Forensic Science International: Genetics 54 (2021) 102529

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsigen

Review article

Forensic proteomics

Glendon J. Parker [a,*], Heather E. McKiernan [b], Kevin M. Legg [c], Zachary C. Goecker [a]

[a] Department of Environmental Toxicology, University of California Davis, Davis, CA 95616, USA
[b] Department of Chemistry and Physics, Arcadia University, Glenside, PA 19038, USA
[c] Center of Forensic Science and Research Education, Willow Grove, PA 19090, USA

ARTICLE INFO

Keywords:
Forensic proteomics
Mass spectrometry
Body fluid identification
Proteomic genotyping
Genetically variant peptides

ABSTRACT

Protein is a major component of all biological evidence, often the matrix that embeds other biomolecules such as polynucleotides, lipids, carbohydrates, and small molecules. The proteins in a sample reflect the transcriptional and translational program of the originating cell types. Because of this, proteins can be used to identify body fluids and tissues, as well as convey genetic information in the form of single amino acid polymorphisms, the result of non-synonymous SNPs. This review explores the application and potential of forensic proteomics. The historical role that protein analysis played in the development of forensic science is examined. This review details how innovations in proteomic mass spectrometry have addressed many of the historical limitations of forensic protein science, and how the application of forensic proteomics differs from proteomics in the life sciences. Two more developed applications of forensic proteomics are examined in detail: body fluid and tissue identification, and proteomic genotyping. The review then highlights developing areas of proteomics that have the potential to impact forensic science in the near future: fingermark analysis, species identification, peptide toxicology, proteomic sex estimation, and estimation of post-mortem intervals. Finally, the review highlights some of the newer innovations in proteomics that may drive further development of the field. In addition to potential impact, this review also attempts to evaluate the stage of each application in the development, validation and implementation process. This review is targeted at investigators who are interested in learning about proteomics in a forensic context and expanding the amount of information they can extract from biological evidence.

1. Introduction

The handling, transfer and analysis of biological evidence is integral to forensic science [1]. With a few exceptions the dominating component of evidence is protein. It is the matrix that contains other forensically relevant biomolecules, particularly DNA but other molecules as well [2]. The dissolution of the protein matrix and extraction of these molecules is a necessary precursor to subsequent analysis [3]. For most forensic scientists the role of proteins in forensics is historical and as a component to be removed prior to transfer, purification and analysis of the more polymorphic polynucleotides [4,5].

The protein in evidentiary samples however is highly complex and contains significant amounts of information [6]. Some of this information may provide context about the tissue source of the material, or provide a genetic link to the donor of the material, and therefore be useful to investigators [7,8]. The protein population in a sample results from the transcriptional and translational programs and machinery of the originating cells and is a direct result of the sequence of nucleotides in open reading frames. Protein therefore can be used to identify the tissue or body fluid source of a sample, an important piece of forensic context that is still used today [7,8]. It also contains genetic information in the form of amino acid polymorphisms [9]. Prior to the DNA revolution these were the major source of data for individualization. In addition to being more abundant in a sample, by many orders of magnitude, protein is also considerably more stable than DNA [10–14]. This means that protein may still be present when DNA is no longer accessible to investigators in biologically or environmentally degraded material [15,16].

The last few decades have seen a revolution in proteomics that mirrors that experienced in genomics [6,17]. The advances in proteomic technology, particularly mass spectrometry, have had a major impact on the life sciences and the potential also exists to have an impact on forensic investigation. Many of the early limitations of forensic protein science have now been addressed by these technologies and likewise the amount of identifying and contextual information that can now be extracted from proteomic workflows is considerable [6]. How these proteomic capabilities impact the practice of forensic science is an open

---

* Correspondence to: Environmental Toxicology, University of California Davis, 1 Shields Ave, Davis, CA 95616, USA.
E-mail address: gjparker@ucdavis.edu (G.J. Parker).

https://doi.org/10.1016/j.fsigen.2021.102529
Received 24 September 2020; Received in revised form 6 May 2021; Accepted 7 May 2021

Nomenclature

DDA	Data Dependent Acquisition
DIA	Data Independent Acquisition
ESI	electrospray ionization
FN/FP	false negative/false positive
GVP	genetically variant peptide
LC	liquid chromatography
MALDI	Matrix Assisted Laser Desorption Ionization
MRM	multiple reaction monitoring
MS	mass spectrometry
MS1	mass spectrometry of peptide masses
MS2	mass spectrometry of peptide fragments
m/z	mass over charge ratio
PCR	polymerase chain reaction
PRM	parallel reaction monitoring
PSM	Peptide Spectral Matching
QQQ	Triple quadrupole mass spectrometry
RMP	random match probability
SAP	single amino acid polymorphism
SIL	stable isotope labelled
SNP	single nucleotide polymorphism
TN/TP	true negative/true positive
ZooMS	Zooarchaeology using Mass Spectrometry

question that will be addressed over the coming years and decades.

This review examines the history of protein science in forensics, noting the early use of protein in identification of tissue origin, and individualization. It explores the development of proteomic mass spectrometry and the potential this method has as a tool for the forensic investigator. Protein in forensic evidence has many differences and unique considerations when compared to protein analyzed in other areas of proteomics. As is the case with DNA methodologies, forensic proteomic markers and methods move from a discovery phase, through a development and validation phase and, if successful, an adoption phase by the forensic community. We detail six of these applications. Some mirror earlier uses of protein in forensics, but with greater power,

Table 1

Catalog of distinctively expressed body fluid specific proteins.

Fluid specificity	Accession number	Protein name	Yang *et. al.*	Legg *et. al.*	Van Steendam *et. al.*	Zhoa *et. al.*
Seminal Fluid	Q02383	Semenogelin-2	●	●	●	●
	P04279	Semenogelin-1	●	●	●	●
	P15309	Prostatic acid phosphatase	●	●	●	●
	P07288	Prostate-specific antigen	●	●	●	
	Q6W4X9	Mucin-6	●			
	B2R597	Beta-microseminoprotein	●			
	P09466	Glycodelin	●	●		
	P54107	Cysteine-rich secretory protein 1				
	P61916	Epididymal secretory protein E1				
	P12273	Prolactin-inducible protein				
Saliva	P04745	Alpha-amylase 1	●		●	●
	P02814	Submaxillary gland androgen-regulated protein 3B	●	●		
	Q6P5S2	Protein LEG1 homolog	●			
	P09228	Cystatin-SA	●	●		
	P28325	Cystatin-D		●		
	P01037	Cystatin-SN				●
	P15515	Histatin-1	●	●		
	P02808	Statherin		●		
	Q16651	Prostasin				●
	Q8TAX7	Mucin-7				●
Peripheral Blood	P68871	Hemoglobin subunit beta	●	●	●	●
	P69905	Hemoglobin subunit alpha	●	●	●	●
	P02730	Band 3 anion transport protein	●			
	P02549	Spectrin alpha chain, erythrocytic 1	●			
	P01009	Alpha-1-antitrypsin		●		●
	P10909	Clusterin				●
	P01024	Complement C3		●		●
	P02790	Hemopexin		●		
Vaginal Fluid	Q9UBG3	Cornulin		●	●	●
	Q9Y6R7	IgGFc-binding protein		●		
	P07476	Involucrin		●	●	●
	O95274	Ly6/PLAUR domain-containing protein 3		●		
	P14780	Matrix metalloproteinase-9		●		
	P80188	Neutrophil gelatinase-associated lipocalin		●		
	O60437	Periplakin		●		
	Q6UWP8	Suprabasin		●		
	P08670	Vimentin		●		
	Q9HC84	Mucin-5B		●		
	P35321	Cornifin-A			●	●
Menstrual Fluid	P20810	Calpastatin	●	●		
Urine	P10451	Osteopontin		●		●
	P07911	Uromodulin		●	●	●
	P01042	Kininogen-1				●
	P02760	Protein AMBP			●	●

Proteomic analysis of body fluids (Fluid Specificity) identifies gene products (Protein Name, UniProt Accession Number) that have distinctive levels of expression and are candidates for biomarker development. These proteins are described in Yang et al., Legg et al., van Steendam et al., and Zhoa et al. [154,161,233,236] reliability and sensitivity. Other applications are novel or developed for other purposes with potential impacts in forensics. This review undertakes to delineate where each application is in the development and implementation process and the potential utility and impact of each on forensic practice.

2. Historical uses of protein in forensic science

The identification of body fluids is an essential component for crime scene reconstruction, establishing forensic context, and corroboration [1,7]. Each body fluid has a distinctive combination of proteins and small molecules [18]. Significant scientific effort has gone into identifying and characterizing the distinctive protein components, genetic material and other biomolecules of body fluids, particularly in semen and blood (Table 1) [7,19]. A comprehensive survey by Gaensslen in 1983 summarized these efforts going back to 1827 [20]. These foundational studies used technologies that included early advances in inorganic and organic chemistry, biochemistry and protein science. As new technologies became available, such as protein electrophoresis, alternative light source methods, luminescent and fluorescent chemistry, and histology and immunology, these also were used to characterize protein and identification of forensically relevant body fluids [19–22]. Many of these approaches are still in development today and continue to be refined [19,23–27]. Current techniques for presumptive and confirmatory tests depend on specific protein components of body fluids, such as the presence of the protein co-factor heme for the luminol reaction, or protein specific epitopes for immunological detection of body-fluid specific proteins. These protocols are still essential tools for crime scene investigation today [24,28]. Antibodies to specific proteins have been developed and characterized, both for denatured proteins as well as those with intact secondary and tertiary structures, for almost all forensically relevant body fluids [29]. This has resulted in the development of accessible and affordable sandwich-type immunological detection cassettes and test strips for presumptive identification of body fluids [24,26,27]. Over time the development of specialized antibodies allowed for development of more quantitative ELISA assays that were also developed for forensically relevant body fluid proteins including semen, blood and saliva [30]. We argue below that the accuracy and sensitivity of mass spectrometry now allows for these protein assays to be confirmatory (Section 5).

A major historical tool in the forensic analysis of biological material was the discovery and use of ABO blood types [31,32]. Chemically there are intrinsic advantages in these epitopes, since the carbohydrate structures are stable, antigenic and persist in the environment. Reliable, robust and sensitive immunological assays, both for presumptive and quantitative tests and assays were developed for both body fluids and solid tissue [20,33]. The presence of the two glycosyltransferases responsible for the epitopes, the N-acetylgalactosaminyltransferase for the Type A and galactosyltransferase for Type B, is genetically variant [34]. The result is a set of powerful immunological tools for individualization, exclusion and paternity testing [35,36]. The probability that any individual would have a given variant was dependent on their genetic ancestry [34]. The use of these markers in forensics was widespread until the DNA revolution that began with Sir Alec Jeffreys and expanded and became codified with the adoption of STR-typing methodologies [5,37–39].

Some of the tissue-specific proteins described were also discovered to have genetic variants [20,36]. Advances in chromatography, gel and paper electrophoresis, and enzyme kinetics were used to both identify these variants and develop assays for their detection in forensic samples [40]. The distribution of each variant in the population was used to estimate random match probability in early versions of the product rule [36,41]. These assays however were typically insensitive and consumed limited evidentiary material. Many methods required enzyme activity or intact secondary epitope structure resulting in a lack of sensitivity and false negative detection. Forensic technicians had to choose combinations of methods from multiple possible workflows to gain the maximum amount of forensic information and discriminating power from a sample. These choices required experience and superlative technical ability. The levels of discrimination gained roughly ranged across one to three orders of magnitude [41,42]. In spite of these now modest levels of discrimination, early protein-based assays were used for body fluid identification, individualization or exclusion, and species determination [20,36,42].

The DNA revolution resulted in a dramatic change of focus for protein-based forensic analysis [5,7]. The higher levels of DNA polymorphism, development of PCR, and identification of indel loci that contained multiple highly discriminating alleles made a compelling case for developing and using DNA-based methods [5,43–45]. Another major and perhaps underappreciated result of the shift to DNA-based methods was a simplification and consolidation of sample and analytical workflows. Instead of choosing between multiple protein-based assays, forensic technicians would process a sample and conduct a PCR reaction followed by application onto capillary electrophoresis in order to obtain genotypes on multiple loci from a single sample [5]. While forensic science has justifiably focused on DNA-based methods, many of the initial challenges with protein-based forensics have been addressed by advances in protein science, in particular the development of proteomic mass spectrometry.

3. Basics of proteomic mass spectrometry

Advances in proteomic mass spectrometry have revolutionized, and will continue to revolutionize, cell biology and biochemistry [46]. This is due to major innovations that have solved four fundamental and intrinsic challenges in proteomics. The first is the continued evolution of protein and peptide separation and fractionation by high-pressure liquid chromatography (HPLC). Current chromatography instrumentation can generate higher, more consistent pressures and ultra-low and steady flow rates. Coupled with the availability of smaller column beads with more uniform diameters, chromatography has better performance, higher capacity and more consistent retention times. The advantage extends down to narrow nano-flow capillary columns with internal diameters of 75 μm that result in reduced cross-sectional dilution and increased sensitivity in mass spectrometry [47]. These combine to provide data with higher depths of coverage while consuming less material.

The second innovation, and most fundamental, was invention of electrospray ionization (ESI), published by John Fenn in 1989 [48,49]. When peptides are eluted from the chromatography column, they are atomized in a strong electric field with superheated neutral gases. The acidic solvent sublimates from the droplets conferring a charge on peptides within the droplet. The result is the conversion of a large peptide ion (often about 2000 Daltons) from a liquid to a gas phase, which is the ideal substrate for manipulation using electric fields. The process is chemically gentle and occurs without degrading the peptide, preserving the structure for subsequent analysis. Up to this point it was not possible to volatilize large peptides without introducing intense chemistries that would degrade larger molecules.

The third advance that promoted proteomic mass spectrometry was improved engineering of mass analyzers that manipulate peptide ions and measure them with higher accuracy, precision, frequency and sensitivity with improved detectors. There are several approaches used and platforms available [50]. Quadrupoles select and direct peptides in a given mass range. Electromagnetic fields can then be used to move the ion cloud into different chambers to measure the mass of all peptides using either an ion trap (IT), orbitrap, or time-of-flight mass detector (ToF) [51]. Typically a single peptide is selected and moved to a collision chamber to be fragmented, mainly at the stable but more brittle peptide bonds (Fig. 1) [52]. The pieces can then have their masses measured, resulting in a fragmentation spectrum. The result is highly specific and reproducible; the spectrum pattern is a direct result of the amino acid sequence. The investigator gains two pieces of matched

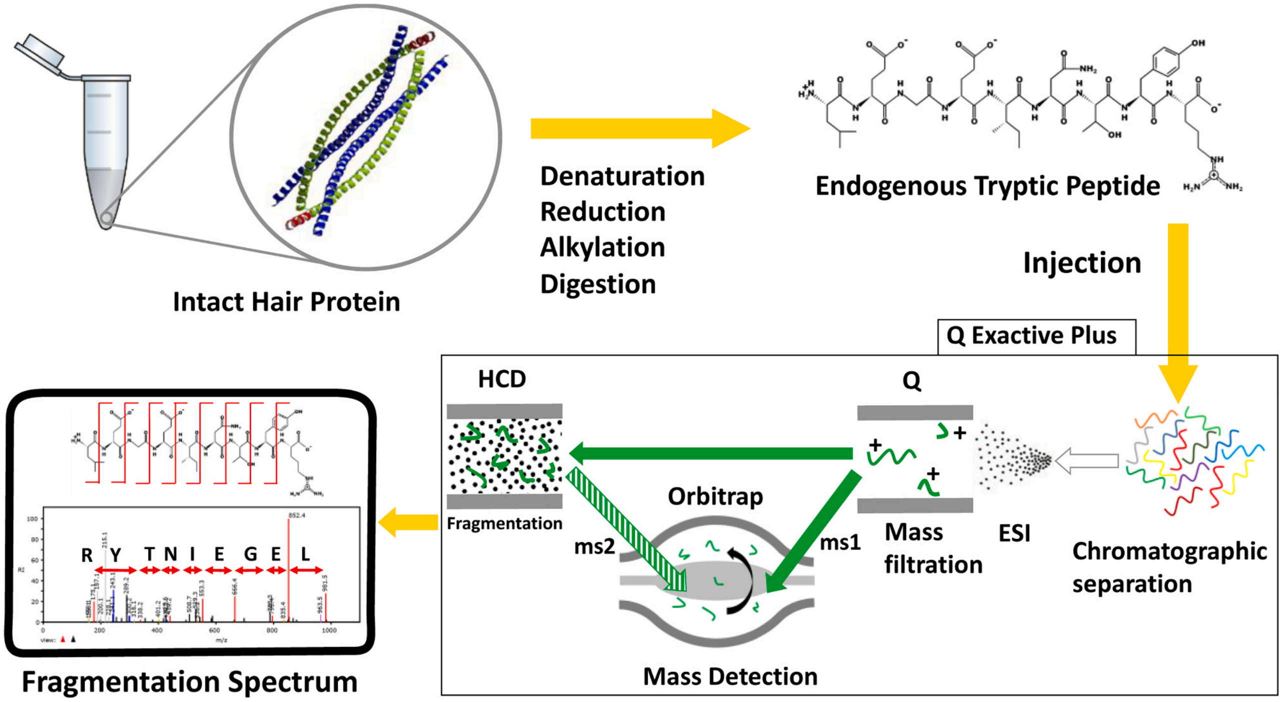

Fig. 1. Shotgun Proteomics. In shotgun proteomic mass spectrometry proteins are digested using trypsin and treated chemically with reduction and alkylation. The resulting complex peptide mixture are resolved using chromatography. As the peptides elute they are volatilized using electrospray ionization (ESI) and then manipulated using mass spectrometry to obtain both the mass of the peptide (MS1) and the mass of the peptide fragments (MS2) by injection into an orbitrap. The spectra of MS2 fragments are then compared to predicted spectra based on a protein database containing expected peptides from the human genome project.

information: a peptide mass and the fragmentation pattern (MS/MS) of the peptide pieces [53]. The MS/MS spectrum also contains several layers of information that can now be accurately predicted using deep learning approaches including which fragment ions are seen (not all are) and the intensity ratio between these fragment ions (spectral angle) [54]. The combination of all of these data points, retention time, peptide mass, fragmentation patterns and fragmentation ratios, is highly specific and, when validated, considered unique for a given peptide.

The fourth innovation was sequencing of the human genome. In addition to accelerating the DNA revolution, this also allowed the complete human proteome to be delineated [55]. While the peptide fragmentation pattern is unique for a given peptide, matching the resulting spectrum to a given amino acid sequence can be highly complex and uncertain when analyzed in isolation. By restricting the number of possible predicted peptide sequences to be present in a sample, the matching process becomes simpler and more confident [56]. In proteomic and bioinformatics terms this reduces the size of the 'search space' [57]. The process is called peptide spectral matching (PSM) [52,58]. Each match has a score that improves with increased alignments between a theoretical mass from a database and a fragmentation mass spectrometry spectrum mass from sample data. Statistical scores can be calculated to estimate the probability and range of uncertainty that a given score would occur randomly. However, the because the relationship is statistical, even spectra with confident sequence assignment will have a formal, if often minute, level of uncertainty and possibility of mismatch. Peptide spectral matches consequently require additional validation as described below.

These innovations combined to create a revolution in proteomics. The sequence of innovations described above is broadly described as shotgun proteomics, 'bottom-up' proteomics, or data-dependent acquisition, and it accounts for the bulk of proteomic studies in the life sciences (Fig. 1). The common orbitrap platform produced by ThermoFisher, a Q-Exactive mass spectrometer, can detect 10 peptides per second over a 2 h gradient using 2 μg of HeLa protein to generate 48,000 spectra and detect 27,000 unique sequences from 3900 proteins in a single run [59]. The latest generation instrument, a Bruker timsToF mass spectrometer, can detect 100 spectra per second using 200 ng of HeLa protein to generate 600,000 spectra and detect 32,000 unique sequences from 5700 proteins [60]. These are highly complex, rich datasets that provide a comprehensive survey of a protein sample. These values increase with deep proteome mapping that uses combined runs or pre-fractionation of any samples [61,62]. The premise of forensic proteomics is that information on this scope and scale will be useful to investigators.

The goal of shotgun proteomics is to achieve the highest depth of coverage of the proteome and identify as many proteins and peptides as possible with limited a priori expectations for sample composition. Conceptually it is similar to large scale shotgun DNA sequencing using untargeted, unbiased DNA libraries. There are limitations to this approach however. There are millions of different peptides in a trypsin digest, from over 10,000 genes that are expressed in any tissue, and that have abundances that range over seven orders of magnitude [62,63]. It is unavoidable that there will be run-to-run variation. Selection of peptides for fragmentation is highly stochastic in shotgun proteomics, and some peptides may not be detected in every run [64]. Mass detection is concentration dependent and scanner acquisition is rate limited. This has significant implications for forensic proteomics, since the lack of detection cannot be taken as evidence of absence. This analytical challenge has many similarities with genetic analysis of low-read, low coverage genomes [65,66]. There are several innovations and data acquisition approaches to address the problem of stochasticity. One approach increasing in popularity is Data-Independent Acquisition (DIA) where fragment ion (MS2) spectra are continuously acquired in an unbiased fashion [67]. DIA has the advantage of collecting fragmentation data that allows for subsequent analysis of previously collected raw data. Analysis is currently challenging because it is difficult to match fragmentation data with the broader window of possible primary peptide masses. Improved bioinformatic tools are being continually developed [54,67–69].

The most effective approach to deal with the problem of stochastic effects is to use targeted data acquisition that focuses on a known set of highly characterized peptides [70]. This focuses the resources of the instrument on a limited set of targets and results in dramatically increased sensitivity. Targeted acquisition can be carried out on a variety of platforms including the common "discovery" based systems, such as Orbitrap mass spectrometers operating in Parallel Reaction Mode (PRM), as well as dedicated targeted mass spectrometers, such as the triple quadrupole systems (QQQ) operating in Multiple Reaction Monitoring (MRM) mode as discussed below (Fig. 2). When operating in PRM-mode an Orbitrap has a predefined schedule of peptide masses that are targeted at scheduled retention times [71–73]. During data acquisition, the targeted peptides are specifically selected in a quadrupole mass filter, transferred and accumulated in the instrument, and then fragmented with the resulting full scan fragment ions being measured in the Orbitrap [71]. PRM is flexible, easy to set up, and is highly sensitive [71–73]. It is limited however by the variable retention times in capillary liquid chromatography and the expense of obtaining a proteomics-grade instrument.

The most popular platform for targeted proteomics is the triple quadrupole mass spectrometer (QQQ) that is used in many clinical and forensic toxicology laboratories. In this configuration peptides are isolated based on *m/z* values in quadrupole 1, fragmented in quadrupole 2, and specific fragment ions selected for detection in quadrupole 3 (Fig. 2). Both targeted mass spectrometry approaches are quantitative. In the case of MRM the peak area of the targeted peptide pieces can be measured against a stable isotope labelled (SIL) peptide standard, or previously run standard curve run in the same sample sequence. An SIL peptide in the sample will have the identical retention time, peptide mass, fragmentation masses (transitions) and ratio of intensities as the known synthetic standard and these can be directly compared using an extracted ion chromatogram. In analytical chemistry and forensic toxicology when these measurements are matched the analyst can conclude that the endogenous peptide is in the sample [74,75]. Using pre-defined retention time windows and precursor and fragment mass information, these targeted assays may be developed to simultaneously and reproducibly detect hundreds of peptides in a single run, with high sensitivity, precision and accuracy [50,70,76,77]. At a practical level QQQ mass spectrometers have lower running costs, daily maintenance is more straightforward, and they are easier to use [76,78,79]. Modern QQQ mass spectrometers are capable of attomole-level (10^{-18}) limits of detection at approximately half the cost in capital investment and significantly lower maintenance costs [80,81]. Because each run typically occurs over a 5–15-minute gradient, instead of the 60–90-minute gradients on proteomics-grade instruments, sample costs are much lower. Finally, QQQ instruments are broadly accessible to forensic scientists since they are commonly used in forensic toxicology as well as clinical laboratories.

In an integrated research program forensic proteomics requires the use of both targeted and untargeted mass spectrometry approaches. Discovery of forensically relevant peptides requires an untargeted bottom-up, shotgun proteomic, survey of the protein population in a tissue in order to identify, characterize and validate relevant peptides. The subsequent analysis of these peptides in a forensic laboratory setting requires the development of more robust, sensitive and cheaper targeted mass spectrometry protocols on triple quadrupole (MRM) or high precision proteomic-grade mass spectrometry systems (PRM).

