SUPPLEMENTARY MATERIAL

Functional hyperspectral imaging captures subtle details of cell metabolism in olfactory neurosphere cells, disease-specific models of neurodegenerative disorders

Martin E. Gosnell[1,3], Ayad G. Anwer[1], Juan C. Cassano[2], Carolyn M. Sue[2] and Ewa M. Goldys[1*]

[1]ARC Centre of Excellence in Nanoscale BioPhotonics, Macquarie University, Sydney, 2109, New South Wales, Australia.

[2]Department of Neurogenetics, Kolling Institute, Royal North Shore Hospital/Northern Clinical School, University of Sydney, New South Wales, Australia.

[3] Quantitative Pty Ltd, ABN 17 165 684 186, www.quantitative.net.au, tel. +614 22 498 630

**Corresponding author ewa.goldys@mq.edu.au*

A: SUPPLEMENTARY MATERIALS AND METHODS

In sections S1-S4 below we provide additional details to the method presented in the main body of our paper, extended descriptions and protocols.

S1. Biological methods

S1.1. Preparation of Olfactory Neurospheres

In this study we used MELAS patient-derived olfactory-derived neurospheres (ONS) with varying levels of mutational generated from olfactory primary cultures. To prepare these ONS, nasal biopsies were collected into DMEM/F12 containing Dispase (3mg/ml) to separate the olfactory epithelial layer from the lamina propria. Olfactory cell cultures were then established in Dulbecco's modified Eagle's medium (DMEM; Gibco, Grand Island, NY) containing 10% fetal calf serum (FCS; Gibco), penicillin/streptomycin (P/S; 100 U/ml; Gibco). To generate olfactory neurospheres, cells were detached using TrypLE Select (Gibco), and plated on Poly-L-Lysine (0.38µg/cm^2; Sigma-Aldrich, St. Louis) coated 92 mm petri dishes (Nunc, Rochester, NY) at a cell density of 1x10^3 cells/cm^2 in serum-free media supplemented with both basic fibroblast growth factor (bFGF; 25 ng/ml; Chemicon, Millipore, Billerica, MA) epidermal growth factor (EGF; 50ng/ml; Chemicon, Millipore, Billerica, MA), insulin/transferring/selenium (ITS; 1g/l insulin, 0.55 g/l transferring, 0.67 mg/ml sodium selenite; Gibco) and P/S 100 U/ml. After reaching a size of approximately 100 µm in diameter, olfactory neurospheres (ONS) detached and were collected every second day thereafter. ONS were then expanded by passaging in DMEM, 10% FCS and 100 U/ml P/S (complete medium). ONS cells were then used for all experiments described. ONS cells were cultured in 92 mm petri dishes, in DMEM containing 10% FCS and 100 U/ml P/S, in a 37^{o}C incubator with 5% CO_2

S1.2. Galactose treatment

For galactose treatment, confluent plates were incubated with glucose and pyruvate free-DMEM (Gibco) with 10% FCS, 25 mM galactose, 1 mM pyruvate and 100 U/ml P/S. Cells were kept in this media for 4 days, with one media change at 2 days. At the end of the treatment cells were cultured in complete media and mutational load was analysed.

S1.3. Cell culture for imaging:

Cells used in this study originated from four patients, two MELAS patients with varying mutational load of 11% (patient 04019, referred to as 11% MELAS) and 44% (patient 04020, 44% MELAS) before, and after 4 days galactose treatment (referred to as 11% or 44% MELAS treated) and their healthy individuals representing age/gender matched controls: 09038 to 04019 and 03005 to 04020.

Before the experiment, all main dishes of cells were subcultured at the same time once they reached more than 80 % confluence. We present results for two subsequent passages for cells 090038+, 090038-, 04019+, 04019. We also report results for a single passage of cells 03005+ ,03005-, 04020+ ,04020 – cells. Here +/- indicates the presence or absence of galactose treatment These cells were imaged over 5 days at an approximately constant density of ~ 5000 cells/ cm^2). Each sample was prepared and imaged in duplicate. Altogether, over 1500 cells were imaged in this study.

S1.4. Mutational load

Mutational load of A3243G in mtDNA was analysed using last hot cycle PCR/RFLP. Using PCR, mtDNA was amplified using the forward primer 3243F (nucleotide positions 3116-3134) 5'-CCTCCCTGTTGTACGAAAGGAC-3' and the reverse primer 3243R (nucleotide positions 3353-3333) 5'-GCGATTAGAATGGGTACAATG-3' with an annealing temperature 50^{o}C. After 25 cycles, 3 μCi of α-phosphorous 32-labeled adenosine triphosphate was added to each reaction, and a last PCR cycle was performed. The resultant fragment size generated was 238bp in size, containing the m.3243A>G mutation site of interest, was digested with *HaeIII*. The digested PCR products were then electrophoresed on a 12% polyacrylamide gel (280V, 150mA, 2.5 hours) and placed in a Fuji imaging plate (IP, Fuji) overnight. The IP was then scanned in a FLA-3000 Fuji scanner (Fuji) the percentage of MELAS A3243G mutation was quantified.

S1.5. Citrate Synthase Assay for determination of Mitochondrial Mass

The exclusive mitochondrial matrix marker citrate synthase was assayed as a measure of mitochondrial mass using the Citrate Synthase Assay Kit (Sigma-Aldrich). Cells were collected and lysed using CellLytic M Cell Lysis Reagent (125 μL per 10^6-10^7 cells). After 15 minute incubation on a shaker, protein-containing supernatant was collected by centrifuging cells for 15 minutes at 14,000g and storing at 4^{o}C and assayed for protein. Citrate Synthase activity was measured in a 96 well plate as described by the supplier. Briefly, a mixture containing assay buffer, 30 mM Acetyl CoA solution and 10mM DNTB was aliquoted in each well, the sample was added and the absorbance was followed over 90 seconds in a BioRad Benchmark (BioRad) set at 412 nm on kinetic program to measure the baseline reaction. 10mM OAA solution was then added and the absorbance of the reaction

was followed for 90 seconds to measure total activity. To calculate total activity, the following formula was used, were *dil* is the dilution factor of the original sample, *V(ml)* is the reaction volume in ml, $V_{enz}(ml)$ is the volume of the enzyme sample in ml, ε is the extinction coefficient of TNB at 412nm which is 13.6, and *L(cm)* is the path length for absorbance measurement.

$$\text{units (µmole/ml/min)} = \frac{(\Delta A_{412})/\text{min} \times V(\text{ml}) \times \text{dil}}{\varepsilon^{mM} \times L(\text{cm}) \times V_{enz}(\text{ml})}$$

S1.6. Lactate Production

To determine lactate production by cells, after collection, cells were deprived from glucose by resuspending them in PBS for 1 hour at 37°C. Protein was assayed using the BCA protein assay kit (ThermoFischer). After glucose deprivation, cells were resuspended in PBS containing 10mM glucose at 0.25 mg/ml. Cells were shaken for 1 hour at 37°C, then spun and assayed for lactate using the Lactate assay kit (Roche, Basel, Switzerland). Briefly, 100 µL of cell sample was aliquoted into a 96 well plate with 100 µL of assay reaction mix, and after 20 minutes incubation at 37°C, absorbance was read at 570nm on the BioRad Benchmark spectrophotometer. Lactate concentration was calculated using the following formula:

Lactate mmol/mg/hr = (Abs@570nm -0.32) x 26.4

S1.7. ATP production assay

The ATP production was measured as follows.Briefly, cells were collected in PBS and protein was assayed for each cell line. Cells were then lysed in cell suspension buffer (150mM KCl, 25mM Tris HCl [pH 7.75], 2mM EDTA, 8mM K_2PO_4, 2mM KH_2PO_4, 0.1mM $MgCl_2$ and 0.1% BSA) at 1mg protein/ml. To initiate ATP synthesis, 750 µl of substrate buffer (10 mM malic acid, 10 mM pyruvate, 100 mM adenosine 5'-diphosphate, 0.004% digitonin and 0.15 mM P1, P3-di[adenosine 5'] pentaphosphate in cell suspension buffer) was then added to 250 µL of cell suspension (Reaction A). A 50 µL aliquot was removed (Reaction B), to which 450 µL of boiling quenching buffer (100 mM Tris HCl [pH 7.75] and 500 mM EDTA) was added. Reaction A was incubated at 37°C for 10 minutes, after which a 50 µL was again added to 450 µL (Reaction C). Reactions b and C were then assayed for ATP using the ATP Bioluminescence Assay kit CLS II (Roche). Briefly, ATP standards were made, and 100 µL of luciferase reagent was added to 100 µL of standard or quenched cell sample. Luciferase production was measured in a luminometer (Berthold, Detection system), and ATP was determined from the standard curve generated.

