Forensic Science International: Genetics 41 (2019) 19–23

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Comparison of protein expression levels and proteomically-inferred genotypes using human hair from different body sites

Jennifer A. Milan[a,1], Pei-Wen Wu[a,1], Michelle R. Salemi[b], Blythe P. Durbin-Johnson[c], David M. Rocke[c], Brett S. Phinney[b], Robert H. Rice[a,1], Glendon J. Parker[a,1,*]

[a] Forensic Science Graduate Program and Department of Environmental Toxicology, University of California, Davis, CA, USA
[b] Proteomics Core Facility, University of California, Davis, CA, USA
[c] Division of Biostatistics, Department of Public Health Sciences, Clinical and Translational Science Center, Biostatistics Core, University of California, Davis, CA, USA

ARTICLE INFO

Keywords:
Genetically variant peptides
Single amino acid polymorphism
Hair shafts
Protein profiling
Body sites

ABSTRACT

The microanatomy of human hair differs as a function of the site of origin on the body. This was a major consideration when anatomical features of hair were used as a means of comparison and human identification. Recent advances have demonstrated that proteomics of the hair shaft can be used to develop profiles of protein abundance and genetically variant peptides, the latter in turn being used to infer genotypes of SNP alleles. Because the profile of proteins would be expected to change as hair anatomy changes, it is an open question if the profile of genetically variant peptides will also change. While some sample to sample variation is expected, a potential drawback of using genetically variant peptides to infer an individual genotype is that the proteomic profile might change as a function of body site origin as well as an individual's genotype. The hypothesis in this study is that the profile of hair shaft genetically variant peptides depends more on an individual's genotype than on the site of hair shaft origin. To test this an analysis of both protein expression levels and genetically variant peptides was conducted on 4 body sites (scalp, axillary, beard and pubic hair) from 5 individuals with 4 biological replicates. Levels of protein expression were estimated using label-free quantification on resulting proteomic mass spectrometry datasets. The same datasets were then also analyzed for the presence of genetically variant peptides. This study demonstrates that the protein profiles of hair shafts varied as a function of somatic origin. By contrast the profile of genetically variant peptides, and resulting inferred genotype of SNP alleles, were more dependent on the individual. In this study random match probabilities ranged up to 1 in 196. Individual identification based on genetically variant peptides therefore can be obtained from human hair without regard to the site of origin. If the site of hair shaft origin was legally relevant then microscopic analysis is still necessary. This study demonstrates the utility of proteomic analysis for extracting forensic information from hair shaft evidence.

1. Introduction

Recent work has demonstrated that proteomic datasets of hair shafts can be a rich source of information for forensic human identification [1]. Genetically variant peptides (GVPs), which contain single amino acid polymorphisms (SAPs), can be used to infer the presence of the corresponding non-synonymous SNP allele in an individual's genome [1,2]. Before this technique may be used in forensic casework however, it needs to be established that the profile will not vary as a function of body location [3]. Shed hair can come from many body locations and can display distinctly different microanatomy [4–6]. Given these differences in hair anatomy from different body locations, it is reasonable to predict that there would also be differences in protein composition and expression levels. It has been an open question whether these anatomical changes will systematically impact the detection of genetically variant peptides.

Examination of protein profiles showed that inbred mouse strains were distinguishable by their hair [7]. This finding suggested that humans, an outbred population, could also be distinguished by their hair protein profiles. The suggestion was confirmed, a phenomenon that was shown to have a genetic basis through studies of monozygotic twins [8,9]. However, preliminary data indicated that the anatomic site of

Abbreviations: GVP, genetically variant peptide; SAP, single amino acid polymorphism; SNP, single nucleotide polymorphism; MS, mass spectrometry

* Corresponding author at: Department of Environmental Toxicology, University of California, One Shields Avenue, Davis, CA, 95616–8588, USA.
E-mail address: gjparker@ucdavis.edu (G.J. Parker).

