OCR-Stats: Robust estimation and statistical testing of mitochondrial respiration activities using Seahorse XF Analyzer

Vicente A. Yépez M.[a,b], Laura S. Kremer[c,d], Arcangela Iuso[c,d], Mirjana Gušić[c,d], Robert Kopajtich[c,d], Eliška Koňaříková[c,d], Agnieszka Nadel[c,d], Leonhard Wachutka[a], Holger Prokisch[c,d] and Julien Gagneur[a,b,*]

a. Department of Informatics, Technical University Munich, Boltzmannstr. 3, 85748 Garching, Germany.
b. Quantitative Biosciences Munich, Gene Center, Department of Biochemistry, Ludwig-Maximilians Universität München
c. Institute of Human Genetics, Helmholtz Zentrum München, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany.
d. Institute of Human Genetics, Klinikum rechts der Isar, Technical University Munich, Ismaninger Str. 22, 81675 München, Germany.

*Correspondence to: J.G. (gagneur@in.tum.de)

Abstract

Accurate quantification of cellular and mitochondrial bioenergetic activity is of great interest in many medical and biological areas. Mitochondrial stress experiments performed with Seahorse Bioscience XF Analyzers allow estimating 6 bioenergetics measures by monitoring oxygen consumption rates (OCR) of living cells in multi-well plates. However, detailed statistical analyses of OCR measurements from XF Analyzers have been lacking so far. Here, we performed 126 mitochondrial stress experiments involving 203 fibroblast cell lines to understand how OCR behaves across different biosamples, wells, and plates; which allowed us to statistically model OCR behavior over time. We show that the noise of OCR is multiplicative and that outlier data points can concern individual measurements or all measurements of a well. Based on these insights, we developed a novel statistical method, OCR-Stats, that: i) models multiplicative noise, ii) automatically identifies outlier data points and outlier wells, and iii) takes into account replicates both within and between plates. This led to a significant reduction of the coefficient of variation across experiments of basal respiration by 36% (P = 0.004), and of maximal respiration by 32% (P = 0.023). Also, we propose an optimal experimental design with a minimum number of well replicates needed to obtain confident results. Finally, we use statistical testing taking into account the inter-plate variation to compare the bioenergetics measures of two samples.

Keywords: Oxygen Consumption Rate (OCR); mitochondrial respiration; bioenergetics; statistical testing; outlier detection.

1. Introduction

Mitochondria are double membrane enclosed, ubiquitous, maternally inherited, cytoplasmic organelles present in most eukaryotic organisms (Gorman et al., 2016). They are the powerhouses of the cell (Bhola et al., 2016; Sun et al., 2016), and are also involved in regulating reactive oxygen species (Wallace, 2007), apoptosis (Bhola et al., 2016), amino acid synthesis (Birsoy et al., 2015; Sullivan et al., 2015), cell proliferation (Sullivan et al., 2015), cell signaling (Zong et al., 2016), and in the regulation of innate and adaptive immunity (Weinberg et al., 2015). It follows that a decline in mitochondrial function, reflected by a diminished electron transport chain activity, is implicated in many human diseases ranging from rare genetic disorders (Titov, Cracan et al., 2016) to common disorders such as cancer (Wallace, 2012; Zong et al., 2016), diabetes (Dunham-Snary et al., 2014), neurodegeneration (Yao et al., 2009), and aging (Sun et al., 2016). One of the most informative assessments of mitochondrial function is the quantification of cellular respiration, as it directly reflects electron transport chain impairment (Titov, Cracan et al., 2016) and depends on many sequential reactions from glycolysis to oxidative phosphorylation (Koopman et al., 2016). Estimations of oxygen consumption rates (OCR) expressed in pmol/min, which are mainly driven by mitochondrial respiration through oxidative phosphorylation, and extracellular acidification rates (ECAR) expressed in mpH/min, which reflect glycolysis (Divakaruni et al., 2014; Ferrick et al., 2008; Koopman et al., 2016), are more conclusive for the ability to synthesize ATP and mitochondrial function than measurements of intermediates (such as ATP or NADH) and potentials (Brand et al., 2011; Dmitriev et al., 2012).

OCR was classically measured using a Clark-type electrode, which required a substantial amount of purified mitochondria, was time consuming, and did not allow automated injection of compounds (Wu et al., 2007). Also, experimentation with isolated mitochondria is ineffective because cellular regulation of mitochondrial function is removed during isolation (Hill et al., 2012). In the last few years, a new technology using fluorescent oxygen sensors (Gerencser et al., 2009) in a microplate assay format has been developed by the company Seahorse Bioscience (now part of Agilent Technologies) (Ribeiro et al., 2015). It allows simultaneous, real-time measurements of both OCR and ECAR in multiple cell lines and conditions, reducing the amount of required sample material and increasing the throughput (Divakaruni et al., 2014; Ribeiro et al., 2015).

Typically, OCR and ECAR are measured using the Seahorse XF Analyzer in 96 (or 24) well-plates at multiple time steps under three consecutive treatments (Fig. 1B), in a procedure known as mitochondrial stress test (Agilent Technologies, 2017). Under basal conditions, complexes I-IV exploit energy derived from electron transport to

pump protons across the inner mitochondrial membrane. The thereby generated proton gradient is subsequently harnessed by complex V to generate ATP. Blockage of the proton translocation through complex V by oligomycin represses ATP production and prevents the electron transport throughout complexes I-IV due to the unexploited gradient. Administration of FCCP, an ionophor, subsequently dissipates the gradient uncoupling electron transport from complex V activity and increasing oxygen consumption to a maximum level. Finally, mitochondrial respiration is completely halted using the complex I inhibitor Rotenone. This approach is label-free and non-destructive, so the cells can be retained and used for further assays (Ferrick et al., 2008). OCR differences between different stages of these procedures provide estimation of six different bioenergetics measures: basal respiration, proton leak, non-mitochondrial respiration, ATP production, spare respiratory capacity, and maximal respiration (Brand et al., 2011; Divakaruni et al., 2014) (Figure 1). Increase in proton leak and decrease in maximum respiratory capacity are indicators of mitochondrial dysfunction (Brand et al., 2011). ATP production, basal respiration, and spare capacity alter in response to ATP demand, which is not necessarily mitochondrion-related as it may be the consequence of deregulation of any cellular process altering general cellular energy demand.

