Analyst

Recent Advances of Vibrational Spectroscopy and Chemometrics for Forensic Biological Analysis

Journal:	*Analyst*
Manuscript ID	AN-CRV-09-2021-001637.R1
Article Type:	Critical Review
Date Submitted by the Author:	12-Nov-2021
Complete List of Authors:	Takamura, Ayari; RIKEN, Center for Sustainable Resource Science Ozawa, Takeaki; The University of Tokyo, Department of Chemistry

Critical Review

Recent Advances of Vibrational Spectroscopy and Chemometrics for Forensic Biological Analysis

Ayari Takamura[1, 2*], Takeaki Ozawa[1**]

[1] Department of Chemistry, Graduate School of Science, The University of Tokyo
7-3-1, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
[2] RIKEN Center for Sustainable Resource Science
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Correspondence should be addressed to:
Ayari Takamura (*)
E-mail: ayari.takamura@riken.jp
Tel.: +81-045-503-9490, Fax: +81-045-503-9489

Takeaki Ozawa (**)
E-mail: ozawa@chem.s.u-tokyo.ac.jp
Tel.: +81-3-5841-4351, Fax: +81-3-5802-2989

Abstract

Biological materials found at a crime scene are crucially important evidence for forensic investigation because they provide contextual information about a crime and can be linked to the donor-individuals through combination with DNA analysis. Applications of vibrational spectroscopy to forensic biological analysis have been emerging because of its advantageous characteristics such as the non-destructivity, rapid measurement, and quantitative evaluation, compared to most current methods based on histological observation or biochemical techniques. This review presents an overview of recent developments in vibrational spectroscopy for forensic biological analysis. We also emphasize chemometric techniques, which can elicit reliable and advanced analytical outputs from highly complex spectral data from forensic biological materials. The analytical subjects addressed herein include body fluids, hair, soft tissue, bones, and bioagents. Promising applications for various analytical purposes in forensic biology are presented. Simultaneously, future avenues of study requiring further investigation are discussed.

Introduction

Forensic science, a discipline of applied science contributing to criminal investigations and judicial systems, involves extremely broad analytical subjects encompassing illicit drugs, explosives, toxicology, biological tissues, questioned documents, and ballistic[1, 2]. Especially, biological materials (e.g., hair, soft tissues, body fluids, and bone) found at a crime scene can be crucially important clues for forensic investigations because analysis of biological evidence can indicate how a crime was committed. Moreover, such materials, when examined using DNA analysis, can narrow down or identify source individuals. Conventionally, forensic analyses of biological evidence have been implemented via a sequence of visual and microscopic observations, and via serological and biochemical techniques. Such techniques have been proven to be effective. They have contributed greatly to forensic exams to date.

However, some fundamental characteristics of conventional techniques have been critiqued in recent years. Destructive testing consumes the limited amounts of examined materials, which can deter practitioners from conducting some important examinations, directly proceeding instead to DNA analysis. Also, some current experiments are time-consuming, taking up to several days. The consequent delays to investigation can be problematic when quick results are demanded. Furthermore, the practitioners' subjective assessments can bias qualitative findings[3]. Therefore, alternative methods have been demanded consistently so that forensic biology can conduct more reliable and efficient examinations of forensic biological materials. To satisfy such emerging demands, vibrational spectroscopy (i.e., Raman spectroscopy and Fourier-transformed infrared (FT-IR) spectroscopy) has been applied extensively to support forensic biology. The quantitative output of these non-destructive

and rapid spectroscopic techniques can be used with statistical evaluation, producing an automatic and cost-effective means of efficient forensic examination.

This review is aimed at summarizing recent studies elucidating vibrational spectroscopy and biological materials with results that are relevant to forensic concerns. Muro et al. presented a comprehensive review about the development of vibrational spectroscopy for forensic purposes up to 2014[1]. Subsequently, the forensic applications of Raman spectroscopy were reviewed by Doty et al. in 2016[4] and by Khandasammy et al. in 2018[5]. Surface-enhanced Raman scattering (SERS) applications to various forensic fields were summarized by Muehlethaler et al. in 2016[6] and by Fikiet et al. in 2018[7]. Developments of body fluid analysis were discussed for FT-IR applications by Mistek et al. in 2018[8] and for both Raman spectroscopy and IR spectroscopy in a part of a review by Silva et al. in 2019[9]. Since the publication of these reports, various important studies in the field of forensic biology have been reported. The review presented herein includes more recently reported studies as well as discussion that is more specific to the field of forensic biology. Additionally, this review is intended to emphasize chemometric techniques that have been adopted for the developed methods. The applications of vibrational spectroscopy have been investigated most intensively for body fluid samples, among other biological materials. Consequently, a considerable part of the discussion in this review is devoted to body fluid analysis. However, other subjects including hair analysis, forensic anthropology, and bioagent analysis have also shown considerably important progress in their analytical approaches based on vibrational spectroscopy and chemometrics.

Vibrational spectroscopy for forensic biological materials

Both Raman spectroscopy and FT-IR spectroscopy are typical vibrational spectroscopic techniques used in modern science. These two spectroscopic techniques, based on different light-induced phenomena, involve different properties to be considered in measurements.

Raman spectroscopy observes inelastic scattering, or Raman scattering, of light after irradiation on a sample. Raman scattering reveals the transitions between two molecular vibrational states in the corresponding energy shift of photons. The vibrational modes involving change of the molecular polarizabilities are Raman active. Typically, Raman scatterings are weak phenomena. The signal intensity is determined by the cross-section (max. 10^{-28} cm^2) and the power of incident light, and by the relative abundance of chemical components[10, 11]. The cross section is positively dependent on the light frequency. Consequently, generally speaking, a shorter excitation light wavelength yields stronger Raman signals. However, short wavelength excitation can simultaneously cause a strong fluorescence background, particularly for biological samples. Therefore, Raman spectroscopic analysis often uses near-infrared (NIR) excitation lasers for biological samples, including forensic biology samples. Additionally, the resonance effect and enhancement on metal surfaces contribute considerably to the increase of Raman signal intensities. Furthermore, although Raman spectroscopy

is applicable to various gas, liquid, and solid samples, the upright and back-scattering arrangement of Raman optical systems is usually preferred for analysis of forensic biomaterials because they are mostly solid (or dried).

Another mode of vibrational spectroscopy, IR spectroscopy, is based on absorption of infrared light by the corresponding molecular vibrational modes. The vibrational modes changing the dipole moments are IR active, which are complementary to the Raman active modes. Because infrared irradiation is used rather than the visible or near-infrared lasers of Raman spectroscopy, thermal damage or photo-damage to samples fundamentally does not occur with IR spectroscopy. As described below, FT-IR spectroscopy for forensic relevant biological materials has been conducted mostly using the attenuated total reflection (ATR) method. Few studies have used other FT-IR measurement methods such as transmission and external reflection. Actually, ATR method has become increasingly popular. It is increasingly applicable to various samples, even those with highly absorptive properties. The ATR phenomenon at a boundary between an ATR prism and a contacted sample generates evanescent waves toward the interior of the sample. Then, IR absorption occurs only on the sample surface (within several micrometers), which corresponds to the depth at which the evanescent wave can penetrate. The penetration depth of an evanescent wave is dependent on the light wavelength, incident angle, and refractive index of both the ATR prism and the sample[12].

Chemometrics

Spectra of biological materials observed by both Raman and FT-IR spectroscopy are typically complex because of mixed signals from various chemical compositions. In addition, the relative abundances of the respective components can vary depending on heterogeneous spatial distributions, donor individualities, and experiment conditions. Therefore, the vibrational spectra of biological materials often require evaluation based on multivariate statistical methods. Here, chemometric techniques have been indispensable for obtaining the analytical outputs of interest. Chemometrics is a discipline that uses mathematical and statistical methods to elucidate and correlate external parameters from complex chemical data. Since the burgeoning of the discipline in the late 1960s and 1970s, spectroscopic data have been a central subject of chemometric investigation[13].

Exploration of spectral pre-processing techniques has been an interest in chemometrics along with development of analytical models based on spectral data [13, 14]. Purposes of data pre-processing are data quality adjustment, data format organization, and data feature enhancement. Pre-processing for vibrational spectral data includes baseline correction, removal of outliers or low-quality (e.g., low signal-to-noise ratio) data, normalization, smoothing, differentiation, and binning[15, 16]. Raman spectra also require processing of wavelength calibration, spectral axis alignment, cosmic ray/spike removal, and removal of the fluorescence background. Variable selection techniques based on genetic algorithms (GA) and interval partial least squares (iPLS) are also available. The choice and

implementation of the spectral pre-processing are of crucial importance for subsequent predictive modeling.

Chemometric modeling is aimed at predicting qualitative or quantitative properties of multivariate data, which correspond respectively to classification and calibration analysis. Various chemometric techniques have been developed according to analytical purposes and data properties. Principal component analysis (PCA), hierarchical clustering, and self-organization map are useful for pattern recognition. Moreover, PCA is used for reducing the dimensionality of multivariate datasets and for de-noising before detailed modeling. Multivariate deconvolution techniques such as multivariate curve resolution (MCR) and independent component analysis are other approaches to decompose spectral datasets into a linear combination of a small number of significant spectral components. Classification is applied to assign the examined data to one or more classes. Multivariate one-class classification techniques involve soft independent modeling of class analogy (SIMCA), Hotelling's T^2, D statistics, and Q statistics. These techniques are designed to evaluate similarities among elements of the same class. However, conventionally used two or multi-class classification techniques determine the boundaries separating data of different classes in multi-dimensional space. Examples of such techniques are Euclidean distance to centroids, linear discriminant analysis (LDA), and quadratic discriminant analysis (QDA). For a dataset with a higher number of variables than samples, partial least squares discriminant analysis (PLS-DA) is an effective method because it involves the process of dimension reduction. Multivariate calibration is aimed at predicting quantitative properties (e.g., concentration of a constituent) by regression of spectral data. One relevant method is multivariate linear regression. For high-dimensional data, principal component regression (PCR) and partial least squares regression (PLSR) are used. The techniques described earlier are designated as linear modeling methods: they assume a linear relation between the response variables (categorical values for classification, and continuous values for calibration) and explanatory variables (i.e., spectral intensity at each wavenumber). For finer modeling of complex analytical data, nonlinear modeling techniques such as support vector machine (SVM), random forest (RF), and artificial neural networks (ANN) (e.g., recurrent neural network (RNN), convolutional neural network (CNN)) can be alternative means of constructing classification or calibration models.

Body Fluids

Body fluid identification

Body fluids (e.g., blood, saliva, and semen) are commonly collected in various contexts of a crime scene, particularly involving violence such as sexual assault and homicide. Analysis of body fluid evidence contributes to provision of contextual information of the crime. Moreover, identifying a source of DNA is critically important to demonstrate relevance between identified individuals and the evidence. The primary purpose of body fluid analysis in forensic investigations is to determine

body fluid types. Conventionally, forensic examinations of body fluid evidence have been conducted using biochemical and serological techniques designed to detect composition characteristics of various body fluid types. Today's forensic exams comprise presumptive tests and subsequent confirmatory tests. For example, blood evidence is presumptively discovered via detection of hemoglobin's oxidation activity based on luminol chemiluminescence or chemical colorimetric testing[17, 18]. Unfortunately, such easy and rapid presumptive tests can yield false-positive results. Subsequent confirmatory tests are more selective, providing more reliable results. Blood evidence is then conclusively identified by detection of human hemoglobin using the specific antigen, which is usually conducted with commercial immunochromatography kits. However, most current techniques for body fluids are destructive, which is disadvantageous because the amounts of examined evidence are usually limited. They must be preserved to the greatest extent possible for subsequent DNA analysis and future re-investigations.

Vibrational spectroscopic approaches are quite preferable for forensic body fluid analysis. Vibrational spectroscopy allows rapid, versatile and nondestructive examinations. The last decade has been a revolutionary era of forensic body fluid analysis using vibrational spectroscopy, for which numerous pertinent studies have been reported (Table 1). Investigations of Raman spectroscopic analysis for forensic body fluid samples were launched in 2008 by Virkler and Lednev[19]. They first reported the potential for discriminating typical body fluid types (i.e., human semen, canine semen, vaginal fluid, saliva, sweat, and blood) based on the characteristic Raman peaks and their corresponding components. Thereafter, the Lednev research group at the University of Albany has made great contributions to the development of this field to date. From early investigations, Raman spectral signatures were developed for various body fluid types (i.e., blood[20], saliva[21], semen[22], sweat[23], and vaginal fluid[24]) to identify them and simultaneously describe the heterogeneities of the traces and donor-dependent variations. Herein, multivariate deconvolution techniques of significant factor analysis (SFA), principal component analysis (PCA), and multivariate curve resolution (MCR) using alternating least squares (ALS) algorithm in chemometrics were used to find the Raman profiles of significant components. Raman spectroscopic analysis has been subsequently developed to discriminate body fluid samples based on advanced statistical or chemometric techniques. Discrimination of body fluid types has been studied, beginning with discrimination of three body fluid types (blood, saliva, and semen)[25] and peripheral/menstrual blood[26]. In 2016, Muro et al. demonstrated discrimination of five common body fluids: peripheral blood, saliva, semen, sweat, and vaginal fluid[27]. They developed discriminant models using PLS-DA and support vector machine discriminant analysis (SVM-DA) algorithms. In addition, more informative spectral regions were selected using interval PLS-DA and GA, thereby achieving nearly perfect discrimination. More recently, Vyas et al. have expanded the Muro's model for five body fluids including urine. They demonstrated 100% accuracy for the identification of all body fluid types[28].

