Chemometrics and Intelligent Laboratory Systems 182 (2018) 172–179

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

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

Confocal Raman spectroscopy and multivariate data analysis for evaluation of spermatozoa with normal and abnormal morphology. A feasibility study

R.V. Nazarenko [a,*], A.V. Irzhak [b], A.L. Pomerantsev [c,d], O. Ye. Rodionova [c,d]

[a] EKO Infertility Treatment Centre, Moscow, 115191, Russia
[b] Institute of Microelectronics Technology and High Purity Materials RAS, Chernogolovka, 142432, Russia
[c] Semenov Institute of Chemical Physics RAS, Moscow, 119991, Russia
[d] Institute of Natural and Technical Systems RAS, Sochi, 354024, Russia

ARTICLE INFO

Keywords:
Confocal Raman spectroscopy
Sperm morphology
DNA fragmentation
Data driven soft independent modeling of class analogies

ABSTRACT

This paper investigates a feasibility of using confocal Raman spectroscopy (CRS) and multivariate analysis for classification of sperm cells. The spectral based classification is compared with the morphological analysis, which is the main criterion for sperm selection in intracytoplasmic sperm injection procedure. The spectral analysis is conducted using the data driven soft independent modeling of class analogies method. The supervised classification reveals numerous outliers that pass from the 'normal' class to the 'abnormal' class, and vice versa. The ultimate result shows that the initial morphological discrimination overlaps with the spectral classification only partly. It is shown that CRS provides additional information regarding the nuclear DNA stability and helps to reveal spermatozoa with fragmented and defective DNA. This can be a promising direction for future evaluation of spectra from live, unfixed cells.

1. Introduction

Over the past 25 years, Raman spectroscopy has proven to be an effective and reliable method for characterization of the intermolecular bonds [1]. This method has perspectives in the biological and medical applications. The confocal Raman spectroscopy (CRS) does not need any special conditions or sample preparation to acquire spectra. It can be used for the structural and compositional analysis of a sample before application of some destructive analytical methods [2]. The amount of the CRS applications for the investigation of tissues [3], cells [4], subcellular organelles and intercellular metabolic processes grows subsequently [5]. A particular interest deserves the CRS usage in the areas of oncology [6,7], cardiology [8], and reproductive medicine [9]. Despite the fact that the results of these studies are still far from the clinical practice implementation, the available data allow us to conclude that this method has a great potential. In biology and reproductive medicine the researchers are interested in using CRS for evaluating quality of gametes. Confocal microscopes allow Raman spectroscopy to be performed with very high lateral spatial resolution and minimal depth of field (i.e., below 1 μm), thereby permitting the identification of molecules in organelles [10]. The CRS approach is especially valuable in the analysis and selection of sperm for ICSI (intracytoplasmic sperm injection) procedure, because spermatozoon can be explored without harming.

The studies of sperm using CRS are few so far. The sperm is a specialized, highly differentiated cell bearing the haploid number of chromosomes. The sperm main function is to deliver genetic material into the oocyte. To fulfill this task the sperm undergoes a number of significant structural and functional changes ensuring tighter condensation of nuclear DNA. However, those changes have their negative features. With synthetic and metabolic processes coming to a halt in the cell, DNA is not transcribed and ready mRNA transcripts are absent due to the extremely small amount of cytoplasm. Thus, the mature sperm cell with 90–95% of its histones replaced with protamines is unable to repair the DNA damage. The DNA condensation provides a sufficiently reliable degree of protection, but if the process of histone replacement with protamines is not completed, the vulnerable DNA sites are exposed to damaging agents. Oxidative stress can lead to the single- and double-stranded DNA breaks, and the lack of DNA repair mechanisms in the sperm prevents DNA restoration. This sperm can have good morphological features and motility characteristics; nevertheless, it gives rise to an embryo of a poor quality and low implantation potential. The corresponding embryos often stop in development [11]. Sperm DNA damage can be partially restored by the oocyte, but once the damage reaches a critical mass the oocyte DNA repair mechanism is insufficient. This fact

* Corresponding author. EKO Infertility Treatment Centre, Moscow, 115191, Holodilny per.2, Bld.2, Russia.
E-mail address: zdan@list.ru (R.V. Nazarenko).

https://doi.org/10.1016/j.chemolab.2018.10.002
Received 18 July 2018; Received in revised form 3 October 2018; Accepted 5 October 2018
Available online 6 October 2018
0169-7439/

leads to the growth of interest to the study of sperm nuclear DNA and assessment of its functional state. Since standard methods of ejaculate evaluation using light microscopy cannot provide relevant data, additional sperm evaluation methods have recently got widespread. In particular, this applies to methods used for studying sperm DNA fragmentation [12–14]. Unfortunately, beyond determining the percentage of sperm with fragmented DNA in the ejaculate, these methods provide little information relevant for use in assisted reproductive technologies. In ICSI practice, we need a method that can test a single sperm cell without damaging it, for the integrity of sperm nuclear DNA and its functional state. We think that the CRS approach is quite promising in solving the above mentioned tasks.

The first study of Raman spectra (RS) of salmon spermatozoa was conducted in 1986 by Kubasek et al. [15]. In 2009 Huser and co-authors [16] compared RS of sperms with normal and abnormal morphology and concluded that the sperm morphology does not always correlate with a proper nuclear DNA packaging. Meister et al. [17] investigated the spectra of subcellular organelles and the impact of UV irradiation upon them. Further work on the use of Raman spectroscopy in visualization of the sperm nuclear DNA damage was continued by Mallidis et al. [18], who used UV irradiation for damaging the structure of sperm nuclear DNA. Sánchez et al. [10] induced oxidative damage by Fenton's reaction with hydrogen peroxide. Spectra obtained from the sperms damaged in such a way fully correspond to those obtained in study by Mallidis et al. [18]. It should be emphasized that until recently all studies were conducted on fixed cells. The article by Li et al. [19], who used CRS on unfixed live buffalo spermatozoa, was published only in 2016. To keep the sperms in focus, the so-called optical (laser) tweezers were used. It was shown that spectra acquired from the motile spermatozoa were similar to those obtained from the fixed sperm cells. This method seems to be very promising for the introduction of CRS into clinical practice of in vitro fertilization centers.

