

# Analyst

Accepted Manuscript



This article can be cited before page numbers have been issued, to do this please use: J. Xu, M. Potter, C. Tomas, J. Elson, K. Morten, J. Poulton, N. Wang, H. Jin, Z. Hou and W. Huang, *Analyst*, 2018, DOI: 10.1039/C8AN01437J.



This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the [author guidelines](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the ethical guidelines, outlined in our [author and reviewer resource centre](#), still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.

**A new approach to find biomarkers in chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) by single-cell Raman micro-spectroscopy**

*Jiabao Xu<sup>1</sup>, Michelle Potter<sup>2</sup>, Cara Tomas<sup>3</sup>, Jo Elson<sup>3</sup>, Karl J. Morten<sup>2</sup>, Joanna Poulton<sup>2</sup>, Ning Wang<sup>4</sup>, Hanqing Jin<sup>4</sup>, Zhaoxu Hou<sup>4</sup> and Wei E. Huang<sup>1\*</sup>*

1. Department of Engineering Science, University of Oxford, Begbroke Science Park, Woodstock Road, Oxford, OX5 1PF, United Kingdom.
2. Nuffield Department of Women’s and Reproductive Health, University of Oxford, the Women Centre, John Radcliffe Hospital, Headley Way, Headington, Oxford, OX3 9DU, United Kingdom
3. Institute of Cellular Medicine, Newcastle University, Newcastle Upon-Tyne, NE2 4HH, United Kingdom
4. Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, United Kingdom

\*Corresponding author: Wei E. Huang

[wei.huang@eng.ox.ac.uk](mailto:wei.huang@eng.ox.ac.uk)

Telephone: +44 (0)1865 283786, Fax: +44 (0)1865 3749

**Keywords:** Raman micro-spectroscopy, single cell, chronic fatigue syndrome, myalgic encephalomyelitis, fatigue, mitochondria, adenosine triphosphate (ATP), phenylalanine, biomarker, oxidative phosphorylation

Analyst Accepted Manuscript

## Abstract

Chronic fatigue syndrome (CFS), also called myalgic encephalomyelitis (ME), is a debilitating disorder characterized by physical and mental exhaustion. Mitochondrial and energetic dysfunction has been investigated in CFS patients due to a hallmark relationship with fatigue, however, no consistent conclusion has yet been achieved. Single-cell Raman spectra (SCRS) are label-free biochemical profiles, indicating phenotypic fingerprints of single cells. In this study, we applied a new approach using single-cell Raman microspectroscopy (SCRM) to examine  $\rho^0$  cells that lack mitochondrial DNA (mtDNA), and peripheral blood mononuclear cells (PBMCs) from CFS patients and healthy controls. The experimental results show that Raman bands associated with phenylalanine in  $\rho^0$  cells and CFS patient PBMCs were significantly higher than wild type model and healthy controls. Remarkably, an increase in intensities of Raman phenylalanine bands were also observed in CFS patients. As similar changes were observed in the  $\rho^0$  cell model with a known deficiency in the mitochondrial respiratory chain as well as in CFS patients, our results suggest that the increase in cellular phenylalanine may relate to mitochondrial/energetic dysfunction in both systems. Interestingly, phenylalanine can be used as a potential biomarker for diagnosis of CFS by SCRM. A machine learning classification model achieved an accuracy rate of 98% correctly assigning Raman spectra to either the CFS group or the control group. SCRM combined with machine learning algorithm therefore has the potential to become a diagnostic tool for CFS.

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

**Introduction**

Chronic fatigue syndrome (CFS), also called myalgic encephalomyelitis (ME), is a debilitating complicated disorder characterized by extreme fatigue that is not relieved by rest and other disabling symptoms including neurocognitive impairment and post-exertional malaise<sup>1</sup>. CFS affects a population prevalence of at least 0.2% in the UK and represents an extensive burden on the patients, their families and carers, and hence on the society<sup>2</sup>. However, it is a challenge to physicians and researchers and remains an incompletely characterized illness, in part due to its controversial definition, pathogenesis and diagnosis<sup>1</sup>. Therefore, finding potential biomarkers is of great importance for understanding the disease and developing targeted treatment<sup>3</sup>.

Mitochondria have been of great interest to CFS research due to emerging evidences of mitochondrial dysfunction as a putative biological mechanism for fatigue<sup>4</sup>. The underlying hypothesis is that fatigue and other accompanying symptoms are in part due to impaired energy metabolism at the cellular level which is largely controlled by mitochondria as an energy power plant and its ATP production<sup>5</sup>. The ATP profile of peripheral blood mononuclear cells (PBMCs) has been used as a diagnostic tool for CFS, however, controversial results were obtained. Several studies have shown a decreased level of ATP in patients' cohorts<sup>6-8</sup>, while others failed to detect any differences or have surprisingly found an elevated energy production in CFS patients<sup>9, 10</sup>.

Here, in an attempt of linking mitochondrial dysfunction and CFS pathogenesis, we sought to analyze single-cell Raman spectra (SCRS) and identify universal biomarkers in cells. A SCRS can be regarded as an intrinsic chemical fingerprint of a cell containing highly resolved Raman bands for major cellular building blocks such as proteins, amino acids, lipids and phospholipids, and carbohydrates<sup>11</sup>. Hence, SCRS are label-free biochemical profiles of individual cells reflecting physiological states and metabolic changes<sup>12-14</sup>. With rich and semi-quantitative metabolic information of single cells, SCRS can be used to indicate cellular metabolism and display disease-related biomarkers.