Surface-enhanced Raman spectroscopy (SERS) can be a powerful tool for forensic body fluid analysis, especially because of its high sensitivity. Premasiri et al. reported SERS spectra of whole blood, red blood cells, and blood plasma using a 785 nm excitation laser and Au particles covering a SiO_2 substrate[29]. They demonstrated that SERS spectra of whole blood are dominated by signals from the blood plasma component, whereas normal Raman spectra of whole blood are almost entirely derived from oxyhemoglobin. Additionally, they demonstrated time-dependent changes of SERS spectra of whole blood, which were attributed to an increase of hypoxanthine leaked from cellular components after approx. 15 h of storage. Bonifacio et al. reported a systematic comparison of SERS spectra of blood plasma and serum using various Ag and Au aqueous colloids with three laser wavelengths (i.e., 514, 633, and 785 nm)[30]. After the authors assessed various sample preparation procedures, they concluded that only the combination of filtering proteins and the use of Ag nanoparticles and 785 nm excitation laser provided repeatable spectra. Recently, Shaine et al. reported confirmative detection of dried bloodstains using SERS[31]. They reported detailed assignments of SERS spectra of dried blood with Au and Ag nanoparticles on SiO_2 substrates based on DFT calculations. They showed the SERS detection sensitivity for blood as similar or higher than that of common immunochromatographic kits such as the RSID test, hemoglobin HemDirect test, and HemaTrace test. Moreover, using the PLS-DA algorithm, they demonstrated the discrimination of SERS spectra of blood and four other body fluids that are commonly found at crime scenes.

In 2011, Fourier transform infrared (FT-IR) spectra of body fluids were first reported in the context of forensic analysis by Elkin[32]. Elkin used attenuated total reflection (ATR) equipment to observe and compare infrared spectra of various human body fluids and materials. Results of that study indicated that amide I peaks and fingerprint regions were characteristic for the respective examined samples. That combination of distinctive peaks potentially enables differentiation of various samples. Orphanou et al. investigated the ATR FT-IR spectra of human blood, saliva, semen, and vaginal secretions, which are frequently obtained at scenes of suspected violent or sexual offenses[33]. The methods described in those studies determined detectable components and vibrational modes that were characteristic to each body fluid type, particularly in the spectral regions of lipids (3000–2800 cm^{-1}), proteins (1700–1600 cm^{-1}), and nucleic acids (1250–1000 cm^{-1}). Based on the spectral patterns, combinations of peaks, and peak frequencies, the observed spectra of these body fluids were inferred as distinguishable. Zapata et al. reported discrimination of stains of semen, vaginal fluid, and urine using external reflection FT-IR spectroscopy and chemometric analysis[34]. The spectra of the respective body fluids were classified correctly from other body fluids and potential false positive substances by PCA and SIMCA. Takamura et al. reported discrimination between antemortem blood and postmortem blood using ATR FT-IR spectra and PLS-DA modeling[35]. Discrimination of postmortem blood is crucially important for forensic investigations to reveal a crime sequence. Using GA, contributive spectral regions are indicated, which represent a signal increase of lactic acids in the

spectra of postmortem blood. Takamura et al. also demonstrated a discriminant model for ATR FT-IR spectra of various body fluids (e.g., peripheral blood, saliva, semen, urine, and sweat) combining PLS-DA, LDA, and Q-statistics[36]. They proposed a model architecture of a dichotomous classification tree based on hierarchical clustering analysis results. This model architecture enabled robust discrimination even for disturbed spectra of body fluid samples aged over several months. In addition, insertion of Q-statistics models functioned as outlier analyses for non-body fluid samples.

Species identification and phenotype profiling of body fluid traces

Forensic body fluid analysis using vibrational spectroscopy combined with chemometrics has been developed further for more advanced purposes: species identification and phenotyping of human donors (Table 1). Determination of origins of body fluid traces and narrowing down of donor candidates based on phenotypes are beneficial for forensic investigation. Components of a certain body fluid type are remarkably similar among various species and human individuals. Therefore, chemometric techniques have been invaluable for investigating similar spectral patterns. Species identification has been explored mainly for bloodstains. The Lednev research group has investigated discriminant analysis between human and animal (non-human) blood using Raman spectroscopy by stepwisely increasing the number of examined animal species[37, 38]. A recent report of their work by Doty et al. demonstrated the use of PLS-DA modeling for discrimination of the Raman spectra of blood from humans and 16 animal species[39]. The examined animal species included cat, chicken, dog, horse, mouse, opossum, pig, rabbit, raccoon, rat, chimpanzee, deer, elk, ferret, fish, and macaque, which were presumed to be forensically relevant. Receiver operating characteristic (ROC) analysis of the constructed PLS-DA model demonstrated 99% discrimination accuracy between human and non-human subjects for an external dataset. Most recently, Wang et al. reported a recurrent neural network model for discriminating the Raman spectra of blood from 20 kinds of species including human[40]. These models were designed to classify a Raman spectrum into one of the examined species. The total classification accuracy of the constructed RNN model reached 97.7%. Additionally, they assessed resistance to wavenumber drift of −5 to 5 cm^{-1} and cross-instrumental modeling based on Raman spectra measured using two different Raman spectrometers. The RNN models trained by considering these variations also showed comparable classification performance.

Species identification has also been achieved using FT-IR spectroscopy. Mistek and Lednev developed a discriminant model for ATR FT-IR spectra between human, cat, and dog bloodstains[41]. They recently expanded their work using blood samples from 11 animal species[42]. The two-class discriminant model based on PLS-DA showed superior discrimination accuracy between humans and animals. They also emphasized that none of the blood spectra from external animal species outside of model training was assigned as human, satisfying the forensic requisites. Wang and coworkers have explored species identification based on ATR FT-IR spectra, considering forensic practical challenges,

in their subsequent two studies. Lin et al. reported PLS-DA models to differentiate human bloodstains from the bloodstains of five animal species. They assessed the models' performance for bloodstains subjected to indoor and outdoor conditions in addition to an aging process[43]. Wei et al. recently demonstrated the use of ATR FT-IR spectroscopy for species identification of semen stains based on the supernatants[44]. They compared effects of substrate types and time since deposition on the observed spectra. The developed PLS-DA model successfully discriminated the supernatants of human semen from those of nonhuman ones. In addition, species identification of semen supernatant without sperm cells indicated the applicability to azoospermia samples.

Phenotype profiling based on body fluid stains using vibrational spectroscopy has been developed for identifying the sex, race, and age of donors. Here also, the Lednev research group pioneered these analytical interests. Both Raman spectroscopy and FT-IR spectroscopy have exhibited promising potential. Phenotype profiling of bloodstains has been explored intensively for determining sex[45, 46], race[46, 47], and age[48, 49]. Sikirzhytskaya et al. reported sex determination of bloodstain using Raman spectroscopy[45]. They developed a classification model based on an artificial neuron network (ANN) coupled with GA for wavenumber selection. The obtained selectivity and sensitivity for sex discrimination were 95% for a training dataset and >80% for a test dataset, which were superior to the compared SVM-DA models. Mistek et al. demonstrated discrimination of both sex and race (i.e., Caucasian, African American, and Hispanic) based on ATR FT-IR spectra of bloodstains[46]. They used a PLS-DA algorithm. The main results of the external model validation were 92% prediction accuracy for both sex and race based on individual spectra, and 100% accuracy at donor-level classification. Mistek et al. also reported race differentiation of bloodstains by Raman spectroscopy coupled with SVM-DA[47]. The latest progress for phenotype profiling of bloodstains is to differentiate groups by the chronological age (CA) of donors. Analytical methods used for determination of CA are quite useful for forensics because such information cannot be obtained via DNA profiling. Doty et al. constructed a SVM-DA model based on Raman spectra of bloodstains to differentiate a donor's CA between newborns (<1 year), adolescents (11–13 years), and adults (43–68 years)[48]. Cross-validation of the model showed high sensitivity and specificity of more than 0.96 and 0.97, respectively, for the donors' age groups. After using ATR FT-IR and a PLS-DA algorithm, Giuliano et al. recently reported similar findings for bloodstains from the same CA groups[49]. The discrimination accuracy, as evaluated by subject-wise leave-one-out cross-validation, was 92% for the individual spectra and 95% for the donors.

Phenotype profiling has been explored for body fluids of other types as well as bloodstains. Sex determination of saliva traces was investigated by Muro et al. using Raman spectroscopy and SVM-DA algorithms[50]. Muro and Lednev also reported race differentiation of semen traces using Raman spectroscopy[51]. For that study, SVM-DA combined with GA was adopted for the wavenumber selection. All 18 donors for internal CV and 7 donors for external tests were classified correctly

according to their corresponding races. Takamura and coworkers demonstrated sex determination of urine traces using ATR FT-IR[52]. Because DNA is often found only in small amounts, this technique is especially useful because genome-based analysis for urine traces is often unavailable. After the authors selected informative spectral regions using GA, they developed a PLS-DA and ANN discriminant model. The evaluated accuracy was 0.97 for donor-wise discrimination in the use of the PLS-DA algorithm. In addition, the constructed ANN model showed comparable discrimination performance to that of the PLS-DA model.

Prediction of time since deposition of body fluid traces

Development of analytical techniques to estimate time since deposition (TSD) of body fluid traces has long persisted as a great concern in forensic science[53, 54]. Such techniques are beneficial to indicate the relevance of an evidence to the crime, and to estimate when and how a crime was committed. Especially, the techniques for bloodstains have the most intensively explored. This is because bloodstains are frequently found in crime scene, and because hemoglobin (Hb), which is a representative component in blood, shows distinctive chemical changes by autoxidation during the aging. Specific autoxidation of hemoglobin, which involves oxyhemoglobin (oxyHb), methemoglobin (metHb), and hemichrome (HC), has been expected to be an indicator of the bloodstain aging[53, 54]. These hemoglobin changes are visible as changes of the visible color from red to brown during bloodstain aging, whereas the rate of the color change is affected by environmental conditions[55-57]. Other components in blood, such as RNAs in white blood cells and blood plasma, have also been analyzed during aging. They might change at different rates to that of hemoglobin autoxidation. However, currently proposed techniques targeting these compositions are fundamentally destructive and are rather costly[54]. Despite enormous research efforts undertaken over the last century, no analytical method has been established to predict the TSD of bloodstains in forensic practice. Nevertheless, techniques using vibrational spectroscopy have emerged, showing promising potential and the exclusive benefit of non-destructivity (Table 1).

Raman spectra of blood excited by visible-wavelength or NIR-wavelength lasers, are dominated by signals from hemoglobin and the oxidized variants because of resonance and semi-resonance effects[58, 59]. The characteristic spectral patterns and the detailed peak assignment of hemoglobin variants have been explored intensively. For example, bands observed at 1638 cm^{-1} ($\nu(C_\alpha C_m)_{asym}$), 1225 cm^{-1} (ν_{13} or ν_{42} of $\delta(C_mH)$), 570 cm^{-1} ($\nu(Fe–O_2)$), and 419 cm^{-1} ($\delta(Fe–O–O)$) are characteristic of oxyHb, which has Fe^{2+}–O_2 binding at Heme groups[60-63]. MetHb is known to show specific bands at 376 cm^{-1} ($\delta(C_\beta C_c C_d)$), 1629 cm^{-1} ($\nu(C_\alpha C_m)_{asym}$), 1372 cm^{-1} (ν_4), and 1212 cm^{-1} ($\nu_5+\nu_{18}$ or ν_{13})[29, 60, 61, 64]. Moreover, earlier work has demonstrated that short wavelength excitation lasers, high laser power, and prolonged irradiation can cause photo-denaturation and thermal denaturation of hemoglobin as well as increasing of fluorescence background and signal-to-noise ratio[61, 65]. These denaturation

processes brings increases of some peaks at 1396, 1365, 1248, 972, and 662 cm^{-1}[65], potentially accompanying an increase of HC concentration[66]. Such laser-induced denaturation produces similar spectra of bloodstains, which can hinder reliable TSD analysis[61]. Therefore, for forensic purposes of predicting TSD, the use of an NIR excitation laser (i.e., 785 nm) and low laser power have been regarded as suitable to avoid overestimation of TSD and to minimize prediction errors[54, 61]. Predictive methods for TSD of bloodstains using Raman spectroscopy were first reported by the Lednev research group. Doty et al. demonstrated chemometric models for prediction of TSD of bloodstains up to one week[67], and further extended them up to two years[68]. Their second report described application of two linear regression methods, PLSR and PCR algorithms, to the observed Raman spectra of bloodstains excited by a 785 nm laser. The overall accuracies of approximately 70% for predicting the TSD at each time point were obtained for the respective regression methods. At the same time, they used a discriminant model that their group established earlier to assess whether an aged bloodstain is identifiable as "blood"[27]. The discrimination accuracy was 89%. It is noteworthy that the rate of 100% was obtained for bloodstains that had been aged for up to one month. Takamura et al. recently reported a multivariate spectral deconvolution model to describe spectral changes during bloodstain aging using Raman spectroscopy with 785 nm excitation[66] (Fig. 1). They specifically examined the kinetics latent in the spectral changes derived from autoxidation of Hb variants and denaturation of other components. The identified kinetic formulas were used as constraints to deconvolute the Raman spectra into five significant spectral components, corresponding to the respective blood components. Consequently, based on the decomposed spectral components, they proposed an index for the relative degree of bloodstain age, which is available independently of the environmental conditions.