The aim of this study is to assess feasibility of the CRS and multivariate analysis methods in studying the sperm nuclear DNA. Additional goal is to perform a comparative analysis of the RS obtained from the morphologically normal and abnormal spermatozoa.

2. Materials and methods

2.1. Samples

We performed analysis of semen taken from 21 healthy donors. The period of sexual abstinence was 2–7 days. In each case, a standard semiological study of the ejaculate was carried out using a light microscopy. The results were interpreted according to the 2010 World Health Organization guidelines [20] and corresponded to normozoospermia (≥4% of normal spermatozoa according to Kruger's strict criteria).

2.2. Compliance with ethical standards

This study was approved by the Independent Local Committee for Ethics of the Regional Research Institute of Obstetrics and Gynecology, Moscow, Russia (Registration number JRB № 06004245), and written informed consent was obtained from all participants. The authors declare that they have no conflict of interest.

2.3. Sample preparation for spectra acquisition

The ejaculate was subjected to a double density gradient centrifugation (FertiPro, Breenem, Belgium) for 20 min at 415 g. After removal of the supernatant, the pellet was resuspended with Phosphate Buffered Saline (PBS, Sigma-Aldrich, Saint Louis MO, USA) at 37 °C and re-centrifuged for 10 min at 415 g. Thereafter 10 μl of the sperm suspension was placed onto the aluminum foil. The preparation was fixed in alcohol (ethanol 70%) and air-dried.

2.4. Reference method

Using standard light microscopy, we assembled two sets of samples (each sample corresponds to 1 sperm cell): morphologically normal sperm cells (group N, 125 samples), and morphologically abnormal sperm cells (group A, 36 samples). We selected only the sperm cells with clearly visible tail, nucleus and acrosome.

2.5. Confocal Raman spectroscopy

The SENTERRA confocal Raman microscope, Bruker Optics, Germany was used. The spectral analysis of the sperm nucleus was carried out at the 532 nm laser excitation wavelength and the 10 mW power in the range of 280–1730 cm^{-1} at the resolution of 3–5 cm^{-1}. Each sample was measured for 20 s integration time averaged over 3 scans. The spectra acquisition mode was chosen in a way that did not damage the sperm cell according to Eidengeiser et al. [21], who confirmed that laser power less than 15 mW did not induce the photo damage of the sperm cell membrane. The more detailed information about the influence of laser light on sperm cell can be found in the article [21].

Fig. 1 shows a typical image of morphologically normal spermatozoon with designation of the regions of analysis (a) and the RS obtained from the different regions (b). 1 – nucleus, 2 – neck, 3 – acrosome. Spectra acquired from different regions have specific features. This demonstrates that the laser beam has a local impact, and, therefore, confirms suitability of the CRS method in our studies. In general, it can be seen that the spectra obtained at acrosome and neck have less intensities in comparison with the nucleus one. A high band round 749 cm^{-1}, which appears in the spectra of the neck region, refers to mitochondrial DNA according to Tesarik et al. [13]. In our study, the spectra were collected from nucleus region only.

2.6. Preliminary spectra analysis

Each spectrum is analyzed in the range from 680 cm^{-1} to 1700 cm^{-1} after polynomial baseline fitting (4th order), smoothing using the moving average filter (21 points). We also perform the standard normal variate correction for all spectra. Using the preprocessed spectra, we compile the data matrix **X** (161 x 510), where each row contains the Raman readings in the interval from 680 to 1700 cm^{-1} at the 2 cm^{-1} increment; the number of rows corresponds to the number of spectra.

2.7. Chemometric methods

Principal component analysis (PCA) [22] is used for the data dimensionality reduction, for separation of the significant information from noise, and for elucidation of hidden relationships between the samples.

The DD-SIMCA (Data driven - Soft independent modeling of class analogies) method is used for the outlier detection and the final classification. DD-SIMCA is a one-class classifier (OCC). Such type of methods is also called the class modeling approach [23]. For application of OCC, a unique target class should be defined by the properties of its representative members. The set of representative target samples is called the training set. The DD-SIMCA method consists of three steps. At the first step, the PCA decomposition is applied to the $(I \times J)$ training matrix **X**.

$$X = TP^t + E \quad (1)$$

where $\mathbf{T} = \{t_{ia}\}$ is the $(I \times A)$ scores matrix; $\mathbf{P} = \{p_{ja}\}$ is the $(J \times A)$ loadings matrix; $\mathbf{E} = \{e_{ij}\}$ is the $(I \times J)$ matrix of residuals; and A is the number of principal components (PCs). At the second step, for each object $i = 1, \ldots, I$ from the training set, two distances are calculated. They are the score distance h_i, and the orthogonal distance v_i:

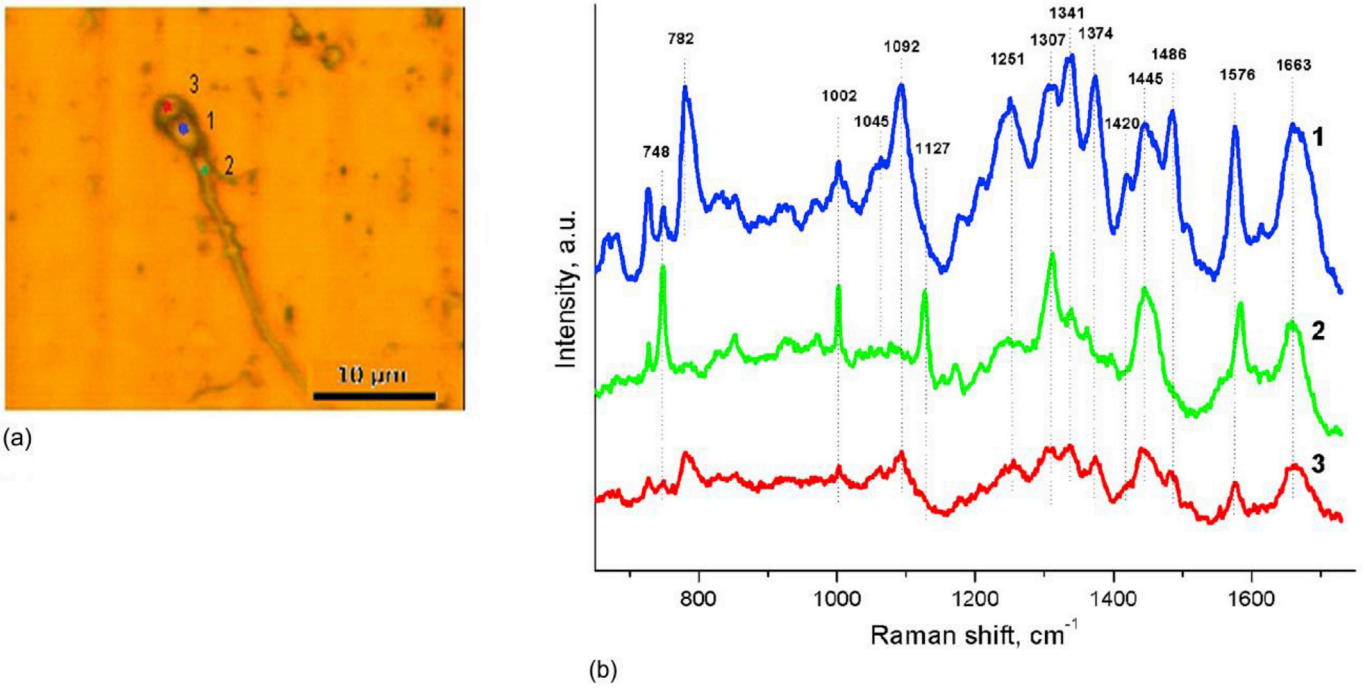

Fig. 1. (a) Microimage of a morphologically normal spermatozoon with the regions of CRS analysis. (b) Exemplars of the RS obtained at the corresponding regions. 1-nucleus; 2 –neck; 3 – acrosome.

$$h_i = \mathbf{t}_i^t(\mathbf{T}^t\mathbf{T})^{-1}\mathbf{t}_i = \sum_{a=1}^{A}\frac{t_{ia}^2}{\lambda_a},\ v_i = \sum_{j=1}^{J} e_{ij}^2 \tag{2}$$

In this formula elements $\lambda_a = \sum_{i=1}^{I} t_{ia}^2$ are the eigenvalues of matrix $\mathbf{X}^t\mathbf{X}$ ranked in descending order. In the process control area [24] these distances are often denoted as Q (for v) and T^2 for (h). The total distance for the sample is calculated as:

$$c = N_h\frac{h}{h_0} + N_v\frac{v}{v_0} \propto \chi^2(N_h + N_v) \tag{3}$$

where parameters v_0 and h_0 are the scaling factors, N_h and N_v are the numbers of the degrees of freedom (DoF). These parameters are unknown *a priori*, and they are estimated using a data driven approach with application of a classical (method of moments) and/or a robust (median and interquartile) statistics. Details are presented in the works by Pomerantsev et al. [25,26]. This feature of DD-SIMCA provides the possibility to apply the method for detection of outliers. In case the abovementioned parameters of the χ^2-distribution coincide being calculated by both techniques, we can conclude that the training set does not contain outliers.

At the third step, the acceptance area or a threshold (the green curve in Fig. 2) for the target class is defined. Given the type I error α, the acceptance area is determined as

$$c \le c_{\text{crit}}(\alpha), \tag{4}$$

where

$$c_{\text{crit}} = \chi^{-2}(1-\alpha,\ N_h + N_v) \tag{5}$$

is the (1–α) quantile of the chi-squared distribution with $N_v + N_h$ DoF. This cut-off level forms the border for the acceptance area, which covers (1 - α) 100% of all population. If an object belongs to this area, it is assessed as a regular one (the green dots in Fig. 2).

The second border is determined as the outlier cut-off level (the red curve in Fig. 2) constructed for a given γ-value. This value specifies the probability that at least one regular object from the data set will be

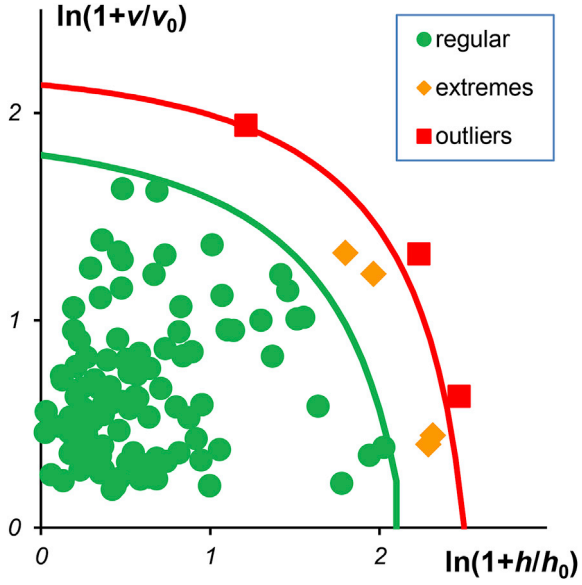

Fig. 2. Successive partitioning of the group N into NN and NA classes, α = 0.01, γ = 0.05. Samples marked with red squares are considered as outliers and are removed from the training set. The yellow diamonds represent the extreme objects. The green dots are the regular samples. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

erroneously considered as an outlier. Objects located beyond the outlier threshold are considered as outliers (the red squares in Fig. 2). The area located between the acceptance and the outlier areas is the extreme objects area. Extreme objects or 'extremes' (the yellow diamonds in Fig. 2) are always present in the data and they should not be confused with outliers [26]. Elimination of extremes is undesirable, since this can distort the analysis results.

In this work, the DD-SIMCA method is used for two purposes. First

purpose is the search for outliers in the training set. The second purpose is the OCC model building [27].

Calculations are performed using Chemometrics Add-In for the Microsoft Excel [28] and Matlab GUI tool for DD-SIMCA [29], which is freely available from GitHub: https://github.com/yzontov/dd-simca.git.

3. Results

3.1. Analysis of morphologically normal spermatozoa (group N)

To build an OCC model we have to be sure that the target class (normal samples) does not contain outliers. Suspicious samples are revealed by application of the DD-SIMCA procedure to the spectra of the morphologically normal sperm (group N). For this purpose, we develop the SIMCA-models applying two methods of estimation: classical and robust. The number of PCs is selected based on the principle of parsimony [30], i.e. the minimal A that provides a desired solution corresponding to the given α value. In our case, two values of A are in focus, 2 or 3. When outliers are detected (Fig. 2, red squares), these samples are excluded from the target class, and the DD-SIMCA procedure is repeated for the reduced training set. The process of the outlier elimination is over when the results obtained using the classical and robust methods become similar. It took 5 steps of the DD-SIMCA application to clean the training set from the outliers.