In this study, firstly, we used single-cell Raman micro-spectroscopy (SCRM) to obtain phenotypic profiles of human  $\rho^0$  cell lines completely depleted of mitochondrial DNA (mtDNA) to study the effect of severe mitochondrial dysfunctions. Secondly, we compared SCRS of PBMCs from 5 CFS patients and 5 healthy controls as a pilot study to find potential biomarkers for CFS diagnosis. The comparison of two results provides insights into the

association of mitochondrial dysfunction with the pathophysiology of CFS and demonstrates a coherent and consistent approach to diagnosis of the illness.

## Experimental

### Cell culture, WT cybrid generation and PicoGreen staining

143B  $\rho^0$  cell line lacking mtDNA was originally generated by King and Attardi<sup>15</sup>. It has been extensively characterized and used in many studies investigating the pathogenic effect of mtDNA mutations<sup>16</sup>. The  $\rho^0$  line was generated by treating the original 143B osteosarcoma line with low concentrations of ethidium bromide (EtBr) for several months. The EtBr intercalated with mtDNA resulting in its removal from the cell. The high glucose concentration in tissue culture media with uridine and pyruvate supplementation allowed  $\rho^0$  mtDNA-free cells to be propagated indefinitely. The wild-type (WT) line used in this study was generated by fusing 143B  $\rho^0$  cells with normal mtDNA from human platelets. Repopulating the  $\rho^0$  line is a better control system than the original 143B wild type cells as the long-term treatment with EtBr to generate the  $\rho^0$  line could have introduced genetic changes not found in the original 143B  $\rho^0$  cell line. Platelets do not contain nuclei and provide a simple system for the polyethylene glycol (PEG) induced fusion of mitochondria into the  $\rho^0$  cells. The protocol for PEG induced fusion was carried out as previously described<sup>17</sup>. The  $\rho^0$  cells were removed by culturing on Dulbecco's Modified Eagle Medium (DMEM) (25 mM) containing dialyzed serum lacking uridine and pyruvate. Under these conditions the  $\rho^0$  cells died, and the culture contained only new cybrids (cytoplasmic-hybrids) with WT mtDNA. Clones were generated by ringing growing colonies, removing cells with trypsin and re-plating.

Once cultures were established the presence of mtDNA was confirmed by PicoGreen staining. PicoGreen staining was carried out on live cells at 37 °C for 30 min in standard culture media using a 3  $\mu$ L/mL dilution of PicoGreen. Cells were examined using a FITC filter and a Leica DM IRE2 fluorescence microscope (Leica Ltd, UK). The mtDNA copy number of the  $\rho^0$  cells and the WT cells was also assessed by real-time PCR comparing nuclear and mtDNA markers. Results from the WT line being repopulated with mtDNA<sup>18</sup> were consistent with results from the original 143B line (data not shown).

### ATP measurements of WT and $\rho^0$ cells

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Cells were added with 50,000 or 75,000 cells per well in a 96-well white clear bottom plate, in which each well contained high concentration of glucose (25 mM) in DMEM. Cells were incubated at 37 °C allowing to adhere for 8 hr. Once cells had attached media was removed and replaced with 100 µL of fresh media containing various concentrations of glucose from 0 mM to 25 mM and incubated overnight at 37 °C and 5% CO<sub>2</sub>. ATP was measured using the Abcam ATP assay kit (Ab113849). Before use 1 vial of lyophilized substrate was reconstituted in 5 mL of substrate buffer. Reconstituted substrate and the detergent were equilibrated to room temperature. Once equilibrated 50 µL of detergent was added to the wells and placed on a shaker for 5 min at 700 rpm for cell lysis and ATP stabilizing. Next 50 µL of substrate was added and the plate was agitated for a further 5 min at 700 rpm. The plate was dark-adapted for at least 10 min before luminescence was read. ATP levels were calculated from the standard curve that was run alongside the samples.

**Ethics**

This work involving human subjects was done under ethics number MS-IDREC-C1-2015-173 and 12/NE/0146.

**Isolation of PBMCs from blood samples**

PBMCs were separated from the whole blood using the Histopaque method as described by Tomas et al<sup>19</sup>. Briefly, whole blood was centrifuged at 700 g for 10 min. Plasma was removed and blood was made up to its original volume with sterile phosphate buffered saline (PBS). Histopaque 1.077 was carefully layered on top of a layer of Histopaque 1.119 and the blood layer added on top of the Histopaque gradient. The tube was centrifuged at 700 g for 30 min with the break off. The PBMC layer was collected and washed with fresh PBS. Two types of cell morphology in the PBMC fraction were observed under the microscope with different cell sizes (5 – 7 µm and 15 – 20 µm). Only cells with larger sizes were selected in the Raman experiments after comparison of results (data not shown) to minimize heterogeneity generated from different cell types. Further experiments should be done to confirm the identity of the cells, which is now believed to be monocytes according to the cell population and cell size.