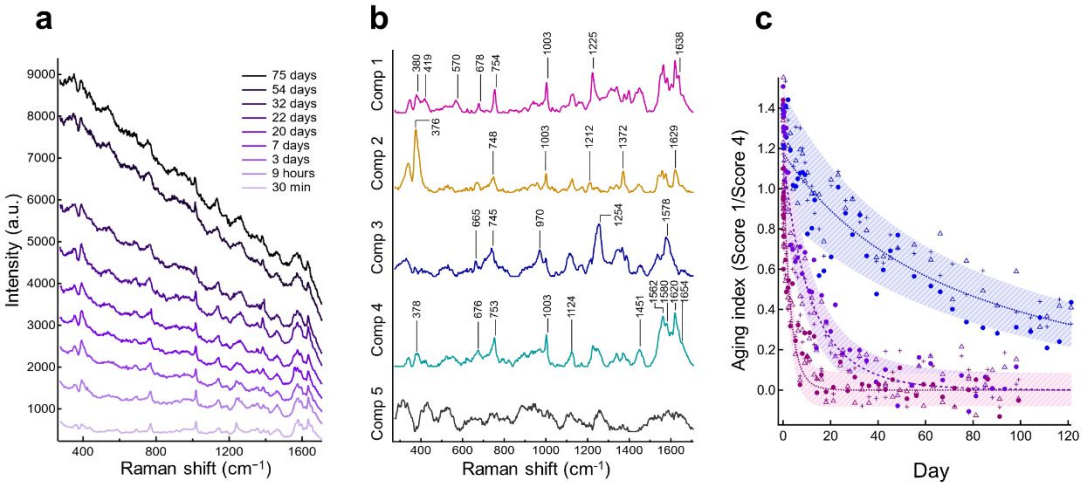

Fig. 1 Deconvolution of Raman spectra of bloodstains during aging. (a) Spectral series observed during bloodstain aging at 24 °C. The spectra are shown before subtraction of fluorescence background. (b) Decomposed Raman spectral profiles of bloodstains by multivariate curve-resolution alternating least squares. (c) Index for bloodstain aging at 30 °C (magenta), 24 °C (purple), and 16 °C (blue). The index was defined as the ratio scores between the first and the fourth spectral components yielded via multivariate deconvolution technique based on the latent kinetics. The circles, crosses, and triangles represent data from three donors. The modeled ratios are shown as dashed lines with the standard deviations (shaded areas) for each temperature.

In contrast to Raman spectroscopy, ATR FT-IR spectroscopy presents some advantages for TSD analysis of bloodstains. The ATR FT-IR spectra are not adversely affected by fluorescence. Also, measurement via infrared incident does not cause photo-degradation or thermal degradation of blood compositions. Lin et al. used ATR FT-IR spectroscopy to investigate bloodstains that had been exposed to indoor and outdoor conditions[69]. They constructed PLSR models to predict the age of bloodstains for durations of 0.25–7 days, 7–85 days, and 0.25–107 days based on indoor and outdoor spectral data. The PLSR models showed reliable estimation for the samples under the same environmental conditions, but poor performance under different environmental conditions. Alternative models used to distinguish fresh (age<1 day) and older (age> 1 day) bloodstains were established via PLS-DA. The model trained using outdoor data represented good discrimination performance even for indoor data. Kumar et al. recently demonstrated estimation of bloodstain age based on the ATR FT-IR spectroscopy using different models including curve estimation, multiple linear regression (MLR), and PLSR[70]. The ATR FT-IR spectra were collected from bloodstains that had been aged for 1–175 days. After 25 distinctive peaks on the first derivative spectra were selected as independent variables, they were used one by one for model building. The MLR model composed of the three variables was inferred as the best, with prediction error of approx. 3 ± 1 days. The PLSR model showed comparable performance, providing prediction error of approx. 4 ± 1 days.

More recently, prediction methods for TSD of body fluids other than bloodstains have been

explored, although they are still in an early stage of development. Zha et al. reported a preliminary study to predict TSD of semen stains using ATR FT-IR[71]. Semen stains were prepared on glass, tissue, and fabric made of regenerated cellulose. The semen stains were stored for up to six days, then the semen supernatants via extraction were assessed. In this work, PCA results indicated interference from the substrates as negligible compared to spectral variations derived from sample aging. The PLSR models were constructed using the absorbance and second derivative of the "bio-fingerprint" region of 1800–900 cm^{-1}, respectively. The model based on the second derivative spectra demonstrated better prediction accuracy, with evaluated coefficient of determination (R^2) values of 0.81 for cross-validation and 0.74 for external validation. Work particularly addressing the drying processes of three body fluids (semen, saliva, and urine) was undertaken by Das et al.[72]. Changes of ATR FT-IR spectra of these body fluid traces were observed during the drying of phase I (water dominant) and phase II (rapid water evaporation) up to 42 min. Phase II showed more drastic changes of the spectra during drying. The PCR and PLSR models were constructed using several age-linked peaks, indicating the potential of this analytical approach for estimation of the TSD, particularly during the initial drying process, for each body fluid.

Practical challenges: substrate interference and body fluid mixtures

Challenging issues presented by the practical use of vibrational spectroscopy for forensic body fluid analyses are mixtures of different body fluid sources and interference of the substrate on which the body fluid has been deposited. Body fluid evidence is often discovered as mixed with body fluids of other types. Such mixed body fluids can exhibit complicated spectral patterns: the spectral characteristics of the respective body fluids are overlapped, simultaneously involving donor-dependent variations. The substrate interference also causes mixed signals of body fluids and the substrate materials. How much the interference affects the body fluid spectra is dependent on substrate properties such as absorptiveness, thickness, porosity, and surface roughness[73, 74]. Moreover, heterogeneous substrates (e.g., cloth of denim and blended fabrics) increase the difficulty of spectral treatment[75, 76]. These tendencies of body fluid mixtures and substrate interference have been studied using both Raman spectroscopy and FT-IR spectroscopy with assessment of various body fluid types and substrate materials (Table 2).

Elkin reported identification of various body fluid types based on characteristic bands in ATR FT-IR spectra, even when deposited on interfering substrates such as a white T-shirt and white copier paper[32]. Later, Quinn and Elkins showed ATR FT-IR spectra of venous blood, menstrual blood, semen, saliva, and breastmilk in a neat state or after deposition on various substrates including cotton, nylon, wood, paper, and glass[73]. They demonstrated that the characteristic signals of body fluids were still detectable on the substrates. However, their intensity was decreased considerably, and decreased especially when pipetted onto a porous substrate (e.g., cotton, nylon, paper, or wood) into which body fluids were absorbed. They also showed that the substrate roughness (i.e., weave of fabrics) can lessen

body fluid signals. Recently, Sharm et al. reported the use of ATR FT-IR spectroscopy for discrimination of menstrual blood and peripheral blood[74], seminal fluid[77], and vaginal fluid stains[78]. In that work, they comprehensively studied practically challenging factors such as substrate interference, mixture with other body fluids, differentiation of look-alike non-biological materials, dilutions, washing, and chemical treatments. Regarding substrate interference, substrates of various types were assessed including white cotton, denim, polyester, paper, wood, skin, plastic, grass, glass, condoms, sanitary napkins, and floor tiles. The non-porous substrates provided spectral profiles of body fluids that were almost consistent with those in the neat states. However, the porous substrate alleviated or altered body fluid signals, whereas only some distinctive peaks of proteins such as Amide I and Amide II were still recognized. Body fluid mixtures of seminal fluid or peripheral/menstrual blood with vaginal fluid showed overwhelming signals of seminal fluid or peripheral/menstrual blood, which hindered observation of vaginal fluid signals[77, 78]. Hager demonstrated identification of urine in a liquid state and on fabrics such as white cotton, blue jeans, a white lab coat, and a blue uniform (polyester) as well as sweat-contaminated cotton shirts[79]. For this study, they used a hand-held Raman spectrometer to facilitate on-site analysis. Actually, PLS-DA algorithms for the observed Raman spectra showed high discrimination accuracies between the presence and the absence of urine on all studied fabric types.

Chemometric strategies have also been explored to examine vibrational spectra of body fluids with interference by a substrate and after mixture with other body fluids. Sikirzhytski et al. combined Raman spectroscopy and SVM-DA to demonstrate classification of pure blood, pure semen, and a blood–semen mixture[80]. The constructed model enabled classification with high accuracy, whereas only mixture samples with a small portion of blood (5%) were misclassified as semen. The Lednev research group also investigated treatments of the Raman spectra of body fluid stains (blood[75, 81] and semen[76]) with interference by substrate materials. They manually subtracted substrate contribution from the interfered body fluid spectra. Subsequently, they performed multivariate fitting of the recovered body fluid signals using the corresponding Raman signatures derived from pure body fluid spectra. Goodness of fit scores (i.e., sum of squares due to error (SSE), R^2 and root mean squared error of prediction (RMSE)) showed reliable results for identifying bloodstains, irrespective of substrate and contaminant types[75, 81]. Scores for semen identification were lower than expected, indicating the possibility of semen component separation because of capillary effects of fabrics[76]. Takamura et al. demonstrated discrimination of antemortem and postmortem blood deposited on clothes of cotton, denim, and polyester[35]. The ATR FT-IR spectra of the bloodstains were subjected to multivariate fitting using the PCA loadings of both pure blood and substrates. Additionally, a weight factor developed based on the PLS-DA loadings were also incorporated into the fitting calculation. The extracted spectral profiles of blood signals allowed the discrimination of blood origins with high accuracy (max. 95%). More recently, McLaughlin et al. reported automatic extraction of body fluid

signals from Raman spectra in the presence of substrate interference[82]. For this study, hypothetical addition multivariate analysis with numerical differentiation (HAMAND) was used. It is a spectral processing technique used to separate and quantify the contributions of a known spectral component in a mixture or overlapping spectrum[83, 84]. The spectral components of semen were identified successfully in the Raman spectra of semen stains on glass and polyester, even though the net contribution was estimated as only less than 0.01%.

During the last few years, the applications of vibrational spectroscopy and chemometrics to forensic body fluid analysis have been explored intensively: the identification techniques have covered most of common body fluid types; the phenotype profiling methods have developed to determine sex, age, and age-group. The techniques for TSD prediction have begun to target various body fluid types as they have for bloodstains. At the same time, challenging aspects of body fluid analysis based on vibrational spectra have been revealed. Issues of substrate interference and body fluid mixture remain. Other factors such as low concentration and degradation of body fluid evidence must also be settled. Through continuous efforts undertaken to date, the promising potential of vibrational spectroscopy for forensic body fluid analysis has been widely explored and proven. Methods to be developed will move from proof-of-concept studies to in-depth validation studies to address the full scope of pragmatic difficulties encountered when examining forensic cases.

Hair

Hair is ubiquitous in ordinary environments. It is commonly discovered at crime scenes. Hair evidence discovered at crime scenes can provide important information about source individuals. Although it is beyond the scope of this review, hair evidence has also been used to detect drug use and to obtain a DNA profile from a hair root, if present, or mitochondrial DNA from the hair shaft. Among hairs found on various regions of human body, mostly scalp and pubic hair can be important evidence for forensic investigation of crimes such as intrusion into homes by a sneak thief, violence, and sexual assault. Forensic analysis of hair basically aims at determination of whether the sample is natural hair rather than a synthetic fiber, whether it is from a human or an animal, and if human, whether the examined unknown hair is matched to hair collected from a known source, usually from the suspect. Microscopic observation is used currently to evaluate attributes of hair, including the length, diameter, cross-section shape, medullary index, chemical treatment, and somatic region. The microscopic analysis of hair by experienced examiners has proven its worth and has contributed to forensic investigation to date. However, its scientific validity has been questioned in recent years[3, 85]. First, subjectivity is an inherent flaw of microscopic observation. Even for well-trained examiners, their conclusion can be biased on information received before or during the exam. In addition, absence of clear criteria can result in inconsistent conclusions, depending on the examiner. Therefore, other

analytical techniques to characterize hair are worth exploring to improve the reliability of forensic hair analysis. Here, vibrational spectroscopy can offer alternative methods of forensic hair analysis in a nondestructive manner.

