Vibrational Spectroscopy 126 (2023) 103529

Contents lists available at ScienceDirect

Vibrational Spectroscopy

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

Label-free discrimination of T and B lymphocyte activation based on vibrational spectroscopy – A machine learning approach

Luis Ramalhete [a,b,c,*,1], Ruben Araújo [b,2], Aníbal Ferreira [b,d,g,3], Cecília R.C. Calado [e,f,4]

[a] Blood and Transplantation Center of Lisbon, Instituto Português do Sangue e da Transplantação, Alameda das Linhas de Torres, nº 117, 1769 – 001 Lisboa, Portugal
[b] NOVA Medical School, Faculdade de Ciências Médicas, Universidade NOVA de Lisboa, 1169-056 Lisbon, Portugal
[c] iNOVA4Health - Advancing Precision Medicine, RG11: Reno-Vascular Diseases Group, NOVA Medical School, Faculdade de Ciências Médicas, Universidade NOVA de Lisboa, 1169-056 Lisbon, Portugal
[d] Centro Hospitalar Universitário Lisboa Central, Hospital Curry Cabral, Serviço de Nefrologia, Lisboa, Portugal
[e] ISEL-Instituto Superior de Engenharia de Lisboa, Instituto Politécnico de Lisboa, R. Conselheiro Emídio Navarro 1, 1959-007 Lisboa, Portugal
[f] CIMOSM — Centro de Investigação em Modelação e Otimização de Sistemas Multifuncionais, ISEL, Portugal
[g] Nova Medical School, Lisboa, Portugal

ARTICLE INFO

Keywords:
B lymphocytes
T lymphocytes
MIR spectroscopy
Cellular activation
Molecular fingerprint
Machine learning

ABSTRACT

B and T-lymphocytes are major players of the specific immune system, responsible by an efficient response to target antigens. Despite the high relevance of these cells' activation in diverse human pathophysiological processes, its analysis in clinical context presents diverse constraints. In the present work, MIR spectroscopy was used to acquire the cells molecular profile in a label-free, simple, rapid, economic, and high-throughput mode.

Recurring to machine learning algorithms MIR data was subsequently evaluated. Models were developed based on specific spectral bands as selected by Gini index and the Fast Correlation Based Filter. To determine if it was, possible to predict from the spectra, if B and T lymphocyte were activated, and what was the molecular fingerprint of T- or B- lymphocyte activation.

The molecular composition of activated lymphocytes was so different from naïve cells, that very good prediction models were developed with whole spectra (with AUC=0.98). Activated B lymphocytes also present a very distinct molecular profile in relation to activated T lymphocytes, leading to excellent prediction models, especially if based on target bands (AUC=0.99). The identification of critical target bands, according to the metabolic differences between B and T lymphocytes and in association with the molecular mechanism of the activation process highlighted bands associated to lipids and glycogen levels.

The method developed presents therefore, appealing characteristics to promote a new diagnostic tool to analyze and discriminate B from T-lymphocytes.

1. Introduction

B and T-lymphocytes are major key players of the adaptive (specific) immune system, where an efficient response to target antigens implies its activation mediated by the innate immune system. Lymphocytes are triggered through antigen-specific receptors present at their surface, leading to their activation and proliferation into a large number of specialized identical effector lymphocytes [1,2]. The initial trigger of

Abbreviations: LyB-PHa_yes, Activated B lymphocytes; LyT-PHa_yes, Activated T lymphocytes; ATR-FTIR, Attenuated Total Reflection – Fourier Transform Infrared Spectroscopy; AUC, Area Under the Curve; CA, Classification Accuracy; FCBF, Fast Correlation Based Filter; FTIRS, ourier-Transform Infrared Spectroscopy; k-NN, k-Nearest Neighbors; LDA, Linear Discriminant Analysis; MIR, Mid-infrared Spectroscopy; PHa, Phytohemagglutinin; PCA, Principal Component Analysis; LyB-PHa_no, Resting B lymphocytes; LyT-PHa_no, Resting T lymphocytes; SVM, Support-vector machine.

* Corresponding author at: Blood and Transplantation Center of Lisbon, Instituto Português do Sangue e da Transplantação, Alameda das Linhas de Torres, nº 117, 1769 – 001 Lisboa, Portugal.
E-mail addresses: luis.m.ramalhete@edu.nms.unl.pt (L. Ramalhete), cecilia.calado@isel.pt (C.R.C. Calado).
[1] https://orcid.org/0000–0002-8911–3380
[2] https://orcid.org/0000–0002-9369–6486
[3] https://orcid.org/0000–0002-3300–6033
[4] https://orcid.org/0000–0002-5264–9755

https://doi.org/10.1016/j.vibspec.2023.103529
Received 30 January 2023; Received in revised form 12 March 2023; Accepted 12 April 2023
Available online 12 April 2023

Table 1
Examples of studies exploring Raman and MIR spectroscopy to discriminate or detect activation of immune cells. Assays based on primary cells are highlighted in bold.

Study goal & sample dimension & activation mode	Technique: detection mode Model used: predicting values	Reference
Discriminate non-activated blood primary B (n = 25) from **primary** T-lymphocytes (n = 25)	MIR spectroscopy: transmission SVM: accuracy 98 %, precision 95 %	[31]
Activation of blood **primary** T-lymphocytes (n = 8), after 60 min incubation with phytohaemagluttinin	MIR spectroscopy: transmission HCA: Sens. & Spec. 100 %	[32]
Activation of Jurkat T cells (n = 2), after 75 min incubation with anti-CD3 antibody	MIR spectroscopy/ ATR Two-tailed paired student's t-test (absorbance at 1367 and 1358 cm^{-1}): $p < 0.02$	[33]
Activation of murine macrophage (n = 20), after 60 min incubation with lipopolysaccharide and gamma interferon	MIR spectroscopy: ATR LDA: accuracy 80 %	[34]
Activation of Raji B (n = 1), Jurkat T cell (n = 1), THP1 monocytes (n = 1), **PBMC** healthy volunteer (n = 2), after 24–72 hr incubation with phytohaemagluttinin or phorbol 12-myristate 13-acetate or lipopolysac- charide	Raman Spectroscopy PCA and LDA: Raji B: Sens. 95.2 % Spec. 81.6 % Jurkat: Sens. 91 % Spec. 70 %; Monocyte: Sens. 91 % Spec. 91 %	[27]
Discriminate mouse **primary** B (n = 60) from T-lymphocyte (n = 96) and mouse activation of **primary** T-lymphocyte after 48 hr incubation anti-CD3 and anti-CD28	Raman spectroscopy PCA and LDA B vs T-lymphocytes error rate 2.2 %; T-lymphocytes error rate 2.3 %	[28]
Activation of **primary** T-lymphocytes (n = 37) in kidney transplanted, mitomycin C inactivated (n = 28), resting T-lymphocytes (n = 35) and after 7-day incubation in mixed lymphocyte culture	Raman spectroscopy DFA 785 nm Sens. 95.7 %, Spec. 100 % 514.5 nm Sens. 89.3 %, Spec. 93.8 %,	[29]
Activation of **primary** T-lymphocytes, after 180 min incubation with phytohemagglutinin	MIR spectroscopy: transmission PCA	[35]

SVM, Support Vector Machine; HCA, Hierarchical Cluster Analysis; Sens., Sensitivity; Spec., Specificity; ATR, Attenuated Total Reflection; LDA, Linear Discriminant Analysis; PCA, Principal Component Analysis; DFA, Discriminant Function Analysis; k-NN; K-Nearest Neighbors

Table 2
Machine Learning algorithm models to discriminate activated from resting lymphocytes (B and T).

