Discrimination of human and nonhuman blood using Raman spectroscopy with self-reference algorithm

Haiyi Bian
Peng Wang
Jun Wang
Huancai Yin
Yubing Tian
Pengli Bai
Xiaodong Wu
Ning Wang
Yuguo Tang
Jing Gao

Discrimination of human and nonhuman blood using Raman spectroscopy with self-reference algorithm

Haiyi Bian,[a] **Peng Wang,**[a] **Jun Wang,**[b] **Huancai Yin,**[b] **Yubing Tian,**[a] **Pengli Bai,**[b] **Xiaodong Wu,**[a] **Ning Wang,**[a,c] **Yuguo Tang,**[a,b] **and Jing Gao**[a,*]
[a]Chinese Academy of Sciences, Jiangsu Key Laboratory of Medical Optics, Suzhou Institute of Biomedical Engineering and Technology, Suzhou, China
[b]Chinese Academy of Sciences, CAS Key Lab of Biomedical Diagnostics, Suzhou Institute of Biomedical Engineering and Technology, Suzhou, China
[c]Shanghai University, School of Mechatronic Engineering and Automation, Shanghai, China

Abstract. We report a self-reference algorithm to discriminate human and nonhuman blood by calculating the ratios of identification Raman peaks to reference Raman peaks and choosing appropriate threshold values. The influence of using different reference peaks and identification peaks was analyzed in detail. The Raman peak at 1003 cm^{-1} was proved to be a stable reference peak to avoid the influencing factors, such as the incident laser intensity and the amount of sample. The Raman peak at 1341 cm^{-1} was found to be an efficient identification peak, which indicates that the difference between human and nonhuman blood results from the C–H bend in tryptophan. The comparison between self-reference algorithm and partial least square method was made. It was found that the self-reference algorithm not only obtained the discrimination results with the same accuracy, but also provided information on the difference of chemical composition. In addition, the performance of self-reference algorithm whose true positive rate is 100% is significant for customs inspection to avoid genetic disclosure and forensic science. © *2017 Society of Photo-Optical Instrumentation Engineers (SPIE)* [DOI: 10.1117/1.JBO.22.9.095006]

Keywords: blood species discrimination; Raman spectroscopy; self-reference algorithm.

Paper 170302RR received May 8, 2017; accepted for publication Sep. 5, 2017; published online Sep. 21, 2017.

1 Introduction

The identification of interspecies blood is important and necessary for forensic science, customs inspection, and veterinary purposes.[1–4] A variety of immunological methods, including enzyme-linked immunosorbent assay,[5] flow cytometry,[6] latex agglutination,[7] and chemiluminescence[8] have been developed for quantitative detection of low concentrations of biomarkers in human blood. The use of microcolumn gel immunoassay can further differentiate the red blood cell types.[9] Even molecular biology detection method can be used for the analysis of human race.[10] The analytical methods aforementioned are indirect methods requiring the aid of reagents. In the face of an unknown sample, the reagents may not work if it does not cover the unknown sample. Therefore, the information of the species to be analyzed should be provided first.

High performance liquid chromatography (HPLC) is used to identify human blood because of its advantage of resolution, sensitivity, and speed. Several HPLC methods have been developed, which could simultaneously identify a substance, such as blood and determine its origin of species. Fresh blood and blood stains from human and animal species were analyzed using this method by Inoue et al.[11] The sensitivity of this approach can be improved using fluorescence detection with ultraviolet excitation. Mass spectrometry (MS) was demonstrated as an effective tool in identifying species by Espinoza et al.[12] They use the properties of hemoglobin in blood samples as a means of differentiating one species from another. Several different North

American mammals were identified via the different molecular weights of their α- and β-chains. For HPLC and MS, the blood samples were distinguished according to their differences in hemoglobin, which were accomplished in a destructive manner. Moreover, these methods require laborious and complex preparation–separation and extraction procedures. Therefore, it is desirable to develop a rapid and noninvasive analytical method for distinguishing human and nonhuman blood samples.

