Investigating sex determination through MALDI MS analysis of peptides and proteins in natural fingermarks through comprehensive statistical modelling

HEATON, Cameron, BURY, Charles, PATEL, Ekta, BRADSHAW, Robert <http://orcid.org/0000-0003-1533-2166>, WULFERT, Florian, HEEREN, Ron, COLE, Laura <http://orcid.org/0000-0002-2538-6291>, MERCHANT, Leeanna, DENISON, Neil, MCCOLM, Richard and FRANCESE, Simona <http://orcid.org/0000-0002-1381-1262>

Published version

HEATON, Cameron, BURY, Charles, PATEL, Ekta, BRADSHAW, Robert, WULFERT, Florian, HEEREN, Ron, COLE, Laura, MERCHANT, Leeanna, DENISON, Neil, MCCOLM, Richard and FRANCESE, Simona (2020). Investigating sex determination through MALDI MS analysis of peptides and proteins in natural fingermarks through comprehensive statistical modelling. Forensic Chemistry, 20 (100271), 1-12.

Copyright and re-use policy

Investigating sex determination through MALDI MS analysis of peptides and proteins in natural fingermarks through comprehensive statistical modelling

Cameron Heaton[a#], Charles S. Bury[b#], Ekta Patel[b], Robert Bradshaw[a], Florian Wulfert[c], Ron M. Heeren[d], Laura Cole[a], Leeanna Marchant[e], Neil Denison[f], Richard McColm[g] and Simona Francese[a*]

[a]Centre for Mass Spectrometry Imaging, Sheffield Hallam University, UK

[b]Medicines Discovery Catapult, UK

[d]Biosciences and Chemistry Department, Sheffield Hallam University, UK

[d]The Maastricht MultiModal Molecular Imaging Institute (M4I), Maastricht University, The Netherlands

[e]Forensic Enhancement Laboratory, Yorkshire & Humber Regional Scientific Support Services, UK

[f]Identification Services Yorkshire and the Humber Region, West Yorkshire Police, UK

[g]Defence Science & Technology Laboratory, Porton Down, UK

[#] these authors have equally contributed to the work presented in this paper

*Corresponding author: Simona Francese, Sheffield Hallam University, Howard Street S1 1WB, Sheffield, UK s.francese@shu.ac.uk

Abstract

In the last decade, Matrix Assisted Laser Desorption Ionisation Mass Spectrometry (MALDI MS) has proven to be a valuable analytical tool in forensic research as it can detect and map molecular information of forensic relevance in trace evidence such as fingermarks and hair. The first published proof of concept demonstrating that it was possible to differentiate males and females from the peptide and protein content of their fingermarks was published in 2012. In that work, MALDI MS was used in Profiling mode (MALDI MSP) to quickly obtain spectral profiles of ungroomed marks. These were submitted to Partial Least Square Discriminant Analysis (PLS-DA) yielding sex discrimination with an accuracy between 67.5% and 74.4%, if harsh classification criteria were applied. Since then, this research has progressed to investigate the opportunity to increase the accuracy of prediction in natural marks (obtained with no preparation of the fingertip prior to deposition) either unenhanced or enhanced prior to matrix application and MALDI analysis. Extensive statistical modelling has been employed to determine the model with the highest sex predictive accuracy. Results show that in natural marks the presence of polymers (as external

contaminants) in fingermarks affects the peptide/protein signals to various degrees and only by using one type of scoring system a method has been identified to provide up to 86.1% predictive power in discriminating female from male marks.

Keywords: fingermark, natural, MALDI MS, sex, machine learning

1.0 Introduction

Fingermarks are the most commonly collected evidence from crime scenes and remain the major contributor to suspect identification. Any intelligence that can be extracted from the chemical make-up of a fingermark could greatly inform investigations, contributing to narrowing down the pool of suspects. The sex of an individual is one such type of important intelligence which can only be currently gathered, if requested, through DNA. However, DNA is sensitive to environmental conditions, susceptible of degradation and in some cases, found in too low abundance to yield any information. By 2012, a few groups had unsuccessfully attempted the determination of the sex of an individual by exploiting the lipid content of fingermarks and using Laser Desorption Ionisation (LDI) Mass Spectrometry [1,2].

In recent years, Halamek's group used a different approach to the problem by exploiting the presence of amino acids in fingermarks. In a first publication, the authors made an encouraging claim that by using a biocatalytic assay to measure the absorbance of the complex formed between amino acids and the L-AAO/HRP bioassay, it was possible to obtain a 99% predictive value for correct discrimination of male and female fingermarks [3]. However these results were obtained for artificial 'fingermark mixtures' and no predictive value is reported for actual natural fingermarks or previously enhanced marks. In a second study by the same group [4], enhanced marks were considered by using 1,2-Indanedione, an amino acid reagent which is gaining *momentum* for fingermark enhancement on paper [5,6]. Fluorescence measurements, taken upon formation of a complex between 1,2-Indanedione and amino acids, were used to discriminate the female and male donors' fingermarks. This method would have the benefit of being quick and simple; it would also be multi-informative as, in one measurement, it could provide the biometric information (the ridge pattern) and the sex of the donor. The authors have acted on the prior knowledge that the quantities of amino acids in sweat are expressed twice as much in females than in males. Although this 'prior knowledge' has not been referenced in the literature by the authors, it seems to be confirmed as the fluorescence signal obtained in female marks is, in fact, twice as much as that obtained in male marks.

Whilst the approach appears to be no doubt interesting, it does require further investigation as this second study only employed 4 fingermarks and no actual

statistical analysis details are reported. In addition, further understanding the relationship between fluorescence signal and pressure variability in fingermark deposition is required. Recently, the application of Desorption Electrospray Ionisation (DESI) Mass Spectrometry Imaging (MSI) yielded the discrimination of female and male donors' fingermarks (2 female and 2 male donors for a total of 12 fingermarks) exploiting their lipid composition [7]. Principal Component Analysis (PCA) was applied to statistically treat the data. Although the first two components demonstrated good separation of male and female marks, only cross validation data have been reported and, crucially, only 4 donors were used, with no prior application of an enhancement technique. Additionally, no details are reported on the type of marks employed.

Therefore, to date there is no method that can actually be considered viable to determine sex from fingermarks.