Model	AUC	CA	F1	Precision	Recall
2^{nd} derivative of complete spectra					
k-NN	0.95	0.85	0.83	0.86	0.85
SVM	0.975	0.864	0.85	0.87	0.86
2^{nd} derivative with target bands defined by the Gini index					
k-NN	0.98	0.94	0.94	0.94	0.94
SVM	0.94	0.89	0.89	0.90	0.89

that response is usually associated to pathogen-associated immunostimulants (dependent on receptors recognizing molecular patterns associated to pathogens), leading to the production of extracellular signaling that promote inflammation and help activate specific immune responses [2].

The activation of lymphocytes occurs in stages that will, inevitably, impact the cellular metabolism. In each transition, the cell will pass through specific metabolic states, orchestrated in a succession of transcriptional factors and transduction signals working in a network, leading to cellular differentiation and expansion. For example, whenever a naïve lymphocyte is activated, the change in the cellular state leads to a more rapid intake of glucose and an increase in demand for metabolic precursors necessary to the biogenesis of proteins, nucleic acids and lipids, where, understandably, these changes will be transmitted to the cell metabolomic profile [3–6].

The assessment of activation of lymphocytes is essential in several clinical scenarios, in the diagnosis and prognosis of diseases, e.g., in auto-immune disorders, cancer progression and in allograft rejection. Notwithstanding the knowledge of lymphocyte activation, it is also relevant to define what type of lymphocytes are activated. For example, allograft rejection may result from immune-cellular mechanisms, mediated by T-lymphocytes and primed by donor-derived antigens, or be associated to an acute or chronic antibody-based rejection mediated by B-lymphocytes [7–9]. Therefore, the identification of the type of activated lymphocytes (B or T) will enable the identification of the immune-rejection mechanism and, consequently, to define the immune or rescue therapy. The monitoring of the type of activated lymphocyte will result in more efficient and personalized therapy [1,7,10–13].

Despite the importance of lymphocyte activation in diverse pathophysiological processes, conventional clinical laboratories do not routinely conduct these analysis, due to diverse limitations associated to the conventional analytical techniques [10,12–16]. For example, most techniques based on interaction of antigens among a high diversity of the cells polyclonal populations (e.g., enzyme-linked immunosorbent assay (ELISA), enzyme-linked immune absorbent spot (ELISPOT), flow cytometry and proliferation assays, may present low sensitivity and/or specificity [17]. One of the most used techniques is ELISPOT due to its high sensitivity and due to allowing automatic analysis.

The method usually rapid, since it does not require a period for cell activation and proliferation. This is due to the fact that ELISPOT, is based on the analyses of produced cytokines. Therefore, a major limitation is the indirect prediction of the immune cell's activation, carrying only a somewhat reasonable specificity since these assays lack the ability of providing information on the ability of antigen-specific lymphocytes to mediate other effector functions. Proliferation or cytotoxicity-based assays, such as mixed lymphocyte culture, limiting dilution assays or functional assays as the ImmunoKnow® assay, measuring ATP on CD4 cells, are usually expensive and time consuming, while others resort to radioactive isotopes [10,12–16,18]. Therefore, it is relevant to develop a new method enabling to predict these immune cells activation, in a highly specific, sensitive, reproducible, rapid, simple and economic mode. Vibrational spectroscopy, including Raman and InfraRed (IR) spectroscopy, measures vibrational modes of biomolecules and presents diverse characteristics that could enable achieving that goal. In Raman and Mid-infrared (MIR) spectroscopy, it is possible to analyze diverse human samples, e.g., blood, blood cells, serum, and urine, resulting in the identification of diverse pathologies with a high degree of specificity and sensitivity, e.g., in neurology [19–21], nephrology [22,23], or oncology [24–26].

Due to the high potential of vibrational spectroscopy based-methods, many researchers have explored Raman and MIR spectroscopy to identify immune-cells or activated immune-cells. Unfortunately, some works are based on animal cell lines, which lack the natural variability associated with in vivo models, i.e., to primary lymphocytes as acquired from human peripheral blood (Table 1). From the experiments based on primary cells, it is worthy to highlight 3 Raman spectroscopy-based works, that, however, were based on an very low dimension sample (with only 2 samples from peripheral blood monocyte cells, PBMC) [27], or where the activation was based on a long period between 2 and 7 days of incubation [28,29]. Theoretically, MIR spectroscopy may be advantageous over Raman spectroscopy, due to the low probability of a molecule to undergo Raman state transition, usually, higher concentrations of the target analyte is required, which can decrease the analysis sensitivity and specificity in relation to MIR spectroscopy [30]. Probably

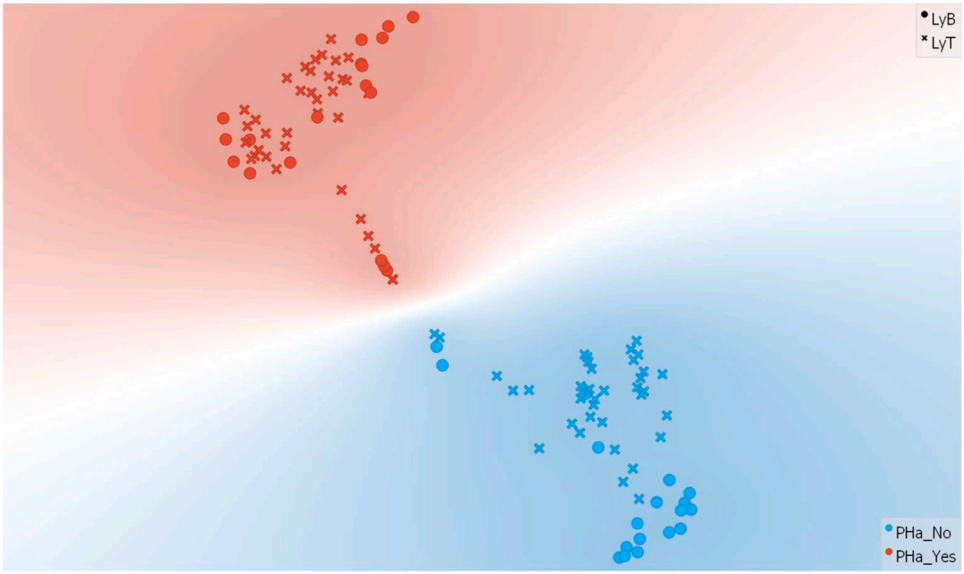

Fig. 1. *t*-SNE of second derivative spectra from non-activated (Blue) and activated B- and T-lymphocytes (red).

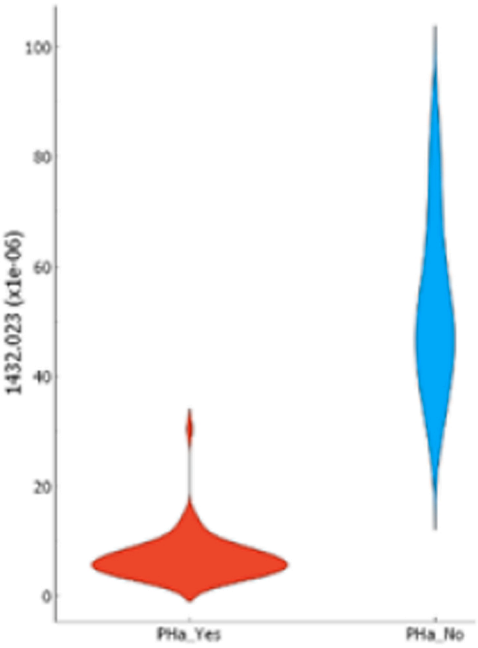

Fig. 2. Violin plot of the 1432 cm^{-1} spectral band relative to activated (by PHA incubation) and non-activated T and B lymphocytes.

due to this, both works previous referred needing 2–7 days incubation, in order to detect cells activation based on Raman spectroscopy.