[1] These authors contributed equally.

https://doi.org/10.1016/j.fsigen.2019.03.009
Received 20 December 2018; Received in revised form 14 February 2019; Accepted 10 March 2019
Available online 11 March 2019
1872-4973/

origin influenced the hair shaft protein profile [9].

Hair shafts are frequently found at crime scenes and potentially constitute valuable evidence [3]. Comparative microscopic hair examination has generated evidence for human identifications in the past, but quantitative analyses are now sought for greater accuracy and reliability [6,10]. Microscopic hair comparisons are still used to identify the site of hair shaft origin [3,4,11]. Mitochondrial DNA in the hair shaft, in the absence of nuclear DNA amplicons, has proven valuable in this regard, but gives information about only the maternal genetic lineage [12–15]. The hair protein constituents are a direct reflection of an individual's genomic sequence and expression levels and thus are a potential rich source of genetic and identifying information [1].

With proper processing, proteins of the hair shaft are efficiently trypsinized to yield peptides derived from the constituent proteins. Matching the peptides to the human database identifies the proteins digested, and numbers of matched peptide spectral counts indicates the relative amounts of a given protein. However, since the peptide sequences also reflect variations in coding DNA sequence, the latter can be inferred from the former. From databases of human genomic sequences, the known distribution of non-synonymous single nucleotide polymorphisms in the genome among human populations can be exploited to target variant peptides derived from hair proteins [1]. This approach previously permitted matching a collection of hair samples to the correct monozygotic twin donors [8].

2. Materials and methods

2.1. Human subjects and sample collection

A total of 5 subjects provided samples of hair shafts from four body sites (25 mg each, typically in tufts of hair about 2 in. in length) with informed consent and approval from the Institutional Review Board of the University of California, Davis (Protocol #741750-1). The sampled hair had not received previous chemical treatment (coloring, bleaching, straightening). The subjects were all male in order to include facial beard hair in the cohort of different hair types. Other hair types collected include shafts from the scalp, pubis and axillary region.

2.2. Hair processing

Hair samples were processed as previously described [8]. Briefly, 4 mg of hair shafts from each subject (n = 5) and site (scalp, pubic, beard and axillary) were processed as 4 biological replicates in parallel for a total of 80 samples. Hair samples were chopped into 2 mm lengths, weighed (4 mg), rinsed twice in 2% sodium dodecanoate (SD) w/v in 0.05 M ammonium bicarbonate to remove dust and debris, heated overnight at 70 °C in 0.4 mL of the same buffer adjusted to 25 mM in dithioerythritol, and then alkylated with iodoacetamide for 45 min at room temperature with magnetic stirring in the dark. To avoid denaturation of trypsin, SD was removed by ethyl acetate extraction after acidification to pH 2–3 with TFA (0.75%). After the upper organic phase was removed, the lower aqueous phase was readjusted to pH 8 with ammonium hydroxide and bicarbonate, and the protein was digested by daily addition of 40 µg of reductively methylated trypsin [13]. The digests were clarified by centrifugation (5 min x 14,300 g) and submitted for mass spectrometric analysis.

2.3. Mass spectrometry and protein identification

As previously described, digests were analyzed by separation on a Proxeon Easy-nLC II HPLC followed by an in-line electrospray ionization with a Proxeon nanospray source and mass spectrometry on a Thermo Scientific Q Exactive Orbitrap (Thermo Scientific, Asheville, NC, USA) [8]. Tandem mass spectral data were analyzed using X!Tandem Vengeance (2015.12.15.2), and peptide/protein identifications were validated using Scaffold version 4.4.3. Protein identifications (decoy false discovery rate 2.8%) were based on a minimum of two peptides (peptide decoy false discovery rate 0.7%). For two-way comparisons of expression level at different sites, data were analyzed as weighted-normalized spectral counts, where exclusive peptide counts were used to delete proteins from the weighted that were identified primarily through shared peptides. Two reagent blank samples with no hair were incorporated into the analysis, with no consistent hair specific genes identified and no genetically variant peptides detected. Each acquisition was separated by a 'sawtooth gradient' wash.