In a different approach published in 2012, Ferguson *et al* [8] provided the first encouraging proof of concept employing Matrix Assisted Laser Desorption Ionisation Mass Spectrometry (MALDI MS) to detect endogenous peptides and small proteins in fingermarks. These species belong to eccrine secretions and, as such, peptides and proteins exist in an aqueous environment composed of ~98% water and ~2% inorganic species (electrolytes) as well as organic molecules (amino acids and urea [9]). Recently, the quantity of peptides and proteins in a fingermark has been estimated as ranging between 0.2 - 51.0 µg [10]. Despite the relatively low abundance, especially compared to that of lipids, Ferguson *et al*. [8] demonstrated that by combining MALDI MS detection of peptides and proteins and Partial Least Square Discriminant Analysis (PLS-DA), sex discrimination was possible with an accuracy of prediction between 67.5 and 74.4% (or 85% if less harsh classification criteria were applied). However, this study only collected a relatively small number of donors (80). It also excluded participants from the study if falling outside of the 18-45 years age range and if they had taken drugs or medications within the two weeks prior to the finger smear collection. Furthermore, *ungroomed* fingermark smears were employed, that is, artificially prepared marks with depleted lipid and contaminant content [11]. These types of marks are a useful model to investigate feasibility of a method but they do not reflect the variable molecular composition of the real crime scene marks species [12]). In the present study, it has been investigated whether a comparable achievement of high predictive power can be

achieved by employing a larger, more representative cohort of donors, and by using: (i) natural marks (deposited without prior preparation of fingertips) and (ii) from enhanced natural marks. In the current investigation, the results of two consecutive studies are reported (study 1 and study 2). Firstly, in study 1, the sample size was expanded from 80 [8] to 199 donors employing natural marks, with no exclusion criteria for donor participation (except for individuals below 18 years of age) and both with and without prior fingermark enhancement. In order to enhance fingermarks, gold vacuum metal deposition (VMD) was employed. Although this study did not yield meaningful results as later discussed in this paper, it informed the design of subsequent research (study 2). In study 2 172 donors donated natural marks which were analysed by MALDI MS both with and without prior enhancement by white powder.

MALDI mass spectra from study 2 have been subjected to extensive statistical modelling, and the sex discriminating power of a range of machine learning classification models has been evaluated. Furthermore, a systematic analysis of multiple peak picking strategies, classification model choices, feature selection and scoring criteria approaches has been performed. Ultimately, this processing strategy has proved the overall approach to be practically applicable in the triaging of crime scene marks.

2.0 Experimental

2.1 Materials

ALUGRAM® SIL G/UV254 pre-coated aluminium sheets, α-cyano-4-hydroxycinnamic acid (α-CHCA) and trifluoroacetic acid (TFA) were purchased from Sigma-Aldrich (Gillingham, UK). Acetone, acetonitrile and methanol were purchased from Fisher Scientific (Loughborough, UK). Doubled sided conductive copper foil shielding tape was purchased from 3M (St. Paul, MN, USA). Indium tin oxide (ITO)-coated slides were purchased from Delta Technologies Ltd. (Loveland, CO, USA). Sirchie Indestructible White “Hi-Fi” Volcano latent fingerprint enhancement powder was purchased from WA Products (Essex, UK). The Breeze™ single-use fibreglass zephyr brushes were purchased from SceneSafe (Essex, UK).

2.2 Instrumentation and instrumental parameters

For study 1 data, three acquisitions were made per mark on each matrix spot on an Autoflex MALDI TOF-TOF mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany) equipped with a 200 Hz SmartBeam™ laser, selecting the laser spot size as 100 x 150 µm. Calibration was performed between each separate set of donor samples using Bruker Protein Calibration Standard I. Acquisition was carried out in the mass range 1,000 - 12,474 Da in positive linear mode. Each spectrum was obtained through a total of 900 shots (300 shots from each of the three replicates).

For study 2, data were acquired using two Bruker Rapiflex MALDI TOF-TOF mass spectrometers (Bruker Daltonik GmbH, Bremen, Germany) equipped with a neodymium-doped yttrium aluminium garnet (Nd:YAG) 355 nm SmartBeam^{TM} 3D laser operating at a repetition rate of 10 kHz. In this instance, the laser spot size was 95 x 95 µm. Calibration was performed between each separate set of donor samples using a Bruker Protein Calibration Standard I mixed with adrenocorticotropic hormone corticotropin-like intermediate peptide (ACTH CLIP) fragment [18-39] to provide a calibration range from 2,465 - 12,361 Da. Calibration was performed in quadratic mode with at least 4 calibration points. Acquisition was carried out in the mass range up to 12,600 Da in positive linear mode. Each spectrum was obtained through a total of 30,000 shots (10,000 shots from each of the three replicates across the mark) in a random walk mode.

2.3 Methods

2.3.1 Fingermark collection

For study 1, 6 natural fingermarks were collected from each of the 199 participants rubbing their fingertips together prior to deposition on pre-cleaned TLC sheets (aluminium slides), totalling 1,194 samples. The marks were frozen at -80ºC in microscope slide boxes. For study 2, 6 natural fingermarks were collected from each of the 172 participants rubbing their fingertips together before depositing three prints from both their right and left hands on two separate aluminium slides, totalling 1,032 samples. The slides were immediately placed inside microscope slide holder boxes and kept at room temperature until analysis. For both study 1 and 2, fingermarks were collected from participants from West Yorkshire Police and Sheffield Hallam

University in accordance with approved SHU ethics applications HWB-BRERF014/11 (study 1) and ER17244422 (study 2).

2.3.2 Fingermark enhancement

In study 1, the marks were cut in half length ways, and one half visualised with VMD, using gold nano-powder according to the protocol detailed by the Fingermark Visualisation Manual [5]. In study 2, 516 marks remained unenhanced, whereas 516 were enhanced using Sirchie Indestructible White enhancement powder. Each donor fingermark set was enhanced using a separate disposable zephyr brush which was then disposed of to avoid cross-contamination.

2.3.3 Fingermark preparation for MALDI MSP

For study 1, samples were defrosted prior to analysis. As the aluminium sheets were removed from the glass slider containers, they were placed in a vacuum desiccator for 10 minutes and then subsequently mounted on the MALDI plate using double sided carbon tape for analysis on an Autoflex MALDI TOF-TOF (Buker Daltonics GmbH, Bremen, Germany). α-CHCA matrix (5 mg/ml in 25:25:50 ACN:Eth:0.5% TFA_{aq}) was spotted in three locations directly onto the mark. 0.5 µL calibration standard was pipette mixed with matrix directly on a blank area of the slide. For study 2, the room temperature stored aluminium sheets were removed from their glass slide containers and mounted on conductive ITO slides with double-sided copper tape. 0.5 µL calibration standard was pipette mixed with 0.5 µL α-CHCA matrix at 5 mg/mL in 70:30 ACN:0.5% TFA_{aq} in the top centre of the slide. For each set of marks (non-enhanced and enhanced), three spots of 0.5 µL of α-CHCA were then deposited down the middle of each mark.