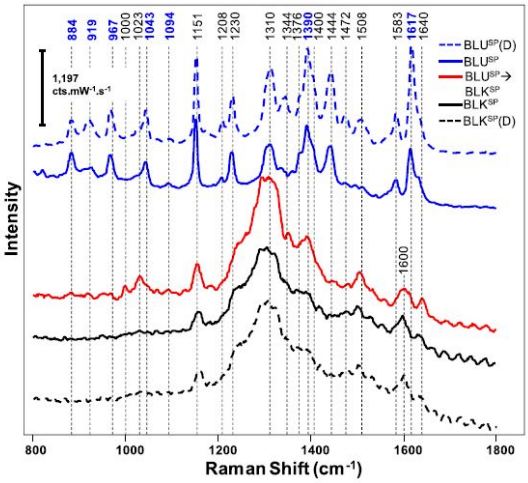

Fig. 2 SERS spectra of hair colored by blue semipermanent dye (BLU^{sp}) and re-dyed after by black semipermanent dye ($BLU^{sp} \rightarrow BLK^{sp}$). The spectra of hair colored by BLKsp and the dyes themselves (BLK^{sp} (D) and BLU^{sp} (D)) are also included. Reproduced from ref. 90 with permission from American Chemical Society, copyright 2019.

Hair analysis using Raman spectroscopy has not been common, primarily because high fluorescence from melanin granules, hair pigment, interferes with the spectra. However, some studies have demonstrated the great potential of Raman spectroscopy to evaluate detailed chemical states of hair components. Kuzuhara and coworkers have investigated Raman spectroscopic features of human hair exposed to permanent wave treatments including steps of reduction, stress relaxation, and oxidation[86]. Raman spectra were collected from cross-sections of white human hair untreated and treated via each step. That study revealed that disulfide (-SS) bonds of gauche-gauche-gauche (GGG) and gauche-gauche-trans (GGT) conformations decreased by reduction processing with thioglycolic acid, whereas trans-gauche-trans (TGT) conformations showed no marked change. Subsequent stress relaxation processing further promoted cleavage of the GGT and GGT conformations. The final oxidation treatment caused an increase of the S–O band assigned to cysteic acid and a greater decrease of GGG contents, indicating molecular disorganization in the cuticle and cortex cell. In a recent study, Kuzuhara reported damage effects on hair by reduction treatment with thioglycerol[87]. The cross-sections of white human hair reduced with thioglycerol shows less decreases of GGG and GGT disulfide contents, or less damage, than those reduced with thioglycolic acid. Nevertheless, thioglycerol provided higher waving efficiency than thioglycolic acid. dos Santos et al. compared Caucasian and Afro hair, applying different treatments of heating, bleaching, and straightening[88]. Raman spectra of intact Caucasian and Afro hairs were similar, but the Afro hair showed slightly higher intensity of the Amide III band. I Also, changes at some characteristic bands, such as the S–S, C–C, S–O, and Amide III, after each and all three treatments were observed differently between the Caucasian and Afro hair.

SERS techniques to detect colorants in dyed hair have been investigated for forensic purposes. Kurouski et al. used gold nanorods and 785 nm excitation laser to observe SERS spectra of artificially

dyed hairs[89]. They demonstrated that the distinctive bands in the SERS spectra of dyed hair exactly matched those of chemical colorants contained in the commercial dyes. However, normal Raman spectra showed no colorant signal. Importantly, the SERS spectra were observed without interference from natural hair pigments because the pigments exist in the inner structure of hair (i.e., the cortex), thereby separated from SERS effects. Based on the characteristic SERS bands, hair that had been dyed using different commercial brands were distinguishable even when they were the same color. They also showed the usability of a portable Raman spectrometer to measure SERS spectra with a high signal-to-noise ratio, offering potential for field examinations of hair at a crime scene. Recently, Esparza et al. expanded the work described above. SERS were used to examine the underlying colorants in re-dyed hair[90]. Results indicated that SERS enabled detection of blue semipermanent colorant on hair that had been re-dyed with both black semipermanent and permanent dye (Fig. 2). Meanwhile, black permanent colorant could not be detected if the hair had been re-dyed with another permanent dye because of the presence of similar oxidation products. Furthermore, as long as nine weeks before analysis, SERS were able to sense colorant on dyed hair that had been subjected to normal daily washing.

Forensic hair analysis using FT-IR spectroscopy has also been explored. Boll et al. used ATR FT-IR spectra to (1) determine if the sample is of dyed or non-dyed hair, (2) distinguish brands of dye, and (3) discriminate between the dye color (i.e., black or medium brown), combined with the PLS-DA algorithm[91]. They examined hair collected from donors of different ages, biological sexes, and races. The natural color of the donated hair included blonde, brown, black, and red. The constructed three classification models provided an average prediction accuracy of 98.1% ± 3.0% for the spectrum level, and at least 90.0% confidence for donor-wise classification. Pienpinijtham et al. demonstrated ATR FT-IR spectra of human hair, which were measured using a homemade dome-shaped Ge μIRE accessory[92]. This equipment enabled detection of cosmetic residues on a single hair surface as well as the hair compositions. The ATR FT-IR spectra from hair from the same person were found to be identical even between black and white hairs. Hairs from different individuals are distinguishable because of differences in bands of peptide (Amide I, II, and III) and cosmetics particularly involving disulfide bond variations and the signals of silicone oil. Moreover, they showed that split hairs exhibited red-shifts of amide bands and a new peak at 1575 cm^{-1} compared because of peptide bond cleavage. Most recently, Contreras et al. showed techniques using ATR FT-IR spectroscopy to elucidate the hair history of bleaching and dyeing[93]. The PLS-DA models were constructed based on the ATR FT-IR spectra to discriminate whether hair had been bleached or not, and if bleached then to determine what bleaching agent had been used: a commercially available bleaching agent or professional bleach. Furthermore, the PLS-DA demonstrated the promising capability to assess whether hair had been colored before being bleached; then to assess whether the colorant used was permanent or semi-permanent.

In recent years, vibrational spectroscopy has demonstrated that it provides detailed chemical information for hair. Especially, cosmetic treatments of different types applied to hair (i.e., hair dyes and bleach) have been proved to exhibit characteristic changes of the spectral pattern. Actually, SERS has shown highly specific and strong bands derived from artificial colorants on the hair surface, allowing determination of the dye brand. Combination with chemometric modeling has enabled the objective discrimination and identification of hair treatments. Although further research and validation are necessary, vibrational spectroscopy is anticipated as a helpful tool to complement current microscopic observations for forensic hair examinations.

Forensic Anthropology

Forensic anthropology is a sub-discipline of forensic science for recovering, analyzing, and identifying human remains. The remains examined in criminal investigation are recoverable in various conditions such as decomposing, mummified, and skeletonized, depending on the surrounding environment and time since death. Regarding completeness, it ranges from a piece of bone, tooth and a part of body to an entire individual. Based on the recovered soft tissues and skeleton, forensic analysis of the remains is aimed at determining the biological profiles of the individuals (e.g., sex, age at death, and race), cause of death, and time since death, or postmortem interval (PMI). Such analysis engenders indication of deceased individuals and events before and after the death, which are helpful for forensic investigations. Traditionally, analytical methods applied by forensic anthropologists have consisted of observations and measurements of the remains. Observations of the remaining soft tissues can result in estimation of the cause of death and PMI. Measurements of the size and morphology of the skeletons, including distances between osteological landmarks, are useful to predict sex. The individual's age at the death can be estimated by epiphyseal fusion and dental development in addition to bone measurements. These current methods have been effectual. However, most of the techniques are qualitative. They can thereby be influenced by the examiners' bias. Moreover, the manual measurements might involve considerable error and might not be reproducible. Therefore, development of novel techniques based on an objective evaluation has been demanded to improve the accuracy and reliability of forensic anthropological analysis.

Vibrational spectroscopy selectively provides information about chemical states of the composite molecules in tissues[94] and bones[95]. Its quantitative and nondestructive nature is advantageous compared to current methods of forensic anthropology and other destructive analytical methods such as mass spectroscopy and protein or DNA analysis[94, 96]. Because of the nondestructive manner, vibrational spectroscopy of tissues has also been applied widely to medical diagnosis. The use of vibrational spectroscopy for the remains has been investigated for various previously described analytical interests and was described in 2015 in an earlier review by Muro[1]. A literature search has revealed that the recent pertinent reports were mostly about studies of estimation of age at death, PMI,

and analysis of burned bones for pre-burned metric estimation. Herein, vibrational spectroscopy is particularly advantageous because DNA analysis is unable to offer the information. Herein, we review recent works, particularly studies involving samples of human remains.

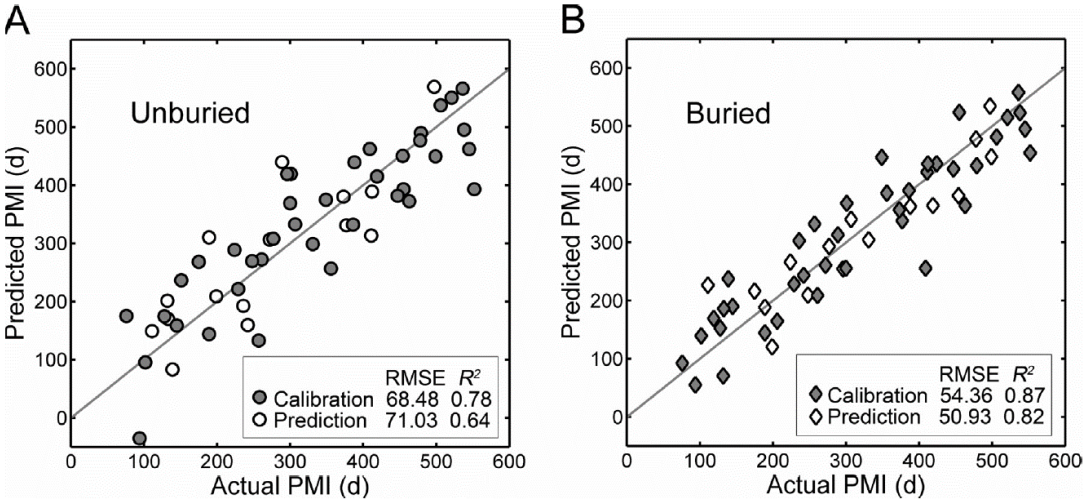

Fig. 3 Plots of predicted versus actual PMI from the calibration and prediction datasets by PLSR models coupled with GA. Gray lines represent the perfect prediction. Reproduced from ref. 100 with permission from Elsevier, copyright 2017.

In 2015, Pezzotti et al. assessed the use of Raman spectroscopy to predict human cadaver skin donor age[97]. Skin samples were obtained from five donors aged from a few months to 62 years. Raman spectra with a 532 nm excitation laser were measured from the top surface to deeper zones (approx. 700 μm) of skin. They attempted to assign all bands observed in the Raman spectra of skin via spectral deconvolution. Then they correlated the intensities with the donor age. Results show that Raman bands corresponding to protein folding were sensitive to infants and young individuals. The bands of lipid crystallization particularly varied with the age of adult individuals. Pedrosa et al. showed correlative relationships between distinctive bands of ATR FT-IR spectra and donor's age of bones[98]. The assessed bone samples were femora and humeri collected from 44 females and 36 males. They evaluated the bands of bone collagen (Am/P), carbonate type A (API), carbonate type B (BPI), the relation between the carbonate content (types A and B) to type B carbonate (C/C), carbonate-phosphate ratio (C/P), and the crystallinity index (CI). The femora of female bones showed increased CI and decreased BPI with donor age because of a crystalline structure disorder. Consequently, the potentials of these variables for age estimation were found, particularly for females. Most recently, Bonicelli et al. evaluated 113 rib cortical bones from subjects of 12–84 years old, using various methods including ATR FT-IR spectroscopy and other physicochemical and mechanical analyses[99]. Multivariate linear equations to predict the age of bone were developed by selectively combining variables extracted the respective analytical data. The best accuracy demonstrated was $R^2 = 0.863$ and

mean absolute error of 4.64 years.