**Single-cell Raman spectra (SCRS) measurements**

Cells were fixed in 4% paraformaldehyde in PBS (v/v) at room temperature for 15 min. The fixed cells were washed twice with Milli-Q water and dropped onto a specifically-coated microscopic slide (which gives no background Raman signal) to be air-dried. SCRS were

acquired using an HR Evolution confocal Raman microscope (Horiba Jobin-Yvon, UK Ltd) equipped with a 532 nm neodymium-yttrium aluminum garnet laser. The laser power on cells was 5 mW, attenuated by neutral density filters. An objective with a magnification of 50× was used to focus single cells with a laser spot size of  $\sim 1 \mu\text{m}^2$  and Raman scattering was detected by a charge coupled device (CCD) cooled at  $-70^\circ\text{C}$ . The spectra were acquired in the range of 330 to 1900  $\text{cm}^{-1}$  with a 600 grooves/mm diffraction grating. A mapping mode was used to characterize single cells and the acquisition parameters were 5 s per spectrum, 10 spectra per cell and 20 - 30 single cells per each condition.

### Raman data pre-processing and analysis

All spectra were pre-processed by cosmic ray correction and polyline baseline fitting with LabSpec 6 (Horiba Scientific, France). Spectral normalization was done by vector normalization of the entire spectral region. Data analysis, statistics and visualization were done under an R environment using in-house scripts. Quantification of intracellular biomolecule was done by integrating corresponding Raman bands in SCRS. Bands associated with phenylalanine were integrated in the range of 993 – 1013  $\text{cm}^{-1}$  and 1022 – 1036  $\text{cm}^{-1}$  to quantify its intracellular concentration. Quantification results were represented as box plots. The rectangle in the box plots represents the second and third quartiles with a line inside representing the median. The lower and upper quartiles are drawn as lines outside the box. Sample means were compared by using Welch's two sample t-test for unequal variance.

### Machine learning classification model

A machine learning classification model was built based on SCRS of 5 CFS patients and 5 controls (80 and 126 SCRS respectively). Non-linear support vector machine (SVM) algorithm was used to build the model due to its best performance among all models tested (data not shown). In SVM, non-linear hyperplanes (Radial Basis Function kernel) were used to separate data in a high-dimensional space. Leave-one-out-cross-validation (LOOCV) was used to evaluate the performance of the model and performance measures were computed as sensitivity for each condition as well as an overall accuracy rate.

### Extracellular flux analysis

OXPHOS of PBMCs was determined using the Seahorse XFe96 extracellular flux analyzer as described by Tomas et al<sup>19</sup>. Briefly, PBMCs were seeded in quadruplicate on a 96-well microplate (Agilent Technologies), coated with poly-D-lysine, at a density of 500,000 cells per well. The plate was incubated overnight at  $37^\circ\text{C}$  and 5%  $\text{CO}_2$ . Oxygen consumption rate



(OCR) was recorded by the Seahorse XFe96 extracellular flux analyzer following the sequential addition of 1  $\mu$ M oligomycin, 3  $\mu$ M carbonyl cyanide-4-(trifluoromethoxy) phenylhydrazone (FCCP), and a combination of 0.5  $\mu$ M rotenone and 0.5  $\mu$ M antimycin A. Three basal measures of OCR were made before the addition of the first injection (oligomycin) and three measures of OCR were made after the injection of each of the compounds. OCR values were normalized for protein concentration using a bicinchoninic acid assay. Analysis was conducted using Wave software version 2.2.0.276 and Microsoft Excel 2013. Graphs were drawn using Graphpad Prism 7 software.

Results and discussion

Cells lacking mtDNA have distinct single-cell Raman spectra (SCRS)

PicoGreen can be used specifically to stain mitochondrial DNA (mtDNA) in living cells<sup>20</sup>. After incubation with PicoGreen,  $\rho^0$  cells repopulated with wild-type mtDNA (WT cells) showed brightly stained nuclei surrounded by numerous bright cytoplasmic speckles (Fig. 1A), whilst  $\rho^0$  cells only displayed brightly stained nuclei without surrounding speckling (Fig. 1B), confirming the depletion of mtDNA in  $\rho^0$  cells.

As mitochondria are the primary energy power plant of most eukaryotic cells and supply the cell with metabolic energy in the form of ATP<sup>21</sup>, we measured the ATP production of WT cells and  $\rho^0$  cells with either 50,000 or 75,000 cells (Fig. 2). Surprisingly, at both cell loadings,  $\rho^0$  cells produced a similar amount of ATP compared to WT cells, when the supplemented glucose concentrations were high (11 mM and 25 mM). On the other hand, ATP concentration in  $\rho^0$  cells was significantly lower than WT cells, when the glucose concentrations were low (0 mM to 5 mM). Our results suggest that, when glucose is sufficient,  $\rho^0$  cells can adapt to stimulate ATP production via non-mitochondrial glycolysis to compensate for having a poor mitochondrial respiratory chain. When the glucose is low and ATP production from glycolysis is restricted, the WT cell line is able to switch to a mitochondrial mode of ATP production using the respiratory chain via electron transport coupled phosphorylation (ETCP). However, the  $\rho^0$  cells are unable to use ETCP due to defected mitochondria<sup>22</sup>.