2.4 Statistical analysis

2.4.1 Study 1

2.4.1.1 Pre-processing

For study 1, the three spectra for 1 donor's fingermark were kept separate and used as technical replicates. Smoothing, baseline correction and peak picking parameters were selected in FlexAnalysis (Bruker's proprietary data processing software) with S/N >3.5. Spectra were visually assessed for polymer contamination and minimum intensity threshold and outliers were removed. FlexAnalysis spectra files were exported and converted to CSV-format files and the 2 header lines removed with a

FlexAnalysis method script. The CSV files were converted to TXT files and imported into Markerview mass spectrometry processing software (Sciex, AB Sciex Pte. Ltd, Warrington, Cheshire) as centroided spectra with a mass tolerance of 0.5 Da. The filtering minimum response was 0.1 au and the maximum number of peaks limited to 1,000. PCA was applied with Pareto scaling. The whole dataset was then exported to Matlab (The Mathworks, Inc., Natick, MA, USA) for statistical analysis with the PLS_Toolbox (Eigenvector Research, Inc., Manson, WA, USA).

2.4.1.2 Statistical processing

Normalized and mean centred spectra were used to build a PLS-DA classification model for the classes male and female. The data was split into a training and validation set, ensuring that all spectra of each fingermark were either in one or the other and that the male/female distribution was comparable for both. Internal cross-validation (4-way split venetian blinds) was applied to assess ideal model complexity at 4 latent variables. Classification errors were quantified for model fit and cross validation accuracy during the cross-validation stage and the independent validation set was then used to estimate the prediction error for unknown spectra. To increase stability and accuracy, another model was built, using the same datasets but based on only those spectral variables with higher importance to the first model (Variable Importance on Projection selection method, VIP). The same assessment of model fit, CV error and prediction accuracy was applied.

2.4.2 Study 2

2.4.2.1 Pre-processing

For study 2, three marks were obtained per donor, with each mark corresponding to a single mass spectrum, obtained by averaging three individual mass spectra acquired from three distinct matrix spots. The effects of two pre-processing strategies have been compared: (a) the three marks per donor have been kept separate for statistical processing, and (b) the three individual spectra per donor have instead been combined in a single averaged spectrum.

Prior to statistical analysis with SCiLS Lab software (v 5.12.0, Bruker Daltonics, GmbH, Bremen, Germany), the mass spectra were converted to imzML-format files using R (v 3.6.1, R Foundation for Statistical Computing, Vienna, Austria) and the

MALDIquant package, and then imported into SCiLS Lab software. All spectra were baseline corrected (tophat) at 200 interval width.

Each of the data sets (natural and natural enhanced fingermarks) were averaged and imported into mMass (v 5.5.0) [13], where peak picking, smoothing and baseline correction were carried out to identify signals of interest. A S/N threshold of 1:1 was selected to enable peak picking of a sufficient number of peaks to perform statistical processing. This peak list was then imported into SCiLS Lab and used as reference *m/z* positions for LDA classification. An overall peak list was generated using mMass, and, again, separate trial peak lists were exported to SCiLS Lab for both the S/N >2:1 and 1:1 levels.

Prior to further analysis of sex classification models using Python (and separate to the pre-processing steps for SCiLS Lab described above), an R script was written to perform initial sample processing with the MALDIquant package, including baseline correction, TIC normalisation and spectral smoothing. Peak picking was also performed here using MALDIquant. By varying both the required S/N parameter and required minimum occurrence rate of each peak across the sample spectra (denoted “minFreq” in the following text), the effect of different peak picking strategies on the performance of the resulting sex classification models was monitored (Table 1). Peak picked cross-sample spectra arrays were then exported to CSV-format for subsequent model training.

	Required fraction of spectra needed for a peak to be included in the analyses (minFreq)			
S/N	**1%**	**10%**	**50%**	**90%**
2	2867	2764	141	6
3	2570	1773	61	-
5	869	428	27	-
10	335	90	6	-
20	168	19	-	-

Table 1. Variable peak picking strategies used in MALDIquant yield different counts of peaks for downstream analysis. S/N values vary between 2 to 20 (vertical axis), and the required fraction of spectra needed for a peak to be included vary between 1% and 90% (horizontal axis) denoted 'miniFreq'.

2.4.2.2 Statistical processing

A 10-fold supervised Linear Discriminant Analysis (LDA) classification was performed in SCiLS Lab on both the unenhanced and enhanced datasets. As a requirement of the LDA method, it has been assumed that all data is sampled from a multivariate normal distribution, with different means in each sex group, but identical covariance across the groups. In the training phase, the group means and dataset covariance are calculated. For classification, the spectra are subsequently assigned to the group that maximises the probability.

In a parallel approach that was conducted separate to the SCiLS Lab analysis, the relative performance of a range of different classification model types has been systematically investigated (Table 2). All machine learning has been implemented in Python, utilising the scikit-learn and XGBOOST packages, with k-fold cross validation (CV) being used to assess the relative performance of each model. CV performance has been performed with k set to 5, 10, 25, and 50, to determine the robustness of the reported accuracy scores between different k *values*. During each individual CV k-fold split, care was taken to ensure that all three marks (technical replicates) from a donor were assigned to either the train or test set.

Model name	Reason for inclusion
LDA	Standard classification model available in the SCiLS Lab software
Random forest	Decision tree-based ensemble learning classification method
Naïve Bayes	A standard baseline model, typically used for benchmarking purposes
XGBOOST	A tree-based ensemble method using gradient boosting that has recently gained large popularity
PLS-DA	Classification method used in the previous study by Ferguson *et al*., (2012)[4]

Table 2. Overview of classification models used in the current analysis

In addition to using all picked *m/z* positions as inputs to the classifiers in Table 2, the impact of several feature selection strategies (Table S1) on the performance of each model has been investigated. The following feature selection methods have been included: (a) the PLS-DA VIP (Variable Importance in Projection) score, (b) random forest feature importance (based on the decrease in Gini impurity associated with each feature across trees) and (c) univariate feature selection via a chi-squared test

to identify *m/z* peaks that are most likely to be dependent on sex. For each k-fold CV split, only the training subset of samples was used for feature selection, to ensure that each test set of samples remained hidden throughout cross validation.

Neural network implementation for multi-class classification - A dense neural network architecture was implemented for joint prediction of sex and contaminants for the unenhanced fingerprint dataset using the Keras python package (with Tensorflow backend). The architecture consisted of two fully connected dense layers (of 100 and 10 nodes respectively), incorporating 20% dropout and Relu activation layers between each dense layer. Two contrasting output layer designs have been compared: (i) an output layer consisted of two nodes with softmax activation for sex prediction only, and (ii) an output layer consisted of two pairs of two nodes, with each pair having an independent softmax normalisation for sex and contamination state prediction. In both cases, the model was trained for 100 epochs with a categorical cross-entropy loss function (one separate loss function for each pair in (ii)).