Raman spectroscopy has been widely applied in forensic field as a simple, rapid, and direct analytical method, including identification of body fluids,[13] quantitation of saxitoxin,[14] detection of minute particles in blood,[15] and discrimination of bone samples.[16] Raman spectroscopy was first applied in identification of blood samples by De Wael et al.[15] They reported that there was no spectral difference between human, feline, and canine samples by analyzing the spectra of blood particles obtained by microspectrophotometry, Raman spectroscopy, or Fourier transform infrared spectroscopy. Elkins[17] also reported that the infrared spectra of human and sheep blood could not be visually distinguished. Principal component analysis method was demonstrated effective to discriminate human, cat, and dog blood samples by Virkler and Lednev.[18] Lednev et al.[19] used a partial least squares discriminant analysis classification model to distinguish human and nonhuman blood samples. Their model demonstrated 100% accuracy for differentiation between human and nonhuman blood samples. However, the physical meanings of each principal component could not be explained clearly, which means that the difference of chemical

---

compositions between human and nonhuman blood could not be found using these chemometric methods.

In this paper, we proposed a self-reference algorithm to discriminate human and nonhuman blood samples. The characterization of the ratio between two Raman peaks was used to confirm the blood species. The Raman spectra were analyzed by this method and compared with partial least squares (PLS) algorithm. The experimental results demonstrated that our method was as effective as PLS algorithm in discriminating human and nonhuman blood specimens. Moreover, according to the identification peak used in this paper, we found that the primary difference between the Raman spectra of human and the nonhuman blood was at 1341 cm^{-1}. This result indicates that the chemical composition of tryptophan (C–H bend) was one of the main factors inducing the small difference of Raman spectra in human and nonhuman blood.

2 Experimental Setup

2.1 Raman Microscope

A Renishaw inVia micro-Raman spectroscopy system (the United Kingdom) with the laser excitation wavelength of 785 nm was used to collect the Raman spectra of liquid blood specimens.[20,21] The incident light (785-nm diode laser) was focused into the sample and the backscattered Raman signal was collected with a Leica microscope with 50× objective. The laser power incident on the sample was about 0.6 mW, which is safe for the blood samples. The wave number was calibrated with silicon whose Raman spectral peak is at 520 cm^{-1}. The spectra were measured with an integration time of 10 s. The Raman spectra at the same sample position were accumulated seven times. The wave number range of the Raman spectra for blood samples was 300 to 1700 cm^{-1}. 785 nm was chosen as the excitation wavelength because more characteristic Raman peaks can be acquired.[20]

2.2 Blood Samples

Human blood samples were obtained from the volunteers in Suzhou and animal blood samples obtained from dog, rat, and rabbit were provided by Laboratory Animal Center of Soochow University. All experiments were performed in compliance with the law. All the blood samples were contained in vacuum blood collection tubes with anticoagulants (ethylenediaminetetraacetic acid dipotassium salt dihydrate) and then preserved in icebox immediately after they were acquired. The spectra were obtained between 24 and 72 h after the samples were acquired. For each liquid blood sample, ~20 μL was placed on an aluminum-covered microscope slide for detection. The aluminum cover was used to reduce the fluorescence background.

2.3 Principle of Self-Reference Algorithm

The intensity of Raman scattering signal is decided by the intensity of excitation light of specific wavelength, the property and quantity of the sample, and the sensitivity of spectrometer. This relationship has the following form[22]

$$I_{\text{Raman shift}} \propto K \times c \times (\partial \alpha / \partial Q)^2, \qquad (1)$$

where $I_{\text{Raman shift}}$ is the intensity of Raman spectrum; K is related to the intensity incident on the sample, the amount of the sample, and the parameter of the spectrometer. c is the molecular

concentration. α is the polarizability of the molecules. Q is the vibrational amplitude of the molecule. The parameters c, α, and Q are all related to the molecule measured. The parameter c' is defined as $c' = c \times (\partial \alpha / \partial Q)^2$. Then, Eq. (1) can be rewritten as

$$I_{\text{Raman shift}} \propto K \times c'. \qquad (2)$$

The traditional single intensity identification method can be used to identify different materials with different Raman peaks. But for materials with similar Raman spectra, single intensity identification method is incapable of action, because the Raman intensity is related to many factors such as the intensity incident on the sample and the amount of sample. To solve this problem, we introduce a method defined as the self-reference algorithm. In this method, two Raman shifts are used to discriminate the human and nonhuman blood. The one named as reference Raman peak is used to eliminate the influence of the interference factors. The other named as identification Raman peak is used to demonstrate the essential difference of human and nonhuman blood. The ratio of the identification and reference peaks' intensities is calculated as

$$\text{Ratio} = I_{\text{Raman shift 1}} / I_{\text{Raman shift 2}}. \qquad (3)$$

Using Eq. (1), we can get the following form of Ratio:

$$\text{Ratio} \propto c_1' / c_2'. \qquad (4)$$

Since the Raman spectra are obtained simultaneously, the influence factors in K are removed by performing the division of two Raman spectra. Consequently, the calculated Ratio is only related to the self-reference of molecules 1 and 2. This method would be effective if the difference between molecules 1 and 2 is bigger than the noise.