Estimation of time since death or PMI has been a main subject of recent searches for human remains using vibrational spectroscopy. Wang et al. used FT-IR spectroscopy to construct prediction models of PMI of human bones[100]. For 76 to 552 days, bone samples collected from 56 human corpses were exposed to two conditions: buried (i.e., placed in soil) and unburied (i.e., exposed to the air). The PLSR models with GA provided prediction results of $R^2 = 0.64$ and RMSE = 71.03 days for unburied bones, and $R^2 = 0.82$ and RMSE = 50.93 days for buried bones (Fig. 3). Buried bones showed superior prediction accuracy probably because the rapid decomposition rate led to more significant spectral changes. Moreover, GA determined that amide I band was especially important, corresponding to protein degradation and the decrease of nitrogen. Woess et al. observed human bones using reflection and ATR IR microscopy and Raman microscopy with 785 nm excitation[101]. Assessed bone samples were archaeological ($n = 2$, PMI = 650 ± 870 years, 1030 ± 1260 years) and forensic ($n = 4$, PMI = 1 day – 85 years). The reflection and ATR IR spectra showed an increase of bone mineralization represented at 1042 cm^{-1} and a decrease of organic compounds (e,g., phospholipids, proteins, carbohydrates) in archaeological bones compared to forensic bones. Raman spectra provided a similar result: reduction of bands assigned to phospholipids, proteins, and carbohydrates in bone with higher PMI (> 3 years), especially in bones associated with archaeological sites. They also demonstrated that PCA of the observed spectral set enabled them to distinguish bone samples with different PMIs. Using a micro FT-IR spectrometer, Li et al. demonstrated PMI estimation based on annular cartilage samples[102]. Annular cartilage was assumed to be preferable for PMI estimation because of the slower degradation rate than those of other soft tissues and biofluids. For this study, annular cartilage samples were obtained from human remains for which PMIs were within 30 days. The samples were fixed in formalin and were then sliced into 4 μm thickness using a microtome. PLSR models to predict PMI were constructed while varying spectral processing. The best result was $R^2 = 0.95$ and RMSE = 1.49 days for the test dataset. Moreover, evaluation of "variable importance in projection" indicated significant contributions of the bands of collagen proteins and carbohydrates for PMI estimation. More recently, Baide et al. reported the use of a hand-held Raman spectrometer with an 830 nm laser to analyze PMI based on dental enamel surfaces of incisors[103]. A hand-held Raman spectrometer allows *in situ* analysis without risks to the preservation of bones and quick measurements in intervals of 30 s. Furthermore, incisor samples provide easy accessibility and involve less risk of detachment during exhumation than anterior dentition. Teeth in the early stage of decomposition (PMI of 22–42 days) showed a distinctive spectral profile and specifically higher intensities at 1402 cm^{-1}. The teeth in PMI of 64–84 days provided higher intensities of bands at 1134 and 1180 cm^{-1}. These features are expected to be useful to distinguish them from further decomposed teeth.

Heat exposure of remains induces shrinkage or warping of bones as well as destruction of surrounding soft tissues. Morphological alteration in burned bones hampers metric evaluation in

conventional anthropological analysis, causing difficulty of identifying sex, status, and age at death of

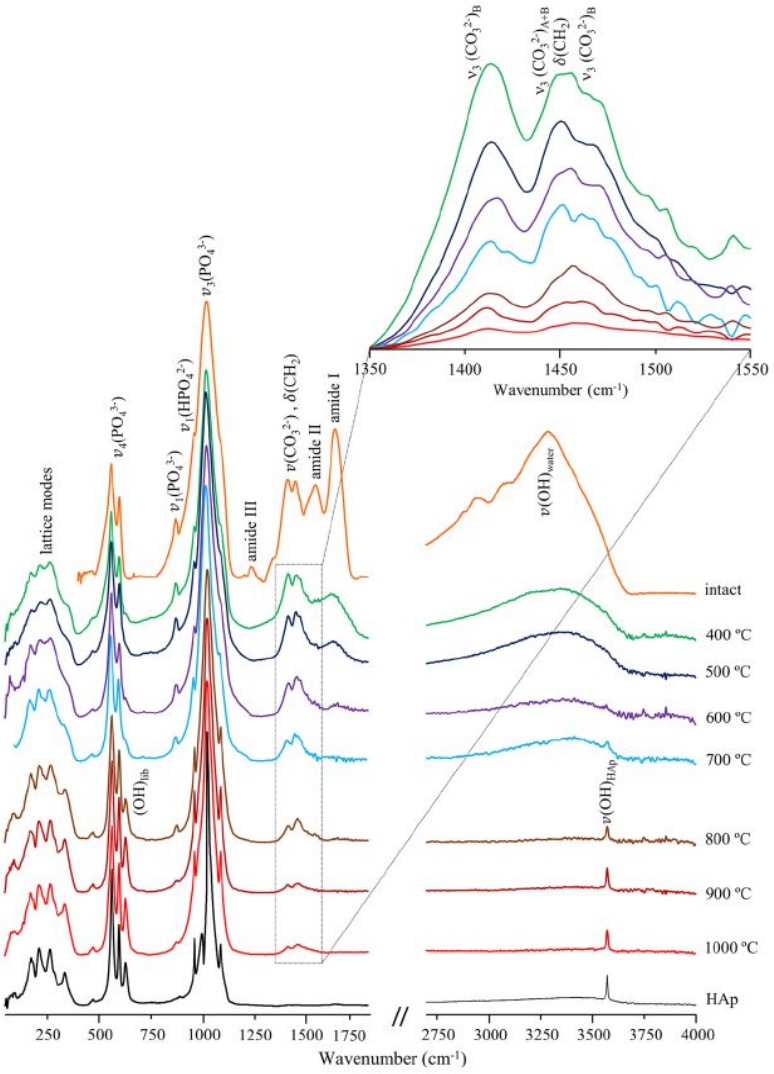

Fig. 4 FTIR-ATR spectra of human humerus: intact and burned at different temperatures (400–1000 °C). The inset represents a region of the asymmetric stretching carbonate bands $\nu_3(CO_2^{3-})$. The spectrum of reference calcium hydroxyapatite (HAp, SPM 2910b) is also included for comparison. Reproduced from ref. 107 with permission from Springer Nature, copyright 2018.

remains[104, 105]. Therefore, alternative methods would be useful to diagnose changes induced by a certain heat condition and subsequently facilitate estimation of the pre-burnt condition or metrics of bones. During the past few years, research groups including de Carvalho and Gonçalves et al. have consecutively reported analysis of burned bones using vibrational spectroscopy. In 2016, Vassalo et al. assessed the dependence of occurrence of bone warping on various bone attributes and experimental parameters of burning[106]. FT-IR spectroscopy was used to evaluate collagen contents of human bones. A logistic model to predict occurrence of bone warping revealed a significant contribution of collagen contents, but it was partial. This result indicated that bone warping is affected complexly by various factors such as maximum temperature and burning time or other un-assessed factors. In the following

study, Marques et al. assessed spectral features of FT-IR, Raman, and inelastic neutron scattering (INS) spectroscopy for human bones burned under controlled temperature conditions of 400–1000 °C (Fig. 4)[107]. Results of INS spectroscopy revealed detailed assignments of bans observed in respective spectroscopy. Particularly, bands of OH_{lib}, ν(OH) and $\nu_4(PO_4^{3-})$, which are all infrared-active, showed distinctive dependence on temperature. Consequently, it was inferred that these bands are useful as spectral biomarkers for routine analyses such as those using a bench-top ATR FT-IR spectrometer. Festa et al. also demonstrated the use of INS, FT-IR and Raman spectroscopy for ancient burned bones to investigate heating conditions[108]. Most recently, Gonçalves et al. reported a technique to predict pre-burned metric based on ATR FT-IR spectra of burned bones[109]. Measurements of bone metrics were conducted both before and after burning at various temperatures and for various durations. Indices to describe the heat-induced osteometric changes were generated based on the distinctive band intensities in the ATR FT-IR spectra of burned bones. Several multivariate regression models built for metric change prediction provided accuracy of R^2 = 0.19–0.54 and RMSE = 3.41–4.76. Furthermore, the constructed model showed better accuracy of sex determination of burned bones than other osteometric methods, and showed comparable accuracy to that found for estimation based on pre-burned bones.

As for miscellaneous research topics, Wang et al. used FT-IR spectroscopy to demonstrate discrimination between human and non-human (pig, goat, and cow) bones[110]. To simulate practical forensic cases, bone samples were prepared by boiling and decomposition, along with fresh ones. A PLS-DA model showed discrimination accuracies between human and non-human bones: 99.72% and 99.53%, respectively, for internal and external validation. Moreover, the loading plots of PLS-DA (and PCA) represented the diversity of inorganic portion (i.e., carbonates and phosphates), which is preferable for forensic practices because of long-term stability under various conditions. Lin et al. reported a technique for postmortem diagnosis to determine fetal anaphylactic shock based on FT-IR spectroscopic measurements of pulmonary edema[111]. In fact, PCA determined that the FT-IR spectra of the fetal anaphylactic shock group showed more contents of turn and α-helix protein and less tyrosine-rich protein than the control group including mechanical asphyxia, brain injury, and acute cardiac death. The PLS-DA model built by combining GA demonstrated good separation between these two groups.

Vibrational spectroscopy has been applied to various analytical subjects in forensic anthropology. Chemometric evaluation has contributed significantly to realize reliable and advanced analyses based on the observed spectra of bones, teeth, and soft tissues. The conditions of samples addressed in forensic anthropology can be changed drastically depending on the surrounding environments and the temporal interval as well as donors' physical properties. Such conditional effects are expected to be reflected to a considerable degree in variations of the spectral characteristics. Here, evaluating and describing experimental and simulating conditions precisely are anticipated as key factors to share

findings and to promote further progress in this research field. Although actual cases in forensic anthropology are quite diverse and complex, vibrational spectroscopy is expected to be a useful analytical tool for more objective and quantitative evaluation.

Bioagents

Bioagents (biological agents) are pathogens and their toxic products that threaten human health or even cause death of infected persons[112]. The diversity of bioagents includes viruses, bacteria, protozoa, fungi, and toxin from biological sources. In the context of forensics or homeland security, bioagents usable as biological weapons are an important concern. Bioagents involve severe properties in terms of public security and criminal investigation: they are often reproduced rapidly, easily transported, usually odorless, and are visually indistinguishable from ubiquitous materials such as powdered medicine and food; the recognition of outbreaks or appearance of a victim's symptoms lag behind the attack because of an incubation period; moreover, bioagents might be genetically modified for weaponization to increase transmissibility and lethality. Consequently, rapid and reliable identification of bioagents is a quite important task for defense against potential attacks.

Current methods used to characterize and detect bioagents are based on immunological assays including enzyme-linked immunosorbent assay (ELISA)[113], polymerase chain reaction (PCR), whole-genome sequencing[114], and mass spectrometry[112, 115]. However, these methods are constrained by their low sensitivity and time-consuming processes[112], which are entirely unsuitable for bioagent detection considering possible scenarios of bioterror. Here, vibrational spectroscopic techniques have some potential to fulfill the demands of rapid and reliable detection of bioagents. Furthermore, they might enable on-site analysis using portable instruments.

The Popp research group at Friedrich Schiller University Jena has been investigating analytical techniques of pathogenic microorganisms or bacteria using Raman spectroscopy for clinical diagnosis, food safety, agriculture, and environmental science[116, 117]. In recent work, Arend et al. demonstrated detection of infected neutrophils based on Raman spectra observed with 532 nm excitation[118]. They used an RF algorithm to classify the Raman spectra. The results showed discrimination between infected and non-infected neutrophils with 90% accuracy, and further demonstrated determination of pathogen species (i.e., bacteria or fungi) with 92% accuracy. Lorenz et al. reported prediction of *E.coli* pathogenicity[119]. They prepared seven strains of non-pathogenic *E.coli* and seven strains of pathogenic *E.coli* for training of a PCA-SVM classification model. The constructed model was assessed for discrimination of external *E.coli* strains: two pathogenic and one non-pathogenic. The average sensitivity was 77%, indicating its usefulness as a rapid screening method of *E.coli* pathogenicity.

Walper et al. provided an excellent and comprehensive review about detection methodologies for biothreat agents[112]. Although FT-IR spectroscopy for microorganism can be hampered by water interference, a few works for applications of FT-IR spectroscopy were examined. They included detections of foodborne pathogens such as *E.coli*, *Salmonella*, and *Listeria*.

SERS has been regarded as advantageous for bacteria detection because of its higher sensitivity and lower degree of interference by the fluorescence background than normal Raman spectroscopy. Villa et al. demonstrated SERS analysis to discriminate bacteria genera and species using SERS[120].