Here we applied single-cell Raman microspectroscopy (SCRM) to examine the WT cells and the  $\rho^0$  cells on high glucose and explain the differences in bioenergetic pathway despite a



similar ATP production in two lines. Fig. 3A shows SCRS of  $\rho^0$  and WT cells with 25 mM glucose, averaged from 30 single cells each. Variations from the single-cell measurements at each Raman peak position was represented by the shallow shade. Relatively low standard deviation was seen due to the small heterogeneity in the *in vitro* study of cell lines (Fig. 3A). Compared with WT cells,  $\rho^0$  cells have a distinct spectral pattern in the fingerprint region of their SCRS (600 – 1800  $\text{cm}^{-1}$ ), which typically can be used as a phenotypic fingerprint of cells that contains the most important biochemical information.

Unsupervised principal component analysis (PCA) was then used to reduce the high dimensionality of SCRS due to the presence of over 1500 Raman bands. A PCA plot along PC1 and PC2 illustrates two clearly separable clusters representing the WT cells and the  $\rho^0$  cells (Fig. 3B). As the differences between two cell types are most significant along PC1, a loading plot was made for feature extraction of important Raman bands (Fig. 3C). The most prominent feature was observed at 1003  $\text{cm}^{-1}$ , which can be assigned to aromatic ring vibrations of phenylalanine<sup>23</sup>. Other bands related to phenylalanine were also found to have high contribution including 1609 (phenyl ring bond-stretching vibrations of phenylalanine), 1030 (C-H/C-C in-phase motion of phenylalanine) and 618  $\text{cm}^{-1}$  (phenyl ring breathing vibrations of phenylalanine)<sup>23</sup>. All bands related to phenylalanine were higher in the  $\rho^0$  cells compared with the WT cells (Fig. 3A), which indicates an important role of the aromatic amino acid, phenylalanine, in the metabolism of cells with mitochondrial dysfunction.

Other differences include bands centered at 1658 (Amide I of proteins)<sup>24</sup> and 1440  $\text{cm}^{-1}$  ( $\text{CH}_2$  and  $\text{CH}_3$  deformation vibrations of lipids)<sup>25</sup> (Fig. 3C), both of which are higher in WT cells than in  $\rho^0$  cells (Fig. 3A). This suggests that, despite a higher accumulation of cellular phenylalanine,  $\rho^0$  cells have an overall reduced intracellular concentration of proteins and lipids due to an impaired metabolism. As phenylalanine alone might not be sufficient to characterize the pathomechanism, further biomarkers should be identified to simultaneously and more reliably identify mitochondrial dysfunction.

### PBMCs of CFS patients can be distinguished by a single Raman marker

Fig. 4A and 4B elucidate the SCRS of the peripheral blood mononuclear cells (PBMCs) from 5 CFS patients and 5 healthy controls, averaged either for each of the 10 individuals (20-30 SCRS per individual) (Fig. 4A) or for the 2 groups with 80 SCRS for the CFS group and 126 SCRS for the control group (Fig. 4B). As phenylalanine has been found to be a potential biomarker in  $\rho^0$  cells, it is hypothesized that it could be a suitable candidate for investigation

in CFS patients that are believed to have similar bioenergetic dysfunction. Out of 5 patients, 4 of them demonstrated Raman spectra with elevated phenylalanine band at 1003 cm<sup>-1</sup> (CFS 1–4) while the other one (CFS 5) exhibited similar intensity compared to the controls (Fig. 4A). By averaging 80 SCRS from the patients and 126 spectra from the controls, it is more visible that the phenylalanine band shows a marked increase in patients’ cells (Fig. 4B).

PCA of SCRS from 10 samples showed a considerable separation of the CFS group and the control group (Fig. 4C). While CFS 1-4 have smaller ellipsoids with little overlap with the controls, CFS 5 has a higher single-cell scattering and largely overlap with the control ellipsoids, which correlates with the observation of phenylalanine in Fig. 4A. In order to verify that the separation between groups in the PCA clustering was attributable to phenylalanine, we plotted the Raman wavenumber loadings along PC4 and PC6 which showed the largest separation between groups. Raman bands centered at 1003 and 1030 cm<sup>-1</sup> were identified for describing the maximum variances along both dimensions (Fig. 4D). The intracellular concentration of phenylalanine was semi-quantified by integrating the Raman bands at 1003 and 1030 cm<sup>-1</sup>, respectively (Fig. 5A and 5B). Both of the signature bands of phenylalanine were found to be significantly higher in the patients compared to the controls (p < 0.0001).

Previous research supports the possibility that the pathomechanism of CFS is linked with changes in amino acids. Reductions in the concentrations of certain amino acids including phenylalanine were reported in the serum and urine of CFS patients<sup>26-29</sup>. These findings support a possible metabolic defect in CFS patients related to amino acid metabolism by analyzing the metabolomics of the biofluids from the patients. Our work, to the best of our knowledge, is the first to report the changes in single peripheral blood cells of CFS patients. Our results suggest that the patients’ cells could be accumulating more amino acids (e.g. phenylalanine) which results in a reduction in the fluids observed in other studies. This may be due to an induction of secondary rescue mechanisms to intracellularly accumulate more amino acids and maintain a normal ATP production in the metabolically dysfunctional patients’ cells<sup>28</sup>. Furthermore, our finding demonstrates the impact of phenylalanine in both ρ<sup>0</sup> cells lack mtDNA and blood cells from CFS patients, which could also link mitochondrial defects with CFS. Nevertheless, how the raised phenylalanine levels in PBMCs relates to other tissues in the patients is still unclear. Raman approaches which could provide phenotypic spectra from deeper penetration into tissues would be very useful in future investigations.

