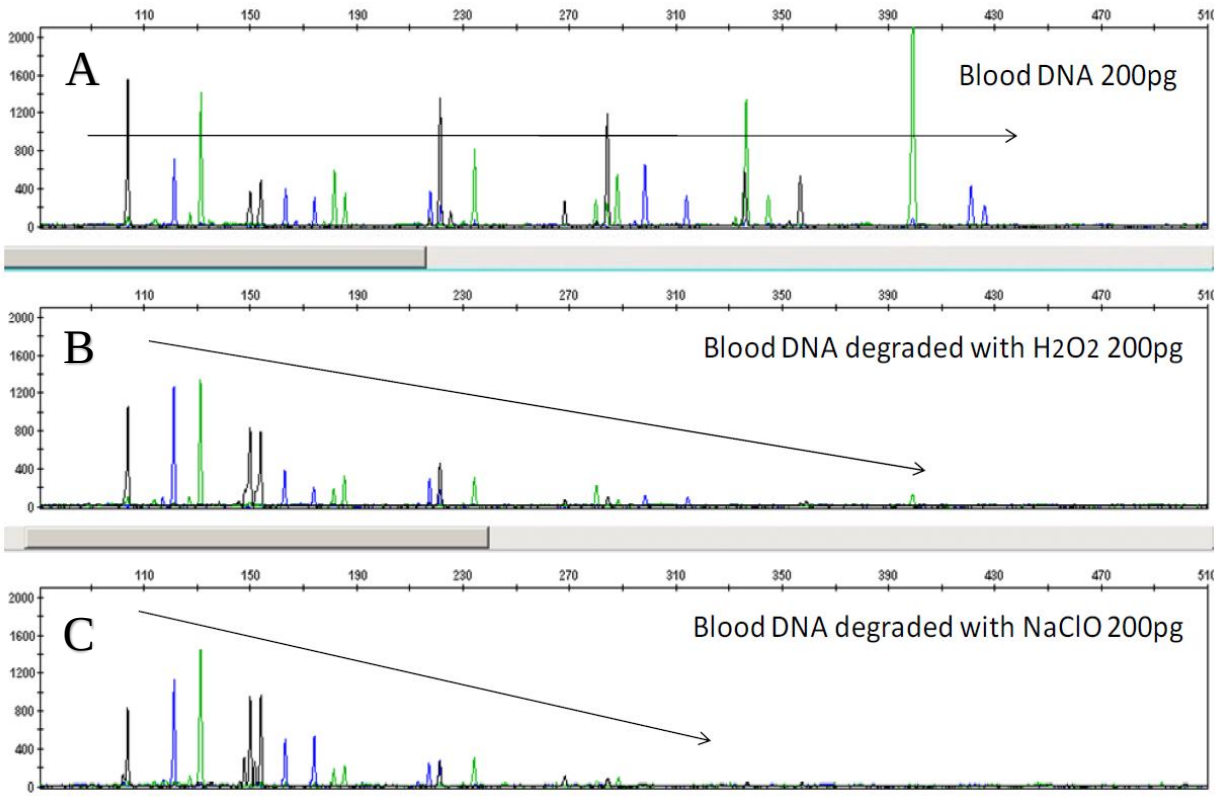

Figure 3. "Ski slope" pattern in electropherogram due to DNA template degradation. A. Electropherogram showed of 200pg of DNA extracted from a control blood sample without degradation. B. Electropherogram shows the blood sample DNA degraded with hydrogen peroxide. C. Electropherogram shows blood sample DNA degraded with bleach. The samples were typed using the Powerplex® 16 STR multiplex kit and ran on an ABI 310 using the manufacturer's suggested protocols showing the regression pattern of PCR amplified peaks in degraded DNA samples (McCord, 2011).

Advancements in the sensitivity of DNA typing kits to detect minute amounts of DNA for full STR profiles (0.5 ng) [15], with more current research obtaining the same using as little as 0.125 ng [34], have resulted in more mixture profiles in frequent casework [15]. The complex mixtures of 3 or more possible contributors further convolutes the electropherogram by introducing allele drop-outs and/or severe peak imbalances that may decrease the statistical weight or result in inconclusive outcomes altogether [2, 14,15]. Probabilistic genotyping has been employed to deconvolute complex mixtures using software programs such as STRmix™. This method uses fully continuous models based on Markov chain Monte Carlo resampling to

calculate the probability of observed mixed DNA profiles, given a genotype combination, within the interval [0,1] and displayed as a Likelihood Ratio. Still, limitations due to allele signals under the stochastic threshold or allele drop-outs can affect interpretation with false exclusions, as ultimately, the issues are caused by the quality of the sample themselves [6].

1.4 Salivary Proteins for Forensic Identification

Previous studies showed that some salivary protein and enzyme concentrations increase from youth to adult ages until a rapid decline when approaching older ages in conjunction with variable saliva flow rates and protein concentrations dependent on sex [37]. Ethnicity and geographical region for an individual can be determined using salivary protein profiling by looking at ethnicity-specific proteins [28, 29]. Some salivary proteins have been found only within Korean populations, determined by comparing the data to the integrated human salivary proteome, which hints at the implication of race determination solely through saliva evidence [28]. Individuals residing in higher altitude regions experience induced hypobaric hypoxia, causing oxidative stress that influences the upregulation or downregulation in the production of proteins found in the saliva of subjects exposed to that environment [28]. Lastly, salivary biomarkers have already been linked to diseases like dental caries, diabetes, and breast cancer [21]. Discriminating salivary protein profiling has also been used to identify conditions such as periodontitis with an increased level of albumin and hemoglobin in obese patient saliva, sex by measuring different concentrations of X-linked salivary proteins, and diet using food biomarkers [20, 21]. Furthermore, some pathologies, including leukemia, multiple myeloma, and pernicious anemia, can deposit trace amounts of symptom-specific biomarkers to the mouth, altering the saliva composition of an individual. Identification of these biomarkers in salivary evidence may

infer diseases inflicting the perpetrator or victim, adding another level of certainty when confirming identity [21].

Taken together, these studies suggest the potential forensic applications using protein abundance/profiles for subject identification. Complex DNA mixtures are challenging to resolve [1-3, 13, 14, 18], whereas many salivary proteins are highly conserved among individuals [22-24], with a few key discriminating ones, which are vital in identification [30, 31].

When sampling the salivary proteome of 11 individuals living in the Sacramento/Davis area, the protein differences were exploited using PCA to create clusters specific for each subject. Shotgun proteomics identified 973 salivary proteins, and the cohort was reduced to 94 more discriminatory proteins. Because the total proteome collected included conserved proteins, the PCA cluster was less dispersed compared to the PCA cluster of the 94 cohort [30]. Each sample had its own unique plot on the PCA hinting at the prospect of a novel identification method.

1.5 DNA vs Proteomics

The quality of the DNA extracted determines most of the quality of the results. The issue of DNA degradation or inefficient amplification may cause unreliable analyses due to the malleability of the DNA's molecular structure. The ribose and nitrogenous bases are more susceptible to depurination, oxidation, and hydrolysis [19] making peptide structures intrinsically more stable under some conditions [49].

Although DNA typing is powerful when clear contributors can be resolved [1-3, 15, 18] STR repeats used in forensic casework only allow identification because those biomarkers are within noncoding regions of the chromosomes [27]. Using proteomic profiling as an alternate or complementary identification method from salivary evidence found at crime scenes can assist

DNA typing's limitations. Saliva transfer from an individual, like bite marks, can then suggest the sex, age, geographical region, smoking habits/drug use, and diseases of the perpetrator by looking at the salivary contents [7, 20-23, 25, 28-30, 35], however this information can be an infringement on privacy, and access to these details must be discussed.

One of the appeals of STR DNA typing is the visual discrimination provided by the electropherogram peaks determined by the alleles found within each locus. Each peak allows the analyst to determine which allele was present in the same, the contributions of other individuals, if any, and the number of contributors by looking at peak height ratios at each locus [1, 2, 4, 14].

Comparing proteome profiling and DNA Short Tandem Repeat (STR) profiling is paramount in forensic science. While DNA STR profiling has long been the gold standard for human identification, proteome profiling introduces a novel dimension to forensic investigations. DNA analysis provides valuable information about genetic relatedness but cannot elucidate the impact of environmental factors or the dynamic nature of an individual's biology. Proteome profiling, on the other hand, offers insights into an individual's unique protein expression patterns, which can be influenced by age, lifestyle, and even pathological conditions. Integrating both approaches can enhance the accuracy and comprehensiveness of forensic investigations, ensuring more robust identifications and shedding light on the complex interplay between genetics and the environment. This holistic approach holds the promise of a more refined and precise understanding of human identity and behavior, ultimately contributing to the advancement of forensic science in solving crimes and ensuring justice.

2. Materials and Methods

All procedures regarding saliva collection were conducted by Thomas & Giulivi, 2021 and Smith & Giulivi, 2023. Samples were collected under informed consent policies by the

institutional review board (approved by the IRB (IRBNet ID: 1544585-1, 4/17/2020). Both sets (total of 27) were used for salivary proteomes, whereas half of Smith & Giulivi's batch (15) and the perpetrator's sample was used for DNA analysis (16 total). All methods and sample collection related to proteins were published under Thomas & Giulivi, 2021 and Smith & Giulivi, 2023. All methods regarding DNA extraction, quantification, amplification, typing, and subsequent RMP calculations were original work.

2.1 Sample Collection

Previous graduate student Ms. Hannah Smith collected 15 saliva samples for DNA or protein analysis. They were obtained from all female subjects, aged between 20 and 61 years old, and collected between 11 a.m. and 2 p.m. on February 4th, 2020. The subjects did not eat, drink, smoke, or perform oral hygiene routines for at least 15 minutes before sample collection. They rinsed their mouths with water, pooled the saliva for 60 seconds, tilted their heads back, and spat into a sterile container [31]. Using whole, unstimulated saliva decreases the amount of squamous epithelial buccal cells that may interfere with the liquid saliva proteome. The samples were labeled HS1-15. The samples were stored in a -20 °C freezer [30]. Samples HS1-15 were renamed A1-A15 for the DNA portion of this research. Sample A16 was collected on April 19th, 2023, around 8 a.m., from one of the subjects who donated saliva on February 4th, 2020 (sample A14). The other 11 samples used for proteome testing were collected as described by Thomas & Giulivi, 2021.

2.2 Saliva Sample Preparation

The whole saliva samples were subjected to a pre-modified acetone treatment to not only precipitate and concentrate the proteins, but also to inactivate any viral proteins to prevent COVID-19 or SARS-COV infection [39]. Four volumes of -20 °C analytical grade Sigma-

Aldrich acetone was added to each sample and left overnight at 4 °C. The samples were then centrifuged at 16,000 x g for 10 minutes at 4 °C. After the supernatant was discarded, the pellet was washed twice with -20 °C acetone and spun at 16,000 x g for 10 minutes at 4 °C. The new proteinaceous pellet was then dried under vacuum for 15 minutes [30].

The pellet was then solubilized in 100 μL of 6 M urea/50 mM ammonium bicarbonate, pH 8 and was subsequently treated with 2.5 μL of 5 mM dithiothreitol (DTT), and the solution was incubated for 30 minutes at 37 °C. Iodoacetamide (IAA) of 20 μL at 5 mMM was then added to the solution and incubated for 30 minutes in the dark. Afterwards, 20 μL of DTT was added incubated for 10 minutes at 20-22 °C to quench the IAA. Promega® mass spectrometry grade Lys-C/trypsin mix was added in a 1:25 ratio and incubated for 4 hours at 37 °C. To dilute the urea concentration to >1 M, 600 μL of 50 mM ammonium bicarbonate was added and incubated overnight at 37 °C. The following day, the digests were placed in a Macro Spin Column (The Nest Group, Inc.) to desalt and around 10-100 μg of the digests were prepared for subsequent mass spectrometry analysis [30].

2.3 Liquid Chromatography and Tandem Mass Spectroscopy

The protein digests were randomized and analyzed using a Thermo Scientific Q Exactive Orbirtrap mass spectrometer fitted with a Thermo Scientific Proxeon Easy-nLC II HPLC and a Proxeon nanospray source at the Proteomics Facility at the University of California, Davis. The digests were loaded on a 100 μm x 25 mm Magic C18 200 Å 5U reverse phase trap where they were desalted online before being separated using a 75 μm x 150 mm Magic C18 200 Å 3U reverse phase column. The peptides were eluted using a 90-min gradient with a 300 nL/min flow rate. An MS survey scan was obtained for the 300-1600 *m/z* range, MS/MS spectra acquired where the top 15 ions in the MS spectra were subjected to high energy collisional dissociation,

an isolation mass window of 2.0 *m/z* was used for precursor ion selection, and a 27% normalized collision energy for fragmentation with a 5-s duration for dynamic exclusion [30].

2.4 Protein Identification

Using X! Tandem (THE GPM, thegpm.org; version X! Tandem Alanine (2017.2.1.4)) database searching all MS/MS samples to search the HumanFR_crap05292020_rev database (unknown version, 149,657 entries) assuming digestion enzyme: trypsin, and searched with a fragment ion mass tolerance of 20 PPM and a parent ion tolerance of 20 PPM, carbamidomethyl of cysteine, selenocysteine, Glu->pyro-Glu of the N-terminus, ammonia-loss of the N-terminus, Gln->pyro=Glu of the N-terminus, deamidated of Asn and Gln, oxidation of Met and Trp and deoxidation of Met and Trp were all specified in X! Tandem as variable modifications. Weighted spectral counting was used to determine protein abundance [30].

The criteria for protein identification-Scaffold (version Scaffold_4.11.1, Proteome Software Inc., Portland, OR) to validate MS/MS based peptide and protein identification were accepted if they could be established at greater than 88.0% probability to achieve a less than 0.5% False Discovery Rate (FDR) according to the Scaffold Local FDR algorithm. Acceptable protein identification established at greater than 5.0% probability to achieve an FDR less than 5.0% contained at least 2 identifiable peptides, in which the probabilities were assigned using the Protein Prophet algorithm. Proteins with similar peptides that were not discernable with MS/MS were grouped together to satisfy parsimony principles [30].