From works conducted in MIR spectroscopy, it is important to highlight our laboratory previous work, enabling to discriminate primary non-activated T- from B-lymphocyte with a 98 % accuracy [31], and the T-lymphocyte activation (n = 8) after 1 hr. incubation with 100 % accuracy [32]. The main goal of the present work was to develop predictive methods based on MIR spectra of the activation of lymphocytes (B and T lymphocytes) obtained from PBMC of 18 volunteers, and to predict which lymphocytes (B-or T) were activated. Furthermore, the MIR spectra acquisition was conducted with use of a 96-wells microplate to enable an automatic and high-throughput analysis. Therefore, we aim to promote a reliable method of predicting, from the MIR spectra of PBMC, in a rapid, economic, and high-throughput mode, the lymphocyte activation, and in the case of activation, be able to quickly ascertain as to which type of lymphocyte population was activated.

2. Materials and methods

2.1. Biological assay

Ten milliliters of heparinized sodium whole venous blood was collected from 18 healthy volunteers, whose informed consent was collected from. Whole blood was diluted at 1:1 (v/v) with RPMI medium (Lonza) and centrifuged at 800xg for 10 min. Buffy coat was extracted and used for T and B lymphocyte isolation, being negatively sorted with immunomagnetic beads using an EasySep® Kit (StemCell Technologies, Vancouver, British Columbia, Canada), according to the manufacturer's recommendations and resuspended in RPMI medium until usage. Cells were counted in a Neubauer chamber and their viability was verified with trypan blue.

2.2. T and B Lymphocyte activation

Lymphocyte populations were subdivided into two groups resting and activated, resuspended at a 4×10^5 cells per 50 μL of saline phosphate buffer (PBS). T and B lymphocytes activation was conducted by incubation with the lectin mitogen phytohemagglutinin (PHa, Sigma Aldrich, St. Louis, MO) at 12.5 μg/ml, for 60 min at 37 °C. Cells were washed with PBS and resuspended in 50 μL PBS, according to Kwack et al. [36]. Resting T- and B- lymphocytes were submitted to the same protocol, but without PHa. In the end 72 samples were obtained, 18 activated T-cell (LyT-PHa_yes), 18 resting T-cell (LyT-PHa_no), 18 activated B-cell (LyB-PHa_yes) and 18 resting B-cell (LyB-PHa_no). Samples were resuspended in PBS and adjusted to a concentration of 8×10^3/μL.

2.3. MIR spectra acquisition

Samples of 25 μL with 2×10^5 of T or B lymphocytes were plated into a 96-wells Si microplate and then dehydrated in a desiccator for 150 min under vacuum. Spectra were collected using a FTIR spectrometer (Vertex 70, Bruker), equipped with an HTS-XT (Bruker) accessory. Each spectrum represented 64 coadded scans, acquired in transmission mode between 400 and 4000 cm^{-1}, with a resolution of 2 cm^{-1}. The first well of the 96-wells microplate was left without a sample and the corresponding spectra used as background, according to the HTS-XT

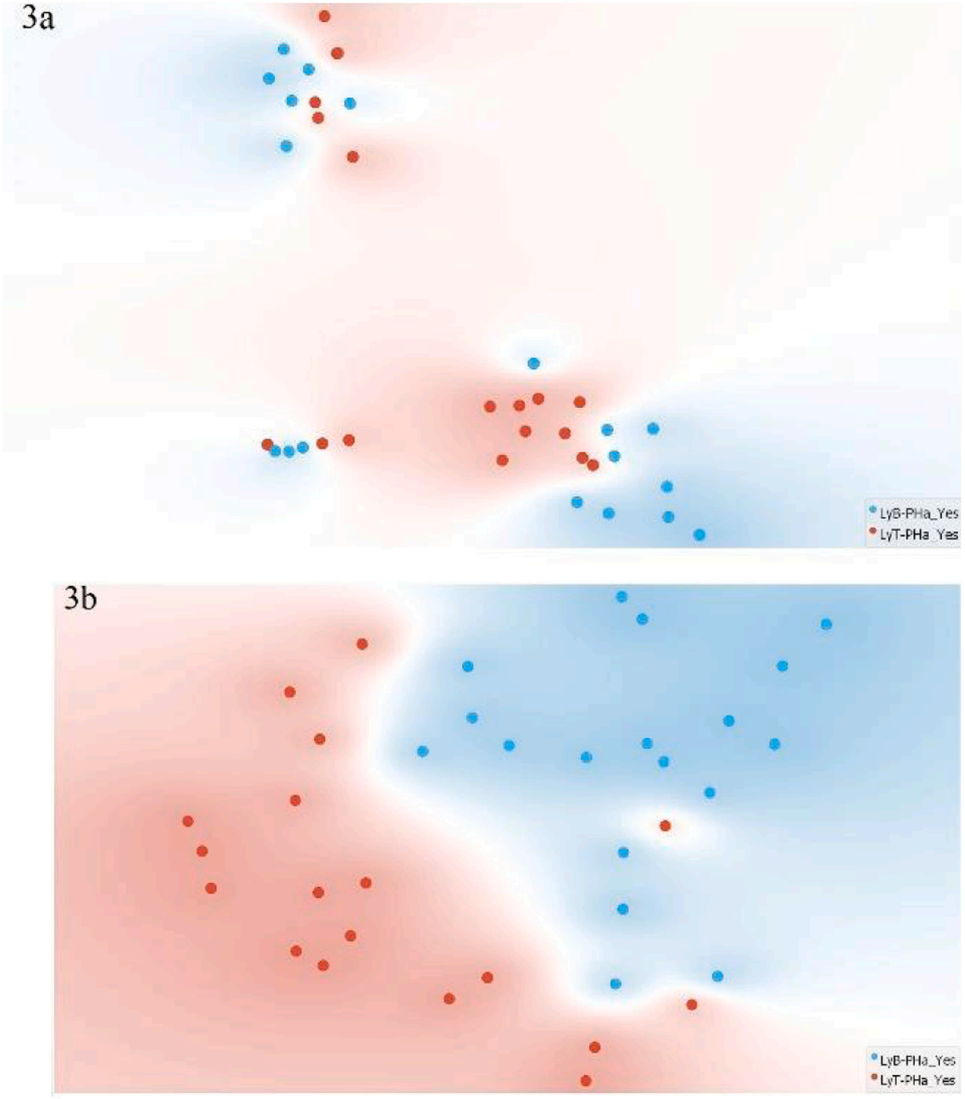

Fig. 3. *t*-SNE of activated T- and activated B- lymphocytes, based on whole second derivative spectra (3a) or based on bands selected by FCBF (3b).

Table 3
Machine Learning algorithms performance based on different pre-processing methods for discriminating activated T-lymphocytes from activated B-lymphocytes.

Model	AUC	CA	F1	Precision	Recall
Activated T vs B Lymphocyte					
2nd derivative					
k-NN	0.77	0.76	0.68	0.82	076
SVM	0.91	0.88	0.87	0.88	0.88
2nd derivative based on wavelengths as defined by FCBF (n = 10)					
k-NN	0.95	0.88	0.88	0.88	0.88
SVM	**0.99**	**0.96**	**0.96**	**0.96**	**0.96**

manufacturer.