3.0 Results and Discussion

The study published by Ferguson *et al*. [8], offered an encouraging proof of principle that MALDI spectral profiles of peptides and proteins could differentiate between "female and male fingermarks". However, the study design did not permit the evaluation of the potential for implementation in the forensic fingerprinting workflow. The study employed a small donor cohort (80 donors), and lipid and contaminant depleted marks. It excluded participants below 18 and over 45 years and those taking medications up to two weeks prior to the fingermark collection. Also, there was no consideration of the prior application of a CSI fingermark enhancement technique. All of these design features prevented any evaluation on the robustness of the method and on the real opportunity for casework application. For these reasons, study 1 was subsequently devised where marks from 199 donors were collected with no exclusion criteria and no prior treatment of the fingertip (thus producing "natural" marks). In addition, half of the marks were subjected to gold vacuum metal deposition (VMD), a technique normally employed by crime labs in high profile cases to visualise marks on semi-porous and non-porous surfaces.

Gold nanoparticles have been previously found to enhance the MALDI signal due to laser energy trapping by gold nano-islands [14]. Bradshaw *et al*. [15] found that gold deposited through VMD boosted the ion signals of small molecules and enabled higher quality of the molecular images of the fingermark ridge pattern. MALDI MS Imaging compatibility with a commonly applied fingermark enhancement technique was proved, as well as its ability to increase the clarity of the ridge pattern (compared to that obtained by the sole application of VMD) for older marks.

However, within study 1, VMD caused ion suppression of peptides and proteins (Figure 1). Around 20% of the marks exhibited the peptide and protein spectral profile expected (Fig. 1A) although the ion population was reduced in comparison with spectra observed in Ferguson *et al.* [8] Around 30% of the spectra were dominated by gold nanoparticle clusters where the typical difference of 187 mass units were observed between the peaks (Fig. 1B) and around 50% of the spectra exhibited complete suppression of the protein/peptide ion signals with just gold nanoparticle clusters being detected (Fig. 1C).

It is speculated that these instances reflect the distribution in the participants' cohort of what are defined "good, average and bad donors" [16]. These attributes refer to the abundance (and nature) of endogenous species secreted into the sweat that is then transferred in a fingermark upon contact of the fingertip with a surface. It may be possible that a different gold layer thickness would improve the detection of lower abundance peptides and proteins in certain donors. However, even if this was possible, it would not be applicable in an operational context due to the required compliance with accredited processes. Given the above results, only the natural unenhanced marks were subjected to statistical analysis.

During internal cross validation, the classification errors varied between ~20% for those spectra used to build the models and 40% for the left out cross-validation samples. The prediction error of the independent validation set was worse again at 45%, approaching accuracy of random guess. Receiver operating characteristic (ROC) analysis was used to evaluate the performance of the method and estimate the probability of distinguishing between the male and female samples groups. ROC analysis involves changing the threshold (for example, above which the mass spectral characteristics fit the female group) and observing the effect on the predictive power of the model to produce a ROC curve. The area under the ROC curve (AUC) represents the probability that the diagnostic test will correctly

distinguish between the male and female samples so that the larger the AUC, the higher the probability that each sample will be identified correctly. As it can be seen from Figure 2, the predictive power was only slightly better than "a flip of a coin" as males and females were classified with a 55% accuracy of prediction only.

As the PLS-DA model used less than 20% of the spectral variance, a second model was built by selecting only those spectral variables that had most influence on the first model, using the VIP selection method. The resulting model did not show any improvement on the first model. This led to the conclusion that across all spectral variables there was too much random variance between spectra from the same fingermark to be able to build stable and sex relevant models. Individual inspection of the raw spectral data appeared to confirm this conclusion. These results were disappointing considered the encouraging proof of concept published in 2012 [8]. Failure to correctly classify donor's sex could be due to higher variability in the fingermark composition, typical of natural marks, including the variable presence and amount of potentially ion suppressing contaminants. It was also possible that the removal of restriction criteria to donors' participation could have contributed to fingermark compositional variability to a level which the statistical model was not able to cope with. If these were the reasons, then this study showed that sex discrimination from fingermarks is not possible in "uncontrolled" conditions using endogenous peptides and proteins.

However, additional variability may have been unintentionally introduced in the system due to: (i) two different operators acquiring data from fingermarks, (ii) the variability in instrumental performance and environmental conditions over an extended 4 month period (the overall time needed to complete the analysis of all of the marks) and (iii) sample compositional changes or molecular migration possibly determined by rapid condensation during the fingermark sample defrosting step.

It was speculated that the presence of ion suppressing contaminants could be attenuated by higher sensitivity instruments than the Autoflex MALDI TOF-TOF (Bruker Daltonik GmbH, Bremen, Germany). From a processing point of view, in this study, the three MALDI MS spectra acquired in three different areas of the same fingermark were used as technical replicates. This approach might have introduced further variability in the system as the matrix could have well been deposited in a fingermark area of lower protein/peptide content. Ion intensity variability, due to a combination of relatively low number laser shot accumulation and the known "sweet

spot" phenomenon in MALDI (lack of shot to shot reproducibility due to uneven analyte-matrix co-crystallisation), would have further impacted spectral reproducibility. As there was a significant possibility that failure of study 1 was determined by a combination of sub-optimal study design and the time taken to analyse all the samples, a further study (study 2) was outlined.

In study 2, 172 donors donated a total of 1,032 natural marks (6 per donor) which, this time, were kept at room temperature prior to analysis, to avoid any possible freeze-thaw or condensation issues upon defrosting. Furthermore, a single operator acquired all the MALDI MS spectral profiles and over the course of just over 1 month to minimise environmental and fingermark compositional variability. From a processing point of view, given that the three matrix spots had been acquired from the same donor, a reasonable approach was to average the three MALDI MS spectra to account for both the "sweet spots" phenomenon and fingermark compositional variability. In an additional approach, the spectra from the individual donor's marks (3 in total) were also combined in one averaged spectrum. In study 2 VMD enhancement was replaced with an enhancing powder to produce a set of enhanced marks in addition to a set of unenhanced marks. Fingermark powders are commonly used at crime scenes. Amongst the three powders tested, namely white (titanium dioxide-based), black (carbon based) and aluminium powders, the titanium dioxide-based white powder enabled the detection of the ions of interest, the ion population of which rivalled that of the same unenhanced split fingermark in preliminary investigations. In addition, this powder was found to even increase the ion intensity with respect to its unenhanced counterpart in split marks (data not shown) and was therefore selected for this study. It is important to highlight that identification of biomarkers of sex by MALDI MS/MS or by any better suited technique such as LC MS/MS was outside the scope of this study, both in the initial design and on the basis of the results obtained.