Extracellular flux analysis of four CFS patients who were most disparate from the controls in the SCRS results (CFS 1 – 4) was performed in order to investigate whether a correlation exists between abnormalities in oxidative phosphorylation (OXPHOS) and the SCRS analysis when PBMCs were incubated in low and high glucose conditions. Two of the patients have very low OXPHOS in both the low and high conditions (CFS 2 and 4) and the other two samples had much higher OXPHOS profiles (CFS 1 and 3). While mitochondrial deficiency was markedly detected in two of the four patients by OXPHOS analysis, SCRS of PBMC cells were able to pick up the abnormalities in all of them. Nevertheless, the sample size of patients and controls is small considering the possibly high heterogeneity in CFS individuals. Our results serve as a pilot study to explore the potential of SCRM as a tool to identify biomarkers in CFS. A larger-scale study is currently under preparation to consolidate the results from the current work.

### **Machine learning models can successfully classify CFS patients with a high accuracy rate**

In Raman approaches, besides feature extraction to find informative biomarkers, classification based on samples' spectra is often desirable for diagnostic purposes. Machine learning approaches are usually suitable to resolve the complicated Raman dataset<sup>30</sup>. Non-linear support vector machine (SVM) is a supervised learning algorithm using higher dimensional spaces to separate different classes which are non-separable in linear classification. A classification model was trained by non-linear SVM to distinguish between the CFS group and the control group based on their SCRS (80 SCRS for the CFS group, 126 SCRS for the control group). Leave-one-out-cross-validation was used to evaluate the model by using any 205 spectra to train the model and then testing if the model can correctly classify the one spectrum being left out. The results are summarized in Table 1. This model successfully classified the CFS group with a 96.3% sensitivity (77 out of 80) and the control group with a 99.2% sensitivity (125 out of 126). The lower sensitivity observed in the patient group can be explained by the presence of high heterogeneity in the CFS patients. A total accuracy of 98.1% was achieved based on 206 Raman spectra. Notably, with an increasing sample size and number of Raman spectra to construct the reference database, classification models with better robustness can be achieved. As one Raman spectrum can be obtained within seconds, one patient sample that consists of multiple spectra and multiple cells can be characterized and classified within a few minutes, which implies an enormous potential and feasibility in clinical practice.

1

2

3

4

5 **Conclusions**

6

7 This study is to evaluate the feasibility of single-cell Raman analysis for the detection of

8 biomarkers related to mitochondrial dysfunction and CFS. Accordingly, we identified the

9 aromatic amino acid, phenylalanine, has an elevated intracellular concentration and can be

10 used as a potential biomarker in  $\rho^0$  cells lacking mitochondrial DNA, as well as in peripheral

11 blood mononuclear cells of CFS patients. Moreover, a machine learning model achieved an

12 accuracy of 98.1% correctly classifying patients and controls based on their Raman spectra.

13 The combination of Raman biomarkers and classification models might lead to improvements

14 in our understanding of CFS pathogenesis and have the potential to be used as a diagnostic

15 tool of CFS.

16

17 **Acknowledgement**

18 MP and CT were supported by the ME association. KM was supported by Diabetes UK. CT

19 was supported by the Medical Research Council. WEH acknowledges support from EPSRC

20 (EP/M002403/1 and EP/M02833X/1) and NERC (NE/M002934/1) in the UK.

21

22 **Conflicts of Interests**

23 There are no conflicts to declare.

24

25 **Statement**

26 All experiments were performed in accordance with the guidelines of Newcastle and Oxford

27 University for carrying out clinical studies. Blood samples were collected from patients and

28 controls following approval by the National Research Ethics Committee North East –

29 Newcastle & North Tyneside. Informed consents were obtained from human participants of

30 this study.

31

32 **References**

33 1. D. B. Fischer, A. H. William, A. C. Strauss, E. R. Unger, L. Jason, G. D. Marshall, Jr.

34 and J. D. Dimitrakoff, *Fatigue*, 2014, **2**, 93-109.

35 2. L. C. Nacul, E. M. Lacerda, D. Pheby, P. Campion, M. Molokhia, S. Fayyaz, J. C.