Current literature describing the Seahorse technology addressed experimental aspects regarding sample preparation (Dranka et al., 2011; Zhang et al., 2012), the amount of cells to seed (Zhang et al., 2012; Zhou et al., 2012), and compound concentration in different organisms (Dranka et al., 2011; Koopman et al., 2016; Shah-Simpson et al., 2016). However, studies regarding statistical best practices for determining OCR levels and testing them against another are lacking. The sole definition of bioenergetic measure varies between authors, as well as the number of time points in each interval (one time point in (Dranka, Hill, & Darley-Usmar, 2010), two time points in (Chacko et al., 2014) and four or more time points in (Dunham-Snary et al., 2014)); and whether differences (Invernizzi et al., 2012; Koopman et al., 2016; Sullivan et al., 2015), ratios (Yao et al., 2009; Zhang et al., 2011), or both (Shah-Simpson et al., 2016; Zhou et al., 2012) should be computed. Consequently, comparison of results across studies is difficult. Moreover, statistical power analyses for experimental design are often not provided. Differences in OCR between distinct biosamples (e.g. patient vs. control, or gene knockout vs. WT) can be as low as 12 – 30% (Almontashiri et al., 2014; Mitsopoulos et al., 2015; Stroud et al., 2016). Therefore, to design experiments with appropriate power to significantly detect such differences, it is important to know the source and amplitude of the variation within each sample, and reduce it as much as possible.

Here, we developed statistical good practices to support experimentalists in designing, analyzing, and reporting results of Seahorse mitochondrial stress experiments. To this end, we analyzed a large dataset of 126 mitochondrial stress experiments in 96-well plate format involving 203 different fibroblast cell lines (Table S1). The large amount of between-plate and within-plate replicates allowed us to statistically characterize the nature and amount of biases and random variations in these data. Based on these insights, we developed a statistical procedure, called OCR-Stats, to extract robust and accurate oxygen consumption rates for each well, which translates into robust summarized values of the multiple replicates inside one plate and across plates. OCR-Stats includes normalization of raw data and outlier identification and controls for well and plates biases, which led to significant increased in accuracy over state-of-the-art methods. Between-well and between-plate biases, as well as random variations, were found to be essentially multiplicative. This motivated for a definition of bioenergetics measures based on ratios. We formally defined 5 such measures: ETC-dependent OC proportion, ATPase-dependent OC proportion, ETC-dependent proportion of ATPase-independent OC, and Maximal OC fold change (Fig. 1A). We provide estimators for each one that were empirically normally distributed, which permitted using linear regression models for assessing statistical significance of bioenergetics measures comparisons. Furthermore, our study provides experimental design guidance by i) showing that between-plate variation largely dominates within-plate variation, implying that it is important to seed the same biosamples in multiple plates, and ii) providing estimates of variances within and between plates for each bioenergetic measure allowing for statistical power computations. A free and pose source implementation of OCR-stats in the statistical language R is provided at github.com/gagneurlab/OCR-Stats.

2. Results

2.1 Experimental design and raw data

We derived OCR, ECAR, and cell number for 203 dermal fibroblast cultures derived from patients suffering from rare mitochondrial diseases, and control cells from healthy donors (normal human dermal fibroblasts - NHDF, Methods, Table S1). These were assayed in 126 plates, all using the same protocol (Methods). We grew 27 cell lines multiple times and placed them in more than one plate. We will refer to these growth replicates as different biosamples. The NHDF cell line was seeded in all plates for assessment of potential systematic plate biases. All four corners of each plate were left as blank, i.e. filled with media but no cells to control for changes in temperature (Dranka et al., 2011). The typical layout of a plate is depicted in Fig. 1C,

showing how each biosample is present in many well replicates. We seeded between 3 and 7 biosamples per plate (median = 4). This variation reflects typical set-ups of experiments in a lab performed over multiple years.

We used the standard mitochondrial stress test assay (Fig. 1A, (Agilent Technologies, 2017)) leading to four time intervals with three time points each and denoted by Int_1 (resting cells), Int_2 (after oligomycin), Int_3 (after FCCP) and Int_4 (after Rotenone). Wells for which the median OCR level did not follow the expected order, namely, median(OCR(Int_3)) > median(OCR(Int_1)) > median(OCR(Int_2)) > median(OCR(Int_4)), were discarded (977 wells, 10.47%). We also excluded from the analysis contaminated wells and wells in which the cells got detached (461 wells, 4.94%, Methods).

2.2 Random and systematic variations between replicates within plates

Typical replicate time series are shown in Fig. 2A, with data from 12 wells for a single biosample in a single plate. It shows the kinds of variations that we observed.

First, outlier data points occurred frequently. We distinguished two different types of outliers: entire series for a well (e.g., well G5 in Fig. 2A) and individual data points (e.g., well B6 at time point 6 in Fig. 2A). In the latter case, eliminating the entire series for well B6 would be too restrictive, and would result in losing valuable data from the other 11 valid time points. Therefore, methods to find outliers considering these two possibilities must be devised.

Second, we noticed that the higher the OCR value, the higher the variance between replicates, suggesting that the error is multiplicative. Unequal variance, or heteroscedasticity, can strongly affect the validity of statistical tests and the robustness of estimations. We therefore suggest modeling OCR on a logarithmic scale, where the dependency between variance and mean disappears (Figs. 2B, 2C). Respiratory chain enzyme activities such as NADH-ubiquinone reductase have already been shown to obey log-normal distributions (Hautakangas et al., 2016).