3 Results and Discussion

3.1 Necessity of Self-Reference Algorithm for Analysis of Raman Spectra of Human and Nonhuman Blood

The average Raman spectra of human and nonhuman's blood are shown in Fig. 1. Even though there is a difference between
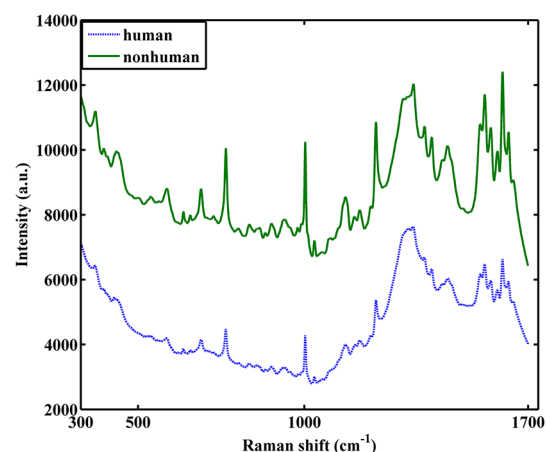

Fig. 1 The average Raman spectra of human and nonhuman blood.

the intensities of human and nonhuman blood, the Raman shifts of human and nonhuman blood are almost the same. Therefore, discrimination of human and nonhuman blood only according to the original Raman spectra is obviously not suitable.

The Raman spectra were preprocessed with the following procedures: (1) smooth the signal with Savitzky–Golay algorithm and (2) remove the baseline of the smoothed signal with adaptive iteratively reweighted penalized least squares (airPLS).[23,24] Figure 2 shows the normalized average intensity peaks of nonhuman and human blood at 676 cm^{-1} (pyrrole symmetric bending), 1003 cm^{-1} (aromatic ring breathing), and 1449 cm^{-1} (CH_2 and CH_3 bend) are approximately equal. The average intensity from 1310 cm^{-1} to 1342 cm^{-1} is obviously different. Therefore, we decided to make an evaluation on the discrimination of human and nonhuman blood using the intensity peaks of 1341 cm^{-1}. Since the Raman shift at 1341 cm^{-1} is corresponding to C–H bend in tryptophan,[25,26] we could speculate that human blood contain larger amount of tryptophan than nonhuman blood.

Figure 3 shows the intensity at 1341 cm^{-1} of each blood sample. Samples from 1 to 35 are human blood and the ones from 36 to 79 are nonhuman blood, respectively. The human blood's mean value of intensity at 1341 cm^{-1} is 1536.8 and the root mean square (RMS) is 70.5. The nonhuman blood's mean value of intensities at 1341 cm^{-1} is 1604.1 and the RMS is 282.8, which is relatively large. Therefore, even though the average intensities at 1341 cm^{-1} of human and nonhuman blood are different, the discrimination of human and nonhuman blood according to the single intensity peak is still difficult. The intensity peak of nonhuman blood varies large. This large RMS induces that the intensity peak at 1341 cm^{-1} alone is not suitable for the discrimination of blood specimens. As mentioned in Sec. 2.3, the Raman intensity is influenced by many factors such as the intensity incident on the sample and the amount of sample even for the same spectrometer. Therefore, we introduce the self-reference algorithm. Two intensity peaks are used: one is named reference intensity peak, such as 1449 cm^{-1} and the other is named identification intensity peak, such as 1341 cm^{-1}. By calculation the ratio of identification intensity peak to reference intensity peak, discrimination of human and nonhuman blood can be realized.
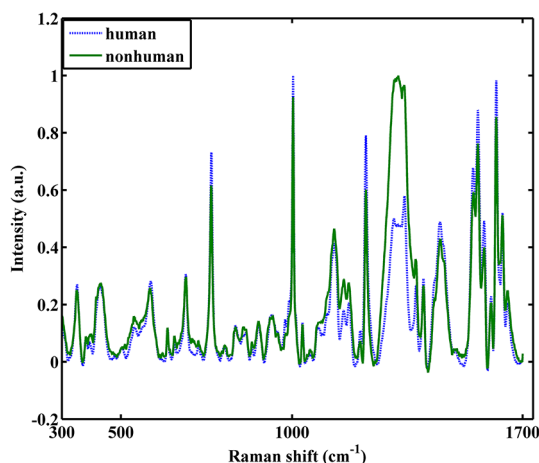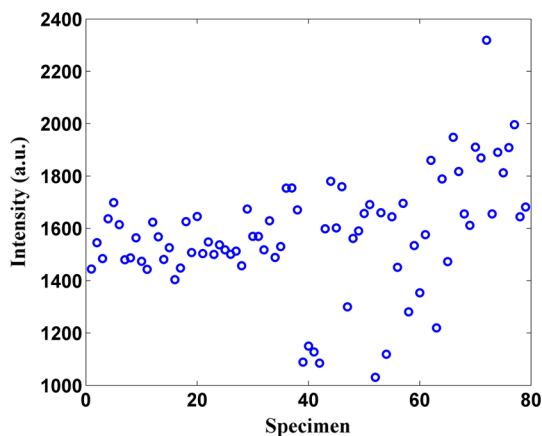