The development of shotgun and targeted proteomic mass spectrometry solved a major problem for protein-based forensic analysis; the need to have multiple tests and assays for multiple markers that consumed highly limited forensic material. Workflows have simplified and coalesced into preparing a sample and applying it to mass spectrometry [82]. This parallels the effect of PCR on DNA based methods, where a single analysis can result in a genotype from many loci. Current proteomic methods are considerably more systematic, easier to implement, and more powerful in terms of information gained. The primary objectives of forensic proteomics however remain the same as earlier uses of forensic protein science: identifying the tissue source of the material, analyzing identifying genetic information, and species

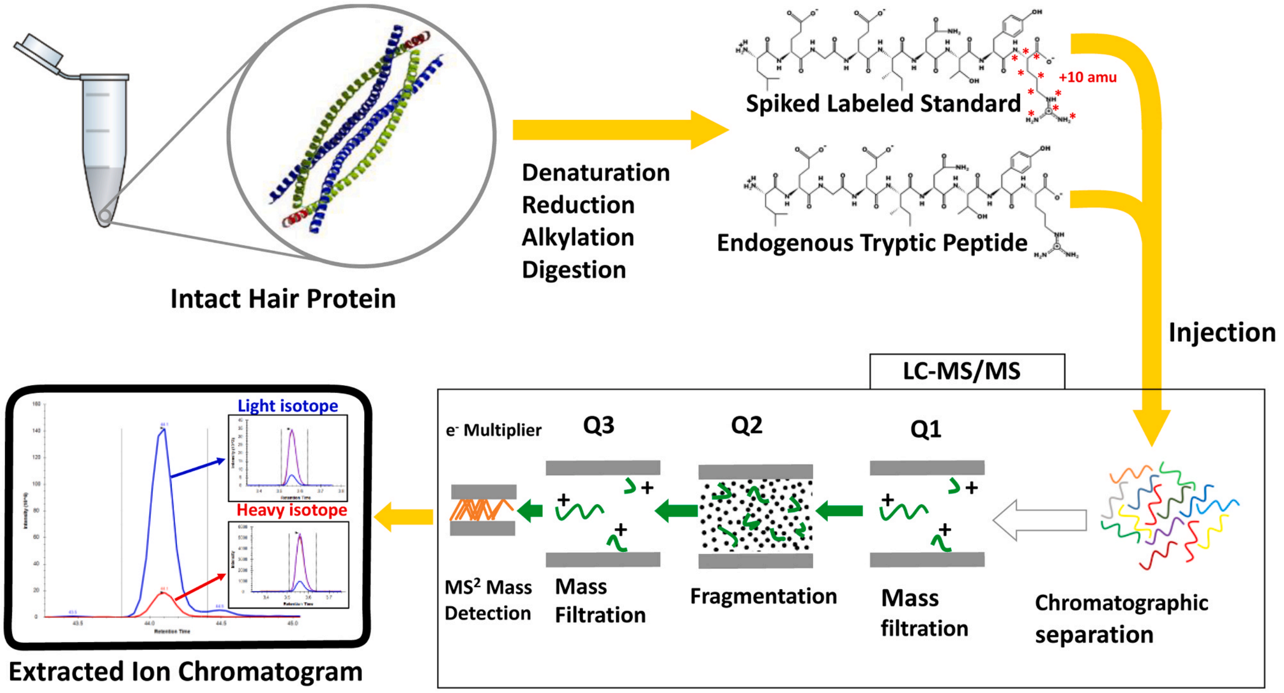

Fig. 2. Targeted Proteomics. In targeted mass spectrometry using a triple quadrupole mass spectrometer, proteins are digested using trypsin and treated chemically with reduction and alkylation. The resulting complex peptide mixture are resolved using chromatography. As the peptides elute at known retention times, they are volatilized using electrospray ionization (ESI), and the target peptide is filtered using a precise mass window using the first quadrupole (Q1), fragmented through collision induced dissociation in a second quadrupole (Q2), and two or three targeted fragments measured using the final quadrupole (Q3). Peptide validation occurs through simultaneous analysis of matching stable isotope labeled synthetic target peptides.

determination [7].

4. Proteomics of forensic samples

Protein is intrinsically more stable than DNA. The DNA backbone is rich in slightly reactive electrophilic oxygen and nucleophilic amines [16]. The amines and oxygen in the peptide backbone however, are adjacent across the amide peptide bond and therefore less reactive [2]. The persistence of protein in a forensic sample makes proteomic mass spectrometry an ideal tool for analysis of compromised samples. This is demonstrated by the use of paleoproteomics for archaeological species identification [12,83,84]. The oldest DNA on record is from permafrost mammoth samples that are over a million years old [85]. There are several paleoproteomic dates beyond that point [86–89], with the current oldest date being struthiocalcin protein detected in a 3.8 million year old Ostrich egg [90,91]. While controversial, there are even claims of collagen detected from even deeper time frames in *Tyrannosaurus rex, Brachylophosaurus canadensis* fossils [92–98]. The ability of paleoproteomics to discriminate between different species is dependent on extracting genetically relevant information from highly compromised protein samples. The same properties make protein an ideal substrate for analysis of forensic samples as well.

Before proteomics can be widely applied to forensic contexts, practitioners need to first recognize and account for the unique biological, chemical and analytical contexts, as well as legal requirements, of processing material from a crime scene. A primary difference in forensic proteomics is the type of sample analyzed: forensic contexts result in degraded samples that are variable and often highly limited. As with DNA-based methods, forensic proteomic protocols need to accommodate large variation in the amount of material available for analysis. Proteomic samples in forensic science will tend to be solid substrates, such as hair or bone, or dried, such as body fluid or touch samples from skin or clothing. Forensic proteomic methods should be compatible with other methodologies such as DNA-typing, particularly since they are destructive. In addition, identified peptide markers need to be thoroughly validated to withstand legal and scientific scrutiny.

4.1. The forensic environment and protein structure

Biological samples located at a crime scene will either be removed from the body, be from a deceased individual, or will be in high variable, uncontrolled environments. These may include higher temperatures, availability of water, oxidative chemistry, extremes of pH, and endogenous and environmental small molecules and metabolites [1]. The repair mechanisms available in living cells and organisms will not be available to reverse the accumulation of denaturation and chemical modification to protein that accumulate over time prior to collection. The amount of intact, correctly shaped proteins is reduced. As discussed above, this compromises the sensitivity of body fluid detection and reduced the sensitivity of many, now obsolete, protein-based assays of genetic variation. Many chemistries are intrinsic and promoted by protein movement and ambient temperature. The most characterized and commonly detected of these is deamidation, where the side chains of glutamine and asparagine react with the peptide bond and convert to glutamic or aspartic acid or their gamma derivatives [99–102]. Another process is racemization of amino acids from L- into D-enantiomers [103–105]. Common environmental reactants, such as oxygen, can modify sulfur containing amino acid side chains, particularly methionine and cysteine [106–109]. These modifications are, for the most part, highly predictable and easily incorporated into search algorithms. Other non-enzymatic chemistries are less predictive. Maillard reactions including Schiff Base formation, Amadorri rearrangements and aldol condensation reactions are highly diverse and become difficult to incorporate into search algorithms [110–112]. Eventually, these and other hydrolytic processes begin to cleave the peptide backbone. Maillard reaction will increase cross-linking within and between polypeptides, polynucleotides, fats and carbohydrates [110,111,113–116]. Over time these two processes increase the diversity of peptide masses, or peptidoforms, and reduce the level of unmodified peptides available for analysis [117].

4.2. Forensic sample preparation

Most sample preparation for proteomics follows a standardized workflow: protein extraction, proteolytic digestion and sample cleanup. This mirrors the workflows in DNA sample preparation[3]. Sample preparation may include rehydration, tissue homogenization, cell lysis, sample solubilization and centrifugation. The protein matrix needs to be denatured so that proteases may enter and digest the exposed peptide backbones within [118,119]. This occurs by the use of chaotropic agents, such as urea or guanidine, and strong detergents. The disulfide bonds in the matrix are opened up with reductants and capped by alkylation so that cysteines cannot reform randomly and reform the matrix [82,120]. The most common protease used is trypsin that cleaves the peptide backbone on the C-terminal side of the positively charged arginines and lysines [121–123]. The resulting peptides typically are 8–15 residues long and have a positive charge at both ends of the peptide, both ideal properties for chromatography, ionization, and manipulation and fragmentation in mass spectrometry [60,121,124]. Intact proteins are large and accumulate many chemical and biological modifications, which complicate mass spectral analysis. Consistent predictable digestion with a protease however will limit mass variation, charge state variations, and produce good chromatography. Uniform proteolysis will maximize the potential that any one peptide form will be sufficiently concentrated to be detected. Following digestion, samples can be cleaned up and desalted prior to mass spectral analysis using solid-phase extraction [125]. Targeted protein or peptide capture with immobilized antibodies may also be used to enrich a sample prior to analysis, concentrating biomarker targets and mitigating the effects of complex matrices [126–128]. As discussed below (Section 8) some innovative work with magnetic beads and microwaves may further advance sample processing and analysis [129,130]. Alternatively, abundant proteins that would normally dominate the proteomic data can be depleted using commercially available immobilized antibodies [131,132]. Different substrates, such as bone, teeth or hair and skin may require specialized deviations from these standard protocols in order to accommodate the mineral content of bone [118,129,133–135] or the high sulfur content and level of disulfide bonds in keratinized hair, nails and skin [9,136–139]. In spite of these variations, these methods are straightforward and automatable, and allow for development of routine high-throughput operations favored by forensic laboratories.

4.3. Multi-omic sample preparation

The parallels between protein and DNA sample processing raise the possibility of obtaining both nucleotide and proteomic information from the same sample and processing workflow [140–143]. As mentioned above, both proteomic and DNA-based sample preparation follow a common pattern: biomolecule extraction, proteolytic digestion and sample cleanup. However, there are currently steps in both workflows that are incompatible for the other type of biomolecule. For example, proteinase K is not compatible with proteomics since its specificity is too broad, the number of peptides too variable, and peptide length too short. Likewise, the detergent sodium dodecyl sulfate is difficult to remove from a sample and is incompatible with mass spectrometry instrumentation [119]. Iodoacetamide alkylation in proteomics releases iodine that has oxidative chemistry that may react with, and therefore degrade, polynucleotides [120,140,141,144,145]. These reagents can be easily substituted however. Proteinase K can be substituted by trypsin and LysC enzymes with no loss of DNA yields [140–142]. Chloroacetamide, that also may be used at lower concentrations, can be a substitute for iodoacetamide [82]. Sodium dodecyl sulfate can be substituted with sodium dodecanoate, or even acid labile surfactants [136,137,146–148]. Magnetic beads may make workflows more compatible [130,149,150]. The end result of respective workflows, purified polynucleotides or peptides, have significantly different chemical properties. A double stranded polynucleotide that is 100 base pairs long is about 66 thousand Daltons and is highly anionic. The typical peptide length, ranging from 8–15 amino acids, is about 900–1700 Daltons and is highly diverse in terms of ionic and hydrophobic properties. Separation of these two types of substrate should be achievable with minimum loss of either polynucleotide- or protein-based information [140–142].

4.4. Forensic proteomic tissues and body fluids

Forensic proteomics also differs from other proteomic methods by the types of sample that are typically analyzed. Body fluids are often easily accessible and as such have been used by proteomic specialists in the life sciences to demonstrate advances in mass spectrometry instrumentation, methodology and analysis [151]. This has resulted in an ultra-deep analysis of the salivary proteome and a demonstration that the salivary proteome is dynamic in response to smoking and daily fluctuations in the microbiome [152,153]. The proteomic composition of semen and vaginal fluid is also relevant for clinical as well as forensic applications [154–165]. The same is also true for sweat (Section 7.1) [166,167] and vomit [168]. Forensically relevant solid tissues such as bone [130,134,135,169–173], teeth (Section 7.4) [174–185], hair (Section 6.6) [9,186–192], nail plate [139], and skin cells (Section 7.1) [148,193,194], have also been thoroughly processed and analyzed for both forensic as well as clinical purposes. Mineralized and keratinized tissues have also been examined from a paleontological, paleoproteomic perspective that also has forensic relevance for analysis of highly degraded human remains [9,13,130,185,195–207].

4.5. Validation in forensic proteomics

A basic question in proteomics is whether the spectra or other signals obtained through mass spectrometry are caused by peptides with the claimed amino acid sequences. The assignment of a peptide sequence with a fragmentation spectrum is intrinsically a statistical match that results in a possibility, if only a minute one, of false assignment. Increased confidence in peptide assignment can be achieved by increasing the stringency of peptide spectral matches [66]. These include, but are not limited to, ensuring that only tryptic peptides are analyzed, that the monoisotopic mass of the observed peptide must be close to the mass of the theoretical peptide, to only analyzing spectra above a given threshold signal level, and restricting the fragmentation spectra analyzed to only peptides with 2 or 3 charges [66].

There are three basic strategies to validate that the reported sequence of a peptide is the correct one. The first approach is to internally and statistically validate the efficacy of the peptide spectral matching process in shotgun proteomic datasets. This occurs through the use of decoy databases that are comprised of an equivalently sized database of incorrect peptide sequences that are known to not occur in the human proteome [208,209]. This approach assumes that incorrect matches to the decoy, or false, reference proteome database will occur at the same rate as incorrect matches to the human reference proteome database [208]. The occurrence of false decoy assignment will increase for poorer quality spectra. Peptide spectral matches are graded on a continuum of excellent to poor assignment scores based on the degree of alignment between theoretical and actual spectra. There are many variations on the theme and different scoring schemes [58,208–213]. Nevertheless, these scores allow us to create a gradient of low to high quality spectral assignments, with low quality assignments having a greater proportion of decoy assignments [208]. In proteomics the consensus is that a 1% level of false peptide assignment is acceptable, and the statistical quality score at which this occurs is the threshold for excluding poorer quality scoring spectra from the analysis.

The second validation approach is to compare observed peptide masses, and mass spectra with an assigned peptide sequence, with known standards; synthetic peptides that will have identical chemical properties to the putative endogenous peptide in the forensic sample. Described above (Section 3), these standards have the same chromatography and chemical metrics, except maybe for mass, as the endogenous peptide in the sample [126,214,215]. Adding an identical synthetic peptide to the sample that contains an endogenous peptide and observing the resulting changes in intensity, known as standard addition, is common even though it obscures the endogenous peptide signal [216]. In targeted mass spectrometry that use analytical columns, retention time is very consistent and so standards can be added in separate runs. In shotgun proteomic platforms this is not a good option because retention times when using nanocapillary columns tend to vary and heavier stable isotope-labelled (SIL) peptides should be added to the sample [126,214,215]. Because of the heavier mass the standards, while behaving identically to endogenous peptides, will not interfere with mass detection. This form of validation is identical in concept to currently used and accepted approaches in forensic and other forms of toxicology [74,75]. A final use of standard peptides is to systematically compare the spectra of known peptides with spectra from within a sample [217,218]. This approach exploits the spectral libraries of peptides generated by the National Institute of Standards and Technology or Global Proteome Machine [217,219]. The use of spectral libraries in this manner is a major tool for identifying and validating unknown compounds in forensic toxicology [220,221].

The final validation used in forensic proteomics is genetic [9]. Genetically variant peptides, the result of non-synonymous SNPs, are used to infer the presence of SNP alleles in a donor's genome. Peptide spectral matching for this small but informative subset of peptides can be challenging because most of the sequence is identical to the corresponding peptide derived from the other allele, many fragments will be the same. Also, many single amino acid polymorphisms may have the similar mass shifts as chemically variant peptides. Without a high level of characterization and use of filtering criteria discussed below (Section 6.3) many peptide sequences containing single amino acid polymorphisms will be false positives. Genetic inference however is easily tested by direct DNA sequencing of the predicted inferred SNP allele and the performance of individual GVP inference can be measured [9]. This mode of validation is unique to genetically variant peptides; sequencing invariant DNA is not informative.

Validation in forensic contexts also requires additional measures. The Scientific Working Group on DNA Analysis Methods (SWGDAM) and FBI Quality Assurance Standards guidance require that forensic methods should conducted on samples from foreseeable real-world scenarios [222,223]. An example of these additional validation measures is described below in Section 6.6.

The remainder of the review will focus on applications of forensic proteomics: body fluid and tissue identification and proteomic genotyping, as well as the developing methods of post-mortem interval estimation, species identification, peptide toxicology, sex estimation, and fingermark analysis. Each area will show a progression from discovery of relevant biomarkers to forensic applications. Forensic application generally requires a focus on quantitative methodology, efforts to demonstrate reliability, robustness, replication and reproducibility and to be thoroughly validated on real world samples [222]. In order for forensic proteomics to be practical for investigators the methods must provide useful confirmatory information and investigative leads as well as be accessible in terms of cost, ease of use, and access to the appropriate mass spectrometry platforms.

5. Proteomic body fluid and tissue identification

The classification of biological fluids of forensic interest (i.e. blood, semen, saliva, vaginal/menstrual fluid, etc.) found in relation to a crime is referred to as forensic serology. Biological fluid detection and identification provides important contextual information to a forensic investigation. While genetic testing can help establish from whom DNA may have come from, only serological testing can provide an indication of the body fluid or tissue from which a DNA profile may have originated [5]. The ability to obtain the most probative value from a biological stain in a criminal investigation, requires both the development of an interpretable DNA profile and the identification of the biological substance from which the profile originated.

The proteins present in a biological sample provide forensic contextual information. Transcriptional programs are tissue, or body fluid, specific. Proteomes consequently reflect tissue structure and physiological function. The gene products present in a sample, and their relative amounts, can therefore be used to infer the tissue origin of a sample. Contemporary forensic tools for the identification of biological fluids, however, are based on the same fundamental methods that have been employed for much of the history of forensic science. Namely, these are chemical reactions with components of a body fluid; detection of enzymatic activity characteristic of a body fluid; or, in the case of semen, direct visualization of spermatozoa by microscopy [20,224]. While these techniques have value to forensic investigations, they also suffer from a variety of substantial test-specific limitations, most notable are limitations associated with specificity. Operational forensic laboratories still performing serological analyses have almost uniformly shifted from relying heavily on enzymatic and colorimetric assays in favor of immunological reactions in the form of immunochromatographic assays. While sensitive, fast, and easy to use, these assays can be costly and suffer from a lack of body fluid specificity. Target protein biomarkers present at lower concentrations in other biological fluids can also generate a positive reaction [29]. For example, the sensitivity of the ABAcard HemaTrace® can be as low as 0.07 μg/mL of hemoglobin, making the assay more sensitive than chemical color reactions for blood. As a result, however, the test has been shown to produce positive results with seminal fluid stains, and oral, vaginal, anal, and rectal swabs [225]. At this level of sensitivity, it is thought that the assay is detecting very low amounts of hemoglobin in these non-target fluids; thereby generating false positive reactions in regard to the biological fluid being targeted. However, unpredictable cross-reactivity with non-target molecules having similar conformational epitopes is also possible as are non-specific binding events due to extremes of pH, the presence of organic acids or other sample-specific chemical compounds [226,227]. Environmental contaminants also have the possibility of interfering with antibody binding [29]. Degraded samples will not work with these assays due to loss of conformational integrity of the target protein. Additionally, at high concentrations of target antigen, these tests suffer from a phenomenon called the high-dose hook effect, leading to false negative reactions when the target antigen is in excess [228]. Finally, an inability to target multiple fluids within one assay leads to additional sample consumption and low throughput creating a bottleneck in the laboratory. In contrast to traditional methods, proteomic body fluid identification allows for the multiplexing of multiple fluid detection within a single assay, requiring significantly less sample consumption as compared to traditional serological approaches and most crucially, provides a confirmatory identification of a target fluid.

5.1. *Discovery and validation of body fluid and tissue biomarkers*

The proteomics research and development pipeline are characterized by a discovery phase aimed at identifying candidate biomarkers of interest, followed by a verification phase to assess their forensic utility. Following candidate verification, bioanalytical method validation must take place to assess the reliability and performance limits of the analytical approach. The initial discovery phase requires an extensive analysis of the target tissues, usually body fluids or other comparison tissues [70]. As previously stated, proteomes can be highly complex and with up to 10,000 gene products being expressed in any one cell type, many distinctive proteins that are apparently unique in one tissue may be present at low levels in other tissues [62,229,230]. For example, PSA can be detected in other body fluids including at minute but detectable levels in females [231]. Ideally biomarker proteins are unique for a given tissue or body fluid. An ideal protein will also be abundant and not vary greatly as a result of individual genetic background or physiology. Peptides from abundant and tissue-specific proteins are a good starting point for biomarker discovery [70]. However, forensically useful information may also be expressed as ratios and relative abundances in comparison with other tissues and body fluids [154,232]. A number of untargeted proteomics studies have led to an extensive, somewhat overlapping list of biomarker targets for forensically relevant biological fluids including peripheral blood, vaginal/menstrual fluid, semen, urine, nasal secretion, vomit and saliva [154,161,162,165,168]. Overall, 44 fluid specific biomarkers were identified via several approaches (Table 1) [154,161,165,233].

While many protein biomarkers have been proposed for forensic use, not all peptides from the same protein are equal. Some peptides digest more readily, and other peptides ionize more efficiently in ESI. Other peptides may be from an area of the protein that may degrade faster in the environment or be subject to other endogenous protease activities. If a peptide is from an area of a biomarker protein that is environmentally modified, then that will change the mass of the peptide creating a chemical variant and proportionally reduce the abundance of the target peptide [136]. In order to establish the forensic utility of candidate protein biomarkers to correctly infer a biological fluid, a series of verification or validation steps must be conducted to establish the specificity and reliability of each peptide consistent with SWGDAM and SWGTOX/OSAC guidelines [74,222]. Studies designed to distinguish two closely related biological fluids, namely peripheral and menstrual blood have been undertaken [234,235]. Population studies have been conducted to demonstrate the specificity, detection consistency, and relative interindividual variability in protein abundance of biomarker targets in a cohort of 50 individuals [236]. Results of these studies have shown that multiple fluid specific protein markers are both specific to their intended fluid as well as consistently detected across a large population. Finally, several of these studies successfully identified the biomarkers of interest in a wide array of casework-type samples.

5.2. *Targeted proteomic body fluid and tissue identification*

Once candidate protein biomarkers have been identified and verified, targeted assays can facilitate the detection and quantitation of even low abundance biomarker protein targets against a background of other non-target molecules in complex biological matrices. To date, all targeted assays have utilized QQQ mass spectrometers operating in MRM mode [237]. Studies evaluating targeted assay performance for forensic biological fluid identification have been reported using casework-type samples as well as authentic casework samples [236,238,239]. In one such application, suspected bloodstains from multiple crime scenes were tested both with immunochromatographic assays as well as with a targeted proteomic method. The targeted proteomic assay demonstrated greater sensitivity than traditional immunological test strips, especially when analyzing degraded and compromised samples. In another authentic application, researchers were able to couple machine-learning algorithms with novel proteomic biotargets to accurately predict organ tissue origin from various bovine samples. While successful in test cases, this assay was only partially successful in analyzing a bullet from a homicide investigation in order to predict the path it took through a victim [240]. While interpretation using machine-learning in this case was complicated by the fact that a bullet may pass through several organs in a body, the ability to detect proteomic makers from such a challenging trace sample shows promise for future applications [236,240]. These studies have shown that targeted proteomic body fluid identification methods are rigorous, sensitive and amenable to challenging casework samples. Additional studies have demonstrated that targeted proteomic workflows are fully compatible with downstream genetic testing [140–142].

6. Proteomic genotyping

Protein contains genetic information in the form of single amino acid polymorphisms (SAPs), the result of non-synonymous SNPs [43, 241–243]. Proteomic genotyping is the detection of genetically variant peptides (GVPs) that contain SAPs and then inferring the presence of the corresponding SNP alleles in the genome of the individual who produced the protein sample (Fig. 3). In aggregate the resulting profile of inferred SNP alleles, like any profile of nucleotide variation, can be used to calculate the statistical association of an individual and a protein sample; the probability that a given combination of inferred SNP alleles would randomly occur in the reference population [9]. Currently almost 500 non-synonymous SNP alleles have been shown to be accurately inferred by GVPs in the literature [9,136,142,148,186,244]. GVPs will be present in any protein sample. The samples where proteomic genotyping provides benefit for a forensic investigator are those where obtaining DNA-based information is problematic. This includes telogenic hair shafts, degraded and compromised bones and teeth, fingermarks, and sexual assault evidence. The method however can be used for any protein source [9,242].

6.1. Genetic information in the proteome

The amount of genetic information in a proteomic sample can be considerable. At the individual level a typical genome contains 3.5–4.3 million SNPs, depending on their genetic background [43]. A subset of these variants, non-synonymous SNPs, change the codon-assignment in an open reading frame, resulting in a single amino acid polymorphism (SAP) that occurs in a genetically variant peptide (GVP) (Fig. 3) [9,245]. Like the originating SNP alleles these peptides occur at a given frequencies in the population. There are between 10.2 and 12.2 thousand non-synonymous SNPs per genome, depending on ancestral background [43], roughly a common variant locus in every second gene product. Most of these variants are commonly distributed across the population, with about 70% having a minor allele frequency greater than 5% [44]. While this is the 'snapshot' that occurs within an individual, the scope of potential non-synonymous SNPs is much larger. At the time of writing, the dbSNP database (ncbi.nlm.nih.gov) contained 720 million SNPs of which 7.9 million are missense SNPs. By far most of these variants are rare and even private [43]. However, as many as 122 thousand had a minor allele frequency greater than 0.5%, the threshold used in proteomic genotyping [9], and 69 thousand were common with global minor allele frequencies above 5%, the rough equivalent of 3.5 common GVPs per protein in the population. Given a typical proteome of 400 proteins, such as in the hair shaft proteome, an investigator could anticipate 1400 GVP-inferred common loci that would be useful for human identification [136]. This value will decrease as samples degrade and detectable proteomes become less complex, and the value will increase as sample processing improves and mass spectrometers and bioinformatic algorithms become more sensitive. The potential for increase is high. In living human cells about 10–11 thousand genes are expressed, which is about 50% of the total number of available genes [55]. Much higher than the 400 most detectable proteins. The level of protein expression varies widely, with a dynamic range across many orders of magnitude from less than 1000–20 million per cell, with a median value of around 300 thousand copies per cell, the equivalent of 18 cycles of PCR [62,246]. The premise of proteomic genotyping is that this level of genetic information will be informative to the forensic investigator.

6.2. Discovery of genetically variant peptides

Discovery to this point has focused on discovery of minor allelic peptides and has been limited by the sample size and populations examined for GVPs. The largest cohort published to date was 66 European American subjects [9]. GVP discovery has predominantly occurred from the most abundant proteins and those with higher minor allele frequencies. As samples from more individuals are examined, and at greater sensitivity, GVP discovery will increase. The discovery process to this point has primarily depended on empirical analysis of peptide spectra with mass shifts consistent with amino acid substitution. This 'bottom up' approach focuses on highly performing peptides that ionize well and come from proteins with a high copy number per cell. It has been the major mode of GVP discovery to this point.