Third, we observed systematic biases in OCR between wells (e.g., OCR values of well C6 are among the highest while OCR values of well B5 are among the lowest at all time points, Fig. 2A). Variations in cell number, initial conditions, treatment concentrations, and fluorophore sleeve calibration can lead to systematic differences between wells, which we refer to as well biases. To investigate whether well biases could be mostly corrected using cell number as suggested in (Dranka et al., 2010),

we counted the number of cells after the experiments using Cyquant (Methods). As expected, median OCR for each interval grows linearly with cell number measured at the end of the experiment (Spearman rho between 0.32 and 0.47, $P < 2.2e-16$, Fig. S1A). However, the relation is not perfect reflecting important additional sources of variations, and also possible noise in measuring cell number. Strikingly, dividing OCR by cell count led to a higher coefficient of variation (standard deviation divided by the mean) between replicate wells than without that correction (Fig. S1B). This analysis showed that normalization for cell number should not be done simply by a blunt division by raw cell counts and motivated us to derive another method to capture well biases.

2.3 A statistical model of OCR

Building on these insights, we next introduced a statistical model of OCR within plate. For a given biosample in one plate, we modeled the logarithm of OCR $y_{w,t}$ of well *w* at time point *t* as a sum of well bias, interval effects and noise, i.e.,:

$$y_{w,t} = \alpha_{i(t)} + \beta_w + \varepsilon_{w,t}. \quad (1)$$

The term $\alpha_{i(t)}$ is the effect of the interval *i(t)* of time point *t*. The term β_w is the relative bias of well *w* compared to a reference well, which is set arbitrarily and corresponds to the first well in alphabetical order. The term $\varepsilon_{w,t}$ is the error.

We defined the OCR levels (θ_i) as the expected log OCR per interval, averaged over all wells:

$$\hat{\theta}_i = \alpha_i(t) + \frac{\sum_w \beta_w}{n}, \quad (2)$$

where *n* is the number of wells.

Note that the well bias is modeled independently for each plate, i.e., the bias of a certain well in one plate is different from the bias of the well at the same location in another plate.

We present now our OCR-Stats algorithm, for a given plate:

1. Fit the log linear model (1) using the least-squares method, which consists in minimizing $\sum_w \sum_t (y_{w,t} - \alpha_{i(t)} - \beta_w)^2$, thus obtaining the coefficients α_i, β_w; and $\hat{\theta}_i$ using (2).
2. For each time point *t* in interval *i* and well *w*, define the OCR residual: $e_{w,t} = y_{w,t} - \hat{\theta}_{i(t)}$, which is used to identify outliers (Methods).
3. Identify and remove well level outliers, fit again, iteratively, until no more are found.
4. Identify and remove single point outliers, fit again, iteratively, until no more are found.

5. Scale back to natural scale in order to compute the bioenergetics measures (e.g.: Basal respiration = $e^{\theta_1} - e^{\theta_4}$, Maximal respiration = $e^{\theta_3} - e^{\theta_4}$, etc.), or take the difference in the logarithmic scale to obtain the metrics from Table 1.

2.4 Variations within plates

We were then interested in determining the amplitude of the variance inside each plate in order to compute the number of wells needed to obtain robust estimates $\hat{\boldsymbol{\theta}}$. Using only the controls NHDF, we computed the standard deviation $\sigma_{i,j}^{w}$ of the logarithm of OCR across all wells for each plate *j* and interval *i*. Then, we computed the median across plates, thus obtaining one value σ_i^w per interval ($\sigma_1^w = 0.10, \sigma_2^w = 0.13, \sigma_3^w = 0.12, \sigma_4^w = 0.16$). As we worked in the logarithmic scale, the error in the natural scale becomes multiplicative and relative. The standard error of the estimates $\hat{\boldsymbol{\theta}}$ can be expressed as $\sigma_{\hat{\theta}_i} = \sigma_i^w / \sqrt{n_w}$, where n_w is the number of wells. The highest value of σ_i^w was 0.16, so in order to get a relative error of 5%, cells should be seeded in 10 wells. This result comes from a variation after removing outliers, so considering that around 16.5% of wells were found to be outliers, then ideally we should use $10/(1 - 0.165) \approx 12$ wells per biosample.

2.5 Variations between plates

After analyzing the variation among wells inside plates, we set up to study the variation across multiple plates. Using data from the controls NHDF, we found that the variability between plates for all four intervals is much larger than between wells (Table S2, Fig. S4). We next asked whether there exists a systematic plate bias that could be corrected for. We indeed observed a similar increase in OCR on the interval 1 for both biosamples on plate #20140430 with respect to #20140428 (Fig. 3A). To test whether this tendency held across every repeated biosample, we compared all replicate pairings with their respective NHDF controls and found a positive correlation (Fig. 3B). These differences can come from changes in temperature or the use of different sensor cartridges (Koopman et al., 2016). Because the plate biases are systematic, we can correct for them using a log linear model (Methods). Nonetheless, the biases do not explain all the between plate variation as the remaining variance is large (relative variance of the residuals: I_1: 49.8%, I_2: 51.6%, I_3: 65.6% and I_4: 55.9%). It is therefore important to perform multiple plate analyses to be able to conclude for a reproducible systematic difference between biosamples.

2.6 Statistical comparison between biosamples

In order to compare the bioenergetics measures of two biosamples, we first need to decide if it is better evaluating differences or ratios of the OCR levels in the natural scale. Even after correcting for well biases, there is a remaining cell number effect (Fig. 3C); therefore, we recommend working with ratios of OCR levels (or differences in the logarithmic scale). We propose the following definitions:

OCR ratios	**Abbr.**	**Metrics**	**Tested differences *d***	**Equivalent**
ETC-dependent OC proportion	E/I – proportion	$\frac{OCR_1 - OCR_4}{OCR_1} = 1 - \exp(\theta_{Ei} - \theta_I)$	$(\theta_{I,f} - \theta_{Ei,f}) - (\theta_{I,\mathrm{ctrl}} - \theta_{Ei,\mathrm{ctrl}})$	Basal Respiration
ATPase-dependent OC proportion	A/I – proportion	$\frac{OCR_1 - OCR_2}{OCR_1} = 1 - \exp(\theta_{Ai} - \theta_I)$	$(\theta_{I,f} - \theta_{Ai,f}) - (\theta_{I,ctrl} - \theta_{Ai,ctrl})$	ATP Production
ETC-dependent proportion of ATPase-independent OC	E/Ai - proportion	$\frac{OCR_2 - OCR_4}{OCR_2} = 1 - \exp(\theta_{Ei} - \theta_{Ai})$	$(\theta_{Ai,f} - \theta_{Ei,f}) - (\theta_{Ai,ctrl} - \theta_{Ei,ctrl})$	Proton Leak
Maximal OC fold change	M/I – fold change	$\frac{OCR_3}{OCR_1} = \exp(\theta_M - \theta_I)$	$(\theta_{M,f} - \theta_{I,f}) - (\theta_{M,ctrl} - \theta_{I,ctrl})$	Spare Capacity
Maximal over ETC-independent OC fold change	M/Ei – fold change	$\frac{OCR_3}{OCR_4} = \exp(\theta_M - \theta_{Ei})$	$(\theta_{M,f} - \theta_{Ei,f}) - (\theta_{M,ctrl} - \theta_{Ei,ctrl})$	Maximal Respiration