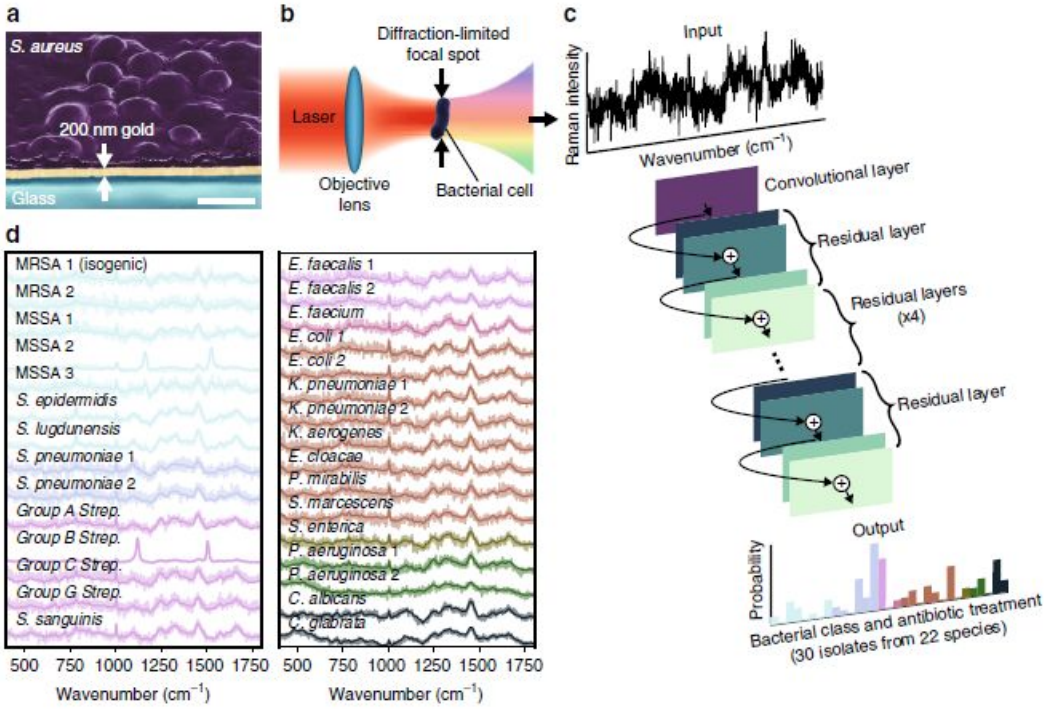

Fig. 5 A convolutional neural network (CNN) for identifying bacteria based on Raman spectra. (a) An SEM cross-section of the sample which bacterial cells are deposited on to gold-coated silica substrates. Scale bar is 1 μm. (b) Measurement schematic: Raman signal from single cells can be acquired from a diffraction-limited spot size by laser focusing. (c) A schematic of classification of low-signal Raman spectra via a one-dimensional residual network with 25 total convolutional layers into one of 30 isolates and antibiotic treatment groups. (d) Average Raman spectra of 30 bacterial isolates (bold lines) overlaid on representative noisy single spectra for each isolate. The spectra are colored according to antibiotic treatment group. Reproduced from ref. 124 with permission from Springer Nature, copyright 2019.

For their study, filter paper coated with gold nanoparticles was used as the SERS substrate. The substrate is inexpensive for fabrication; it provides numerous hot spots for SERE with high porosity. The PLS-DA model for the SERS spectra of bacteria achieved complete classification at the genus level, with correct classification at the species level except for one sample among 60 test samples. Moreover, potential identification of new species was demonstrated based on the proposed model by outlier analysis with Q-residuals and Hotelling's T^2 values. Liu et al. reported detection of pathogenic bacteria using a silver nanorod (AgNR)-based SERS substrate optimized for bacterial analysis[121]. The AgNR monolayer formed by air-liquid interface self-assembly method showed sensitive SERS signals of 22 bacterial strains. A *t*-test analysis unveiled the most distinctive regions in SERS spectra of each strain from those of the other 21 strains. Consequently, ROC analysis for the selected spectral features determined that 20 out of 22 strains could be discriminated. More SERS studies have been reported recently for efficient pathogen detection using other types of SERS substrate: detection of atmospheric bioaerosol by a commercial SERS substrate (i.e., Klarite)[122] and virus detection by a novel substrate

composed of hollow nanocones at the bottom of microbowls[123].The combination of Raman spectroscopy and deep learning technologies is promising for the identification of various bacteria species. Ho et al. applied a CNN to classification of 30 common pathogenic bacteria (Fig. 5)[124]. Herein, an extensive dataset of the bacteria Raman spectra was collected using 633 nm illumination; then it was used for model training. Classification into 30-class bacteria strains was achieved with more than 82% accuracy, even for the noisy Raman spectra observed in 1 s measurement times. Antibiotic treatment for bacteria was also identified with 97.0 ±0.3% classification accuracy. The author finally validated their approaches for empirical clinical strains obtained from patients, then demonstrated eminent accuracy (approx. 99.7%) of identifying appropriate treatment corresponding to the respective pathogens.

Vibrational spectroscopy has shown sufficient capability to represent characteristics of bioagents based on their spectra. The observation is applicable without sample culturing as well as in the rapid and non-destructive manner. Data analysis approaches involving chemometrics and artificial neural network have been strongly advancing the reliability and versatility of detection techniques of pathogenic organism. Additional research for bioagent identification should progress considering selectivity, rapidity, and applicability for various bioagent types, sample conditions and possible genetic modifications.

Conclusions and Future Perspectives

In contrast to current methods used in the field of forensic biology, such as visual and microscopic observations and serological and biochemical techniques, vibrational spectroscopic analysis presents several important advantages of non-destructivity, rapidness, and quantitative observation. Additionally, when used in combination with chemometric techniques, it offers objective and statistical evaluation of complex vibrational spectra, which is especially demanded for the development of modern forensic science.

Throughout the extensive studies conduced during past decades, the potential applications of vibrational spectroscopy have been demonstrated for various forensic biological subjects. At the same time, the fundamental issues and demands to be investigated further before practical implementation have been indicated. Particularly, they are sample degradation, substrate interference, mixture and contamination with other materials, and effects of environmental conditions. Moreover, the analytical evaluation needs to be accompanied by error rate estimation. Inter-laboratory validation is also necessary to confirm the independence of executed conditions and instruments. In that sense, a spectral database or library is requested: it can be expected to facilitate inter-laboratory comparison of the spectral data. Then it enables develop spectral processing techniques to correct differences among laboratories. In addition, the database can be expected to help to construct more versatile chemometric

models by training with large amounts of spectral data observed by different experimental conditions. Another direction of future development is integration of chemometric tools. As this Review has discussed, a single vibrational spectrum can provide various information that is expected to be useful for different analytical purposes in forensic biology. Therefore, the establishment of a chemometric platform to provide manifold analytical outputs from a single examined datum (or dataset) can be an ultimate goal of developments in these research fields.

As for the instruments, the possibility of *in-situ* analysis using portable or hand-held instruments is another advantage of vibrational spectroscopy. Whereas most current studies have been conducted using desktop instruments, the usability of spectra observed using hand-held Raman spectrometers has already been demonstrated in several works for analyzing hair, body fluid stains, and teeth. Further expansion of applications of the portable instruments is being pursued. Furthermore, precise comparisons of the analytical qualities between desktop and portable instruments are demanded.

Despite the research progress discussed herein, the position of vibrational spectroscopic analysis in the sequence of forensic biological exams has not been determined. Whereas some developed methods have aimed to provide confirmatory or conclusive results, such techniques still require improvements and more rigorous validation, considering various points of difficulty encountered in forensic casework. Supplementary use to the current methods is a possible method in practice, such as for screening or exploratory analysis and corroboration of results using current methods. For either usage, whether confirmatory or supplementary, defining suitable experimental conditions or limitations for the reliable use of new techniques is expected to be necessary.

Requirements for further study and improvement remain, but vibrational spectroscopy has promising potential for application to forensic biological analysis. In addition, its combination with chemometrics has been proven to be invaluable for completing reliable and advanced analyses of the complex spectra of forensic biological materials. The developments of vibrational spectroscopic analysis are expected to contribute to increased reliability and efficiency of practical examinations in forensic biology. Moreover, portable instruments can provide alternative means and modes of forensic biological investigation in the field. Continued investigation is anticipated by effectively integrating knowledge and perspectives gained from vibrational spectroscopy, chemometrics, and forensics.

Author contributions

Ayari Takamura: Conceptualization, Investigation, Writing – original draft, Writing – review & editing. Takeaki Ozawa: Conceptualization, Writing – review & editing.

Conflicts of interest

There are no conflicts to declare.

Acknowledgments

This work was supported by CREST (JPMJCR1752 to T.O.) from Japan Science and Technology (JST), Japan Society for the Promotion of Science (JSPS) KAKENHI (Grant-in-Aid for Scientific Research (A) 19H00900 to T.O, and the RIKEN Special Postdoctoral Researcher Program to A.T.

References

1. C. K. Muro, K. C. Doty, J. Bueno, L. Halámková and I. K. Lednev, *Anal. Chem.*, 2015, **87**, 306-327.
2. R. Kumar and V. Sharma, *Trends Anal. Chem.*, 2018, **105**, 191-201.
3. Center for National Research, *Strengthening forensic science in the United States: a path forward*, National Academies Press, 2009.
4. K. C. Doty, C. K. Muro, J. Bueno, L. Halámková and I. K. Lednev, *J. Raman Spectrosc.*, 2016, **47**, 39-50.
5. S. R. Khandasammy, M. A. Fikiet, E. Mistek, Y. Ahmed, L. Halámková, J. Bueno and I. K. Lednev, *Forensic Chem.*, 2018, **8**, 111-133.
6. C. Muehlethaler, M. Leona and J. R. Lombardi, *Anal. Chem.*, 2016, **88**, 152-169.
7. M. A. Fikiet, S. R. Khandasammy, E. Mistek, Y. Ahmed, L. Halámková, J. Bueno and I. K. Lednev, *Spectrochim. Acta A: Mol. Biomol. Spectrosc.*, 2018, **197**, 255-260.
8. E. Mistek and I. K. Lednev, *Spectrosc.*, 2018, **33**, 8-19.
9. C. S. Silva, A. Braz and M. F. Pimentel, *J. Braz. Chem. Soc.*, 2019, **30**, 2259-2290.
10. A. C. Albrecht and M. C. Hutley, *J. Chem. Phys.*, 1971, **55**, 4438-4443.
11. H. Hamaguchi and K. Iwata, *Raman Spectroscopy*, Kodansha, Tokyo, 2015.
12. Y. Furukawa, *Infrared Spectroscopy*, Kodansha, Tokyo, 2018.
13. R. G. Brereton, J. J. Jansen, J. Lopes, F. Marini, A. Pomerantsev, O. Rodionova, J. M. Roger, B. J. Walczak and R. Tauler, *Anal. Bioanal. Chem.*, 2017, **409**, 5891-5899.
14. R. Brereton, J. Jansen, J. Lopes, F. Marini, A. Pomerantsev, O. Rodionova, J. Roger, B. Walczak and R. Tauler, *Anal. Bioanal. Chem.*, 2018, **410**, 6691-6704.
15. R. Gautam, S. Vanga, F. Ariese and S. Umapathy, *EPJ Tech. Instrum.*, 2015, **2**, 8.
16. P. Lasch, *Chemom. Intell. Lab. Syst.*, 2012, **117**, 100-114.
17. S. S. Tobe, N. Watson and N. N. Daéid, *J. Forensic Sci.*, 2007, **52**, 102-109.
18. F. Barni, S. W. Lewis, A. Berti, G. M. Miskelly and G. Lago, *Talanta*, 2007, **72**, 896-913.
19. K. Virkler and I. K. Lednev, *Forensic science international*, 2008, **181**, e1-e5.
20. K. Virkler and I. K. Lednev, *Anal. Bioanal. Chem.,* 2010, **396**, 525-534.
21. K. Virkler and I. K. Lednev, *Analyst*, 2010, **135**, 512-517.
22. K. Virkler and I. K. Lednev, *Forensic Sci. Int.*, 2009, **193**, 56-62.