2.4. Detection of genetically variant peptides

Acquired proteomic datasets from each sample were converted into MGF format using MSconvertGUI (Proteowizard 2.1, http://proteowizard.sourceforge.net) and submitted to the Global Proteome Machine webserver (www.thegpm.org; X!Tandem Alanine (2017.2.1.3)) and GPM Fury (X!Tandem Alanine (2016.10.15.2)). Default search settings were used with the exception of a 20 ppm error for the primary scan, inclusion of complete cysteine carbamidomethylation (C + 57), and partial modifications of oxidized methionine (M + 16), and deamidation (N + 1, Q + 1). Detected variant spectra were filtered for quality and specificity. Quality measures include the log(e) score of less than −2.0, a mass error of fragment ions of less than 0.04 Da, exclusion of peptides based on non-tryptic cleavage and the presence of unexpected modifications and fragmentation masses that also occurred on the alternative allele. Characterized genetically variant peptides (GVPs) were identified in each dataset [1]. When combining datasets from two biological replicates of a sample, presence of a GVP was determined by detection in either or both datasets, with no additional weighting for the second identification.

2.5. Label-free quantitation

Intensity-based absolute quantitation (iBAQ) values of samples analyzed with a Q-Exactive Plus mass spectrometer were used to calculate the relative molar amount of each protein [16]. iBAQ values were calculated by importing MaxQuant (version 1.5.7.4). MS1 intensity values and identification results were transferred into Scaffold version 4.7.5 and iBAQ values calculated. Values were calculated for the protein clusters identified by MaxQuant. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifiers PXD011732 and PXD012072, and the Center for Computational Mass Spectrometry (massIVE) with the dataset identifier MSV000083233 (massive.ucsd.edu) [17].

2.6. Statistical analysis

For analysis of protein profiles, weighted spectral count data from the Scaffold algorithm [9] were transformed using a variance stabilizing transformation for negative binomial data, which takes the form $f_\theta(x) = \ln[x + \sqrt{x^2 + \frac{x}{\theta}} + \frac{1}{2\theta}]$. This transformation, when θ is selected to minimize the correlation between the variance and standard deviation of the transformed data, removes mean-variance dependency from the data so that they may be analyzed using methods that assume constant variance across the range of the data. Data were then analyzed using the Bioconductor package for gene expression analysis limma, version 3.28.17 [18], which fits linear models to each protein separately and then applies empirical Bayes shrinkage to the estimated variances to increase power. Application of the above transformation to RNA-Seq data has been discussed [19]. The linear models used for analysis include terms for body area, subject, and the subject-body area interaction. Analyses were conducted using R, version 3.3.1 (R Core Team, 2015). Multidimensional scaling plotting was performed as described in Torgerson [20].

For statistical analysis, GVPs that were detected in all subjects (n = 10) were removed. Remaining informative GVPs with uneven distribution across the cohort were assigned a value of "1" for detected and "0" for non-detected, then multiplied by the inverse of the genotype frequency of the corresponding SNP allele in the 1000 Genomes Project EUR population (phase 3) to give more weight to rarer GVPs. An agglomerative hierarchical clustering dendrogram was constructed based on the Euclidean distance matrix of the weighted detection data. Analyses were conducted using R, version 3.4.3 (R Core Team, 2017). The hierarchical clustering dendrogram was constructed using the R function hclust (with default options) and the contributed R package dendextend, version 1.6.0.

The probability of the proteomically inferred non-synonymous SNP genotype occurring in the European population, also called the random match probability (RMP), was calculated with the assumption of complete dependence within an open reading frame and complete independence outside of the gene boundaries [1]. When multiple single amino acid polymorphisms were detected within a gene boundary, the combination of each possible allele was counted in the individual genotypes of Europeans in the 1000 Genome Project (phase 3) to determine the appropriate genotype frequency [1,21]. The population distribution of the GVP profile was estimated as the product of inferred SNP allele or allele combinations (Pr(GVP profile|EUR)). The use of Jeffreys prior probability to estimate population distribution in the lack of an observed SNP allele in the European population was not necessary since all inferred SNPs were represented in reference populations of the 1000 Genome Project [1,21,22].