2.4. Spectra pre-processing and processing

Spectra were pre-processed by atmospheric compensation and a second derivative spectra based on a Savitzky-Golay filter, with a 2nd polynomial and a corresponding window size of 19 points [31,37]. Atmospheric compensation was conducted with OPUS® software, version 6.5 (Bruker, Germany), while second derivative spectra and following spectra processing methods were conducted by Orange3 version 3.19.0 (Bioinformatics Lab, University of Ljubljana, Slovenia).

2.5. Machine learning models and statistical analysis

Machine learning models, namely, *t*-distributed stochastic neighbor embedding (*t*-SNE), k-nearest neighbors (k-NN) and support-vector machine (SVM) models were applied. k-NN and SVM models used a 3rd degree polynomial kernel with a cross-validation random sampling, based on 80 % and 20 % of data for model training and validation, respectively, were conducted by Orange3 version 3.19.0 (Bioinformatics Lab, University of Ljubljana, Slovenia). Models were assessed by the area under the curve (AUC), classification accuracy (CA), F-1 score, precision and recall. Models were also developed based on specific spectral bands as selected by Gini index and the Fast Correlation Based Filter (FCBF) [38,39]. The number of selected spectral bands, i.e., features, was based on the lower number of features needed to obtain the highest accuracy (CA). ANOVA was conducted to analyze the impact of specific spectral bands between data sets.

3. Results and discussion

Considering a future application of the technique, two important

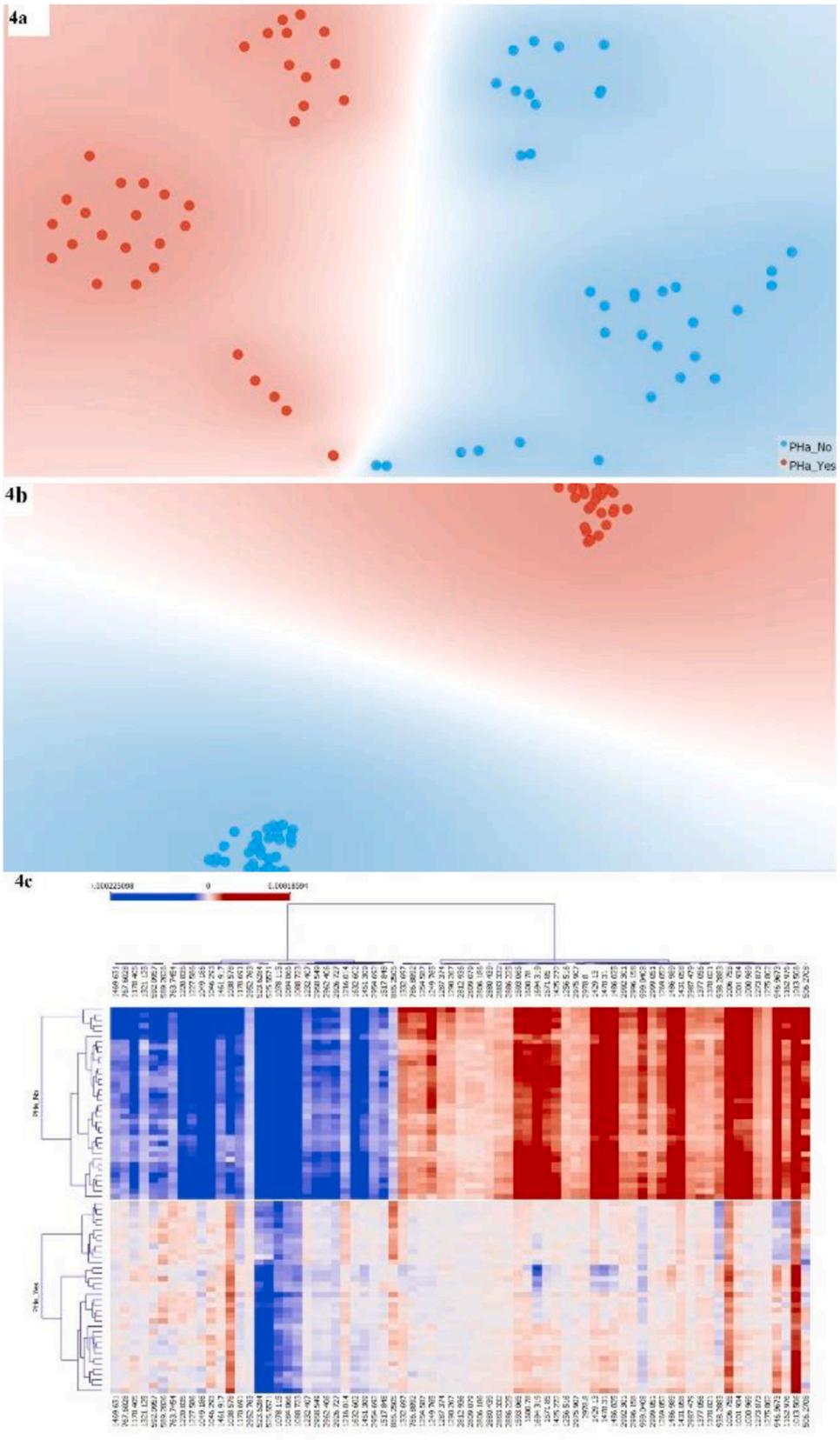

Fig. 4. Discriminating activated from resting T-lymphocytes. 4a) t-SNE plot based on the cells complete spectra, 4b) t-SNE plot based on bands selected by FCBF, 4c) Heatmap of differentially absorbed wavenumbers of 2nd derivative 73 spectral bands identified by FCBF.

Table 4
Spectral bands selected by FCBF as contributing to the separation between naïve and activated T or B- lymphocytes.

wavelengths cm^{-1}	B- lymphocytes % bands	T- lymphocytes % bands
0500–1000	43	16
1000–1500	35	52
1500–2000	3	8
2500–3000	15	23
3000–3500	3	0

Table 5
Machine Learning algorithms performance based on different pre-processing methods for discriminating activated from resting T cell and activated from resting B lymphocytes.

Model	AUC	CA	F1	Precision	Recall
T-Lymphocytes					
2nd derivative complete spectra					
k-NN	1.0	1.0	1.0	1.0	1.0
SVM	1.0	1.0	1.0	1.0	1.0
2nd derivative of selected bands based on FCBF					
k-NN	1.0	1.0	1.0	1.0	1.0
SVM	1.0	1.0	1.0	1.0	1.0
B-Lymphocytes					
2nd derivative complete spectra					
k-NN	1.0	0,96	0,96	0,96	0,96
SVM	1.0	0,96	0,96	0,96	0,96
2nd derivative of selected bands based on FCBF					
k-NN	1.0	0,96	0,96	0,96	0,96
SVM	1.0	1.0	1.0	1.0	1.0

questions first need to be answered, based on the patient PBMC spectra:

– Are these cells activated?
– If so, are these activated B- or T-lymphocytes?

To answer these questions, after lymphocyte isolation (based on negatively magnetic-activated cell sorting), from the PBMC of 18 volunteers, cells were subdivided into 4 populations (2 of T- and 2 B-lymphocytes), one of T-lymphocytes and the other of B- lymphocytes, were submitted to 1 h. PHa stimulus, and the other two cellular populations were submitted to the same process without PHa stimulation. Regarding these cells, it was evaluated by MIR spectra:

- The activation of lymphocytes (considering both B- and T- lymphocytes), in relation to resting lymphocytes.
- The discrimination between activated T- lymphocytes and activated B- lymphocytes

The individual analysis of B- and T-lymphocyte activation to further identify molecular features associated to each of these cells was also conducted.