Unenhanced natural marks - a single average spectrum obtained from each of the three fingermark areas analysed via MALDI was viewed in mMass where baseline subtraction, smoothing and peak picking were applied. A signal-to-noise threshold of 3:1 generated a peak list of 29 peaks. This was not enough signals to attempt a discriminant analysis (LDA). When the signal-to-noise threshold was lowered to 2:1, automatic peak picking generated a peak list of 48 peaks which could

be used for linear discriminant analysis in SCiLS Lab. The spectra from the individual donor's marks (3 in total) were combined in one averaged spectrum (3 mass spectra per donor). The 10-fold LDA classification yielded an accuracy score of ~60%. It was noted that the separate treatment of the three marks per donor has limited the validity of the SCiLS Lab cross-validation calculation (illustrated in Figure S1): the information that there are three sample marks per donor has not been supplied to SCiLS Lab, and the separate marks of one person can be assigned to both the train and test splits simultaneously during cross validation (Figure S1, middle row). This is not representative of the ability of the model to generalise to unseen populations.

To bypass this limitation, further sex classification models have been developed externally to SCiLS Lab, using a modified CV approach that ensured that all three marks per donor were assigned to either the train or test set, per k-fold split. During cross validation, two accuracy measures have been calculated to assess performance: (i) *the separate mark scoring strategy*, in which each mark from each donor is treated separately and the ability of the model to correctly label each mark in the test fraction of samples is assessed and (ii) *the majority vote scoring strategy*, in which for each donor in the test set, a sex prediction is made for each mark separately and the majority vote of the donor's three marks is then taken. As such, if donor *X* has 3 marks {M1, M2, M3}, which a model then assigns the labels of {Male, Male, Female}, then overall the person is labelled as 'Male', and this prediction contributes only once to the accuracy measure. The application of the majority vote scoring scheme was motivated by the previous study by Ferguson *et al.* [8]. However, it was noted that, since the number of spectra per donor is less in study 2 (3 spectra per donor) compared to the previous study (9 spectra per donor), this majority voting scheme is not directly comparable.

Ferguson *et al.,* [8] have previously demonstrated that feature selection via VIP scoring, in conjunction with a PLS-DA classification scheme, yielded superior predictive performance to PLS-DA alone. As reported in the methods section, study 2 extends this analysis to compare a range of distinct feature selection strategies to determine their potential benefit on model predictive performance (Table S1). Figure 3 illustrates the effect of feature selection for the case of the VIP scoring scheme. For each peak picking strategy, the *m/z* positions that attain the highest VIP scores

(in red) are those that are least correlated between the male (*x*-axes) and female (*y*-axes) samples, on average.

Figure 4 and Figure S2 illustrate the overall k-fold accuracy scoring results for each model type and distinct peak picking strategy. The computed k-fold cross validation accuracy scores appear to be relatively robust to: (i) the choice of k (5, 10, 25 or 50) for cross-validation, (ii) the choice of feature selection strategy and (iii) the choice of scoring metric to account for the presence of technical replicates. However, a positive trend is observed across models between numbers of included *m/z* peak positions and mean CV accuracy scores. Overall, XGBOOST appears to perform superior to other trialled methods, however the performance boost is minimal, with a k-fold accuracy of 60-70% attainable by all model types for particular peak picking strategies (Table 3). The maximum reported k-fold CV score was attained by the XGBOOST model at 70.9% (3 s.f.), with k=25, and a peak-picking strategy of {S/N=5, minFreq=0.1}, under the majority scoring score and VIP feature importance selection strategy (as used in [8]) .

Model	**Peak picking strategy**	**Number of k-folds**	**Feature selection strategy**	**Accuracy score (mean calculated over k-fold repeats)**
XGBOOST	minFreq = 0.1 S/N = 5	25	PLS-DA VIP score	0.709
XGBOOST	minFreq = 0.1 S/N = 5	50	PLS-DA VIP score	0.705
PLS-DA	minFreq = 0.01 S/N = 10	50	None	0.703
XGBOOST	minFreq = 0.01 S/N = 5	10	PLS-DA VIP score	0.7
XGBOOST	minFreq = 0.1 S/N = 5	50	None	0.7

XGBOOST	minFreq = 0.5 S/N = 2	10	None	0.699
XGBOOST	minFreq = 0.01 S/N = 20	10	PLS-DA VIP score	0.696
Random forest	minFreq = 0.01 S/N = 5	10	None	0.695
XGBOOST	minFreq = 0.5 S/N = 2	10	PLS-DA VIP score	0.694
XGBOOST	minFreq = 0.01 S/N = 5	50	PLS-DA VIP score	0.692
PLS-DA	minFreq = 0.1 S/N = 5	25	None	0.690
PLS-DA	minFreq = 0.01 S/N = 10	10	None	0.688
PLS-DA	minFreq = 0.01 S/N = 5	50	Random forest feature importance	0.688
PLS-DA	minFreq = 0.1 S/N = 5	10	None	0.688
Random forest	minFreq = 0.01 S/N = 5	25	None	0.688
XGBOOST	minFreq = 0.1 S/N = 3	25	None	0.686
XGBOOST	minFreq = 0.01 S/N = 10	5	None	0.683
LDA	minFreq = 0.5 S/N = 3	25	Univariate feature selection	0.683
XGBOOST	minFreq = 0.01 S/N = 3	10	None	0.683

XGBOOST	minFreq = 0.1 S/N = 5	10	PLS-DA VIP score	0.682

Table 3. Top 20 scoring models (average score calculated over k-fold repeats, with k specified in the ‘Number of k-folds’ column). The scoring scheme was "Majority" for all models. Where the feature selection strategy is specified as “None”, all peak-picked peaks have been used for model training.

Within the current analysis using 3 spectra per donor, the maximal XGBOOST accuracy (~71%) roughly corresponds to the 2/3 majority vote scoring employed as one of the harsher classification criterion in Ferguson *et al.,* [8]. In the 2012 study [8] this harsher classification criterion yielded a score of 67.5% for the PLS-DA classifier employing 9 spectra per donor, with the distinction that, due to the larger number of available spectra per donor, it was additionally possible to define an ‘undecided’ sex label in addition to ‘male’ and ‘female’.

It was further assessed whether sex-classification performance could be improved by instead training on 1 averaged mass spectrum per donor (i.e. 1 training instance per person), as opposed to the 3 separate spectra per donor. This alterative training strategy did not yield better model performance than the original training strategy (data not shown).