2.5 DNA Extraction from Saliva using QIAGEN QIAamp® Blood Mini Kit

The DNA extraction procedure followed a modified version of the QIAGEN® Blood Mini Kit supplementary protocol for whole saliva. The QIAGEN® protease was reconstituted in 1.2 mL of Invitrogen UltraPure sterile distilled water and all pipette tips, microcentrifuge tubes,

and beakers were dry autoclaved to sterilize. Saliva sample volumes of 1- 0.1 mL were transferred into their respective sterile centrifuge tubes along with 4 mL of Gibco™ 1X Dulbecco's Phosphate-Buffered Saline (DPBS), vortexed, and centrifuged at 1800 x g for 5 minutes at 20-22 °C. The supernatants were decanted, the pellet fractions were resuspended in 180 µL of 1X DPBS, and then transferred to sterile 2 mL microcentrifuge tubes. 20 µL of the QIAGEN® Protease and 200 µL of the QIAGEN® Buffer AL were pipetted into each microcentrifuge tube. Samples were vortexed immediately for 15 seconds, and incubated at 56 °C for 1 hour, with vortexing every 20 minutes. Afterwards, 200 µL of 100% ethanol was added to the samples and vortexed. The samples were transferred, without wetting the rim, to a QIAamp® Spin Column and placed atop a 2 mL QIAGEN® collection tube and centrifuged at 6000 x g for 1 minute. The QIAamp® Spin Column was then transferred to a new 2 mL collection tube. The old collection tube discarded, and 500 µL of the QIAGEN® Buffer AW1 was pipetted into the spin column and centrifuged at 6000 x g for 1 minute. The spin column was transferred to a new 2 mL collection tube, and the old collection tube discarded. A 500 µL volume of the QIAGEN® Buffer AW2 was pipetted into the spin column, and the sample was centrifuged at full speed for 3 minutes. The spin columns were placed on each sterilized 1.5 mL microcentrifuge tube and their caps were opened to allow any residual ethanol to evaporate completely. An UltraPure™ water volume of 150 µL was pipetted into the spin column to elute the DNA, incubated at 20-22 °C for 1 minute, and centrifuged at 6000 x g for 1 minute. The eluted DNA extracts were kept at -20 °C for the subsequent quantification.

2.6 DNA Quantification using Real-Time qPCR

The DNA extracts were prepared for quantification using the reagents from the ThermoFisher Quantifiler™ Trio DNA Quantification Kit. To prepare the standards, five sterile

microcentrifuge tubes were assembled, and 10 μL of the Quantifiler™ THP DNA dilution buffer was pipetted into the first tube, and 18 μL was added to each of the rest of the tubes. To prepare the serial dilutions, 10 μL of 100 ng/μL DNA stock was added into the first tube, vortexed, and quickly spun down before pipetting 2 μL of the combined solution from tube 1 into tube 2. From tube 2, 2 μL of solution was pipetted into tube 3, etc. until tube 5, and achieved a final DNA concentration of 0.005 ng/μL. In a Thermofisher MicroAmp™ 96-well reaction plate, 2 μL of the standards were pipetted into their respective standard wells, 2 μL of the DNA dilution buffer to their respective negative control (NTC) wells, and 2 μL of each DNA extract sample was added to their respective sample wells. The 96-well plate was sealed with a MicroAmp™ Optical Adhesive Cover and centrifuged for 1 minute. The plate was then placed into the QuantStudio™ 5 Real-Time PCR System and ran following the Sacramento County District Attorney Lab protocol: 95 °C, 2 minutes, 95 °C for 9 seconds, 60 °C for 30 seconds for 40 cycles.

After receiving the analyzed DNA quant data, the DNA extract samples were pipetted into their respective PCR tubes. The samples' volumes were adjusted to contain approximately 0.5-1.0 ng of DNA within a 15 μL reaction volume, with Promega® Amplification-Grade water. To each prepared PCR tube, 5 μL of the Promega PowerPlex® Fusion 6C reaction mix and 5 μL of the PowerPlex® Fusion 6C primer set were added. The prepared samples were then placed into a Veriti™ 96-Well Thermal Cycler for PCR amplification, following the protocol from the Sacramento County District Attorney Lab: 96 °C for 1 minute, 96 °C for 5 seconds, 60 °C for 1 minute for 29 cycles, 60 °C for 10 minutes, and a 4 °C infinite hold.

2.7 DNA STR Typing using Capillary Electrophoresis

The PCR amplified samples were typed using the Applied Biosystems™ 3500xL Genetic Analyzer with 36 cm capillary arrays, POP-4 polymer, and an Anode Buffer Container

containing 1x running buffer. On a MicroAmp® Optical 96-well reaction plate, each sample well and the allelic ladder received 0.5 μL of the WEN ILS 500 and 9.5 μL of Hi-Di™ Formamide. A 1 μL volume of the amplified sample was pipetted into their respective wells, 1 μL of the Allelic Ladder was added to its respective well, and 10 μL of the Formamide and 1 μL of the WEN Lane Standard were added into each blank well. The 96-well plate was then covered with a rubber MicroAmp® septa and centrifuged for 1 minute.

After centrifugation, the 96-well plate was placed on an Applied Biosystems™ 9700 thermal cycler to denature at 95 °C for 3 minutes. Immediately after, the 96-well plate was placed on an Eppendorf® cooling block to snap-cool for 3 minutes. The 96-well plate was then placed into the Applied Biosystems™ 3500xL Genetic Analyzer and ran using the HID36_POP4XL module.

2.8 Subject Differentiation and DNA Profile Interpretation on GeneMapper™

GeneMapper ID-X v1.6 software was used to evaluate the data. The PowerPlex® Fusion 6C kit ran on an Applied Biosystems™ 3500xL utilizes a validated 100 RFU analytical threshold. Because the sample donors were anonymous and not asked consent to publish their profiles, the STR electropherogram and genotypic frequency data cannot be included in any publication.

2.9 Statistical Analysis of DNA Alleles using Random Match Probabilities

The statistical calculations were conducted following the NRC II recommendations: The Balding-Nichols equation $[p^2+p(1-p)\theta]$ was used for homozygous loci, with a θ-value of 0.01 to account for inbreeding, and the Hardy-Weinberg equation [2pq] for heterozygous loci. The allele frequencies were multiplied across all loci for every profile to calculate the Random Match Probability (RMP) and 1/RMP to convert the values into a LR on an Excel tool. This statistical

method was exclusively used because each profile was single source [3, 33]. The allele frequency data used for calculations were retrieved from the revised STR database [32] provided by the National Institute of Standards & Technology (NIST) (https://strbase-archive.nist.gov/NISTpop.htm).

2.10 Analysis of DNA genotypic frequencies using PCA

Principal component analysis (PCA), a multivariate clustering algorithm, can identify similar structures in large data sets, already employed in gene expression studies [26]. Proteomic PCA allows for another visual representation by portraying profile-specific clusters by plotting discriminatory proteins against each other, which have already been proven viable in single-source samples under a controlled setting [30].

This two-dimensional PCA was conducted using the web tool ClustVis (https://biit.cs.ut.ee/clustvis/) [40] and loading either proteomes for each subject (295 proteins/subject), the proteomes normalized each to their respective sum, or the genotypic frequencies of samples at each locus (23 loci). The individual locus allele frequencies were calculated using the respective Balding-Nichols or Hardy-Weinberg equations collected from [15], but instead of using population-specific frequencies, the entire U.S. population (cumulative) frequencies were implemented to reduce discrepancy during the PCA analysis by certifying a genotypic frequency at each locus. Committing to cumulative population frequencies slightly reduces the accuracy but simplifies the PCA plot. The first PCA plot was created using the data from all loci. The second PCA plot was generated by using only the most discriminating ones. To achieve this, the absolute value of the two highest PCA loading values was chosen, up to sample A7, to achieve an 83.332% cutoff for better cluster separation on the PCA plot.

3. Results

3.1 Analysis of Salivary Proteome PCA Plot

Proteomic PCA plot displays all 26 samples including two from the same perpetrator. Samples CG1-11 were not used for DNA typing because the samples were unavailable after a previous graduate student Mr. Hogan's thesis project. The X-axis shows the principal component 1 with a variance of 30.5% while PC2 scored a 26% with clear separation. These two samples from the perpetrator (CG02, taken on October 22nd, 2020) and (HS14, taken on February 4th, 2022). CG02 was plotted at the coordinate (x,y: -0.1453, 0.5237) while HS14 was (-0.4927, 0.5149). While the spots were close, the distance between the two points measured on the x-axis was larger than that on the y-axis, signaling that proteins within PC1 were more variable with than those belonging to the PC2.

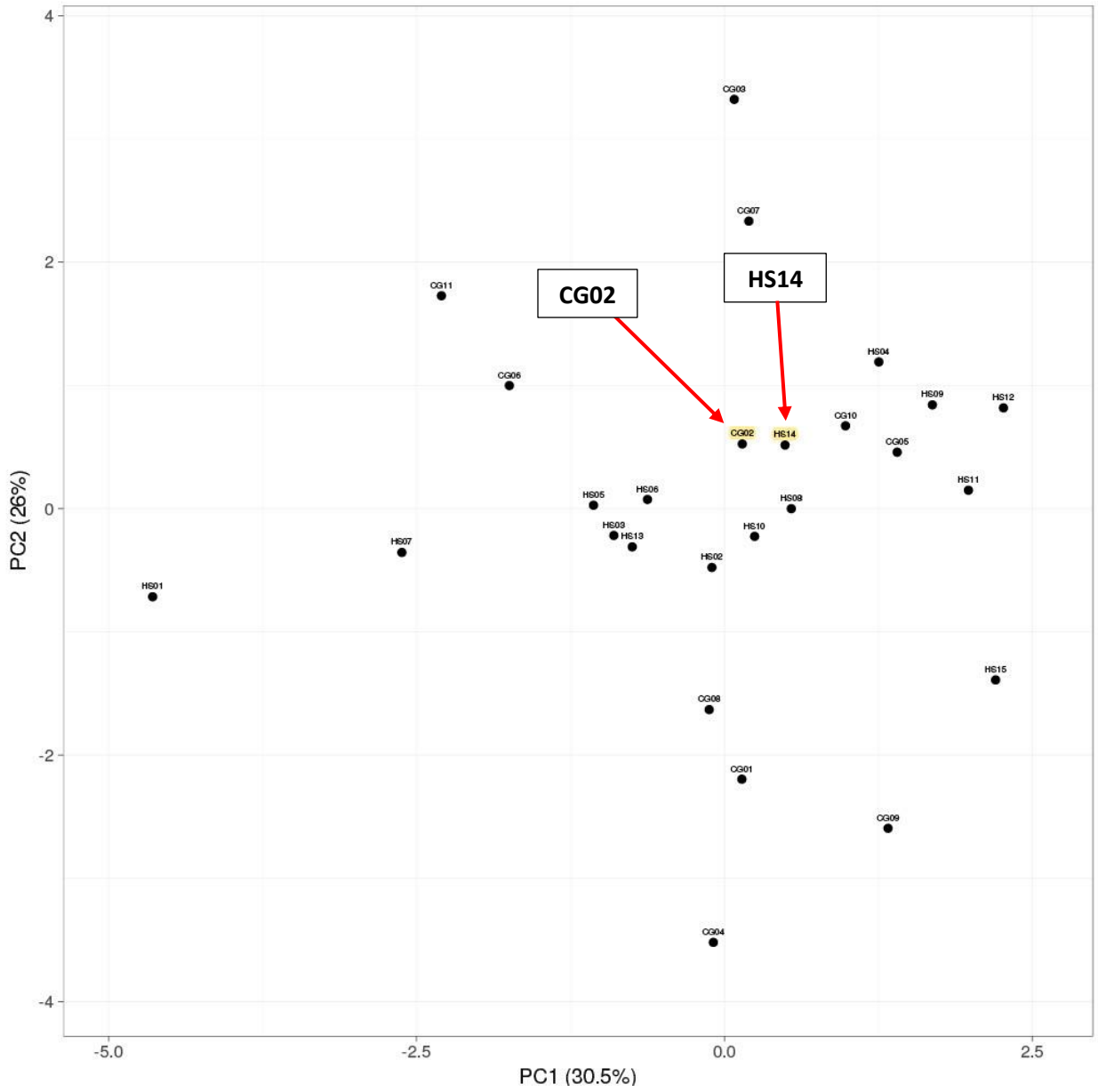

Figure 4. Data recalculated from Mr. Hogan, Ms Smith and Thomas' MS dissertations. PCA plot using salivary proteomic data from 26 samples. Samples HS1-15 correspond to A1-A15, and all CG samples were not used for DNA typing. CG samples were taken on October 22, 2020, and HS on February 4, 2022.

3.2 Analysis of DNA PCA Plot

DNA allele frequencies from [32] using the cumulative population generated a two-dimensional PCA with PC1 scoring 18% and PC2 scoring 16.5% **Figure 5**. This shows sub-optimal dispersion between clusters A1 and A2; A3 and A11; A10 and A5; and most

dramatically A7 and A16. Sample A13 had the genotypic frequencies most unique from the rest of the samples. The perpetrator's samples, A14 and A16, generated the same allele frequencies at every locus, explaining the perfect overlap having the exact coordinates: (-1.0180, 1.7283) with no difference between the axes. However, a significant overlap was noticed with A7, having a coordinate of (-1.0890, 1.8692). The NIST allele frequencies were generated using a population size of 1036 individuals [32]. The individual genotypic frequencies are relatively close to each other with a low average variance of 0.003686, which may be an explanation for the clusters.

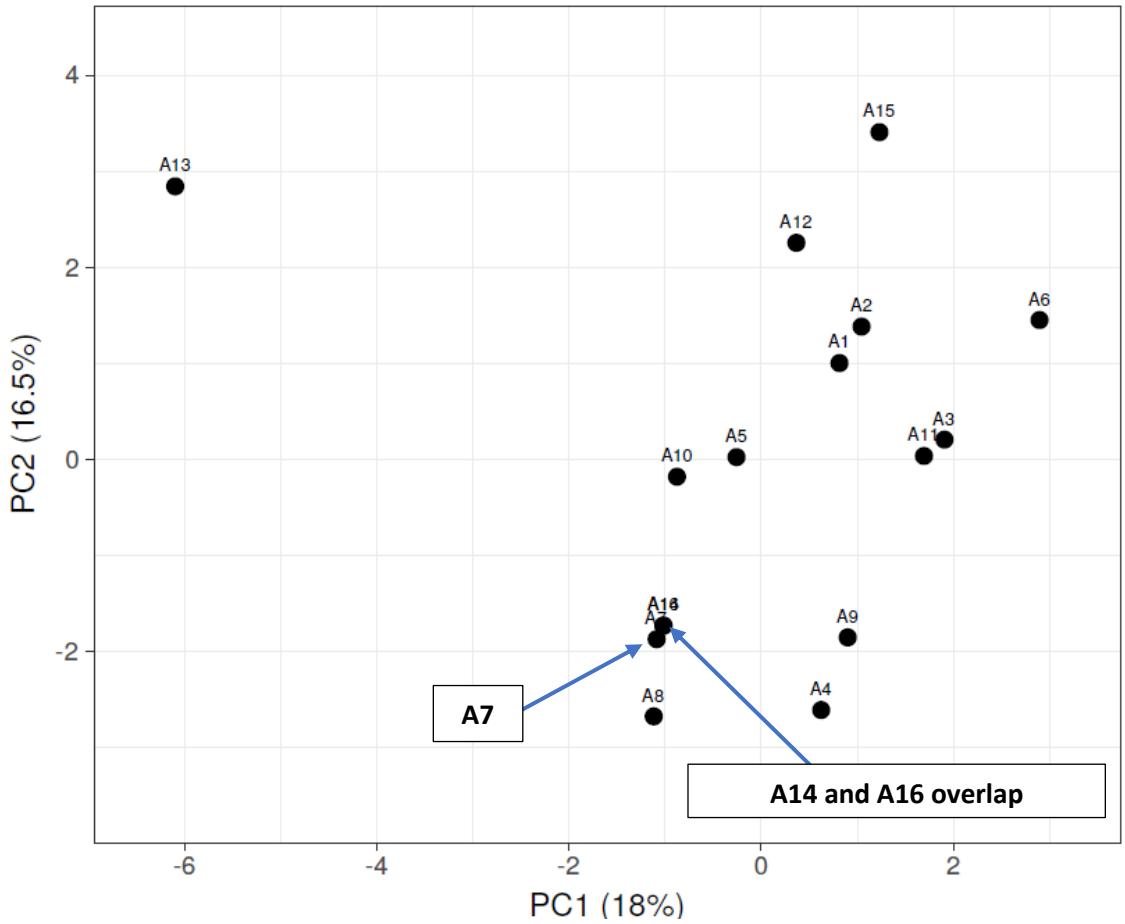

Figure 5. PCA of sample alleles plotted against the loci and the genotypic frequencies with all loci included. This shows poor separation between samples A7 and A16; A3 and A11; and A1 and A2. A4 and A16 overlap because they were contributed samples from the perpetrator.