3. Results

3.1. Profile comparisons among individuals

Previous work demonstrated that profiles of protein expression levels in scalp hair can distinguish among human donors in two-way comparisons [8,9]. Present work investigated whether profiles of hair from other body sites could also distinguish among individuals. To this end hair samples from 4 different body sites (scalp, beard, axillary, pubic) from each of 5 unrelated male individuals were collected and analyzed with 4 biological replicates. Protein and peptide yields averaged 634 ± 54, and 21,721 ± 1254 (average ± sd.) respectively across all samples. There was no difference in proteomic yields from hair shafts across different individuals or body locations (Figure S1). The protein expression levels (weighted spectral counts) were compared protein by protein between individuals separately for each body site. Numbers of proteins with significantly different levels (adjusted $p < 0.05$) are given in Table 1. As indicated, hair from any of the 4 sites was similarly discriminating among the 5 subjects. A multidimensional scaling plot, similar to a principal component analysis, summarizes the fold differences between samples in pairwise comparisons (Figure S2).

Table 1
Differences in protein expression levels at 4 bodily sites among 5 subjects according to two-way comparisons. Shown are the numbers of proteins that differ significantly in weighted spectral counts between subjects at a given site. The subjects are numbered 1-5. Sampled are from axillary (A), beard (B), pubic (P) and scalp (S) regions.

	2A	**3A**	**4A**	**5A**		**2B**	**3B**	**4B**	**5B**
1A	21	54	54	28	**1B**	32	58	50	51
2A		52	46	29	**2B**		44	31	55
3A			59	60	**3B**			74	87
4A				45	**4B**				53
	2P	**3P**	**4P**	**5P**		**2S**	**3S**	**4S**	**5S**
1P	17	29	48	20	**1S**	73	51	44	40
2P		19	46	34	**2S**		96	62	63
3P			19	38	**3S**			63	70
4P				30	**4S**				28

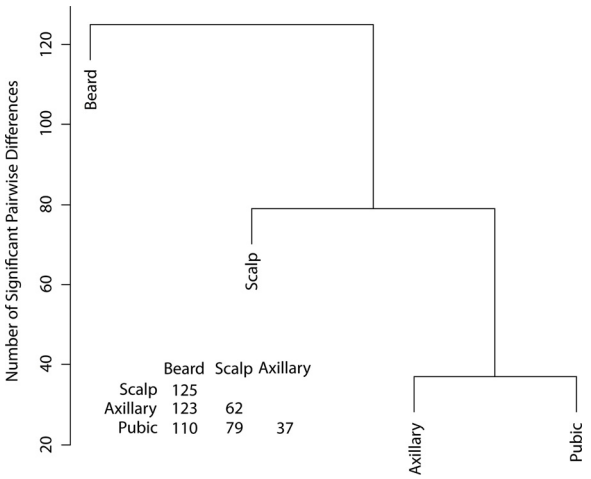

Fig. 1. Hierarchical clustering of significant pairwise differences in protein levels from different anatomic sites. Data from all the samples at each site (n = 20) were grouped and compared by site. The numbers of differences used for the comparison are tabulated.

Hair from the same individual tended to cluster together. However, comparing two individuals using profiles of protein expression with hair from different sites appeared not suitable for such testing because the information from different sites in a given individual were substantially different.

3.2. Protein expression levels vary with anatomic site

The weighted spectra count data above were employed to discern consistent differences between hair samples from the different sites. Fig. 1 shows the table of differences calculated by two-way comparisons among the pooled values for the sites and a hierarchical clustering based on these values. As indicated, hair from axillary and pubic regions were the closest in profile of protein expression levels, while beard hair was most different from those, and scalp hair was in between. Fold differences in protein level for each comparison are given in Supplementary Table S1.