3.1. Discriminating activation of whole (B and T) lymphocytes

The *t*-SNE score plot, based on the lymphocytes 2nd derivative of the complete spectra (400–4000 cm^{-1}), points to a perfect separation between naïve and activated lymphocyte, highlighting a significant different biochemical composition of activated immune cells in relation to resting cells. This is in accordance to the significant metabolic change occurring during the activation process [40,41]. For instance the activation of T-lymphocytes can result in expression changing of approximately 27,000 out of 39,500 genes compared to resting cells [42], that leads, for example, to the change of around 20 % of the proteome (1119 out of 5237 proteins) [40].

To quantify and develop predictive models of the activation process based on the cells spectra, k-NN and SVM models were built based on the complete second derivative spectra. According to the previous *t*-SNE analysis, excellent K-NM and SVM models were developed (with AUCs between 0.95 and 0.98), enabling to predict if the lymphocyte population under evaluation was a naïve or an activated population (Table 2). From the Gini index and the FCBF methods to select the most significant wavelength bands between naive and activated lymphocytes, the Gini index based on the following regions, resulted in the best accuracy to predict lymphocytes activation (from the highest to the lowest impact): 1047, 1469, 1043, 2880, 1432, 768, 1461, 589, 1252, 787 and 1486 cm^{-1}. Fig. 2 displays a violin plot of one of these bands as an example. The k-NN based on these bands presented an increased AUC (from 0.95 to 0.98), and the model's accuracy, F1, precision and recall, also increased from values between 0.83 and 0.86–0.94. The SVM model predictability slightly decreased from an AUC of 0.98–0.94 (Table 2).

The band at 2880 cm^{-1}, associated with lipids (C–H stretching), and the band at 1252 cm^{-1} associated with phosphate groups such as from phospholipids[43], are according to the effect of immune cells activation on the membrane lipid rafts, as pointed either on T-lymphocyte[3, 5, 44] either on B-lymphocyte[3,6]. Titus et al. [33], also observed the ~1250 cm^{-1} band as relevant in the T- lymphocyte activation. The 1043 cm^{-1} band, associated to glycogen, is highly affected during the cell's activation, due to the process needed for energy.

3.2. Discriminating activated T cell from activated B lymphocytes

The scores between activated T and activated B-lymphocytes are not separated in the *t*-SNE score plot based on the cells whole second derivative spectra (Fig. 3a). This highlights a more similar biochemical composition between these two (B and T) lymphocytes, than observed between non-activated and activated B- and T-lymphocytes.

A very good separation between scores in t-SNE were, however achieved, based on the following target second derivative bands, as identified by the FCBF (Fig. 3b):1043, 1049, 2850, 1488, 1197 and 604 cm^{-1}, respectively. Interestingly, the bands at 2850 and 1043 cm^{-1} are highlighted as the ones with higher values for activated B-lymphocytes when compared to activated T-lymphocytes (data not shown). While 2850 cm^{-1} band is most probably reflecting lipids constitution and organization at the cells membrane, 1043 cm^{-1} band reflects the cells glycogen levels [45]. B-lymphocytes are known to be lower energy demanding than T- lymphocytes [46], which could explain the higher glycogen levels on these cells as detected by MIR spectroscopy.

An excellent SVM predicting model was obtained based on the cells complete second derivative spectra, enabling to predict if the activated cells are lymphocytes B or T-lymphocytes (AUC=0.91) (Table 3). The model's predictability was improved when based on second derivative bands pointed by FCBF, according to the observed on the *t*-SNE (Fig. 3, Table 3). A SVM model was achieved, with an AUC of 0.99 and with an accuracy, F1, precision and recall of 0.96.

The excellent models predicting which lymphocyte population was presented is according to the fact that the mitogen-independent cell cycle progression in B-lymphocytes are different from T-lymphocytes [47,48].

3.3. Assessing the molecular fingerprint of T lymphocyte activation

In a previous work, with a smaller population (n = 8), it was possible to discriminate, based on hierarchical cluster analysis of MIR spectra, naïve from activated T- lymphocytes [32]. In the present work, we have tested a higher dimension population (duplicate assays from samples obtained from 18 patients, i.e., 36 assays were conducted) and developed predictive models of -lymphocytes activation.

The *t*-SNE plots based on the complete second derivative spectra, points to a separation between naïve and activated T-lymphocytes (Fig. 4a), further increased in the score-plot based on bands as selected by FCBF (Fig. 4b). The heatmap shows distinct bands associated to T-lymphocytes activation, including decreased values associated to 30