The PCA was redone to achieve a cut-off of 83.332%. In this process, the top two components of each principal component were taken up to PC6 shown in **Table 1**, and removing

the less discriminating loci: D10S1248, D16S539, D2S1338, vWA, D5S818, SE33 (beyond PC6) from the data to improve cluster separation in **Figure 5**.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
1	0.089367	0.035282	0.000891	-0.22153	-0.36615	-0.01774	0.358324
2	-0.09704	0.282639	0.268114	0.266225	-0.20981	0.019129	-0.04689
3	0.009122	-0.16237	0.251384	0.202411	-0.24076	-0.45998	0.002323
4	-0.29568	0.271663	-0.10179	-0.03046	0.050057	-0.25659	0.124154
5	0.247363	-0.11258	0.209448	-0.4203	0.048104	-0.03613	-0.14186
6	-0.3396	-0.15186	0.01878	-0.02906	0.138516	-0.34847	-0.22086
7	-0.07482	0.047317	0.027743	-0.22874	-0.34779	0.243434	0.245503
8	-0.14089	-0.11878	0.035719	0.061096	0.086191	0.503705	0.134442
9	-0.0314	-0.38386	-0.13966	0.173777	-0.00132	-0.08433	0.02166
10	-0.24844	-0.38442	0.105243	-0.10331	-0.02871	0.142569	0.031895
11	-0.03672	0.132266	0.449335	0.080665	0.158741	-0.00712	0.047969
12	-0.21288	-0.39803	-0.07103	-0.13252	-0.05781	-0.08478	0.141347
13	-0.3737	-0.09168	0.191679	-0.25231	0.073739	-0.04431	-0.09194
14	0.299908	-0.19335	-0.06282	0.130451	0.195205	-0.13757	0.25822
15	0.163949	-0.16985	-0.05509	-0.44545	-0.02368	-0.24299	0.076327
16	-0.38182	0.053418	0.002919	0.048083	-0.18939	0.101319	0.210469
17	-0.06187	0.066752	0.016504	-0.23777	-0.49902	0.029126	-0.38646
18	0.105899	0.048943	0.520003	-0.05488	-0.04284	-0.08087	0.138463
19	0.118179	-0.16402	-0.21887	0.230242	-0.33354	-0.02526	0.251565
20	0.161349	-0.24524	0.146013	-0.03488	0.082172	0.3859	-0.28481
21	-0.05061	-0.15872	0.379095	0.037703	0.144768	0.033692	0.402231
22	-0.29824	0.158407	-0.19743	-0.2111	0.290415	0.030148	0.171708
23	0.186885	0.271845	-0.06665	-0.31381	0.167533	-0.0481	0.231337

Table 1. ClustVis Loading values for DNA genotypic frequency PCA. The highlighted values depict the top two components of each principal component. This was done up to PC7 to achieve an 83.332% cutoff.

The PC1 variance was increased from 18% to 18.3%. However, the PC2 variance remained the same at 16.5%. A14 and A16 still displayed a complete overlap having the coordinates (1.6360, 1.6848), but now the rest of the samples identified in **Figure 6** were more dispersed with one spaced-out cluster of A4, A7, A8, A9, and A15 located on the upper right corner. The average genotypic frequency variance had also improved to 0.003831.

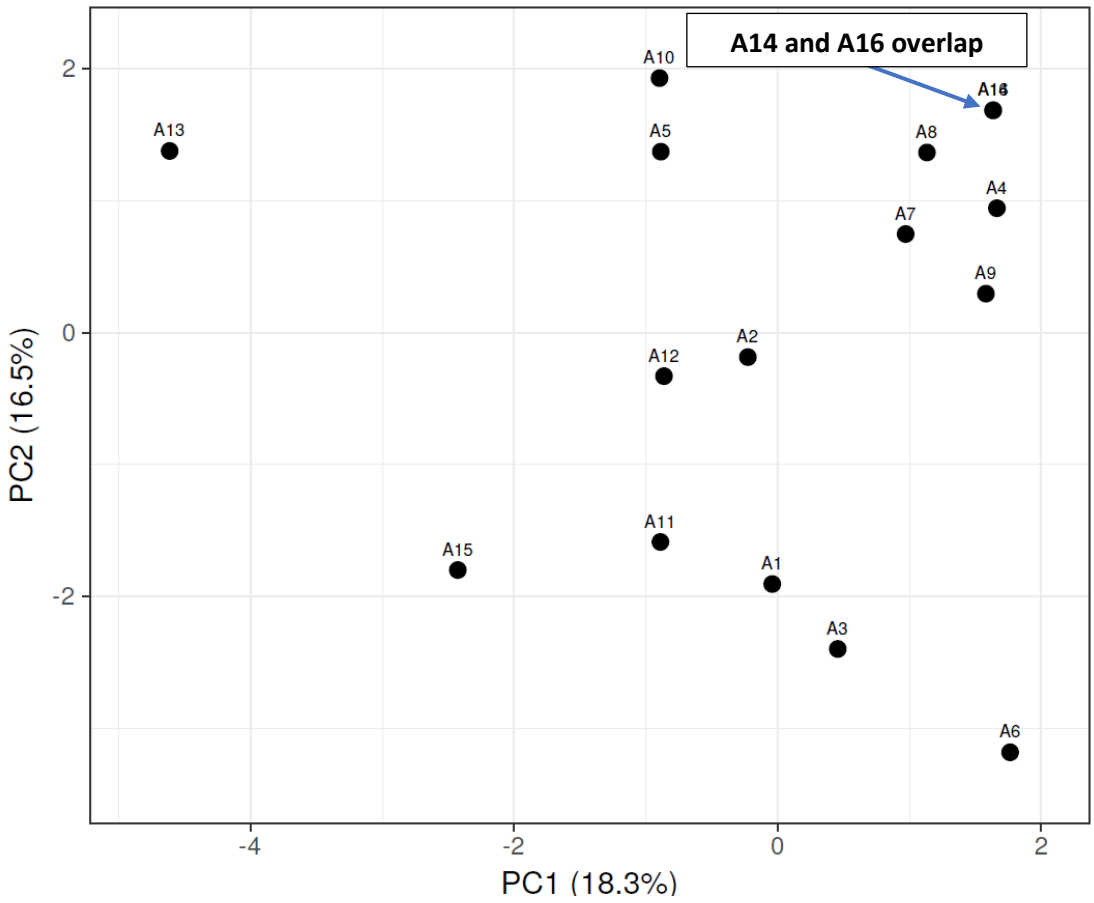

Figure 6. PCA of sample alleles plotted against loci and genotypic frequencies with the least significant loci excluded (D10S1248, D16S539, D2S1338, vWA, D5S818, SE33). Dimensions were reduced to two, with PC1 variance at 18.3% and PC2 variance at 16.5%. A4 and A16 overlap perfectly because those were contributed samples from the same perpetrator.

3.3 Analysis of STR Electropherogram Data

The starting sample volumes of saliva used in the DNA extractions range between 1mL to 0.1mL dependent on saliva availability for testing. The QIAGEN® Blood Mini Kit supplementary protocol for whole saliva DNA extraction required 1 mL, but 0.1 mL generated sufficient DNA for typing. Samples A2, A3, A5, A6, A9, and A12 utilized 1mL of saliva, yielding concentrations of 16.34 ng/µL, 18.12 ng/µL, 10.91 ng/µL, 13.31 ng/µL, 5.06 ng/µL, 4.76 ng/µL, of DNA, respectively. Samples A7 and A13 each contained 0.25 mL of saliva yielding 2.91 ng/µL, and 8.70 ng/µL of DNA, respectively. Samples A1, A4, A8, A10, A11,

A14, and A15, utilized 0.1 mL of saliva and yielded 0.32 ng/µL, 0.32 ng/µL, 0.50 ng/µL, 0.47 ng/µL, 2.74 ng/µL, and 1.39 ng/µL of DNA, respectively. Any DNA concentration between 0.5 ng to 1.0 ng within a reaction volume amplified and typed will yield a full profile.

Sample	Volume (mL)	[DNA] (ng/µL)	Sample Reaction Volume (µL)	Water Reaction Volume (µL)	Reaction [DNA] (ng)	Dilution Factor
A1	0.1	0.3183	2.00	13.00	0.64	Neat
A2	1	16.3355	1.00	14.00	0.82	1:10
A3	1	18.1208	1.00	14.00	0.91	1:20
A4	0.1	0.3205	2.00	13.00	0.64	Neat
A5	1	10.9063	1.00	14.00	1.09	1:10
A6	1	13.3147	1.00	14.00	0.67	1:20
A7	0.25	2.9078	2.50	12.50	0.73	1:10
A8	0.1	0.5033	1.50	13.50	0.75	Neat
A9	1	5.0592	1.50	13.50	0.76	1:10
A10	0.1	0.4679	1.60	13.40	0.75	Neat
A11	0.1	2.7419	4.00	12.00	0.82	1:10
A12	1	4.7557	1.50	13.50	0.76	1:10
A13	0.25	0.7289	1.00	14.00	0.91	1:20
A14	0.1	0.4422	2.00	13.00	0.88	Neat
A15	0.1	1.3887	5.40	9.60	0.75	Neat
A16	0.125	0.6823	1.00	14.00	0.68	Neat
(+) Control / DNA Standard	0.5	1.0	0.5	14.50	0.5	1:10
(-) Control / Reagent Blank	0.125	Undetermined	1.00	15.00	-	-

Table 2. PowerPlex® Fusion 6C PCR amplification sample preparation.

All saliva samples were single-source, full profiles. All profiles also showed stutter peaks that were N ± 4 nucleotide repeats from the true allele peak, which is to be expected. Sample A3 showed a pull-up peak at locus D12S391 at 102 RFUs, distinct from the >1800 RFUs true allele peak and influenced by the tall peak at locus FGA. Sample A13 displayed three pull-ups: a 133 RFU peak at locus Amelogenin caused by a tall TH01 allele, a 126 RFU peak at D3S1358 due to a tall allele peak at D18S51, and a 216 RFU at SE33 due to a tall allele at TPOX. Sample A1

showed a potential tri-allele between loci D12S391 and D19S433 displayed as an off-ladder allele. Replating the sample reproduced the same result with a tri-allele peak of 243 RFUs with an N – 4 stutter peak and GeneMapper^{TM} ID-X v1.6 flagged locus D12S391. Typing the sample on a different PCR amplification kit, i.e., Globalfiler, may be able to resolve the allele, although no database has reported a tri-allelic pattern at that locus yet.

The perpetrator's saliva sample (A14) was among the 15 and was identified accurately using DNA STR typing.

Using the genotypic frequencies provided by [32] each profile was given a LR by multiplying each genotypic frequency across all loci. Because the ethnicities of the saliva donors were never given, a LR was calculated for the four main US populations displayed in **Table 2**.

The least discriminatory LRs were calculated from A1 using Caucasian population allele frequencies with a LR probability of 1.2817 x10^{28}. Using African American population allele frequencies, A1 gave a LR probability of 2.0010 x10^{30}. Using Hispanic population allele frequencies, A1 gave a LR probability of 2.4295 x10^{28}. Using the Asian population allele frequencies, A11 gave a LR probability of 9.0905 x10^{27}. The most discriminatory profiles came from using the Caucasian population allele frequencies, A4 gave a LR probability of 4.7743 x10^{34}, using the African American population allele frequencies, A4 gave a LR probability of 6.6779 x10^{38}, Using the Hispanic population allele frequencies, A3 gave a LR probability of 9.8506 x10^{35}, and using the Asian population allele frequencies, A12 gave a LR probability of 2.9997 x10^{36}.

Samples	Caucasian	African American	Hispanic	Asian
A1	1.28171E+28	2.00096E+30	2.42954E+28	6.47071E+30
A2	1.82281E+29	1.12163E+34	3.26085E+32	8.65518E+30
A3	6.05595E+33	1.11861E+36	9.85056E+34	2.43456E+31
A4	4.77428E+34	6.67785E+38	5.60075E+32	4.98622E+34
A5	6.81143E+30	1.16794E+36	2.52866E+30	4.45224E+30
A6	6.13122E+30	9.87717E+32	2.83883E+31	2.21972E+32
A7	2.77243E+30	1.42379E+33	1.81458E+30	5.11083E+30
A8	2.89114E+34	2.59315E+35	7.81551E+31	6.90424E+35
A9	2.34357E+32	2.04731E+35	1.7132E+33	3.75393E+31
A10	3.43593E+34	7.54163E+37	3.17406E+33	3.3218E+32
A11	3.22714E+34	2.94982E+35	6.85039E+30	9.09049E+27
A12	1.61736E+31	5.69583E+32	3.77128E+31	2.99974E+36
A13	2.93083E+28	1.07749E+37	2.08994E+31	1.18445E+31
A14	8.23354E+29	1.86102E+34	1.18419E+30	1.46761E+31
A15	9.83687E+30	2.68229E+34	7.343E+29	5.57828E+33
A16	8.23354E+29	1.86102E+34	1.18419E+30	1.46761E+31

Table 3. Likelihood Ratio (1/RMP) from calculated allele frequencies of 4 different populations for all 15 samples. Because the ethnicity of the subjects is unknown, LRs were calculated for the four different U.S. populations.

A second set of RMP calculations were done with one US population instead of the four ethnic subpopulations also provided by Hill et al. 2017. The least discriminating profile came from sample A1 with a LR of 3.2971 $x10^{28}$ and the most discriminating profile came from sample A4 with a LR probability of 4.7898 $x10^{36}$. These values were used to construct a PCA plot to compare with the salivary proteome clusters. This simplifies each profile by allowing for one PCA plot per profile without ethnic concerns but reduces the accuracy of each sample's LR, causing the cumulative population LR to fall in between its highest and lowest population-specific allele frequencies.

Samples	LR
A1	3.2971E+28
A2	1.9724E+32
A3	2.0539E+34
A4	4.7898E+36
A5	3.13986E+32
A6	4.5184E+31
A7	3.5618E+30
A8	1.1484E+34
A9	5.74E+32
A10	6.02E+34
A11	1.0703E+34
A12	1.33478E+35
A13	1.22388E+29
A14	6.65986E+30
A15	1.97824E+33
A16	6.65986E+30

Table 4. Likelihood Ratio (1/RMP) from calculated allele frequencies using a single incorporated US population for all 15 samples. Calculations done with values from a single population produced one LR for each sample.