***Table 1**: OCR ratios, metrics and equivalents*

Then, for any given OCR ratio *b* (eg. M/Ei - fold change), we test differences of log OCR ratios of patient versus a control cell line (Table 1) using the following linear model:

$$d_{b,f,p} = \mu_{b,f} + \epsilon_{b,f,p}, (3)$$

where $d_{b,f,p}$ corresponds to the difference of ratio *b* of a cell line *f* and the respective control on plate *p*. We solve it using linear regression, thus obtaining one value $\mu_{b,f}$ per each ratio *b* and cell line *f*. We then compare these $\mu_{b,f}$ values (which follow a t-Student distribution) against the null hypothesis $\mu_{b,f} = 0$ to obtain p-values and confidence intervals (Figs. 4A, 4B, Methods).

2.7 Benchmark of OCR-Stats algorithm

In order to benchmark the OCR-Stats algorithm, we computed the coefficient of variation (standard deviation divided by mean) of the six bioenergetics measures in the natural scale of all repeated biosamples across plates. The lower the coefficient of variation among replicates, the better the method. We cannot test using the final estimates $\hat{\theta}^f$ after correcting for plate effect, because we would fall into circularity as correcting using $\beta_{i,p}$ forces replicates to have a closer value. Therefore, just for benchmarking purposes, we corrected for plate effect using only the data from the controls NHDF *c* of each plate, namely:

$$y_{i,p}^c = \beta_0^c + \beta_i^c + \beta_p^c + \varepsilon_{i,p}. \ (4)$$

We solved (4) using linear regression and used the effects β_p^c as offsets in (1), and recomputed $\hat{\theta}_i$ values accordingly. We scaled back to natural scale to calculate the bioenergetics measures and the coefficient of variation of all repeated biosamples (except the control to avoid circularity) using: i) the default Extreme Differences (ED) method (Methods) provided by the vendor, ii) the log linear (LL) corresponding to steps 1 and 2 of the OCR-Stats algorithm, iii) complete OCR-Stats (LL + outlier removal), and iv) OCR-Stats after correcting for plate effect (OCR-PE) using (4). Each step contributed to lowering the coefficient of variation, obtaining a final significant reduction of 36% and 32% in basal and maximal respiration, respectively, from OCR-PE with respect to ED ($P < 0.03$, one-sided Wilcoxon test) (Fig. 5).

2.8 Benchmark of OCR-Stats statistical testing method

We applied OCR-Stats, Extreme Differences with Wilcoxon test within each plate (within-plate ED), and Extreme Differences with Wilcoxon test across plates (across-plate ED) to obtain the M/Ei ratio and maximal respiration (MR) of all the 26 cell lines that were seeded in more than one plate (Methods). For every approach we

computed p-values for significant fold-changes against the controls. Six of these cell lines come from patients with rare variants in genes associated with an established cellular respiratory defect, allowing for assessing the sensitivity of each approach (Table S3, (Haack et al., 2013; Hildick-Smith et al., 2013; Kremer et al., 2017; Pronicka et al., 2016; Van Haute et al., 2016)). Also, two cell lines (#73901 and #91410) that showed no significant respiratory defects in earlier studies (Powell et al., 2015) (Kremer et al., 2016) served as negative controls.

The within-plate ED method reported significantly higher or lower MR for 56/69=81.2% biosamples (Figs. 4A, 4B, Table S3). Moreover, every cell line was found to be significant on at least one plate, despite large variation in M/Ei fold change between plates (Fig. 4A). Also, for 11 cell lines, one plate at least also gave non-significant differences. These results show the importance of assessing differences using multiple plates and advocate for a more robust approach than within-plate ED.

One approach to take multiple plates into accounts is to perform a Wilcoxon test based on per plate average ED values (across-plate ED, Methods). However, this approach requires samples to be seeded in at least five plates in order to obtain significant results. Here, only one cell line, #78661, was found significant this way.

In contrast, significance with the OCR-Stats statistical algorithm can be reached by seeding a biosample in one plate only; provided there were other between-plate replicates to compute the inter-plate variance. On this data, OCR-Stats was much more conservative than within-plate ED and found only 7/26=27% cell lines to have aggregated significantly lower M/Ei than the control. There was no evidence against the normality and homoscedasticity assumption of OCR-stats as the quantile-quantile plots of the residuals aligned well along the diagonal (Figs. 4C, S4). All the 6 positive control cell lines were reported to have significantly lower M/Ei than control by OCR-Stats (Figs. 4A, 4B, Table S3). Moreover, OCR-Stats did not report significant M/Ei differences for the two negative controls. Altogether, these results show that OCR-Stats successfully identifies and removes variation within and between plates, providing more stable results which translates into less false positives.

Discussion and conclusion

Mitochondrial studies using extracellular fluxes (specifically the XF Analyzer from Seahorse) are gaining popularity; therefore, it is of paramount importance to have a

proper statistical method to estimate the OCR levels from the raw data. In this paper, we have developed such a model, which includes approaches to control for well and plate biases, and automatic outlier identification. By doing so, we were able to significantly reduce the coefficient of variation of replicates across plates. After analyzing the intra-plate variation, we found that the minimum number of wells per biosample should be 12.