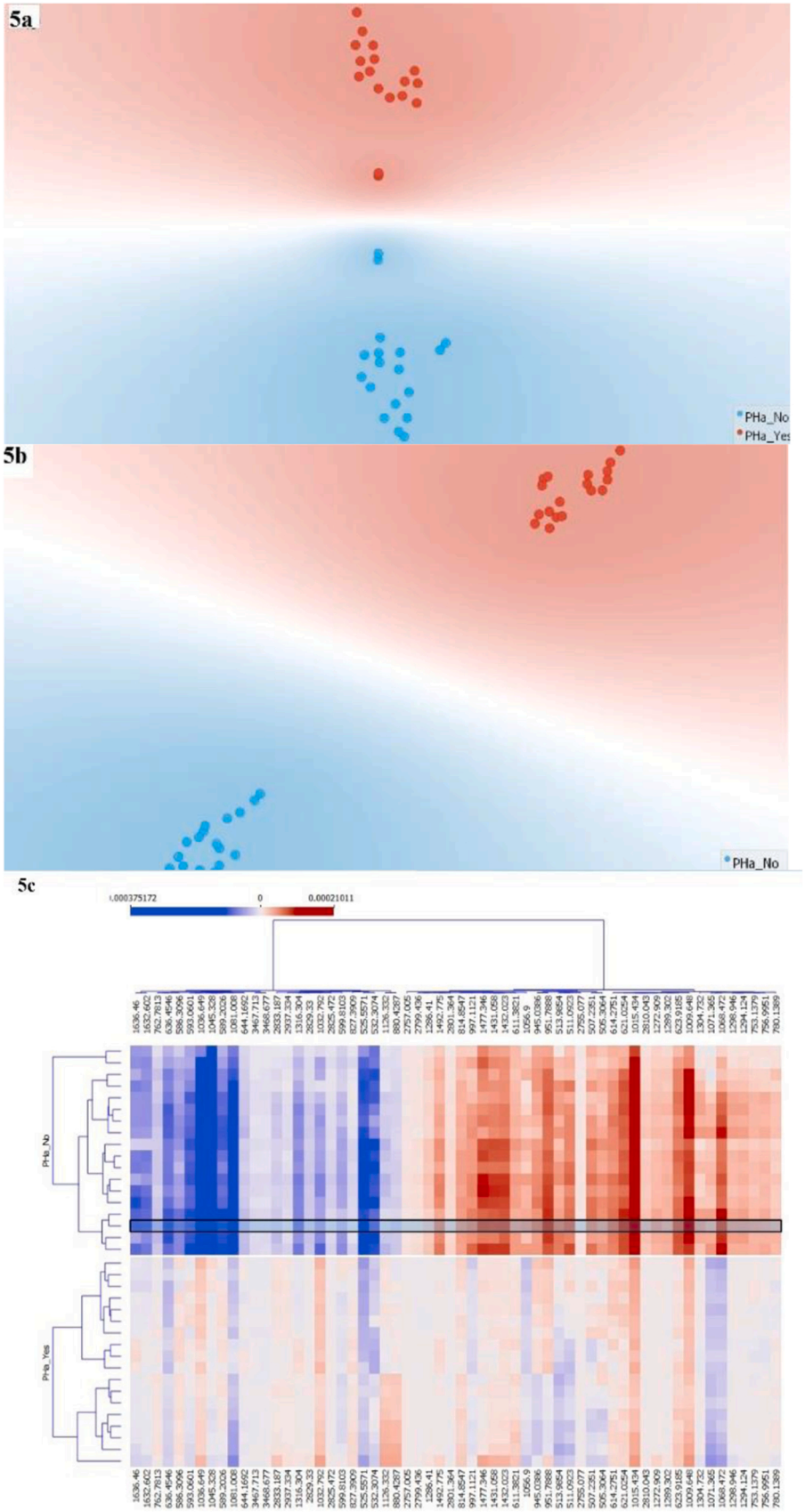

Fig. 5. Discriminating activated from resting B- lymphocytes. 5a) t-SNE plot based on cells second whole spectra, 5b) t-SNE plot based on bands selected by FCBF, 5c) Heatmap of differentially absorbed wavenumbers of 2nd derivative 60 spectral bands identified by FCBF.

bands (e.g., 1464, 767, 1178 and 1321 cm^{-1}) and increased values associated with 43 bands (e.g., 506, 1013, 1162 and 946 cm^{-1}).

To note that 68 % of the second derivative bands pointed by FCBF, as represented in the heatmap, are comprised between 500 cm^{-1} and 1500 cm^{-1} (Table 4). The fingerprint region includes a high diversity of molecular vibrations, including C-O, C-C, and C-N single bond stretches, C-H bending vibrations, and bands related to benzene rings, primarily reflecting differences in phosphate and sugar absorptions in DNA structures. Despite the complexity of these regions as reviewed by Liu et al. [49], and Bunaciu et al. [50], this is a very important region in terms of lymphocyte activation and, as observed by Wood et al. [35] major spectral changes can be observed within this region after activation, as a result from an increase in overall RNA synthesis. The region between 2500 and 3000 cm^{-1}, due to vibration bond of lipids, has also impact in the discrimination between naïve and activated T-lymphocytes, most likely reflecting the relevance of the change at the cells membrane occurring during cell activation, and general lipids metabolism relevance in the differentiation, effector function and survival of T-lymphocyte [3,51,52].

Excellent k-NN and SVM models were achieved, enabling the prediction of T- lymphocytes activation, based either on the complete second derivative spectra or based on features selected by FCBF, resulting in models with AUC of 1.0 (Table 5).

3.4. Assessing the molecular fingerprint of B lymphocyte activation

Due to lower quantities of B-lymphocytes obtained from the isolation process from PBMC (usually, circulating B-lymphocytes represents approximately 10–20 % of lymphocyte population [53]), only one assay was conducted per sample of each patient.

The *t*-SNE plots presents a good separation between naïve and activated B-lymphocytes based on whole second derivative spectra (Fig. 5a). The separation between these two types of samples increases as based on bands selected by FCBF (Fig. 5b). The heatmap shows distinct bands associated to B-lymphocytes activation, including decreased values associated to 35 bands (e.g., 780, 756, 1294 and 1298 cm^{-1}) and increased values associated with 25 bands (e.g., 1636, 762, 635 and 586 cm^{-1}).

From the bands pointed by FCBF as relevant for lymphocytes activation, 78 % are comprised in the fingerprint region (Table 4.). Once again, 15 % of relevant bands were from the lipids region, according to the relevance of the organization of the membrane lipids rafts during the cells activation process [52].

Excellent k-NN and SVM models predicted B-Lymphocyte activation based either on whole second derivative spectra or based on bands selected by FCBF, pointing models with AUC of 1.0 (Table 5).

4. Conclusions

The ability to identify lymphocyte activation in a simple procedure, economically and based on a high-throughput mode is a fundamental need in varied clinical scenarios, e.g., solid organ transplantation and evaluation of immunosuppressive regimes. To achieve that goal, in the present work, it was considered B and T-lymphocytes obtained from PBMC of 18 volunteers. The lymphocytes activation was conducted on 1 h. incubation with PHA, a potent mitogen inducing activation and proliferation of lymphocytes. MIR spectra of activated and non-activated lymphocytes were acquired in a 96-well Si microplate. It was subsequently evaluated if it was possible to predict from the spectra, if lymphocytes (B and T) were activated, and if so, which cells population (B or T) were activated. To evaluate this, non-supervised methods (i.e., *t*-SNE) and supervised methods of machine learning algorithms (i.e., SVM and k-NN) were conducted based on the complete second derivative spectra or, in order to improve the classification, models were based on specific bands selected by the Gini index and FCBF. The molecular composition of activated lymphocytes was so different from naïve cells that very good prediction models were developed while making use of the complete spectra (with AUC=0.98). Activated B-lymphocytes also present a very distinct molecular profile in relation to activated T-lymphocytes, leading to excellent prediction models, especially if based on target bands (AUC=0.99). It was also possible to predict, for individual B and T-lymphocytes, the process activation (AUC=1.0). When considering isolated cells, a higher dimension of bands (i.e., $n > 60$) as selected by the FCBF, it was possible to identify as relevant, according to the high impact of the activation process on the gene's expression. To discriminate activated B from activated T-lymphocytes, a fewer number of bands were identified (i.e., $n < 10$) as significant, pointing more similar compositions between activated B-lymphocytes and activated T-lymphocytes, when compared to the corresponding naïve cells. Interestingly, from these critical bands, the bands associated to lipids and glycogen levels were highlighted, revealing the metabolic differences between B and T-lymphocytes. This work has made some interesting findings and laid down the basis for future work to improve on the classification models and understanding of the spectral difference detected.

Ethical Approval

Not applicable, all data was provided without any identifier or group of identifiers which would allow attribution of private information to an individual.

Funding

This research was funded by project grant DSAIPA/DS/0117/2020 and PTDC/EQU-EQU/3708/2021 supported by Fundação para a Ciência e a Tecnologia. Portugal, and by the project grant NeproMD/ISEL/2020 financed by Instituto Politécnico de Lisboa. Portugal. The present work was conducted in the Engineering & Health Laboratory, the result from a collaboration protocol established between Universidade Católica Portuguesa and Instituto Politécnico de Lisboa.

CRediT authorship contribution statement

LR. and CC. conceptualized the draft. LR and CC. contributed equally to writing and reviewing of the original draft through interpretation of the literature. RA. reviewed and edited the original draft. AF. critically commented and edited the article. All authors have read and agreed to the published version of the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The participants of this study did not give written consent for their data to be shared publicly, so due to the legal context the research supporting data is not available.

References

[1] B.V. Kumar, T.J. Connors, D.L. Farber, Human T cell development, localization, and function throughout life, Immunity 48 (2018) 202–213, https://doi.org/10.1016/j.immuni.2018.01.007.
[2] A.K. Abbas, C.A. Janeway, Immunology, Cell 100 (2000) 129–138, https://doi.org/10.1016/S0092-8674(00)81689-X.
[3] R.E. Lamerton, A. Lightfoot, D.J. Nieves, D.M. Owen, The role of protein and lipid clustering in lymphocyte activation, Front. Immunol. 12 (2021) 1–8, https://doi.org/10.3389/fimmu.2021.600961.
[4] L. Tuosto, C. Xu, Editorial: membrane lipids in T cell functions, Front. Immunol. 9 (2018) 569–572, https://doi.org/10.3389/fimmu.2018.01608.