A systematic and simpler method for discovery is to identify non-synonymous SNP alleles through prior exome or genomic sequencing [148,247]. In this 'top-down' discovery approach, potential alleles are identified in genomic datasets, and the presence of predicted GVPs confirmed in a matching proteome, using a customized reference protein database. This discovery process is not limited by peptide ionization efficiency or protein abundance. This process also has the potential to identify rare, sometimes private GVPs that may be unique to an individual [148,247]. This approach raises the possibility that very rare, private GVPs may be used to resolve an individual contribution from a complex mixture [148,244]. This individualization strategy raises significant statistical and genomic challenges. Validation for this 'top-down' approach is different, since the same set of matching proteomic and genomic data cannot be used for both discovery and validation. This could easily be addressed through proteomically processing another

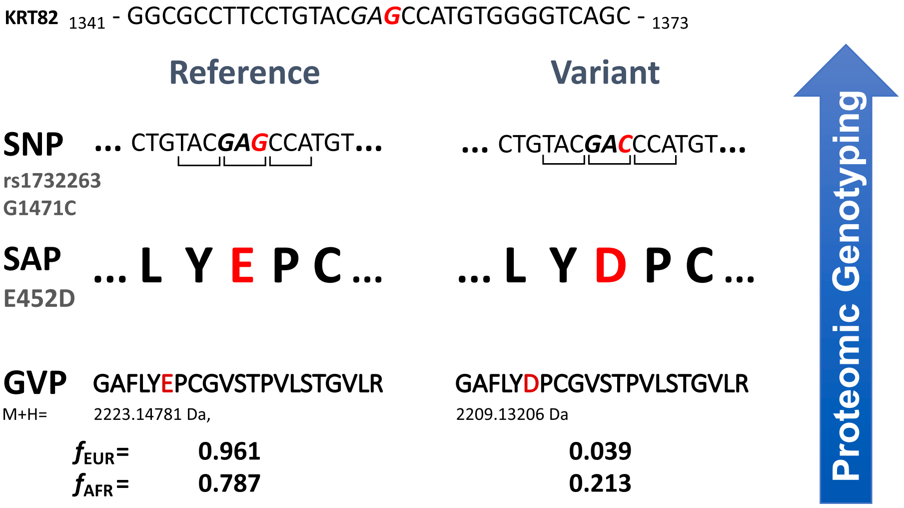

Fig. 3. Proteomic Genotyping is the Use of Proteomic Data to Infer SNP Genotype. A non-synonymous SNP results in a change in codon assignment, in this case G1471C in the gene KRT82 (rs1732263). This variant results in a change in codon assignment from a glutamate (E) to an aspartate (D) at amino acid in position 452. These changes can be detected in proteomic mass spectrometry, since the trypsin digested peptide results in a shift in peptide mass and changes in the resulting fragmentation spectrum. Detection of this peptide allows the investigator to infer the presence of the corresponding SNP allele in the donor's genotype. (reproduced with permission from [245]).

sample set and correctly confirming inferred SNP allele directly with DNA sequencing. As discussed above (Section 4.5) other validation mechanisms, such as standard peptides, are also available.

At a practical level there are factors that limit the effective number of usable GVPs for human identification. Proteomic genotyping is a "read" of the nucleotide information that is filtered by the translational and transcriptional machinery of the cell. Different tissues, with different proteins present, will have a different set of relevant GVPs. Another primary limitation is ionization efficiency. As with any potential peptide biomarker some GVPs may be too long or hydrophobic to efficiently ionize during electrospray ionization or be resolved on chromatography [248]. Some peptides are subject to matrix effects, too hydrophilic or too modified to bind efficiently to chromatography columns. Some peptides come from more easily degraded portions of a protein. Some GVPs may be too small or too similar to other proteins, and so have peptide sequences that occur elsewhere in other gene products that are non-variant. This is common among the highly paralogous keratins and keratin-associated protein families. For a GVP to be useful for identification and individualization the sequence can only be unambiguously derived from a single point on the genome. Homologous sequences from other genes will most likely be invariant and present in every individual. For many loci, only one of the paired GVPs derived from a SNP locus may be appropriate or useable for allele inference [9].

There are other chemical limitations to GVP spectra interpretation. Firstly, there are polymorphisms that have the same or similar mass. These include isoleucine to leucine and lysine to glutamine and *vis versa*. Secondly there are chemical modifications that have the same mass shift as a genetic variant. Common modifications, or chemical variants, include asparagine to aspartate, glutamine to glutamate (deamidation), methionine to phenylalanine, and proline to leucine or isoleucine (oxidation). These genetic polymorphisms should be excluded because they can be accounted for by common chemical variants. Depending on the context and on a case to case basis this may also include polymorphisms accounted for by formylation (lysine to arginine) or incomplete alkylation with iodoacetamide (minus 57 daltons) or overalkylation may be confused as arginine to valine or *vis versa*. A final category is spectra of diverse, sometimes multiple, chemical variants where a peptide spectral match can be made by using a genetic amino acid variant. These spectra look and align differently to a peptide with true genetic variation, even though they have the same total mass. There will often be 'blank' regions of the peptide sequence where there are no alignments, and they will often have a major, but unaligned fragmentation masses, that corresponds to an amino acid in the alternative allele. These falsely attributed GVPs almost always correspond to very rare non-synonymous SNPs, for the simple reason that there are considerably more to choose from. The likelihood of a GVP being correctly assigned can be broadly described mathematically as:

$$\mathrm{LR} = \Pr(\mathrm{GVP})/\Pr(\mathrm{CVP}) \tag{1.0}$$

where Pr(GVP) is the genotype frequency of the genetic variant, and Pr (CVP) is the combined probability of all chemical modifications that result in the same mass. The Pr(CVP) is intangible and cannot be accurately calculated outside of highly controlled laboratory conditions, although it presumably will increase in older more degraded samples. The numerator however is quantifiable because the GVP distribution would be the same, sampling and population structural issues aside, to known genotype frequencies in extant genetic databases. More common genetic variants consequently are more likely to account for mass shifts in a peptide. Empirically, GVP-based inference tends to fail below 1% genotype frequency, or 0.5% allelic frequency [9]. In addition, a surprisingly high percentage (10–30%) of spectra have minor contributions from other peptides resulting in chimeric spectra that have the potential to contribute to ambiguous spectra and incorrect sequence assignment [249].

6.3. *SNP allele inference from proteomic data*

Using a peptide sequence to infer a SNP allele assumes that a single amino acid polymorphism is correlated to a single nucleotide polymorphism. However, because of degeneracy in the genetic code, the 'wobble hypothesis', there are sometimes multiple potential nucleotide changes that can result in an observed novel amino acid. At a practical level however, only one of the possible non-synonymous SNP alleles is common in the population based on extant DNA databases [43,250]. The loci used in proteomic genotyping are common and well represented, for the most part, in every continental population. The originating mutation event therefore would have occurred in a common ancestor, deep in human evolutionary history, and become commonly incorporated into human germlines prior to global migration events [250]. While other nucleotide possibilities may exist for any one amino acid change, SNPs are typically biallelic and only one possibility is recorded in extant genomic databases. The likelihood of the correct proteomic inference being the recorded non-synonymous SNP allele can be described as:

$$LR = \frac{\mathrm{Pr}\boxed{f()}(annotated\ \ SNP\ \ allele)}{\mathrm{Pr}\boxed{f()}(unannotated\ \ SNP\ \ allele)} \tag{2.0}$$

The probability of an annotated allele is equivalent to the genotype frequency, also described as the population distribution, of a given allele [9,247]. The probability of an unannotated allele resulting in the same change in codon assignment being present in a genome at a specific point is vanishingly low, particularly since extant aggregated nucleotide databases now contain more than 71 thousand genomes and 125 thousand exomes (gnomad.broadinstitute.org) [251]. In the unlikely event of alternative nucleotide variants that result in the same codon reassignment do occur then the most common allele could be used in the product rule to provide the most conservative random match probability estimate. The same considerations also apply to mistranslation of mRNA that can occur in conditions of cellular stress and tRNA depletion, with the additional provision that any change would occur as a minor contributor to the pool of translated peptides from a nucleotide sequence [252,253]. As discussed above, any inferences can also be validated by direct confirmation of the SNP allele in a subject's genome or inclusion of SIL peptides in the sample. Using these techniques the performance of individual GVP inferences can be quantified in terms of false discovery rates (FP/(FP+TP)), sensitivity (TP/(TP+FN)) and accuracy (TP+TN)/(TP+FP+TN+FN) [254,255].

6.4. *Proteomic Genotyping and the Product-Rule*

A profile of inferred SNP alleles, like any nucleotide profile, can be used to estimate the distribution of the given profile of variants in a reference population [9,256]. Even if the fidelity of SNP allele inference can be assumed, the profile of non-synonymous SNP alleles still has unique features that need to be considered when estimating random match probabilities. The first factor is the stochastic nature of GVP detection in shotgun-proteomics [63]. This acquisition mode samples as many peptides as possible out of a complex population that may exceed a million different peptides in live cells, and potentially even more in forensic tissues subject to environmental degradation [62]. Because of this GVP-based inference is intrinsically positive in that assumptions can only be made from what is detected, because genotyping is binary in this model it is intrinsically difficult to separate a homozygotic and heterozygotic genotype. It is also difficult to set phase, because peptides are small and it is not easy to assign an allele to a haplotype [256]. The only assumption that can be made is that the individual is not homozygotic for the other allele [9,256]. The appropriate estimate of the population distribution for any one GVP therefore is:

$$\Pr(\mathrm{GVP_A}) = 1\text{-}f_{\mathrm{B}}{}^{2} \tag{3.0}$$

where ‘B’ is the alternative allele and *f* is the allelic frequency. Because the 1000 Genomes Project provides individual genotypes [43], equation 3 can be rewritten as:

$$\Pr(\mathrm{GVP_A}) = 1\text{-}f_{\mathrm{B|B}} \tag{3.1}$$

to provide an empirically based value, where ‘$f_{\mathrm{B}\square\mathrm{B}}$’ is the proportion of the reference population that is homozygote for the other allele. These values can be framed as the proportion of the population that has at least one copy of allele ‘A’. The limitation of binary detection of GVPs however is confined to shotgun proteomics. Quantitative mass spectrometry, by measuring the amount of both peptides deriving from a SNP locus, can determine the range of respective peptide signals at which heterozygote peptides occur and establish homozygosity or heterozygosity based off these signal ratios [257]. This approach incorporates established limits of detection and quantification [257]. This approach is analogous to calls for homozygosity in STR loci based off relative fluorescence signals [5]. A similar approach is also used to estimate the probability of females sex based of the amelogenin X-isoform signal in the event that the Y-isoform is not detected in proteomic sex estimation using logistical regression curves [183,258]. The potential to identify homozygosity based off proteomic signals would considerably increase the discriminating power of proteomic genotyping.

Genetically variant peptides are not evenly distributed across the proteome and can be subject to linkage disequilibrium. Some genes, such as those encoding catalytic activities, are under higher levels of selective pressure and have less or no common single amino acid polymorphisms. The density of GVPs in other open-reading frames appears sporadic and ranges from zero to multiple common non-synonymous SNP alleles. Multiple GVP loci within an open reading frame are subject to linkage disequilibrium. Automatically treating each GVP locus as statistically independent is therefore inappropriate. The initial model for application of the product-rule treated each open reading frame as a single locus, assuming complete linkage within the gene and complete independence outside of the gene boundary [9,244]. The expansion of locus boundaries to incorporate a whole gene has the advantage of substantially expanding the functional number of GVP-inferred alleles from 2 to 4, 8 or even more. The other advantage is that a frequentist counting of each combination of non-synonymous SNP allele combinations make no assumptions about linkage disequilibrium [259]. In order to determine the population distribution of each SNP combination in a gene locus matching individual genotypes are needed [9,244]. These are available using the 1000 Genomes Project, but not when using larger aggregated nucleotide databases [43,251]. Gene-based frequencies of GVP-inferred combinations can be calculated by counting the number of observations of a given combination in respective major reference populations in the 1000 Genomes Project [9,244]. By treating each gene as a GVP-inferred locus, random match probabilities can now be estimated as the product of the frequencies of each gene combination. This can be described mathematically as:

$$RMP = \prod (fGVP_{comb}|population) \tag{4.0}$$

where *RMP* is the random match probability, ‘$f\mathrm{GVP_{comb}}$’ is the population distribution of the combination of inferred non-synonymous SNP alleles within a gene boundary, and ‘population’ is the respective major reference population in the 1000 Genomes Project. The earlier method of counting GVP combinations was effectively limited to three SNP loci because this resulted in 27 different possible combinations, given that heterozygotic combinations need to be treated equivalently [9,244]. Recently a study published an approach that directly counted given combinations against all genotypes in the respective major populations in the 1000 Genomes Project and incorporated it into a ‘GVPFinder’ algorithm, bypassing the need to restrict analysis to only three loci per gene [136].

The major assumption of the above product rule model is that the effects of linkage disequilibrium fade beyond the boundaries of the open reading frame. For many GVP-inferred loci this is a reasonable assumption and one that is easily verified [260]. This approach is a temporary solution however since about 40% of inferred SNP alleles from the hair shaft and skin cell proteome occur in keratin and keratin-associated protein gene clusters on chromosomes 12 and 17 [261–263]. The assumption of equilibrium outside of the gene boundary therefore cannot be taken as a given and needs to be individually tested. If there are residual levels of linkage disequilibrium beyond the open reading frame that include other GVP-inferred SNPs then the locus boundaries would need to be expanded to cover these additional inferred SNP alleles. or Alternatively, statistical dependency may be incorporated directly into a modified product rule or some inferred SNP alleles removed from analysis. One suggested solution is to arbitrarily set a value of 200,000 bp instead of the open reading frame [247]. This would significantly reduce linkage disequilibrium, although it may be difficult to apply in a gene cluster. There also may be instances of residual linkage equilibrium that would require incorporation into product rule calculations. Another recent study proposed that the boundaries of a GVP-inferred locus should be expanded to incorporate a whole chromosome; frequentist approaches to estimate population distribution would be taken within the chromosome and Bayesian approaches between chromosomes [256]. This expansive approach solves the problem of linkage disequilibrium but introduces additional problems. It does not take many inferred SNP alleles to observe genotypes that are not represented in any extant genetic databases, resulting in null observations. Null observations are difficult to handle statistically and use of upper bound estimates using the Clopper-Pearson, Jeffreys beta prior, or 5/2 n rule all result in overly conservative RMP estimates, and in the case of the 5/2 n rule, introduce additional biases [264–266]. This is similar to the problems commonly associated with random match probability estimates of mitochondrial DNA or Y-chromosome STR haplotypes [265,266]. The conservative errors in mitochondrial and Y-STR haplotype RMP estimates are considered appropriate in the United States, but overly conservative in European jurisdictions [265,266]. For now, the appropriate nucleotide boundaries of genetic loci, where the assumption of total linkage can be made internally and no linkage occurs outside of these boundaries, require definitive definition and empirical demonstration.

Precise measurement of linkage disequilibrium opens up the possibility of incorporating GVP-inferred SNP genotypes with other genotypes. Proteomic genotyping is more likely to make a positive contribution to forensic practice in highly degraded evidence where only partial or null STR profiles and mitochondrial haplotypes are available for forensic analysis [266,267]. GVP-inferred loci are autosomal [148]. In aggregate multiple types of genotypes, STR, identifying SNPs, Alu-elements and GVP-inferred SNPs, may combine into a single potentially highly discriminating genotype, provide linkage effects be addressed [267,268]. A recent publication measuring the nucleotide distance between skin cell GVP-inferred loci and STR loci on the same chromosome showed that the closest STR was 2.2 Mb away and 90% were more than 32 Mb away [148]. These nucleotide distances are well beyond any residual effects from linkage disequilibrium and combined complex genotypes safe from linkage effects and risk of statistical dependencies [245,269,270]. The same is true for combining mitochondrial haplotypes and autosomal genotypes. In spite of the fundamentally different nature of inheritance and population distribution, inheritance of nucleotide variation is biologically independent at the mechanistic level. The combination of mitochondrial haplotype and GVP-inferred genotypes into a single random match probability is consequently an attractive possibility [9,271]. Any errors resulting from overestimation of combined haplotype and genotype frequency are unlikely to be greater than those that result from a conservative application of upper-bound priors such as those obtained with the Clopper-Pearson estimate. Nevertheless, even justified theoretical models need to be validated by suitably powerful matching datasets, as was generated by the Budowle laboratory on a Chilean population of matching haplotypes and STR profiles [266,272]. That study observed low levels of dependency but concluded that these were not functional. To quote: "[b]ased on autosomal loci, Y-STRs, and mtDNA, these three systems are mutually independent" [272].

Protein-based product-rule estimates also need to accommodate standard corrections that are necessary in other forms of forensic genotyping. Because GVPs are within open reading frames, presumably they are also subject to positive and negative selective pressures in addition to genetic drift. Hardy-Weinberg Equilibrium need to be tested for and GVP-inferred SNP alleles that deviate may need to be eliminated from forensic analysis. Population structures, both between and within populations would need to be addressed just as they are for STR-typing. The incorporation of F_{ST} values into the product rule is the subject of a recent publication [256]. Just as SNP loci are used for different forensic purposes, depending on their properties, it is reasonable to expect GVP-inferred SNP loci to also become specialized [191,256,273–276]. For example, GVP-inferred loci with higher F_{ST} values may be used specifically for ancestral estimation.

6.5. Likelihood of biogeographic ancestry

Differences in F_{ST} values for GVP-inferred SNPs, as observed for other SNP loci, result in RMP estimates that vary widely as a function of the variable genotype frequencies that are observed in different reference populations. The result is RMP values that vary by many orders of magnitude [9,148,273,277]. These differences have the potential to provide likelihood values for genetic background and preliminary values have been calculated in the literature [9,136]. These estimates are derived as the quotient of random match probabilities derived using genotype frequency values from different reference populations [9]. Describing this mathematically:

$$LR = \frac{Pr\left(GVP_{profile}\middle|population_1\right)}{Pr\left(GVP_{profile}\middle|population_2\right)} \tag{5.0}$$

where the numerator is the random match probability estimated using genotype frequency values obtained from one population, and the denominator is the random match probability using values from another population. When comparing European and African populations the results have been high, with likelihood values ranging over 6 orders of magnitude. Profiles of genetically variant peptides from European samples were more common in European reference populations less frequent in reference African populations and vice versa [136]. The resulting likelihood is a value of relative relationship of the sample between two populations. These estimates are not equivalent to absolute ancestry and cannot in isolation be taken as proof of genetic background. In order to move beyond a tentative estimate, the reliability of likelihood values need to be measured on an appropriately sized sample of known ancestries. Likewise, the variability of likelihood values needs to be established within each population. The resolution of these values also needs to be established. When estimates of random match probability were calculated from all major reference populations in the 1000 Genomes Project using GVP-inferred SNP allele profiles from human skin corneocytes, no effective difference could be determined among non-African populations [148]. This calculation of likelihood ratio is not equivalent to a probability that a given individual belongs to a given genetic ancestry. The likelihood values vary widely, if with different distributions, within populations. The proportion of a given population having a given LR value remains to be established. Ancestral population are not equivalent to geographic populations. Nevertheless, these experiments are achievable and the potential exists where two types of appropriately framed information can potentially be provided to investigators: individualization through estimation of RMPs and biogeographic ancestry through use of likelihood ratios.

6.6. Proteomic genotyping in hair shafts

Hair shafts are a part of every crime scene but are underutilized forensically [278,279]. Hair shafts are discrete, robust, and easy to obtain. Because of this they have been the initial focus of proteomic genotyping development as a usable and practical method for investigators [9]. More than 90% of healthy individuals shed more than 75 cranial hair shafts per day, roughly one every 15–20 min [279–281]. Proteomic processing of hair shafts is a challenge. The tissue is robust, due to the high level of isopeptide and disulfide cross-linking that effectively makes the hair shaft a single protein molecule [282]. These factors make hair a forensically persistent tissue [283]. Hair can be a vehicle for DNA transfer and potentially associate an individual to a crime scene [278,284]. While the dominating proteins are a range of paired Type-I and Type-II keratin gene products and keratin-associated proteins, the proteome of hair shafts is surprisingly complex [9,188–190,247,285–287]. The final stages of the cornification process, transglutaminase reactions, entrap and cross-react with house-keeping genes and other essential gene products required to maintain cell viability up to the final stage of terminal differentiation [9,284,288]. A method demonstrating a multi-omic isolation of both mtDNA and GVPs from hair shafts, with minimal consequences for yields of either type of biomolecule, has also been published [142]. After processing and application to shotgun mass spectrometry, datasets from hair shafts are interrogated for the presence of GVPs [9,136]. The profile of GVPs are then collated and used to estimate random match probability at this point using the RMP method introduced in the initial demonstration paper, although some assumptions regarding linkage disequilibrium need to be addressed [9,256]. Using the current iteration of the product rule the estimated probabilities that a given profile of SNP alleles occurs in the 1000 genome European population ranges up to 1 in 1.1 million and up to 1 in 640 million in African populations [136]. This is roughly consistent with other groups that also estimate random match probabilities from proteomic data [247,256,271].

Implementation of hair shaft proteomic genotyping requires a rigorous validation process analogous to that used for other novel forensic technologies [222]. The validation process in hair shafts has been demonstrated in work by at least six independent laboratories [136,142,191,247,289,290]. Most publication to date has focused on various aspects of sample preparation, particularly the demonstration of the method on single hair shafts, and compatibility with mitochondrial DNA isolation [136,142,191,247,289,290]. However research effort has also been spent meeting the guidelines for validation laid out by SWGDAM, which require performance to be evaluated on samples exposed to real world situations and scenarios [222]. Proteomic genotyping consequently has been applied to hair from different body sites, pigmentation status, sample ages and post-explosion, to determine if these introduce a systematic bias onto the composition of GVP-inferred genotypes. These scenarios each result in distinctive changes to proteomic protein profiles [186,187,285,291,292]. However, the detection of GVP profiles in these samples was more robust and GVP-inferred SNP allele profiles were more correlated to the individual in these different contexts, showing that genotype was more determinative than the biological or chemical context [186,187,291,292]. The one exception was a study that focused on a limited subset of GVPs and where a single peptide was differentially detected in hair from different body sites, and was inconsistent with an earlier study that worked with more GVPs [285,291]. Robustness in GVP-inferred genotypes is not unexpected. GVP profiles are binary and detection or non-detection is a more robust outcome than continuous changes in protein profiles that are a result of biological or chemical conditions. As observed above however, there would be limits to GVP-profile stability [285]. With more proteomic information extracted from a forensic sample, or with greater difference in tissue type or chemical treatment, changes in relative amounts of different protein will start to be reflected as systematic differences in the profiles of detected GVPs. Partial profiles are accepted within certain limits for STR-typing [5]. As with SNPs and STRs loci it would be a matter for regulatory bodies and legal precedent to determine acceptable tolerances for changes in profiles of GVP-inferred loci. The accepted level of sensitivity for GVP detection has not been systematically addressed by researchers in the field, although it has been a major focus of research conducted by the FBI laboratory [286,290]. The efforts to this point have focused on increasing the sensitivity of GVP detection through sample processing and instrumentation, as opposed to setting definitive guidelines for inclusion of a given GVP in legally accepted forensic datasets and analyses.

Implementation of proteomic genotyping, as with any novel forensic technology, is driven by a range of factors. The major is the utility it provides to investigators in terms of useful identifying information and lead generation, particularly if other established methods provide partial or null data. Hair proteomics is attractive here because of the ubiquity of hair shafts and amount of information that may be gained proteomically when other methods, such as mtDNA haplotyping or DNA-typing provide partial or inconclusive findings. Additional priorities also play a role, such as compatibility with other methods, sample consumption, and amount and type of information gained. The method would need to pass hearings, in a *Daubert*, *Frye* or *Kelly* context, on the legal applicability of evidence obtained. These have strict and well-established guidelines, SWGDAM requires a thorough analysis of the replicability, reproducibility, sensitivity, detection limits and whether the human GVPs also occur in other species [222,286,290,293]. These will start to accumulate as the priorities of the field move from increasing sensitivity and maximizing detection, a binary outcome, to a more quantitative approach that incorporates limits of detection and quantitation. This will mirror the work conducted on proteomic body fluid determination [224,232,234–236,239]. Cross-laboratory studies have been conducted, although the field is yet to consolidate around a single mass spectrometry platform [142,293]. This would help to establish standardization, guidelines for interpretation, and appropriate levels of consistency, reproducibility, robustness, sensitivity and discriminatory power. Costs will also become standardized. The issue of linkage disequilibrium in the application of the product rule for GVP-inferred SNP profiles will need to be addressed [9,256].

We anticipate that application of proteomic genotyping in a forensic context will comprise of proteomic analyses of evidentiary material, followed by DNA sequencing of the DNA material from the reference sample of the suspect, or the individual thought to be the donor. This does not preclude the possibility of proteomic analysis of both samples, but a mixed analysis of this type will remove one of the points of variation between the evidentiary and reference sample (Fig. 4). This also means that implementation of hair shaft proteomics as a tool of human identification will also ultimately require development of SNP assays that correspond to the GVP-inferred non-synonymous SNP genotypes [9].

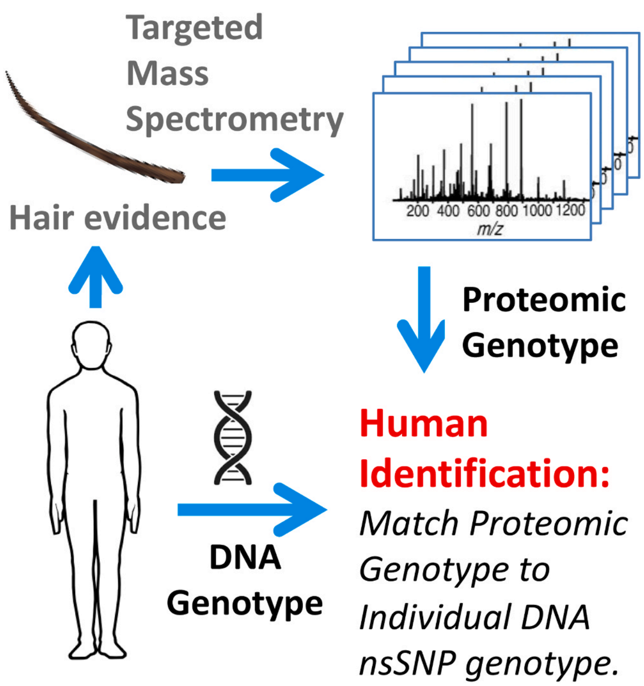

Fig. 4. Proteomic Genotyping of Hair Shafts in a Forensic Context. An evidentiary hair shaft found at a crime scene can be proteomically analyzed using mass spectrometry and endogenous GVPs in the hair shaft detected. These can be used to develop an inferred genotype of SNP alleles. The resulting profile can be directly matched to a suspect's DNA genotype and a statistical measure of association estimated.