36 Leite, F. Poland, A. Howe and M. L. Drachler, *BMC Med.*, 2011, **9**, 91.

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

Analyst Accepted Manuscript

3. J. C. Edwards, S. McGrath, A. Baldwin, M. Livingstone and A. Kewley, *Fatigue*, 2016, **4**, 63-69.
4. K. Filler, D. Lyon, J. Bennett, N. McCain, R. Elswick, N. Lukkahatai and L. N. Saligan, *BBA Clin.*, 2014, **1**, 12-23.
5. D. Mitochondrial Medicine Society's Committee on, R. H. Haas, S. Parikh, M. J. Falk, R. P. Saneto, N. I. Wolf, N. Darin, L. J. Wong, B. H. Cohen and R. K. Naviaux, *Mol. Genet. Metab.*, 2008, **94**, 16-37.
6. S. Myhill, N. E. Booth and J. McLaren-Howard, *Int. J. Clin. Exp. Med.*, 2013, **6**, 1-15.
7. N. E. Booth, S. Myhill and J. McLaren-Howard, *Int. J. Clin. Exp. Med.*, 2012, **5**, 208-220.
8. J. Castro-Marrero, M. D. Cordero, N. Saez-Francas, C. Jimenez-Gutierrez, F. J. Aguilar-Montilla, L. Aliste and J. Alegre-Martin, *Antioxid. Redox. Signal.*, 2013, **19**, 1855-1860.
9. N. Lawson, C. H. Hsieh, D. March and X. Wang, *J. Nat. Sci.*, 2016, **2**.
10. R. C. Vermeulen, R. M. Kurk, F. C. Visser, W. Sluiter and H. R. Scholte, *J. Transl. Med.*, 2010, **8**, 93.
11. J. Xu, I. Webb, P. Poole and W. E. Huang, *Anal. Chem.*, 2017, **89**, 6336-6340.
12. W. E. Huang, R. I. Griffiths, I. P. Thompson, M. J. Bailey and A. S. Whiteley, *Anal. Chem.*, 2004, **76**, 4452-4458.
13. W. E. Huang, S. Ude and A. J. Spiers, *Microb. Ecol.*, 2007, **53**, 471-474.
14. Y. Wang, W. E. Huang, L. Cui and M. Wagner, *Curr. Opin. Biotechnol.*, 2016, **41**, 34-42.
15. M. P. King and G. Attardi, *Science*, 1989, **246**, 500-503.
16. H. M. Wilkins, S. M. Carl and R. H. Swerdlow, *Redox Biol.*, 2014, **2**, 619-631.
17. A. Chomyn, S. T. Lai, R. Shakeley, N. Bresolin, G. Scarlato and G. Attardi, *Am. J. Hum. Genet.*, 1994, **54**, 966-974.
18. K. J. Morten, N. Ashley, F. Wijburg, N. Hadzic, J. Parr, S. Jayawant, S. Adams, L. Bindoff, H. D. Bakker, G. Mieli-Vergani, M. Zeviani and J. Poulton, *Mitochondrion*, 2007, **7**, 386-395.
19. C. Tomas, A. Brown, V. Strassheim, J. L. Elson, J. Newton and P. Manning, *PLoS One*, 2017, **12**, e0186802.
20. N. Ashley, D. Harris and J. Poulton, *Exp. Cell Res.*, 2005, **303**, 432-446.
21. L. D. Osellame, T. S. Blacker and M. R. Duchon, *Best. Pract. Res. Clin. Endocrinol. Metab.*, 2012, **26**, 711-723.
22. M. Potter, E. Newport and K. J. Morten, *Biochem Soc Trans*, 2016, **44**, 1499-1505.
23. B. Hernández, F. Pflüger, S. G. Kruglik and M. Ghomi, *J. Raman Spectrosc.*, 2013, **44**, 827-833.
24. A. Rygula, K. Majzner, K. M. Marzec, A. Kaczor, M. Pilarczyk and M. Baranska, *J. Raman Spectrosc.*, 2013, **44**, 1061-1076.
25. K. Czamara, K. Majzner, M. Z. Pacia, K. Kochan, A. Kaczor and M. Baranska, *J. Raman Spectrosc.*, 2015, **46**, 4-20.
26. C. W. Armstrong, N. R. McGregor, J. R. Sheedy, I. Buttfield, H. L. Butt and P. R. Gooley, *Clin. Chim. Acta.*, 2012, **413**, 1525-1531.
27. C. W. Armstrong, N. R. McGregor, D. P. Lewis, H. L. Butt and P. R. Gooley, *Metabolomics*, 2015, **11**, 1626-1639.
28. O. Fluge, O. Mella, O. Bruland, K. Risa, S. E. Dyrstad, K. Alme, I. G. Rekeland, D. Sapkota, G. V. Rosland, A. Fossa, I. Ktoridou-Valen, S. Lunde, K. Sorland, K. Lien, I. Herder, H. Thurmer, M. E. Gotaas, K. A. Baranowska, L. M. Bohnen, C. Schafer, A. McCann, K. Sommerfelt, L. Helgeland, P. M. Ueland, O. Dahl and K. J. Tronstad, *JCI Insight*, 2016, **1**, e89376.

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

29. S. H. Niblett, K. E. King, R. H. Dunstan, P. Clifton-Bligh, L. A. Hoskin, T. K. Roberts, G. R. Fulcher, N. R. McGregor, J. C. Dunsmore, H. L. Butt, I. Klineberg and T. B. Rothkirch, *Exp. Biol. Med.*, 2016, **232**, 1041-1049.

30. H. J. Butler, L. Ashton, B. Bird, G. Cinque, K. Curtis, J. Dorney, K. Esmonde-White, N. J. Fullwood, B. Gardner, P. L. Martin-Hirsch, M. J. Walsh, M. R. McAinsh, N. Stone and F. L. Martin, *Nat. Protoc.*, 2016, **11**, 664-687.

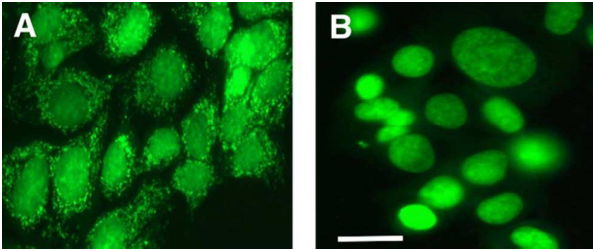
Analyst Accepted Manuscript

**Table 1** Evaluation of machine learning classification based on SCRS of PBMCs from CFS patients (n = 80) and healthy controls (n = 126).