S1.8. Flow cytometry for analysis of mitochondrial parameters

Mitochondrial parameters, as described below, were acquired using a FACS Calibur flow cytometer (BD Biosciences Australia), and data was analysed using CellQuest Pro software. Mitochondrial mass in ONS was measured using Mitotracker Green, a dye that accumulates in mitochondria regardless of their energetic state. Approximately 1x 10^6 cells were stained with 5 nM Mitotracker Green in DMEM for 30 minutes at 37°C. Mitochondrial superoxide generation in ONS was measured using MitoSOX Red superoxide indicator. MitoSOX is

colourless, but in the presence of superoxide, it is oxidized and exhibits red fluorescence. Cells were stained with 1uM MitoSOX in Hanks buffer with calcium and magnesium for 30 minutes at 37^{o}C, washed twice in PBS and resuspended in PBS before analysis via flow cytometry. Mitochondrial ROS levels were expressed as a ratio of MitoSOX fluorescence intensity divided by mitochondrial mass (MitoSOX signal/Mitotracker Green signal). To determine mitochondrial transmembrane potential ($\Delta\psi_m$), cells were incubated with the cyanine dye $DiIC_1(5)$ (1,1'3,3,3',3'-hexamethylindodicarbocyanine iodide) which accumulates in mitochondria with active membrane potentials. Cells were stained 10nM $DiIC_1(5)$ in PBS for 30 minutes at 37^{o}C and then analysed.

S2. Imaging protocols

S2.1. Cell preparation

S2.1.1. Human Olfactory Cell culture protocol

Media: 1- Culturing media: DMEM, 10% FBS, 1% P/S; 2- Tryp LE; 3- PBS
All cell cultures are carried out in an incubator at 37^{o}C / 5% CO_2.

S2.1.2. Harvesting of adherent cells

Reagents: Tryple (GIBCO 12563/12604, PBS**,** Culturing media: for 500 mls: 445 ml of DMEM, 50 ml of FBS and 5 ml of P/S (10 000U/ml)

Procedure: Procedures are carried out using aseptic techniques.

1. Once the cell confluency is about 80% remove media from the plate by gentle aspiration.
2. Rinse the plates once with 2ml of PBS.
3. Add 2ml of Tryp LE to just cover the plate surface area.
4. Incubate for 5min at 37^{o}C in the incubator.
7. Dislodge the cell monolayer by gently tapping the plate and observe under an inverted microscope to confirm that the cells are off the flask surface.
8. If the majority of the cells are not in suspension place the flask back in the incubator for another couple of minutes and repeat tapping.
9. Using a plastic pipette, squirt the Tryp LE over the surface of the plate to remove the last of the cells and transfer the solution into a 15ml labeled falcon tube.
10. Using 2-3ml of PBS, repeat step 9 to ensure most of the cells are collected. Transfer the suspension into the falcon tube.
12. Centrifuge at 500g for 5min.
13. Remove the supernatant being careful not to disrupt the cell pellet. This can be done by gentle aspiration or pipetting.
14. Re-suspend the cell pellet in media depending on what the cells are needed for: Freezing, Splitting or Sphering.

Perform cell counts if needed.
We usually split one10cm plate into three10cm plates.

S2.1.3. Passaging cells:

1. Once the cell confluency is about 80-90% remove media from the plate by gentle aspiration.
2. Rinse the plates once with 2ml of DPBS.
3. Add 2ml of Tryp LE to just cover the plate surface area.
4. Incubate for 7-9 min at 37^{o}C in the incubator.
5. Dislodge the cell monolayer by gently tapping the plate and observe under an inverted microscope to confirm that the cells are off the flask surface.
6. If less than 90% of cells are trypsinized, place the flask back in the incubator for another couple of minutes and repeat tapping.
7. Using a sterile pipette, add double volume of the medium to neutralize the Tryp LE effect and squirt the Tryp LE and the medium over the surface of the plate to remove the last of the cells and transfer the solution into a 15ml labelled falcon tube.
8. Using 2-3ml of PBS, repeat step 9 to ensure most of the cells are collected. Transfer the suspension into the falcon tube.
9. Centrifuge at 500g for 5min.
10. Remove the supernatant being careful not to disrupt the cell pellet. This can be done by gentle aspiration or pipetting.
11. Re-suspend the cell pellet in media depending on what the cells are needed for: Freezing, Splitting or Sphering.
12. Perform cell counts and Trypan blue viability.
13.Cells are subcultured into 100mm by 15mm petri dishes to be used as the main source stock.

S2.1.4. Preparation of triplet imaging dishes

At the time of passaging we took 1000 ul of cell suspension containing 5000 cells and placed it into a 22 mm petri dish with low autofluorescence cover glass bottom, and incubated for 48hrs. The procedure was optimised so that we ideally have at least 4-6 cells per image at x40 on the spectral microscope. All dishes used have laser ablated grids (400 and 200 nm diameter) for cell relocation.

S2.1.5. Imaging

Each small petri dish is rinsed in Hanks balanced solution (HBS), prior to imaging. Using the phase contrast microscope at x 4 or x 10 a moderate to high density region of cells is located centred about a specific range of grid references; these will be noted down as the areas to be spectrally imaged. Selecting the image areas at low magnification and under phase contrast helps ensure imaging of cells at consistent density regions.

S2.1.6. Co-localization experiments

Cells (1000 ul) are seeded on small petri dishes with grids (400 and 200 nm diameter). Incubation is continued for 48 hours at 37 C , 5% CO2 and 90% humidity to allow cells to adhere on small petri dishes. Cells are rinsed with HBSS, covered with HBS, and then they located on grids.

S2.1.7. Viability testing

Viability testing was carried out at the beginning of experiment and at the end. The viability testing procedure include mixing 300 ul of cells suspension + 300 ul of Trypan blue dye and incubation for 5 minutes then calculating the percentage of live cells (bright green) .

S3. Fluorescence characterisation

For reference, pure NADH was measured at a concentration of ~1 μM in the same system as the cell images. The fluorophores identified with the endmembers were measured at approximately physiological concentrations in a Fluorolog Tau3 (Jobin-Yvon-Horiba).

S4. Data analysis

In this paper we use three different methods of data analysis: for the separation of cell classes, for the analysis of individual fluorophore content and for identifying subpopulations.

S4.1. Separation of cell classes

We employ a two stage analysis using Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). The use of a pre-processing PCA step is one way to avoid numerical problems [1] in the later LDA stage when calculating a within group sum of squares and cross products matrix which turns out to be singular if the input variables are linearly correlated. This PCA stage is unsupervised, while the LDA stage uses prior knowledge of class assignment through data labelling and attempts to find a projection that optimally separates the data based on second order statistics through the use of Fishers statistical distance criterion.

In our experiment we take N images of cells, segmented from the dark background, from all patients comprising M pixels at 11 different excitation wavelengths. This means that we measure the spectra comprising fluorescence intensities at each of these 11 wavelengths (vectors in an 11-dimensional space) for each pixel in each image.

We now bring together all data we have taken and produce a matrix with 11 rows comprising $N \times M$ fluorescence intensities at each pixel and at each wavelength (thus each column is an 11-point vector representing the spectrum at some pixel). The fluorescence intensity at each of these pixels at a single wavelength λ_i is treated as a random variable $X_{i,q}$, where $0<q<NxM$. The rows in this matrix are, generally, correlated (because, for example images taken at adjacent wavelengths vary very little). We will need to produce a de-correlated version of this data to avoid numerical problems in later stages of the analysis, described later. We now carry out a transformation of our data which leaves all the important information about cell differences but it removes data correlation between the rows. This is done in a standard way by using the 11x11 covariance matrix. This procedure is transforming the original 11-dimensional basis vectors $\boldsymbol{e_i}$ (representing spectra that are 0 for all wavelengths except for wavelength λ_i where the fluorescence intensity is 1) into specific new basis vectors $\boldsymbol{f_i}$ (i=1-11)that are rotated with respect to the original $\boldsymbol{e_i}$ (those with positive coefficients in the $\mathbf{e_i}$ bass represent spectra (or complex colours in layman terms). The original dataset (NxM vectors with 11 coordinates each representing fluorescence intensities at each pixel) is transformed in such a way that the images are now maximally de-correlated. This dataset still retains the meaning and the information content of the original images.

We take averages of these spectra (vectors) for all pixels in each cell, thus producing a representative spectrum (11-dimensional vector). Thus we analyse as many 11-dimensional vectors as we have cells. The mean spectrum of each cell is assigned a type or class label.

This data can now be projected into a space that optimally discriminates or separates the data based on this class assignment. The dimensionality of this new space and hence the number of new canonical variables is *N-1* where *N* is the number of unique classes assigned to the data.

S4.2. Obtaining fluorophore content from the hyperspectral data

The hyperspectral image set of 11 images taken at each of our 11 excitation wavelength is first processed by using a principal component analysis (PCA). This unsupervised method projects our data originally containing highly correlated variables into a space such that the newly formed variables are uncorrelated and selected to maximally explain data variance. When a set of correlated variables are thus transformed, the new set of variables have diminishing amounts variance, the first containing most and the last containing very little. We then take the top three components which in our case are shown to represent over 98% of data variance. The principal components or eigenvectors of spectral data can be thought of as artificial spectra. They do not necessarily have any physical relevance, rather they are vectors or directions that maximally describe data variance.

Such uncorrelated images are now subjected to unsupervised unmixing. In this procedure we express the spectra of all pixels as a linear combinations of individual component spectra, these components are called "endmembers". The procedure does not assume the number of these endmembers, it is obtained through the following approach. The cluster of pixel spectra is represented as points in an *N*-dimensional space. In this case the endmembers reside at the extreme points or vertices of this convex set or hull of points. The exact position of these vertices is then inferred geometrically by producing a fit to the data to a simplex. The coordinates of these vertices are then noted.