7. Other applications of forensic proteomics

The above applications of proteomics reflect the major focus of forensic proteomics and are closest to implementation. The two above applications, body fluid and tissue identification and individualization through genotyping, reflect the historical priorities of forensic protein science. However, other applications of forensic proteomics are also under active investigation and development. Some of these also reflect historical priorities of forensic science, such as fingermark analysis and species determinations. Other applications reflect new priorities in the field: peptide toxicology, sex estimation and estimation of the post-mortem interval.

7.1. Proteomic analysis of fingermarks

Fingermarks are an important source of forensic information that can provide a unique link between an individual and evidence or crime scene [294–301]. Fingermarks are created when material is transferred from the fingers or palms onto a solid surface. The transfer often contains forensically useful information, either in the form of a two-dimensional friction ridge pattern, or informative molecules such as small compounds, DNA and protein [294,297,301–306]. The transfer is highly variable, often resulting in low levels of material available for analysis [307]. A dominant feature of fingermarks is the difficulty in obtaining DNA [308]. The deposition of biological material depends on many variables, such as the length and pressure of contact, the porosity and adhesiveness of the deposition surface, the level of transferable material on the donor's skin, and environmental factors such as humidity [307]. This often results in failure to obtain useable STR-profiles and this has driven the development of low copy number DNA processing methods to maximize the amount of template present in analysis [308]. Because the fingermarks are on exposed two-dimensional surfaces, they are subject to environmental degradation [305,309–312]. In a scenario where one particular type of information may not provide sufficient information, such as a partial STR-profile, incorporation of additional information types may provide additional identifying information or investigative context [143,148,313–315]. Analytical methods therefore need to be multi-omic, where sample processing does not compromise other modes of analysis [140,141,143,194,304,316–318]. Particularly emphasis is placed on efficient transfer of material so the maximum amount of material enters into the analytical workflow [140,143,194].

The nature of fingermarks allows MALDI mass spectrometry to be conducted in a two-dimensions resulting in images of multiple masses [295,304,316,319–321]. This is a powerful approach to identification of relevant compounds in a fingermark, providing information about an individual's biology, metabolome and chemical exposure [295]. This analysis also maintains and potentially enhances digital information in the latent print, resolve overlapping proteins, detect fingermarks below bloodstains [319], and can be adopted for proteomics [313,314,322–324]. Proteomics has been used to identify body fluids in fingermarks, such as vaginal fluid or blood [162,314,315]. Fingermark proteins, like any protein source, are a source of genetically variant peptides that can be used for human identification and potentially estimates of biogeographic background [143,148,194]. The GVP- inferred SNP genotypes are compatible with other genotypes, such as STRs and *Alu*-element profiles [148]. The forensic role of proteins in fingermarks is an active area of investigation and promises to be at the forefront of developments in sample capture and transfer into analytical workflows, sample processing, proteomic genotyping, and development of complex genotypes into single values for human identification and lead generation.

7.2. Protein markers for species identification

There are many forensic scenarios where it is necessary to identify the originating species of biological evidence as part of standard criminal investigation, enforcement of fish and wildlife regulations, customs and border protection, food security, and analysis of foreign protein contaminants and adulterants. Just as proteins contain information of genetic variation in human populations, they also contain phylogenetic information. As animal populations differentiate, and evolutionary biodistance increases, variants within a population may become the major allele through selective pressure or genetic drift. Eventually, novel alleles may become invariant with nucleotides and amino acids that are no longer polymorphic, but rather population- and species-specific. Many peptides from different species will remain identical in ontologically related proteins, particularly from related species, but commonality decreases with evolutionary distance [198]. However, well annotated faunal genomes allow for simple identification of species-specific amino acids and resulting peptide biomarkers. The immediate forensic application is the species identification of bone pieces with no remaining DNA or osteological markers [12,90,201,325]. Many of these protocols use a variation of mass spectrometry called Matrix-Assisted Laser Desorption Ionization (MALDI) that generates a precise and accurate peptide mass fingerprint off a two-dimensional surface. Described as Zooarchaeology using Mass Spectrometry (ZooMS), MALDI is commonly used in species identification using bone proteomes. ZooMS is based on the principle that closely related species will have more overlapping peptide masses while also containing combinations of species- and genus specific peptide masses [326,327]. The identification of species from degraded bone material using proteomics is an active area of investigation within the paleoproteomic and bioarchaeological communities and attempts have been made to systematize analytical workflows and set standards for their communities [90]. Necessarily, these criteria differ to those required of the forensics community. Investigators engaged in wildlife forensics often need to analyze samples that are degraded or highly processed, both of which can result in DNA removal with protein preservation.

7.3. Peptide toxicology

Another related application is toxicological detection of introduced biological material into a sample [328,329]. DNA-based methods will not be helpful when a purified protein is added to a sample, such as a protein toxin. Forensic toxicology now also includes protein and peptide analyses using targeted proteomics methods [127,330,331]. A prominent example is the biological warfare agent ricin protein [125,127,332–334]. It is part of a wide family of protein toxins that need to be detected in food production and a wide range of other legal and security contexts [335]. Other foreign peptides that are targeted in toxicology include dermorphin, a 7-amino acid peptide from the skin of the South American *Phyllomedusa* frog, and a potent binder to the opioid receptor [336]. It is part of a family of peptides used in horse racing [337,338]. A related application has been developed for detection of synthetic insulins, where suspected overdoses have been implicated in homicides as commercial insulins have specific amino acid modifications [328,329,339].

7.4. Proteomic sex estimation

Forensic proteomics can also be used to sex an individual. Amelogenins, the most characterized sex specific gene family, are expressed as a gene product in enamel, the most robust human tissue [184,340]. The stability of protein in enamel is much higher than other tissues [89,183], due to the two dimensional microcrystal interface of calcium apatite that increases protein stability, deep into archaeological time [10,90,91]. There are two forms of the protein, one from each sex-chromosome [183]. Amelogenin protein is integral to enamel biosynthesis, a process that degrades the protein with endogenous proteases. The result is multiple peptides representing many of the amino acids in the protein [183]. These peptides are a rich source of biomarkers for the X- and Y-chromosome. These can be quantified and used to partition samples into male and female samples. The detection of these respective markers may consequently be used to estimate the sex-chromosome karyotype of the enamel donor [183,205]. Proteomic sex estimation has successfully been applied to enamel from a 9400 year old *Homo sapiens* samples and even an 800,000 year old *Homo antecessor* sample [179,183].

Proteomic sex estimation is a third fundamental method of sex estimation. Osteological sex estimation is easily conducted in the field by a trained expert but is limited by the preservation status of human remains and the age of the deceased individual, male and female skeletons only become dimorphic after puberty [257,341,342]. Genomic sex estimation has the advantage of not being constrained by skeletal preservation or individual development [343–345]. The sex chromosome amelogenin loci are in all commercial forensic STR-based analysis kits [5]. Current forensic sex-estimation methods rely on both STR-typing and quantitative PCR, such as used in the Quantifiler® HP and Quantifiler® Trio DNA Quantification Kits [312,343,346–349]. Detection of male DNA is sensitive in current forensic methods because it uses a multi-copy locus on the Y-chromosome [265,350]. Environmental contamination using qPCR is an issue, but this is addressed by subsequent measurement of amelogenin allele balance and mixture analysis [265]. Highly degraded samples may still, theoretically, pose a challenge if DNA template levels become too low. The sensitivity levels at which sex-determination using forensic DNA methods begin to fail appears not to be demonstrated, although this has been demonstrated using ancient DNA methods that are more sensitive [258,351–354]. Proteomic sex estimation outperforms both osteological and ancient genomic methods [258]. Proteomic sex estimation of an archaeological sample of 55 individuals resulted in confident sex estimates in all samples, with genomic and osteologic methods providing confident estimates in 64% and 25% of individuals respectively [258]. In the comparison study proteomic sex estimation signals did not decline over time, yet DNA reads decreased by roughly an order of magnitude per millennium [258]. Proteomic sex estimation also was successful on fetuses, infants and partially cremated individuals. Genomic and proteomic sex estimates sometimes conflicted [258]. These conflicts were all in lower quality DNA samples, below 100,000 reads. Conditional, ‘consistent with.’ DNA -based sex estimates below this level were incorrect 31–38% of the time depending on the DNA-method used [258]. No such study directly comparing proteomic and forensic DNA sex estimation has been attempted.

Proteomic sex estimation is a validated technique with clear applicability in archaeological contexts [353,354]., the utility in forensic contexts needs to be established. Any new method requires rigorous validation tests to assess the overall performance, advantages, and possible caveats, in order to evaluate its range of applications within forensic workflows. Sex estimation using forensic DNA sample processing and analytical workflows is considered robust, yet the potentially more accurate ancient DNA methods proved to be less reliable than proteomic methods. A systematic analysis of current forensic sex-estimation as a function of DNA quality and direct comparison with proteomic sex estimation remains to be demonstrated.

7.5. Estimation of post-mortem interval

Environmental chemistries change the character of protein as a function of time. Proteomic degradation therefore may be a means to proteomically calibrate the post-mortem interval [355–357]. This has been applied to predict the short-term post-deposition interval of fingermarks [305]. Over longer periods, changes in the bone proteome have been used to identify proteins that degrade at different rates over a period of months or years. This data has been used to develop algorithms for prediction of post-mortem intervals [170,358]. Housekeeping globular proteins, such as glycolytic enzymes, are more flexible and consequently more subject to degradative chemical processes [355]. A recent study used proteomics to study the release of decompositional fluid post-mortem in pigs [359]. Changes in the profile of proteins in fingermarks have also been used to estimate the post-deposition interval [305]. This process may be even more consistent in bones left in aqueous environments [355]. The advantage of proteomics is that these methods are not dependent on anatomical integrity [170]. The role of protein modifications also plays a potential role here. For instance, deamidation of hair proteins has been measured to increase over a period of decades [292] and has also been used to estimate the post-mortem interval [355], date archaeological material and separate out archaeological protein from modern contaminating keratins [88,179,360–363]. Deamidation is known to occur at different rates depending on the amino acid, flanking protein sequences and protein flexibility [364–368]. These complex and variable chemical and environmental factors result in a high level of variance of deamidation at different time points, indicating it is more of a preservation and exposure signal than an indicator of time [102,183,363]. These processes will be accentuated and possibly dominated by the introduction of environmental microflora that will use the protein as a supply of nutrient and introduce non-specific proteases. While the sample will increase in complexity due to these diverse modifications, the portion of the proteome available for analysis will decrease in complexity, resulting in information loss. The contingent nature of chemical modifications, large number of potential modifications, and limited ability to fully identify all proteoforms [117], all present challenges to estimation of the post-mortem interval. The necessary experiments are achievable however and within the scope of current technology.

8. Looking forward

Just as the last 30 years have introduced fundamental change in proteomic technology, further innovations will continue to shape and spur development in the forensics proteomic field. Instrumentation will continue to evolve. For example, as described above, the latest generation of mass spectrometer, a timsToF mass spectrometer from Bruker, is able to obtain 600,000 spectra from a 2 h run and 200 ng of protein [60]. When compared to yields from a Q-Exactive-Plus orbitrap mass spectrometer from Thermo, the state of the art only 5 years previously, this is more than ten-fold the amount of data using one tenth the amount of protein [59]. The acquisition of data on mass spectrometers, particularly the use of stable isotope chemistries, the use of sample multiplexing, and data independent acquisition strategies, will have an impact on forensics just as they will on the rest of the proteomic field. In terms of sample preparation, the use of magnetic beads promises to dramatically expand the applicability of proteomics to degraded and environmentally exposed samples [130,149,369]. Robotics and shortened MS gradients times will expand the number of that can be applied to proteomic workflows. Another potential breakthrough is the use of nanopores for sequencing of single protein molecules [370]. Similar in many ways to the Oxford Nanopore third generation DNA sequencing platform, the potential of single protein molecule sequencing has obvious potential for forensics. In addition to increased sensitivity, this approach directly sequences protein and theoretically removes the need to match complex fragmentation spectra with sequences from reference protein databases [371]. There are still significant technical challenges [371]. Instead of 4 nucleotides there are 21 amino acids with many potential post-translational modifications, resulting in many different hydrodynamic volumes and current blockages [371,372]. Many of these resistivities are very similar [371]. Sequencing of homopolymeric regions, long a problem in DNA nanopore sequencing, are also a problem with protein sequencing [371]. Elimination of secondary structure in proteins may prove more challenging than in DNA and greater control of peptide translocation is also required [371]. Nevertheless, peptide sequencing has been accomplished, including identification of peptide sequence variants [372]. Advances in proteomics also extend to bioinformatic analysis. The use of machine learning in creating and differentiating proteomic profiles is a promising and powerful area of future investigation [373–375].

9. Conclusion

Protein is intrinsic to biological evidence. It is highly stable and is the most abundant biomolecule in forensic samples. The protein population in a biological sample is numerous, and highly complex with useful chemical, biological and genetic properties for the investigator. The protein information in biological evidence has historically been used for body fluid and tissue determination, human identification, and species determination. The advent of powerful, highly sensitive, informative and robust DNA-based methodologies changed the focus of forensic protein science towards body fluid identification. In recent years proteomic technology, through mass spectrometry, has developed to extract more information from the protein component of evidence, and to do so using simpler, more consistent and reproducible workflows. Protein has always been, and continues to be, a major focus of body-fluid identification. Now due to mass spectrometry reliable peptide biomarkers for body fluid identification are becoming more quantitative, sensitive and robust, performing at higher levels of sensitivity and accuracy than other antibody-dependent platforms. Forensic protein sample processing methods and mass spectrometry are beginning to be developed that are compatible with DNA-based methods, consume less sample, and can be automated. This provides a powerful tool for understanding the forensic context of a crime scene; the tissue source of biological evidence can be a critical piece of investigative information. The development of proteomic mass spectrometry and improved sample processing methodologies also provide new opportunities for extracting identifying genetic information from a forensic sample. The detection and generation of genetically variant peptide profiles from protein can still occur after all DNA in a sample has been degraded. Initially focused on hair shafts the technology for GVP detection continues to evolve and challenges such as application of the product-rule are being actively investigated. Validation studies such as application to real life scenarios and cross-laboratory performance studies are beginning to occur. Forensic proteomics also extends to other potential applications. Proteomic sex estimation, post-mortem interval estimation, species determination and fingermark analysis are all being actively investigated. Each of these areas are at different stages of development, and the applicability of each method to forensic investigation and how, and how much, they become adopted by the forensic community remains to be established. The mandate of forensic science is to increase the integrity of the criminal justice system. This means, among other things, that investigators need to obtain the maximum amount of information from a sample with a minimum of sample consumption. Forensic proteomics has reached a point where it can, or can soon, create additional tools that help meet these mandates. The more efficient and effective analysis of the protein content of evidence presents opportunities for those seeking to maximize the amount of information that can be extracted from a crime scene.

Acknowledgements

The authors would like to acknowledge the advice and editing of Drs Robert Rice and Brett Phinney. The authors also acknowledge the significant and thorough editorial contribution provided by the peer-reviewers. GJP acknowledges the support of the National Institute of Justice, USA (2015-DN-BX-K065; 2017-DN-BX-0159) and the National Science Foundation, USA (#BCS-1825022). ZCG acknowledges the support of the National Institute of Justice, USA (2019-R2-CX-0051). Any opinions enclosed in this article are those of the authors and are not endorsed by these funding agencies.

References

[1] S. Ballou, M. Kline, M. Stolorow, M. Taylor, S. Williams, P. Bamberger, Y. Burney, L. Brown, C. Jones, R. Keaton, W. Kiley, K. Thiessen, G. LaPorte, J. Latta, L. Ledray, R. Nagy, L. Schwind, S. Stoiloff, B. Ostrom, in: The Biological Evidence Preservation Handbook: Best Practices for Evidence Handlers, National Institute of Standards and Technology, Gaithersburg, MD, 2013 https://doi.org/https://nvlpubs.nist.gov/nistpubs/ir/2013/NIST.IR.7928.pdf.

[2] J.M. Berg, J.L. Tymoczko, L. Stryer. Biochemistry, 6th ed., W.H. Freeman and Company, 2007.

[3] S.C. Tan, B.C. Yiap, DNA, RNA, and protein extraction: the past and the present, J. Biomed. Biotechnol. 2009 (2009), 574398, https://doi.org/10.1155/2009/574398.

[4] National Research Council. DNA Technology in Forensic Science, The National Academies Press, Washington, DC, 1992, https://doi.org/10.17226/1866.

[5] J.M. Butler. Advanced Topics in Forensic DNA Typing: Methodology, Elsevier Inc, Waltham, MA, 2012.

[6] R. Aebersold, M. Mann, Mass-spectrometric exploration of proteome structure and function, Nature 537 (7620) (2016) 347–355, https://doi.org/10.1038/nature19949.

[7] G. Sensabaugh, Serology: overview, in: J. Payne-James, R.W. Byard (Eds.), Encyclopedia of Forensic and Legal Medicine, Second ed., Elsevier, Oxford, 2016, pp. 254–266, https://doi.org/10.1016/B978-0-12-800034-2.00337-2.

[8] K. Sakurada, K. Watanabe, T. Akutsu, Current methods for body fluid identification related to sexual crime: focusing on saliva, semen, and vaginal fluid, Diagnostics 10 (9) (2020) 693.

[9] G.J. Parker, T. Leppert, D.S. Anex, J.K. Hilmer, N. Matsunami, L. Baird, J. Stevens, K. Parsawar, B.P. Durbin-Johnson, D.M. Rocke, C. Nelson, D. J. Fairbanks, A.S. Wilson, R.H. Rice, S.R. Woodward, B. Bothner, B.R. Hart, M. Leppert, Demonstration of protein-based human identification using the hair shaft proteome, PLoS One 11 (9) (2016), e0160653, https://doi.org/10.1371/journal.pone.0160653.

[10] M.J. Collins, C.M. Nielsen–Marsh, J. Hiller, C.I. Smith, J.P. Roberts, R. V. Prigodich, T.J. Wess, J. Csapò, A.R. Millard, G. Turner–Walker, The survival of organic matter in bone: a review, Archaeometry 44 (3) (2002) 383–394, https://doi.org/10.1111/1475-4754.t01-1-00071.

[11] I. Kontopoulos, K. Penkman, V.E. Mullin, L. Winkelbach, M. Unterlander, A. Scheu, S. Kreutzer, H.B. Hansen, A. Margaryan, M.D. Teasdale, B. Gehlen, M. Street, N. Lynnerup, I. Liritzis, A. Sampson, C. Papageorgopoulou, M. E. Allentoft, J. Burger, D.G. Bradley, M.J. Collins, Screening archaeological bone for palaeogenetic and palaeoproteomic studies, PLoS One 15 (6) (2020), e0235146, https://doi.org/10.1371/journal.pone.0235146.

[12] E. Cappellini, A. Prohaska, F. Racimo, F. Welker, M.W. Pedersen, M.E. Allentoft, P. de Barros Damgaard, P. Gutenbrunner, J. Dunne, S. Hammann, M. Roffet-Salque, M. Ilardo, J.V. Moreno-Mayar, Y. Wang, M. Sikora, L. Vinner, J. Cox, R. P. Evershed, E. Willerslev, Ancient biomolecules and evolutionary inference, Annu Rev. Biochem 87 (2018) 1029–1060, https://doi.org/10.1146/annurev-biochem-062917-012002.

[13] C. Wadsworth, N. Procopio, C. Anderung, J.M. Carretero, E. Iriarte, C. Valdiosera, R. Elburg, K. Penkman, M. Buckley, Comparing ancient DNA survival and proteome content in 69 archaeological cattle tooth and bone samples from multiple European sites, J. Proteom. 158 (2017) 1–8, https://doi.org/10.1016/j.jprot.2017.01.004.

[14] H.N. Poinar, B.A. Stankiewicz, Protein preservation and DNA retrieval from ancient tissues, Proc. Natl. Acad. Sci. USA 96 (15) (1999) 8426–8431.

[15] E. Cappellini, M.J. Collins, M.T. Gilbert, Biochemistry. Unlocking ancient protein palimpsests, Science 343 (6177) (2014) 1320–1322, https://doi.org/10.1126/science.1249274.

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

[17] C. Manzoni, D.A. Kia, J. Vandrovcova, J. Hardy, N.W. Wood, P.A. Lewis, R. Ferrari, Genome, transcriptome and proteome: the rise of omics data and their integration in biomedical sciences, Brief. Bioinform. 19 (2) (2016) 286–302, https://doi.org/10.1093/bib/bbw114.

[18] S. Hu, J.A. Loo, D.T. Wong, Human body fluid proteome analysis, Proteomics 6 (23) (2006) 6326–6353, https://doi.org/10.1002/pmic.200600284.

[19] K. Virkler, I.K. Lednev, Analysis of body fluids for forensic purposes: from laboratory testing to non-destructive rapid confirmatory identification at a crime scene, Forensic Sci. Int. 188 (1–3) (2009) 1–17.

[20] R.E. Gaensslen. Sourcebook in Forensic Serology, Immunology, and Biochemistry, US Department of Justice, National Institute of Justice, Washington, DC, 1983.

[21] R.C. Marshall, J.M. Gillespie, V. Klement, Methods and future prospects for forensic identification of hairs by electrophoresis, J. Forensic Sci. Soc. 25 (1) (1985) 57–66.

[22] E.T. Blake, G.F. Sensabaugh, Haptoglobin typing in bloodstains I. Electrophoresis of immune precipitated haptoglobin, J. Forensic Sci. Soc. 18 (1978) 237–244.

[23] S.S. Tobe, N. Watson, N.N. Daeid, Evaluation of six presumptive tests for blood, their specificity, sensitivity, and effect on high molecular-weight DNA, J. Forensic Sci. 52 (1) (2007) 102–109.

[24] A. Greenfield, M.M. Sloan, Identification of biological fluids and stains, in: S. H. James, J.J. Nordby (Eds.), Forensic Science: An Introduction to Scientific and Investigative Techniques, CRC, 2009, pp. 285–302.

[25] J.M. Cecka, S. Breidenthal, P.I. Terasaki, Direct blood group typing of forensic samples using a simple monoclonal antibody assay, Forensic Sci. Int. 34 (3) (1987) 205–216, https://doi.org/10.1016/0379-0738(87)90075-2.

[26] R.H. Shyu, H.F. Shyu, H.W. Liu, S.S. Tang, Colloidal gold-based immunochromatographic assay for detection of ricin, Toxicon 40 (3) (2002) 255–258, https://doi.org/10.1016/s0041-0101(01)00193-3.

[27] B.A. Schweers, J. Old, P.W. Boonlayangoor, K.A. Reich, Developmental validation of a novel lateral flow strip test for rapid identification of human blood (Rapid Stain IdentificationTM-Blood), Forensic Sci. Int. Genet. 2 (3) (2008) 243–247, https://doi.org/10.1016/j.fsigen.2007.12.006.

[28] L.J. Blum, P. EsperanÇA, S. Rocquefelte, A new high-performance reagent and procedure for latent bloodstain detection based on luminol chemiluminescence, Can. Soc. Forensic Sci. J. 39 (3) (2006) 81–99, https://doi.org/10.1080/00085030.2006.10757139.

[29] S. Johnston, J. Newman, R. Frappier, Validation study of the Abacus Diagnostics ABAcard® Hematrace® membrane test for the forensic identification of human blood, Can. Soc. Forensic Sci. J. 36 (3) (2003) 173–183.

[30] J.C. Herr, T.A. Summers, R.S. McGee, W.M. Sutherland, M. Sigman, R.J. Evans, Characterization of a monoclonal antibody to a conserved epitope on human seminal vesicle-specific peptides: a novel probe/marker system for semen identification, Biol. Reprod. 35 (3) (1986) 773–784, https://doi.org/10.1095/biolreprod35.3.773.

[31] W.M. Watkins, W.T.J. Morgan, Possible genetical pathways for the biosynthesis of blood group mucopolysaccharides, Vox Sang. 4 (2) (1959) 97–119.

[32] G.S. Williams, Blood grouping, in: J. Payne-James (Ed.), Encyclopedia of Forensic and Legal Medicine, Elsevier, Oxford, 2005, pp. 283–288, https://doi.org/10.1016/B0-12-369399-3/00041-0.

[33] H. Mukoyama, S. Seta, The determination of blood groups in tissue samples. Forensic Science Progress, Springer, 1986, pp. 37–90.

[34] B. Castillo, A. Dasgupta, K. Klein, H. Tint, A. Wahed, Chapter 5 - red cell antigens and antibody, in: B. Castillo, A. Dasgupta, K. Klein, H. Tint, A. Wahed (Eds.), Transfusion Medicine for Pathologists, Elsevier, 2018, pp. 69–112, https://doi.org/10.1016/B978-0-12-814313-1.00005-8.

[35] L.N. Sussman, Blood grouping tests in disputed paternity proceedings: studies with A-B-O, M-N, and Rh-Hr factors, J. Am. Med. Assoc. 155 (13) (1954) 1143–1145, https://doi.org/10.1001/jama.1954.03690310021005.

[36] R.E. Gaensslen, S.C. Bell, H.C. Lee, Distributions of genetic markers in United States populations: II. Isoenzyme systems, J. Forensic Sci. 32 (5) (1987) 1348–1381, https://doi.org/10.1520/JFS11184J.

[37] A.J. Jeffreys, V. Wilson, S.L. Thein, Individual-specific ‘fingerprints’ of human DNA, Nature 316 (6023) (1985) 76–79, https://doi.org/10.1038/316076a0.

[38] P. Gill, A.J. Jeffreys, D.J. Werrett, Forensic application of DNA ‘fingerprints’, Nature 318 (6046) (1985) 577–579.

[39] D.H. Kaye. The Double Helix and the Law of Evidence, Harvard University Press, 2010.

[40] G.B. Divall, The application of electrophoretic techniques in the field of criminology, Electrophoresis 6 (6) (1985) 249–258.