An "on the hoof" PCA analysis identified four samples of the dataset as outliers, presenting themselves in a tight cluster in the PCA plot (data not shown). Upon inspection, the corresponding mass spectra showed to be dominated by a number of PEG-based polymers completely suppressing protein and peptide signals in these marks in the mass range 2,000-5,000 Da (Figure S3). These four samples were all from female donors. In their participant questionnaires they reported the use of toiletries including shower gel, soap, moisturiser, foundation, hair products and body lotions. Seventy-eight out of ninety females reported the use of some type of toiletry/cosmetic products but only these four samples did not have a sufficient number of relevant protein/peptide ion signals for classification. To assess the contribution of these fingerprints to the classification model performances, these four

outlier fingerprint samples were removed from the dataset and the model training/evaluation repeated. Overall omission of these spectra did not lead to any clear improvement in model sex classification performance (data not shown).

A more careful inspection of the spectra revealed a higher number of instances in which polymers were present, although to a lesser extent such that they did not hamper classification. Therefore, it was assessed whether improved classifier performance could be attained by directly accounting for such contaminants during model training. Here, each sample was assigned a binary state of ‘contaminated’ or ‘non-contaminated’ (a categorisation that split the dataset approximately into two equal sized subsets), and each classifier in Table 2 was re-trained to classify donor samples as one of four categories: (i) male sample and not contaminated, (ii) female sample and not contaminated; (iii) male sample and contaminated and (iv) female sample and contaminated (Figure 5). As such, the classifier was explicitly presented with a distinction between contaminated and non-contaminated samples during training, as opposed to being required to learn this distinction unaided. Figures S4 and S5 provide the resulting sex classification performances for the XGBOOST and LDA classifiers, respectively (other classifiers exhibited similar results). Overall, the inclusion of contamination state could not be concluded to perform consistently worse/better than the sex classification models reported in Figure 4.

The classification schemes in Table 2 rely on the output class labels being disjoint (as shown in Figure 5). If one disjoint class (e.g. contaminated female samples) is underrepresented, or if the spectral dissimilarity is low between contaminated and non-contaminated instances, it is suggested that any benefit from the inclusion of explicit contamination information in model design could be mitigated. It was investigated whether a multi-label neural network-type architecture (depicted in Figure 6), which can instead be constructed to make predictions against multiple, non-disjoint properties simultaneously, might yield superior predictive power through the incorporation of contamination information.

It is important to note that whereas Figure 6a is analogous to Figure 5a, the 4 output states in Figure 6a are not identical to Figure 5b. Figure S6 illustrates a direct comparison of neural network model performances for the model architectures presented in Figure 5a (in blue) and 5b (in orange). The data displayed in Figure S6

are the results from k=5 cross validation; other k-fold results were qualitatively similar and are not shown. Overall, there is insufficient evidence to indicate that a neural network architecture that accounts for sample contamination can surpass the previously investigated model schemes presented above in Table 2.

Enhanced natural marks- the statistical modelling applied to the unenhanced fingerprint set was subsequently applied to the twin set that was enhanced by white powder (except those employing a 4-output classification and the multi-class classification using the neural network architecture). In particular, each model has been trained using 3 separate fingerprints (3 averaged spectra) per donor. Each model has been trained/evaluated using k-fold cross validation, with variable k. The range of peak picking strategies in Table S1 have been trialled, based on variation of the S/N cut-off and minimum required peak occurrence rates through use of the MALDIquant R package. The donors that had previously shown polymer contamination in the natural marks did not show this level of contamination in the enhanced marks. It could be speculated that as the natural marks were deposited first with one hand, and the enhanced marks second with the other hand, donors are likely to use their dominant hand first; this would also be likely to be the same hand they would apply toiletries and cosmetics with. All the samples from this dataset had been subjected to statistical modelling. The data shown in Figure S7 indicate that the powdering of fingerprints does not generally lead to either an improvement or a reduction in the performance of models, when evaluated over a range of peak-picking strategies and model types.

In Ferguson *et al.* [8], 9 spectra per fingermark (rather than 3 per mark in study 2) were acquired (and treated as technical replicates differently from study 2 in which the 3 spectra per fingermark per donor have instead been averaged). The accuracy of prediction reported by Ferguson *et al*. [8] was 85% during cross validation. However, for most of the donors some of their 9 spectra were predicted as the "opposite sex"; therefore, If the method accuracy is expressed as accuracy for each single spectrum, one would arrive at a correct prediction of 68.9% of female and 74.4% of male spectra.

In conclusion, despite the larger compositional variability that can be anticipated in natural marks, the accuracy of prediction for the second of the two studies reported here (study 2), using the best performing classification model is similar to that

yielded by Ferguson et al [8] (although accuracies are not completely comparable due to slight differences in the number of spectra per mark and how these were treated).

The use of much more sensitive and rapid MALDI instrumentation (Bruker Daltonik Rapiflex (study 2) versus Autoflex (study 1)) has been observed to generate spectra containing a higher number of peptides/protein peaks with an intensity of at least 2 orders of magnitude higher. This is partly due to the higher laser repetition rate (10,000 Hz) enabling the accumulation of 30,000 laser shots in this study versus 300 in study 1 with the Autoflex, which has improved the S/N (Figure 7). It is also possible that a larger spot size would have led to an even ion higher intensity [16] but this hypothesis was not tested. Though the higher sensitivity of the Rapiflex instrumentation may have mitigated the much higher compositional variability exhibited by natural fingermarks, it did not allow a substantial improvement in the accuracy of prediction yielded by Ferguson *et al.* study [8]. Therefore, whilst on a strictly technical point of view the method appears to be robust and resilient to the variable nature of natural marks and the application of a prior fingermark enhancement technique, operationally it would now allow an exclusion of a suspect/suspect sex from a investigation.

Due to the use of personal care products, the presence of the polymers can be more or less prevalent in the spectrum, in some cases partially or completely suppressing the peptides/protein signals. Therefore it was hypothesised that for natural marks the presence of polymers was the major factor leading to operationally low predictive power. Indeed, it has been calculated that, on average ~40-45% of polymer-contaminated spectra are incorrectly predicted (this value range was computed for the XGBOOST classifier, k=5 folds, across each of the peak picking S/N strategies). However, within the fraction of non-polymer contaminated spectra, the median percentage of those incorrectly classified is also between 40-45% depending on the peak picking strategy. As such, there is insufficient evidence to suggest that, in the 30% of misclassified samples by the model, there is an enrichment of polymer contaminated spectra. Even if machine learning algorithms "can learn" to detect and exclude the influence of polymer signals during prediction, incorrect classification still occurs if the remaining ion signals in the spectrum are too few and not discriminatory.