We found that dividing cellular OCR by cell number was introducing more noise than was seen for uncorrected data. Here, we seeded always the same number of cells. Hence, the variations we observed in cell number at the end of the experiments are largely overestimated by noise in measurements. In other experimental settings, in which different numbers of cells are seeded, we suggest to include an offset term to the model (1) equal to the logarithm of the seeded cell number to control for this variation by design. Also, the Seahorse XF Analyzer can be used on isolated mitochondria and on isolated enzymes, where a normalization approach is to divide OCR by mitochondrial proteins or enzyme concentration (Seahorse Bioscience, 2014). However, as described here for cellular assays, robust normalization procedures require careful analysis.

To use XF Seahorse Analyzers for large-scale experiments, one needs to be able to compare biosamples measured on different plates. Our investigation showed that there is roughly multiplicative bias between plates that can be controlled to some extent by including control biosamples across plates, as we did here with NHDF. We proposed an extension of our intra-plate robust linear regression approach to multiple plates that can handle model this plate bias. However, we also noticed that the assumption of a multiplicative plate bias is not sufficient as there are other sources of variation. Therefore, for comparing two biosamples statistically, they need to be placed on the same plate, and repeated multiple times. We demonstrated that it is better to compare OCR ratios rather than differences as this eliminates sources of variation like cell number. We proposed another linear model that takes into account the inter-plate variation, which we showed to agree with previous results of patients diagnosed with mitochondrial disorders.

We also encourage users to understand the biological meaning of each OCR ratio (Table 1). For example, cell line #73387 was found to have a lower, but non significantly ($P < 0.10$), M/Ei ratio (the most common metric used throughout the literature, Table S3), but when analyzing its E/I proportion, we found that it was drastically lower than the control ($P < 1.2\text{x}10^{-7}$). This result is consistent with its

genetic diagnosis (Table S1, (Oláhová et al., 2015)). For visualizing OCR ratios, raw OCR vs. time plots are useful in both logarithmic and natural scales.

In principle, OCR-Stats should be able to estimate ECAR levels. Nevertheless, similar analyses as performed here should be done beforehand in order to guarantee that the method is indeed applicable. Preliminary investigations suggest that the nature of noise (outliers, multiplicative) is similar than for OCR.

Finally, it is important to understand further sources of variations between plates, cell cultures, treatments and other factors in order to correct for them. Here, we found that gender does not significantly influence OCR levels (Fig. S5), but age (for which we have no register), may play a role.

Methods

Biological material

All biosamples come from primary fibroblast cell lines of humans suffering from rare mitochondrial diseases, established in the framework of mitoNet and GENOMIT. The controls used are primary patient fibroblast cell lines, normal human dermal fibroblasts (NHDF) from neonatal tissue, commercially available from Lonza, Basel, Switzerland.

Measure of extracellular fluxes using Seahorse XF96

We seeded 20,000 fibroblasts cells in each well of a XF 96-well cell culture microplate in 80 ml of culture media, and incubated overnight at 37°C in 5% CO_2. The four corners were left only with medium for background correction. Culture medium is replaced with 180 ml of bicarbonate-free DMEM and cells are incubated at 37°C for 30 min before measurement. Oxygen consumption rates (OCR) were measured using a XF96 Extracellular Flux Analyzer (Agilent Technologies, 2017). OCR was determined at four levels: with no additions, and after adding: oligomycin (1 µM); carbonyl cyanide 4-(trifluoromethoxy) phenylhydrazone (FCCP, 0.4 µM); and rotenone (2 µM) (additives purchased from Sigma at highest quality). After each assay, manual inspection was performed on all wells using a conventionally light microscope.

Cell number quantification

Cell number was quantified using the CyQuant Cell Proliferation Kit (Thermo Fisher Scientific, Waltham, MA, USA) according to the manufacturer's protocol. In brief, cells were washed with 200 µL PBS per well and frozen in the microplate at -80°C to ensure subsequent cell lysis. Cells were thawed and resuspended vigorously in 200 µL 1x cell-lysis buffer supplemented with 1x CyQUANT GR dye per well.

Resuspended cells were incubated in the dark for 5 min at RT whereupon fluorescence was measured (excitation: 480 nm, emission: 520 nm).

Extreme Differences (default) Method to compute bioenergetics measures

On every plate independently, for each well, on interval 1 take the OCR corresponding to the last measurement, on intervals 2 and 4 take the minimum and on interval 3 the maximum OCR value (Divakaruni et al., 2014). Then, do the corresponding differences to estimate the bioenergetics measures. Report the results per patient as the mean across wells plus standard deviation or standard error, separately for each plate.

Outlier Removal

For each sample *s* and well *w*, compute the mean across time points of its squared residuals: $r_w \coloneqq \mathrm{mean}_t({e_{w,t}}^2)$, thus obtaining a distribution $\boldsymbol{r}$. Identify as outliers the wells whose $r_w > \mathrm{median}(\boldsymbol{r}) + 5 \cdot \mathrm{mad}(\boldsymbol{r})$, where mad, median absolute deviation, is a robust estimation of the standard deviation (Fig. S2A). We found that deviations by 5 mad from the median were selective enough in practice. Compute the vector of estimates $\hat{\boldsymbol{\theta}}$ using the remaining wells and iterate this procedure until no more wells are identified as outliers. It required 8 iterations until convergence and around 16.5% of all the wells were found to be outliers (Fig. S2B).

Single point outliers are identified in a similar way. After discarding the wells that were found to be outliers in the previous step, categorize as outliers single data points whose ${e_{w,t}}^2 > \mathrm{median}_t\left({e_{w,t}}^2\right) + 5 \cdot \mathrm{mad}_t({e_{w,t}}^2)$ (Fig. S2C). Iterate until no more outliers are found. It required 19 iterations until convergence and approximately 6.1% of single points were found to be outliers (Fig. S2D).

Plate effect model

In an attempt to correct for plate effect, we propose a log linear model where the levels θ' depend on interval *i*, samples *s* and plate *p*:

$$\theta'_{i,s,p} = \alpha_{i,s} + \beta_{i,p} + \varepsilon_{i,s,p},$$

thus obtaining one coefficient $\beta_{i,p}$ for each plate-interval combination. These effects are added to the previous estimates: $\hat{\theta}^f_{i,s,p} = \hat{\theta}_{i,s} - \beta_{i,p}$, obtaining the final estimates $\hat{\theta}^f$. As for (1), the model is solved using linear regression.