4. Discussion

Identification of the perpetrator was accurate when using the salivary proteome and the DNA profile. Regarding the salivary proteomic profiling, the salivary proteome PCA (**Figure 2)** displayed those samples as the closest points on the graph with a difference of -0.34732 on the PC1 axis and -0.0088 on the PC2 axis. Comparing these values to another close cluster, HS03 at coordinates (0.8980, -0.2182) and HS13 at coordinates (0.74858, -0.31095), showed a difference of -0.14944 on PC1 and -0.09275 on PC2; while HS05, (1.06337, 0.02696) and HS06, (0.62480, 0.07352) showed a difference of 0.43856 on PC1 and -0.4656 on PC2. Regarding the DNA genotypic frequency PCA plots, the perpetrator's samples overlapped perfectly on both **Figure 3** and **Figure 4** as her samples, A14 and A16 were both located at on the plot (-1.01795, 1.72830) with all loci included and (-1.63605, 1.68480) with the least significant loci excluded, both achieving a difference of 0 on both axes. The perpetrator's DNA profile matched at all alleles,

allowing the same frequencies to be used during the calculations resulting in the same PCA scores. Sample A7 was still a close match on the PCA plot with all loci at (-1.08897, 1.869153) compared to the A14/A16 at (-1.01795, 1.728301).

When comparing cluster separation between the protein and DNA PCA plots, **Figure 2** showed more dispersion with a PC1 variance of 30.5% and a PC2 variance of 26%. **Figure 3** showed a 12% difference at PC1 and a 19.5% at PC2. **Figure 4** showed a 12.2% difference at PC1 and a 9.5% difference at PC2, being the PC2 in **Figure 3**. Therefore, the proteome profile offers more variability than the genotypic frequencies. Even after the adjustment shown in **Figure 4**, the separation was still inadequate compared to the protein data.

Although both methods were successful in identifying the perpetrator's sample, DNA STR profiling is still the most reliable technique, given it has approximately 40 years of development, optimization, and validation (Butler, 2010). DNA typing revealed an exact allele match between samples A14/HS14 and A16, taken about one year and 2 months apart, both with the same LR listed in **Table 3** and **Table 4**. The proteome PCA plot showed samples CG02 and HS14, taken about 1 year and 3 months apart and as close matches relative to their PCA coordinates, determined through process of elimination within a controlled setting. The variation cannot be explained by the batch effect [42, 43] – technical variation or non-biological differences between measurements of different sample groups, as the samples were corrected by that using the Combat algorithm. It is more likely that the variability in salivary composition with age [36-39] generated the close match.

It is important to note that this comparison was done between STR typing with full profiles and a pilot study done with the salivary proteomic profiling. The individual proteomic points still contained more discrimination than individual genotypic frequencies and. However,

many DNA evidence samples contain low-template DNA. Further research may determine if salivary proteomic profiling exceeds DNA STR identification in samples that mimic typical forensic conditions.

Comparing identification accuracy between degraded DNA samples with salivary proteomic profiling is the next step in research. Future salivary proteomic studies can also include analyzing the PCA coordinate changes with respect to sample age as well as interpreting sample mixtures to see if the subject could be identified. This research has already determined PCA using proteomes to be more discriminating individually than single genotypic frequencies for an individual, and with a starch treatment, discriminatory proteins can be concentrated [31].

To be viable in court, forensic proteomics will also require statistical weight to give credibility to the results. By using LC/MS to analyze the specific amino acid structures of the salivary proteins, the peptide sequence can be genomically sequenced to search for genetically variable peptides [5, 44, 45, 47-49]. Along with frequency data, RMPs or likelihood ratios can be calculated using frequency data provided by the 1000 Genomes Project [50] to place proteomic data on an equivalent statistical level as DNA STR typing [44, 49].

To add statistical weight equivalent to the RMPs that STRs allow, the more discriminatory proteins can be sequenced and reverse transcribed to ascertain subject-specific nonsynonymous single-nucleotide polymorphisms (nsSNPs) to calculate LRs by using data from future genetic variant genotype frequencies, similar to the NIST database used for STR RMPs [49].

Appendix A: ClustVis DNA PCA Loadings Data for All Loci

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16
1	0.089367	0.035282	0.000891	-0.22153	-0.36615	-0.01774	0.358324	-0.29662	0.466928	-0.0639	-0.11429	0.244329	-0.15663	-0.19824	-0.29729	-0.00381
2	-0.09704	0.282639	0.268114	0.266225	-0.20981	0.019129	-0.04689	-0.18612	-0.30407	-0.20317	0.145499	-0.04016	-0.00446	0.165198	0.241912	-0.21002
3	0.009122	-0.16237	0.251384	0.202411	-0.24076	-0.45998	0.002323	-0.16726	-0.10959	-0.13659	0.067482	-0.1965	-0.21638	-0.20498	-0.09252	0.081493
4	-0.29568	0.271663	-0.10179	-0.03046	0.050057	-0.25659	0.124154	-0.06414	-0.27629	-0.26023	-0.15886	0.324351	0.025763	-0.08013	0.004507	0.464197
5	0.247363	-0.11258	0.209448	-0.4203	0.048104	-0.03613	-0.14186	0.101994	0.032893	-0.16685	0.142887	0.217899	-0.36308	0.009549	0.529133	-0.06312
6	-0.3396	-0.15186	0.01878	-0.02906	0.138516	-0.34847	-0.22086	-0.09943	0.014028	0.23253	0.243427	0.140782	-0.03976	0.046772	-0.09918	0.083622
7	-0.07482	0.047317	0.027743	-0.22874	-0.34779	0.243434	0.245503	0.484826	-0.3355	0.083404	-0.0924	-0.04303	0.006652	-0.14756	0.057991	0.199793
8	-0.14089	-0.11878	0.035719	0.061096	0.086191	0.503705	0.134442	-0.52499	-0.10551	0.074113	-0.17491	0.14082	0.038349	0.277266	0.124037	0.156861
9	-0.0314	-0.38386	-0.13966	0.173777	-0.00132	-0.08433	0.02166	0.236202	0.334595	-0.32311	-0.11867	0.189061	0.421223	0.012103	0.302418	0.111163
10	-0.24844	-0.38442	0.105243	-0.10331	-0.02871	0.142569	0.031895	0.150146	-0.22731	0.085422	0.01105	0.003201	-0.156	-0.1425	-0.26286	0.018428
11	-0.03672	0.132266	0.449335	0.080665	0.158741	-0.00712	0.047969	0.37208	0.097716	-0.00803	-0.21027	0.295231	-0.18187	0.404524	-0.27745	-0.06978
12	-0.21288	-0.39803	-0.07103	-0.13252	-0.05781	-0.08478	0.141347	0.022206	-0.05476	-0.24531	0.232446	-0.00942	0.000529	0.509691	-0.1689	-0.01338
13	-0.3737	-0.09168	0.191679	-0.25231	0.073739	-0.04431	-0.09194	-0.06471	-0.10936	0.032312	-0.22841	0.054441	0.391775	-0.33637	0.042486	-0.45884
14	0.299908	-0.19335	-0.06282	0.130451	0.195205	-0.13757	0.25822	-0.05721	-0.30639	-0.23294	-0.3828	0.078552	-0.09541	-0.0176	-0.12224	-0.42473
15	0.163949	-0.16985	-0.05509	-0.44545	-0.02368	-0.24299	0.076327	-0.21537	-0.20302	0.112058	-0.25321	-0.27049	0.013843	0.231793	0.154481	0.170995
16	-0.38182	0.053418	0.002919	0.048083	-0.18939	0.101319	0.210469	0.049326	0.195913	-0.40034	-0.00011	-0.40155	-0.15568	0.032739	0.158282	-0.12895
17	-0.06187	0.066752	0.016504	-0.23777	-0.49902	0.029126	-0.38646	-0.07601	0.068138	0.037526	-0.07512	0.125175	0.118979	0.244412	-0.10998	-0.23239
18	0.105899	0.048943	0.520003	-0.05488	-0.04284	-0.08087	0.138463	-0.11311	0.002259	-0.0754	0.079169	0.209642	0.328437	-0.0587	0.083962	0.173032
19	0.118179	-0.16402	-0.21887	0.230242	-0.33354	-0.02526	0.251565	0.018513	-0.23793	0.227348	0.356984	0.406324	0.046281	-0.00813	0.125723	-0.19813
20	0.161349	-0.24524	0.146013	-0.03488	0.082172	0.3859	-0.28481	-0.12056	-0.1169	-0.39697	0.243012	0.04261	-0.03459	-0.27653	-0.25394	0.150767
21	-0.05061	-0.15872	0.379095	0.037703	0.144768	0.033692	0.402231	-0.0302	0.171742	0.334051	0.184987	-0.20417	0.054738	-0.04084	0.128357	-0.00674
22	-0.29824	0.158407	-0.19743	-0.2111	0.290415	0.030148	0.171708	-0.07529	0.046202	-0.06397	0.195091	0.230658	-0.29689	-0.13156	0.126242	-0.2162
23	0.186885	0.271845	-0.06665	-0.31381	0.167533	-0.0481	0.231337	0.027366	-0.0897	-0.20442	0.413788	-0.13387	0.393282	0.114988	-0.26902	-0.10488

Appendix A: ClustVis Individual and Cumulative DNA PCA Values for All Loci

	Individual	Cumulative
PC1	0.176065	0.176065
PC2	0.157205	0.33327
PC3	0.136085	0.469355
PC4	0.104117	0.573473
PC5	0.100579	0.674051
PC6	0.081844	0.755895
PC7	0.068347	0.824242
PC8	0.046935	0.871177
PC9	0.038171	0.909348
PC10	0.030532	0.93988
PC11	0.022186	0.962067
PC12	0.015045	0.977111
PC13	0.012808	0.989919
PC14	0.00678	0.996699
PC15	0.003301	1
PC16	0	1
PC17	0	1

Appendix A: ClustVis DNA PCA Scores for All Loci

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16
A1	0.814629	-1.00591	0.912269	-0.50429	-1.1185	0.437227	4.073455	-0.55775	-0.72587	0.584957	-0.02446	0.521986	0.304411	0.035345	-2.78E-16	1.80E-16
A2	1.044324	-1.38599	-0.16288	1.592914	-0.952	-0.8071	-1.27177	0.991957	0.248571	0.335836	0.98866	0.331834	1.213945	0.253201	-5.00E-16	-6.94E-17
A3	1.907825	-0.2086	0.346102	-4.03374	0.983017	1.388236	-0.83639	1.165971	-0.92958	0.294711	0.278271	-0.43477	0.225873	-0.39936	-3.89E-16	-6.11E-16
A4	0.624009	2.607247	-1.68567	1.001779	1.389323	-1.61659	0.147929	-0.17374	-0.6987	-0.03774	-1.63691	-0.31848	0.654622	-0.23059	1.11E-16	9.02E-16
A5	-0.25743	-0.0267	-1.41814	1.210311	3.344488	1.788139	-0.14173	-0.85222	-0.31633	-0.76697	0.668139	1.103099	-0.074	-0.11084	8.88E-16	-1.60E-16
A6	2.897433	-1.45394	2.833265	-0.95119	-0.02403	-0.4377	-0.75621	-0.07808	0.866033	-1.29738	-1.0935	0.631216	-0.16587	0.35359	1.53E-16	-5.20E-16
A7	-1.08897	1.869153	-0.13787	0.637313	-1.88923	-0.50117	0.449644	1.195484	-1.46286	-1.94988	0.586206	-0.22532	-0.3346	0.008439	-1.67E-16	2.50E-16
A8	-1.11877	2.673005	-1.76365	-1.67224	-0.24863	0.693379	1.017295	-0.10562	2.585629	-0.50958	0.305655	-0.38443	0.243119	0.171807	1.11E-16	2.78E-16
A9	0.900298	1.850794	-0.11281	0.032907	-0.90304	-1.97777	-0.37399	0.616251	0.753379	1.024278	0.302956	1.051575	-0.68896	-0.54557	-4.96E-16	1.86E-16
A10	-0.87715	0.177495	-1.27987	1.215646	-0.89208	2.696699	-0.50164	1.573881	-0.17071	1.006984	-0.98495	0.176211	-0.41166	0.429938	1.11E-16	-8.33E-17
A11	1.694498	-0.03859	-2.39049	-1.26303	-0.09055	-1.45116	-0.68533	-1.58369	-0.90457	0.570556	0.486496	-0.31995	-0.47743	0.671955	-5.27E-16	-5.83E-16
A12	0.365924	-2.25815	0.918238	1.600819	2.669015	-1.2186	1.293164	1.297027	0.641865	0.234533	0.34175	-1.03882	-0.48263	0.046805	6.11E-16	-1.39E-16
A13	-6.10157	-2.84629	-0.01743	-1.70941	0.126477	-1.31233	-0.50709	-0.14046	-0.14603	0.033105	-0.39369	0.312488	0.086238	-0.02429	-3.33E-16	-3.33E-16
A14	-1.01795	1.728301	2.927025	0.806812	-0.15211	0.668127	-0.6968	-1.0656	-0.11342	0.438884	0.201315	-0.40114	0.022335	-0.03072	-4.16E-17	5.07E-16
A15	1.230858	-3.41014	-1.89512	1.228578	-2.09005	0.982482	-0.51374	-1.21782	0.486011	-0.40118	-0.22725	-0.60435	-0.13773	-0.59898	8.88E-16	1.11E-16
A16	-1.01795	1.728301	2.927025	0.806812	-0.15211	0.668127	-0.6968	-1.0656	-0.11342	0.438884	0.201315	-0.40114	0.022335	-0.03072	-4.16E-17	5.07E-16