[41] R.N. Jonakait, When blood is their argument: probabilities in criminalcases, genetic markers, and, once again, Bayes Theorem, University of Illinois Law Review, 1983, p. 369.

[42] R.E. Gaensslen, Forensic analysis of biological evidence, in: C.H. Wecht (Ed.), Forensic Sciences, Matthew Bender and Co, New York, NY, 2000.

[43] Genomes Project Consortium, A. Auton, L.D. Brooks, R.M. Durbin, E.P. Garrison, H.M. Kang, J.O. Korbel, J.L. Marchini, S. McCarthy, G.A. McVean, G.R. Abecasis, A global reference for human genetic variation, Nature 526 (7571) (2015) 68–74, https://doi.org/10.1038/nature15393.

[44] Genomes Project Consortium, G.R. Abecasis, A. Auton, L.D. Brooks, M. A. DePristo, R.M. Durbin, R.E. Handsaker, H.M. Kang, G.T. Marth, G.A. McVean, An integrated map of genetic variation from 1,092 human genomes, Nature 491 (7422) (2012) 56–65, https://doi.org/10.1038/nature11632.

[45] R.K. Saiki, D.H. Gelfand, S. Stoffel, S.J. Scharf, R. Higuchi, G.T. Horn, K.B. Mullis, H.A. Erlich, Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase, Science 239 (4839) (1988) 487–491.

[46] M. Gstaiger, R. Aebersold, Applying mass spectrometry-based proteomics to genetics, genomics and network biology, Nat. Rev. Genet. 10 (9) (2009) 617–627, https://doi.org/10.1038/nrg2633.

[47] S.R. Wilson, T. Vehus, H.S. Berg, E. Lundanes, Nano-LC in proteomics: recent advances and approaches, Bioanalysis 7 (14) (2015) 1799–1815, https://doi.org/10.4155/bio.15.92.

[48] S. Banerjee, S. Mazumdar, Electrospray ionization mass spectrometry: a technique to access the information beyond the molecular weight of the analyte, Int. J. Anal. Chem. 2012 (2012), 282574, https://doi.org/10.1155/2012/282574.

[49] J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong, C.M. Whitehouse, Electrospray ionization for mass spectrometry of large biomolecules, Science 246 (4926) (1989) 64–71.

[50] E.L. de Graaf, A.F. Altelaar, B. van Breukelen, S. Mohammed, A.J. Heck, Improving SRM assay development: a global comparison between triple quadrupole, ion trap, and higher energy CID peptide fragmentation spectra, J. Proteome Res. 10 (9) (2011) 4334–4341, https://doi.org/10.1021/pr200156b.

[51] A.M. Haag, Mass analyzers and mass spectrometers, in: H. Mirzaei, M. Carrasco (Eds.), Modern Proteomics – Sample Preparation, Analysis and Practical Applications, Springer Nature, 2016, 157 to 169.

[52] H. Steen, M. Mann, The ABC's (and XYZ's) of peptide sequencing, Nat. Rev. Mol. Cell Biol. 5 (9) (2004) 699–711, https://doi.org/10.1038/nrm1468.

[53] C. Escher, L. Reiter, B. MacLean, R. Ossola, F. Herzog, J. Chilton, M.J. MacCoss, O. Rinner, Using iRT, a normalized retention time for more targeted measurement of peptides, Proteomics 12 (8) (2012) 1111–1121, https://doi.org/10.1002/pmic.201100463.

[54] S. Gessulat, T. Schmidt, D.P. Zolg, P. Samaras, K. Schnatbaum, J. Zerweck, T. Knaute, J. Rechenberger, B. Delanghe, A. Huhmer, U. Reimer, H.C. Ehrlich, S. Aiche, B. Kuster, M. Wilhelm, Prosit: proteome-wide prediction of peptide tandem mass spectra by deep learning, Nat. Methods 16 (6) (2019) 509–518, https://doi.org/10.1038/s41592-019-0426-7.

[55] E.S. Lander, L.M. Linton, B. Birren, C. Nusbaum, M.C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J.P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J.C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R.H. Waterston, R.K. Wilson, L.W. Hillier, J. D. McPherson, M.A. Marra, E.R. Mardis, L.A. Fulton, A.T. Chinwalla, K.H. Pepin, W.R. Gish, S.L. Chissoe, M.C. Wendl, K.D. Delehaunty, T.L. Miner, A. Delehaunty, J.B. Kramer, L.L. Cook, R.S. Fulton, D.L. Johnson, P.J. Minx, S.W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J.F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R. A. Gibbs, D.M. Muzny, S.E. Scherer, J.B. Bouck, E.J. Sodergren, K.C. Worley, C. M. Rives, J.H. Gorrell, M.L. Metzker, S.L. Naylor, R.S. Kucherlapati, D.L. Nelson, G.M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, D.R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, H.M. Lee, J. Dubois, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R.W. Davis, N.A. Federspiel, A.P. Abola, M.J. Proctor, R.M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D.R. Cox, M.V. Olson, R. Kaul, C. Raymond, N. Shimizu, K. Kawasaki, S. Minoshima, G.A. Evans, M. Athanasiou, R. Schultz, B.A. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W.R. McCombie, M. de la Bastide, N. Dedhia, H. Blocker, K. Hornischer, G. Nordsiek, R. Agarwala, L. Aravind, J.A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D.G. Brown, C.B. Burge, L. Cerutti, H.C. Chen, D. Church, M. Clamp, R.R. Copley, T. Doerks, S. R. Eddy, E.E. Eichler, T.S. Furey, J. Galagan, J.G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L.S. Johnson, T. A. Jones, S. Kasif, A. Kaspryzk, S. Kennedy, W.J. Kent, P. Kitts, E.V. Koonin, I. Korf, D. Kulp, D. Lancet, T.M. Lowe, A. McLysaght, T. Mikkelsen, J.V. Moran, N. Mulder, V.J. Pollara, C.P. Ponting, G. Schuler, J. Schultz, G. Slater, A.F. Smit, E. Stupka, J. Szustakowski, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y.I. Wolf, K.H. Wolfe, S.P. Yang, R.F. Yeh, F. Collins, M.S. Guyer, J. Peterson, A. Felsenfeld, K.A. Wetterstrand, A. Patrinos, M.J. Morgan, P. de Jong, J.J. Catanese, K. Osoegawa, H. Shizuya, S. Choi, Y. J. Chen, Initial sequencing and analysis of the human genome, Nature 409 (6822) (2001) 860–921, https://doi.org/10.1038/35057062.

[56] T. Addona, K. Clauser, De novo peptide sequencing via manual interpretation of MS/MS spectra, Curr. Protoc. Protein Sci. Chapter 16 (2002), https://doi.org/10.1002/0471140864.ps1611s27 (Unit 16 11).

[57] M.J. MacCoss, C.C. Wu, J.R. Yates 3rd, Probability-based validation of protein identifications using a modified SEQUEST algorithm, Anal. Chem. 74 (21) (2002) 5593–5599.

[58] J.K. Eng, A.L. McCormack, J.R. Yates, An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database, J. Am. Soc. Mass Spectrom. 5 (11) (1994) 976–989, https://doi.org/10.1016/1044-0305(94)80016-2.

[59] R.A. Scheltema, J.P. Hauschild, O. Lange, D. Hornburg, E. Denisov, E. Damoc, A. Kuehn, A. Makarov, M. Mann, The Q Exactive HF, a Benchtop mass spectrometer with a pre-filter, high-performance quadrupole and an ultra-high-field Orbitrap analyzer, Mol. Cell. Proteom. 13 (12) (2014) 3698–3708, https://doi.org/10.1074/mcp.M114.043489.

[60] F. Meier, A.D. Brunner, S. Koch, H. Koch, M. Lubeck, M. Krause, N. Goedecke, J. Decker, T. Kosinski, M.A. Park, N. Bache, O. Hoerning, J. Cox, O. Rather, M. Mann, Online Parallel Accumulation-Serial Fragmentation (PASEF) with a novel trapped ion mobility mass spectrometer, Mol. Cell. Proteom. MCP 17 (12) (2018) 2534–2545, https://doi.org/10.1074/mcp.TIR118.000900.

[61] N. Nagaraj, J.R. Wisniewski, T. Geiger, J. Cox, M. Kircher, J. Kelso, S. Pääbo, M. Mann, Deep proteome and transcriptome mapping of a human cancer cell line, Mol. Syst. Biol. 7 (548) (2011) 1–8, https://doi.org/0.1038/msb.2011.81.

[62] T. Geiger, A. Wehner, C. Schaab, J. Cox, M. Mann, Comparative proteomic analysis of eleven common cell lines reveals ubiquitous but varying expression of most proteins, Mol. Cell. Proteom. 11 (3) (2012), https://doi.org/10.1074/mcp.M111.014050.

[63] A. Michalski, J. Cox, M. Mann, More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS, J. Proteome Res. 10 (4) (2011) 1785–1793, https://doi.org/10.1021/pr101060v.

[64] C. Wichmann, F. Meier, S. Virreira Winter, A.-D. Brunner, J. Cox, M. Mann, MaxQuant.Live enables global targeting of more than 25,000 peptides, Mol. Cell. Proteom. 18 (5) (2019) 982–994, https://doi.org/10.1074/mcp.TIR118.001131.

[65] S.H. Vohr, C.F. Buen Abad Najar, B. Shapiro, R.E. Green, A method for positive forensic identification of samples from extremely low-coverage sequence data, BMC Genom. 16 (2015) 1034, https://doi.org/10.1186/s12864-015-2241-6.

[66] J.K. Eng, B.C. Searle, K.R. Clauser, D.L. Tabb, A face in the crowd: recognizing peptides through database search, Mol. Cell. Proteom. 10 (11) (2011).

[67] C. Ludwig, L. Gillet, G. Rosenberger, S. Amon, B.C. Collins, R. Aebersold, Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial, Mol. Syst. Biol. 14 (8) (2018), e8126, https://doi.org/10.15252/msb.20178126.

[68] S. Tiwary, R. Levy, P. Gutenbrunner, F. Salinas Soto, K.K. Palaniappan, L. Deming, M. Berndl, A. Brant, P. Cimermancic, J. Cox, High-quality MS/MS spectrum prediction for data-dependent and data-independent acquisition data analysis, Nat. Methods (2019), https://doi.org/10.1038/s41592-019-0427-6.

[69] Y.S. Ting, J.D. Egertson, J.G. Bollinger, B.C. Searle, S.H. Payne, W.S. Noble, M. J. MacCoss, PECAN: library-free peptide detection for data-independent acquisition tandem mass spectrometry data, Nat. Methods 14 (9) (2017) 903–908, https://doi.org/10.1038/nmeth.4390.

[70] P. Picotti, R. Aebersold, Selected reaction monitoring–based proteomics: workflows, potential, pitfalls and future directions, Nat. Methods 9 (6) (2012) 555–566.

[71] S. Gallien, E. Duriez, C. Crone, M. Kellmann, T. Moehring, B. Domon, Targeted proteomic quantification on quadrupole-orbitrap mass spectrometer, Mol. Cell. Proteom. MCP 11 (12) (2012) 1709–1723, https://doi.org/10.1074/mcp.O112.019802.

[72] N. Rauniyar, Parallel reaction monitoring: a targeted experiment performed using high resolution and high mass accuracy mass spectrometry, Int. J. Mol. Sci. 16 (12) (2015) 28566–28581, https://doi.org/10.3390/ijms161226120.

[73] A.C. Peterson, J.D. Russell, D.J. Bailey, M.S. Westphall, J.J. Coon, Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics, Mol. Cell. Proteom. MCP 11 (11) (2012) 1475–1488, https://doi.org/10.1074/mcp.O112.020131.

[74] T. Scientific Working Group for Forensic, Scientific Working Group for Forensic Toxicology (SWGTOX) standard practices for method validation in forensic toxicology, J. Anal. Toxicol. 37 (7) (2013) 452–474, https://doi.org/10.1093/jat/bkt054.

[75] WADA Laboratory Committee, Identification Criteria For Qualitative Assays Incorporating Column Chromatography And Mass Spectrometry, WADA Tech. Doc., 2010, pp. 1–9 (TD2010IDCR).

[76] M.Y. Brusniak, C.S. Chu, U. Kusebauch, M.J. Sartain, J.D. Watts, R.L. Moritz, An assessment of current bioinformatic solutions for analyzing LC-MS data acquired by selected reaction monitoring technology, Proteomics 12 (8) (2012) 1176–1184, https://doi.org/10.1002/pmic.201100571.

[77] A.J. Sherier, R.E. Kieser, N.M.M. Novroski, F.R. Wendt, J.L. King, A.E. Woerner, A. Ambers, P. Garofano, B. Budowle, Copan microFLOQ® Direct Swab collection of bloodstains, saliva, and semen on cotton cloth, Int. J. Leg. Med. 134 (1) (2020) 45–54, https://doi.org/10.1007/s00414-019-02081-6.

[78] J.A. Cham Mead, L. Bianco, C. Bessant, Free computational resources for designing selected reaction monitoring transitions, Proteomics 10 (6) (2010) 1106–1126, https://doi.org/10.1002/pmic.200900396.

[79] C.M. Colangelo, L. Chung, C. Bruce, K.H. Cheung, Review of software tools for design and analysis of large scale MRM proteomic datasets, Methods 61 (3) (2013) 287–298, https://doi.org/10.1016/j.ymeth.2013.05.004.

[80] K.R. McLeish, M.L. Merchant, J.B. Klein, R.A. Ward, proteomic approaches to fundamental questions about neutrophil biology, J. Leukoc. Biol. 94 (4) (2013) 683–692.

[81] L. Wu, C.A. Miller, J. Hsiao, T. Chu, B. Zekavat, A. Fandino, Extended Mass Range Triple Quadrupole for Routine Analysis of High Mass-to-charge Peptide Ions, Application Note, Agilent Technologies, Inc, Santa Clara, CA, 2018.

[82] N.A. Kulak, G. Pichler, I. Paron, N. Nagaraj, M. Mann, Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells, Nat. Methods 11 (3) (2014) 319–324, https://doi.org/10.1038/nmeth.2834.

[83] D.E. Briggs, R.P. Evershed, M.J. Lockheart, The biomolecular paleontology of continental fossils, Paleobiology 26 (sp4) (2000) 169–193.

[84] A. Torti, M.A. Lever, B.B. Jørgensen, Origin, dynamics, and implications of extracellular DNA pools in marine sediments, Mar. Genom. 24 (2015) 185–196.

[85] T. van der Valk, P. Pecnerova, D. Diez-Del-Molino, A. Bergstrom, J. Oppenheimer, S. Hartmann, G. Xenikoudakis, J.A. Thomas, M. Dehasque, E. Saglican, F. R. Fidan, I. Barnes, S. Liu, M. Somel, P.D. Heintzman, P. Nikolskiy, B. Shapiro, P. Skoglund, M. Hofreiter, A.M. Lister, A. Gotherstrom, L. Dalen, Million-year-old DNA sheds light on the genomic history of mammoths, Nature (2021), https://doi.org/10.1038/s41586-021-03224-9.

[86] N. Rybczynski, J.C. Gosse, C.R. Harington, R.A. Wogelius, A.J. Hidy, M. Buckley, Mid-Pliocene warm-period deposits in the High Arctic yield insight into camel evolution, Nat. Commun. 4 (2013) 1550, https://doi.org/10.1038/ncomms2516.

[87] M. Pal Chowdhury, R. Wogelius, P.L. Manning, L. Metz, L. Slimak, M. Buckley, Collagen deamidation in archaeological bone as an assessment for relative decay rates, Archaeometry 61 (6) (2019) 1382–1398, https://doi.org/10.1111/arcm.12492.

[88] F. Welker, J. Ramos-Madrigal, M. Kuhlwilm, W. Liao, P. Gutenbrunner, M. de Manuel, D. Samodova, M. Mackie, M.E. Allentoft, A.M. Bacon, M.J. Collins, J. Cox, C. Lalueza-Fox, J.V. Olsen, F. Demeter, W. Wang, T. Marques-Bonet, E. Cappellini, Enamel proteome shows that Gigantopithecus was an early diverging pongine, Nature 576 (2019) 262–265, https://doi.org/10.1038/s41586-019-1728-8.

[89] E. Cappellini, F. Welker, L. Pandolfi, J.R. Madrigal, A.K. Fotakis, D. Lyon, V.J. Moreno Mayar, M. Bukhsianidze, R.R. Jersie-Christensen, M. Mackie, A. Ginolhac, R. Ferring, M. Tappen, E. Palkopoulou, D. Samodova, P.L. Rüther, M. R. Dickinson, T. Stafford, Y.L. Chan, A. Götherström, S.K.S.S. Nathan, P. D. Heintzman, J.D. Kapp, I. Kirillova, Y. Moodley, J. Agusti, R.-D. Kahlke, G. Kiladze, B. Martínez–Navarro, S. Liu, M.S. Velasco, M.-H.S. Sinding, C. D. Kelstrup, M.E. Allentoft, A. Krogh, L. Orlando, K. Penkman, B. Shapiro, L. Rook, L. Dalén, M.T.P. Gilbert, J.V. Olsen, D. Lordkipanidze, E. Willerslev, Early Pleistocene enamel proteome sequences from Dmanisi resolve Stephanorhinus phylogeny, Nature 574 (7776) (2018) 103–107, https://doi.org/10.1038/s41586-019-1555-y.

[90] J. Hendy, F. Welker, B. Demarchi, C. Speller, C. Warinner, M.J. Collins, A guide to ancient protein studies, Nat. Ecol. Evol. 2 (5) (2018) 791–799, https://doi.org/10.1038/s41559-018-0510-x.

[91] B. Demarchi, S. Hall, T. Roncal-Herrero, C.L. Freeman, J. Woolley, M.K. Crisp, J. Wilson, A. Fotakis, R. Fischer, B.M. Kessler, R. Rakownikow Jersie-Christensen, J.V. Olsen, J. Haile, J. Thomas, C.W. Marean, J. Parkington, S. Presslee, J. Lee-Thorp, P. Ditchfield, J.F. Hamilton, M.W. Ward, C.M. Wang, M.D. Shaw, T. Harrison, M. Dominguez-Rodrigo, R. MacPhee, A. Kwekason, M. Ecker, L. Kolska Horwitz, M. Chazan, R. Kroger, J. Thomas-Oates, J.H. Harding, E. Cappellini, K. Penkman, M.J. Collins, Protein sequences bound to mineral surfaces persist into deep time, Elife 5 (2016), e17092, https://doi.org/10.7554/eLife.17092.

[92] M.H. Schweitzer, Z. Suo, R. Avci, J.M. Asara, M.A. Allen, F.T. Arce, J.R. Horner, Analyses of soft tissue from Tyrannosaurus rex suggest the presence of protein, Science 316 (5822) (2007) 277–280, https://doi.org/10.1126/science.1138709.

[93] M. Buckley, A. Walker, S.Y. Ho, Y. Yang, C. Smith, P. Ashton, J.T. Oates, E. Cappellini, H. Koon, K. Penkman, B. Elsworth, D. Ashford, C. Solazzo, P. Andrews, J. Strahler, B. Shapiro, P. Ostrom, H. Gandhi, W. Miller, B. Raney, M. I. Zylber, M.T. Gilbert, R.V. Prigodich, M. Ryan, K.F. Rijsdijk, A. Janoo, M. J. Collins, Comment on "Protein sequences from mastodon and Tyrannosaurus rex revealed by mass spectrometry", Science 319 (5859) (2008) 33, https://doi.org/10.1126/science.1147046 (author reply 33).

[94] J.M. Asara, M.H. Schweitzer, L.M. Freimark, M. Phillips, L.C. Cantley, Protein sequences from mastodon and Tyrannosaurus rex revealed by mass spectrometry, Science 316 (5822) (2007) 280–285, https://doi.org/10.1126/science.1137614 (https://doi.org/316/5822/280 [pii]).

[95] M.H. Schweitzer, W. Zheng, C.L. Organ, R. Avci, Z. Suo, L.M. Freimark, V. S. Lebleu, M.B. Duncan, M.G. Vander Heiden, J.M. Neveu, W.S. Lane, J.S. Cottrell, J.R. Horner, L.C. Cantley, R. Kalluri, J.M. Asara, Biomolecular characterization and protein sequences of the Campanian hadrosaur B. canadensis, Science 324 (5927) (2009) 626–631, https://doi.org/10.1126/science.1165069 (https://doi.org/324/5927/626 [pii]).

[96] M. Bern, B.S. Phinney, D. Goldberg, Reanalysis of Tyrannosaurus rex mass spectra, J. Proteome Res. 8 (9) (2009) 4328–4332, https://doi.org/10.1021/pr900349r.

[97] T.P. Cleland, E.R. Schroeter, L. Zamdborg, W. Zheng, J.E. Lee, J.C. Tran, M. Bern, M.B. Duncan, V.S. Lebleu, D.R. Ahlf, P.M. Thomas, R. Kalluri, N.L. Kelleher, M. H. Schweitzer, Mass spectrometry and antibody-based characterization of blood vessels from Brachylophosaurus canadensis, J. Proteome Res. 14 (12) (2015) 5252–5262, https://doi.org/10.1021/acs.jproteome.5b00675.

[98] E.R. Schroeter, C.J. DeHart, T.P. Cleland, W. Zheng, P.M. Thomas, N.L. Kelleher, M. Bern, M.H. Schweitzer, Expansion for the Brachylophosaurus canadensis Collagen I sequence and additional evidence of the preservation of cretaceous protein, J. Proteome Res. 16 (2) (2017) 920–932, https://doi.org/10.1021/acs.jproteome.6b00873.

[99] X. Li, C. Lin, P.B. O'Connor, Glutamine deamidation: differentiation of glutamic acid and gamma-glutamic acid in peptides by electron capture dissociation, Anal. Chem. 82 (9) (2010) 3606–3615, https://doi.org/10.1021/ac9028467.

[100] T.P. Cleland, E.R. Schroeter, M.H. Schweitzer, Biologically and diagenetically derived peptide modifications in moa collagens, Proc. Biol. Sci. R. Soc. 282 (1808) (2015), 20150015, https://doi.org/10.1098/rspb.2015.0015.

[101] N.E. Robinson, Protein deamidation, Proc. Natl. Acad. Sci. USA 99 (8) (2002) 5283–5288, https://doi.org/10.1073/pnas.082102799.

[102] E.R. Schroeter, T.P. Cleland, Glutamine deamidation: an indicator of antiquity, or preservational quality? Rapid Commun. Mass Spectrom. 30 (2) (2016) 251–255, https://doi.org/10.1002/rcm.7445.

[103] T. Geiger, S. Clarke, Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation, J. Biol. Chem. 262 (2) (1987) 785–794.

[104] M.J. Collins, E.R. Waite, A.C. van Duin, Predicting protein decomposition: the case of aspartic-acid racemization kinetics, Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 354 (1379) (1999) 51–64, https://doi.org/10.1098/rstb.1999.0359.

[105] K. Kaiser, R. Benner, Hydrolysis-induced racemization of amino acids, Limnol. Oceanogr. Methods 3 (8) (2005) 318–325.

[106] G. Kim, S.J. Weiss, R.L. Levine, Methionine oxidation and reduction in proteins, Biochim. Biophys. Acta 1840 (2) (2014) 901–905, https://doi.org/10.1016/j.bbagen.2013.04.038.

[107] C.H. Lai, C.N. Lee, K.J. Bai, Y.L. Yang, K.J. Chuang, S.M. Wu, H.C. Chuang, Protein oxidation and degradation caused by particulate matter, Sci. Rep. 6 (2016) 33727, https://doi.org/10.1038/srep33727.

[108] K. Artemenko, J. Mi, J. Bergquist, Mass-spectrometry-based characterization of oxidations in proteins, Free Radic. Res. 49 (5) (2015) 477–493, https://doi.org/10.3109/10715762.2015.1023795.

[109] J. Jeong, Y. Jung, S. Na, J. Jeong, E. Lee, M.S. Kim, S. Choi, D.H. Shin, E. Paek, H. Y. Lee, K.J. Lee, Novel oxidative modifications in redox-active cysteine residues, Mol. Cell. Proteom. MCP 10 (3) (2011), https://doi.org/10.1074/mcp.M110.000513 (M110 000513).

[110] R.H. Nagaraj, I.N. Shipanova, F.M. Faust, Protein cross-linking by the Maillard reaction. Isolation, characterization, and in vivo detection of a lysine-lysine cross-link derived from methylglyoxal, J. Biol. Chem. 271 (32) (1996) 19338–19345.

[111] Q. Zhang, J.M. Ames, R.D. Smith, J.W. Baynes, T.O. Metz, A perspective on the Maillard reaction and the analysis of protein glycation by mass spectrometry: probing the pathogenesis of chronic disease, J. Proteome Res. 8 (2) (2009) 754–769, https://doi.org/10.1021/pr800858h.

[112] H.N. Poinar, The genetic secrets some fossils hold, Acc. Chem. Res. 35 (8) (2002) 676–684.

[113] J.W. Baynes, The Maillard hypothesis on aging: time to focus on DNA, Ann. N. Y. Acad. Sci. 959 (1) (2002) 360–367, https://doi.org/10.1111/j.1749-6632.2002.tb02107.x.

[114] R. Bucala, A. Cerami, DNA-advanced glycosylation, in: R. Ikan (Ed.), The Maillard Reaction: Consequences for the Chemical and Life Sciences, John Wiley and Sons, 1996, pp. 161–182.

[115] R. Zamora, F.R. Hidalgo, The Maillard reaction and lipid oxidation, Lipid Technol. 23 (3) (2011) 59–62, https://doi.org/10.1002/lite.201100094.

[116] J. Meltretter, S. Seeber, A. Humeny, C.M. Becker, M. Pischetsrieder, Site-specific formation of Maillard, oxidation, and condensation products from whey proteins during reaction with lactose, J. Agric. Food Chem. 55 (15) (2007) 6096–6103, https://doi.org/10.1021/jf0705567.