[5] S. Zumerle, B. Molon, A. Viola, Membrane rafts in T cell activation: a spotlight on CD28 costimulation, Front. Immunol. 8 (2017) 1–7, https://doi.org/10.3389/fimmu.2017.01467.

[6] E. Sezgin, I. Levental, S. Mayor, C. Eggeling, The mystery of membrane organization: composition, regulation and roles of lipid rafts, Nat. Rev. Mol. Cell Biol. 18 (2017) 361–374, https://doi.org/10.1038/nrm.2017.16.

[7] E. Ingulli, Mechanism of cellular rejection in transplantation, Pediatr. Nephrol. 25 (2010) 61–74, https://doi.org/10.1007/s00467-008-1020-x.

[8] P. Randhawa, T-cell-mediated rejection of the kidney in the era of donor-specific antibodies, Curr. Opin. Organ Transpl. 20 (2015) 325–332, https://doi.org/10.1097/MOT.0000000000000189.

[9] P.I. Terasaki, J. Cai, Humoral theory of transplantation: further evidence, Curr. Opin. Immunol. 17 (2005) 541–545, https://doi.org/10.1016/j.coi.2005.07.018.

[10] J.D. Folds, J.L. Schmitz, Clinical and laboratory assessment of immunity, J. Allergy Clin. Immunol. 111 (2003) 702–711, https://doi.org/10.1067/mai.2003.122.

[11] R.K. Sahu, U. Zelig, M. Huleihel, N. Brosh, M. Talyshinsky, M. Ben-Harosh, S. Mordechai, J. Kapelushnik, Continuous monitoring of WBC (biochemistry) in an adult leukemia patient using advanced FTIR-spectroscopy, Leuk. Res. 30 (2006) 687–693, https://doi.org/10.1016/j.leukres.2005.10.011.

[12] S.M. Bradley, Cytotoxic T lymphocyte responses in allogeneic radiation bone marrow chimeras. The chimeric host strictly dictates the self-repertoire of Ia-restricted T cells but not H-2K/D-restricted T cells, J. Exp. Med. 156 (1982) 1650–1664, https://doi.org/10.1084/jem.156.6.1650.

[13] R.J. Kowalski, A. Zeevi, R.B. Mannon, J.A. Britz, L.M. Carruth, Immunodiagnostics: evaluation of functional T-cell immunocompetence in whole blood independent of circulating cell numbers, J. Immunotoxicol. 4 (2007) 225–232, https://doi.org/10.1080/15476910701385638.

[14] S.H. Lim, W.N. Patton, S. Jobson, T.A. Gentle, M. Baynham, I.M. Franklin, B. J. Broughton, Mixed lymphocyte reactions do not predict severity of graft versus host disease (GVHD) in HLA-DR compatible, sibling bone marrow transplants. J. Clin. Pathol. 41 (1988) 1155–1157. ⟨https://doi.org/10.1136/jcp.41.11.1155⟩.

[15] A. van der Meer, I. Joosten, J. Ruiter, W. Allebes, A single [3H]thymidine-based limiting dilution analysis to determine HTLp and CTLp frequencies for bone marrow donor selection, Bone Marrow Transpl. 20 (1997) 149–155, https://doi.org/10.1038/sj.bmt.1700869.

[16] C. Möbs, T. Schmidt, Research techniques made simple: monitoring of T-cell subsets using the ELISPOT assay, J. Invest. Dermatol. 136 (2016) e55–e59, https://doi.org/10.1016/j.jid.2016.04.009.

[17] R.S. Abraham, Lymphocyte Activation, in: Man. Mol. Clin. Lab. Immunol., ASM Press, Washington, DC, USA, 2016, pp. 269–279, https://doi.org/10.1128/9781555818722.ch28.

[18] J. Pratschke, D. Dragun, I.A. Hauser, S. Horn, T.F. Mueller, P. Schemmer, F. Thaiss, Immunological risk assessment: The key to individualized immunosuppression after kidney transplantation, Transplant. Rev. 30 (2016) 77–84, https://doi.org/10.1016/j.trre.2016.02.002.

[19] M. Griebe, M. Daffertshofer, M. Stroick, M. Syren, P. Ahmad-Nejad, M. Neumaier, J. Backhaus, M.G. Hennerici, M. Fatar, Infrared spectroscopy: a new diagnostic tool in Alzheimer disease, Neurosci. Lett. 420 (2007) 29–33, https://doi.org/10.1016/j.neulet.2007.03.075.

[20] P. Carmona, M. Molina, M. Calero, F. Bermejo-Pareja, P. Martínez-Martín, A. Toledano, Discrimination analysis of blood plasma associated with Alzheimer's disease using vibrational spectroscopy, J. Alzheimers Dis. 34 (2013) 911–920. ⟨https://doi.org/10.3233/JAD-122041⟩.

[21] T.D. Payne, A.S. Moody, A.L. Wood, P.A. Pimiento, J.C. Elliott, B. Sharma, Raman spectroscopy and neuroscience: from fundamental understanding to disease diagnostics and imaging, Analyst 145 (2020) 3461–3480, https://doi.org/10.1039/D0AN00083C.

[22] M. Khanmohammadi, A.B. Garmarudi, M. Ramin, K. Ghasemi, Diagnosis of renal failure by infrared spectrometric analysis of human serum samples and soft independent modeling of class analogy, Microchem. J. 106 (2013) 67–72, https://doi.org/10.1016/j.microc.2012.05.006.

[23] V. Frochot, D. Bazin, E. Letavernier, C. Jouanneau, J.-P. Haymann, M. Daudon, Nephrotoxicity induced by drugs: The case of foscarnet and atazanavir—a SEM and μFTIR investigation, Comptes Rendus Chim. 19 (2016) 1565–1572, https://doi.org/10.1016/j.crci.2016.08.007.

[24] C. Paluszkiewicz, W.M. Kwiatek, A. Banaś, A. Kisiel, A. Marcelli, M. Piccinini, SR-FTIR spectroscopic preliminary findings of non-cancerous, cancerous, and hyperplastic human prostate tissues, Vib. Spectrosc. 43 (2007) 237–242, https://doi.org/10.1016/j.vibspec.2006.08.005.

[25] D. Sheng, X. Liu, W. Li, Y. Wang, X. Chen, X. Wang, Distinction of leukemia patients' and healthy persons' serum using FTIR spectroscopy, Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 101 (2013) 228–232, https://doi.org/10.1016/j.saa.2012.09.072.

[26] K. Liu, Q. Zhao, B. Li, X. Zhao, Raman spectroscopy: a novel technology for gastric cancer diagnosis, Front. Bioeng. Biotechnol. 10 (2022) 1–11, https://doi.org/10.3389/fbioe.2022.856591.

[27] N. Chaudhary, T.N.Q. Nguyen, D. Cullen, A.D. Meade, C. Wynne, Discrimination of immune cell activation using Raman micro-spectroscopy in an in-vitro & ex-vivo model, Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 248 (2021), 119118, https://doi.org/10.1016/j.saa.2020.119118.

[28] T. Ichimura, L. Chiu, K. Fujita, H. Machiyama, T. Yamaguchi, T.M. Watanabe, H. Fujita, Non-label immune cell state prediction using Raman spectroscopy, Sci. Rep. 6 (2016) 37562, https://doi.org/10.1038/srep37562.

[29] K.L. Brown, O.Y. Palyvoda, J.S. Thakur, S.L. Nehlsen-Cannarella, O.R. Fagoaga, S. A. Gruber, G.W. Auner, Raman spectroscopic differentiation of activated versus non-activated T lymphocytes: an in vitro study of an acute allograft rejection model, J. Immunol. Methods 340 (2009) 48–54, https://doi.org/10.1016/j.jim.2008.10.001.