Appendix B: ClustVis DNA PCA Loading Data with Cutoff

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16
D3S1358	-0.10541	-0.2158	0.095133	0.076478	0.392715	0.235809	-0.52678	0.287369	0.163564	0.382038	-0.10793	0.041185	0.09231	-0.0381	0.183954	-0.35718
D1S1656	-0.27812	0.282387	-0.31504	0.138261	0.073578	0.103591	-0.13438	-0.46606	-0.22163	-0.18088	-0.21573	0.308376	0.066536	0.138302	0.437588	-0.13354
D2S441	0.017756	-0.14601	-0.45078	0.208392	0.207946	-0.32286	-0.07897	-0.26317	0.063668	0.266353	-0.16891	0.011832	0.257652	0.174937	-0.50048	0.078773
D13S317	-0.11577	-0.44367	-0.05104	-0.39475	0.025276	0.040024	0.148007	0.092525	-0.15364	0.129321	0.184179	0.385888	0.104491	0.358893	0.268244	0.291225
Penta E	0.330932	0.20177	-0.15745	-0.23498	0.197432	-0.34466	0.058854	-0.20974	0.074363	0.231449	0.499812	0.023861	-0.21373	-0.07515	0.217246	-0.38193
D18S51	0.201292	0.10761	-0.05055	-0.02764	-0.29319	0.483756	-0.42989	-0.24698	0.287464	-0.10414	0.231482	-0.00511	-0.16571	0.282255	-0.15978	0.075323
CSF1PO	0.441126	-0.04992	-0.25444	-0.19586	-0.06495	0.068956	-0.11081	0.184912	-0.18102	-0.3855	-0.1165	-0.0524	0.527261	-0.26329	0.095635	-0.18177
Penta D	-0.28244	0.131306	-0.32439	-0.26948	-0.17872	-0.17431	-0.12894	0.46201	-0.00355	-0.2215	-0.0191	0.219845	-0.24643	0.21561	-0.31386	-0.34288
TH01	0.477215	-0.13612	-0.09565	-0.09818	0.083903	-0.08581	-0.20637	0.029151	-0.31876	0.050189	-0.44234	-0.11215	-0.52975	0.157526	0.100556	0.187748
D21S11	0.005583	-0.36818	0.046755	0.260857	-0.38817	-0.26316	-0.10358	-0.09756	0.199221	-0.05899	-0.12419	0.432079	-0.24221	-0.4307	0.103705	-0.09024
D7S820	0.094652	-0.45897	0.149705	-0.16125	0.196761	-0.12142	-0.06646	-0.30192	0.301991	-0.44518	0.065655	-0.01096	0.087413	0.192662	-0.07823	-0.14011
TPOX	-0.02043	-0.01594	-0.10312	-0.07113	0.556391	0.397903	0.278424	-0.04325	0.023821	-0.22992	-0.04053	0.253056	-0.31075	-0.33632	-0.26887	0.020887
D8S1179	-0.31104	-0.15742	-0.38145	-0.18941	-0.00323	-0.02801	-0.355	-0.05241	-0.08378	-0.01672	0.300825	-0.29616	-0.08143	-0.44081	0.05263	0.380394
D12S391	0.156756	-0.14423	0.011177	0.558749	0.070784	0.042012	-0.10811	0.154999	-0.49915	-0.13179	0.50209	0.180235	-0.01445	0.101044	-0.1319	-0.02606
D19S433	0.072718	-0.25421	-0.18824	-0.14594	-0.35208	0.409117	0.255103	-0.20699	-0.20289	0.401196	-0.03211	0.014607	0.029003	-0.12076	-0.18765	-0.32947
D22S1045	0.133967	0.301759	0.342455	-0.31178	0.010964	-0.09862	-0.31335	-0.11292	-0.15154	0.159764	-0.01257	0.493373	0.198943	-0.19794	-0.29337	0.219881
FGA	-0.3033	-0.13132	0.381899	-0.18299	0.004273	-0.10272	-0.15146	-0.28981	-0.47214	-0.10197	-0.03479	-0.27428	-0.08428	-0.00744	-0.17645	-0.30933

Appendix B: ClustVis Individual and Cumulative DNA PCA Values with Cutoff

	Individual	Cumulative
PC1	0.182556	0.182556
PC2	0.16471	0.347266
PC3	0.1501	0.497366
PC4	0.118042	0.615408
PC5	0.111317	0.726725
PC6	0.092028	0.818753
PC7	0.064598	0.883351
PC8	0.03657	0.919921
PC9	0.030771	0.950692
PC10	0.019204	0.969896
PC11	0.013315	0.983211
PC12	0.00938	0.99259
PC13	0.005321	0.997911
PC14	0.002089	1
PC15	0	1
PC16	0	1

Appendix B: ClustVis DNA PCA Scores with Cutoff

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16
A1	0.040836	-1.90693	0.058323	1.258425	-0.17687	0.197019	-3.16435	0.353007	-0.44865	-0.69947	0.351179	-0.10477	0.022952	-0.06248	1.39E-17	-3.89E-16
A2	0.225999	-0.18652	-1.18105	1.340214	0.49806	-0.38167	1.442267	0.252848	-1.03875	0.011474	0.524997	-0.70423	-0.56403	-0.12819	2.78E-17	-1.94E-16
A3	-0.45664	-2.39931	1.586047	-2.84658	-0.53672	0.779013	1.220647	-0.07619	-0.60323	-0.6586	0.103631	-0.256	0.430607	-0.13011	0	1.11E-16
A4	-1.66467	0.942629	1.826015	1.2581	-0.58588	-1.72366	0.324961	-0.63407	1.316913	-0.57292	-0.23798	-0.06862	-0.07539	-0.34077	-1.11E-16	-4.72E-16
A5	0.886597	1.371242	1.416921	-0.44415	-3.26605	0.64498	-0.3199	-0.53499	-0.56523	0.893415	0.458672	0.470666	-0.18898	-0.10553	4.51E-16	4.72E-16
A6	-1.7649	-3.18286	-2.07501	-1.33947	-0.3655	0.299701	0.059764	0.558415	0.799609	0.359732	-0.3792	0.444156	-0.49526	-0.02906	-1.26E-15	1.36E-15
A7	-0.97203	0.746826	-0.70571	0.895152	0.657862	-0.42626	0.115633	-0.53971	-1.62557	-0.16451	-1.10386	0.488818	0.09167	-0.03258	-3.47E-16	-1.39E-16
A8	-1.13365	1.364968	3.138955	-1.05744	1.675224	0.913137	-0.87933	0.677181	0.044668	0.749838	-0.49079	-0.50627	-0.21329	0.085552	6.66E-16	-7.22E-16
A9	-1.58135	0.294581	0.152661	0.505583	1.957399	-1.19761	0.421934	1.058347	-0.06895	0.584507	0.82386	0.566563	0.410786	-0.06151	-6.07E-17	1.87E-16
A10	0.896661	1.928767	0.279134	0.937325	-0.86156	1.698275	0.911436	1.529859	0.371519	-1.03943	-0.07974	0.316465	-0.07584	0.184291	-3.33E-16	7.22E-16
A11	0.889591	-1.58796	1.837589	0.501705	1.2615	-0.49442	0.446144	-1.46183	0.204046	-0.25729	0.336988	0.301227	-0.22982	0.417356	4.16E-16	-1.67E-16
A12	0.862561	-0.33067	-0.58183	-0.15219	-2.12656	-2.86943	-0.02978	0.711244	0.053508	0.315131	-0.32535	-0.43869	0.256818	0.295496	1.80E-16	8.33E-17
A13	4.615276	1.376515	-1.15699	-2.19929	1.645428	-0.88281	-0.50014	-0.07486	0.307983	-0.17813	-0.067	0.144697	-0.08912	-0.17918	9.44E-16	-5.55E-17
A14	-1.63605	1.684799	-2.13303	-0.57531	0.011537	0.817639	-0.28458	-0.77416	0.317241	-0.03242	0.21262	-0.23658	0.159337	0.093482	-4.86E-17	-1.80E-16
A15	2.427813	-1.80087	-0.32901	2.493247	0.2006	1.808473	0.519884	-0.27093	0.617646	0.77509	-0.34063	-0.20084	0.400213	-0.10025	-2.78E-16	-3.33E-16
A16	-1.63605	1.684799	-2.13303	-0.57531	0.011537	0.817639	-0.28458	-0.77416	0.317241	-0.03242	0.21262	-0.23658	0.159337	0.093482	-4.86E-17	-1.80E-16

Appendix C: NIST Allele Frequencies for U.S. Caucasian Population, N=361

	CSF1PO	D10S1248	D12S391	D13S317	D16S539	D18S51	D19S433	D1S1656	D21S11	D22S1045	D2S1338	D2S441	D3S1358	D5S818	D6S1043	D7S820	D8S1179	F13A01	F13B	FESFPS	FGA	LPL	Penta_C	Penta_D	Penta_E	SE33	TH01	TPOX	vWA
	361	361	361	361	361	361	361	361	361	361	361	361	361	361	361	361	361	361	361	361	361	361	361	361	361	361	361	361
Allele																												
2.2																								0.0042				
3.2																		0.0637										
4.0																		0.0263										
4.2																												
5.0																		0.1939		0.0014			0.0083		0.0762		0.0014	0.0014	
6.0																		0.3490	0.0956					0.0042			0.2355	0.0014	
6.3																													
7.0														0.0028		0.0277		0.3158	0.0166			0.0014	0.0028	0.0042	0.1690		0.1939		
8.0	0.0055			0.1205	0.0180									0.0055	0.0014	0.1440	0.0139	0.0042	0.2465	0.0249			0.0055	0.0208	0.0139		0.0956	0.5249	
8.1																0.0014													
9.0	0.0139			0.0776	0.1066							0.0014		0.0416		0.1676	0.0055		0.2465			0.0388	0.1496	0.2216	0.0125		0.1191	0.1274	
9.1												0.0014																	
9.3																											0.3449		
10.0	0.2202			0.0471	0.0568	0.0083	0.0014	0.0028				0.2105		0.0554	0.0166	0.2562	0.1025	0.0014	0.3892	0.2825		0.4224	0.0665	0.1150	0.0859		0.0083	0.0499	
10.1																													
10.2																													
10.3																													
11.0	0.3089	0.0014		0.3255	0.3144	0.0097	0.0055	0.0776		0.1399		0.3435	0.0014	0.3560	0.2964	0.2050	0.0762		0.0055	0.4114		0.2645	0.3947	0.1260	0.0873	0.0014	0.0014	0.2521	
11.2																													
11.3												0.0609																	
12.0	0.3601	0.0319		0.2687	0.3144	0.1136	0.0706	0.1163		0.0125		0.0471		0.3878	0.2355	0.1593	0.1676	0.0014		0.2355		0.2327	0.2105	0.2327	0.1994	0.0069		0.0416	0.0014
12.2							0.0014																						
12.3												0.0042																	
13.0	0.0817	0.3075		0.1163	0.1634	0.1233	0.2548	0.0665		0.0069		0.0291	0.0014	0.1427	0.0873	0.0346	0.3296	0.0014		0.0429		0.0346	0.1427	0.1967	0.0859	0.0166		0.0014	0.0014
13.2							0.0069																						
13.3																													
13.4																													
14.0	0.0097	0.2978		0.0429	0.0263	0.1343	0.3615	0.0789		0.0568		0.2410	0.1066	0.0069	0.0554	0.0042	0.1662	0.0111		0.0014		0.0042	0.0180	0.0609	0.0623	0.0249			0.0928
14.2						0.0014	0.0235																			0.0028			
14.3								0.0028																					
15.0		0.1967	0.0319	0.0014		0.1704	0.1565	0.1496		0.3213	0.0014	0.0596	0.2729	0.0014	0.0125		0.1039	0.0249				0.0014	0.0014	0.0097	0.0429	0.0402			0.1053
15.2							0.0360																						
15.3								0.0582																					
15.4																									0.0014				
16.0		0.1330	0.0222			0.1468	0.0568	0.1357		0.3823	0.0374	0.0014	0.2382		0.0042		0.0332	0.0069						0.0028	0.0512	0.0402			0.2008
16.2						0.0014	0.0152																			0.0014			
16.3								0.0609																					
17.0		0.0277	0.1274			0.1385	0.0069	0.0471		0.0748	0.1856		0.2105		0.0609		0.0014							0.0014	0.0485	0.0623			0.2839
17.1																													
17.2							0.0014																						
17.3			0.0208					0.1330																		0.0014			
18.0		0.0014	0.1717			0.0776		0.0055		0.0055	0.0734		0.1510		0.0886						0.0249				0.0332	0.0720			0.2022
18.1																													
18.2							0.0014																						
18.3			0.0249					0.0499																		0.0014			
19.0		0.0028	0.1247			0.0402					0.1205		0.0166		0.0983						0.0499				0.0152	0.0720			0.1039
19.1																													
19.2																										0.0042			
19.3			0.0042					0.0152																					
19.4																													
20.0			0.1108			0.0180					0.1565		0.0014		0.0319						0.1233				0.0097	0.0582			0.0069
20.1																													
20.2																										0.0097			
20.3																													
21.0			0.1288			0.0097					0.0374				0.0097						0.1787				0.0028	0.0249			0.0014
21.2																					0.0055					0.0235			
21.3																													
22.0			0.0956			0.0069					0.0346				0.0014						0.2050				0.0014	0.0139			
22.2																					0.0125					0.0374			
22.3																													
23.0			0.0693								0.1053										0.1524					0.0028			
23.2																					0.0028					0.0360			
23.3																													
24.0			0.0471								0.1150										0.1343				0.0014	0.0014			
24.2																					0.0014					0.0222			
24.3																													
25.0			0.0166								0.1025										0.0789								
25.2									0.0014																	0.0416			
26.0			0.0028								0.0305										0.0263								
26.2																										0.0416			
27.0			0.0014						0.0222												0.0042								
27.2																										0.0942			
27.3																													
28.0									0.1593																				
28.2																										0.0762			
28.3																										0.0014			
29.0									0.2022																				
29.2									0.0028																	0.0554			
29.3																													
30.0									0.2825																				
30.2									0.0291																	0.0568			
30.3																													
31.0									0.0720																	0.0028			
31.2									0.0983																	0.0235			
32.0									0.0055																	0.0014			
32.2									0.0900																	0.0125			
33.0									0.0014																	0.0014			
33.1																													
33.2									0.0263																	0.0042			
34.0																										0.0083			
34.2									0.0042																	0.0014			
35.0									0.0014																				
36.0									0.0014																				
37.0																													
38.0																													
39.0																													
43.2																													
H_{obs}	0.7341	0.7645	0.8975	0.7618	0.7645	0.8560	0.7673	0.9252	0.8227	0.7535	0.8698	0.7867	0.7562	0.7064	0.7978	0.8310	0.7839	0.7590	0.7008	0.6731	0.8670	0.7175	0.7701	0.8587	0.8920	0.9501	0.7424	0.6537	0.8061
P_I	0.1285	0.0989	0.0237	0.0777	0.0983	0.0305	0.0838	0.0211	0.0512	0.1239	0.0276	0.0884	0.0758	0.1486	0.0513	0.0628	0.0617	0.1165	0.1288	0.1445	0.0399	0.1510	0.1029	0.0588	0.0243	0.0079	0.0931	0.1811	0.0660
PPE	0.4708	0.5346	0.7794	0.5877	0.5436	0.7481	0.5698	0.7957	0.6661	0.4888	0.7606	0.5616	0.5842	0.4503	0.6609	0.6330	0.6302	0.5035	0.4699	0.4332	0.7173	0.4403	0.5407	0.6509	0.7830	0.8954	0.5476	0.3988	0.6206
HWE	0.8706	0.6522	0.8859	0.1028	0.1647	0.1988	0.9663	0.8772	0.6072	0.9797	0.5097	0.7125	0.3803	0.7578	0.0272	0.5050	0.8675	0.8813	0.8313	0.8566	0.0441	0.9722	0.0213	0.1491	0.4988	0.2478	0.2253	0.5509	0.4328

Appendix C: July 19, 2017, NIST Allele Frequencies for African American Population, N=342