[117] R. Aebersold, J.N. Agar, I.J. Amster, M.S. Baker, C.R. Bertozzi, E.S. Boja, C. E. Costello, B.F. Cravatt, C. Fenselau, B.A. Garcia, Y. Ge, J. Gunawardena, R. C. Hendrickson, P.J. Hergenrother, C.G. Huber, A.R. Ivanov, O.N. Jensen, M. C. Jewett, N.L. Kelleher, L.L. Kiessling, N.J. Krogan, M.R. Larsen, J.A. Loo, R. R. Ogorzalek Loo, E. Lundberg, M.J. MacCoss, P. Mallick, V.K. Mootha, M. Mrksich, T.W. Muir, S.M. Patrie, J.J. Pesavento, S.J. Pitteri, H. Rodriguez, A. Saghatelian, W. Sandoval, H. Schluter, S. Sechi, S.A. Slavoff, L.M. Smith, M. P. Snyder, P.M. Thomas, M. Uhlen, J.E. Van Eyk, M. Vidal, D.R. Walt, F.M. White, E.R. Williams, T. Wohlschlager, V.H. Wysocki, N.A. Yates, N.L. Young, B. Zhang, How many human proteoforms are there? Nat. Chem. Biol. 14 (3) (2018) 206–214, https://doi.org/10.1038/nchembio.2576.

[118] T.P. Cleland, K. Voegele, M.H. Schweitzer, Empirical evaluation of bone extraction protocols, PLoS One 7 (2) (2012), e31443, https://doi.org/10.1371/journal.pone.0031443.

[119] J.L. Proc, M.A. Kuzyk, D.B. Hardie, J. Yang, D.S. Smith, A.M. Jackson, C.E. Parker, C.H. Borchers, A quantitative study of the effects of chaotropic agents, surfactants, and solvents on the digestion efficiency of human plasma proteins by trypsin, J. Proteome Res. 9 (10) (2010) 5422–5437, https://doi.org/10.1021/pr100656u.

[120] J.R. Wisniewski, A. Zougman, N. Nagaraj, M. Mann, Universal sample preparation method for proteome analysis, Nat. Methods 6 (5) (2009) 359–362, https://doi.org/10.1038/nmeth.1322.

[121] S. Heissel, S.J. Frederiksen, J. Bunkenborg, P. Hojrup, Enhanced trypsin on a budget: Stabilization, purification and high-temperature application of inexpensive commercial trypsin for proteomics applications, PLoS One 14 (6) (2019), e0218374, https://doi.org/10.1371/journal.pone.0218374.

[122] R.H. Rice, G.E. Means, W.D. Brown, Stabilization of bovine trypsin by reductive methylation, Biochim. Biophys. Acta 492 (2) (1977) 316–321.

[123] J.V. Olsen, S.-E. Ong, M. Mann, Trypsin cleaves exclusively C-terminal to arginine and lysine residues, Mol. Cell. Proteom. 3 (6) (2004) 608–614.

[124] D.F. Hunt, J.R. Yates, J. Shabanowitz, S. Winston, C.R. Hauer, Protein sequencing by tandem mass spectrometry, Proc. Natl. Acad. Sci. 83 (17) (1986) 6233–6237.

[125] F. Guan, M.A. Robinson, Comprehensive solid-phase extraction of multitudinous bioactive peptides from equine plasma and urine for doping detection, Anal. Chim. Acta 985 (2017) 79–90, https://doi.org/10.1016/j.aca.2017.07.005.

[126] M.A. Gillette, S.A. Carr, Quantitative analysis of peptides and proteins in biomedicine by targeted mass spectrometry, Nat. Methods 10 (1) (2013) 28–34, https://doi.org/10.1038/nmeth.2309.

[127] M. Dupre, B. Gilquin, F. Fenaille, C. Feraudet-Tarisse, J. Dano, M. Ferro, S. Simon, C. Junot, V. Brun, F. Becher, Multiplex quantification of protein toxins in human biofluids and food matrices using immunoextraction and high-resolution targeted mass spectrometry, Anal. Chem. 87 (16) (2015) 8473–8480, https://doi.org/10.1021/acs.analchem.5b01900.

[128] N.L. Anderson, N.G. Anderson, L.R. Haines, D.B. Hardie, R.W. Olafson, T. W. Pearson, Mass spectrometric quantitation of peptides and proteins using Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA), J. Proteome Res. 3 (2) (2004) 235–244.

[129] C. Colleary, N.C. Little, T.P. Cleland, Microwave-assisted acid hydrolysis for whole-bone proteomics and paleoproteomics, Rapid Commun. Mass Spectrom. RCM 34 (2) (2020), e8568, https://doi.org/10.1002/rcm.8568.

[130] T.P. Cleland, Human bone paleoproteomics utilizing the single-pot, solid-phase-enhanced sample preparation method to maximize detected proteins and reduce

humics, J. Proteome Res. 17 (11) (2018) 3976–3983, https://doi.org/10.1021/acs.jproteome.8b00637.

[131] E. Jankovska, M. Svitek, K. Holada, J. Petrak, Affinity depletion versus relative protein enrichment: a side-by-side comparison of two major strategies for increasing human cerebrospinal fluid proteome coverage, Clin. Proteom. 16 (1) (2019) 9.

[132] H. Keshishian, T. Addona, M. Burgess, E. Kuhn, S.A. Carr, Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution, Mol. Cell. Proteom. MCP 6 (12) (2007) 2212–2229, https://doi.org/10.1074/mcp.M700354-MCP200.

[133] T.P. Cleland, Solid digestion of demineralized bone as a method to access potentially insoluble proteins and post-translational modifications, J. Proteome Res. 17 (1) (2018) 536–542, https://doi.org/10.1021/acs.jproteome.7b00670.

[134] N. Procopio, M. Buckley, Minimizing laboratory-induced decay in bone proteomics, J. Proteome Res. 16 (2) (2017) 447–458, https://doi.org/10.1021/acs.jproteome.6b00564.

[135] X. Jiang, M. Ye, G. Liu, S. Feng, L. Cui, H. Zou, Method development of efficient protein extraction in bone tissue for proteome analysis, J. Proteome Res. 6 (6) (2007) 2287–2294, https://doi.org/10.1021/pr070056t.

[136] Z.C. Goecker, M.R. Salemi, N. Karim, B.S. Phinney, R.H. Rice, G.J. Parker, Optimal processing for proteomic genotyping of single human hairs, Forensic Sci. Int.: Genet. 47 (102314) (2020), https://doi.org/10.1016/j.fsigen.2020.102314.

[137] R.H. Rice, B.P. Durbin-Johnson, S.M. Mann, M. Salemi, S. Urayama, D.M. Rocke, B.S. Phinney, J.P. Sundberg, Corneocyte proteomics: applications to skin biology and dermatology, Exp. Dermatol. 27 (8) (2018) 931–938, https://doi.org/10.1111/exd.13756.

[138] R.H. Rice, Proteomic analysis of hair shaft and nail plate, J. Cosmet. Sci. 62 (2011) 229–236.

[139] R.H. Rice, Y. Xia, R.J. Alvarado, B.S. Phinney, Proteomic analysis of human nail plate, J. Proteome Res. 9 (12) (2010) 6752–6758, https://doi.org/10.1021/pr1009349.

[140] S.A. Sterling, K.E. Mason, D.S. Anex, G.J. Parker, B. Hart, M. Prinz, Combined DNA typing and protein identification from unfired brass cartridges, J. Forensic Sci. Press (2019), https://doi.org/10.1111/1556-4029.14042.

[141] S. Kranes, S.A. Sterling, K. Mason, D. Anex, B. Hart, G. Parker, M. Prinz, Simultaneous DNA and protein extraction using trypsin, Forensic Sci. Int. Genet. Suppl. Ser. 6 (2017) e203–e204, https://doi.org/10.1016/j.fsigss.2017.09.081.

[142] L.A. Catlin, R.M. Chou, Z.C. Goecker, L.A. Mullins, D.S.B.S.S. Silva, R.R. Spurbeck, G.J. Parker, C.M. Bartling, Demonstration of a mitochondrial DNA-compatible workflow for genetically variant peptide identification from human hair samples, Forensic Sci. Int. Genet. 43 (2019), 102148, https://doi.org/10.1016/j.fsigen.2019.102148.

[143] K.Q. Schulte, F.C. Hewitt, T.E. Manley, A.J. Reed, M. Baniasad, N.C. Albright, M. E. Powals, D.S. LeSassier, A.R. Smith, L. Zhang, L.W. Allen, B.C. Ludolph, K. L. Weber, A.E. Woerner, M.A. Freitas, M.W. Gardner, Fractionation of DNA and protein from individual latent fingerprints for forensic analysis, Forensic Sci. Int. Genet. 50 (2020), 102405, https://doi.org/10.1016/j.fsigen.2020.102405.

[144] P.G. Hains, P.J. Robinson, The impact of commonly used alkylating agents on artifactual peptide modification, J. Proteome Res. 16 (9) (2017) 3443–3447, https://doi.org/10.1021/acs.jproteome.7b00022.

[145] T. Müller, D. Winter, Systematic evaluation of protein reduction and alkylation reveals massive unspecific side effects by iodine-containing reagents, Mol. Cell. Proteom. (2017), https://doi.org/10.1074/mcp.M116.064048.

[146] T. Masuda, M. Tomita, Y. Ishihama, Phase transfer surfactant-aided trypsin digestion for membrane proteome analysis, J. Proteome Res. 7 (2) (2008) 731–740, https://doi.org/10.1021/pr700658q.

[147] Y.Q. Yu, M. Gilar, P.J. Lee, E.S. Bouvier, J.C. Gebler, Enzyme-friendly, mass spectrometry-compatible surfactant for in-solution enzymatic digestion of proteins, Anal. Chem. 75 (21) (2003) 6023–6028, https://doi.org/10.1021/ac0346196.

[148] T. Borja, N. Karim, Z. Goecker, M. Salemi, B. Phinney, M. Naeem, R. Rice, G. Parker, Proteomic genotyping of fingermark donors with genetically variant peptides, Forensic Sci. Int. Genet. 42 (2019) 21–30, https://doi.org/10.1016/j.fsigen.2019.05.005.

[149] S. Moggridge, P.H. Sorensen, G.B. Morin, C.S. Hughes, Extending the compatibility of the SP3 paramagnetic bead processing approach for proteomics, J. Proteome Res. 17 (4) (2018) 1730–1740, https://doi.org/10.1021/acs.jproteome.7b00913.

[150] S. Berensmeier, Magnetic particles for the separation and purification of nucleic acids, Appl. Microbiol. Biotechnol. 73 (3) (2006) 495–504, https://doi.org/10.1007/s00253-006-0675-0.

[151] A. Kwasnik, C. Tonry, A. McArdle, A. Qasim Butt, R. Inzitari, S.R. Pennington, Proteomes, their composition and their sources, in: H. Mirzaei, M. Carrasco (Eds.), Modern Proteomics – Sample Preparation, Analysis and Practical Applications, Springer Nature, 2016, pp. 3–22.

[152] N. Grassl, N.A. Kulak, G. Pichler, P.E. Geyer, J. Jung, S. Schubert, P. Sinitcyn, J. Cox, M. Mann, Ultra-deep and quantitative saliva proteome reveals dynamics of the oral microbiome, Genome Med. 8 (1) (2016) 44, https://doi.org/10.1186/s13073-016-0293-0.

[153] A. Murr, C. Pink, E. Hammer, S. Michalik, V.M. Dhople, B. Holtfreter, U. Volker, T. Kocher, M. Gesell Salazar, Cross-sectional association of salivary proteins with age, sex, body mass index, smoking, and education, J. Proteome Res. 16 (6) (2017) 2273–2281, https://doi.org/10.1021/acs.jproteome.7b00133.

[154] H. Yang, B. Zhou, H. Deng, M. Prinz, D. Siegel, Body fluid identification by mass spectrometry, Int. J. Leg. Med. 127 (6) (2013) 1065–1077, https://doi.org/10.1007/s00414-013-0848-1.

[155] P. Danielson, Isolation of Highly Specific Protein Markers for the Identification of Biological Stains: Adapting Comparative Proteomics to Forensics, in: Do Justice (Ed.) Department of Justice, 2011, pp. 1–37.

[156] M. Jodar, E. Sendler, S.A. Krawetz, The protein and transcript profiles of human semen, Cell Tissue Res 363 (1) (2016) 85–96, https://doi.org/10.1007/s00441-015-2237-1.

[157] K.Y. Fung, L.M. Glode, S. Green, M.W. Duncan, A comprehensive characterization of the peptide and protein constituents of human seminal fluid, Prostate 61 (2) (2004) 171–181, https://doi.org/10.1002/pros.20089.

[158] E.P. Diamandis, W.P. Arnett, G. Foussias, H. Pappas, S. Ghandi, D.N. Melegos, B. Mullen, H. Yu, J. Srigley, K. Jarvi, Seminal plasma biochemical markers and their association with semen analysis findings, Urology 53 (3) (1999) 596–603.

[159] B. Pilch, M. Mann, Large-scale and high-confidence proteomic analysis of human seminal plasma, Genome Biol. 7 (5) (2006) R40, https://doi.org/10.1186/gb-2006-7-5-r40.

[160] A.P. Drabovich, P. Saraon, K. Jarvi, E.P. Diamandis, Seminal plasma as a diagnostic fluid for male reproductive system disorders, Nat. Rev. Urol. 11 (5) (2014) 278–288, https://doi.org/10.1038/nrurol.2014.74.

[161] K. Van Steendam, M. De Ceuleneer, M. Dhaenens, D. Van Hoofstat, D. Deforce, Mass spectrometry-based proteomics as a tool to identify biological matrices in forensic science, Int. J. Leg. Med. 127 (2) (2013) 287–298, https://doi.org/10.1007/s00414-012-0747-x.

[162] A. Igoh, Y. Doi, K. Sakurada, Identification and evaluation of potential forensic marker proteins in vaginal fluid by liquid chromatography/mass spectrometry, Anal. Bioanal. Chem. 407 (23) (2015) 7135–7144, https://doi.org/10.1007/s00216-015-8877-x.

[163] S. Kamanna, J. Henry, N. Voelcker, A. Linacre, K.P. Kirkbride, A complementary forensic 'proteo-genomic' approach for the direct identification of biological fluid traces under fingernails, Anal. Bioanal. Chem. 410 (2018) 6165–6175, https://doi.org/10.1007/s00216-018-1223-3.

[164] A. Thurman, T. Jacot, J. Melendez, T. Kimble, M. Snead, R. Jamshidi, A. Wheeless, D.F. Archer, G.F. Doncel, C. Mauck, Assessment of the vaginal residence time of biomarkers of semen exposure, Contraception 94 (5) (2016) 512–520, https://doi.org/10.1016/j.contraception.2016.05.012.

[165] K.M. Legg, R. Powell, N. Reisdorph, R. Reisdorph, P.B. Danielson, Discovery of highly specific protein markers for the identification of biological stains, Electrophoresis 35 (21–22) (2014) 3069–3078, https://doi.org/10.1002/elps.201400125.

[166] Y. Yu, I. Prassas, C.M. Muytjens, E.P. Diamandis, Proteomic and peptidomic analysis of human sweat with emphasis on proteolysis, J. Proteom. 155 (2017) 40–48, https://doi.org/10.1016/j.jprot.2017.01.005.

[167] M.M. Raiszadeh, M.M. Ross, P.S. Russo, M.A. Schaepper, W. Zhou, J. Deng, D. Ng, A. Dickson, C. Dickson, M. Strom, C. Osorio, T. Soeprono, J.D. Wulfkuhle, E. F. Petricoin, L.A. Liotta, W.M. Kirsch, Proteomic analysis of eccrine sweat: implications for the discovery of schizophrenia biomarker proteins, J. Proteome Res. 11 (4) (2012) 2127–2139, https://doi.org/10.1021/pr2007957.

[168] M. Pieri, A. Silvestre, M. De Cicco, G. Mamone, E. Capasso, F. Addeo, G. Picariello, Mass spectrometry-based proteomics for the forensic identification of vomit traces, J. Proteom. 209 (2019), 103524, https://doi.org/10.1016/j.jprot.2019.103524.

[169] R.D. Alves, J.A. Demmers, K. Bezstarosti, B.C. van der Eerden, J.A. Verhaar, M. Eijken, J.P. van Leeuwen, Unraveling the human bone microenvironment beyond the classical extracellular matrix proteins: a human bone protein library, J. Proteome Res. 10 (10) (2011) 4725–4733, https://doi.org/10.1021/pr200522n.

[170] N. Procopio, A. Williams, A.T. Chamberlain, M. Buckley, Forensic proteomics for the evaluation of the post-mortem decay in bones, J. Proteom. 177 (2018) 21–30, https://doi.org/10.1016/j.jprot.2018.01.016.

[171] N. Procopio, A.T. Chamberlain, M. Buckley, Intra- and interskeletal proteome variations in fresh and buried bones, J. Proteome Res. 16 (5) (2017) 2016–2029, https://doi.org/10.1021/acs.jproteome.6b01070.

[172] E.R. Schroeter, C.J. DeHart, M.H. Schweitzer, P.M. Thomas, N.L. Kelleher, Bone protein "extractomics": comparing the efficiency of bone protein extractions of Gallus gallus in tandem mass spectrometry, with an eye towards paleoproteomics, PeerJ 4 (2016), e2603, https://doi.org/10.7717/peerj.2603.

[173] S.M. Lyon, A. Mayampurath, M.R. Rogers, D.J. Wolfgeher, S.M. Fisher, S. L. Volchenboum, T.C. He, R.R. Reid, A method for whole protein isolation from human cranial bone, Anal. Biochem. (2016), https://doi.org/10.1016/j.ab.2016.09.021.

[174] M. Jagr, P. Ergang, S. Pataridis, M. Kolrosova, M. Bartos, I. Miksik, Proteomic analysis of dentin-enamel junction and adjacent protein-containing enamel matrix layer of healthy human molar teeth, Eur. J. Oral Sci. 127 (2) (2019) 112–121, https://doi.org/10.1111/eos.12594.

[175] C. Froment, M. Hourset, N. Saenz-Oyhereguy, E. Mouton-Barbosa, C. Willmann, C. Zanolli, R. Esclassan, R. Donat, C. Theves, O. Burlet-Schiltz, C. Mollereau, Analysis of 5000 year-old human teeth using optimized large-scale and targeted proteomics approaches for detection of sex-specific peptides, J. Proteom. 211 (2020), 103548, https://doi.org/10.1016/j.jprot.2019.103548.

[176] V.C. Wasinger, D. Curnoe, S. Bustamante, R. Mendoza, R. Shoocongdej, L. Adler, A. Baker, K. Chintakanon, C. Boel, P.S.C. Tacon, Analysis of the preserved amino acid bias in peptide profiles of iron age teeth from a tropical environment enable sexing of individuals using amelogenin MRM, Proteomics 19 (5) (2019), e1800341, https://doi.org/10.1002/pmic.201800341.

[177] M.J. Hubbard, J.C. Kon, Proteomic analysis of dental tissues, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 771 (1–2) (2002) 211–220, https://doi.org/S1570023202000429 [pii].

[178] D.R. Green, F. Schulte, K.H. Lee, M.K. Pugach, M. Hardt, F.B. Bidlack, Mapping the tooth enamel proteome and amelogenin phosphorylation onto mineralizing porcine tooth crowns, Front. Physiol. 10 (2019) 925, https://doi.org/10.3389/fphys.2019.00925.

[179] F. Welker, J. Ramos-Madrigal, P. Gutenbrunner, M. Mackie, S. Tiwary, R. Rakownikow Jersie-Christensen, C. Chiva, M.R. Dickinson, M. Kuhlwilm, M. de Manuel, P. Gelabert, M. Martinon-Torres, A. Margvelashvili, J.L. Arsuaga, E. Carbonell, T. Marques-Bonet, K. Penkman, E. Sabido, J. Cox, J.V. Olsen, D. Lordkipanidze, F. Racimo, C. Lalueza-Fox, J.M. Bermudez de Castro, E. Willerslev, E. Cappellini, The dental proteome of Homo antecessor, Nature 580 (7802) (2020) 235–238, https://doi.org/10.1038/s41586-020-2153-8.

[180] N.A. Stewart, G.F. Molina, J.P.M. Issa, N.A. Yates, M. Sosovicka, A.R. Vieira, S.R. P. Line, J. Montgomery, R.F. Gerlach, The identification of peptides by nanoLC-MS/MS from human surface tooth enamel following a simple acid etch extraction, RSC Adv. 6 (66) (2016) 61673–61679, https://doi.org/10.1039/C6RA05120K.

[181] G.A. Castiblanco, D. Rutishauser, L.L. Ilag, S. Martignon, J.E. Castellanos, W. Mejia, Identification of proteins from human permanent erupted enamel, Eur. J. Oral Sci. 123 (6) (2015) 390–395, https://doi.org/10.1111/eos.12214.

[182] E.S. Park, H.S. Cho, T.G. Kwon, S.N. Jang, S.H. Lee, C.H. An, H.I. Shin, J.Y. Kim, J. Y. Cho, Proteomics analysis of human dentin reveals distinct protein expression profiles, J. Proteome Res. 8 (3) (2009) 1338–1346, https://doi.org/10.1021/pr801065s, 10.1021/pr801065s [pii].

[183] G.J. Parker, J.M. Yip, J.W. Eerkens, M. Salemi, B. Durbin-Johnson, C. Kiesow, R. Haas, J.E. Buikstra, H. Klaus, L.A. Regan, D.M. Rocke, B.S. Phinney, Sex estimation using sexually dimorphic amelogenin protein fragments in human enamel, J. Archaeol. Sci. 101 (2019) 169–180, https://doi.org/10.1016/j.jas.2018.08.011.

[184] C.M. Nielsen-Marsh, C. Stegemann, R. Hoffmann, T. Smith, R. Feeney, M. Toussaint, K. Harvati, E. Panagopoulou, J.-J. Hublin, M.P. Richards, Extraction and sequencing of human and Neanderthal mature enamel proteins using MALDI-TOF/TOF MS, J. Archaeol. Sci. 36 (8) (2009) 1758–1763.

[185] I.M. Porto, H.J. Laure, F.B. de Sousa, J.C. Rosa, R.F. Gerlach, Techniques for the recovery of small amounts of mature enamel proteins, J. Archaeol. Sci. 38 (2011) 3596–3604.

[186] F. Chu, K.E. Mason, D.S. Anex, A.D. Jones, B.R. Hart, Proteomic characterization of damaged single hairs recovered after an explosion for protein-based human identification, J. Proteome Res. (2020), https://doi.org/10.1021/acs.jproteome.0c00102.

[187] R.N. Franklin, N. Karim, Z.C. Goecker, B.P. Durbin-Johnson, R.H. Rice, G. J. Parker, Proteomic genotyping: using mass spectrometry to infer SNP genotypes in pigmented and non-pigmented hair, Forensic Sci. Int. 310 (2020), 110200, https://doi.org/10.1016/j.forsciint.2020.110200.

[188] S.S. Adav, R.S. Subbaiaih, S.K. Kerk, A.Y. Lee, H.Y. Lai, K.W. Ng, S.K. Sze, A. Schmidtchen, Studies on the proteome of human hair - identification of histones and deamidated keratins, Sci. Rep. 8 (1) (2018) 1599, https://doi.org/10.1038/s41598-018-20041-9.

[189] L. Wang, W. Xu, L. Cao, T. Tian, M. Yang, Z. Li, F. Ping, W. Fan, Differential expression of proteins associated with the hair follicle cycle - proteomics and bioinformatics analyses, PLoS One 11 (1) (2016), e0146791, https://doi.org/10.1371/journal.pone.0146791.

[190] Y.J. Lee, R.H. Rice, Y.M. Lee, Proteome analysis of human hair shaft: from protein identification to posttranslational modification, Mol. Cell. Proteom. MCP 5 (5) (2006) 789–800.

[191] F. Lei, J. Li, L. Shan-Fei, Z. Jian, L. Hai-Bo, J. An-Quan, Y. Jian, W. Gui-Qiang, L. Cai-Xia, Development and validation of protein-based forensic ancestry inference method using hair shaft proteome, Prog. Biochem. Biophys. 46 (1) (2019) 81–88, https://doi.org/10.16476/j.pibb.2018.0179.

[192] H.A. Adeola, J.C. Van Wyk, A. Arowolo, R.M. Ngwanya, K. Mkentane, N. P. Khumalo, Emerging diagnostic and therapeutic potentials of human hair proteomics, Proteom. Appl. 12 (2) (2018), 1700048.

[193] N. Karim, B. Durbin-Johnson, D.M. Rocke, M. Salemi, B.S. Phinney, M. Naeem, R. H. Rice, Proteomic manifestations of genetic defects in autosomal recessive congenital ichthyosis, J. Proteom. 201 (2019) 104–109, https://doi.org/10.1016/j.jprot.2019.04.007.

[194] D.S. LeSassier, K.Q. Schulte, T.E. Manley, A.R. Smith, M.L. Powals, N.C. Albright, B.C. Ludolph, K.L. Weber, A.E. Woerner, M.W. Gardner, F.C. Hewitt, Artificial fingerprints for cross-comparison of forensic DNA and protein recovery methods, PLoS One 14 (10) (2019), e0223170, https://doi.org/10.1371/journal.pone.0223170.

[195] L.O. Brandt, A.L. Schmidt, U. Mannering, M. Sarret, C.D. Kelstrup, J.V. Olsen, E. Cappellini, Species identification of archaeological skin objects from Danish bogs: comparison between mass spectrometry-based peptide sequencing and microscopy-based methods, PLoS One 9 (9) (2014), e106875, https://doi.org/10.1371/journal.pone.0106875.

[196] K. Hollemeyer, W. Altmeyer, E. Heinzle, C. Pitra, Species identification of Oetzi's clothing with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry based on peptide pattern similarities of hair digests, Rapid Commun. Mass Spectrom. RCM 22 (18) (2008) 2751–2767, https://doi.org/10.1002/rcm.3679.

[197] L.L. Meilloura, A. Zazzo, J. Lesur, S. Cersoy, A. Marie, M. Lebon, D. Pleurdeau, S. Zirah, Identification of degraded bone and tooth splinters from arid environmentsusing palaeoproteomics, Palaeogeogr. Palaeoclimatol. Palaeoecol. 551 (2018) 472–482.