To help describe the differences among individual proteins in hair from the different sites, the proteins were subjected to label free quantitation [16]. For this purpose, the quantitative values for each protein were averaged among the 5 subjects. The result is given in Table 2, where the top 25 proteins at each site are shown. Although semi-quantitative, these values are useful in judging the relative prevalence of constituent proteins. For example, as previously found using an early method of label-free quantitation, keratins comprise > 90% of the protein in the hair shaft [23].

3.3. Analysis of genetically variant peptides across body sites

Proteomes of hair shafts from the four body sites (scalp, beard, axillary, pubic) from five subjects were generated with four biological replicates, as described above. Genetically variant peptide (GVP) profiles were collated from polymorphic peptides identified in the proteomic datasets. Two biological replicates from each body site were pooled for each individual. The resulting GVP profiles were collated for each analyzed dataset and pooled for each set of replicate samples. Figure S2 shows a heatmap summary of the GVPs detected in the datasets. Due to the diversity in trypsin proteolysis, occasionally one SAP is represented in several semi-tryptic partially redundant peptides, although in this study that was not the case.

For statistical analysis, the values were weighted as the inverse of inferred SNP genotype frequency from the European populations. This follows the assumption that correlation between rarer alleles should

Table 2
Relative protein levels in hair from different anatomic locations. Estimates obtained by label free quantitation (iBAQ) of all the samples from each site (n = 20) were averaged, normalized to 100%, and listed as percentages in order of prevalence in scalp samples. The top 25 proteins account for > 98% of the calculated total amounts.

Gene Name	Axillary	Beard	Pubic	Scalp
KRT85[a]	25.46	25.19	25.42	26.04
KRT33B[b]	21.63	23.28	21.98	24.28
KRT32	9.91	10.26	9.90	10.58
KRT35	7.93	8.57	7.99	8.57
KRT38[c]	9.85	6.65	10.28	6.32
KRT36	4.94	4.93	5.01	4.99
KRT75	2.66	2.42	2.52	3.07
KRT7	2.17	2.38	2.22	2.42
KRT10	3.47	2.34	3.08	2.41
GRIPAP1	0.99	1.30	1.50	2.10
KRT82	1.94	1.97	1.90	1.90
HIST1H4A	1.92	1.72	1.72	1.67
HIST1H2BC	1.14	1.06	0.99	1.06
VSIG8	0.90	0.75	0.83	0.78
hCG_2039566	0.36	0.45	0.40	0.38
KRT39	0.36	0.33	0.33	0.37
KRT14[d]	1.12	1.86	0.32	0.35
S100A3	0.35	0.29	0.32	0.28
HIST2H3PS2	0.35	0.27	0.29	0.27
LGALS3	0.20	0.23	0.20	0.22
CALML3	0.27	0.33	0.28	0.21
UBB	0.27	0.32	0.26	0.20
LGALS7	0.17	0.22	0.18	0.14
SFN	0.10	0.15	0.10	0.09
DSG4	0.08	0.08	0.08	0.08

[a] Cluster including KRT81, KRT83, KRT86.
[b] Cluster including KRT31, KRT33A, KRT34, KRT40.
[c] Cluster including KRT37.
[d] Cluster including KRT16, KRT17.

have a greater weight in the analysis. From the data an agglomerative hierarchical clustering dendogram was constructed based on the Euclidean distance matrix of the weighted detection data. The GVP profiles clustered by subject not by body site (Fig. 2). This indicates that GVP profiles obtained from different body sites from one individual are more similar than those profiles obtained from other unrelated individuals. While sample-to-sample variation occurs, and some GVPs are not always detected, the differences observed are explained more by underlying genotype than biological differences in protein expression that track with anatomical changes to the hair shaft. The GVP profile therefore appears dependent on individual genotype and is not systematically affected by the anatomical source of the hair shafts.