However, interestingly, a third scoring strategy, denoted here as '*full consensus scoring*', has been trialled in addition to the previously discussed majority vote and the separate mark scoring strategies, in the attempt to achieve improved resulting accuracy of prediction. Here, the three fingermark sex predictions made by a trained classifier for a single individual are only considered when all three marks are predicted to be the same sex; if a subset of the 3 marks are predicted to have conflicting sexes, the output is instead discarded. *This leads to an observably greater confidence in the model output in the remaining cases of up to a maximum of 86.1%* (for XGBOOST, Figure 8a, Table S2). However, it inevitably also leads to no valid predictions for over 50% of the individuals on average (computed across peak picking strategies, feature selection strategies) for each k-fold cross validation assessment (Figure 8b). Other model types also exhibited trends similar to Figure 8a, however with lower maximum full consensus scores (Table S2).

Although lack of classification for a large subset of samples is clearly undesirable, generally, in an operational context, it would be better to have a method through which "no classification" is achieved for many individuals but a high accuracy of sex prediction can be obtained for a small subset of individuals, than a method by which all individuals can be classified but with a lower accuracy of prediction such that it is not operationally useful. Therefore, the use of full consensus scoring system is highly recommended.

However, in the case of one recovered crime scene mark only, there will be an insufficient number of marks to use the majority vote or the full consensus scoring system in order to determine a sex prediction. Furthermore, the reported performance statistics for the cases of the majority vote and full consensus scoring schemes (or even for the separate marks scoring system which treats the 3 marks per individual as 3 entirely independent test set instances) cannot be assumed to be appropriate indicators of the expected reliability of the presented predictive models when predicting sex based on only one single mark.

In order to better ascertain model performances in the practical setting where only one fingermark may be available, a further session of k-fold CV was performed for each model type, but now with only a random 1/3 marks included per individual in each test subset per k-fold. Model training remained consistent to previously described, with all 3/3 marks per individual from each CV training subset being used.

As illustrated in Figure S8, there is a strong positive correlation between the mean CV scores, when only one mark is present per individual in each CV test subset, and the three previously discussed scoring schemes that have all utilised 3 fingermarks per test set individual. As such, this clearly demonstrates the applicability of the current predictive models for cases where only one fingermark may be available. Furthermore, when using only 1 mark per test set individual, the maximum model performance was reported as 70.9% attained by the XGBOOST classifier under k=25 cross validation (with a peak picking strategy of S/N=3 and min peak frequency=50%, under the random forest feature selection strategy) which is exactly on par with the top majority voting scheme scores using 3/3 marks. This circumstance very much indicates that the models are capable of dealing with just 1 mark, but it also may suggest that the benefit of submitting 3 marks only truly comes apparent when the full consensus scheme is employed.

Indeed, at crime scenes, investigators recover a variable number of marks ranging from one to a few and it is not unusual that a few are identified to an individual thus enabling the use of the full consensus scoring system and a predictive power of 86.1%.

Although even a predictive power of 86.1% may not render the method operational such that a suspect or a gender can be excluded from the investigations, the Police Force co-authoring this study suggests an alternative use of the method; in the case of multiple marks to screen during police investigations of high profile crimes, this method could inform the forensic strategy and act as some sort of triage to select the mark(s) to prioritise for examination.

4.0 Conclusions

The accuracy of prediction of sex of an individual from fingermarks using XGBOOST and the full consensus scoring system (86.1%) is essentially the same as that provided by a previously published proof of concept, when more relaxed classification criterion was adopted (85%). However, this result is important because it confirms the predictive power of this system *in natural marks* and with and without the prior application of an enhancing powder. The lack of fundamental improvement in the predictive power has been ascribed to the ion suppressing presence of PEG

based polymers which are found routinely in personal care products and toiletries and that are transferred in natural fingermarks (versus absence in model ungroomed marks which are depleted from lipids and contaminants). However, because of the most representative conditions used, the study presented here provides significant information for the forensic community. The authors believe this to be the case because, with respect to a previous proof of concept, it now clarifies what type of operational use can be made of the method developed for sex determination from fingermarks; as the accuracy of sex prediction (86.1%) does not allow exclusion of suspects from an investigation of serious crimes, in the case of multiple marks to screen during police investigations, this method could inform the forensic strategy and act as triage to select the mark(s) to prioritise for examination.

Additionally, this study presents, to date, the most comprehensive statistical modelling amongst all the studies attempting to discriminate sex from fingermarks. Furthermore, the accuracy of prediction of any other statistical modelling reported prior to the present study (however high) cannot be assumed to be maintained for only one submitted crime scene mark. Crucially, dissimilar to previous studies, the implementation of statistical models and its accuracy of prediction, has also been investigated for the circumstance in which only one fingermark sample is retrieved from a potential crime scene.

The authors believe that the comprehensiveness of the statistical modelling explored provides the community with clear direction as to the strategy and the highest performing combination of model/classification criterion for future studies attempting to investigate biomarkers of sex, other than peptides and proteins.

Code availability

Python and R scripts written for all sex classification model training and validation are available upon request to the authors.

Acknowledgements

Data acquisition for Study 2 (172 natural marks) was carried out at the M4I Institute, Maastricht University, Maastricht, The Netherlands. This work was supported by COST Action CA16101 “MULTI-modal Imaging of FOREnsic SciEnce Evidence - tools for Forensic Science” (MULTIFORESEE) supported by COST (European

Cooperation in Science and Technology) via a Short-Term Scientific Mission (STSM). The overall work presented here has been supported by The Home Office Innovation Fund (funding Dr R, Bradshaw), BBSRC-GlaxoSmithKline PhD studentship (awarded to Dr E. Patel) and a VC PhD scholarship funded by DSTL and Sheffield Hallam University (awarded to Mr C. Heaton). Dr Benjamin Baluff is gratefully acknowledged for the training given to Mr Cameron Heaton. Miss Katie Kennedy is gratefully acknowledged for her assistance with the fingerprint collection.

Figure Legends

Figure 1. MALDI MS spectra of VMD enhanced natural marks from study 2. A: typical mass spectrum obtained when gold clusters did not interfere with peptide/protein ionisation. B: donors' spectra showing prevalence of gold clusters and scarcity of peptide/proteins signals. C: spectra showing complete peptide/protein ion suppression and only presence of gold clusters. Where appropriate, peptide/proteins signals are labelled with a star symbol.

Figure 2. ROC analysis and AUC calculation from the application of PLSA to peptide and protein MALDI mass spectral profiles.