	CSF1PO	D10S1248	D12S391	D13S317	D16S539	D18S51	D19S433	D1S1656	D21S11	D22S1045	D2S1338	D2S441	D3S1358	D5S818	D6S1043	D7S820	D8S1179	F13A01	F13B	FESFPS	FGA	LPL	Penta_C	Penta_D	Penta_E	SE33	TH01	TPOX	WA
	342	342	342	342	342	342	342	342	342	342	342	342	342	342	342	342	342	342	342	342	342	342	342	342	342	342	341	342
Allele																												
2.2																								0.1140				
3.2																		0.1184						0.0088				
4.0																		0.0687										
4.2																		0.0015										
5.0					0.0015													0.3392					0.0249	0.0439	0.0950		0.0044		
6.0																0.0015		0.1272	0.3626	0.0015				0.0102	0.0015		0.1316	0.0894	
6.3																			0.0015										
7.0	0.0556													0.0029		0.0117		0.1974	0.1681	0.0029		0.0161	0.0161	0.0439	0.1038	0.0015	0.4079	0.0176	
8.0	0.0556	0.0029		0.0278	0.0322					0.0073				0.0468		0.2281	0.0073	0.0570	0.1082	0.0994		0.0044	0.0512	0.1082	0.1667		0.1959	0.3680	
8.1																													
9.0	0.0395	0.0029		0.0336	0.1827	0.0029						0.0029		0.0322	0.0015	0.1155	0.0044	0.0088	0.2281	0.0278		0.1301	0.1725	0.1681	0.0512		0.1594	0.1950	
9.1																													
9.3																											0.0965		
10.0	0.2500	0.0073		0.0307	0.1170	0.0044	0.0102	0.0146		0.0409		0.0848		0.0731	0.0058	0.3363	0.0307	0.0058	0.1287	0.2251		0.3480	0.0702	0.0994	0.0468		0.0044	0.0865	
10.1																							0.0015						
10.2																										0.0015			
10.3																				0.0424									
11.0	0.2485	0.0351		0.3099	0.3143	0.0015	0.0629	0.0453		0.1447		0.3626		0.2339	0.1535	0.2032	0.0526	0.0073	0.0029	0.3275		0.1243	0.2778	0.1798	0.0643			0.2155	0.0029
11.2																										0.0029			
11.3												0.0439								0.0029									
12.0	0.2953	0.1301		0.4181	0.2047	0.0760	0.1228	0.0643		0.0541		0.1652	0.0044	0.3684	0.2237	0.0877	0.1301	0.0044		0.2266		0.2953	0.2500	0.1082	0.1287	0.0029		0.0264	0.0015
12.2							0.0365																			0.0029			
12.3												0.0058																	
13.0	0.0468	0.2339		0.1404	0.1228	0.0409	0.2456	0.1009		0.0029		0.0439	0.0029	0.2237	0.0980	0.0146	0.2193	0.0336		0.0439		0.0746	0.0877	0.0833	0.1038	0.0117		0.0015	0.0088
13.2						0.0044	0.0526																			0.0058			
13.3												0.0029																	
13.4																								0.0015					
14.0	0.0088	0.2763	0.0015	0.0395	0.0249	0.0716	0.2105	0.2573		0.0775		0.2675	0.0906	0.0161	0.0585	0.0015	0.2939	0.0175				0.0073	0.0395	0.0249	0.0687	0.0512			0.0804
14.2							0.0746																			0.0058			
14.3								0.0073				0.0015																	
15.0		0.1974	0.0775			0.1652	0.0804	0.1579		0.2515	0.0015	0.0190	0.3085	0.0029	0.0526		0.1901	0.0132					0.0058	0.0044	0.0556	0.0439			0.1915
15.2						0.0015	0.0614						0.0015													0.0029			
15.3								0.0292																					
15.4																													
16.0		0.0877	0.0673			0.1711	0.0044	0.1096		0.1915	0.0556		0.3187		0.0161		0.0643						0.0029		0.0409	0.0482			0.2500
16.2							0.0263														0.0015					0.0044			
16.3								0.1023																					
17.0		0.0249	0.1667			0.1520		0.0278		0.2091	0.1009		0.2120		0.0570		0.0044							0.0015	0.0439	0.0950			0.2354
17.1			0.0044																										
17.2							0.0088														0.0015					0.0015			
17.3			0.0044					0.0497																		0.0015			
18.0		0.0015	0.2529			0.1213		0.0029		0.0146	0.0424		0.0570		0.1082		0.0029				0.0015				0.0161	0.1199			0.1491
18.1			0.0015																										
18.2							0.0029														0.0175								
18.3			0.0044					0.0234							0.0015														
19.0			0.1477			0.0994				0.0058	0.1389		0.0044		0.1126						0.0512				0.0073	0.1272			0.0629
19.1			0.0088																										
19.2																					0.0029								
19.3			0.0044					0.0073							0.0015														
19.4																													
20.0			0.1038			0.0629					0.1038				0.0716						0.0541				0.0044	0.0980			0.0161
20.1			0.0029																										
20.2																										0.0029			
20.3																													
21.0			0.0643			0.0102					0.1360				0.0175						0.1228					0.0482			0.0015
21.2						0.0015																				0.0102			
21.3															0.0015														
22.0			0.0365			0.0073					0.1374										0.1988					0.0146			
22.2																					0.0044					0.0044			
22.3																					0.0015								
23.0			0.0292			0.0044					0.1038				0.0102						0.1696					0.0073			
23.2																					0.0015					0.0175			
23.3																													
24.0			0.0132			0.0015					0.0833				0.0058						0.1330								
24.2																										0.0132			
24.3																													
25.0			0.0088								0.0775				0.0015						0.1184				0.0015				
25.2																					0.0015					0.0307			
26.0									0.0015		0.0146				0.0015						0.0702					0.0015			
26.2																										0.0629			
27.0									0.0746		0.0044										0.0234								
27.2																										0.0424			
27.3																													
28.0									0.2456												0.0146								
28.2																										0.0453			
28.3																										0.0015			
29.0									0.2047												0.0058								
29.2																										0.0234			
29.3									0.0015																				
30.0									0.1696												0.0015								
30.2									0.0175												0.0015					0.0161			
30.3																													
31.0									0.0789																	0.0015			
31.2									0.0512												0.0015					0.0161			
32.0									0.0088																				
32.2									0.0614																	0.0044			
33.0									0.0044																				
33.1									0.0029																				
33.2									0.0351																	0.0029			
34.0									0.0058																	0.0029			
34.2																													
35.0									0.0219																				
36.0									0.0088																	0.0015			
37.0									0.0029																				
38.0									0.0015																				
39.0									0.0015																				
43.2																													
H_{obs}	0.7895	0.8216	0.8626	0.7047	0.7895	0.8538	0.8889	0.8713	0.8392	0.8041	0.9006	0.7895	0.7632	0.7281	0.8860	0.7719	0.7953	0.8216	0.7719	0.7690	0.8772	0.7573	0.8129	0.8684	0.9035	0.9269	0.7632	0.7595	0.8158
P_I	0.0827	0.0693	0.0373	0.1342	0.0733	0.0280	0.0401	0.0338	0.0435	0.0552	0.0229	0.1051	0.1090	0.0974	0.0296	0.0889	0.0660	0.0637	0.0968	0.0847	0.0318	0.1028	0.0652	0.0268	0.0200	0.0108	0.1050	0.0879	0.0623
PPE	0.5724	0.6138	0.7214	0.4687	0.6018	0.7555	0.7167	0.7356	0.7028	0.6524	0.7832	0.5471	0.5158	0.5310	0.7554	0.5615	0.6186	0.6334	0.5424	0.5705	0.7457	0.5347	0.6368	0.7607	0.8069	0.8606	0.5259	0.5543	0.6283
HWE	0.9734	0.5478	0.2306	0.4413	0.0875	0.8266	0.4372	0.1013	0.2025	0.2738	0.9331	0.051875	0.9097	0.6428	0.4438	0.5644	0.6175	0.3141	0.9197	0.6750	0.3225	0.5141	0.3803	0.3994	0.3972	0.9322	0.7741	0.8472	0.7872

Appendix C: July 19, 2017, NIST Allele Frequencies for U.S. Hispanic Population, N=236

	CSF1PO	D10S1248	D12S391	D13S317	D16S539	D18S51	D19S433	D1S1656	D21S11	D22S1045	D2S1338	D2S441	D3S1358	D5S818	D6S1043	D7S820	D8S1179	F13A01	F13B	FESFPS	FGA	LPL	Penta_C	Penta_D	Penta_E	SE33	TH01	TPOX	WA
	236	236	236	236	236	236	236	236	236	236	236	236	236	236	236	236	236	236	236	236	236	236	236	235	236	236	236	236
Allele																												
2.2																								0.0170				
3.2																		0.2140						0.0021				
4.0																		0.0826										
4.2																												
5.0																		0.1949					0.0169	0.0064	0.0360			
6.0																		0.1674	0.1208					0.0021		0.2394	0.0085	
6.3																									0.0021			
7.0	0.0127													0.0339		0.0106		0.3072	0.0212	0.0021		0.0042	0.0064	0.0021	0.1165		0.2966	0.0064	
8.0	0.0042			0.1102	0.0191									0.0085		0.1208	0.0148	0.0064	0.1864	0.0169			0.0339	0.0191	0.0254		0.0911	0.4852	
8.1																												
9.0	0.0233			0.1653	0.1398		0.0021							0.0530		0.0911	0.0064	0.0042	0.2309	0.0085		0.0424	0.2225	0.2426	0.0169		0.1462	0.0932	
9.1																												
9.3																										0.2182		
10.0	0.2373			0.0996	0.1504	0.0021	0.0021	0.0064		0.0148		0.3369		0.0572	0.0042	0.3072	0.0932		0.4386	0.2034		0.4831	0.0657	0.1638	0.0847		0.0085	0.0487	
10.1																												
10.2																												
10.3																0.0021				0.0021								
11.0	0.2797	0.0042		0.2182	0.2648	0.0148	0.0148	0.0275		0.0636		0.2987		0.3898	0.1780	0.2775	0.0530		0.0021	0.4640		0.1992	0.3305	0.1553	0.0763			0.2542	0.0021
11.2																												
11.3												0.0445																
12.0	0.3750	0.0424		0.2352	0.2775	0.1144	0.0657	0.0890		0.0127		0.0360		0.3390	0.2055	0.1547	0.1292	0.0042		0.2182		0.2140	0.2097	0.1574	0.1695	0.0042		0.1038	
12.2							0.0127																			0.0042			
12.3												0.0021																	
13.0	0.0593	0.2733		0.1059	0.1335	0.1229	0.2225	0.1144		0.0085		0.0233	0.0064	0.1081	0.1017	0.0360	0.2733	0.0064		0.0742		0.0487	0.1038	0.1447	0.0932	0.0085			
13.2							0.0445																					
13.3																												
13.4																												
14.0	0.0064	0.3390		0.0614	0.0127	0.1610	0.3538	0.1165		0.0275		0.2055	0.0784	0.0085	0.1314		0.2627			0.0106		0.0085	0.0106	0.0702	0.0720	0.0275			0.0805
14.2						0.0021	0.0381																			0.0021			
14.3								0.0042																					
15.0	0.0021	0.2119	0.0445	0.0042	0.0021	0.1589	0.1356	0.1377		0.4258		0.0487	0.3220	0.0021	0.0318		0.1292	0.0064						0.0106	0.0953	0.0360			0.1441
15.2							0.0551																			0.0064			
15.3								0.0508																					
15.4																									0.0021				
16.0		0.0996	0.0424			0.1250	0.0254	0.1758		0.3496	0.0297	0.0021	0.2797		0.0021		0.0318	0.0042						0.0043	0.0614	0.0699			0.2839
16.2							0.0275																						
16.3								0.0508																		0.0042			
17.0		0.0254	0.0763			0.1250		0.0424		0.0911	0.1695	0.0021	0.1843		0.0487		0.0042	0.0021			0.0021			0.0021	0.0551	0.0763			0.2458
17.1																													
17.2																													
17.3			0.0169					0.1483																		0.0042			
18.0		0.0021	0.1780			0.0784		0.0064		0.0064	0.0805		0.1229		0.1081		0.0021				0.0127				0.0339	0.1102			0.1801
18.1															0.0021														
18.2																													
18.3			0.0127					0.0254																					
19.0		0.0021	0.1886			0.0466					0.1928		0.0042		0.0784						0.0805				0.0212	0.0890			0.0508
19.1			0.0021																										
19.2																										0.0064			
19.3			0.0064					0.0042							0.0021														
19.4																													
20.0			0.1547			0.0275					0.1271		0.0021		0.0318						0.0847				0.0212	0.0487			0.0106
20.1																													
20.2																										0.0042			
20.3			0.0021												0.0106														
21.0			0.1123			0.0085					0.0318				0.0064						0.1525				0.0064	0.0169			0.0021
21.2																										0.0169			
21.3															0.0424														
22.0			0.0678			0.0106					0.0572										0.1653				0.0021	0.0148			
22.2			0.0021																		0.0042					0.0212			
22.3															0.0127														
23.0			0.0572								0.1398										0.1208				0.0064				
23.2																					0.0042					0.0233			
23.3															0.0021														
24.0			0.0169			0.0021					0.0763										0.1419								
24.2									0.0021																	0.0233			
24.3																													
25.0			0.0064								0.0784										0.1186				0.0042				
25.2																										0.0233			
26.0			0.0064								0.0169										0.0614								
26.2									0.0021																	0.0742			
27.0			0.0064						0.0275												0.0445					0.0021			
27.2																										0.0763			
27.3																										0.0021			
28.0									0.0996												0.0021								
28.2																										0.0678			
28.3																													
29.0									0.2076												0.0021								
29.2									0.0021																	0.0614			
29.3																													
30.0									0.2733												0.0021					0.0021			
30.2									0.0233																	0.0403			
30.3																													
31.0									0.0763																				
31.2									0.0996																	0.0127			
32.0									0.0169																				
32.2									0.1271																	0.0106			
33.0									0.0042																	0.0021			
33.1									0.0021																				
33.2									0.0339																	0.0042			
34.0									0.0021																				
34.2																													
35.0																													
36.0																													
37.0																													
38.0																													
39.0																													
43.2																													
H_{obs}	0.7288	0.7627	0.8898	0.8517	0.7881	0.9068	0.7797	0.8814	0.8644	0.7331	0.8517	0.7542	0.7331	0.7373	0.8898	0.8220	0.7839	0.7966	0.7034	0.6822	0.8814	0.6610	0.7542	0.8681	0.8983	0.9280	0.7585	0.6737	0.7881
P_I	0.1255	0.1050	0.0332	0.0526	0.0762	0.0303	0.0677	0.0284	0.0468	0.1743	0.0297	0.1098	0.0904	0.1457	0.0330	0.0867	0.0621	0.0781	0.1371	0.1416	0.0278	0.1545	0.0788	0.0499	0.0181	0.0099	0.0848	0.1554	0.0730
PPE	0.4756	0.5286	0.7559	0.6657	0.5918	0.7563	0.6172	0.7679	0.6829	0.4388	0.7474	0.5250	0.5427	0.4825	0.7486	0.5761	0.6303	0.5829	0.4635	0.4506	0.7585	0.4374	0.5789	0.6775	0.8199	0.8790	0.5624	0.4452	0.6009
HWE	0.8950	0.4475	0.1050	0.8219	0.6644	0.6691	0.0916	0.2778	0.9872	0.1353	0.7247	0.0397	0.9116	0.4516	0.5547	0.4806	0.1369	0.9116	0.3322	0.8238	0.8784	0.7781	0.5275	0.6644	0.5516	0.7269	0.7919	0.2359	0.6691