In the result, the group N (125 samples) is divided into two classes: class NN (102 samples) and class NA (23 samples). Class NN is later used as the target class. It comprises the samples recognized as the normal both by morphology and by the results of spectral data analysis. Class NA comprises the samples considered normal according to the morphology analysis, but treated as extraneous samples relative to the target class NN. The final acceptance area is designed at 3 PCs using samples of the target class NN (102 samples). In Fig. 3 acceptance area for $\alpha = 0.01$ is represented by the solid green curve; the training samples of class NN are shown as the closed blue dots, the class NA samples are represented by the red dots with the blue border, class AN is given by the blue squares with the red border. In general, the markers border color stands for the original classification, while the interior color symbolizes the ultimate decision. In all cases, the blue color means normality, and the red color stands for abnormality.

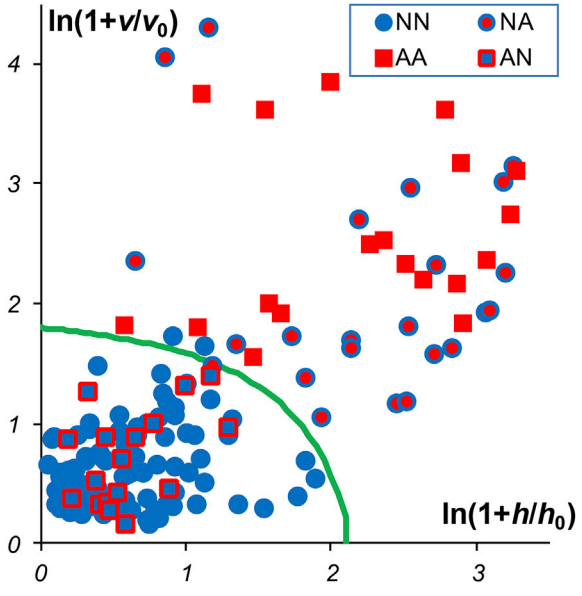

Fig. 3. Distribution of all samples into classes with respect to NN (the blue dots) class, $\alpha = 0.01$. The green curve delineates the acceptance area. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

3.2. Analysis of morphologically abnormal spermatozoa (group A)

We assume that morphological analysis of abnormal spermatozoa samples (group A) may, in some cases, lead to erroneous results, in the same way as for group N samples. The spectra of the second group are analyzed using the model developed for the NN target class. Initially group A contains 36 samples (spectra) of the morphologically abnormal spermatozoa. Application of the DD-SIMCA model results in a split of group A into two classes: class AA (19 samples), morphologically and spectrally anomalous, and class AN (17 samples), morphologically abnormal, but spectrally normal. In Fig. 3 the AN samples (the blue squares with red border) are located inside the acceptance area. The NA samples (the red dots with blue border), comprising the outliers of group N, are located outside the acceptance area, and they are uniformly distributed among the class AA samples (the red squares).

Certainly, we can use the AA samples as the target class and consider all other samples as the alternative classes. The DD-SIMCA model for class AA accepts all alternative samples as the members of the target class AA. Therefore, in respect to class AA all samples are regular and belong to class AA. It is concluded that class NN is a uniform class, which, in fact, includes class AN. Therefore, all the samples in these classes are classified as the spectrally normal.

The application of PCA to the joint data set shows (Fig. 4) that classes NN and AN form a compact group in the centre of the PC score space, whereas the AA and NA samples are located at the periphery. Classes AA and NA are the heterogeneous groups with a wide dispersion of data. This fact allows us to classify such spectra as spectrally abnormal.

The comparison of the average spectra of the spectrally normal (NN + AN) and abnormal (NA + AA) samples is shown in Fig. 5. The average spectrum of the normal samples (the blue curve) and the 95% tolerance corridor (vertical blue bands) are calculated for each wavelength on the entire NN + AN set. The red curve represents the average spectrum computed for the AA + NA classes. Vertical red bands show the corresponding 95% tolerance corridor for this set.

Despite the fact that the abnormal spectrum (Fig. 5) has a high band at 1045 cm^{-1}, the tolerance corridor of this spectrum completely includes the normal spectrum with its corridor. We can see that spectral abnormality is characterized by a much greater spread, which makes it impossible to distinguish visually the normal spectrum from the abnormal one.

Let us analyze the PCA model for the NN + AN set, interpreting the AA + NA data as new samples. Fig. 6 represents the bi-plot [31] for PC2 vs. PC1. In this plot the PCA scores for the first two components are shown together with the corresponding PCA-loadings. This plot helps to interpret the sample properties and the variable relationships simultaneously. In Fig. 6, the blue dots (PCA scores) represent spectrally normal samples, while the red squares show spectrally abnormal samples. Samples located close to each other are similar. The opposite samples are different. The green crosses (PCA loadings) represent variables (Raman bands). Loadings indicated by numbers present the characteristic bands. The samples located near these extreme crosses have the distinctive Raman band. We can conclude that PC1 is mainly determined by the bands at 782 cm^{-1} and 1663 cm^{-1}. PC2 is formed by bands at 980 cm^{-1} and 1374 cm^{-1}. Two very important bands round 1045 cm^{-1} and 1451 cm^{-1} contribute to both first PCs. The spectrally abnormal AA + NA samples (the red squares) are located mainly in peripheral areas of the plot, while the spectrally normal samples NN + AN (the blue dots) occupy the centre. Different loadings values, i.e. bands, are responsible for specific locations of various sub-groups inside the AA + NA set. So in Fig. 6, it can be seen that spectrally abnormal samples (NA1, NA2, AA1, AA2) have Raman specific features (bands) different from uniform class NN + AN.

Concluding the data analysis we can note that they represent a real-world example for the data set of a special type called 'vatrushka'.[1]

[1] Vatrushka is a pastry formed as a ring of dough with quark in the middle.