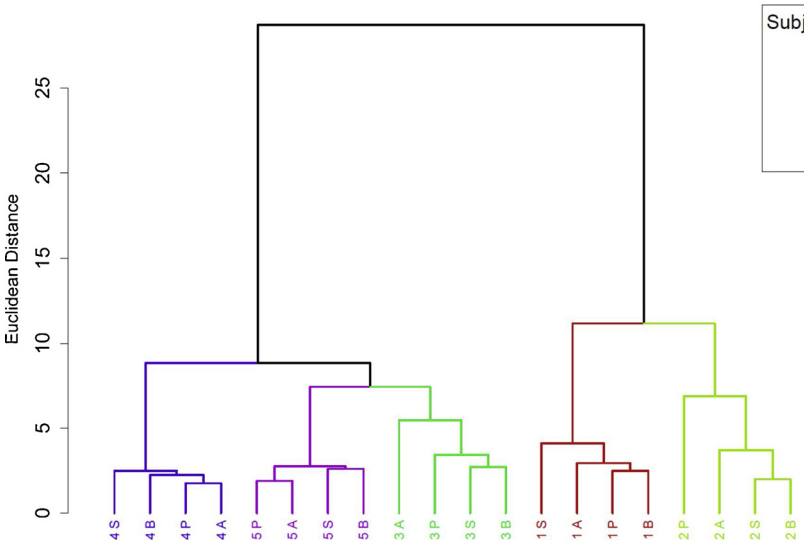

Fig. 2. Hierarchical clustering dendrogram of GVP detection data. Proteomic duplicates were pooled and detected GVPs collated and weighted by the inverse of genotype frequency. A hierarchical clustering dendogram was generated as described in the Methods. Branches corresponding to clusters by subject are shown in a different color for each subject. Sample labels indicate the body area from which the sample originated: S = Scalp, B = Beard, P = Pubic, A = Axillary.

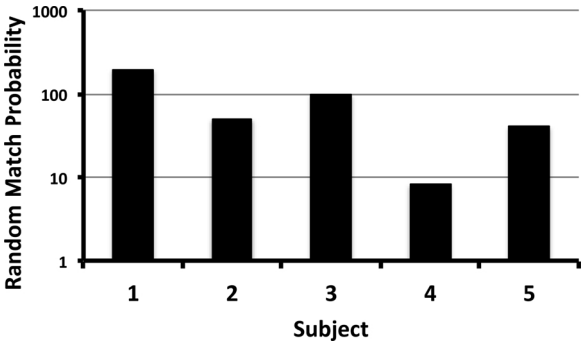

Fig. 3. Random Match Probabilities of the Profile of Genetically Variant Peptides. The cumulative profile of genetically variant peptides was developed for each subject. Each probability of the inferred SNP genotype occurring in the European population of the 1000 genome project was estimated as described in the Methods. Probabilities are presented as the inverse of Pr(GVP profile|EUR).

3.4. Calculation of random match probability

Profiles of genetically variant peptides for each subject were cumulatively combined from datasets of all body sites. These were used to create profiles of inferred SNP alleles. The resulting probability of each genotype occurring in the European population of the 1000 Genome Project (Pr(GVP profile|EUR)) was estimated as described in the Methods [1,21]. Application of the product-rule assumed complete dependence within the open reading frame and complete independence outside of the gene boundary [1]. Resulting random match probabilities (RMP) ranged from a low of 1 in 8 to a high of 1 in 196 (Fig. 3). The number of detected GVP peptides and inferred SNP alleles ranged from 23 to 27, although this is reduced to 12 to 17 when non-discriminatory alleles ($gf \geq 0.99$) are excluded.

4. Discussion

Based on results from inbred mouse strains [7,24], profiles of human hair protein expression were originally pursued to expand the usefulness of hair evidence for distinguishing among individuals. Present results indicate that such comparisons are similarly discriminating regardless of the anatomic origin of the hair provided that samples from the same site are compared. Since the protein profiles depend on the site of origin, analogous to those from epidermal corneocytes from

different sites on the skin, comparing the same sites is critical [24,25]. A corollary is that the site of origin of a hair from a given individual could be deduced by comparing its profile to those from other anatomic sites of interest in that individual. This application, which could be useful in certain forensic contexts, merits further exploration.