Appendix C: July 19, 2017, NIST Allele Frequencies for U.S Asian Population, N=342

	CSF1PO	D10S1248	D12S391	D13S317	D16S539	D18S51	D19S433	D1S1656	D21S11	D22S1045	D2S1338	D2S441	D3S1358	D5S818	D6S1043	D7S820	D8S1179	F13A01	F13B	FESFPS	FGA	LPL	Penta_C	Penta_D	Penta_E	SE33	TH01	TPOX	vWA
	97	97	97	97	97	97	97	97	97	97	97	97	97	97	97	97	97	97	97	97	97	97	97	97	97	97	97	97
Allele																												
2.2																												
3.2																		0.3093										
4.0																		0.1237										
4.2																												
5.0																		0.0619					0.0258		0.0722			
6.0																		0.4897	0.0155					0.0103			0.1701	
6.3																												
7.0	0.0206													0.0155		0.0052		0.0155	0.0052			0.0052	0.0052				0.2680	
8.0				0.2165								0.0052		0.0052		0.1340			0.0464				0.0670	0.0052	0.0052		0.0722	0.5464
8.1																												
9.0	0.0670			0.1443	0.3557									0.0979		0.0464			0.1959	0.0155		0.0052	0.3557	0.3144	0.0258		0.4433	0.0825
9.1												0.0052																
9.3																											0.0412	
10.0	0.2010	0.0052		0.1031	0.1649							0.2680		0.2268	0.0309	0.2629	0.1237		0.7320	0.0928		0.6443	0.0670	0.1856	0.0619		0.0052	0.0309
10.1																												
10.2																												
10.3																				0.0052								
11.0	0.2165			0.2680	0.1907			0.0309		0.2010		0.3505		0.2732	0.1495	0.3505	0.1186		0.0052	0.3608		0.1186	0.3144	0.1804	0.1598			0.2938
11.2																												
11.3												0.0309																
12.0	0.3866	0.0876		0.2113	0.1753	0.0361	0.0361	0.0464		0.0052		0.2165		0.2062	0.1237	0.1753	0.1186			0.3093		0.2113	0.1443	0.1804	0.0928			0.0464
12.2							0.0258																					
12.3												0.0052			0.0052													
13.0	0.0876	0.3196		0.0567	0.0979	0.2165	0.2835	0.1340				0.0258		0.1598	0.1237	0.0258	0.2010			0.1959		0.0155	0.0155	0.1031	0.0515			
13.2							0.0206																					
13.3																												
13.4																												
14.0	0.0155	0.2526			0.0155	0.2371	0.2990	0.0619		0.0103		0.0825	0.0258	0.0155	0.1546		0.2010			0.0206			0.0052	0.0103	0.0515				0.1959
14.2							0.1031																					
14.3																												
15.0	0.0052	0.2062	0.0412			0.1804	0.0619	0.2784		0.3093		0.0103	0.3660		0.0361		0.1289							0.0103	0.1134	0.0103			0.0206
15.2							0.1237																					
15.3																												
15.4																												
16.0		0.0979	0.0103			0.1289	0.0103	0.2010		0.2268	0.0206		0.3299		0.0052		0.0928								0.0567	0.0309			0.1392
16.2							0.0309																					
16.3								0.0155																				
17.0		0.0309	0.0825			0.0670		0.0722		0.2216	0.0567		0.2010		0.0722		0.0103				0.0103				0.0825	0.0309			0.3144
17.1																												
17.2							0.0052																					
17.3								0.0876																				
18.0			0.2629			0.0309		0.0155		0.0206	0.1340		0.0670		0.1598		0.0052				0.0258				0.0773	0.0619			0.2062
18.1																												
18.2																												
18.3								0.0515																				
19.0			0.1753			0.0412				0.0052	0.1804		0.0103		0.0928						0.0567				0.0412	0.0619			0.1082
19.1																												
19.2																												
19.3			0.0052					0.0052																				
19.4																									0.0103			
20.0			0.1959			0.0258					0.1598				0.0412						0.0876				0.0206	0.0928			0.0155
20.1																												
20.2																										0.0155		
20.3																												
21.0			0.0979			0.0103					0.0155				0.0052						0.1031				0.0206	0.0515			
21.2																										0.0258		
21.3																												
22.0			0.0567			0.0155					0.0515										0.2423				0.0309	0.0155			
22.2																										0.0412		
22.3																												
23.0			0.0258			0.0052					0.1649										0.2062				0.0206	0.0103			
23.2																										0.0464		
23.3																												
24.0			0.0103								0.1289										0.1495								
24.2																					0.0052					0.0619		
24.3			0.0052																									
25.0			0.0155								0.0464										0.0722				0.0052				
25.2																										0.0567		
26.0			0.0103								0.0361										0.0309								
26.2																										0.0722		
27.0			0.0052								0.0052										0.0052								
27.2																										0.1031		
27.3																												
28.0						0.0052			0.0567																			
28.2									0.0052																	0.0773		
28.3																												
29.0									0.2010																			
29.2																										0.0619		
29.3																												
30.0									0.3299																			
30.2									0.0052																	0.0361		
30.3									0.0103																			
31.0									0.1237																			
31.2									0.0361																	0.0258		
32.0									0.0567																			
32.2									0.1134																	0.0103		
33.0									0.0155																			
33.1																												
33.2									0.0464																			
34.0																												
34.2																												
35.0																												
36.0																												
37.0																												
38.0																												
39.0																												
43.2																					0.0052							
H_{obs}	0.7835	0.7526	0.8660	0.8041	0.7423	0.8763	0.7835	0.8351	0.7732	0.7010	0.9072	0.8144	0.7423	0.8144	0.8351	0.6804	0.9072	0.6701	0.3711	0.8247	0.8763	0.5670	0.7216	0.7423	0.9175	0.9278	0.6804	0.6186	0.8144
P_I	0.1040	0.0919	0.0473	0.0758	0.0928	0.0507	0.0724	0.0481	0.0566	0.1030	0.0428	0.1187	0.1485	0.0790	0.0324	0.0985	0.0492	0.1935	0.3880	0.1474	0.0528	0.2854	0.1017	0.0766	0.0194	0.0190	0.1361	0.2122	0.0885
PPE	0.5351	0.5619	0.6849	0.6069	0.5580	0.6807	0.6135	0.6916	0.6440	0.5350	0.7382	0.5259	0.4613	0.5958	0.7562	0.5361	0.6976	0.3872	0.2191	0.4855	0.6994	0.2915	0.5235	0.5908	0.8347	0.8767	0.4535	0.3525	0.5872
HWE	0.3394	0.6250	0.8547	0.2150	0.5697	0.9350	0.3331	0.6150	0.8559	0.0256	0.1675	0.4441	0.4916	0.5206	0.4388	0.3403	0.5513	0.1603	0.0628	0.0953	0.0728	0.2134	0.3956	0.5516	0.7950	0.0091	0.8084	0.9791	0.0934

Appendix C: July 19, 2017, NIST Allele Frequencies for Entire U.S. (Cumulative Population), N=1036

	CSF1PO	D10S1248	D12S391	D13S317	D16S539	D18S51	D19S433	D1S1656	D21S11	D22S1045	D2S1338	D2S441	D3S1358	D5S818	D6S1043	D7S820	D8S1179	F13A01	F13B	FESFPS	FGA	LPL	Penta_C	Penta_D	Penta_E	SE33	TH01	TPOX	WA
	1036	1036	1036	1036	1036	1036	1036	1036	1036	1036	1036	1036	1036	1036	1036	1036	1036	1036	1036	1036	1036	1036	1036	1035	1036	1036	1035	1036
Allele																												
2.2																								0.0430				
3.2																		0.1390						0.0034				
4.0																		0.0623										
4.2																		0.0005										
5.0					0.00048													0.2297		0.0005			0.0174	0.0159	0.0729		0.0019	0.0005	
6.0																0.0005		0.2476	0.1819	0.0005				0.0063	0.0005		0.1959	0.0319	
6.3																			0.0005							0.0005			
7.0	0.0232													0.0111		0.0164		0.2466	0.0666	0.0014		0.0072	0.0082	0.0164	0.1197	0.0005	0.2949	0.0072	
8.0	0.0212	0.0010		0.0965	0.0212					0.0024		0.0005		0.0198	0.0005	0.1655	0.0106	0.0217	0.1684	0.0454		0.0014	0.0328	0.0478	0.0661		0.1255	0.4662	
8.1																0.0005													
9.0	0.0294	0.0010		0.0893	0.1626	0.0010	0.0005					0.0014		0.0463	0.0005	0.1216	0.0048	0.0039	0.2321	0.0125		0.0666	0.1931	0.2174	0.0275		0.1689	0.1377	
9.1												0.0010																	
9.3																											0.2056		
10.0	0.2321	0.0029		0.0589	0.1081	0.0048	0.0043	0.0072		0.0169		0.2032		0.0777	0.0116	0.2949	0.0787	0.0024	0.3465	0.2278		0.4324	0.0676	0.1275	0.0705		0.0068	0.0599	
10.1																							0.0005						
10.2																										0.0005			
10.3																0.0005				0.0150									
11.0	0.2736	0.0130		0.2905	0.2915	0.0072	0.0261	0.0512		0.1298		0.3403	0.0005	0.3156	0.2085	0.2346	0.0671	0.0024	0.0039	0.3909		0.1897	0.3340	0.1556	0.0840	0.0005	0.0005	0.2444	0.0014
11.2																										0.0010			
11.3												0.0487								0.0010									
12.0	0.3446	0.0719		0.3050	0.2568	0.0941	0.0835	0.0864		0.0256		0.0994	0.0014	0.3533	0.2143	0.1361	0.1419	0.0029		0.2355		0.2471	0.2172	0.1696	0.1593	0.0043		0.0512	0.0010
12.2							0.0179																			0.0019			
12.3												0.0043			0.0005														
13.0	0.0656	0.2765		0.1163	0.1371	0.1047	0.2471	0.0951		0.0053		0.0323	0.0029	0.1631	0.0975	0.0275	0.2683	0.0130		0.0647		0.0492	0.1038	0.1386	0.0903	0.0116		0.0010	0.0034
13.2						0.0014	0.0319																			0.0019			
13.3												0.0010																	
13.4																								0.0005					
14.0	0.0092	0.2958	0.0005	0.0420	0.0217	0.1293	0.3041	0.1448		0.0526		0.2268	0.0874	0.0111	0.0830	0.0019	0.2336	0.0097		0.0048		0.0058	0.0222	0.0464	0.0656	0.0319			0.0956
14.2						0.0010	0.0512																			0.0034			
14.3								0.0043				0.0005																	
15.0	0.0010	0.2013	0.0507	0.0014	0.0005	0.1670	0.1178	0.1617		0.3209	0.0010	0.0391	0.3045	0.0019	0.0323		0.1404	0.0145				0.0005	0.0024	0.0082	0.0656	0.0376			0.1347
15.2						0.0005	0.0569						0.0005													0.0024			
15.3								0.0415																					
15.4																									0.0010				
16.0		0.1071	0.0405			0.1482	0.0280	0.1424		0.2973	0.0401	0.0010	0.2828		0.0077		0.0487	0.0034					0.0010	0.0019	0.0507	0.0487			0.2302
16.2						0.0005	0.0232														0.0005					0.0019			
16.3								0.0681																		0.0010			
17.0		0.0265	0.1245			0.1332	0.0024	0.0420		0.1366	0.1419	0.0005	0.2042		0.0579		0.0039	0.0005			0.0014			0.0014	0.0516	0.0734			0.2621
17.1			0.0014																										
17.2							0.0039														0.0005					0.0005			
17.3			0.0125					0.1047																		0.0019			
18.0		0.0014	0.2085			0.0878		0.0058		0.0101	0.0705		0.1057		0.1062		0.0019				0.0145				0.0319	0.0956			0.1800
18.1			0.0005												0.0005														
18.2							0.0014														0.0058								
18.3			0.0130					0.0357							0.0005											0.0005			
19.0		0.0014	0.1515			0.0613				0.0024	0.1486		0.0092		0.0980						0.0579				0.0164	0.0931			0.0787
19.1			0.0034																										
19.2																					0.0010					0.0029			
19.3			0.0048					0.0092							0.0010														
19.4																									0.0010				
20.0			0.1264			0.0357					0.1327		0.0010		0.0458						0.0883				0.0116	0.0724			0.0116
20.1			0.0010																										
20.2																										0.0068			
20.3			0.0005												0.0024														
21.0			0.1009			0.0097					0.0666				0.0111						0.1472				0.0043	0.0333			0.0014
21.2						0.0005															0.0019					0.0179			
21.3															0.0101														
22.0			0.0661			0.0087					0.0753				0.0005						0.1974				0.0039	0.0145			
22.2			0.0005																		0.0068					0.0232			
22.3															0.0029						0.0005								
23.0			0.0492			0.0019					0.1182				0.0034						0.1559				0.0034	0.0043			
23.2																					0.0024					0.0280			
23.3															0.0005														
24.0			0.0256			0.0010					0.0970				0.0019						0.1371				0.0005	0.0005			
24.2									0.0005												0.0010					0.0232			
24.3			0.0005																										
25.0			0.0116								0.0835				0.0005						0.1004				0.0019				
25.2									0.0005												0.0005					0.0352			
26.0			0.0034						0.0005		0.0227				0.0005						0.0492					0.0005			
26.2									0.0005																	0.0589			
27.0			0.0024						0.0386		0.0019										0.0198					0.0005			
27.2																										0.0738			
27.3																										0.0005			
28.0						0.0005			0.1646												0.0053								
28.2									0.0005																	0.0642			
28.3																										0.0010			
29.0									0.2042												0.0024								
29.2									0.0014																	0.0468			
29.3									0.0005																				
30.0									0.2476												0.0010					0.0005			
30.2									0.0217												0.0005					0.0376			
30.3									0.0010																				
31.0									0.0801																	0.0014			
31.2									0.0772												0.0005					0.0188			
32.0									0.0140																	0.0005			
32.2									0.0912																	0.0092			
33.0									0.0043																	0.0010			
33.1									0.0014																				
33.2									0.0328																	0.0034			
34.0									0.0024																	0.0039			
34.2									0.0014																	0.0005			
35.0									0.0077																				
36.0									0.0034																	0.0005			
37.0									0.0010																				
38.0									0.0005																				
39.0									0.0005																				
43.2																					0.0005								
H_{obs}	0.7558	0.7819	0.8813	0.7674	0.7761	0.8687	0.8118	0.8890	0.8330	0.7606	0.8793	0.7828	0.7519	0.7307	0.8514	0.7954	0.7992	0.7799	0.6940	0.7210	0.8745	0.7037	0.7761	0.8531	0.8996	0.9353	0.7471	0.6899	0.8060
P_I	0.1054	0.0845	0.0271	0.0768	0.0749	0.0258	0.0559	0.0224	0.0403	0.0921	0.0220	0.0841	0.0915	0.1104	0.0320	0.0727	0.0558	0.0676	0.0972	0.1130	0.0308	0.1334	0.0770	0.0381	0.0146	0.0066	0.0766	0.1355	0.0611
PPE	0.5185	0.5699	0.7597	0.5923	0.5917	0.7624	0.6509	0.7836	0.6966	0.5614	0.7816	0.5739	0.5480	0.5146	0.7352	0.6004	0.6470	0.6085	0.5377	0.5039	0.7423	0.4731	0.5923	0.7144	0.8272	0.8881	0.5804	0.4700	0.6281
HWE	0.4431	0.5144	0.1863	0.0000	0.0709	0.7522	0.1556	0.1281	0.5066	0.0184	0.2078	0.3813	0.6678	0.3272	0.2913	0.4241	0.0656	0.0006	0.0000	0.2675	0.4588	0.0750	0.0075	0.0047	0.2200	0.5163	0.0000	0.0444	0.7644

Appendix D: Cited Figure Use Permissions and Acknowledgements

Figure 1: This chapter was published in John M. Butler, *Fundamentals in Forensic DNA Typing*, Page 23, Copyright Elsevier (2010).