S4.3 Search for cell subpopulations.

In order to find the cell subpopulations we first analyse average cell abundances of the key endmembers (fluorophores) in each cell group. We then specify the number of subpopulations we wish to find in each data group (*K=2* in the presented analysis). Then for all groups within all variables, the data undergo an unsupervised mixture model. We used a "Figueiredo Jain" algorithm, which, as most of these unsupervised mixture model methods, is non deterministic. Each produced solution goes through a series of tests and selection filters, these reject solutions that do not meet the criteria so that, finally the highest scoring solution per data group is selected. The criteria are that the subpopulations cannot be trivial in size (a % size of the whole is specified, here we regard 30% as a minimum), the subpopulations cannot be overlapping (as these are less distinctive) or more specifically they must have a statistical separation greater than unity. Finally the subpopulations must pass a Kolomogrov-Smirnov test $p<0.05$. To apply these selection criteria the data must be broken into respective subpopulations. The mixture model returns only the subpopulation means and covariance so the data needs to be classified in one of the *K* subpopulations, and here we use a Naive Bayes

classifier. Subpopulations are then represented on box plots and on 2D and 3D scattergrams (the latter having the 3rd dimension indicated by colour)

S4.4. Description of selected technicalities in data analysis

S4.4.1. Parzen (kernel) smoothed histograms

Parzen estimations use a kernel function with specified bandwidth to produce a representation of an estimation of the underlying distribution from which the sample distribution may have been taken [2]. This is used here simply as a visual aid to highlight potential subpopulations, no statistics are derived from the smoothed histograms.

S4.4.2. Kolmogrov-Smirnov (KS), non-parametric test of different source distributions

This test is based on a minimum distance estimation used to test the equality of one dimensional probability distributions and can be used to compare two samples by the quantification of distance between sample empirical distribution functions. When testing two samples the null hypothesis states that the samples are drawn from the same distribution, continuity of distribution being the only requirement. It is a very general non-parametric test and sensitive to difference in the shape and position of the empirical distribution functions of the compared samples and is particularly useful where there is doubt regarding the nature of the source population. Completely general tests such as this are usually far stricter. See Reference[3].

S4.4.3. Statistical distance measurement (Fisher distance)

The univariate statistical distance metric *D* we use provides a measure of ratio of between-class distance to within-class dispersion. *D* is defined as:

$$D = \frac{|\mu_1 - \mu_2|^2}{{S_1}^2 + {S_2}^2} \qquad (1)$$

where S_i and μ_i are the mean and standard deviation of data from class *i*

This is a univariate form of the same criterion used in the LDA analysis and data projection.

The Fisher distance metric described gives a good measure of distance between to distributions based on the distance between the mean, but account to some degree for overlap or dispersion. It is not informative in the case where the means of two distributions are same but the population differ in variance.

S4.4.4. Principal Component Analysis (PCA)

Let X be the *n x p* data matrix of observed pixel spectrums by wavelength, where $n >> p$. The data covariance matrix *S* can be expressed by

$$S = X^T \left(I - \frac{1}{n} e_n e_n^T \right) X \qquad (2)$$

where e_ℓ denotes an *l*-th basis vector.

The principal component analysis of X may be obtained through the eigenvalue decomposition of nS.

$$S = V\Sigma^2 V^T \qquad (3)$$

where $\Sigma^2 = diag\left(\sigma_1^2\sigma_2^2, \dots, \sigma_p^2\right)$, or the variance vector of each variable and V represents the orthogonal matrix of eigenvectors forming the basis vectors of the new space onto which the data may be projected forming the new de-correlated variables Y.

$$Y = X \times V \qquad (4)$$

S4.4.5. Linear Discriminant Analysis (LDA)

Falling within the framework of supervised learning techniques, linear discriminant analysis looks to find basis axes which maximise the Maximum Fisher distance of the projected data. This criterion is the ratio of the between class scatter to the within class scatter.

Let us assume we have C pattern classes, $\omega_1, \omega_2, \dots, \omega_C$ in a pattern space of N dimensions, where ι_i is the number of training samples in class i; x_{ij} denotes the j-th training sample of class i; μ_i the mean vector of the samples in class i, where we have assumed the mean value is the expected value $\mu_i = E\left(x_{ij}|\omega_i\right)$ of the population of class i; and μ_o is the expected value (mean) of the entire data set. Then the between class scatter is defined as

$$S_b = \frac{1}{M}\sum_{i=1}^{C} \iota_i(\mu_i - \mu_o)(\mu_i - \mu_o)^T \qquad (5)$$

and the within-class scatter as

$$S_w = \frac{1}{M}\sum_{i=1}^{C}\sum_{j=1}^{\iota_i}\left(x_{ij} - \mu_i\right)\left(x_{ij} - \mu_i\right)^T \qquad (6)$$

The classes or groups of observations can be chosen arbitrarily (for example we could consider two classes: “cells from healthy controls ” and “cells from sick patients”).

An estimate of the class mean μ_i is obtained through calculating the class sample average, and similarly

$$m_i = \frac{1}{\iota_i}\sum_{j=1}^{\iota_i} x_{ij} \qquad (7)$$

the average of all samples is used to estimate the expected value (mean) of the entire data set.

$$m_o = \frac{1}{M}\sum_{i=1}^{C}\sum_{j=1}^{\iota_i} x_{ij} \qquad (8)$$

The Fisher criterion we wish to maximise is expressed as

$$J_F(w) = \frac{w^T S_b w}{w^T S_w w} \qquad (9)$$

The eigenvectors $\boldsymbol{w_1, w_2, \ldots w_d}$ of $S_b w = \lambda S_w w$ form the new coordinate system. The corresponding eigenvalues indicate the ratio of between/within variance and since the rank of S_b will be $C - 1$, then only that number of eigenvalues are non zero.

Thus we can be certain that projections of the input data into this space will provide a close to optimal class discrimination. There are now C-1 new canonical variables and for ease of visualization it is useful to choose three classes of data in the discriminant analysis so that the result may visualized in a 2D scatter plot. The pixel observations are further grouped according to their origin on a cell basis, and a statistical metric calculated on each such group and used to represent the cell as a single point.

B: SUPPLEMENTARY RESULTS

S5.1. Statistical distances for all cell classes

Supplementary Table 1 summarises the results of statistical distances for all the cell classes under investigation.

Supplementary Table 1. Summary of statistical distances for 11% and 44% MELAS

Maximum Fisher Statistical Distances using LDA between cell clasess		
MELAS LOADING	***11%***	***44%***
Healthy - MELAS	*1.87*	*1.98*
Healthy - Treated	*1.17*	*1.42*
MELAS - Treated	*1.05*	*8.89*

This Table should be viewed together with Fig. 1 b-g which compare MELAS cells and their healthy controls analysed by using histograms.

S5.2. Comparison of the spectra of five key fluorophores for 11% MELAS cells with endmembers identified in our images by unsupervised unmixing.

The corresponding figure for 44% MELAS is included in the main body of the paper (Fig. 2). The reader should note that although five fluorophores have been identified for 11% MELAS and 44% MELAS cells, these are not the same five. In 11% we observed FAD while in 44% we see red-shifted flavins. The remaining four fluorophores are the same.