Fig. 3 The intensity of the spectra at 1341 cm^{-1}.

3.2 First Example of Blood Discrimination Using Self-Reference Algorithm and its Cross Validation

Figure 4 shows the results obtained by self-reference algorithm. The ratio of the intensity peak at 1341 cm^{-1} to that at 1132 cm^{-1} for human and nonhuman blood is listed. The mean value of human blood is 2.3118 and its RMS is 0.3457. The mean value of nonhuman blood is 1.1868 and its RMS is 0.1936. The threshold for discriminating human blood and nonhuman blood is set as the average of the minimal ratio of human blood and maximal ratio of nonhuman blood, 1.7 in this case as shown in Fig. 4. The value above this line is human blood and below this line is nonhuman blood.

Leave-one-out cross-validation method[27,28] was used to validate the self-reference algorithm. The cross-validation results were summarized in Table 1. This table shows the number of positive identified samples, along with a total percentage of true positive classifications for each individual species.

To avoid genetic disclosure and protect biodiversity, the performance of self-reference algorithm whose internal true positive rate (TPR) is 100% is significant for customs inspection. These results demonstrate that self-reference algorithm works well in discriminating human and nonhuman blood, which has real-world applicability in customs inspection, forensic
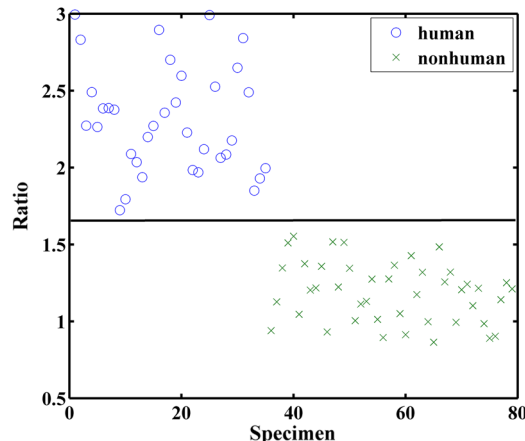

Fig. 2 The normalized average spectra of human and nonhuman blood.

Fig. 4 The ratio of the intensity peak at 1341 cm^{-1} to that at 1132 cm^{-1}.

Table 1 Internal cross-validation results for self-reference algorithm.

Actual class	Predicted samples	
	Nonhuman blood	Human blood
Dog	10	0
Rabbit	9	0
Rat	25	0
Human	0	35
True positive rate	100%	100%

Table 3 The identification rates when the reference peak is 1132 cm^{-1}.

Wave number (cm^{-1})	TPR (%)	FPR (%)	TNR (%)	FNR (%)	Threshold
1311	100	0	100	0	1.7
1316	100	0	100	0	1.7
1321	100	0	100	0	1.7
1326	100	0	100	0	1.7
1331	100	0	100	0	1.63
1336	100	0	100	0	1.7
1341	100	0	100	0	1.68

Note: TPR, human blood identified as human blood/total human blood. FPR, nonhuman blood identified as human blood/total human blood. TNR, nonhuman blood identified as nonhuman blood/total nonhuman blood. FNR, human blood identified as nonhuman blood/total nonhuman blood.

science, and veterinary setting. To demonstrate the ratio used to discriminate human and nonhuman blood is not coincidental, different peak choices were performed: (1) the same reference peak and all the identification peaks; (2) another reference peak and all the identification peaks.