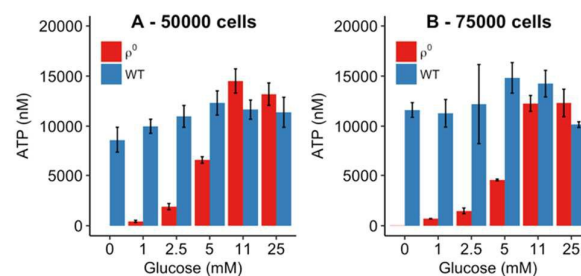
	Successful classification	Sensitivity
CFS group	77/80	96.3%
Control group	125/126	99.2%
Overall Accuracy	98.1%	



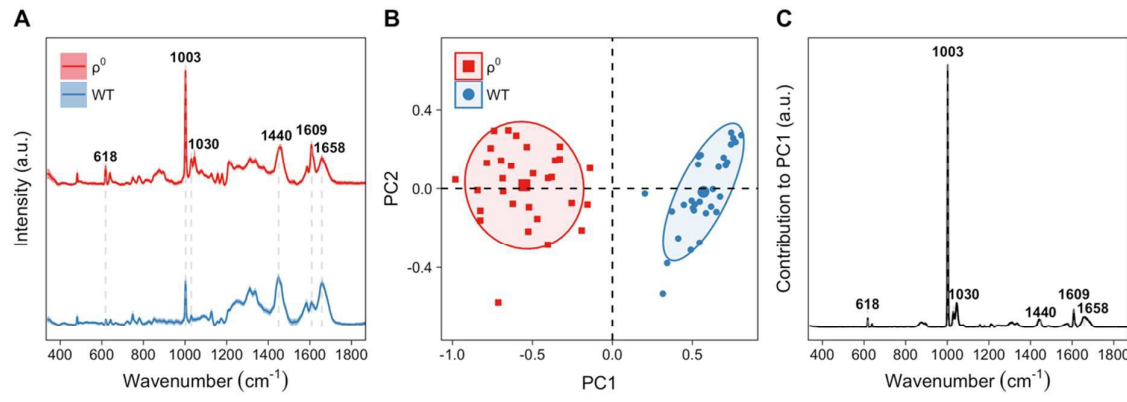
**Figures and Legends:**



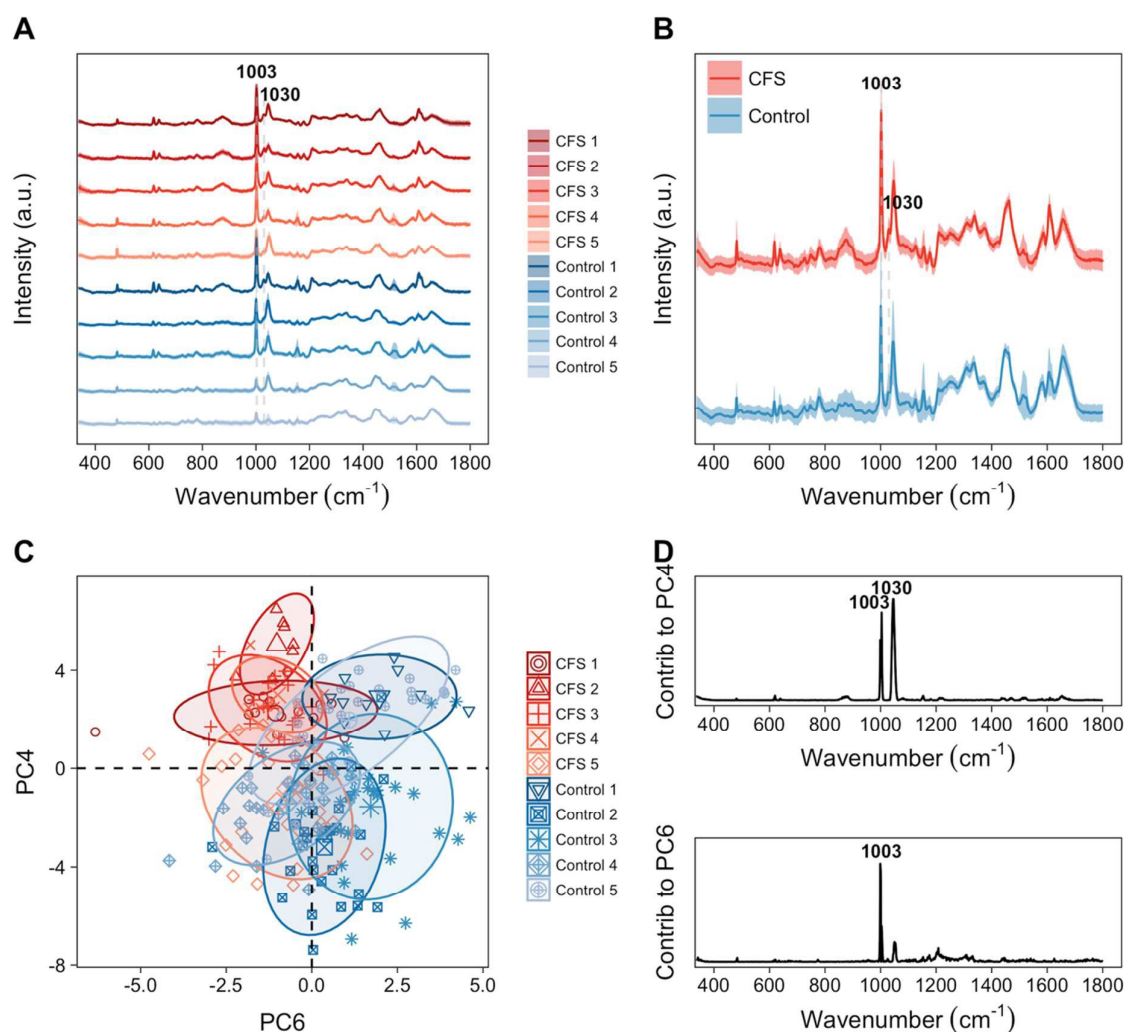
**Figure 1.** (A) Live WT cells stained with PicoGreen show bright nuclear and cytoplasmic staining. (B) Live  $\rho^0$  cells only show apparent nuclear staining (scale bar 20  $\mu\text{m}$ ).