Multi-plate averaging method

In case of inter-plate comparisons, the multi-plate averaging methods takes the average and standard error of the bioenergetics measures obtained using the ED method of all repeated biosamples across plates (Agilent Technologies, 2016).

Statistical Testing

To evaluate the OCR ratios between a fibroblast *f* and a control, we use the

corresponding tested difference *d* (Table 1). For a fibroblast *f* located on a plate *p*, we define $\mu_{i,j,f,p} := \left[\hat{\theta}_i - \hat{\theta}_j\right]_{f,p} - \left[\hat{\theta}_i - \hat{\theta}_j\right]_{NHDF,p}$, where *i* and *j* are any two different intervals. From there, we can obtain a t-statistic: $t_{\hat{d}} = \frac{\mu - d_0}{se(\mu)}$, where $d_0 = 0$ as that is the value against we want to compare μ against, and *se* is the standard error. The t-statistic follows a t-distribution with $n - 2$ degrees of freedom, from which we can obtain p-values. Moreover, we can obtain confidence intervals: $[\mu - se(\mu)t_{n-2}^{\alpha},\ \mu + se(\mu)t_{n-2}^{\alpha}]$, where $(1-\alpha)$ is the confidence level and t_{n-2}^{α} the $(1-\alpha/2)$ quantile of the t_{n-2} distribution. Note that the normality assumption holds for the residuals $\epsilon_{b,f,p}$ (Figs. 4C, S4).

Acknowledgements

We would like to thank Daniel Bader, Žiga Avsec, Jun Cheng and Paula Fernández-Guerra for valuable discussions and manuscript revision. This study was supported by the German Bundesministerium für Bildung und Forschung (BMBF) through the E-Rare project GENOMIT (01GM1603 and 01GM1207, H.P. and T.M.), through the Juniorverbund in der Systemmedizin ‘mitOmics’ (FKZ 01ZX1405A J.G., L.W. and V.A.Y.M.) and the DZHK (German Centre for Cardiovascular Research, L.S.K.). A Fellowship through the Graduate School of Quantitative Biosciences Munich (QBM) supports V.A.Y.M.. H.P. is supported by EU FP7 Mitochondrial European Educational Training Project (317433). J.G., V.A.Y.M. and H.P. are supported by EU Horizon2020 Collaborative Research Project SOUND (633974). We thank the Cell lines and DNA Bank of Pediatric Movement Disorders and Mitochondrial Diseases of the Telethon Genetic Biobank Network (GTB09003).

Author contributions

J.G. and H.P. planned the project and overviewed the research. H.P. designed the experiments. V.A.Y.M. curated and analyzed the data. J.G. devised the statistical analysis. L.S.K., A.I., E.K., M.G., and A.N. performed the mitochondrial stress test experiments and cell number quantification. V.A.Y.M., L.W. and J.G. made the figures. V.A.Y.M. and J.G. wrote the manuscript. All authors performed critical revision of the manuscript.

Figure legends

Figure 1. Principle of the mitochondrial stress test assay (**A**) Cartoon illustration of OCR levels (y-axis) versus time (x-axis). Injection of the three compounds oligomycin, FCCP and rotenone delimit four time intervals within which OCR is roughly constant. (**B**) Targets of each compound in the electron transport chain. (**C**) Typical layout of a mitochondrial stress test 96-well plate.

Figure 2. OCR behavior over time. (**A**) Typical time series replicates inside a plate. Behavior of OCR expressed in pmol/min (y-axis) of Fibro_VY_017 over time (x-axis). Colors indicate the row, and shape the column of 12 well replicates. Variation increases for larger OCR values, OCR has a systematic well effect and there exist two types of outliers: well-level and single-point. (**B**) Scatterplot of standard deviation (y-axis) vs. mean (x-axis) of all 3 time replicates of each interval, well and plate of OCR of NHDF only, shows a positive correlation (n = 409). (**C**) Same as (B) but for the logarithm of OCR, where the correlation disappears.

Figure 3. Plate bias. (**A**) Log of OCR in interval 3 (y-axis) for the cell lines #65126 and NHDF (x-axis) which were seeded in 2 different plates (color-coded). The similar increase in OCR from plate #20140128 to #20140430 in both biosamples suggests that there is a systematic plate bias. (**B**) Scatterplots of the differences of the logarithm of OCR levels $\boldsymbol{\theta}$ of all possible 2 by 2 combinations of repeated biosamples across experiments (y-axis) against their respective controls (NHDF) (x-axis) show that there exists a positive correlation (I_1: ρ= 0.64, $P < 2.3\text{x}10^{-8}$, I_2: ρ= 0.65, $P < 1.2\text{x}10^{-8}$, I_3: ρ= 0.52, $P < 1.2\text{x}10^{-5}$, I_4: ρ= 0.64, $P < 1.4\text{x}10^{-8}$), suggesting a systematic plate bias (n = 63). **(C)** Scatterplot of the difference of log OCR levels of patients vs. control NHDF (both axes) of every interval with respect to another. All intervals correlate with each other even after removing plate bias (by subtracting control values).

Figure 4. Statistical testing of M/Ei fold change patient vs. control on multiple plates. (A) Ratio of M/Ei fold change (y-axis) of all cell lines repeated across plates (x-axis) and their respective control, sorted by p-value obtained using the OCR-Stats method. Left of the red dashed line are cell lines with significantly lower M/Ei fold change using OCR-Stats. Dots in orange represent cell lines with significantly lower or higher M/Ei fold change using the ED method. Highlighted positive (+) and negative (-) controls. **(B)** Similar as (A), but depicting the p-value in logarithmic scale (y-axis) using OCR-Stats. Red dashed line at P = 0.05. Dots in red represent biosamples with significantly lower M/Ei fold change using the OCR-Stats method. **(C)** Quantile-quantile theoretical (x-axis) vs. observed (y-axis) plot of the residuals of the linear model (3) applied to M/Ei fold change.