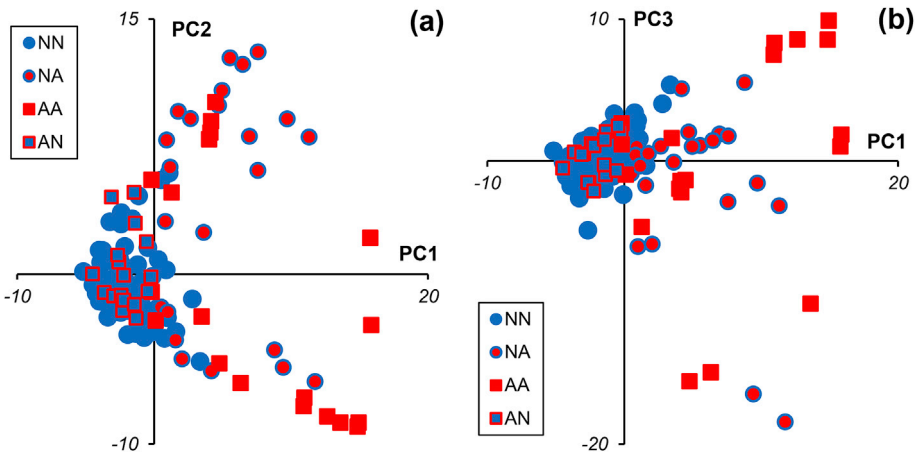

Fig. 4. PCA on the entire data set. The score plots PC1-PC2 (a) and PC1-PC3 (b).

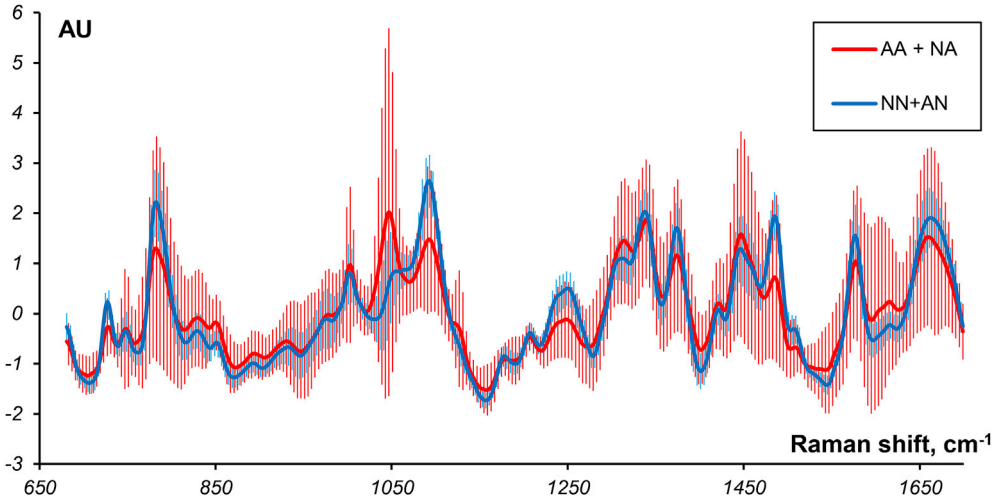

Fig. 5. Average spectra with their 95% tolerance intervals for NN + AN classes (blue) and AA + NA (red). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

This term and analysis of its distinct properties was recently published in paper [32]. The essential feature of 'vatrushka'- type data is that all objects from the target class are located completely in the middle (filling), while the alternatives occupy the periphery area (pastry).

4. Discussion

The reasons, which lead to abnormal development of the sperm, - or its subsequent damage, - are very different. The manifestations of such abnormalities vary widely also. The spectra acquired for the abnormal spermatozoa are always different and they do not form an entire class. It is illustrated in Fig. 7, where the average spectrum of the NN + AN class samples is presented as the black dotted curve. Additionally, the spectra of four abnormal samples (NA1, AA1, AA2, and NA2 in Fig. 6) are shown as the solid curves. It can be seen that sample NA1 (red) has a large band in the region of 1045 cm^{-1}, which is a marker of DNA damage according to the literature [10], [18]. At the same time, the band of sample NA2 (green) is insignificant in this region, but prominent bands appear in the regions of 1002 cm^{-1} and 1374 cm^{-1}.

These findings clearly correspond to the locations of these samples in the bi-plot (Fig. 6) – close to 1045 loadings point (NA1) and between the loadings points 1002 and 1374 (NA2).

In general, plots in Figs. 6 and 7 explain the nature of the spectrally normal and abnormal samples. The normal samples are very similar, while the abnormal ones are very unlike and they, in fact, comprise several subclasses with different loadings, which correspond to the bands round 1045, 1445, etc cm^{-1}. Basing on the information available in the literature, we tried to match the specific Raman bands to the corresponding bond vibrations in given functional groups of atoms (Table 1).

When analyzing a spectrum of the intact sperm nucleus, we distinguish the bands characteristic of DNA, 1092 cm^{-1} and 782 cm^{-1} (Figs. 1 and 5). Ellis et al. [33] showed that the 1092 cm^{-1} band is a marker of B-DNA and Z-DNA. Normally it has a flat or horizontal shoulder in the 1045 cm^{-1} region, which is a characteristic of the nucleus with intact DNA [10] [18]. The results of our study confirm this conclusion. The band round the 782 cm^{-1} region corresponds to thymine and cytosine vibrations, as well as to the DNA backbone. According to Huser et al. [16], this band is a marker of protamine-DNA packaging. The triple band in the 1420 cm^{-1}, 1445 cm^{-1} and 1486 cm^{-1} regions is ladder-shaped, with the highest part in the 1486 cm^{-1} region. The band round 1002 cm^{-1} characterizes the amino acid phenylalanine. It is stable and is presented with different intensities in almost all charts.

Partitioning of the samples into the normal and abnormal ones by their spectral characteristics is possible only employing the multivariate analysis. The summary results obtained using the DD-SIMCA classification are presented in Table 2.