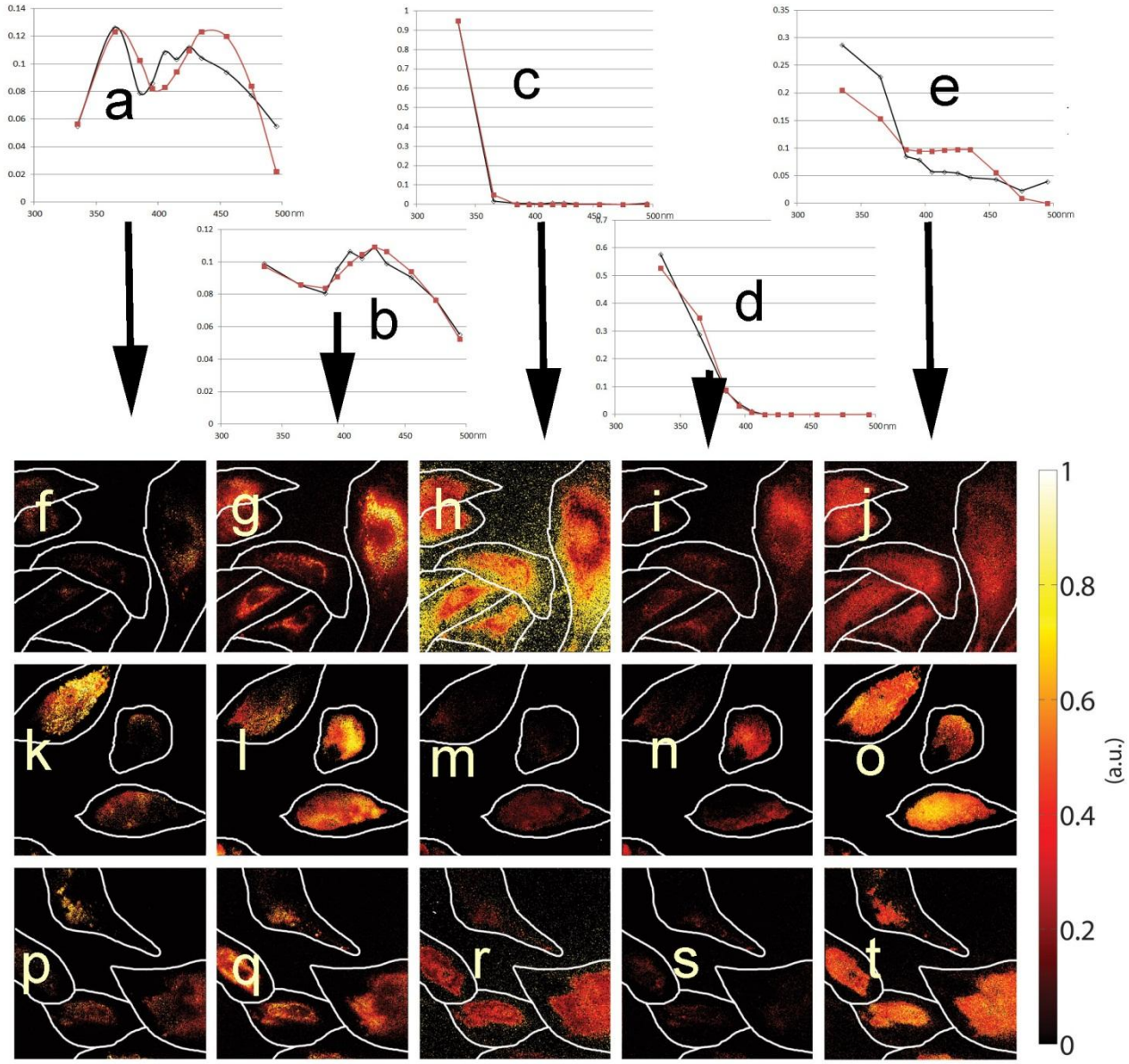


Fig. S1 a-e: Comparison of the spectra of five key fluorophores (red) in 11% MELAS cells with endmembers identified in our images by unsupervised unmixing (black). It shows that the endmembers agree well with (a) FAD, (b) A2E, (c) bound NADH, (d) free NADH and (e) lipofuscin. Images of cells showing the abundance of the five key fluorophores in healthy control (top row: f-j), 11% MELAS (middle row: k-o) and 11% MELAS treated patient (bottom row:p-t). Columns from left to right: f,k,p - FAD, g.l.q- A2E, h,m,r- bound NADH, i,n,s -free NADH, j,o,t - lipofuscin. The relative contribution is shown by the colour scale on the right. The outlines of cells have been marked for clarity. The colour scale reflects fluorophore concentrations.

S5.3. Identification of the organelle

Figure S2 shows MELAS cells labelled with Mitotracker, Lysotracker and the ER-Tracker; the corresponding image for healthy cells is shown in Figure S3 e

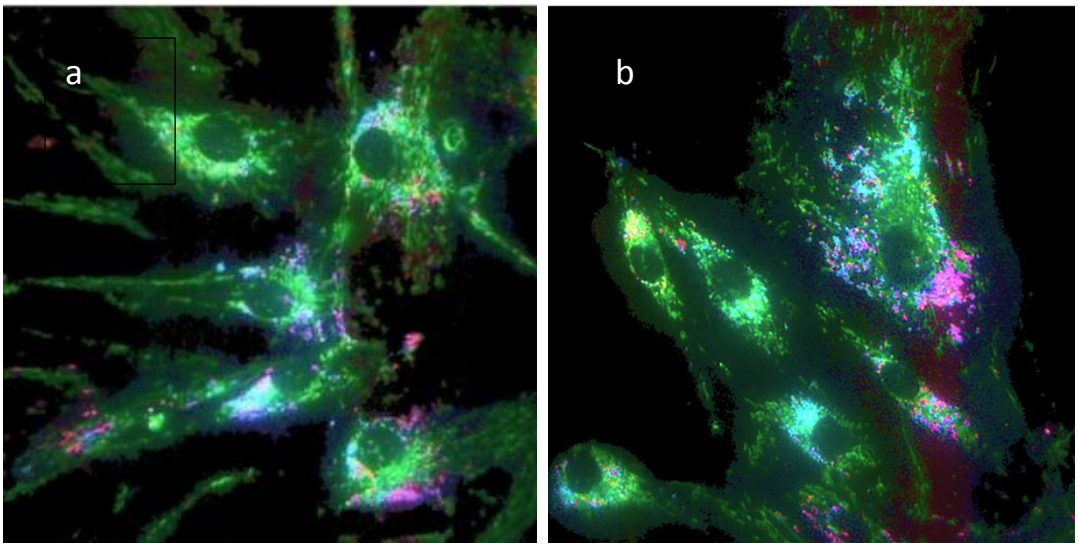


Figure S2. Identification of the organelle. (a) healthy control to 44% MELAS, (b) 44% MELAS cells, clearly we see the fibrous mitochondria highlighted in green, globular lysosomes in red and the network endoplasmic reticulum showing through in a diffuse blue. Stains used: Green Mitotracker, Red Lysotracker, Blue ER-Tracker.

S5.4. Colocalisation with stained organelle

In order to verify the identity of the endmenbers with specific fluorophores we carried out colocalisation of autofluorescence images for each of these fluorophores with corresponding images of cells stained with organelle stains. The organelle staining was carried out subsequent to autofluorescence imaging so that the latter could be recorded in a completely unadulterated state; we then compensated for the motion of the cell by using a custom-built adaptive colocalisation algorithm using linear affine parameters. As an example, Figure S3 shows the spectral component attributed to flavins (top left) in comparison with the Mitotracker stained image of the same cell taken 45 minutes later. The image at the bottom left shows the colocalisation of the two images at the top, and right bottom corner shows the flavin image colocalises with the motion-compensated Mitotracker image. This image clearly shows that the green stained mitochondria are almost 100% colocalised with flavin (yellow pixels, almost no green pixels), but a significant fraction of flavin pixels does not colocalise with the mitochondria in agreement with established knowledge about flavins and bound NADH.

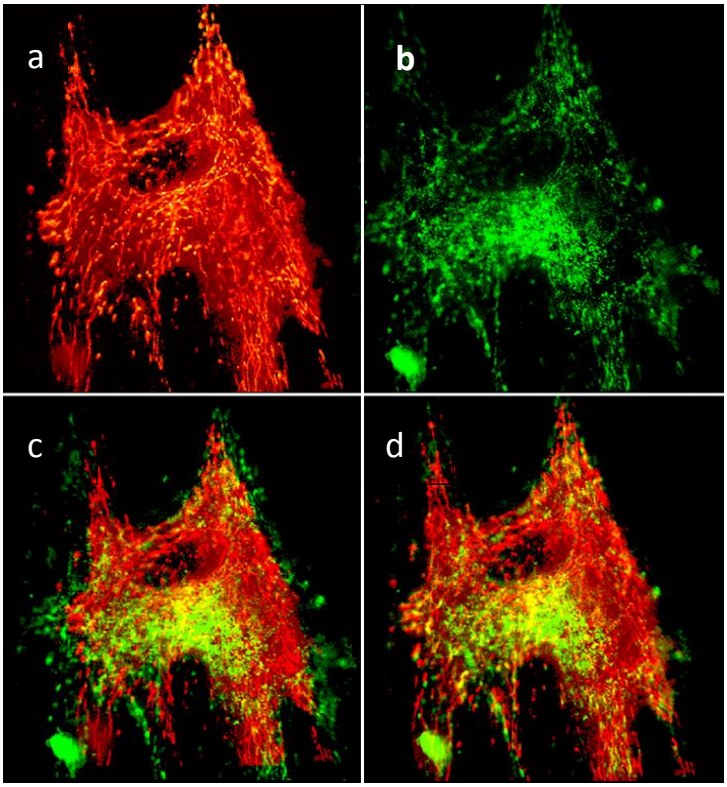


Figure S3. Colocalisation of autofluorescence images separated into individual fluorophores with stained images in a healthy control cell. (a) flavins in red , (b) Mitotracker in green, (c) superimposed flavins and Mitotracker images of the same cell taken 45 minutes later, colocalised pixels in yellow,(d) flavins colocalised with motion-adjusted Mitrotracker;