Figure 5. Benchmark using coefficient of variation. Coefficient of variation (CV = standard deviation / mean, y-axis) of replicates across experiments (n=26) using different methods (x-axis) to estimate the 6 bioenergetics measures. In all, except for Spare Capacity, OCR-Stats with plate effect showed significantly lower variation with

respect to the Extreme Differences method. P-values obtained from one-sided paired Wilcoxon test.

References

Agilent Technologies. (2016). Multi-File XF Report Generator, User Guide. Retrieved from http://www.agilent.com/cs/library/usermanuals/public/Report Generator User Guide_Seahorse XF Cell Mito Stress Test_MultiFile_RevA.pdf

Agilent Technologies. (2017). Mito Stress Test Kit, User Guide, (103016–400). Retrieved from https://www.agilent.com/cs/library/usermanuals/public/XF_Cell_Mito_Stress_Test_Kit_User_Guide.pdf

Almontashiri, N. A. M., Chen, H. H., Mailloux, R. J., Tatsuta, T., Teng, A. C. T., Mahmoud, A. B., … Stewart, A. F. R. (2014). SPG7 Variant Escapes Phosphorylation-Regulated Processing by AFG3L2, Elevates Mitochondrial ROS, and Is Associated with Multiple Clinical Phenotypes. *Cell Reports*, *7*(3), 834–847. http://doi.org/10.1016/j.celrep.2014.03.051

Bhola, P. D., & Letai, A. (2016). Mitochondria-Judges and Executioners of Cell Death Sentences. *Molecular Cell*, *61*(5), 695–704. http://doi.org/10.1016/j.molcel.2016.02.019

Birsoy, K., Wang, T., Chen, W. W., Freinkman, E., Abu-Remaileh, M., & Sabatini, D. M. (2015). An Essential Role of the Mitochondrial Electron Transport Chain in Cell Proliferation Is to Enable Aspartate Synthesis. *Cell*, *162*, 540–551. http://doi.org/10.1016/j.cell.2015.07.016

Brand, M. D., & Nicholls, D. G. (2011). Assessing mitochondrial dysfunction in cells. *The Biochemical Journal*, *435*(2), 297–312. http://doi.org/10.1042/BJ20110162

Chacko, B. K., Kramer, P. a, Ravi, S., Benavides, G. a, Mitchell, T., Dranka, B. P., … Darley-Usmar, V. M. (2014). The Bioenergetic Health Index: a new concept in mitochondrial translational research. *Clinical Science*, *127*(6), 367–373. http://doi.org/10.1042/CS20140101

Divakaruni, A. S., Paradyse, A., Ferrick, D. A., Murphy, A. N., & Jastroch, M. (2014). *Analysis and interpretation of microplate-based oxygen consumption and pH data*. *Methods in Enzymology* (Vol. 547). http://doi.org/10.1016/B978-0-12-801415-8.00016-3

Dmitriev, R. I., & Papkovsky, D. B. (2012). Optical probes and techniques for O 2 measurement in live cells and tissue. *Cellular and Molecular Life Sciences*, *69*, 2025–2039. http://doi.org/10.1007/s00018-011-0914-0

Dranka, B. P., Benavides, G. A., Diers, A. R., Giordano, S., Blake, R., Reily, C., … Darley-Usmar, V. M. (2011). Assessing bioenergetic function in response to oxidative stress by metabolic profiling. *Free Radical Biology and Medicine*, *51*(9), 1621–1635. http://doi.org/10.1016/j.freeradbiomed.2011.08.005

Dranka, B. P., Hill, B. G., & Darley-Usmar, V. M. (2010). Mitochondrial reserve capacity in endothelial cells: The impact of nitric oxide and reactive oxygen species. *Free Radical Biology and Medicine*, *48*(7), 905–914. http://doi.org/10.1016/j.freeradbiomed.2010.01.015

Dunham-Snary, K. J., Sandel, M. W., Westbrook, D. G., & Ballinger, S. W. (2014). Redox Biology A method for assessing mitochondrial bioenergetics in whole white adipose tissues. *Redox Biology*, *2*, 656–660. http://doi.org/10.1016/j.redox.2014.04.005

Ferrick, D. A., Neilson, A., & Beeson, C. (2008). Advances in measuring cellular bioenergetics using extracellular flux. *Drug Discovery Today*, *13*(March). http://doi.org/10.1016/j.drudis.2007.12.008

Gerencser, A. A., Neilson, A., Choi, S. W., Edman, U., Yadava, N., Oh, R. J., … Brand, M. D. (2009). Quantitative Microplate-Based Respirometry with

Correction for Oxygen Diffusion. *Analytical Chemistry*, *81*(16), 6868–6878.

Gorman, G. S., Chinnery, P. F., DiMauro, S., Hirano, M., Koga, Y., McFarland, R., … Turnbull, D. M. (2016). Mitochondrial diseases. *Nature Reviews Disease Primers*, *2*. http://doi.org/10.1038/nrdp.2016.80

Haack, T. B., Rolinski, B., Haberberger, B., Zimmermann, F., Schum, J., Strecker, V., … Prokisch, H. (2013). Homozygous missense mutation in BOLA3 causes multiple mitochondrial dysfunctions syndrome in two siblings. *Journal of Inherited Metabolic Disease*, *36*(1), 55–62. http://doi.org/10.1007/s10545-012-9489-7

Hautakangas, M. R., Hinttala, R., Rantala, H., Nieminen, P., Uusimaa, J., & Hassinen, I. E. (2016). Evaluating clinical mitochondrial respiratory chain enzymes from biopsy specimens presenting skewed probability distribution of activity data. *Mitochondrion*, *29*, 53–58. http://doi.org/10.1016/j.mito.2016.05.004

Hildick-Smith, G. J., Cooney, J. D., Garone, C., Kremer, L. S., Haack, T. B., Thon, J. N., … Paw, B. H. (2013). Macrocytic anemia and mitochondriopathy resulting from a defect in sideroflexin 4. *American Journal of Human Genetics*, *93*(5), 906–914. http://doi.org/10.1016/j.ajhg.2013.09.011

Hill, B. G., Benavides, G. A., Jr, J. R. L., Ballinger, S., & Italia, L. D. (2012). Integration of cellular bioenergetics with mitochondrial quality control and autophagy. *Biological Chemistry*, *393*(12), 1485–1512. http://doi.org/10.1515/hsz-2012-0198