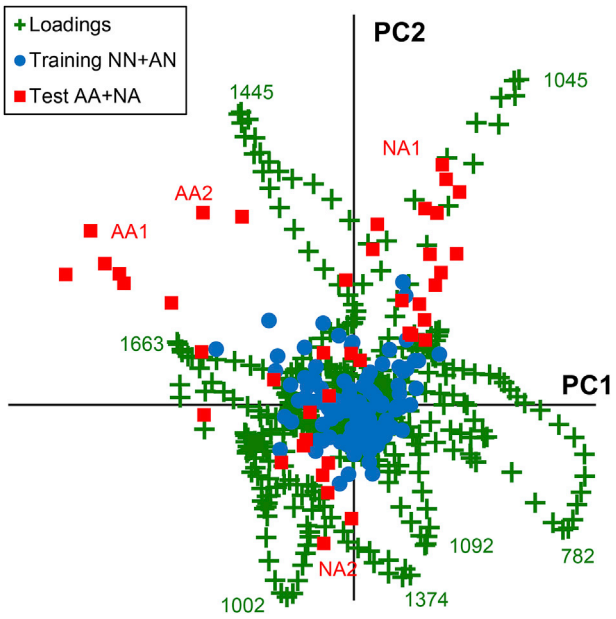

Fig. 6. PCA projection for the NN + AN data. Bi-plot (PC2 vs. PC1). The blue dots represent scores for the training (NN + AN) samples, the red squares show scores for the predicted AA + NA samples, the green crosses are loadings. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Normal and abnormal spectra are present in both groups; however, the normal spectra prevailed (82%) in group N, and abnormal spectra are equal to normal in group A.

Studying the group of morphologically abnormal spermatozoa we did not succeed in isolating the characteristic features inherent in the whole class of anomalous spectra due to its heterogeneity. Apparently, this can be explained by the fact that abnormality is an internal property determined by the multiple damages of the nuclear DNA, which manifests itself not in the specific absorption bands but in the entire spectral range. Thus, the separation of samples into normal and abnormal is a non-trivial task that can only be solved by proper analysis of spectral data.

To clarify this issue we selected the spectral dataset that contains the bands, which are primarily responsible for the DNA damage (thanks to anonymous reviewer for this idea) – 780, 1090, and 1045 cm^{-1} – and performed linear discriminant analysis on this data. The results

Table 1
Main CRS bands and the corresponding bond vibrations in given functional groups of atoms [16] [19], [33–35].

Wavenumber, cm^{-1}	Nucleic acid	Amino acid	Lipids
748	U, C (ring stretch)		
782	T, C (ring stretch) DNA backbone		
1002		Phenylalanine (Phenyl ring)	
1045	DNA damage marker		
1092	PO_2^- marker of B-DNA and Z-DNA (PO_2 symmetric stretch)		
1251	A, C (ring stretch)	N-H and C-H amide III	PO_2^- asymmetric stretch Phospholipids
1307	A (ring stretch)	N-H and C-H amide III	
1341	G (ring stretch)	N-H and C-H amide III	
1374	G (ring stretch)		
1420	A (ring stretch)		
1445			Methylene deformation (CH2δ) (endogenous lipids)
1486	G, A (ring stretch)		
1576	G, A (ring stretch)		
1663	T (C=O stretch)	C=O amide I	C=C Unsaturated lipid bonds

demonstrate that the spectrally normal (NN + AN) and abnormal (AA + NA) classes are clearly separated with accuracy of 94%. This is an indirect confirmation of our main idea that CRS is able to reveal the spermatozoa with damaged DNA.

5. Conclusions

According to the literature [36–38], severe degrees of teratozoospermia are associated not only with sperm aneuploidy, but also with various amounts of nuclear DNA fragmentation. Therefore, the sperm morphology is not an absolute predictor of the nuclear DNA stability. Rather, sperm morphology reflects a degree of probability of the DNA breaks in the morphologically abnormal spermatozoa during fragmentation tests [39]. As a starting point, it was suggested that

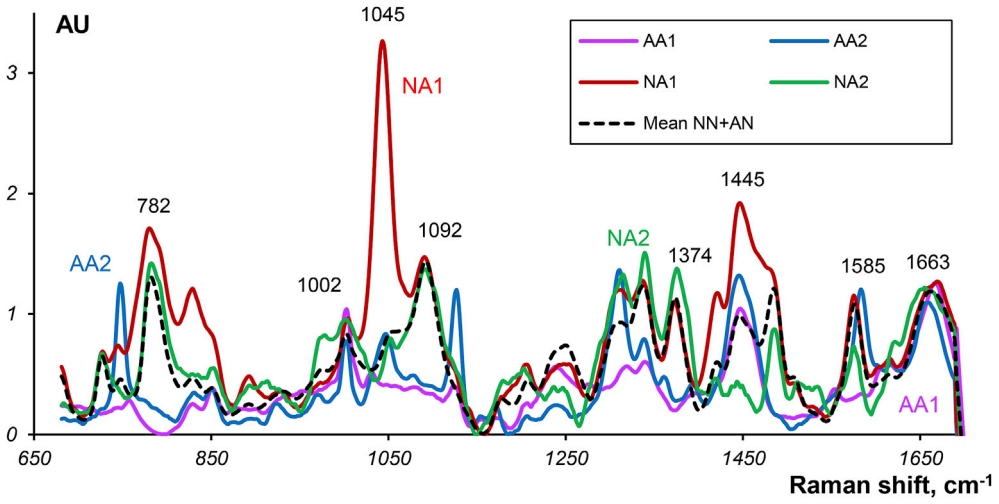

Fig. 7. Average Raman spectra of NN + AN class samples (dotted line) and spectra of four different abnormal samples (solid lines).

Table 2
Relationship between sperm groups and normal and abnormal spectral classes.

Group	Spectral Class	Number of spectra	Proportion, %
Morphologically normal sperm (group N)	Normal	102	82
	Abnormal	23	18
Morphologically abnormal sperm (group A)	Normal	17	47
	Abnormal	19	53

spermatozoa with normal morphology, in general, should have more stable and intact genomes than those of morphologically abnormal spermatozoa.

Investigating the whole collection of cells using CRS, we have found out that Raman spectra provide an additional information to standard microscope study:

(1) Separation samples into normal and abnormal using CRS is only partially consistent with the results of standard light microscopy.
(2) Cells considered normal by both approaches, CRS and standard light microscopy, compose a compact class. We can assume that this class characterizes the nucleus with intact, unfragmented, mature DNA. Using this class as a training set, we are able to differentiate the normal spectra from abnormal in the group of morphologically abnormal spermatozoa.
(3) Spectra that correspond to the abnormal cells have deviations due to various defects in nuclear DNA. This conclusion is made based on the previously published results presented by various laboratories. Of course, this is a preliminary assumption, which requires direct DNA analysis.