[30] B.R. Cunha, L. Ramalhete, L.P. Fonseca, C.R.C. Calado, Fourier-Transform Mid-Infrared (FT-MIR) Spectroscopy in Biomedicine, in: Y. Tutar (Ed.), Essent. Tech. Med. Life Sci. A Guid. to Contemp. Methods Curr. Appl. Part II, Bentham Science Publishers, 2020. ⟨https://doi.org/10.2174/9789811464867120010 1⟩.

[31] L. Ramalhete, R. Araújo, C.R.C. Calado, Discriminating B and T-lymphocyte from its molecular profile acquired in a label-free and high-throughput method, Vib. Spectrosc. 111 (2020), 103177, https://doi.org/10.1016/j.vibspec.2020.103177.

[32] L. Ramalhete, C.R.C. Calado, Assessing the molecular fingerprint of T lymphocyte activation. 2019 IEEE 6th Port. Meet. Bioeng, IEEE,, 2019, pp. 1–3, https://doi.org/10.1109/ENBENG.2019.8692471.

[33] J.A. Timlin, L.E. Martin, C.R. Lyons, B. Hjelle, M.K. Alam, Dynamics of cellular activation as revealed by attenuated total reflectance infrared spectroscopy, Vib. Spectrosc. 50 (2009) 78–85, https://doi.org/10.1016/j.vibspec.2008.07.017.

[34] J.A. Timlin, L.E. Martin, C.R. Lyons, B. Hjelle, M.K. Alam, Dynamics of cellular activation as revealed by attenuated total reflectance infrared spectroscopy, Vib. Spectrosc. 50 (2009) 78–85, https://doi.org/10.1016/j.vibspec.2008.07.017.

[35] B.R. Wood, B. Tait, D. McNaughton, Fourier transform infrared spectroscopy as a method for monitoring the molecular dynamics of lymphocyte activation, Appl. Spectrosc. 54 (2000) 353–359, https://doi.org/10.1366/0003702001949627.

[36] K.H. Kwack, H.-W. Lee, Progranulin inhibits human T lymphocyte proliferation by inducing the formation of regulatory T lymphocytes, Mediat. Inflamm. 2017 (2017) 1–9, https://doi.org/10.1155/2017/7682083.

[37] P. Lasch, Spectral pre-processing for biomedical vibrational spectroscopy and microspectroscopic imaging, Chemom. Intell. Lab. Syst. 117 (2012) 100–114, https://doi.org/10.1016/j.chemolab.2012.03.011.

[38] C.L.M. Morais, F.S.L. Costa, K.M.G. Lima, Variable selection with a support vector machine for discriminating Cryptococcus fungal species based on ATR-FTIR spectroscopy, Anal. Methods 9 (2017) 2964–2970, https://doi.org/10.1039/C7AY00428A.

[39] S.H. Huang, Supervised feature selection: a tutorial, Artif. Intell. Res. 4 (2015), https://doi.org/10.5430/air.v4n2p22.

[40] Y. Subbannayya, M. Haug, S.M. Pinto, V. Mohanty, H.Z. Meås, T.H. Flo, T.S. K. Prasad, R.K. Kandasamy, The proteomic landscape of resting and activated CD4 + T cells reveal insights into cell differentiation and function, Int. J. Mol. Sci. 22 (2020) 275, https://doi.org/10.3390/ijms22010275.

[41] P.A. Szabo, H.M. Levitin, M. Miron, M.E. Snyder, T. Senda, J. Yuan, Y.L. Cheng, E. C. Bush, P. Dogra, P. Thapa, D.L. Farber, P.A. Sims, Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease, Nat. Commun. 10 (2019) 4706, https://doi.org/10.1038/s41467-019-12464-3.

[42] F.B. Stentz, A.E. Kitabchi, Transcriptome and proteome expression in activated human CD4 and CD8 T-lymphocytes, Biochem. Biophys. Res. Commun. 324 (2004) 692–696, https://doi.org/10.1016/j.bbrc.2004.09.113.

[43] A. Rohman, A. Windarsih, E. Lukitaningsih, M. Rafi, K. Betania, N.A. Fadzillah, The use of FTIR and Raman spectroscopy in combination with chemometrics for analysis of biomolecules in biomedical fluids: a review, Biomed. Spectrosc. Imaging 8 (2020) 55–71, https://doi.org/10.3233/BSI-200189.

[44] D. Schieffer, S. Naware, W. Bakun, A.K. Bamezai, Lipid raft-based membrane order is important for antigen-specific clonal expansion of CD4+ T lymphocytes, BMC Immunol. 15 (2014) 58, https://doi.org/10.1186/s12865-014-0058-8.

[45] M. Kołodziej, E. Kaznowska, S. Paszek, J. Cebulski, E. Barnaś, M. Cholewa, J. Vongsvivut, I. Zawlik, Characterisation of breast cancer molecular signature and treatment assessment with vibrational spectroscopy and chemometric approach, PLoS One 17 (2022), e0264347, https://doi.org/10.1371/journal.pone.0264347.

[46] J.K. Khalsa, A.S. Chawla, S.B. Prabhu, M. Vats, A. Dhar, G. Dev, N. Das, S. Mukherjee, S. Tanwar, H. Banerjee, J.M. Durdik, V. Bal, A. George, S. Rath, G. A. Arimbasseri, Functionally significant metabolic differences between B and T lymphocyte lineages, Immunology 158 (2019) 104–120, https://doi.org/10.1111/imm.13098.

[47] T.V. Shankey, R.P. Daniele, P.C. Nowell, Kinetics of mitogen-induced proliferation and differentiation of human peripheral blood B lymphocytes, Clin. Exp. Immunol. 44 (1981) 156–166.

[48] A. Singh, M.H. Spitzer, J.P. Joy, M. Kaileh, X. Qiu, G.P. Nolan, R. Sen, Postmitotic G1 phase survivin drives mitogen-independent cell division of B lymphocytes, Proc. Natl. Acad. Sci. 119 (2022) 1–9, https://doi.org/10.1073/pnas.2115567119.

[49] K.-Z. Liu, M. Xu, D.A. Scott, Biomolecular characterisation of leucocytes by infrared spectroscopy, Br. J. Haematol. 136 (2007) 713–722, https://doi.org/10.1111/j.1365-2141.2006.06474.x.

[50] A.A. Bunaciu, Ş. Fleschin, V.D. Hoang, H.Y. Aboul-Enein, Vibrational spectroscopy in body fluids analysis, Crit. Rev. Anal. Chem. 47 (2017) 67–75, https://doi.org/10.1080/10408347.2016.1209104.

[51] D. Howie, A. Ten Bokum, A.S. Necula, S.P. Cobbold, H. Waldmann, The role of lipid metabolism in T lymphocyte differentiation and survival, Front. Immunol. 8 (2018), https://doi.org/10.3389/fimmu.2017.01949.

[52] S.K. Pierce, Lipid rafts and B-cell activation, Nat. Rev. Immunol. 2 (2002) 96–105, https://doi.org/10.1038/nri726.

[53] E. Kokuina, M.C. Breff-Fonseca, C.A. Villegas-Valverde, I. Mora-Díaz, Normal values of T, B and NK lymphocyte subpopulations in peripheral blood of healthy Cuban adults, MEDICC Rev. 21 (2019) 16–21, https://doi.org/10.37757/MR2019.V21.N2-3.5.