3.3 Study of Influence of Using Other Identification Peaks by Self-Reference Algorithm

From Fig. 2, we can see several wave numbers can be used as the identification intensity peak. The range of identification wave number is from 1310 to 1342 cm^{-1}. To evaluate the self-reference algorithm's generality for all identification wave numbers, we calculated the ratio every five wave numbers from the wave number 1311 cm^{-1}. The results are given in Table 2. The parameter named with 1 represents human blood and 2 represents nonhuman blood. The threshold value can be set as 1.7 to discriminate human and nonhuman blood for all the other wave numbers expect for two wave numbers of 1331 and 1341 cm^{-1}, whose threshold is 1.63 and 1.68, respectively. The TPR, the false positive rate (FPR), the true negative rate (TNR), and the false negative rate (FNR) were then analyzed and listed in Table 3.

3.4 Study of Influence of Using Other Reference Peaks by Self-Reference Algorithm

To evaluate the influence of choosing the reference wave number, 1003 cm^{-1} was chosen as the reference wave number to

compare with the results using peak at 1132 cm^{-1} as reference. The same identification wave numbers were used. The results were shown in Table 4. Compared with Table 2, even though the threshold value is different, the discrimination of blood samples is still effective. The threshold value varies because of the different intensity peak at 1003 and 1132 cm^{-1}. The accuracy of the discriminations based on the data in Table 4 is 100%.

Figure 5 shows the results obtained by self-reference algorithm and normalization. The identification peaks of both are 1341 cm^{-1}. The reference peak in self-reference algorithm is 1003 cm^{-1}. In normalization process, the intensity at 1341 cm^{-1} is divided by the maximal value of the Raman spectrum. The results shown in Fig. 5 indicate that the maximal value for each blood sample is not at the same wavenumber. The data overlapping of self-reference algorithm and normalization means the maximal value is just at the reference Raman peak 1003 cm^{-1}. The change of denominator in normalization is equivalent to the change of reference peak. Therefore, the result

Table 2 The results obtained by the ratio of identification peak at different wave numbers to reference peak at 1132 cm^{-1}.

Wave number (cm^{-1})	Mean ratio 1	RMS 1	Mean ratio 2	RMS 2	Minimum ratio 1	Maximal ratio 2	Threshold
1311	2.2694	0.3207	1.2058	0.1866	1.8560	1.5768	1.7
1316	2.2787	0.3486	1.1804	0.2052	1.7290	1.6520	1.7
1321	2.3081	0.3342	1.1998	0.1983	1.7606	1.6704	1.7
1326	2.2871	0.3319	1.1772	0.1949	1.8424	1.6496	1.7
1331	2.2100	0.3338	1.1878	0.1796	1.6307	1.6256	1.63
1336	2.2198	0.3211	1.2567	0.1833	1.7213	1.6833	1.7
1341	2.2194	0.2665	1.3904	0.1572	1.6848	1.6744	1.68

Table 4 The results obtained by the ratio of reference peak at 1003 cm^{-1} to identification peak at different wave numbers.

| Wave number (cm−1)	Mean ratio 1	RMS 1	Mean ratio 2	RMS 2	Minimum ratio 1	Maximal ratio 2	Threshold
1311	1.0939	0.1548	0.4988	0.0987	0.8313	0.7901	0.8
1316	1.0976	0.1672	0.4886	0.1063	0.8137	0.7755	0.8
1321	1.1116	0.1601	0.4961	0.1021	0.8905	0.8050	0.85
1326	1.1029	0.1656	0.4872	0.1012	0.8587	0.7836	0.8
1331	1.0719	0.1595	0.4918	0.0950	0.8217	0.7531	0.8
1336	1.0448	0.1481	0.5007	0.0917	0.8346	0.7942	0.8
1341	1.0613	0.1300	0.5582	0.0972	0.8644	0.8138	0.84
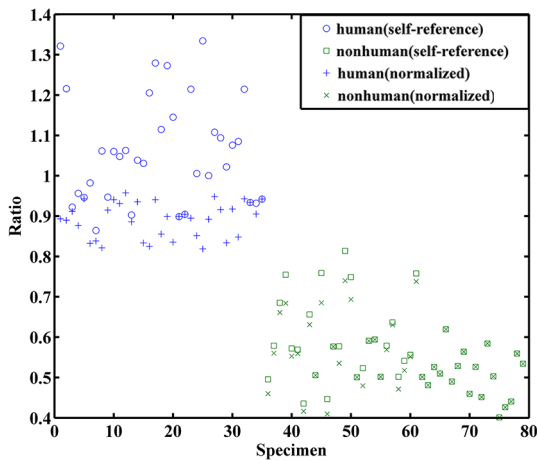
Fig. 5 The discrimination of human and nonhuman blood with self-reference algorithm and normalization.

obtained by normalization is lack of stability theoretically, even though the discrimination could be realized occasionally. The peak intensity at 1341 cm^{-1} in normalized Raman spectrum is not a suitable parameter that can be used to discriminate human and nonhuman blood.