We assembled a spectral library characterizing mature, undamaged chromatin of spermatozoa with normal and abnormal morphology. It is supposed that CRS can make differentiation of spermatozoa with an intact nucleus from spermatozoa with fragmented and defective nuclear DNA possible. However, further research is needed in this area, and the accumulation of scientific data would assist in making objective conclusions. It may be interesting to use CRS for evaluating spectra from living, unfixed spermatozoa, and to select the viable ones. This seems to be a very promising direction for sperm selection in the ICSI procedure.

Declarations of interest

None.

Acknowledgments

O. Rodionova and A. Pomerantsev acknowledge a partial support within the Russian state assignment 0082-2014-0019.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chemolab.2018.10.002.

References

[1] T. Huser, J. Chan, Raman spectroscopy for physiological investigations of tissues and cells, Adv. Drug Deliv. Rev. 89 (2015) 57–70. https://doi.org/10.1016/j.addr.2015.06.011.
[2] E. Boros, P. Bak Porse, I. Nielsen, S. Kucheryavskiy, Using Raman spectroscopic imaging for non-destructive analysis of filler distribution in chalk filled polypropylene, J. Appl. Polym. Sci. 133 (2016) 43016. https://doi.org/10.1002/app.43016.
[3] N. Stone, M.C. Hart Prieto, P. Crow, J. Uff, A.W. Ritchie, The use of Raman spectroscopy to provide an estimation of the gross biochemistry associated with urological pathologies, Anal. Bioanal. Chem. 387 (2006) 1657–1668. https://doi.org/10.1007/s00216-006-0937-9.
[4] A.D. Meade, C. Clarke, F. Draux, G.D. Sockalingum, M. Manfait, F.M. Lyng, H.J. Byrne, Studies of chemical fixation effects in human cell lines using Raman microspectroscopy, Anal. Bioanal. Chem. 396 (2010) 1781–1791. https://doi.org/10.1007/s00216-009-3411-7.
[5] C. Krafft, J. Popp, The many facets of Raman spectroscopy for biomedical analysis, Anal. Bioanal. Chem. 407 (2014) 699–717. https://doi.org/10.1007/s00216-014-8311-9.
[6] T.J. Moritz, D.S. Taylor, D.M. Krol, J. Fritch, J.W. Chan, Detection of doxorubicin-induced apoptosis of leukemic T-lymphocytes by laser tweezers Raman spectroscopy, Biomed. Opt. Express 1 (2010) 1138. https://doi.org/10.1364/boe.1.001138.
[7] J.W. Chan, D.S. Taylor, T. Zwerdling, S.M. Lane, K. Ihara, T. Huser, Micro-Raman spectroscopy detects individual neoplastic and normal hematopoietic cells, Biophys. J. 90 (2006) 648–656. https://doi.org/10.1529/biophysj.105.066761.
[8] F.C. Pascut, H.T. Goh, N. Welch, L.D. Buttery, C. Denning, I. Notingher, Noninvasive detection and imaging of molecular markers in live cardiomyocytes derived from human embryonic stem cells, Biophys. J. 100 (2011) 251–259. https://doi.org/10.1016/j.bpj.2010.11.043.
[9] C. Mallidis, V. Sanchez, J. Wistuba, F. Wuebbeling, M. Burger, C. Fallnich, S. Schlatt, Raman microspectroscopy: shining a new light on reproductive medicine, Hum. Reprod. Update 20 (2013) 403–414. https://doi.org/10.1093/humupd/dmt055.
[10] V. Sánchez, K. Redmann, J. Wistuba, F. Wübbeling, M. Burger, H. Oldenhof, W.F. Wolkers, S. Kliesch, S. Schlatt, C. Mallidis, Oxidative DNA damage in human sperm can be detected by Raman microspectroscopy, Fertil. Steril. 98 (2012) 1124–1129. https://doi.org/10.1016/j.fertnstert.2012.07.1059.
[11] B.E. Speyer, A.R. Pizzey, M. Ranieri, R. Joshi, J.D.A. Delhanty, P. Serhal, Fall in implantation rates following ICSI with sperm with high DNA fragmentation, Hum. Reprod. 25 (2010) 1609–1618. https://doi.org/10.1093/humrep/deq116.
[12] D.P. Evenson, K.L. Larson, L.K. Jost, Sperm Chromatin Structure Assay: its clinical use for detecting sperm DNA fragmentation in male infertility and comparisons with other techniques, J. Androl. 23 (2002) 25–43. https://doi.org/10.1002/j.1939-4640.2002.tb02599.x.
[13] J. Tesarik, R. Mendoza-Tesarik, C. Mendoza, Sperm nuclear DNA damage: update on the mechanism, diagnosis and treatment, Reprod. Biomed. Online 12 (2006) 715–721. https://doi.org/10.1016/s1472-6483(10)61083-8.
[14] H.J. Chi, D.Y. Chung, S.Y. Choi, J.H. Kim, G.Y. Kim, J.S. Lee, H.S. Lee, M.H. Kim, S.I. Roh, Integrity of human sperm DNA assessed by the neutral comet assay and its relationship to semen parameters and clinical outcomes for the IVF-ET program, Clin. Exp. Reprod. Med. 38 (2011) 10. https://doi.org/10.5653/cerm.2011.38.1.10.
[15] W.L. Kubasek, Y. Wang, G.A. Thomas, T.W. Patapoff, K.H. Schoenwaelder, J.H. Van der Sande, W.L. Peticolas, Raman spectra of the model B-DNA oligomer d(CGCGAATTCGCG)2 and of the DNA in living salmon sperm show that both have very similar B-type conformations, Biochemistry 25 (1986) 7440–7445. https://doi.org/10.1021/bi00371a028.
[16] T. Huser, C.A. Orme, C.W. Hollars, M.H. Corzett, R. Balhorn, Raman spectroscopy of DNA packaging in individual human sperm cells distinguishes normal from abnormal cells, J. Biophot. 2 (2009) 322–332. https://doi.org/10.1002/jbio.200910012.
[17] K. Meister, D.A. Schmidt, E. Bründermann, M. Havenith, Confocal Raman microspectroscopy as an analytical tool to assess the mitochondrial status in human spermatozoa, Analyst 135 (2010) 1370. https://doi.org/10.1039/b927012d.
[18] C. Mallidis, J. Wistuba, B. Bleisteiner, O.S. Damm, P. Gross, F. Wubbeling, C. Fallnich, M. Burger, S. Schlatt, In situ visualization of damaged DNA in human sperm by Raman microspectroscopy, Hum. Reprod. 26 (2011) 1641–1649. https://doi.org/10.1093/humrep/der122.
[19] X. Li, M. Wang, H. Chen, Q. Li, H. Yang, H. Xu, Y. Lu, M. Zhang, X. Yang, S. Lu, K. Lu, Flow cytometric and near-infrared Raman spectroscopic investigation of quality in stained, sorted, and frozen-thawed buffalo sperm, Anim. Reprod. Sci. 170 (2016) 90–99. https://doi.org/10.1016/j.anireprosci.2016.04.008.
[20] World Health Organization (WHO), World Health Organization Laboratory Manual for the Examination and Processing of Human Semen, WHO, Geneva, 2010.
[21] E. Edengeiser, K. Meister, E. Bründermann, S. Büning, S. Ebbinghaus, M. Havenith, Non-invasive chemical assessment of living human spermatozoa, RSC Adv. 5 (2015) 10424–10429. https://doi.org/10.1039/c4ra12158a.
[22] S. Wold, K. Esbensen, P. Geladi, Principal component analysis, Chemometr. Intell. Lab. Syst. 2 (1987) 37–52. https://doi.org/10.1016/0169-7439(87)80084-9.
[23] P. Oliveri, M.I. López, M.C. Casolino, I. Ruisánchez, M.P. Callao, L. Medini, S. Lanteri, Partial least squares density modeling (PLS-DM) – a new class-modeling strategy applied to the authentication of olives in brine by near-infrared spectroscopy, Anal. Chim. Acta 851 (2014) 30–36. https://doi.org/10.1016/j.aca.2014.09.013.
[24] S. Joe Qin, Statistical process monitoring: basics and beyond, J. Chemometr. 17 (2003) 480–502. https://doi.org/10.1002/cem.800.
[25] A.L. Pomerantsev, Acceptance areas for multivariate classification derived by projection methods, J. Chemom. 22 (2008) 601–609. https://doi.org/10.1002/cem.1147.
[26] A.L. Pomerantsev, O.Y. Rodionova, Concept and role of extreme objects in PCA/SIMCA, J. Chemom. 28 (2013) 429–438. https://doi.org/10.1002/cem.2506.
[27] O.Ye Rodionova, YaV. Sokovikov, A.L. Pomerantsev, Quality control of packed raw materials in pharmaceutical industry, Anal. Chim. Acta 642 (2009) 222–227. https://doi.org/10.1016/j.aca.2008.08.004.
[28] A.L. Pomerantsev, Chemometrics in Excel, John Wiley & Sons, Inc., 2014. https://doi.org/10.1002/9781118873212.