23. V. Sikirzhytski, A. Sikirzhytskaya and I. K. Lednev, *Anal. Chim. Acta*, 2012, **718**, 78-83.
24. A. Sikirzhytskaya, V. Sikirzhytski and I. K. Lednev, *Forensic Sci. Int*, 2012, **216**, 44-48.
25. V. Sikirzhytski, K. Virkler and I. K. Lednev, *Sens.*, 2010, **10**, 2869-2884.
26. A. Sikirzhytskaya, V. Sikirzhytski and I. K. Lednev, *J. Biophotonics*, 2014, **7**, 59-67.
27. C. K. Muro, K. C. Doty, L. de Souza Fernandes and I. K. Lednev, *Forensic Chem.*, 2016, **1**, 31-38.
28. B. Vyas, L. Halámková and I. K. Lednev, *Forensic Chem.*, 2020, **20**, 100247.
29. W. R. Premasiri, J. C. Lee and L. D. Ziegler, *J. Phys. Chem. B*, 2012, **116**, 9376-9386.
30. A. Bonifacio, S. Dalla Marta, R. Spizzo, S. Cervo, A. Steffan, A. Colombatti and V. Sergo, *Anal. Bioanal. Chem.*, 2014, **406**, 2355-2365.
31. M. L. Shaine, W. R. Premasiri, H. M. Ingraham, R. Andino, P. Lemler, A. N. Brodeur and L. D. Ziegler, *Analyst*, 2020, **145**, 6097-6110.
32. K. M. Elkins, *J. Forensic Sci.*, 2011, **56**, 1580-1587.
33. C.-M. Orphanou, *Forensic Sci. Int.*, 2015, **252**, e10-e16.
34. F. Zapata, M. Á. F. de la Ossa and C. García-Ruiz, *Appl. Spectrosc.*, 2016, **70**, 654-665.
35. A. Takamura, K. Watanabe, T. Akutsu, H. Ikegaya and T. Ozawa, *Anal. Chem.*, 2017, **89**, 9797-9804.
36. A. Takamura, K. Watanabe, T. Akutsu and T. Ozawa, *Sci. Rep.*, 2018, **8**, 8459.
37. K. Virkler and I. K. Lednev, *Anal. Chem.*, 2009, **81**, 7773-7777.
38. G. McLaughlin, K. C. Doty and I. K. Lednev, *Forensic Sci. Int.*, 2014, **238**, 91-95.
39. K. C. Doty and I. K. Lednev, *Forensic Sci. Int.*, 2018, **282**, 204-210.
40. P. Wang, L. Guo, Y. Tian, J. Chen, S. Huang, C. Wang, P. Bai, D. Chen, W. Zhu, H. Yang, W. Yao and J. Gao, *OSA Contin.*, 2021, **4**, 672-687.
41. E. Mistek and I. K. Lednev, *Anal. Bioanal. Chem.*, 2015, **407**, 7435-7442.
42. E. Mistek-Morabito and I. K. Lednev, *Commun. Chem.*, 2020, **3**, 178.
43. H. Lin, Y. Zhang, Q. Wang, B. Li, S. Fan and Z. Wang, *Int. J. Legal Med.*, 2018, **132**, 667-674.
44. X. Wei, K. Yu, D. Wu, P. Huang, Q. Sun and Z. Wang, *Int. J. Legal Med.*,, 2021, **135**, 73-80.
45. A. Sikirzhytskaya, V. Sikirzhytski and I. K. Lednev, *Anal. Chem.*, 2017, **89**, 1486-1492.
46. E. Mistek, L. Halámková and I. K. Lednev, *Forensic Chem.*, 2019, **16**, 100176.
47. E. Mistek, L. Halámková, K. C. Doty, C. K. Muro and I. K. Lednev, *Anal. Chem.*, 2016, **88**, 7453-7456.
48. K. C. Doty and I. K. Lednev, *ACS Cent. Sci.*, 2018, **4**, 862-867.
49. S. Giuliano, E. Mistek-Morabito and I. K. Lednev, *ACS Omega*, 2020, **5**, 27026-27031.
50. C. K. Muro, L. de Souza Fernandes and I. K. Lednev, *Anal. Chem.*, 2016, **88**, 12489-12493.
51. C. K. Muro and I. K. Lednev, *Anal. Chem.*, 2017, **89**, 4344-4348.
52. A. Takamura, L. Halamkova, T. Ozawa and I. K. Lednev, *Anal. Chem.*, 2019, **91**, 6288-6295.

53. R. H. Bremmer, K. G. de Bruin, M. J. C. van Gemert, T. G. van Leeuwen and M. C. G. Aalders, *Forensic Sci. Int.*, 2012, **216**, 1-11.
54. A. R. Weber and I. K. Lednev, *Forensic Chem.*, 2020, **19**, 100248.
55. B. L. Horecker, *J. Biol. Chem.*, 1943, **148**, 173-183.
56. L. Tomellini, *Archives d'antropologie criminelle de Criminologie*, 1907, **14**.
57. D. Patterson, *Nature*, 1960, **187**, 688-689.
58. C. G. Atkins, K. Buckley, M. W. Blades and R. F. Turner, *Appl. Spectrosc.*, 2017, **71**, 767-793.
59. B. R. Wood and D. McNaughton, *J. Raman Spectrosc.*, 2002, **33**, 517-523.
60. S. Hu, K. M. Smith and T. G. Spiro, *J. Am. Chem. Soc.*, 1996, **118**, 12638-12646.
61. P. Lemler, W. R. Premasiri, A. DelMonaco and L. D. Ziegler, *Anal. Bioanal. Chem.*, 2014, **406**, 193-200.
62. M. Asghari‐Khiavi, A. Mechler, K. R. Bambery, D. McNaughton and B. R. Wood, *J. Raman Spectrosc.*, 2009, **40**, 1668-1674.
63. B. R. Wood, P. Caspers, G. J. Puppels, S. Pandiancherri and D. McNaughton, *Anal. Bioanal. Chem.*, 2007, **387**, 1691-1703.
64. B. R. Wood, B. Tait and D. McNaughton, *Biochim. Biophys. Acta, Mol. Cell Res.*, 2001, **1539**, 58-70.
65. B. R. Wood, L. Hammer, L. Davis and D. McNaughton, *J. Biomed. Opt.*, 2005, **10**, 014005.
66. A. Takamura, D. Watanabe, R. Shimada and T. Ozawa, *Commun. Chem.*, 2019, **2**, 115.
67. K. C. Doty, G. McLaughlin and I. K. Lednev, *Anal. Bioanal. Chem.*, 2016, **408**, 3993-4001.
68. K. C. Doty, C. K. Muro and I. K. Lednev, *Forensic Chem.*, 2017, **5**, 1-7.
69. H. Lin, Y. Zhang, Q. Wang, B. Li, P. Huang and Z. Wang, *Sci. Rep.*, 2017, **7**, 13254.
70. R. Kumar, K. Sharma and V. Sharma, *Sci. Justice*, 2020, **60**, 538-546.
71. S. Zha, X. Wei, R. Fang, Q. Wang, H. Lin, K. Zhang, H. Zhang, R. Liu, Z. Li, P. Huang and Z. Wang, *Forensic Sci. Res.*, 2020, **5**, 119-125.
72. T. Das, A. Harshey, A. Srivastava, K. Nigam, V. K. Yadav, K. Sharma and A. Sharma, *Sci. Rep.*, 2021, **11**, 11855.
73. A. A. Quinn and K. M. Elkins, *J. Forensic Sci.*, 2017, **62**, 197-204.
74. S. Sharma, R. Chophi and R. Singh, *Int. J. Legal Med.*, 2020, **134**, 63-77.
75. G. McLaughlin, V. Sikirzhytski and I. K. Lednev, *Forensic Sci. Int.*, 2013, **231**, 157-166.
76. G. McLaughlin and I. K. Lednev, *J. Forensic Sci.*, 2015, **60**, 595-604.
77. S. Sharma and R. Singh, *Int. J. Legal Med.*, 2020, **134**, 411-432.
78. S. Sharma and R. Singh, *Int. J. Legal Med.*, 2020, **134**, 1591-1602.
79. E. Hager, C. Farber and D. Kurouski, *Forensic Chem.*, 2018, **9**, 44-49.
80. V. Sikirzhytski, A. Sikirzhytskaya and I. K. Lednev, *Forensic Sci. Int.*, 2012, **222**, 259-265.
81. A. Sikirzhytskaya, V. Sikirzhytski, G. McLaughlin and I. K. Lednev, *J. Forensic Sci.*, 2013, **58**,

1141-1148.

82. G. McLaughlin, M. A. Fikiet, M. Ando, H. Hamaguchi and I. K. Lednev, *J. Raman Spectrosc.*, 2019, **50**, 1147-1153.
83. M. Ando, I. K. Lednev and H. Hamaguchi, in *Frontiers and Advances in Molecular Spectroscopy*, Elsevier, 2018, pp. 369-378.
84. M. Ando and H. Hamaguchi, *J. Spectrosc. Soc. Jpn.*, 2015, **64**, 280-284.
85. President's Council of Advisors on Science and Technology. *Report to the President Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods*, 2016.
86. A. Kuzuhara, *Int. J. Cosmet. Sci.*, 2016, **38**, 201-209.
87. A. Kuzuhara, *Int. J. Cosmet. Sci.*, 2018, **40**, 34-43.
88. J. D. dos Santos, H. G. M. Edwards and L. F. C. de Oliveira, *Heliyon*, 2019, **5**, e01582.
89. D. Kurouski and R. P. Van Duyne, *Anal. Chem.*, 2015, **87**, 2901-2906.
90. I. Esparza, R. Wang and D. Kurouski, *Anal. Chem.,* 2019, **91**, 7313-7318.
91. M. S. Boll, K. C. Doty, R. Wickenheiser and I. K. Lednev, *Forensic Chem.*, 2017, **6**, 1-9.
92. P. Pienpinijtham, C. Thammacharoen, S. Naranitad and S. Ekgasit, *Spectrochim. Acta A: Mol. Biomol. Spectrosc.*, 2018, **197**, 230-236.
93. F. Contreras, A. Ermolenkov and D. Kurouski, *Anal. Methods*, 2020, **12**, 3741-3747.
94. N. S. Ozek, A. Afnan, J. Bueno, G. McLaughlin, M. Ralbovsky, A. Sikirzhytskaya, V. Sikirzhytski, F. Severcan and I. K. Lednev, *Vibrational Spectroscopy in Diagnosis and Screening*, IOS Press, Amsterdam, 2012.
95. C. d. C. A. Lopes, P. H. J. O. Limirio, V. R. Novais and P. Dechichi, *Appl. Spectrosc. Rev.*, 2018, **53**, 747-769.
96. H. A. Zimmerman, C. J. Meizel-Lambert, J. J. Schultz and M. E. Sigman, *Sci. Justice*, 2015, **55**, 131-138.
97. G. Pezzotti, M. Boffelli, D. Miyamori, T. Uemura, Y. Marunaka, W. Zhu and H. Ikegaya, *J. Biomed. Opt.*, 2015, **20**, 065008.
98. M. Pedrosa, F. Curate, L. A. E. Batista de Carvalho, M. P. M. Marques and M. T. Ferreira, *Int. J. Legal Med.*, 2020, **134**, 1905-1914.
99. A. Bonicelli, P. Zioupos, E. Arnold, K. D. Rogers, B. Xhemali and E. F. Kranioti, *Sci. Rep.*, 2021, **11**, 2086.
100. Q. Wang, Y. Zhang, H. Lin, S. Zha, R. Fang, X. Wei, S. Fan and Z. Wang, *Forensic Sci. Int.*, 2017, **281**, 113-120.
101. C. Woess, S. H. Unterberger, C. Roider, M. Ritsch-Marte, N. Pemberger, J. Cemper-Kiesslich, P. Hatzer-Grubwieser, W. Parson and J. D. Pallua, *PLoS One*, 2017, **12**, e0174552.
102. Z. Li, J. Huang, Z. Wang, J. Zhang and P. Huang, *Forensic Sci. Med. Pathol.*, 2019, **15**, 521-527.
103. A. Baide, C. Farber, M. Krimmer, D. Wescott and D. Kurouski, *Forensic Chem.*, 2020, **20**, 100270.