Figure 2: This chapter was published in John M. Butler, *Advanced Topics in Forensic DNA Typing: Interpretation*, Page 40, Copyright Elsevier (2015).

Figure 3: This article was published in U.S. Department of Justice- NRJC Library, Bruce McCord; Kerry Opel; Maribel Funes; Silvia Zoppis; Lee Meadows Jantz, *An Investigation of the Effect of DNA Degradation and Inhibition on PCR Amplification of Single Source and Mixed Forensic Samples*, Page 13, NJRC Open-Source (2011).

References

1. Butler, J. M. (2010). *Fundamentals of forensic DNA typing*. Academic Press.

2. Butler, J. M. (2012). *Advanced Topics in Forensic DNA Typing: Methodology*. Elsevier. https://doi.org/10.1016/C2011-0-04189-3

3. Butler, J. M. (Ed.). (2015). *Advanced topics in forensic DNA typing: Interpretation*. Academic Press.

4. Bright, J.-A., Taylor, D., Curran, J. M., & Buckleton, J. S. (2013). Degradation of forensic DNA profiles. *Australian Journal of Forensic Sciences*, *45*(4), 445–449. https://doi.org/10.1080/00450618.2013.772235

5. Merkley, E. D. (2019). Introduction to Forensic Proteomics. In E. D. Merkley (Ed.), *ACS Symposium Series* (Vol. 1339, pp. 1–8). American Chemical Society. https://doi.org/10.1021/bk-2019-1339.ch001

6. Buckleton, J. S., Bright, J., Gittelson, S., Moretti, T. R., Onorato, A. J., Bieber, F. R., Budowle, B., & Taylor, D. A. (2019). The Probabilistic genotyping software STR mix: Utility and evidence for its validity. *Journal of Forensic Sciences*, *64*(2), 393–405. https://doi.org/10.1111/1556-4029.13898

7. Chatterjee, S. (2019). Saliva as a forensic tool. *Journal of Forensic Dental Sciences*, *11*(1), 1. https://doi.org/10.4103/jfo.jfds_69_18

8. Myers, J. R., & Adkins, W. K. (2008). Comparison of Modern Techniques for Saliva Screening*. *Journal of Forensic Sciences*, *53*(4), 862–867. https://doi.org/10.1111/j.1556-4029.2008.00755.x

9. Elkins, K. M. (2013a). Serology. In *Forensic DNA Biology* (pp. 9–25). Elsevier. https://doi.org/10.1016/B978-0-12-394585-3.00002-X

10. Mullen, C. (2012). Amylase: Phadebas test, saliva. In A. Jamieson & A. Moenssens (Eds.), *Wiley Encyclopedia of Forensic Science* (1st ed.). Wiley. https://doi.org/10.1002/9780470061589.fsa1026

11. Old, J. B., Schweers, B. A., Boonlayangoor, P. W., & Reich, K. A. (2009). Developmental validation of RSID™-Saliva: A Lateral flow immunochromatographic strip test for the forensic detection of saliva. *Journal of Forensic Sciences*, *54*(4), 866–873. https://doi.org/10.1111/j.1556-4029.2009.01055.x

12. Toennes, S. W., Steinmeyer, S., Maurer, H.-J., Moeller, M. R., & Kauert, G. F. (2005). Screening for drugs of abuse in oral fluid—Correlation of analysis results with serum in forensic cases*. *Journal of Analytical Toxicology*, *29*(1), 22–27. https://doi.org/10.1093/jat/29.1.22

13. McCord, B., Opel, K., Funes, M., Zoppis, S., & Jantz, L. M. (n.d.). *An Investigation of the Effect of DNA Degradation and Inhibition on PCR Amplification of Single Source and Mixed Forensic Samples*.

14. Taupin, J. M. (2020). *Interpreting complex forensic DNA evidence*. CRC Press, Taylor & Francis Group.

15. Gill, P., Gusmão, L., Haned, H., Mayr, W. R., Morling, N., Parson, W., Prieto, L., Prinz, M., Schneider, H., Schneider, P. M., & Weir, B. S. (2012). DNA commission of the International Society of Forensic Genetics: Recommendations on the evaluation of STR typing results that may include drop-out and/or drop-in using probabilistic methods. *Forensic Science International: Genetics*, *6*(6), 679–688. https://doi.org/10.1016/j.fsigen.2012.06.002

16. Elkins, K. M. (2013b). DNA Extraction. In *Forensic DNA Biology* (pp. 39–52). Elsevier. https://doi.org/10.1016/B978-0-12-394585-3.00004-3

17. Elkins, K. M. (2013c). Sampling Biological Evidence for DNA Extraction. In *Forensic DNA Biology* (pp. 27–37). Elsevier. https://doi.org/10.1016/B978-0-12-394585-3.00003-1

18. R. Hughes-Stamm, S., J. Ashton, K., & van Daal, A. (2011). Assessment of DNA degradation and the genotyping success of highly degraded samples. *International Journal of Legal Medicine*, *125*(3), 341–348. https://doi.org/10.1007/s00414-010-0455-3

19. Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature*, *362*(6422), 709–715. https://doi.org/10.1038/362709a0

20. Rangé, H., Léger, T., Huchon, C., Ciangura, C., Diallo, D., Poitou, C., Meilhac, O., Bouchard, P., & Chaussain, C. (2012). Salivary proteome modifications associated with periodontitis in obese patients. *Journal of Clinical Periodontology*, *39*(9), 799–806. https://doi.org/10.1111/j.1600-051X.2012.01913.x

21. Rajshekar, M., Tennant, M., & Thejaswini, B. (2014). Salivary biomarkers and their applicability in forensic identification. *Sri Lanka Journal of Forensic Medicine, Science & Law*, *4*(1), 10. https://doi.org/10.4038/sljfmsl.v4i1.6462

22. Kawas, S. A., Rahim, Z. H. A., & Ferguson, D. B. (2012). Potential uses of human salivary protein and peptide analysis in the diagnosis of disease. *Archives of Oral Biology*, *57*(1), 1–9. https://doi.org/10.1016/j.archoralbio.2011.06.013

23. Sivadasan, P., Gupta, M. K., Sathe, G. J., Balakrishnan, L., Palit, P., Gowda, H., Suresh, A., Kuriakose, M. A., & Sirdeshmukh, R. (2015). Human salivary proteome—A resource of potential biomarkers for oral cancer. *Journal of Proteomics*, *127*, 89–95. https://doi.org/10.1016/j.jprot.2015.05.039

24. Siqueira, W. L., Salih, E., Wan, D. L., Helmerhorst, E. J., & Oppenheim, F. G. (2008). Proteome of human minor salivary gland secretion. *Journal of Dental Research*, *87*(5), 445–450. https://doi.org/10.1177/154405910808700508

25. Lee, Y.-H., Zhou, H., Reiss, J. K., Yan, X., Zhang, L., Chia, D., & Wong, D. T. (2011). Direct saliva transcriptome analysis. *Clinical Chemistry*, *57*(9), 1295–1302. https://doi.org/10.1373/clinchem.2010.159210

26. Khatri, P., Sirota, M., & Butte, A. J. (2012). Ten years of pathway analysis: current approaches and outstanding challenges. *PLoS Computational Biology*, *8*(2), e1002375. https://doi.org/10.1371/journal.pcbi.1002375

27. Fan, H., & Chu, J.-Y. (2007). A brief review of short tandem repeat mutation. *Genomics, Proteomics & Bioinformatics*, *5*(1), 7–14. https://doi.org/10.1016/S1672-0229(07)60009-6

28. Cho, H. R., Kim, H. S., Park, J. S., Park, S. C., Kim, K. P., Wood, T. D., & Choi, Y. S. (2017). Construction and characterization of the Korean whole saliva proteome to determine ethnic differences in human saliva proteome. *PLOS ONE*, *12*(7), e0181765. https://doi.org/10.1371/journal.pone.0181765

29. Jain, S., Ahmad, Y., & Bhargava, K. (2018). Salivary proteome patterns of individuals exposed to High Altitude. *Archives of Oral Biology*, *96*, 104–112. https://doi.org/10.1016/j.archoralbio.2018.09.002

30. Thomas, C., & Giulivi, C. (2021). Saliva protein profiling for subject identification and potential medical applications. *Medicine in Omics*, *3*, 100012. https://doi.org/10.1016/j.meomic.2021.100012

31. Smith, H., & Giulivi, C. (2023). Starch treatment improves the salivary proteome for subject identification purposes. *Forensic Science, Medicine and Pathology*. https://doi.org/10.1007/s12024-023-00629-y

32. Hill, C. R., Duewer, D. L., Kline, M. C., Coble, M. D., & Butler, J. M. (2013). U.S. population data for 29 autosomal STR loci. *Forensic Science International: Genetics*, *7*(3), e82–e83. https://doi.org/10.1016/j.fsigen.2012.12.004

33. National Research Council (U.S.), & National Research Council (U.S.) (Eds.). (1996). *The evaluation of forensic DNA evidence*. National Academy Press.

34. Cisana, S., Cerri, N., Bosetti, A., Verzeletti, A., & Cortellini, V. (2017). PowerPlex® Fusion 6C System: Evaluation study for analysis of casework and database samples. *Croatian Medical Journal*, *58*(1), 26–33. https://doi.org/10.3325/cmj.2017.58.26

35. Ambatipudi, K. S., Lu, B., Hagen, F. K., Melvin, J. E., & Yates, J. R. (2009). Quantitative analysis of age specific variation in the abundance of human female parotid salivary proteins. *Journal of Proteome Research*, *8*(11), 5093–5102. https://doi.org/10.1021/pr900478h

36. Bhuptani, D., Kumar, S., Vats, M., & Sagav, R. (2018). Age and gender related changes of salivary total protein levels for forensic application. *The Journal of Forensic Odonto-Stomatology*, *36*(1), 26–33.

37. Kalipatnapu, P., Kelly, R. H., Rao, K. N., & van Thiel, D. H. (1983). Salivary composition: Effects of age and sex. *Acta Medica Portuguesa*, *4*(7–8), 327–330.

38. Marcotte, H., & Lavoie, M. C. (1998). Oral microbial ecology and the role of salivary immunoglobulin A. *Microbiology and Molecular Biology Reviews*, *62*(1), 71–109. https://doi.org/10.1128/MMBR.62.1.71-109.1998

39. Wang, K., Wang, X., Zheng, S., Niu, Y., Zheng, W., Qin, X., Li, Z., Luo, J., Jiang, W., Zhou, X., Li, W., & Zhang, L. (2018). iTRAQ-based quantitative analysis of age-specific variations in salivary proteome of caries-susceptible individuals. *Journal of Translational Medicine*, *16*(1), 293. https://doi.org/10.1186/s12967-018-1669-2

40. Metsalu, T., & Vilo, J. (2015). ClustVis: A web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Research*, *43*(W1), W566–W570. https://doi.org/10.1093/nar/gkv468

41. Karpievitch, Y. V., Nikolic, S. B., Wilson, R., Sharman, J. E., & Edwards, L. M. (2014). Metabolomics Data Normalization with EigenMS. *PLoS ONE*, *9*(12), e116221. https://doi.org/10.1371/journal.pone.0116221

42. Nyamundanda, G., Poudel, P., Patil, Y., & Sadanandam, A. (2017). A novel statistical method to diagnose, quantify and correct batch effects in genomic studies. *Scientific Reports*, *7*(1), 10849. https://doi.org/10.1038/s41598-017-11110-6

43. Zhan, X., Adnan, A., Zhou, Y., Khan, A., Kasim, K., & McNevin, D. (2018). Forensic characterization of 15 autosomal STRs in four populations from Xinjiang, China, and genetic relationships with neighboring populations. *Scientific Reports*, *8*(1), 4673. https://doi.org/10.1038/s41598-018-22975-6

44. Borja, T., Karim, N., Goecker, Z., Salemi, M., Phinney, B., Naeem, M., Rice, R., & Parker, G. (2019). Proteomic genotyping of fingermark donors with genetically variant peptides. *Forensic Science International: Genetics*, *42*, 21–30. https://doi.org/10.1016/j.fsigen.2019.05.005

45. Jarman, K. H., & Merkley, E. D. (2019). The statistical defensibility of forensic proteomics. In E. D. Merkley (Ed.), *ACS Symposium Series* (Vol. 1339, pp. 203–228). American Chemical Society. https://doi.org/10.1021/bk-2019-1339.ch013

46. Jordan, D., & Mills, D. (2021). Past, present, and future of DNA typing for analyzing human and non-human forensic samples. *Frontiers in Ecology and Evolution*, *9*, 646130. https://doi.org/10.3389/fevo.2021.646130

47. Mason, K. E., Anex, D., Grey, T., Hart, B., & Parker, G. (2018). Protein-based forensic identification using genetically variant peptides in human bone. *Forensic Science International*, *288*, 89–96. https://doi.org/10.1016/j.forsciint.2018.04.016

48. Oppenheim, F. G., Salih, E., Siqueira, W. L., Zhang, W., & Helmerhorst, E. J. (2007). Salivary proteome and its genetic polymorphisms. *Annals of the New York Academy of Sciences*, *1098*(1), 22–50. https://doi.org/10.1196/annals.1384.030

49. Parker, G. J., McKiernan, H. E., Legg, K. M., & Goecker, Z. C. (2021). Forensic proteomics. *Forensic Science International: Genetics*, *54*, 102529. https://doi.org/10.1016/j.fsigen.2021.102529

50. The 1000 Genomes Project Consortium, Corresponding authors, Auton, A., Abecasis, G. R., Steering committee, Altshuler, D. M., Durbin, R. M., Abecasis, G. R., Bentley, D. R., Chakravarti, A., Clark, A. G., Donnelly, P., Eichler, E. E., Flicek, P., Gabriel, S. B., Gibbs, R. A., Green, E. D., Hurles, M. E., Knoppers, B. M., … Abecasis, G. R. (2015). A global reference for human genetic variation. *Nature*, *526*(7571), 68–74. https://doi.org/10.1038/nature15393