3.5 Validation of the Self-Reference Algorithm by Blind Test

In total, 10 human blood samples and 16 animal blood samples (four dog blood samples, four rabbit blood samples, four rat blood samples, and four horse blood samples) were used to validate the self-reference algorithm. 1003 cm^{-1} was chosen as reference peak. 1341 cm^{-1} was chosen as the identification peak. The results were displayed in Table 5 and Fig. 6. Therefore, the self-reference algorithm is proved to be accurate and stable by blind test.

Table 5 The validation of the self-reference algorithm.

| Spectrum	Predicted class	Actual specimen	Spectrum	Predicted class	Actual specimen
Unknown 1	Nonhuman	Rabbit	Unknown 14	Nonhuman	Rat
Unknown 2	Nonhuman	Dog	Unknown 15	Human	Human
Unknown 3	Human	Human	Unknown 16	Human	Human
Unknown 4	Human	Human	Unknown 17	Nonhuman	Horse
Unknown 5	Nonhuman	Rat	Unknown 18	Human	Human
Unknown 6	Nonhuman	Rat	Unknown 19	Nonhuman	Dog
Unknown 7	Nonhuman	Dog	Unknown 20	Human	Human
Unknown 8	Nonhuman	Horse	Unknown 21	Nonhuman	Horse
Unknown 9	Human	Human	Unknown 22	Nonhuman	Rat
Unknown 10	Nonhuman	Rabbit	Unknown 23	Human	Human
Unknown 11	Nonhuman	Dog	Unknown 24	Nonhuman	Rabbit
Unknown 12	Nonhuman	Horse	Unknown 25	Nonhuman	Rabbit
Unknown 13	Human	Human	Unknown 26	Human	Human
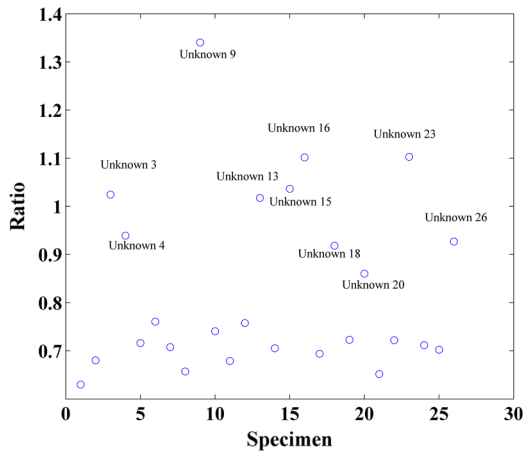
Fig. 6 The validation of self-reference algorithm with 10 human blood samples and 16 animal blood samples.

3.6 Comparison Between Self-Reference Algorithm and Partial Least Square

The PLS[29,30] discrimination analysis was applied and the results were shown in Fig. 7. Figure 7(a) shows the 2-D score distributions of PC1 and PC2. Figure 7(b) shows the 3-D score distributions of PC1, PC2, and PC3. The results indicate that the discrimination of human and nonhuman blood could be realized with two principle components.
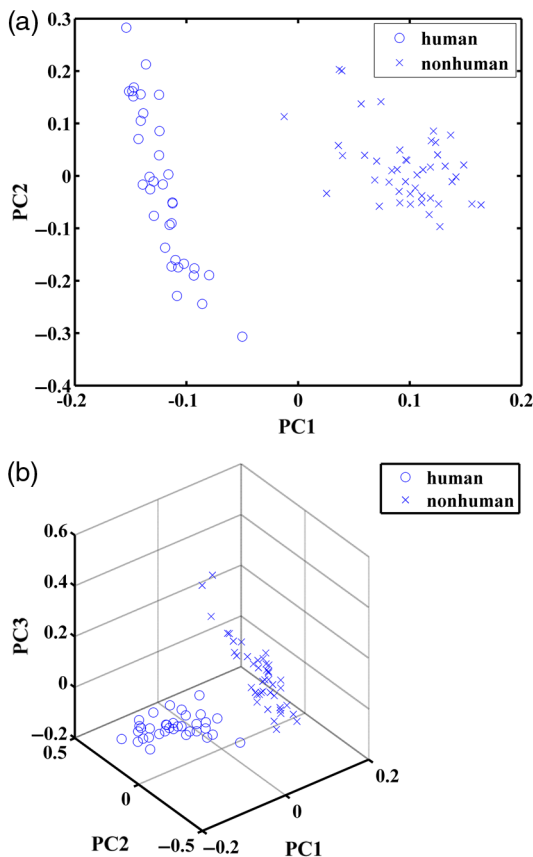