While this use of protein profiles has merit, including an evident genetic basis for the observed differences among individuals, application to personal identification appears less discriminating than using genetically variant peptides (GVPs) [1,8,9]. Present work demonstrates that GVP profiles developed using hair derived from different anatomic sites in an individual are more similar than profiles from different individuals. Thus, the hair donors were easily sorted by their GVPs regardless of the anatomic origin of the sampled hair. These inferred genotypes can be discriminatory with random match probabilities of up to 1 in 196 obtained in this study.

In the present case, some GVPs were not detected in certain profiles, producing false negatives. The detection of GVPs, as with any peptide, is dependent on meeting the chemical thresholds for fragmentation (MS2) in the mass spectrometry instrument [8]. These chemical and analytical factors are intrinsic to proteomic analysis and are unlikely to be due to underlying biological variation. An important caveat in this study is that the inferred genetic SNP alleles were not confirmed by separate DNA-based genotyping and depend on the validation gained in previous studies [1]. Biological variation does play a role, however. Occasionally a common GVP would appear only for a single subject. We hypothesize that such variation is due to protein expression level as well as GVP frequency [8]. Recent improvements in hair processing and instrumentation have increased the sensitivity of GVP detection several fold even for low levels of hair [26]. In our laboratory random match probabilities can now range up to 1 in 100 million for a single hair (data not shown), which is not yet at the level of identity. While these levels of detection exceed the values in this study, more development is still required. We predict that the detection of more GVPs in the sample will have two effects: correlation of GVP profiles within an individual will increase regardless of body location, and correlation between individuals will decrease. Investigators can therefore use hair as a source of protein-based genetic information regardless of the anatomical site of origin. This and other methodological improvements promise to yield dramatically improved estimates of random match probability and thus increase the forensic utility of hair shafts for individual identification.

CRediT authorship contribution statement

Jennifer A. Milan: Data curation, Formal analysis, Investigation, Visualization, Writing - original draft, Writing - review & editing. **Pei-Wen Wu:** Data curation, Investigation. **Michelle R. Salemi:** Investigation. **Blythe P. Durbin-Johnson:** Formal analysis, Software, Methodology, Visualization. **David M. Rocke:** Conceptualization, Funding acquisition, Methodology, Supervision. **Brett S. Phinney:** Methodology, Formal analysis, Resources. **Robert H. Rice:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Supervision, Methodology, Visualization, Writing - original draft, Writing - review & editing. **Glendon J. Parker:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Supervision, Project administration, Methodology, Visualization, Writing - original draft, Writing - review & editing.

Acknowledgements

We thank Zachary Goecker for technical assistance. This work was supported by National Institute of Justice grants 2011-DN-BX-K543 and 2015-DN-BX-K065, and National Institutes of Health grant UL1TR000002 and UL1 TR001860. The authors have declared no conflict of interest, with the exception of GJP who has a patent based on the use of genetically variant peptides for human identification (US 8,877,455 B2, Australian Patent 2011229918, Canadian Patent CA 2794248, and European Patent EP11759843.3). The patent is owned by Parker Proteomics LLC. Protein-Based Identification Technologies LLC (PBIT) has an exclusive license to develop the intellectual property and is co-owned by Utah Valley University and GJP. This ownership of PBIT and associated intellectual property does not alter policies on sharing data and materials. These financial conflicts of interest are administered by the Research Integrity and Compliance Office, Office of Research at the University of California, Davis to ensure compliance with University of California Policy.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.fsigen.2019.03.009.