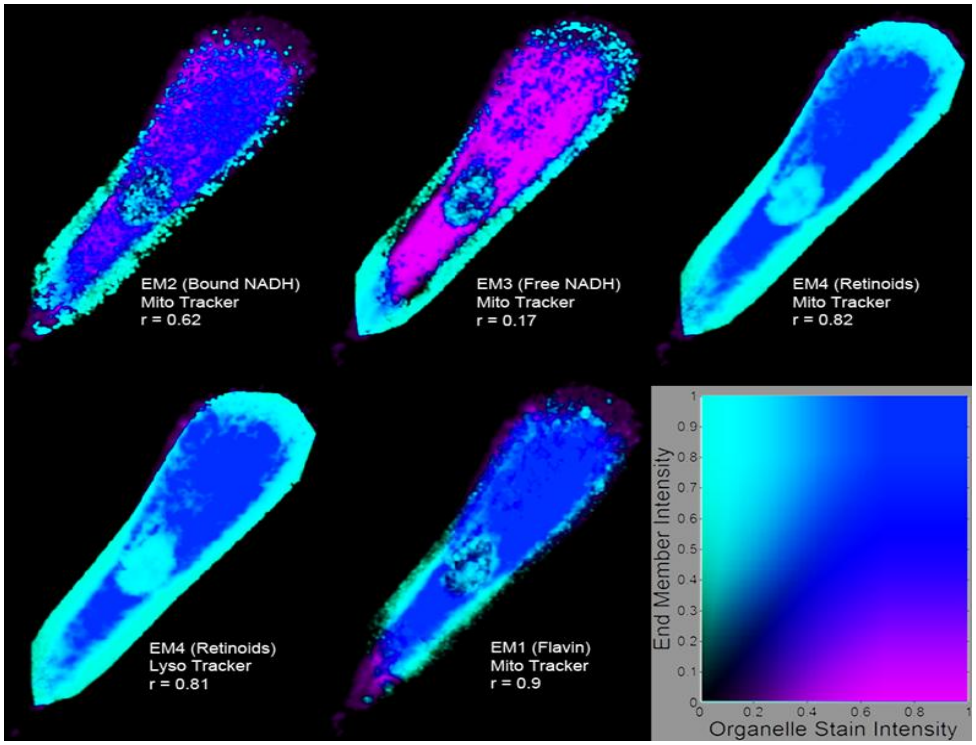


Figure S4. Image correlation between fluorescent components (endmembers) and organelle stains. Light blue: endmembers, purple: organelle stains (both as indicated in the figure), dark blue – colocalisation.

S5.5. Results of chemical quenching of fluorescence

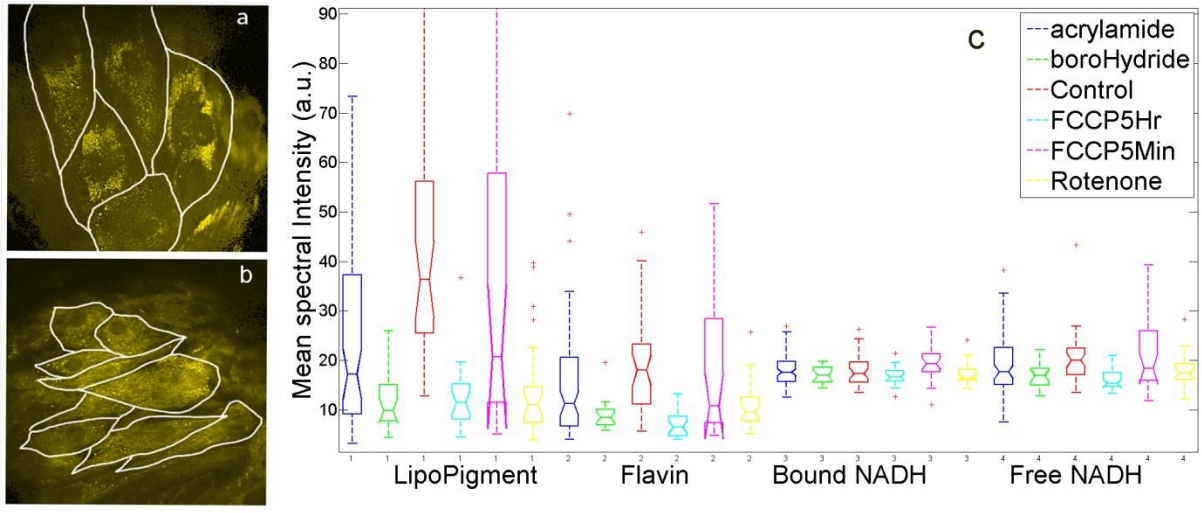


Figure S5. Chemical quenching of fluorophores. (a,b). Illustration of the effect of sodium borohydride in the flavin endmember spectral abundance. (a) control image (b) after application of borohydride. (c) the mean endmember spectral abundance data for all cells and all quenching conditions. Sodium borohydride is a reducing agent and it reduced the fluorescence of oxidised flavins. A reduction by a factor of ~ 2 is observed. Acrylamide is

expected to reduce flavin fluorescence, and this effect is observed here. It should have an insignificant effect on bound NADH and only a slight effect on bound NADH. FCCP should reduce membrane potential and lead to a drop of free NADH, which is immediately observed with an additional drop after 5 hours. FCCP stimulates NADH oxidation, accordingly free NADH intensity is reduced with bound NADH showing a similar but less pronounced effect. Rotenone, an NADH hydrogenase inhibitor appears to have a similar but weaker behaviour.

S5.6. Statistical analysis of fluorophore intensities and selected intensity ratios

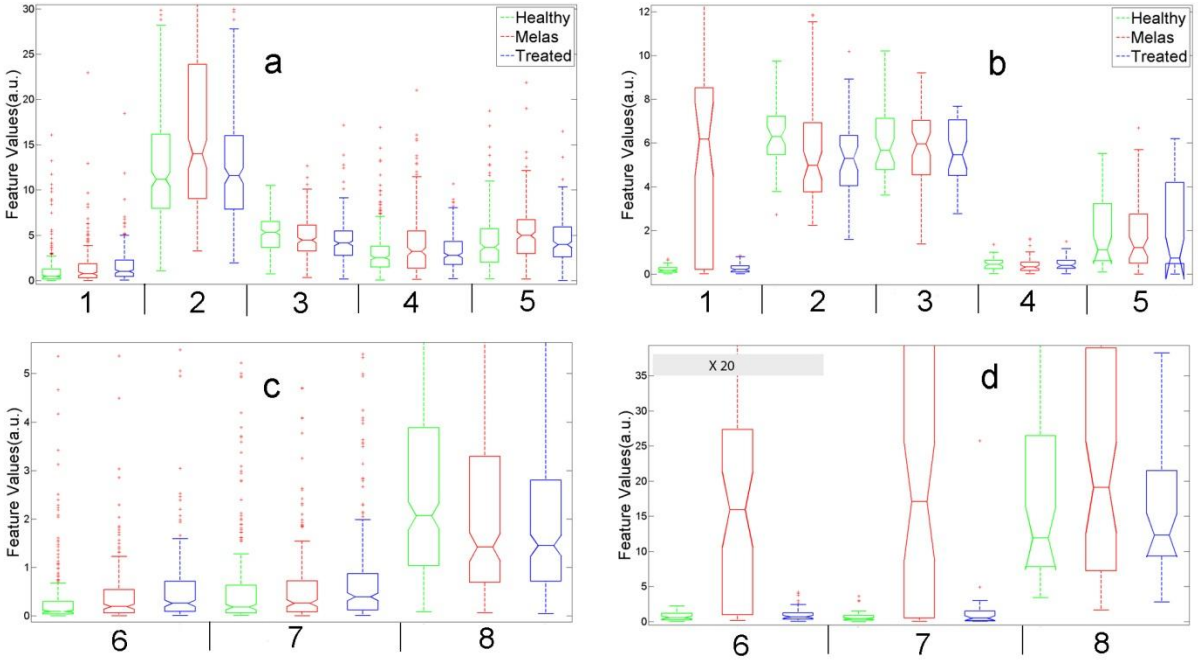
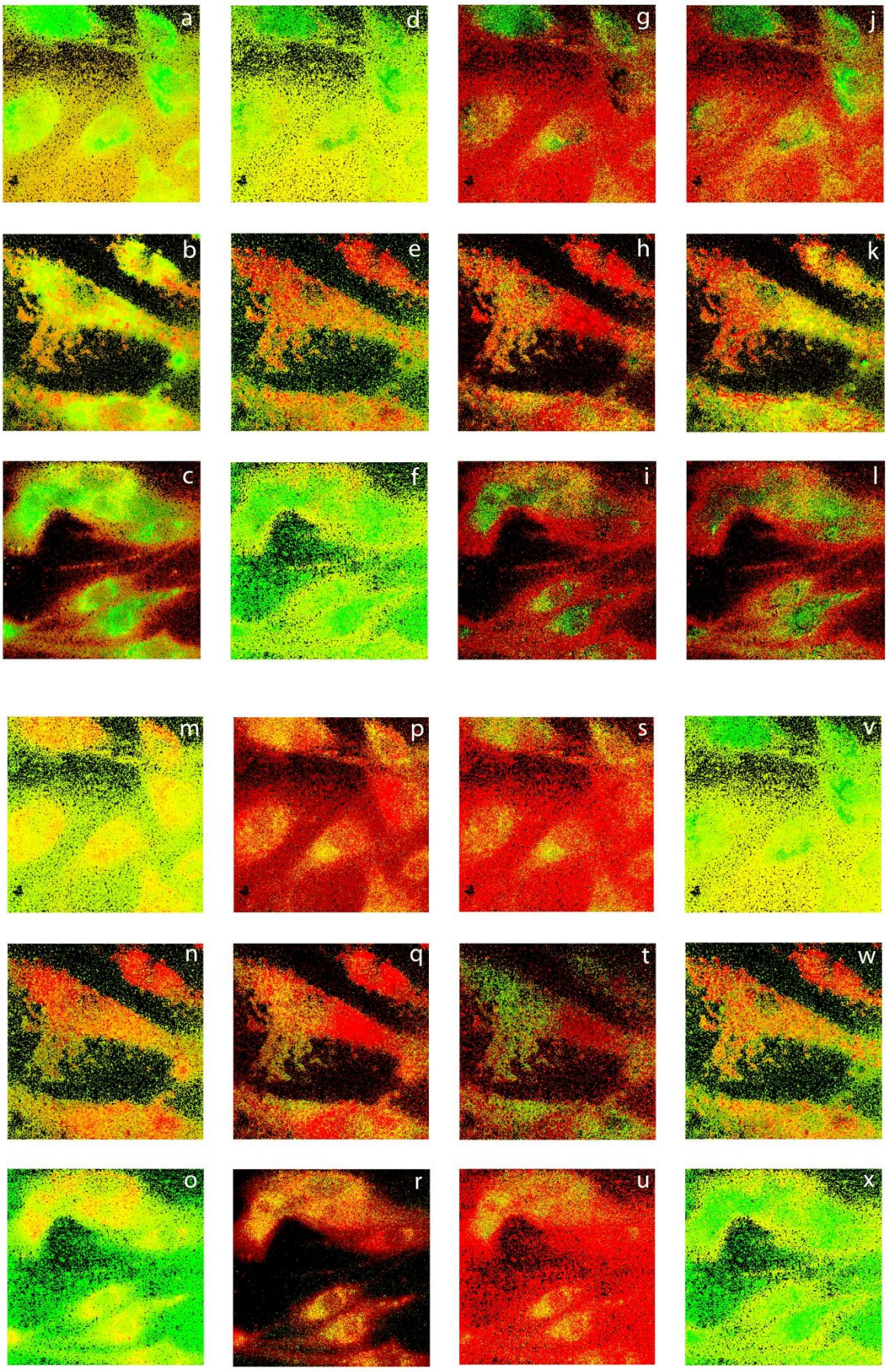