Fig. 7 The score plots of PLS model (a) 2-D score plot with two principal components and (b) 3-D score plot with three principal components.

Table 6 Internal cross-validation results for self-reference algorithm.

Actual class	Predicted samples	
	Nonhuman blood	Human blood
Dog	10	0
Rabbit	9	0
Rat	25	0
Human	0	35
True positive rate	100%	100%

The performance of PLS model was also evaluated with leave-one-out cross-validation method. The results were displayed in Table 6. Both self-reference algorithm and PLS can achieve satisfactory results. However, the model construction by self-reference algorithm is much easier. Moreover, the difference of chemical composition between human blood and animal blood can be well analyzed by self-reference algorithm. We find the C–H bend in tryptophan in blood is one of the main factors that induce the difference of the Raman spectra between human and nonhuman blood. For forensic science and veterinary purposes, we want to know not only the blood species, but also the relative concentration of some chemical components. Self-reference algorithm can provide the relative concentration information. In comparison, for PLS model, the information of the difference in chemical composition and the relative concentration cannot be achieved, although the results are enough for discriminating human and nonhuman blood.

4 Conclusion

In this study, self-reference algorithm was proved to enable an effective discrimination of liquid human and nonhuman blood. The reference Raman peak at 1003 cm^{-1} and identification Raman peak at 1341 cm^{-1} were chosen for more stable and effective performance, in which term the threshold value was determined to be 0.84. The cross validation and blind test demonstrate the accuracy and the TPR of discrimination are both 100%, which indicates potential real-world applications in forensic science and custom inspection. In addition, the selection of Raman peak at 1341 cm^{-1} (corresponding to the C–H bend in tryptophan) as the optimized identification peak indicates there is a significant difference of tryptophan contents in human blood and that in nonhuman blood. In conclusion, the self-reference algorithm could not only discriminate human and nonhuman blood with highly accurate and stable results, but also point out their significant difference of chemical composition.

Disclosures

No conflicts or competing interests, financial or otherwise, apply to any of the authors.

References

1. S. C. Renner et al., "Import and export of biological samples from tropical countries-considerations and guidelines for research teams," *Org. Divers. Evol.* **12**(1), 81–98 (2012).

2. R. Rankin et al., "CpG motif identification for veterinary and laboratory species demonstrates that sequence recognition is highly conserved," *Antisense Nucleic Acid Drug Dev.* **11**(5), 333–340 (2001).

3. E. N. Barker et al., "A novel haemoplasma species identified in archived primate blood smears," *Vet. Microbiol.* **149**(3), 478–481 (2011).

4. S. Unajak et al., "Full length research paper identification of species (meat and blood samples) using nested-PCR analysis of mitochondrial DNA," *Afr. J. Biotechnol.* **10**(29), 5670 (2011).

5. W. Xia et al., "A specific enzyme-linked immunosorbent assay for measuring β-amyloid protein oligomers in human plasma and brain tissue of patients with Alzheimer disease," *Arch. Neurol.* **66**(2), 190–199 (2009).

6. A. F. Orozco and D. E. Lewis, "Flow cytometric analysis of circulating microparticles in plasma," *Cytometry* **77**(6), 502–514 (2010).

7. I. Aly et al., "Novel nanomagnetic beads based-latex agglutination assay for rapid diagnosis of human schistosomiasis haematobium," *Int. J. Med. Health Biomed. Pharm. Eng.* **7**(12), 642–647 (2013).

8. N. He et al., "Chemiluminescence analysis for HBV-DNA hybridization detection with magnetic nanoparticles based DNA extraction from positive whole blood samples," *J. Biomed. Nanotechnol.* **9**(2), 267–273 (2013).