[198] F. Welker, Elucidation of cross-species proteomic effects in human and hominin bone proteome identification through a bioinformatics experiment, BMC Evolut. Biol. 18 (1) (2018) 23, https://doi.org/10.1186/s12862-018-1141-1.

[199] R. Sawafuji, E. Cappellini, T. Nagaoka, A.K. Fotakis, R.R. Jersie-Christensen, J. V. Olsen, K. Hirata, S. Ueda, Proteomic profiling of archaeological human bone, R. Soc. Open Sci. 4 (6) (2017), 161004, https://doi.org/10.1098/rsos.161004.

[200] C. Wadsworth, M. Buckley, Characterization of proteomes extracted through collagen-based stable isotope and radiocarbon dating methods, J. Proteome Res. 17 (1) (2018) 429–439, https://doi.org/10.1021/acs.jproteome.7b00624.

[201] M. Buckley, Ancient collagen reveals evolutionary history of the endemic South American 'ungulates', Proc. Biol. Sci. R. Soc. 282 (1806) (2015), 20142671 https://doi.org/10.1098/rspb.2014.2671.

[202] F. Welker, M. Hajdinjak, S. Talamo, K. Jaouen, M. Dannemann, F. David, M. Julien, M. Meyer, J. Kelso, I. Barnes, S. Brace, P. Kamminga, R. Fischer, B. M. Kessler, J.R. Stewart, S. Paabo, M.J. Collins, J.J. Hublin, Palaeoproteomic evidence identifies archaic hominins associated with the Chatelperronian at the Grotte du Renne, Proc. Natl. Acad. Sci. USA 113 (40) (2016) 11162–11167, https://doi.org/10.1073/pnas.1605834113.

[203] C. Wadsworth, M. Buckley, Proteome degradation in fossils: investigating the longevity of protein survival in ancient bone, Rapid Commun. Mass Spectrom. RCM 28 (6) (2014) 605–615, https://doi.org/10.1002/rcm.6821.

[204] T.H. Schmidt-Schultz, M. Schultz, Bone protects proteins over thousands of years: extraction, analysis, and interpretation of extracellular matrix proteins in archeological skeletal remains, Am. J. Phys. Anthropol. 123 (1) (2004) 30–39, https://doi.org/10.1002/ajpa.10308.

[205] N.A. Stewart, R.F. Gerlach, R.L. Gowland, K.J. Gron, J. Montgomery, Sex determination of human remains from peptides in tooth enamel, Proc. Natl. Acad. Sci. USA (2017), https://doi.org/10.1073/pnas.1714926115.

[206] R.F. Gerlach, S. Line, Mature dental enamel: a diverse repository of past and present information (918.2), FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol. (2014).

[207] I.M. Porto, H.J. Laure, R.H. Tykot, F.B. de Sousa, J.C. Rosa, R.F. Gerlach, Recovery and identification of mature enamel proteins in ancient teeth, Eur. J. Oral Sci. 119 (Suppl 1) (2011) 83–87, https://doi.org/10.1111/j.1600-0722.2011.00885.x.

[208] J.E. Elias, S.P. Gygi, Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry, Nat. Methods 4 (3) (2007) 207–214, https://doi.org/10.1038/nmeth1019.

[209] J. Zhang, L. Xin, B. Shan, W. Chen, M. Xie, D. Yuen, W. Zhang, Z. Zhang, G. A. Lajoie, B. Ma, PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification, Mol. Cell. Proteom. MCP 11 (4) (2012), https://doi.org/10.1074/mcp.M111.010587 (M111 010587).

[210] I. Chernukhin, Pepgrep: a tool for peptide MS/MS pattern matching, Genom. Proteom. Bioinform. 11 (2) (2013) 127–132, https://doi.org/10.1016/j.gpb.2013.02.001.

[211] J. Li, Z. Su, Z.Q. Ma, R.J. Slebos, P. Halvey, D.L. Tabb, D.C. Liebler, W. Pao, B. Zhang, A bioinformatics workflow for variant peptide detection in shotgun proteomics, Mol. Cell. Proteom. MCP 10 (5) (2011), https://doi.org/10.1074/mcp.M110.006536 (M110 006536).

[212] D. Fenyo, R.C. Beavis, A method for assessing the statistical significance of mass spectrometry-based protein identifications using general scoring schemes, Anal. Chem. 75 (4) (2003) 768–774.

[213] H. Choi, A.I. Nesvizhskii, False discovery rates and related statistical concepts in mass spectrometry-based proteomics, J. Proteome Res. 7 (1) (2008) 47–50, https://doi.org/10.1021/pr700747q.

[214] A.J. Percy, J. Yang, A.G. Chambers, Y. Mohammed, T. Miliotis, C.H. Borchers, Protocol for standardizing high-to- moderate abundance protein biomarker assessments through an MRM-with-standard-peptides quantitative approach, in: H. Mirzaei, M. Carrasco (Eds.), Modern Proteomics – Sample Preparation, Analysis and Practical Applications, Springer Nature, 2016, 515 to 530.

[215] B.J. Smith, D. Martins-de-Souza, M. Fioramonte, A guide to mass spectrometry-based quantitative proteomics, Methods Mol. Biol. 1916 (2019) 3–39, https://doi.org/10.1007/978-1-4939-8994-2_1.

[216] M. Bader, A systematic approach to standard addition methods in instrumental analysis, J. Chem. Educ. 57 (10) (1980) 703, https://doi.org/10.1021/ed057p703.

[217] H. Lam, E.W. Deutsch, J.S. Eddes, J.K. Eng, S.E. Stein, R. Aebersold, Building consensus spectral libraries for peptide identification in proteomics, Nat. Methods 5 (10) (2008) 873–875, https://doi.org/10.1038/nmeth.1254.

[218] Z. Zhang, M.C. Burke, W.E. Wallace, Y. Liang, S.L. Sheetlin, Y.A. Mirokhin, D. V. Tchekhovskoi, S.E. Stein, Sensitive method for the confident identification of genetically variant peptides in human hair keratin, J. Forensic Sci. (2019), https://doi.org/10.1111/1556-4029.14229.

[219] D. Fenyo, J. Eriksson, R. Beavis, Mass spectrometric protein identification using the global proteome machine, Methods Mol. Biol. 673 (2010) 189–202, https://doi.org/10.1007/978-1-60761-842-3_11.

[220] S.E. Stein, An integrated method for spectrum extraction and compound identification from gas chromatography/mass spectrometry data, J. Am. Soc. Mass Spectrom. 10 (8) (1999) 770–781.

[221] T. Ogawa, K. Zaitsu, T. Kokaji, K. Suga, F. Kondo, M. Iwai, T. Suzuki, A. Ishii, H. Seno, Development and application of a forensic toxicological library for identification of 56 natural toxic substances by liquid chromatography–quadrupole time-of-flight mass spectrometry, Forensic Toxicol. 38 (1) (2020) 232–242, https://doi.org/10.1007/s11419-019-00506-w.

[222] SWGDAM, in: SWGDAM Validation Guidelines for DNA Analysis Methods, Federal Bureau of Investigation/Department of Justice (Ed.) FBI, Washington DC, 2016.

[223] Federal Bureau of Investigation Quality Assurance Standards For Forensic Science Testing Laboratories, in: United States Department of Justice (ed) U.S. Department of Justice, 2020, pp.1-41.

[224] M. Prinz, Y. Tang, D. Siegel, H. Yang, H. Deng, in: Establishment of a Fast and Accurate Proteomic Method for Body Fluid/Cell Type Identification, Department of Justice (Ed.) National Institute of Justice, 2011.

[225] R. Li. Forensic Biology, 2nd ed., CRC Press, 2015.

[226] M.M. Foley, C.O. Brown, C.G. Westring, P.B. Danielson, H.E. McKiernan, Effects of organic acids and common household products on the occurrence of false positive test results using immunochromatographic assays, Forensic Sci. Int. 308 (2020), 110165.

[227] U. Ricci, I. Carboni, F. Torricelli, False-positive results with amylase testing of citrus fruits, J. Forensic Sci. 59 (5) (2014) 1410–1412, https://doi.org/10.1111/1556-4029.12457.

[228] C. Atkinson, T. Silenieks, C. Pearman, Validation of ABAcard(TM) HemaTrace® KITS: To Determine The Presence Of Higher Primate Hemoglobin In Bloodstains, Evidence Recovery and Biology Analytical Groups Summer Vacation Project, Forensic Science SA, Adelaide, (2003).

[229] P.J. Thul, L. Akesson, M. Wiking, D. Mahdessian, A. Geladaki, H. Ait Blal, T. Alm, A. Asplund, L. Bjork, L.M. Breckels, A. Backstrom, F. Danielsson, L. Fagerberg, J. Fall, L. Gatto, C. Gnann, S. Hober, M. Hjelmare, F. Johansson, S. Lee, C. Lindskog, J. Mulder, C.M. Mulvey, P. Nilsson, P. Oksvold, J. Rockberg, R. Schutten, J.M. Schwenk, A. Sivertsson, E. Sjostedt, M. Skogs, C. Stadler, D. P. Sullivan, H. Tegel, C. Winsnes, C. Zhang, M. Zwahlen, A. Mardinoglu, F. Ponten, K. von Feilitzen, K.S. Lilley, M. Uhlen, E. Lundberg, A subcellular map of the human proteome, Science 356 (6340) (2017), https://doi.org/10.1126/science.aal3321.

[230] E.W. Deutsch, Z. Sun, D. Campbell, U. Kusebauch, C.S. Chu, L. Mendoza, D. Shteynberg, G.S. Omenn, R.L. Moritz, State of the human proteome in 2014/2015 as viewed through peptideatlas: enhancing accuracy and coverage through the AtlasProphet, J. Proteome Res. 14 (9) (2015) 3461–3473, https://doi.org/10.1021/acs.jproteome.5b00500.

[231] D.N. Melegos, H. Yu, M. Ashok, C. Wang, F. Stanczyk, E.P. Diamandis, Prostate-specific antigen in female serum, a potential new marker of androgen excess, J. Clin. Endocrinol. Metab. 82 (3) (1997) 777–780, https://doi.org/10.1210/jcem.82.3.3792.

[232] H. Yang, B. Zhou, D. Siegel, Y. Tang, M. Prinz, A Forensic Assay for Bodily Fluid Identification by Mass Spectrometry, American Society for Mass Spectrometry, Denver, Colarado, 2011.

[233] M. Zhao, Y. Yang, Z. Guo, C. Shao, H. Sun, Y. Zhang, Y. Sun, Y. Liu, Y. Song, L. Zhang, Q. Li, J. Liu, M. Li, Y. Gao, W. Sun, A comparative proteomics analysis of five body fluids: plasma, urine, cerebrospinal fluid, amniotic fluid, and saliva, Proteom. Clin. Appl. 12 (6) (2018), 1800008, https://doi.org/10.1002/prca.201800008.

[234] H. Yang, B. Zhou, M. Prinz, D. Siegel, Proteomic analysis of menstrual blood, Mol. Cell. Proteom. MCP (2012), https://doi.org/10.1074/mcp.M112.018390.

[235] D. Siegel, Development of a Proteomic Assay for Menstrual Blood, Vaginal Fluid and Species Identification, in: Do Justice (Ed.) Department of Justice, Washington D.C., 2018, p. 95.

[236] K.M. Legg, R. Powell, N. Reisdorph, R. Reisdorph, P.B. Danielson, Verification of protein biomarker specificity for the identification of biological stains by quadrupole time-of-flight mass spectrometry, Electrophoresis 38 (6) (2017) 833–845, https://doi.org/10.1002/elps.201600352.

[237] A.P. Drabovich, M.P. Pavlou, I. Batruch, E.P. Diamandis, Proteomic and mass spectrometry technologies for biomarker discovery. Proteomic and Metabolomic Approaches to Biomarker Discovery, Elsevier, 2013, pp. 17–37.

[238] A. Illiano, V. Arpino, G. Pinto, A. Berti, V. Verdoliva, G. Peluso, P. Pucci, A. Amoresano, Multiple reaction monitoring tandem mass spectrometry approach for the identification of biological fluids at crime scene investigations, Anal. Chem. 90 (9) (2018) 5627–5636.

[239] H.E. McKiernan, C.O. Brown, L.C. Arantes, P.B. Danielson, K.M. Legg, NextGen serology: leveraging mass spectrometry for protein-based human body fluid identification. Applications in Forensic Proteomics: Protein Identification and Profiling, ACS Publications, 2019, pp. 47–80.

[240] S. Dammeier, S. Nahnsen, J. Veit, F. Wehner, M. Ueffing, O. Kohlbacher, Mass-spectrometry-based proteomics reveals organ-specific expression patterns to be used as forensic evidence, J. Proteome Res. 15 (1) (2016) 182–192, https://doi.org/10.1021/acs.jproteome.5b00704.

[241] G.R. Abecasis, A. Auton, L.D. Brooks, M.A. DePristo, R.M. Durbin, R.E. Handsaker, H.M. Kang, G.T. Marth, G.A. McVean, An integrated map of genetic variation from 1,092 human genomes, Nature 491 (7422) (2012) 56–65, https://doi.org/10.1038/nature11632.

[242] P.E. Geyer, S.P. Mann, P.V. Treit, M. Mann, Plasma proteomes can be reidentifiable and potentially contain personally sensitive and incidental findings, Mol. Cell. Proteom. 20 (2021), 100035, https://doi.org/10.1074/mcp.RA120.002359.

[243] M.K. Bunger, B.J. Cargile, J.R. Sevinsky, E. Deyanova, N.A. Yates, R. C. Hendrickson, J.L. Stephenson Jr., Detection and validation of non-synonymous coding SNPs from orthogonal analysis of shotgun proteomics data, J. Proteome Res. 6 (6) (2007) 2331–2340, https://doi.org/10.1021/pr0700908.

[244] K.E. Mason, D. Anex, T. Grey, B. Hart, G. Parker, Protein-based forensic identification using genetically variant peptides in human bone, Forensic Sci. Int. 288 (2018) 89–96.

[245] G. Parker, Z. Goecker, R. Franklin, B. Durbin-Johnson, J. Milan, N. Karim, C. De Leon, A. Matzoll, T. Borja, R. Rice, Proteomic genotyping: using mass spectrometry to infer SNP genotypes in a forensic context, Forensic Sci. Int. Genet. Suppl. Ser. 7 (2019) 664–666.

[246] M. Beck, A. Schmidt, J. Malmstroem, M. Claassen, A. Ori, A. Szymborska, F. Herzog, O. Rinner, J. Ellenberg, R. Aebersold, The quantitative proteome of a human cell line, Mol. Syst. Biol. 7 (1) (2011) 549.

[247] K.E. Mason, P.H. Paul, F. Chu, D.S. Anex, B.R. Hart, Development of a Protein-based Human Identification Capability from a Single Hair (Epub ahead of print), J. Forensic Sci. (2019), https://doi.org/10.1111/1556-4029.13995.

[248] T. Kocher, P. Pichler, M. De Pra, L. Rieux, R. Swart, K. Mechtler, Development and performance evaluation of an ultralow flow nanoliquid chromatography-tandem mass spectrometry set-up, Proteomics (2014), https://doi.org/10.1002/pmic.201300418.

[249] S. Houel, R. Abernathy, K. Renganathan, K. Meyer-Arendt, N.G. Ahn, W.M. Old, Quantifying the impact of chimera MS/MS spectra on peptide identification in large-scale proteomics studies, J. Proteome Res. 9 (8) (2010) 4152–4160, https://doi.org/10.1021/pr1003856.

[250] W. Fu, T.D. O'Connor, G. Jun, H.M. Kang, G. Abecasis, S.M. Leal, S. Gabriel, D. Altshuler, J. Shendure, D.A. Nickerson, M.J. Bamshad, J.M. Akey, Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants, Nature 493 (7431) (2013) 216–220, https://doi.org/10.1038/nature11690.

[251] K.J. Karczewski, L.C. Francioli, G. Tiao, B.B. Cummings, J. Alfoldi, Q. Wang, R. L. Collins, K.M. Laricchia, A. Ganna, D.P. Birnbaum, L.D. Gauthier, H. Brand, M. Solomonson, N.A. Watts, D. Rhodes, M. Singer-Berk, E.M. England, E.G. Seaby, J.A. Kosmicki, R.K. Walters, K. Tashman, Y. Farjoun, E. Banks, T. Poterba, A. Wang, C. Seed, N. Whiffin, J.X. Chong, K.E. Samocha, E. Pierce-Hoffman, Z. Zappala, A.H. O'Donnell-Luria, E.V. Minikel, B. Weisburd, M. Lek, J.S. Ware, C. Vittal, I.M. Armean, L. Bergelson, K. Cibulskis, K.M. Connolly, M. Covarrubias, S. Donnelly, S. Ferriera, S. Gabriel, J. Gentry, N. Gupta, T. Jeandet, D. Kaplan, C. Llanwarne, R. Munshi, S. Novod, N. Petrillo, D. Roazen, V. Ruano-Rubio, A. Saltzman, M. Schleicher, J. Soto, K. Tibbetts, C. Tolonen, G. Wade, M. E. Talkowski, C. Genome Aggregation Database, B.M. Neale, M.J. Daly, D. G. MacArthur, The mutational constraint spectrum quantified from variation in 141,456 humans, Nature 581 (7809) (2020) 434–443, https://doi.org/10.1038/s41586-020-2308-7.

[252] A. Moghal, K. Mohler, M. Ibba, Mistranslation of the genetic code, FEBS Lett. 588 (23) (2014) 4305–4310.

[253] M.H. Schwartz, T. Pan, Function and origin of mistranslation in distinct cellular contexts, Crit. Rev. Biochem. Mol. Biol. 52 (2) (2017) 205–219.

[254] R.M. Centor, Signal detectability: the use of ROC curves and their analyses, Med Decis. Mak. 11 (2) (1991) 102–106.

[255] M.H. Zweig, G. Campbell, Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine, Clin. Chem. 39 (4) (1993) 561–577.

[256] A.E. Woerner, F.C. Hewitt, M.W. Gardner, M.A. Freitas, K.Q. Schulte, D. S. LeSassier, M. Baniasad, A.J. Reed, M.E. Powals, A.R. Smith, N.C. Albright, B. C. Ludolph, L. Zhang, L.W. Allen, K. Weber, B. Budowle, An algorithm for random match probability calculation from peptide sequences, Forensic Sci. Int. Genet. 47 (2020), 102295, https://doi.org/10.1016/j.fsigen.2020.102295.

[257] M. Norris Bradley, C.M. Shuford, P.L. Holland, M. Levandoski, R.P. Grant, Quality over quantity: a qualitative, targeted bottom-up proteomics approach to genotyping apolipoprotein L1, Clin. Biochem 82 (2020) 58–65, https://doi.org/10.1016/j.clinbiochem.2020.03.015.

[258] T. Buonasera, J. Eerkens, A. de Famingh, L. Engbring, J. Yip, L. H, R. Haas, D. DiGiuseppe, D. Grant, M. Salemi, C. Nijmeh, M. Arellano, A. Leventhal, B. Phinney, B.F. Byrd, R.S. Malhi, G.J. Parker, A comparison of proteomic, genomic, and osteological methods of archaeological sex estimation, Sci. Rep. (2020) 1–15, https://doi.org/10.1038/s41598-020-68550-w.

[259] I.W. Evert, B.S. Weir. Interpreting DNA Evidence: Statistical Genetics for Forensic Scientists, 1st ed., Sinauer Associates, 1998.

[260] K.G. Ardlie, L. Kruglyak, M. Seielstad, Patterns of linkage disequilibrium in the human genome, Nat. Rev. Genet. 3 (4) (2002) 299–309, https://doi.org/10.1038/nrg777.

[261] M.A. Rogers, H. Winter, L. Langbein, R. Bleiler, J. Schweizer, The human type I keratin gene family: characterization of new hair follicle specific members and evaluation of the chromosome 17q21.2 gene domain, Differ. Res. Biol. Divers. 72 (9–10) (2004) 527–540, https://doi.org/10.1111/j.1432-0436.2004.07209006.x.

[262] M.A. Rogers, L. Edler, H. Winter, L. Langbein, I. Beckmann, J. Schweizer, Characterization of new members of the human type II keratin gene family and a general evaluation of the keratin gene domain on chromosome 12q13.13, J. Investig. Dermatol. 124 (3) (2005) 536–544, https://doi.org/10.1111/j.0022-202X.2004.23530.x.

[263] P.E. Bowden, The human type II keratin gene cluster on chromosome 12q13.13: final count or hidden secrets? J. Investig. Dermatol. 124 (3) (2005) https://doi.org/10.1111/j.0022-202X.2005.23647.x (xv–xvii).

[264] A. Agresti, B.A. Coull, Approximate is better than "exact" for interval estimation of binomial proportions, Am. Stat. 52 (2) (1998) 119–126.

[265] L. Roewer, M.M. Andersen, J. Ballantyne, J.M. Butler, A. Caliebe, D. Corach, M. E. D'Amato, L. Gusmão, Y. Hou, P. de Knijff, W. Parson, M. Prinz, P.M. Schneider, D. Taylor, M. Vennemann, S. Willuweit, DNA commission of the International Society of Forensic Genetics (ISFG): recommendations on the interpretation of Y-STR results in forensic analysis, Forensic Sci. Int. Genet. (2020), 102308, https://doi.org/10.1016/j.fsigen.2020.102308.

[266] W. Parson, L. Gusmao, D.R. Hares, J.A. Irwin, W.R. Mayr, N. Morling, E. Pokorak, M. Prinz, A. Salas, P.M. Schneider, T.J. Parsons, DNA Commission of the International Society for Forensic Genetics: revised and extended guidelines for mitochondrial DNA typing, Forensic Sci. Int. Genet. 13 (2014) 134–142, https://doi.org/10.1016/j.fsigen.2014.07.010.

[267] J. Ge, B. Budowle, R. Chakraborty, DNA identification by pedigree likelihood ratio accommodating population substructure and mutations, Investig. Genet. 1 (1) (2010) 8, https://doi.org/10.1186/2041-2223-1-8.

[268] SWGDAM, Scientific Working Group on DNA Analysis Methods (SWGDAM), Interpretation guidelines for Mitochondrial DNA Analysis by Forensic DNA Testing Laboratories., in: Department of Justice (Ed.) National Institute of Justice, Washington D.C., 2019, p. 26.

[269] I.Z. Mamedov, I.A. Shagina, M.A. Kurnikova, S.N. Novozhilov, D.A. Shagin, Y. B. Lebedev, A new set of markers for human identification based on 32 polymorphic Alu insertions, Eur. J. Hum. Genet. EJHG 18 (7) (2010) 808–814, https://doi.org/10.1038/ejhg.2010.22.

[270] K.S. Grisedale, G.M. Murphy, H. Brown, M.R. Wilson, S.K. Sinha, Successful nuclear DNA profiling of rootless hair shafts: a novel approach, Int. J. Leg. Med. 132 (1) (2018) 107–115, https://doi.org/10.1007/s00414-017-1698-z.

[271] F. Chu, Analytical strategies to investigate the hair proteome for human identification. Chemistry, Michigan State University, 2020.

[272] R.K. Roby, S.D. Gonzalez, N.R. Phillips, J.V. Planz, J.L. Thomas, J.A.P. Astudillo, J. Ge, E.A. Morales, A.J. Eisenberg, R. Chakraborty, Autosomal STR allele frequencies and Y-STR and mtDNA haplotypes in Chilean sample populations, Forensic Sci. Int. Genet. Suppl. Ser. 2 (1) (2009) 532–533.

[273] B. Budowle, A. van Daal, Forensically relevant SNP classes, BioTechniques 44 (5) (2008) 603–608, https://doi.org/10.2144/000112806 (610).

[274] F. Oldoni, K.K. Kidd, D. Podini, Microhaplotypes in forensic genetics, Forensic Sci. Int. Genet. 38 (2019) 54–69, https://doi.org/10.1016/j.fsigen.2018.09.009.

[275] J.M. Butler, M.D. Coble, P.M. Vallone, STRs vs. SNPs: thoughts on the future of forensic DNA testing, Forensic Sci. Med. Pathol. 3 (3) (2007) 200–205, https://doi.org/10.1007/s12024-007-0018-1.

[276] C. Phillips, C. Santos, M. Fondevila, A. Carracedo, M.V. Lareu, Inference of ancestry in forensic analysis I: autosomal ancestry-informative marker sets, Methods Mol. Biol. 1420 (2016) 233–253, https://doi.org/10.1007/978-1-4939-3597-0_18.

[277] B.F. Algee-Hewitt, M.D. Edge, J. Kim, J.Z. Li, N.A. Rosenberg, Individual identifiability predicts population identifiability in forensic microsatellite markers, Curr. Biol. 26 (7) (2016) 935–942.

[278] D.W. Deedrick, S.L. Koch, Microscopy of hair part 1: a practical guide and manual for human hairs, Forensic Sci. Commun. 6 (1) (2004).

[279] D.W. Deedrick, Hairs, fibers, crime, and evidence: part 1 hair evidence, Forensic Sci. Commun. 2 (3) (2000).

[280] M.R. Schneider, R. Schmidt-Ullrich, R. Paus, The hair follicle as a dynamic miniorgan, Curr. Biol. CB 19 (3) (2009) R132–R142, https://doi.org/10.1016/j.cub.2008.12.005 (https://doi.org/S0960-9822(08)01626-6 [pii]).

[281] A.M. Kligman, Pathologic dynamics of human hair loss: I. Telogen effluvium, Arch. Dermatol. 83 (2) (1961) 175–198.

[282] J. Plowman, D.P. Harland, S. Deb-Choudhury. The Hair Fibre: Proteins, Structure and Development, Springer, 2017, https://doi.org/10.1007/978-981-10-8195-8.

[283] A.S. Wilson, Taphonomic alterations to hair and nail, in: E.M.J. Schotsmans, N. Márquez-Grant, S.L. Forbes (Eds.), Taphonomy of Human Remains: Forensic Analysis of the Dead and the Depositional Environment, John Wiley & Sons Ltd, 2017, pp. 81–91.