Figure S6 . (a, b) Average cell intensities for all fluorophore groups and 3 classes of cells a) 11% MELAS, (b) 44% MELAS . (1) FAD, (2) A2E, (3) bound NADH, (4) free NADH and (5) lipofuscin.- (c,d) ratio features for 11% MELAS (c) and 44% MELAS (d). (1) the ratio of FAD for 11% MELAS or flavins for 44% MELAS to bound NADH,(2) FAD for 11% MELAS or flavins for 44% MELAS to free NADH; (3) bound NADH to free NADH. Colours: healthy (green) MELAS (red) and MELAS treated (blue).(

S5.7. Colocalisation of fluorophores


a
d
g
j
b
e
h
k
c
f
i
l
m
p
s
v
n
q
t
w
o
r
u
x

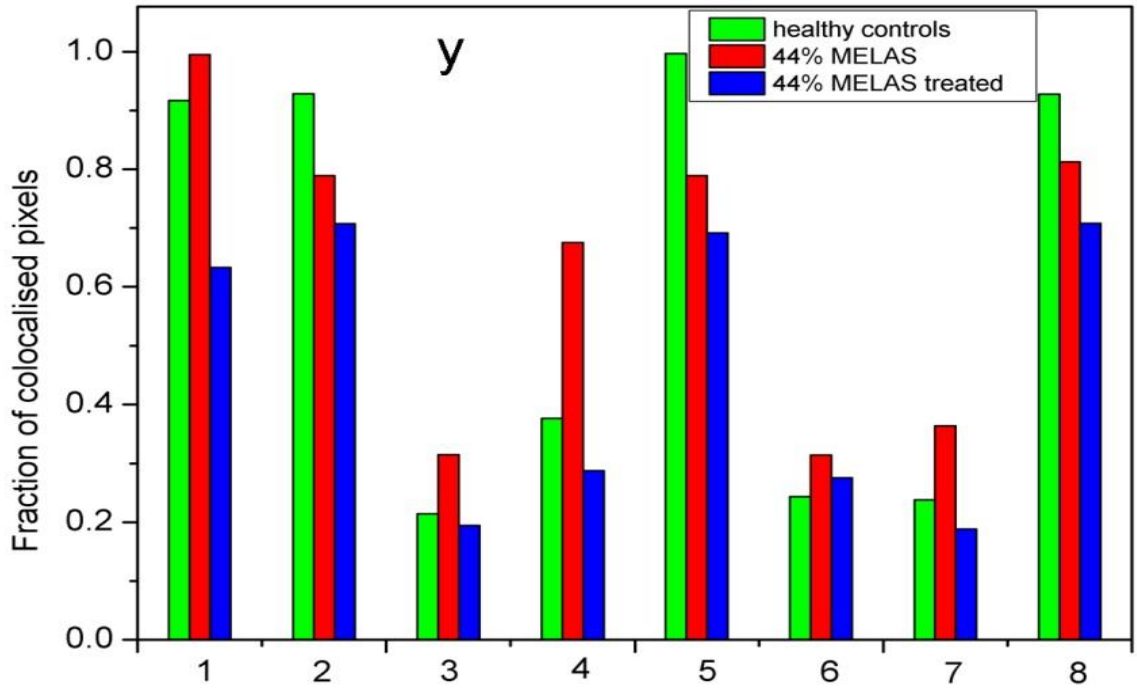


Figure S7. Co-localisation of selected fluorophores. Red and green –original fluorophore content, yellow – colocalised pixels. Row 1 and 4 healthy control cells matched to 44% MELAS, row 2 and 5 44% MELAS, row 3 and 6 -44% MELAS treated cells. (a-c) Flavin (red) and A2E (green) where we see both fluorophores are highly abundant in distinct perinuclear ring formation consistent with literature. (d-e) Flavin (red) and bound NADH (green) in healthy cells is highly co-localised in the mitochondria around the nucleus, however MELAS cells have significantly less of bound NADH presumably due to impaired complex 1 function. (g-h) Flavin (red) and free NADH (green), the free NADH is localised mostly in the cytosol so it co-localises less with flavins, however the MELAS cells with their more glycolytic metabolism appear to have more co-localised pixels. (j-i) Flavins(red) and lipofuscin (green) appear to co-localise slightly, perhaps due to disease-induced accumulation of lipofuscin in the lysosomes. (m-o) total retinoids (A2E+lipofuscin) (red) and bound NADH (green) appear to co-localise far less in MELAS 44% cells.(p-r) Total retinoids (red) and free NADH (green) appear to have similar levels. (s-u) Bound NADH (red) and free NADH (green) – both healthy and treated cells show significant bound NADH in the peripheral regions consistent with normal mitochondrial function, but there is some colocalisation in the cytosol as well, less in MELAS. (v-x) Flavins (red) and total NADH (green) appear to co-localise in healthy cells and also in treated especially closer to nucleus, but colocalisation is reduced in MELAS cells. (y) Co-localisation metrics for the above images:green-healthy,red-44% MELAS,blue: 44% MELAS treated; 1 flavins and A2E; 2-flavins and bound NADH; 3- flavins and free NADH; 4 – flavins and lipofuscin; 5-lipofuscin + A2E and bound NADH; 6 – lipofuscin + A2E and free NADH; 7- NADH bound and free NADH; 8 - flavins and total NADH.

S5.8. Statistical analysis of the existence of subpopulations for all investigated cells