104. T. J. U. Thompson, *Forensic Sci. Int.*, 2004, **146**, S203-S205.
105. T. Thompson, *J. Forensic Sci.*, 2005, **50**, JFS2004297-2004298.
106. A. R. Vassalo, E. Cunha, L. A. E. B. de Carvalho and D. Gonçalves, *Int. J. Legal Med.*, 2016, **130**, 1647-1656.
107. M. P. M. Marques, A. P. Mamede, A. R. Vassalo, C. Makhoul, E. Cunha, D. Gonçalves, S. F. Parker and L. A. E. Batista de Carvalho, *Sci. Rep.*, 2018, **8**, 15935.
108. G. Festa, C. Andreani, M. Baldoni, V. Cipollari, C. Martínez-Labarga, F. Martini, O. Rickards, M. F. Rolfo, L. Sarti, N. Volante, R. Senesi, F. R. Stasolla, S. F. Parker, A. R. Vassalo, A. P. Mamede, L. A. E. Batista de Carvalho and M. P. M. Marques, *Sci. Adv.*, 2019, **5**, eaaw1292.
109. D. Gonçalves, A. R. Vassalo, C. Makhoul, G. Piga, A. P. Mamede, S. F. Parker, M. T. Ferreira, E. Cunha, M. P. M. Marques and L. A. E. B. de Carvalho, *Am. J. Phys. Anthropol.*, 2020, **173**, 734-747.
110. Q. Wang, W. Li, R. Liu, K. Zhang, H. Zhang, S. Fan and Z. Wang, *Int. J. Legal Med.*, 2019, **133**, 269-276.
111. H. Lin, Y. Luo, L. Wang, K. Deng, Q. Sun, R. Fang, X. Wei, S. Zha, Z. Wang and P. Huang, *Int. J. Legal Med.*, 2018, **132**, 477-486.
112. S. A. Walper, G. Lasarte Aragonés, K. E. Sapsford, C. W. Brown, C. E. Rowland, J. C. Breger and I. L. Medintz, *ACS Sens.*, 2018, **3**, 1894-2024.
113. A. H. Peruski and L. F. Peruski Jr, *Clin. Vaccine Immunol.*, 2003, **10**, 506-513.
114. M. Visser, R. Bester, J. T. Burger and H. J. Maree, *Virol. J.*, 2016, **13**, 1-6.
115. E. Duriez, J. Armengaud, F. Fenaille and E. Ezan, *J. Mass Spectrom.*, 2016, **51**, 183-199.
116. S. Stöckel, J. Kirchhoff, U. Neugebauer, P. Rösch and J. Popp, *J. Raman Spectrosc.*, 2016, **47**, 89-109.
117. B. Lorenz, C. Wichmann, S. Stöckel, P. Rösch and J. Popp, *Trends Microbiol.*, 2017, **25**, 413-424.
118. N. Arend, A. Pittner, A. Ramoji, A. S. Mondol, M. Dahms, J. Rüger, O. Kurzai, I. W. Schie, M. Bauer, J. Popp and U. Neugebauer, *Anal. Chem.*, 2020, **92**, 10560-10568.
119. B. Lorenz, N. Ali, T. Bocklitz, P. Rösch and J. Popp, *Anal. Bioanal. Chem.*, 2020, **412**, 8241-8247.
120. J. E. L. Villa, N. R. Quiñones, F. Fantinatti-Garboggini and R. J. Poppi, *Anal. Bioanal. Chem.*, 2019, **411**, 705-713.
121. S. Liu, Q. Hu, C. Li, F. Zhang, H. Gu, X. Wang, S. Li, L. Xue, T. Madl, Y. Zhang and L. Zhou, *ACS Sens.*, 2021, **6**, 2911-2919.
122. M. A. Tahir, X. Zhang, H. Cheng, D. Xu, Y. Feng, G. Sui, H. Fu, V. K. Valev, L. Zhang and J. Chen, *Analyst*, 2020, **145**, 277-285.
123. X. Zhang, X. Zhang, C. Luo, Z. Liu, Y. Chen, S. Dong, C. Jiang, S. Yang, F. Wang and X. Xiao, *Small*, 2019, **15**, 1805516.
121. C.-S. Ho, N. Jean, C. A. Hogan, L. Blackmon, S. S. Jeffrey, M. Holodniy, N. Banaei, A. A. E.

Saleh, S. Ermon and J. Dionne, *Nature Commun.*, 2019, **10**, 4927.

Table 1 Body fluid analysis studies for different analytical interests discussed in this review

Analytical interest		Spectroscopic technique[a]	Body fluid sample	Chemometric technique/Spectral analysis	Model validation	Ref.
Body fluid identification		Confocal RS (785 nm)	Human semen, canine semen, vaginal fluid, saliva, sweat, blood	Visual comparison of characteristic peaks	—	19
			Blood	SFA, PCA, ALS, Goodness of fit statistics	EV	20
			Saliva	SFA, ALS, Goodness of fit statistics	EV	21
			Semen	SFA, ALS, Goodness of fit statistics	EV	22
			Sweat	SFA, PCA, ALS, Goodness of fit statistics	EV*	23
			Vaginal fluid	SFA, PCA, ALS, Goodness of fit statistics	EV	24
			Blood, saliva, semen	SFA, ALS, SIMCA, LDA, PLS-DA	CV, EV*	25
			Peripheral blood, menstrual blood, vaginal fluid	PCA, PLS-DA. SVM-DA, ROC	CV	26
			Peripheral blood, saliva, semen, sweat, vaginal fluid	PLS-DA, SVM-DA, interval PLS-DA, GA	CV, EV	27
			Peripheral blood, saliva, semen, sweat, vaginal fluid, urine	Nonnegative PCA, SVM-DA coupled with GA	CV, EV	28
		SERS (785 nm)	Blood (whole blood, red blood cells, blood plasma)	Visual comparison and quantification of characteristic peaks	—	29
		SERS (514, 633, and 785 nm)	Blood plasma, serum	Visual comparison and quantification of spectral regions of interest	—	30
		SERS (785 nm)	Blood, four other body fluids (saliva, semen, urine, vaginal fluid)	PLS-DA	CV, EV*	31
		ATR FT-IR	Blood, nasal mucus, vaginal mucus, saliva, tears, urine (other human solid material)	Visual comparison of characteristic peaks	—	32
			Blood, saliva, semen, vaginal secretion	Comparison of characteristic peaks	—	33
			Antemortem and postmortem blood	PLS-DA, GA	CV	35
			Peripheral blood, saliva, semen, urine, sweat	HCA, PLS-DA, LDA, Q-statistics	CV	36
		External reflection FT-IR	Semen, vaginal fluid, urine on cotton fabrics	PCA, SIMCA	EV	34
Species identification		Confocal RS (785 nm)	Human, canine and feline blood	SFA, PCA	CV	37
			Human blood, animal blood from 11 species	PLS-DA	CV, EV	38
			Human blood, animal blood from 16 species	PLS-DA, ROC	CV, EV	39
			Human blood, animal blood from 19 species	RNN, ROC	CV, EV	40
		ATR FT-IR	Human, cat, and dog blood	PLS-DA, GA	CV, EV	41
			Human blood, animal blood from 11 species	PLS-DA, GA	CV, EV	42
			Human blood, animal blood from 5 species	PCA, K-means clustering, PLS-DA	CV, EV	43
			Supernatant of human semen and animal semen from 5 animal species	PCA, PLS-DA	CV, EV	44
Phenotype profiling						
	Sex	Confocal RS (785 nm)	Blood from male and female donors	HCA, SVM-DA, ANN, GA	CV, EV	45
			Saliva from male and female donors	SVM-DA, ROC	CV, EV	50
		ATR FT-IR	Urine from male and female donors	PLS-DA, ANN, GA	CV	52
	Sex, race	ATR FT-IR	Blood from male/female Caucasian, African American, and Hispanic donors	PCA, PLS-DA, ROC	CV, EV	46
	Race	Confocal RS (785 nm)	Blood from Caucasian and African American donors	PCA, SVM-DA, GA, ROC	CV	47
			Semen from Black and Caucasian donors	SVM-DA, GA, ROC	CV, EV	51
	Age group	Confocal RS (785 nm)	Blood from newborn, adolescent, and adult donors	SVM-DA	CV, EV	48
		ATR FT-IR	Blood from newborn, adolescent, and adult donors	PLS-DA, GA	CV	49
Prediction of TSD		Confocal RS (785 nm)	Bloodstains aged up to one week	2D correlation spectroscopic analysis, PLSR	CV, EV	67
			Bloodstains aged up to two years	PLSR, PCR	CV, EV	68
			Bloodstains aged at three different	MCR-ALS, Modified ALS for kinetics	EV	66

		temperatures up to 121 days			
	ATR FT-IR	Bloodstains aged up to 107 days in indoor/outdoor conditions	PLSR, PLS-DA	CV, EV	69
		Bloodstains aged for 1–175 days	Curve estimation, MLR, PLSR	CV	70
		Supernatant of semen stains aged up to 6 days on glass, tissue, and fabric	PLSR	CV, EV	71
		Semen, saliva, urine during drying	PCR, PLSR	CV, EV	72

[a] Wavelength used for excitation is shown in parentheses.

*Validated partially rather than for all models reported nor for all sample types assessed.

RS: Raman spectroscopy; SERS: surface enhanced Raman spectroscopy; ATR: attenuated total reflection; FT-IR: Fourier transform infrared spectroscopy; CV: cross-validation; EV: external validation; TSD: time since deposition; SFA: significant factor analysis; PCA: principal component analysis; ALS: alternating least squares; SIMCA: soft independent modeling of class analogy; PLS-DA: partial least squares discriminant analysis; LDA: linear discriminant analysis; SVM-DA: support vector machine discriminant analysis; SFA: significant factor analysis, PCA: principal component analysis; MCR: multivariate curve resolution; ROC: receiver operating characteristic analysis; GA: genetic algorithm; HCA: hierarchical clustering analysis; RNN: recurrent neural network; ANN: artificial neural network; PLSR: partial least squares regression; PCR: principal component regression; MLR: multiple linear regression.

Table 2 Vibrational spectroscopic studies for body fluid mixture and substrate interference[a]

Spectroscopic technique[b]	Body fluid sample	Interfering substrate	Spectral analysis for body fluid mixture/ substrate interference/	Ref.
ATR FT-IR	Blood, nasal mucus, vaginal mucus, saliva, tears, urine (other human solid materials)	White T-shirt, white copier paper	Visual comparison of characteristic peaks	32
ATR FT-IR	Venous blood, menstrual blood, semen, saliva, breastmilk	Cotton, nylon, wood, paper, glass	Visual comparison of characteristic peaks	73
ATR FT-IR	Menstrual blood	White cotton, denim, polyester, paper, wood, plastic, grass, glass, floor tile, sanitary napkin	Visual comparison of characteristic peaks	74
ATR FT-IR	Semen, semen/vaginal fluid mixture	White cotton, denim, polyester, paper, wood, skin, plastic, grass, glass, condom, floor tile	Visual comparison of characteristic peaks	77
ATR FT-IR	Vaginal fluid, vaginal fluid/semen mixture after coitus, vaginal fluid/menstrual or peripheral blood mixture	White cotton cloth, denim, polyester, tissue paper, glass, plastic, floor tiles, polished wood	Visual comparison of characteristic peaks	78
Hand-held RS (1064 nm)	Urine	White cotton, blue jeans, lab coat, uniform shirt, sweat-contaminated cotton shirt	PLS-DA (CV)	79
Confocal RS (785 nm)	Blood	Dust, sand, soil	Multivariate regression with body fluid signatures	75
Confocal RS (406.7, 457.9, 488, 514.5, 647.1 and 785 nm)	Blood	Glass, tile, denim, cotton	Manual subtraction of substrate contribution, multivariate regression with body fluid signatures	81
Confocal RS (785 nm)	Semen	Pig skin, glass, cotton, polyester, blended fabric	Manual subtraction of substrate contribution, multivariate regression with body fluid signatures	76
ATR FT-IR	Antemortem and postmortem blood	Cotton, denim, polyester	Multivariate regression using PCA loadings weighted by a factor derived from PLS-DA loadings	35
Confocal RS (785 nm)	Semen	Glass, polyester	MCR, HAMAND	82
Confocal RS (785 nm)	Blood-semen mixture	—	SVM-regression, SVM-DA (CV, EV)	80

[a] Each abbreviation can be referred to footnotes of Table 1.

[b] Wavelength used for excitation is shown in parentheses.

Figure Legends

Fig. 1 Deconvolution of Raman spectra of bloodstains during aging. (a) Spectral series observed during bloodstain aging at 24 °C. The spectra are shown before subtraction of fluorescence background. (b) Decomposed Raman spectral profiles of bloodstains by multivariate curve-resolution alternating least squares. (c) Index for bloodstain aging at 30 °C (magenta), 24 °C (purple), and 16 °C (blue). The index was defined as the ratio scores between the first and the fourth spectral components yielded via multivariate deconvolution technique based on the latent kinetics. The circles, crosses, and triangles represent data from three donors. The modeled ratios are shown as dashed lines with the standard deviations (shaded areas) for each temperature.

Fig. 2 SERS spectra of hair colored by blue semipermanent dye (BLU^{sp}) and re-dyed after by black semipermanent dye ($BLU^{sp}\rightarrow BLK^{sp}$). The spectra of hair colored by BLKsp and the dyes themselves (BLK^{sp} (D) and BLU^{sp} (D)) are also included. Reproduced from ref. 90 with permission from American Chemical Society, copyright 2019

Fig. 3 Plots of predicted versus actual PMI from the calibration and prediction datasets by PLSR models coupled with GA. Gray lines represent the perfect prediction. Reproduced from ref. 100 with permission from Elsevier, copyright 2017.

Fig. 4 FTIR-ATR spectra of human humerus: intact and burned at different temperatures (400–1000 °C). The inset represents a region of the asymmetric stretching carbonate bands $\nu_3(CO_2^{3-})$. The spectrum of reference calcium hydroxyapatite (HAp, SPM 2910b) is also included for comparison. Reproduced from ref. 107 with permission from Springer Nature, copyright 2018.

Fig. 5 A convolutional neural network (CNN) for identifying bacteria based on Raman spectra. (a) An SEM cross-section of the sample which bacterial cells are deposited on to gold-coated silica substrates. Scale bar is 1 μm. (b) Measurement schematic: Raman signal from single cells can be acquired from a diffraction-limited spot size by laser focusing. (c) A schematic of classification of low-signal Raman spectra via a one-dimensional residual network with 25 total convolutional layers into one of 30 isolates and antibiotic treatment groups. (d) Average Raman spectra of 30 bacterial isolates (bold lines) overlaid on representative noisy single spectra for each isolate. The spectra are colored according to antibiotic treatment group. Reproduced from ref. 124 with permission from Springer Nature, copyright 2019.