[284] D. McNevin, L. Wilson-Wilde, J. Robertson, J. Kyd, C. Lennard, Short tandem repeat (STR) genotyping of keratinised hair. Part 1. Review of current status and knowledge gaps, Forensic Sci. Int. 153 (2–3) (2005) 237–246, https://doi.org/10.1016/j.forsciint.2005.05.006.

[285] F. Chu, K.E. Mason, D.S. Anex, A.D. Jones, B.R. Hart, Hair proteome variation at different body locations on genetically variant peptide detection for protein-based human identification, Sci. Rep. 9 (1) (2019) 7641, https://doi.org/10.1038/s41598-019-44007-7.

[286] T.L. Carlson, M. Moini, B.A. Eckenrode, B.M. Allred, J. Donfack, Protein extraction from human anagen head hairs one millimeter or less in total length, BioTechniques 64 (4) (2018) 170–176, https://doi.org/10.2144/btn-2018-2004.

[287] N.R. Barthelemy, A. Bednarczyk, C. Schaeffer-Reiss, D. Jullien, A. Van Dorsselaer, N. Cavusoglu, Proteomic tools for the investigation of human hair structural proteins and evidence of weakness sites on hair keratin coil segments, Anal. Biochem. 421 (1) (2012) 43–55, https://doi.org/10.1016/j.ab.2011.10.011.

[288] L. Eckhart, S. Lippens, E. Tschachler, W. Declercq, Cell death by cornification, Biochim. Biophys. Acta 1833 (12) (2013) 3471–3480, https://doi.org/10.1016/j.bbamcr.2013.06.010.

[289] Z. Zhang, S.E. Stein, MS Library-based, Sensitive Method for Identification of Genetically-Variant Peptides in Human Hair, 67TH ASMS Conference on Mass Spectrometry and Allied Topics Atlanta, GA, (2019).

[290] K.F. Jones, T.L. Carlson, B.A. Eckenrode, J. Donfack, Assessing protein sequencing in human single hair shafts of decreasing lengths, Forensic Sci. Int. Genet. 44 (2019), 102145, https://doi.org/10.1016/j.fsigen.2019.102145.

[291] J.A. Milan, P.-W. Wu, M.R. Salemi, B.P. Durbin-Johnson, D.M. Rocke, B. S. Phinney, R.H. Rice, G.J. Parker, Comparison of protein expression levels and proteomically-inferred genotypes using human hair from different body sites, Forensic Sci. Int. Genet. 41 (2019) 19–23, https://doi.org/10.1016/j.fsigen.2019.03.009.

[292] T.J. Plott, N. Karim, B.P. Durbin-Johnson, D.P. Swift, R.S. Youngquist, M. Salemi, B.S. Phinney, D.M. Rocke, M.G. Davis, G.J. Parker, R.H. Rice, Age-Related Changes in Hair Shaft Protein Profiling and Genetically Variant Peptides, Forensic Science International: Genetics in press, (2020).

[293] M. Lawas, K.F. Jones, K.E. Mason, D.S. Anex, T.L. Carlson, L.V. Forger, B. A. Eckenrode, B. Hart, J. Donfack, Assessing single-source reproducibility of human head hair peptide profiling from different regions of the scalp, Forensic science international, Forensic Sci. Int. Genet. 50 (2021), 102396, https://doi.org/10.1016/j.fsigen.2020.102396.

[294] R. Ramotowski, Composition of latent print residue, in: A.K. Datta (Ed.), Advances in Fingerprint Technology, CRC Press LLC, 2001, pp. 63–104.

[295] S. Francese, Techniques for fingermark analysis using MALDI MS: a practical overview, in: R. Cramer (Ed.), Advances in MALDI and Laser-Induced Soft Ionization Mass Spectrometry, Springer, 2016, pp. 93–128, https://doi.org/10.1007/978-3-319-04819-2_6.

[296] J.E. Templeton, A. Linacre, DNA profiles from fingermarks, BioTechniques 57 (5) (2014) 259–266, https://doi.org/10.2144/000114227.

[297] A. van Dam, F.T. van Beek, M.C. Aalders, T.G. van Leeuwen, S.A. Lambrechts, Techniques that acquire donor profiling information from fingermarks - a review, Sci. Justice J. Forensic Sci. Soc. 56 (2) (2016) 143–154, https://doi.org/10.1016/j.scijus.2015.12.002.

[298] Scientific Working Group on Friction Ridge Analysis, The Fingerprint Sourcebook, in: National Institute of Justice (Ed.) Department of Justice, Washington D.C., (2011).

[299] I. Hefetz, N. Einot, M. Faerman, M. Horowitz, J. Almog, Touch DNA: the effect of the deposition pressure on the quality of latent fingermarks and STR profiles, Forensic Sci. Int. Genet. 38 (2019) 105–112, https://doi.org/10.1016/j.fsigen.2018.10.016.

[300] S. Fieldhouse, E. Oravcova, L. Walton-Williams, The effect of DNA recovery on the subsequent quality of latent fingermarks, Forensic Sci. Int. 267 (2016) 78–88, https://doi.org/10.1016/j.forsciint.2016.08.003.

[301] A. Girod, R. Ramotowski, C. Weyermann, Composition of fingermark residue: a qualitative and quantitative review, Forensic Sci. Int. 223 (1–3) (2012) 10–24, https://doi.org/10.1016/j.forsciint.2012.05.018.

[302] C. Huynh, J. Halámek, Trends in fingerprint analysis, Trends Anal. Chem. 82 (2016) 328–336, https://doi.org/10.1016/j.trac.2016.06.003.

[303] R.A.H. van Oorschot, B. Szkuta, G.E. Meakin, B. Kokshoorn, M. Goray, DNA transfer in forensic science: a review, Forensic Sci. Int. Genet. 38 (2019) 140–166, https://doi.org/10.1016/j.fsigen.2018.10.014.

[304] S. Francese, R. Bradshaw, L.S. Ferguson, R. Wolstenholme, M.R. Clench, S. Bleay, Beyond the ridge pattern: multi-informative analysis of latent fingermarks by MALDI mass spectrometry, Analyst 138 (15) (2013) 4215–4228, https://doi.org/10.1039/c3an36896c.

[305] S. Oonk, T. Schuurmans, M. Pabst, L. de Smet, M. de Puit, Proteomics as a new tool to study fingermark ageing in forensics, Sci. Rep. 8 (1) (2018) 16425, https://doi.org/10.1038/s41598-018-34791-z.

[306] R.A. van Oorschot, M.K. Jones, DNA fingerprints from fingerprints, Nature 387 (6635) (1997) 767, https://doi.org/10.1038/42838.

[307] R.A. Wickenheiser, Trace DNA: a review, discussion of theory, and application of the transfer of trace quantities of DNA through skin contact, J. Forensic Sci. 47 (3) (2002) 442–450.

[308] S.E. Cavanaugh, A.S. Bathrick, Direct PCR amplification of forensic touch and other challenging DNA samples: a review, Forensic Sci. Int. Genet. 32 (2018) 40–49, https://doi.org/10.1016/j.fsigen.2017.10.005.

[309] L. Ostojic, E. Wurmbach, Analysis of fingerprint samples, testing various conditions, for forensic DNA identification, Sci. Justice J. Forensic Sci. Soc. 57 (1) (2017) 35–40, https://doi.org/10.1016/j.scijus.2016.08.009.

[310] S. Cadd, M. Islam, P. Manson, S. Bleay, Fingerprint composition and aging: a literature review, Sci. Justice J. Forensic Sci. Soc. 55 (4) (2015) 219–238, https://doi.org/10.1016/j.scijus.2015.02.004.

[311] S.A. Sterling, Range of Detection for Proteins and DNA from Fingerprints on Fired and Unfired Cartridge Casings, Forensic Science, CUNY - John Jay College, CUNY Academic Works, 2017.

[312] Z.C. Goecker, S.E. Swiontek, A. Lakhtakia, R. Roy, Comparison of Quantifiler((R)) Trio and InnoQuant human DNA quantification kits for detection of DNA degradation in developed and aged fingerprints, Forensic Sci. Int. 263 (2016) 132–138, https://doi.org/10.1016/j.forsciint.2016.04.009.

[313] S. Francese, C. Russo, Fingermarks as a new proteomic specimen: state of the art and perspective of in situ proteomics, in: E.D. Merkley (Ed.), Applications in Forensic Proteomics: Protein Identification and Profiling, ACS, Washington, DC, 2019, pp. 91–105.

[314] S. Kamanna, J. Henry, N. Voelcker, A. Linacre, K.P. Kirkbride, “Bottom-up” in situ proteomic differentiation of human and non-human haemoglobins for forensic purposes by matrix-assisted laser desorption/ionization time-of-flight tandem mass spectrometry, Rapid Commun. Mass Spectrom. 31 (22) (2017) 1927–1937, https://doi.org/10.1002/rcm.7986.

[315] L. Deininger, E. Patel, M.R. Clench, V. Sears, C. Sammon, S. Francese, Proteomics goes forensic: detection and mapping of blood signatures in fingermarks, Proteomics 16 (11–12) (2016) 1707–1717, https://doi.org/10.1002/pmic.201500544.

[316] K. Scotcher, R. Bradshaw, The analysis of latent fingermarks on polymer banknotes using MALDI-MS, Sci. Rep. 8 (1) (2018) 8765, https://doi.org/10.1038/s41598-018-27004-0.

[317] A. van Dam, M.C. Aalders, T.G. van Leeuwen, S.A. Lambrechts, The compatibility of fingerprint visualization techniques with immunolabeling, J. Forensic Sci. 58 (4) (2013) 999–1002, https://doi.org/10.1111/1556-4029.12147.

[318] A. van Dam, M.C.G. Aalders, K. van de Braak, H.J.J. Hardy, T.G. van Leeuwen, S. A.G. Lambrechts, Simultaneous labeling of multiple components in a single fingermark, Forensic Sci. Int. 232 (1) (2013) 173–179, https://doi.org/10.1016/j.forsciint.2013.07.020.

[319] K. Yeh, W.S. Burr, N.L. Stock, T. Stotesbury, Preliminary analysis of latent fingerprints recovered from underneath bloodstains using matrix-assisted laser desorption/ionization fourier-transform ion cyclotron resonance mass

spectrometry imaging (MALDI FT-ICR MSI), Forensic Chem. 20 (2020), 100274, https://doi.org/10.1016/j.forc.2020.100274.

[320] K. Tanaka, H. Waki, Y. Ido, S. Akita, Y. Yoshida, T. Yoshida, T. Matsuo, Protein and polymer analyses up to *m/z* 100 000 by laser ionization time-of-flight mass spectrometry, Rapid Commun. Mass Spectrom. 2 (8) (1988) 151–153.

[321] J.R. Yates 3rd, Mass spectrometry and the age of the proteome, J. Mass Spectrom. JMS 33 (1) (1998) 1–19 (https://doi.org/10.1002/(SICI)1096-9888(199801)33:1<1::AID-JMS624>3.0.CO;2-9).

[322] L.S. Ferguson, F. Wulfert, R. Wolstenholme, J.M. Fonville, M.R. Clench, V. A. Carolan, S. Francese, Direct detection of peptides and small proteins in fingermarks and determination of sex by MALDI mass spectrometry profiling, Analyst 137 (20) (2012) 4686–4692, https://doi.org/10.1039/c2an36074h.

[323] C. Elsner, B. Abel, Ultrafast high-resolution mass spectrometric finger pore imaging in latent finger prints, Sci. Rep. 4 (1) (2014) 6905, https://doi.org/10.1038/srep06905.

[324] E. Patel, M.R. Clench, A. West, P.S. Marshall, N. Marshall, S. Francese, Alternative surfactants for improved efficiency of in situ tryptic proteolysis of fingermarks, J. Am. Soc. Mass Spectrom. 26 (6) (2015) 862–872, https://doi.org/10.1007/s13361-015-1140-z.

[325] F. Welker, M.J. Collins, J.A. Thomas, M. Wadsley, S. Brace, E. Cappellini, S. T. Turvey, M. Reguero, J.N. Gelfo, A. Kramarz, J. Burger, J. Thomas-Oates, D. A. Ashford, P.D. Ashton, K. Rowsell, D.M. Porter, B. Kessler, R. Fischer, C. Baessmann, S. Kaspar, J.V. Olsen, P. Kiley, J.A. Elliott, C.D. Kelstrup, V. Mullin, M. Hofreiter, E. Willerslev, J.J. Hublin, L. Orlando, I. Barnes, R.D. MacPhee, Ancient proteins resolve the evolutionary history of Darwin's South American ungulates, Nature 522 (7554) (2015) 81–84, https://doi.org/10.1038/nature14249.

[326] K. Douka, S. Brown, T. Higham, S. Pääbo, A. Derevianko, M. Shunkov, FINDER project: collagen fingerprinting (ZooMS) for the identification of new human fossils, Antiquity 93 (e1) (2019) 1–6.

[327] A. Desmond, N. Barton, A. Bouzouggar, K. Doukac, P. Fernandez, L. Humphrey, J. Morales, E. Turner, M. Buckley, ZooMS identification of bone tools from the North African Later Stone Age, J. Archaeol. Sci. 98 (2018) 149–157.

[328] N. Beckett, R. Tidy, B. Douglas, N. Sumner, Insulin related-deaths: analysis of insulin and synthetic analogues in coronial exhibits, 24th Lorne Proteomics Symposium., Lorne, Victoria., (2019).

[329] L.M. Labay, C.P. Bitting, K.M. Legg, B.L. Logan, The determination of insulin overdose in postmortem investigations, Acad. Forensic Pathol. 6 (2) (2016) 174–183.

[330] S.C. McGrath, D.M. Schieltz, L.G. McWilliams, J.L. Pirkle, J.R. Barr, Detection and quantification of ricin in beverages using isotope dilution tandem mass spectrometry, Anal. Chem. 83 (8) (2011) 2897–2905, https://doi.org/10.1021/ac102571f.

[331] E. Duriez, J. Armengaud, F. Fenaille, E. Ezan, Mass spectrometry for the detection of bioterrorism agents: from environmental to clinical applications, J. Mass Spectrom. JMS 51 (3) (2016) 183–199, https://doi.org/10.1002/jms.3747.

[332] S.L. Richards, A.T. Cawley, M.J. Raftery, Detection and quantification of dermorphin and selected analogs in equine urine, Bioanalysis 5 (24) (2013) 2995–3007, https://doi.org/10.4155/bio.13.281.

[333] S.R. Kalb, J.R. Barr, Mass spectrometric detection of ricin and its activity in food and clinical samples, Anal. Chem. 81 (6) (2009) 2037–2042, https://doi.org/10.1021/ac802769s.

[334] R. Steel, M. Timms, V. Levina, J. Vine, A high throughput screen for 17 Dermorphin peptides in equine and human urine and equine plasma, Drug Test. Anal. 6 (9) (2014) 909–921, https://doi.org/10.1002/dta.1585.

[335] M. Duracova, J. Klimentova, A. Fucikova, J. Dresler, Proteomic methods of detection and quantification of protein toxins, in: Toxins, 10, 2018, https://doi.org/10.3390/toxins10030099.

[336] H. Mizoguchi, G. Bagetta, T. Sakurada, S. Sakurada, Dermorphin tetrapeptide analogs as potent and long-lasting analgesics with pharmacological profiles distinct from morphine, Peptides 32 (2) (2011) 421–427, https://doi.org/10.1016/j.peptides.2010.11.013.

[337] F. Guan, Y. You, X. Li, M.A. Robinson, A comprehensive approach to detecting multitudinous bioactive peptides in equine plasma and urine using hydrophilicinteraction liquid chromatography coupled to high resolutionmass spectrometry, Drug Test. Anal. 11 (2019) 1308–1325, https://doi.org/10.1002/dta.2671.

[338] D.D. Nguyen, S.K. Johnson, F. Busetti, V.A. Solah, Formation and degradation of beta-casomorphins in dairy processing, Crit. Rev. Food Sci. Nutr. 55 (2015) 1955–1967.

[339] N. Beckett, B. Douglas, N. Sumner, Vitreous humour as an alternative matrix for the detection of 6-MAM in coronial casework, ANZFSS 24th International Symposium on the Forensic Sciences., Perth, Western Australia., (2018).

[340] E.C. Salido, P.H. Yen, K. Koprivnikar, L.C. Yu, L.J. Shapiro, The human enamel protein gene amelogenin is expressed from both the X and the Y chromosomes, Am. J. Hum. Genet. 50 (2) (1992) 303–316.

[341] J.E. Buikstra, D.H. Ubelaker, Standards for Data Collection from Human Skeletal Remains Arkansas Archaeological Survey Press Fayetteville, (1994).

[342] T. Waldron, The relative survival of the human skeleton: implications for palaeopathology, in: A. Boddington, A. Garland, R. Janaway (Eds.), In "Death, decay and reconstruction: approaches to archaeology and forensic science, Manchester University Press, 1987, pp. 55–64.

[343] A.C. Stone, G.R. Milner, S. Paabo, M. Stoneking, Sex determination of ancient human skeletons using DNA, Am. J. Phys. Anthropol. 99 (2) (1996) 231–238 (https://doi.org/10.1002/(SICI)1096-8644(199602)99:2<231::AID-AJPA1>3.0.CO;2-1).

[344] K.M. Sullivan, A. Mannucci, C.P. Kimpton, P. Gill, A rapid and quantitative DNA sex test: fluorescence-based PCR analysis of XY homologous gene amelogenin, BioTechniques 15 (4) (1993) 636–638, 640.

[345] Y. Nakahori, O. Takenaka, Y. Nakagome, A human X-Y homologous region encodes "amelogenin", Genomics 9 (2) (1991) 264–269, https://doi.org/10.1016/0888-7543(91)90251-9.

[346] K.M. Horsman, J.A. Hickey, R.W. Cotton, J.P. Landers, L.O. Maddox, Development of a human-specific real-time PCR assay for the simultaneous quantitation of total genomic and male DNA, J. Forensic Sci. 51 (4) (2006) 758–765.

[347] S. Deepak, K. Kottapalli, R. Rakwal, G. Oros, K. Rangappa, H. Iwahashi, Y. Masuo, G. Agrawal, Real-time PCR: revolutionizing detection and expression analysis of genes, Curr. Genom. 8 (4) (2007) 234–251.

[348] M. Barbisin, R. Fang, C.E. O'Shea, L.M. Calandro, M.R. Furtado, J.G. Shewale, Developmental validation of the Quantifiler Duo DNA Quantification kit for simultaneous quantification of total human and human male DNA and detection of PCR inhibitors in biological samples, J. Forensic Sci. 54 (2) (2009) 305–319, https://doi.org/10.1111/j.1556-4029.2008.00951.x.

[349] A. Holt, S.C. Wootton, J.J. Mulero, P.M. Brzoska, E. Langit, R.L. Green, Developmental validation of the Quantifiler((R)) HP and Trio Kits for human DNA quantification in forensic samples, Forensic Sci. Int. Genet. 21 (2016) 145–157, https://doi.org/10.1016/j.fsigen.2015.12.007.

[350] M. Kayser, Forensic use of Y-chromosome DNA: a general overview, Hum. Genet. 136 (5) (2017) 621–635, https://doi.org/10.1007/s00439-017-1776-9.

[351] A. Mittnik, C.-C. Wang, J. Svoboda, J. Krause, A molecular approach to the sexing of the triple burial at the upper paleolithic site of Dolní Věstonice, PLoS One 11 (10) (2016), e0163019, https://doi.org/10.1371/journal.pone.0163019.

[352] P. Skoglund, J. Stora, A. Götherström, M. Jakobsson, Accurate sex identification of ancient human remains using DNA shotgun sequencing, J. Archaeol. Sci. 40 (12) (2013) 4477–4482, https://doi.org/10.1016/j.jas.2013.07.004.

[353] C. Xavier, M. Eduardoff, B. Bertoglio, C. Amory, C. Berger, A. Casas-Vargas, J. Pallua, W. Parson, Evaluation of DNA extraction methods developed for forensic and ancient DNA applications using bone samples of different age, Genes 12 (2) (2021) 146.

[354] M.V. Emery, K. Bolhofner, S. Winingear, R. Oldt, M. Montes, S. Kanthaswamy, J. E. Buikstra, L.C. Fulginiti, A.C. Stone, Reconstructing full and partial STR profiles from severely burned human remains using comparative ancient and forensic DNA extraction techniques, Forensic Sci. Int. Genet. 46 (102272) (2020) 1–8.

[355] H. Mizukami, B. Hathway, N. Procopio, Aquatic decomposition of mammalian corpses: a forensic proteomic approach, J. Proteome Res. 19 (5) (2020) 2122–2135, https://doi.org/10.1021/acs.jproteome.0c00060.

[356] G. Prieto-Bonete, M.D. Perez-Carceles, A. Maurandi-Lopez, C. Perez-Martinez, A. Luna, Association between protein profile and postmortem interval in human bone remains, J. Proteom. 192 (2019) 54–63, https://doi.org/10.1016/j.jprot.2018.08.008.

[357] R.C. Griffin, H. Moody, K.E. Penkman, M.J. Collins, The application of amino acid racemization in the acid soluble fraction of enamel to the estimation of the age of human teeth, Forensic Sci. Int. 175 (1) (2008) 11–16, https://doi.org/10.1016/j.forsciint.2007.04.226.

[358] G. Prieto-Bonetea, M.D. Pérez-Cárceles, A. Maurandi-López, C. Pérez-Martínez, A. Luna, Association between protein profile and postmortem interval in human bone remains, J. Proteom. (2018), https://doi.org/10.1016/j.jprot.2018.08.008.

[359] A.-N.D. Nolan, G. Maker, R.J. Mead, S. Bringans, S.J. Speers, Peptide analysis of mammalian decomposition fluid in relation to the post-mortem interval, Forensic Sci. Int. 311 (2020), 110269, https://doi.org/10.1016/j.forsciint.2020.110269.

[360] A. Ramsøe, V. van Heekeren, P. Ponce, R. Fischer, I. Barnes, C. Speller, M. J. Collins, DeamiDATE 1.0: site-specific deamidation as a tool to assess authenticity of members of ancient proteomes, J. Archaeol. Sci. 115 (105080) (2020).

[361] J. Wilson, N.L. van Doorn, M.J. Collins, Assessing the extent of bone degradation using glutamine deamidation in collagen, Anal. Chem. 84 (21) (2012) 9041–9048, https://doi.org/10.1021/ac301333t.

[362] F. Welker, M.A. Soressi, M. Roussel, I. van Riemsdijk, J.-J. Hublin, M.J. Collins, Variations in glutamine deamidation for a Châtelperronian bone assemblage as measured by peptide mass fingerprinting of collagen, STAR: Science & Technology of Archaeological Research, 3(1), (2016) 15–27. ⟨https://doi.org/10.1080/20548923.2016.1258825⟩.

[363] A. Ramsøe, M. Crispin, M. Mackie, K. McGrath, R. Fischer, B. Demarchi, M. J. Collins, J. Hendy, C. Speller, Assessing the degradation of ancient milk proteins through site-specific deamidation patterns, Sci. Rep. 11 (1) (2021) 7795, https://doi.org/10.1038/s41598-021-87125-x.

[364] P. Hao, Y. Ren, A.J. Alpert, S.K. Sze, Detection, evaluation and minimization of nonenzymatic deamidation in proteomic sample preparation, Mol. Cell. Proteom. MCP 10 (10) (2011), https://doi.org/10.1074/mcp.O111.009381 (O111 009381).

[365] A.R. Hipkiss, Accumulation of altered proteins and ageing: causes and effects, Exp. Gerontol. 41 (5) (2006) 464–473, https://doi.org/10.1016/j.exger.2006.03.004.

[366] N.E. Robinson, A.B. Robinson, Molecular clocks, Proc. Natl. Acad. Sci. USA 98 (3) (2001) 944–949, https://doi.org/10.1073/pnas.98.3.944.

[367] P.P. Hurtado, P.B. O'Connor, Deamidation of collagen, Anal. Chem. 84 (6) (2012) 3017–3025, https://doi.org/10.1021/ac202980z.

[368] N.E. Robinson, A.B. Robinson, Deamidation of human proteins, Proc. Natl. Acad. Sci. USA 98 (22) (2001) 12409–12413, https://doi.org/10.1073/pnas.221463198.

[369] C.S. Hughes, S. Foehr, D.A. Garfield, E.E. Furlong, L.M. Steinmetz, J. Krijgsveld, Ultrasensitive proteome analysis using paramagnetic bead technology, Mol. Syst. Biol. 10 (10) (2014) 757, https://doi.org/10.15252/msb.20145625.

[370] N. Varongchayakul, J. Song, A. Meller, M.W. Grinstaff, Single-molecule protein sensing in a nanopore: a tutorial, Chem. Soc. Rev. 47 (23) (2018) 8512–8524, https://doi.org/10.1039/c8cs00106e.

[371] Z.-L. Hu, M.-Z. Huo, Y.-L. Ying, Y.-T. Long, Biological Nanopore Approach for Single-Molecule Protein Sequencing, Angewandte Chemie International Edition n/a(n/a), (2021). https://doi.org/https://doi.org/10.1002/anie.202013462.

[372] H. Ouldali, K. Sarthak, T. Ensslen, F. Piguet, P. Manivet, J. Pelta, J.C. Behrends, A. Aksimentiev, A. Oukhaled, Electrical recognition of the twenty proteinogenic amino acids using an aerolysin nanopore, Nat. Biotechnol. 38 (2) (2020) 176–181, https://doi.org/10.1038/s41587-019-0345-2.

[373] W. Liu, E. Butler, H. Yang, D. Fenyö, D. Siegel, Informatics approaches to forensic body fluid identification by proteomic mass spectrometry. Applications in Forensic Proteomics: Protein Identification and Profiling, ACS Publications, 2019, pp. 81–90.

[374] I.K. Kushner, G. Clair, S.O. Purvine, J.-Y. Lee, J.N. Adkins, S.H. Payne, Individual variability of protein expression in human tissues, J. Proteome Res. 17 (11) (2018) 3914–3922.

[375] L. Huang, D. Shao, Y. Wang, X. Cui, Y. Li, Q. Chen, J. Cui, Human body-fluid proteome: quantitative profiling and computational prediction, Brief. Bioinform. (2020), https://doi.org/10.1093/bib/bbz160.