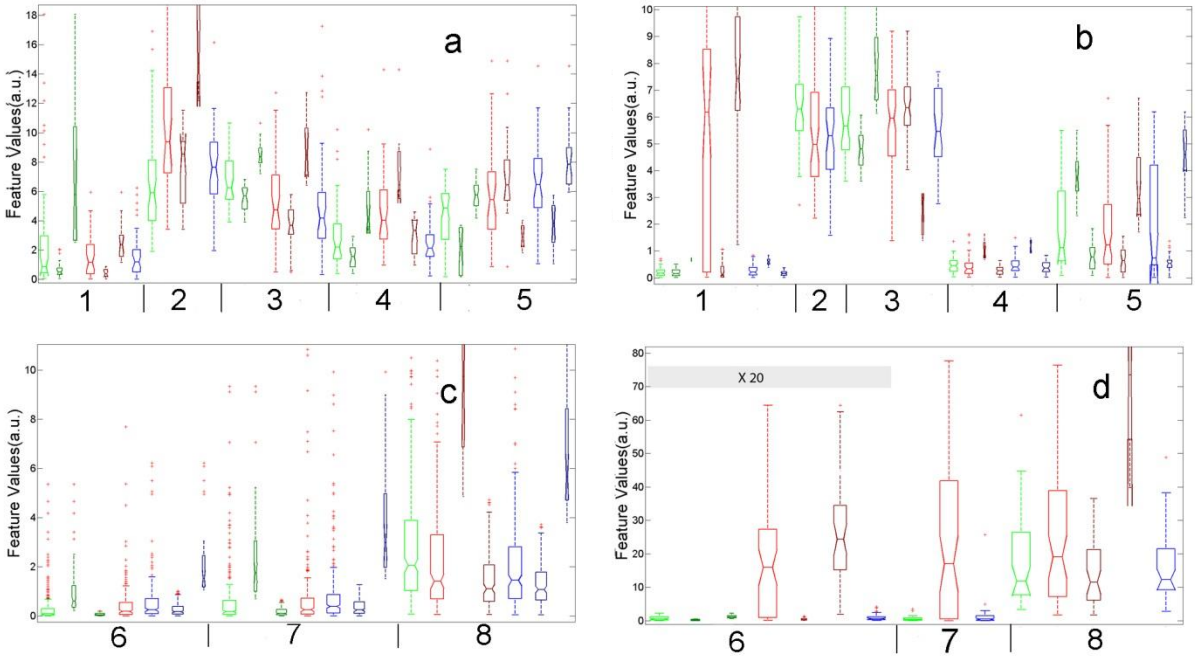


Figure S8 . Subpopulations for all features shown in Fig S6. (a, b) Average cell intensities for all fluorophore groups and 3 classes of cells a) 11% MELAS, (b) 44% MELAS . (1) FAD, (2) A2E, (3) bound NADH, (4) free NADH and (5) lipofuscin.- (c,d) ratio features for 11% MELAS (c) and 44% MELAS (d). (6) the ratio of FAD for 11% MELAS or flavins for 44% MELAS to bound NADH;(7) FAD for 11% MELAS or flavins for 44% MELAS to free NADH; (8) bound NADH to free NADH. Colours: healthy (green) MELAS (red) and MELAS treated (blue).Dark green: subpopulations for healthy cells, dark red – subpopulation for MELAS cells, dark blue: subpopulations for treated cells.

S5.9. Biochemical analysis

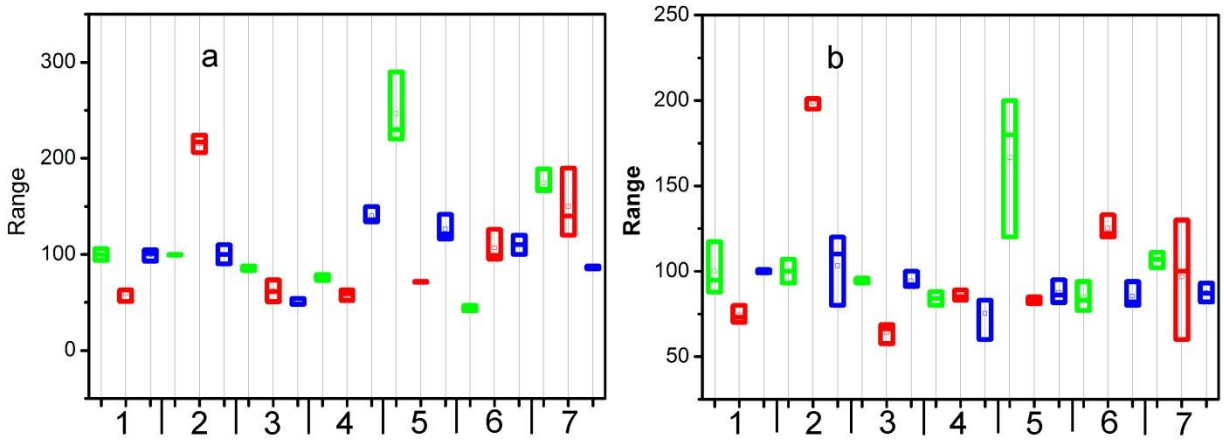


Figure S9 . Biochemical characterisation of the examined cells. a) Data for healthy control, 11% MELAS and 11% MELAS treated; b) healthy control, 44% MELAS, 44% MELAS treated Boxplots represent triplicate measurements of the following quantities: 1 – ATP final %, 2 – lactate x 100, 3- MGT mass, 4 –Citrate synthase, 5 – MTMP, 6- SOX, 7-Cell viability. Green-healthy, red MELAS, blue – treated.

C: SUPPLEMENTARY DISCUSSION

S6.1 Hyperspectral imaging enables biochemical and metabolic imaging of cells

In addition to cell classifications, our hyperspectral data set provides detailed insights into cell biochemistry. The fluorescence of unlabelled cells is produced by endogenous fluorophores including NADH and NAD(P)H, riboflavin and flavin-coenzymes including FAD and FMN, retinol and other retinoids including components of lipofuscin, pyridoxine, vitamin B12, vitamin D, ceroid, cytochromes, vitamin K, tryptophan, tyrosine, phenylalanine, kynurenine and porphyrins, as well as the complexes of these fluorescent molecules with proteins [4-21]. Among these fluorophores protein-bound and free NAD(P)H produces a significant part of the fluorescence signal following excitation in the 330-390 nm wavelength range. Flavins are prominent in the longer wavelength range (excitation peak at 360 nm, emission in the range 420-500 nm) with part of their spectra well separated from the NAD(P)H group. Retinoids, especially retinol also tend to be highly visible [4,5], in particular lipofuscin and A2E [6] . Previous studies of autofluorescence in cells indicate the absence of fluorophore quenching, this means that fluorophore concentration is proportional to fluorescence intensity. In particular, a linear correlation has been shown between total NADH content determined biochemically and NADH fluorescence, [7,8]. The total NADH concentration is typically in 100uM range [9]. Other fluorophores listed here are less abundant, for example flavins are present in cells at ~ 50 uM [10] and the most abundant retinoid, retinol at 1-2 uM concentrations while retinoic acid, anhydroretinol and retinal have lower concentrations; all are below the onset of concentration quenching [11] (the possible exception to this may be lipofuscin which can form highly concentrated granules).

S6.2. Colocalisation of key fluorophores

We carried out a colocalisation analysis between different endmembers which is consistent with the assignment of the fluorophore groups observed in this study (see Figure S7. To this aim, single fluorophore images were artificially assigned red and green colour and superimposed (Fig. S 7 a-x) with yellow colour indicating co-localisation. We also calculated the co-occurrence (the ratio of yellow to total pixels) (Fig S7 y) which indicates that: (a) flavins and bound NADH are highly co-localised. This is as expected because:(a) flavins, in particular FAD, are present in the mitochondria; (b) flavins generally do not strongly colocalise with free NADH; (c) flavins and retinoinds, in healthy cells are well colocalised over the mitochondrial region, in agreement with literature [12]; (d) bound NADH and free NADH in healthy cells are very tightly compartmentalised with no overlap. This is as expected, because free and bound NADH are localised in the cytosol and in the mitochondria, respectively [18]; (e) in healthy cells any free NADH is clearly separated from the retinoids. This is expected as retinoids are localised mostly in the lysosomes and in the mitochondria. Thus the fluorophore colocalisation results in healthy cells are well explained by standard cell biology. However, pathological metabolism in diseased cells presents a more complicated picture, due to a skewed activity of key enzymes, much of this is reversed by the galactose treatment which largely brings back the metabolism into the healthy range (see Fig 3 and Fig S6).

S6.3. Trends fluorophore content and their ratios between healthy, MELAS and treated cells in comparison with traditional biochemistry.

Here, we discuss the trends observed in various cellular characteristics between healthy, MELAS and MELAS treated cells (for 44% mutational load only). We first discuss the fluorophore content (Fig. 2). It shows that the 44% MELAS cells show significantly more flavin fluorescence than the controls, which was brought back to the healthy level by galactose (Fig 2 i). This can be attributed to reactive oxygen species which impair enzymatic complexes of the electron chain leading to oxidation of flavoproteins linked to these enzymes. Alternatively, increased flavins fluorescence may be due to a deficiency in succinate dehydrogenase (SDH), a complex 2 enzyme of the electron transport chain which serves a critical role in the citric acid cycle of reducing fluorescent flavin (FAD) into non fluorescent $FADH_2$. We hypothesise that the 44% MELAS patient had a deficiency (reduced activity or amount or both) in that enzyme caused by mutation leading to the buildup of fluorescent FAD.

Furthermore, we found that the MELAS cells had significantly less bound NADH than the control cells. Following the galactose treatment the bound NADH has increased significantly, but not quite to the healthy level. Conversely, free NADH was significantly higher in MELAS compared to the healthy control. It was brought down by galactose treatment, but not quite to the healthy level. This trend is consistent with increased glycolysis in MELAS which is independently verified by biochemical analyses, including increased lactate and reduced ATP levels as a result of glycolysis, reduced ATP levels (Fig S9).

We now discuss cellular images of fluorophore content. We observe that bound NADH in MELAS cells is located all around the cell while a weakly defined perinuclear ring is most

clearly observed in some MELAS cells (Fig. 2 j). This aligns well with the measurements of MGT mass (Fig S9) which was lower in MELAS compared to healthy and treated cells reflecting that functioning mitochondria in MELAS cells occupy a reduced volume compared to healthy and treated cells. Fig. 2 i also shows that flavins in MELAS cells form clear filamentous structures which colocalise well with the mitochondria (Fig. S4), but significant levels of uncolocalised flavins are also present elsewhere. This agrees with established understanding that a significant proportion of cell flavins are produced in the mitochondria and with the hypothesis of deficient SDH. Free NADH is noticeable in the nuclei especially in MELAS cells (Fig 2 k) similar to that observed in Ref [22] where it was attributed to chromatin modification. The cellular retinoids are also observed, these form perinuclear rings (Fig. 2 b). The rings are consistent with mitochondrial localisation of retinol which is an essential cofactor of kinase Cδ [12,14] They are also consistent with the lysosomal localisation of the key retinoid in lipofuscin, the A2E pigment produced during oxidative metabolism which is implicated in autophagocytotic degradation [23]. Another retinoid, retinoic acid is also present in the nucleus where it participates in nuclear receptor-dependent gene transcription [11], in agreement with this we see some retinoids in the nuclei (Fig 2 b).