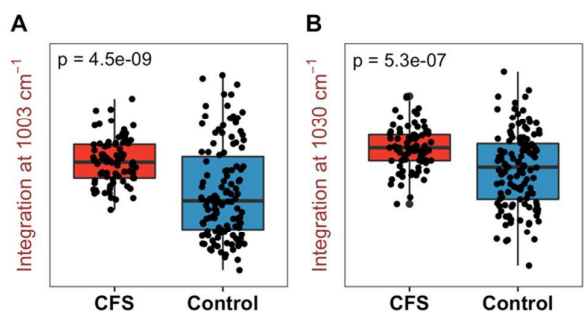
**Figure 2.** Extracellular ATP concentrations of  $\rho^0$  and WT cells were measured in (A) 50,000 cells and (B) 75,000 cells in the presence of different concentrations of glucose. Values are the mean  $\pm$  S.E. of three independent measurements.



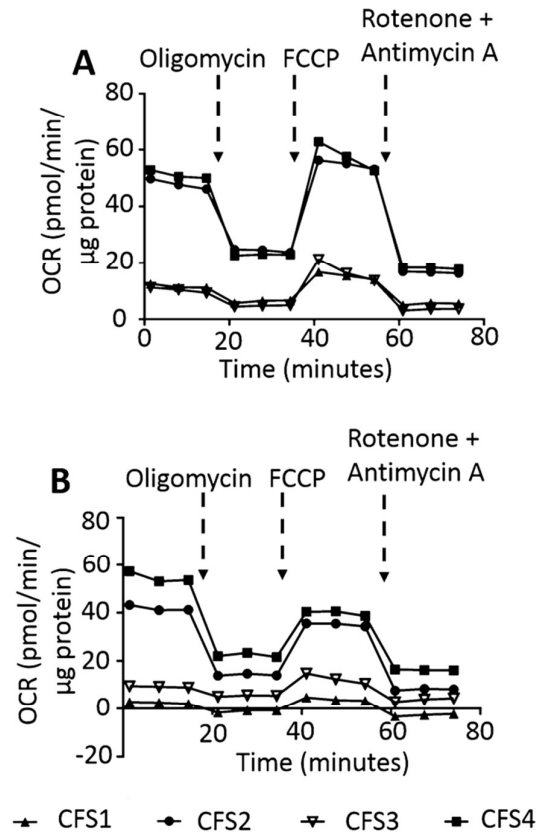
**Figure 3.** (A) Raman spectra of 143B WT cells and  $\rho^0$  cells, averaged from 30 single cells each, show distinct spectral patterns. The shaded area represents the standard deviation of the spread in single-cell measurements. (B) Unsupervised PCA separates the SCRS of WT cells and  $\rho^0$  cells in two clusters. (C) PCA loading plot shows the most significant Raman wavenumbers contributing to PC1 of the PCA.



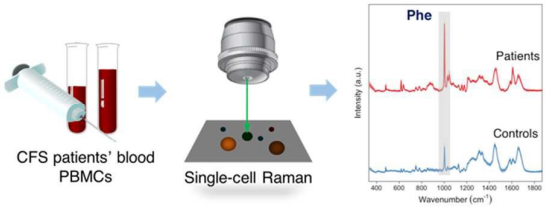
**Figure 4.** (A) Averaged SCRS of PBMCs from 5 CFS patients and 5 healthy controls, grouped by individuals. (B) Averaged SCRS of PBMCs from 5 CFS patients ( $n = 80$ ) and 5 healthy controls ( $n = 126$ ). The shaded area represents standard deviation from single-cell measurements. (C) PCA plot of SCRS of PBMCs from 5 CFS patients ( $n = 80$ ) and 5 healthy controls ( $n = 126$ ), grouped by individuals. (D) Raman wavenumber loading plots against contributions of relevant Raman band to PC4 and PC6 of the PCA.



**Figure 5.** Semi-quantification of intracellular phenylalanine in PBMCs of CFS group and control group by integrating phenylalanine Raman bands centered at (A) 1003 cm<sup>-1</sup> and (B) 1030 cm<sup>-1</sup>.



**Figure 6.** Traces from mitochondrial stress tests performed using PBMCs from 4 CFS patients incubated in (A) Low glucose (1 mM) and (B) high glucose (10 mM).



Single-cell Raman microspectroscopy to detect phenylalanine as a potential biomarker for mitochondrial dysfunction and chronic fatigue syndrome

Analyst Accepted Manuscript