Figure 3. Relationship between the average spectra for the male and female samples for a subset of tested peak picking strategies. Strategies adopted S/N ratios of (a-b) 2:1, (c-d) 3:1, and (e-f) 5:1, and cross-spectra minimum peak occurrence rates (miniFreq) of (a, c, e) 1% and (b, d, f) 10%. Each point represents a peak picked position and the *x* and *y* axis values indicate the average TIC normalised intensity at this peak picked *m/z* position for the female and male samples, respectively. The scatter points have been coloured by the calculated PLS-DA VIP score.

Figure 4. 5-fold cross validation accuracy scores for five model schemes: LDA, random forest, naïve Bayes, XGBOOST and PLS-DA for the (A) majority vote scoring and (B) separate mark scoring schemes. In each subplot, the four different feature selection schemes presented in Table S1 have been displayed separately ('vip' = VIP scoring, 'all' = all *m/z* peaks (i.e. non-feature selection step), 'fs' = random forest feature selection, and 'ufs' = univariate feature selection). In addition, the accuracy scores have been separated by distinct peak picking strategies, and

ordered by the number of *m/z* positions associated with each strategy (see colour scales).

Figure 5. Classification models employed for the non-enhanced dataset. (a) previous 2-output state classification scheme, in which a classification model is trained to label input spectra data as either ‘male’ or ‘female’. (b) Modified 4-output state classification scheme, in which the classifier now must label each input spectra by both sex and contamination state.

Figure 6. Analysis of unenhanced fingerprint dataset via a neural network architecture. (a) previous 2-output state classification scheme (using a fully-connected neural network architecture), with an output layer consisting of two nodes with softmax activation for male/female sex classification. (b) Multi-class output neural network architecture with an output layer consisting of two pairs of two nodes, with each pair having an independent softmax normalisation, such that each input sample is simultaneously classified by sex and contamination level separately.

Figure 7. Comparison between best donors' fingermark MALDI mass spectra on autoflex (top panel) and rapiflex (bottom panel) MALDI Mass Spectrometers

Figure 8. Application of "full consensus scoring" strategy. (a) Correlation between the previously calculated majority vote scores (*x*-axis) and full consensus scoring schemes (*y*-axis), for the example case of the XGBOOST classifier. Each scatter point corresponds to the average cross-validation accuracy score for a specific peak picking strategy (see colour legend), and *k*-fold (see legend), with each of the four feature selection strategies in Table S1 treated as separate scatter points. The diagonal line *y=x* is shown. (b) For each *k*-fold *k* value (*x*-axis), the total number of test set individuals included within the full consensus scoring scheme is indicated (green bars) and compared to the corresponding recorded full test sizes per individual *k*-fold (orange bars). Error bars illustrate 1 standard deviation over each *m/z* peak-picking strategy (Table 1) and the four distinct feature selection strategies (Table S1).

References

[1] K. G. Asano, C. K. Bayne, K. M. Horsman, and M. V. Buchanan, Chemical composition of fingerprints for gender determination, J. Forensic Sci. 46 (2002), 1-3

[2] B. Emerson, J. Gidden, Jr. J.O. Lay and B. Durham, Laser desorption/ionization triacylglycerols and other components in fingermark samples, J. Forensic Sci. 56 (2011) 381-389

[3]. C. Huynh, E. Brunelle, L. Halámková, J. Agudelo and J. Halámek, Forensic Identification of Gender from Fingerprints, Anal. Chem. 87 (2015) 11531-11536

[4]. I. Mekkaoui Alaou and J. Halamek, Fluorescence of 1,2-Indanedione With Amino Acids Present in the Fingerprint Residue: Application in Gender Determination, J. Forensic Sci. 64 (2019) 1495-1499

[5]. H. L. Bandey, S. M. Bleay, V. J. Bowman, R. P. Downham and V. G. Sears, Fingermark Visualisation Manual, Home Office, St Albans, (2014) 1-932

[6] J. Lee, and M. Joullié, Fine-tuning latent fingerprint detection on paper using 1,2-indanedione bi-functional reagents, Tetrahedron, 71 (2015) 7620-7629

[7]. E. Bardin, E. Claude, and Z. Takatz, E. Bardin, E. Claude, and Z. Takatz, Waters Application Note, 2018, 1-3 Waters Application Note, 2018, 1-3

[8]. L. S. Ferguson, F. Wulfert, R. Wolstenholme, J. M. Fonville, M. R. Clench, V. A. Carolan, and S. Francese, Direct Detection of Peptides and Small Proteins in Fingermarks and Determination of Sex by MALDI Mass Spectrometry Profiling, The Analyst, 137 (2012), 4686-4692

[9]. A. M. Knowles, Aspects of physicochemical methods for the detection of latent fingerprints, J. Phys. E: Sci. Instru. 11 (1978) 713–721

[10]. S. Oonk, T. Schuurmans, M. Pabst, L. C. P. M. de Smet and M. de Puit, Proteomics as a new tool to study fingermark ageing in forensics, Scientific Reports 8 (2018) 2045-2322

[11]. R. Wolstenholme, R. Bradshaw, M. R. Clench and S. Francese, Study of Latent Fingermarks by Matrix-Assisted Laser desorption/ionisation Mass Spectrometry Imaging of Endogenous Lipids, Rapid Commun. Mass Spectrom. 23 (2009) 3031-3039

[12]. S. Francese, R. Bradshaw, L. Ferguson, R. Wolstenholme, M. Clench and S. Bleay, Beyond the ridge pattern: multi-informative analysis of latent fingermarks by MALDImass spectrometry, The Analyst 138 (2013) 4215-4228

[13]. M. Strohalm, M. Hassman, B. Košata and M. Kodíček, mMass Data Miner: An Open Source Alternative for Mass, Spectrometric Data Analysis, Rapid Commun Mass Spec. 22 (2008) 905-908

[14]. L. A. McDonnell, R. M. A. Heeren, R. P. J. de Lange and I. Fletcher, Higher sensitivity secondary ion mass spectrometry of biological molecules for high resolution, chemically specific imaging, J. Am. Soc. Mass. Spectrom. 17 (2006) 1195-1202

[15]. R. Bradshaw, S. Bleay, R. Wolstenholme, M.R. Clench and S. Francese, Towards the Integration of Matrix Assisted Laser Desorption Ionisation Mass Spectrometry Imaging Into the Current Fingermark Examination Workflow, Forensic Sci. Int. 232 (2013) 111-24

[16]. K. Dreisewerd, M. Schürenberg, M. Karas, F. Hillenkamp, Influence of the laser intensity and spot size on the desorption of molecules and ions in matrix assisted laser desorption/ionization with a uniform beam profile, International Journal of Mass Spectrometry and Ion Processes 141 (2) (1995) 127–148