Furthermore, we discuss cellular images of the most significant abundance ratios. These have physiological meaning and they reflect metabolic functions. Monitoring the balance of mitochondrial fluorescence from NADH and flavins, in particular lipoamide dehydrogenase (LipDH) and electron transfer flavoprotein (ETF) provides information about the cellular redox ratio. Redox state characterisation allows identification of metabolically active cells as well as cells with different self-renewal characteristics [22]. However, accurate redox ratio analysis by fluorescence microscopy is quite challenging, due to difficulties in identifying specific flavins and bound NADH. A significant fraction of the overall flavin-associated autofluorescence is nonredox-responsive, this is because the oxidation state of free cytosolic flavins is not directly coupled to the electron transport chain [23]. Moreover, many authors could not distinguish free from bound NADH, this presents a problem because free NADH produced by glycolysis is not directly related to the redox status. The added complexity is that free NADH levels may still change as a result of impairment of oxidative phosphorylation as the cell attempts to maintain homeostasis. Most authors do not distinguish autofluorescence NADH from NADPH. The latter is used almost exclusively for reductive biosynthesis of fatty acids and steroids, whereas spectrally very similar NADH has a primary role in ATP generation [9]. Despite these subtleties most authors use the total level of NADH and the cellular redox ratio defined as the concentration ratio of flavins to flavins and NADH. This quantity is regarded as inversely proportional to metabolic activity [16-21, 24, 25] . In this work we present the ratio of all flavins to bound, free and total NADH (Fig. 3).

S6.4. Discussion of specific cell subpopulations

Here we discuss cell populations visualised by scatterplots in Fig 4 and statistically analysed in Fig S8. We start from 44% MELAS which is more distinctive.

For 44% MELAS, Fig 4 j shows two distinctive populations (blue and green symbols) where the “blue” cells have generally higher ratio of free to bound NADH. Fig 4 m shows that these

cells also have a higher flavin content than the "green" cells. and they also have a higher average lipofuscin (Fig 4 l, q) and A2E (Fig 4 k, p). The lipofuscin is very pronounced compared to A2E (Fig 4 q). These cells constitute ~ 41% of the total, consistent with 44% mutational loading, and they form a completely separated cluster in Fig 4 r. We propose that these "blue" cells are carrying the mutations.

Fig 4 a-i displays subpopulations in 11% MELAS cells. In Fig 4 a we can see the same two populations, the "blue" cells with unusually high ratio of free NADH to bound NADH and a majority of cells with a much lower ratio. We do not observe excessive FAD in these cells (Fig 4 b) The levels of A2E in the "blue" cells are somewhat higher than in the second population of "green" cells (Fig 4 e),and also these elevated levels of A2E correlate with high free NADH (Fig 4 g). While lipofuscin and A2E appear to correlate (Fig 4 h), the 3D plot (Fig 4i) makes it possible to separate the cells into two distinctive clusters, the main "blue" one and about 7% outliers (highlighted in red) which appear to fit the criteria for the MELAS cells.

The behaviour of these subpopulations before and after treatment is shown in Fig S8. For example, this Figure shows the plots of flavin versus bound NADH content for healthy, 44% MELAS and 44% MELAS treated cells. We can see here that healthy cells form two fairly tight populations, with flavins' abundance varying in the range of 0.2 to 2.2. In contrast, the MELAS cells form very distinctive populations, one with low flavins (between 0 and 3) tentatively identified as "quasi-healthy" cells and a group of cells which could be still mutated with an unusually high flavins content (between 5 and 20). After treatment, the high flavins group completely disappears and the remaining cells form a single population spreading over the same region of abundances as the healthy cells. Fig S8 also shows the populations with respect to the free and bound NADH content. Again, we observe two very tight subpopulations in healthy cells. The MELAS cells also show two populations but in a different region of space, and with a much higher free NADH content that the healthy cells. After treatment only a single population is observed.

References

1.Jieping, Y. Characterisation of a family of algorithms for generalized discriminant analysis on undersampled problems. *Journal of Machine Learning Research*. **6**,483-502(2005).
2.Bowman, A.W., Azzalini, A. *Applied Smoothing Techniques for Data Analysis*. (Oxford University Press, New York,1997).
3.Massey, F.J. The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the American Statistical Association*. **46**(253), 68-78(1951).
4.Wright, B.K. *et al*.Phasor-FLIM analysis of NADH distribution and localization in the nucleus of live progenitor myoblast cells. *Microsc .Res .Tech*. **75**(12),1717-22(2012).
5.Wright, B.K. *et al*. NADH distribution in live progenitor stem cells by phasor-fluorescence lifetime image microscopy. *Biophys J*. **103**(1),L7-L9(2012).
6.Lamb, L.E. *et al*. Primary Photophysical Properties of A2E in Solution. *The Journal of Physical Chemistry B*. **105**(46), 11507-11512(2001).
7.Chance, B., Williamson, J. R., Jamieson, D & Schoener,B. Properties and kinetics of reduced pyridine nucleotide fluorescence of the isolated and in vivo rat heart. *Biochem. Z*. **341**, 357-377(1965).

8.Nuutinen, E.M. Subcellular origin of the surface fluorescence of reduced nicotinamide nucleotides in the isolated perfused rat heart. *Basic Res. Cardiol*. **79**,49-58(1984).
9.Yu, Q. & Heikal, A.A. Two-photon autofluorescence dynamics imaging reveals sensitivity of intracellular NADH concentration and conformation to cell physiology at the single-cell level. *J Photochem Photobiol B*. **95**(1), 46-57(2009).
10.Perez-Ruiz, T. *et al*.Determination of riboflavin, flavin mononucletide and flavin adenine dinucleotide in biological tissues by capillary zone electrophoresis and laser-induced fluorescence detection. *Electrophoresis*. **22**(6),1170-4(2001).
11.Acin-Perez, R. *et al*. Control of oxidative phosphorylation by vitamin A illuminates a fundamental role in mitochondrial energy homoeostasis. *FASEB J*. **24**(2), 627-36(2010).
12.Gundersen, T.E. Methods for detecting and identifying retinoids in tissue. *J Neurobiol*. **66**(7), 631-44(2006).
13.Huang, S., Heikal, A.A. & Webb, W.W. Two-photon fluorescence spectroscopy and microscopy of NAD(P)H and flavoprotein. *Biophys J*. **82**(5),2811-25(2002).
14.Ruff, S.J. & Ong, D.E. Cellular retinoic acid binding protein is associated with mitochondria. *FEBS Lett*. **487**(2), 282-6(2000).
15.Skala, M.C. et al., In vivo multiphoton microscopy of NADH and FAD redox states, fluorescence lifetimes, and cellular morphology in precancerous epithelia. *Proc. Natl.Acad. Sci ,U S A*. **104**(49), 19494-9(2007).
16. Kirkpatrick, N.D. et al.Endogenous fluorescence spectroscopy of cell suspensions for chemopreventive drug monitoring. *Photochem Photobiol*. **81**(1),125-34(2005).
17.Croce, A.C.et al. Autofluorescence properties of rat liver under hypermetabolic conditions. Photochem Photobiol Sci. **6**(11) 12,02-9(2007).
18.Mayevsky, A. & Rogatsky, G.G. Mitochondrial function in vivo evaluated by NADH fluorescence: from animal models to human studies. *Am J Physiol Cell Physiol*. **292**(2), C615-40(2007).
19.Chance, B. & Thorell, B. Fluorescence measurements of mitochondrial pyridine nucleotide in aerobiosis and anaerobiosis. *Nature*. **184**, 931-4(1959).
20.Perry, R.P.et al. Localization and assay of respiratory enzymes in single living cells. Absorbancy measurements on the Nebenkern. *Nature*. **184**, 929-31(1959).
21.Chance, B. *et al*. Oxidation-reduction ratio studies of mitochondria in freeze-trapped samples. NADH and flavoprotein fluorescence signals. *J Biol Chem*. **254**(11),4764-71(1979).
22.Stringari, C.*et al*. Phasor approach to fluorescence lifetime microscopy distinguishes different metabolic states of germ cells in a live tissue. *Proc. Natl. Acad .Sci, U S A*. **108**(33),13582-7(2011).
23.Rice, W.L., Kaplan, D.L. & Georgakoudi, I. Two-photon microscopy for non-invasive, quantitative monitoring of stem cell differentiation. *PLoS One*. **5**(4),10075(2010).
24.Shiino, A.*et al*.Three-dimensional redox image of the normal gerbil brain. *Neuroscience*. **91**(4), 1581-5(1999).
25.Li, L.Z. *et al*. Quantitative magnetic resonance and optical imaging biomarkers of melanoma metastatic potential. *Proc. Natl. Acad. Sci, U S A*. **106** (16), 6608-13(2009).