[29] Y.V. Zontov, O.Y. Rodionova, S.V. Kucheryavskiy, A.L. Pomerantsev, DD-SIMCA – a MATLAB GUI tool for data driven SIMCA approach, Chemometr. Intell. Lab. Syst. 167 (2017) 23–28. https://doi.org/10.1016/j.chemolab.2017.05.010.
[30] P.B. Harrington, Multiple versus single set validation of multivariate models to avoid mistakes, Crit. Rev. Anal. Chem. 48 (2017) 33–46. https://doi.org/10.1080/10408347.2017.1361314.
[31] K.H. Esbensen, Multivariate Data Analysis – in Practice, fourth ed., CAMO, Oslo, 2000.
[32] A.L. Pomerantsev, O.Ye Rodionova, Multiclass partial least squares discriminant analysis: taking the right way — a critical tutorial, J. Chemometr. 32 (2018) e3030, https://doi.org/10.1002/cem.3030.
[33] D.I. Ellis, D.P. Cowcher, L. Ashton, S. O'Hagan, R. Goodacre, Illuminating disease and enlightening biomedicine: Raman spectroscopy as a diagnostic tool, Analyst 138 (2013) 3871. https://doi.org/10.1039/c3an00698k.
[34] N. Li, D. Chen, Y. Xu, S. Liu, H. Zhang, Confocal Raman micro-spectroscopy for rapid and label-free detection of maleic acid-induced variations in human sperm, Biomed. Opt. Express 5 (2014) 1690. https://doi.org/10.1364/boe.5.001690.
[35] M. Ferrara, G. Di Caprio, S. Managò, A. De Angelis, L. Sirleto, G. Coppola, A. De Luca, Label-free imaging and biochemical characterization of bovine sperm cells, Biosensors 5 (2015) 141–157. https://doi.org/10.3390/bios5020141.
[36] M. Smit, J.C. Romijn, M.F. Wildhagen, R.F.A. Weber, G.R. Dohle, Sperm chromatin structure is associated with the quality of spermatogenesis in infertile patients, Fertil. Steril. 94 (2010) 1748–1752. https://doi.org/10.1016/j.fertnstert.2009.10.030.
[37] S. Brahem, H. Elghezal, H. Ghédir, H. Landolsi, A. Amara, S. Ibala, M. Gribaa, A. Saad, M. Mehdi, Cytogenetic and molecular aspects of absolute teratozoospermia: comparison between polymorphic and monomorphic forms, Urology 78 (2011) 1313–1319. https://doi.org/10.1016/j.urology.2011.08.064.
[38] A. Perrin, N. Louanjli, Y. Ziane, T. Louanjli, C. Le Roy, N. Gueganic, V. Amice, M. De Braekeleer, F. Morel, Study of aneuploidy and DNA fragmentation in gametes of patients with severe teratozoospermia, Reprod. Biomed. Online 22 (2011) 148–154. https://doi.org/10.1016/j.rbmo.2010.10.006.
[39] C. Avendaño, A. Franchi, S. Taylor, M. Morshedi, S. Bocca, S. Oehninger, Fragmentation of DNA in morphologically normal human spermatozoa, Fertil. Steril. 91 (2009) 1077–1084. https://doi.org/10.1016/j.fertnstert.2008.01.015.