9. Y. Liu et al., "Application of microcolum gel test for serodiagnosis of hemolytic disease of newborn," *Chin. J. Clin. Lab. Sci.* **23**(1), 43–45 (2005).

10. E. Birney et al., "Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project," *Nature* **447**(7146), 799–816 (2007).

11. H. Inoue et al., "Species identification of blood and bloodstains by high-performance liquid chromatography," *Int. J. Leg. Med.* **104**(1), 9–12 (1990).

12. E. O. Espinoza et al., "Electrospray ionization mass spectrometric analysis of blood for differentiation of species," *Anal. Biochem.* **268**(2), 252–261 (1999).

13. F. Zapata, M. Á. F. de la Ossa, and C. García-Ruiz, "Emerging spectrometric techniques for the forensic analysis of body fluids," *TrAC, Trends Anal. Chem.* **64**, 53–63 (2015).

14. W. F. Pearman et al., "Characterization of the Ag mediated surface-enhanced Raman spectroscopy of saxitoxin," *Appl. Spectrosc.* **62**(7), 727–732 (2008).

15. K. De Wael et al., "In search of blood—detection of minute particles using spectroscopic methods," *Forensic Sci. Int.* **180**(1), 37–42 (2008).

16. G. McLaughlin and I. K. Lednev, "Spectroscopic discrimination of bone samples from various species," *Am. J. Anal. Chem.* **3**(2), 161–167 (2012).

17. K. M. Elkins, "Rapid presumptive fingerprinting of body fluids and materials by ATR FT-IR spectroscopy," *J. Forensic Sci.* **56**(6), 1580–1587 (2011).

18. K. Virkler and I. K. Lednev, "Blood species identification for forensic purposes using Raman spectroscopy combined with advanced statistical analysis," *Anal. Chem.* **81**(18), 7773–7777 (2009).

19. G. Mc Laughlin, K. C. Doty, and I. K. Lednev, "Raman spectroscopy of blood for species identification," *Anal. Chem.* **86**(23), 11628–11633 (2014).

20. P. Bai et al., "Discrimination of human and nonhuman blood by raman spectroscopy and partial least squares discriminant analysis," *Anal. Lett.* **50**(2), 379–388 (2017).

21. P. Bai et al., "Rapid qualitative identification method of species of blood based on PCA with Raman spectroscopy," *J. Light Scattering* **28**(2), 163–167 (2016).

22. P. Larkin, *Infrared and Raman Spectroscopy: Principles and Spectral Interpretation*, Elsevier, San Diego (2011).

23. Z. M. Zhang, S. Chen, and Y. Z. Liang, "Baseline correction using adaptive iteratively reweighted penalized least squares," *Analyst* **135**(5), 1138–1146 (2010).

24. S. J. Baek et al., "Baseline correction using asymmetrically reweighted penalized least squares smoothing," *Analyst* **140**(1), 250–257 (2015).

25. M. Casella et al., "Raman and SERS recognition of β-carotene and haemoglobin fingerprints in human whole blood," *Spectrochim. Acta, Part A* **79**(5), 915–919 (2011).

26. K. Virkler and I. K. Lednev, "Raman spectroscopy offers great potential for the nondestructive confirmatory identification of body fluids," *Forensic Sci. Int.* **181**(1), e1–e5 (2008).

27. K. Jayawardana et al., "Identification, review, and systematic cross-validation of microRNA prognostic signatures in metastatic melanoma," *J. Invest. Dermatol.* **136**(1), 245–254 (2016).

28. J. Zhang and S. Wang, "A fast leave-one-out cross-validation for SVM-like family," *Neural Comput. Appl.* **27**(6), 1717–1730 (2016).

29. D. C. Maltaş et al., "Rapid classification of pharmaceutical ingredients with Raman spectroscopy using compressive detection strategy with PLS-DA multivariate filters," *J. Pharm. Biomed. Anal.* **80**, 63–68 (2013).

30. M. R. Almeida et al., "Classification of Amazonian rosewood essential oil by Raman spectroscopy and PLS-DA with reliability estimation," *Talanta* **117**, 305–311 (2013).

Haiyi Bian is an assistant researcher at Suzhou Institute of Biomedical Engineering and Technology. He received his PhD in optical engineering from Nanjing University of Science and Technology in 2016. His research interests include Raman spectroscopy and optical coherence tomography.

Biographies for the other authors are not available.