References

[1] G.J. Parker, et al., Demonstration of protein-based human identification using the hair shaft proteome, PLoS One 11 (9) (2016) e0160653.
[2] K.E. Mason, et al., Protein-based forensic identification using genetically variant peptides in human bone, Forensic Sci. Int. 288 (2018) 89–96.
[3] SWGMAT, Forensic Human Hair Examination Guidelines, (2005).
[4] 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).
[5] E.H. Wyatt, J.M. Riggott, Scanning electron microscopy of hair. observations on surface morphology with respect to site, sex and age in man, Br. J. Dermatol. 96 (6) (1977) 627–633.
[6] D.W. Deedrick, Hairs, fibers, crime, and evidence: part 1 hair evidence, Forensic Sci. Commun. 2 (3) (2000).
[7] R.H. Rice, et al., Distinguishing mouse strains by proteomic analysis of pelage hair, J. Invest. Dermatol. 129 (9) (2009) 2120–2125.
[8] P.W. Wu, et al., Proteomic analysis of hair shafts from monozygotic twins: expression profiles and genetically variant peptides, Proteomics 17 (2017) 13–14.
[9] C.N. Laatsch, et al., Human hair shaft proteomic profiling: individual differences, site specificity and cuticle analysis, PeerJ 2 (2014), https://doi.org/10.7717/peerj.506.
[10] J.M. Taupin, Forensic hair morphology comparison–a dying art or junk science? Sci. Justice 44 (2) (2004) 95–100.
[11] Federal Bureau of Investigation, FBI/DOJ Microscopic Hair Comparison Analysis Review, Department of Justice: fbi.gov, 2014.
[12] M.M. Houck, B. Budowle, Correlation of microscopic and mitochondrial DNA hair comparisons, J. Forensic Sci. 47 (5) (2002) 964–967.
[13] T. Melton, C. Holland, M. Holland, Forensic mitochondrial DNA analysis: current practice and future potential, Forensic Sci. Rev. 24 (2) (2012) 102–122.
[14] K.A. Roberts, C. Calloway, Mitochondrial DNA amplification success rate as a function of hair morphology, J. Forensic Sci. 52 (1) (2007) 40–47.
[15] S.Y. Shih, et al., Applications of probe capture enrichment next generation sequencing for whole mitochondrial genome and 426 nuclear SNPs for forensically challenging samples, Genes (Basel) 9 (1) (2018).
[16] B. Schwanhausser, et al., Global quantification of mammalian gene expression control, Nature 473 (7347) (2011) 337–342.
[17] E.W. Deutsch, et al., The ProteomeXchange consortium in 2017: supporting the cultural change in proteomics public data deposition, Nucleic Acids Res. 45 (D1) (2017) D1100–D1106.
[18] M.E. Ritchie, et al., Limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res. 43 (7) (2015) e47.
[19] D.M. Rocke, et al., Excess False Positive Rates in Methods for Differential Gene Expression Analysis Using RNA-seq in bioRxiv, (2015).
[20] W.S. Torgerson, Theory and Methods of Scaling, Wiley, New York, 1958.
[21] Genomes Project Consortium, et al., A global reference for human genetic variation, Nature 526 (7571) (2015) 68–74.
[22] H. Jeffreys, An invariant form for the prior probability in estimation problems, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences 186 (1007) (1946) 453–461.
[23] Y. Shimomura, et al., Founder mutations in the lipase h gene in families with autosomal recessive woolly hair/hypotrichosis, J. Invest. Dermatol. 129 (8) (2009) 1927–1934.
[24] R.H. Rice, et al., Differentiating inbred mouse strains from each other and those with single gene mutations using hair proteomics, PLoS One 7 (12) (2012) e51956.
[25] L. Langbein, et al., Molecular characterization of the body site-specific human epidermal cytokeratin 9: cDNA cloning, amino acid sequence, and tissue specificity of gene expression, Differentiation 55 (1) (1993) 57–71.
[26] Z.C. Goecker, et al., The Optimization of Human Hair Proteomic Processing for Single Hair and Ancestral Analysis in American Association of Forensic Science, (2018) Seattle, WA Poster B-57.