Invernizzi, F., D'Amato, I., Jensen, P. B., Ravaglia, S., Zeviani, M., & Tiranti, V. (2012). Microscale oxygraphy reveals OXPHOS impairment in MRC mutant cells. *Mitochondrion*, *12*(2), 328–335. http://doi.org/10.1016/j.mito.2012.01.001

Koopman, M., Michels, H., Dancy, B. M., Kamble, R., Mouchiroud, L., Auwerx, J., … Houtkooper, R. H. (2016). A screening-based platform for the assessment of cellular respiration in Caenorhabditis elegans. *Nature Protocols*, *11*(10), 1798–1816. http://doi.org/10.1038/nprot.2016.106

Kremer, L. S., Bader, D. M., Mertes, C., Kopajtich, R., Pichler, G., Iuso, A., … Prokisch, H. (2017). Genetic diagnosis of Mendelian disorders via RNA sequencing. *Nature Communications*, *8*, 15824. http://doi.org/10.1038/ncomms15824

Kremer, L. S., Distelmaier, F., Alhaddad, B., Hempel, M., Iuso, A., Küpper, C., … Haack, T. B. (2016). Bi-allelic Truncating Mutations in TANGO2 Cause Infancy-Onset Recurrent Metabolic Crises with Encephalocardiomyopathy. *American Journal of Human Genetics*, *98*(2), 358–362. http://doi.org/10.1016/j.ajhg.2015.12.009

Mitsopoulos, P., Chang, Y.-H., Wai, T., König, T., Dunn, S. D., Langer, T., & Madrenas, J. (2015). Stomatin-like protein 2 is required for in vivo mitochondrial respiratory chain supercomplex formation and optimal cell function. *Molecular and Cellular Biology*, *35*(10), 1838–47. http://doi.org/10.1128/MCB.00047-15

Oláhová, M., Hardy, S. A., Hall, J., Yarham, J. W., Haack, T. B., Wilson, W. C., … Taylor, R. W. (2015). LRPPRC mutations cause early-onset multisystem mitochondrial disease outside of the French-Canadian population. *Brain*, *138*(12), 3503–3519. http://doi.org/10.1093/brain/awv291

Powell, C. A., Kopajtich, R., D'Souza, A. R., Rorbach, J., Kremer, L. S., Husain, R. A., … Minczuk, M. (2015). TRMT5 Mutations Cause a Defect in Post-transcriptional Modification of Mitochondrial tRNA Associated with Multiple Respiratory-Chain Deficiencies. *American Journal of Human Genetics*, *97*(2), 319–328. http://doi.org/10.1016/j.ajhg.2015.06.011

Pronicka, E., Piekutowska-Abramczuk, D., Ciara, E., Trubicka, J., Rokicki, D., Karkucińska-Więckowska, A., … Płoski, R. (2016). New perspective in diagnostics of mitochondrial disorders: two years' experience with whole-exome sequencing at a national paediatric centre. *Journal of Translational Medicine*,

14(1), 174. http://doi.org/10.1186/s12967-016-0930-9

Ribeiro, S. M., Giménez-cassina, A., & Danial, N. N. (2015). Measurement of Mitochondrial Oxygen Consumption Rates in Mouse Primary Neurons and Astrocytes. *Methods in Molecular Biology*, *1241*, 59–69. http://doi.org/10.1007/978-1-4939-1875-1

Seahorse Bioscience. (2014). Normalizing XF metabolic data to cellular or mitochondrial parameters, User Guide. Retrieved from http://hpst.cz/sites/default/files/attachments/appnote-normalizing-metabolic-data.pdf

Shah-Simpson, S., Pereira, C. F. A., Dumoulin, P. C., Caradonna, K. L., & Burleigh, B. A. (2016). Bioenergetic profiling of Trypanosoma cruzi life stages using Seahorse extracellular flux technology. *Molecular and Biochemical Parasitology*, *208*(2), 91–95. http://doi.org/10.1016/j.molbiopara.2016.07.001

Stroud, D. A., Surgenor, E. E., Formosa, L. E., Reljic, B., Frazier, A. E., Dibley, M. G., … Ryan, M. T. (2016). Accessory subunits are integral for assembly and function of human mitochondrial complex I. *Nature*, *538*(7623), 1–17. http://doi.org/10.1038/nature19754

Sullivan, L. B., Gui, D. Y., Hosios, A. M., Bush, L. N., Freinkman, E., & Vander Heiden, M. G. (2015). Supporting Aspartate Biosynthesis Is an Essential Function of Respiration in Proliferating Cells. *Cell*, *162*(3), 552–563. http://doi.org/10.1016/j.cell.2015.07.017

Sun, N., Youle, R. J., & Finkel, T. (2016). The Mitochondrial Basis of Aging. *Molecular Cell*, *61*(5), 654–666. http://doi.org/10.1016/j.molcel.2016.01.028

Titov, D. V, Cracan, V., Goodman, R. P., Peng, J., Grabarek, Z., & Mootha, V. K. (2016). Complementation of mitochondrial electron transport chain by manipulation of the NAD+/NADH ratio. *Science*, *352*(6282), 231–235. http://doi.org/10.1126/science.aad4017

Van Haute, L., Dietmann, S., Kremer, L., Hussain, S., Pearce, S. F., Powell, C. A., … Minczuk, M. (2016). Deficient methylation and formylation of mt-tRNAMet wobble cytosine in a patient carrying mutations in NSUN3. *Nature Communications*, *7*(May), 12039. http://doi.org/10.1038/ncomms12039

Wallace, D. C. (2007). Why do we still have a maternally inherited mitochondrial DNA? Insights from evolutionary medicine. *Annual Review of Biochemistry*, *76*(1), 781–821. http://doi.org/doi:10.1146/annurev.biochem.76.081205.150955

Wallace, D. C. (2012). Mitochondria and cancer. *Nature Reviews Cancer*, *12*(10), 685–698. http://doi.org/10.1038/nrc3365

Weinberg, S. E., Sena, L. A., & Chandel, N. S. (2015). Mitochondria in the regulation of innate and adaptive immunity. *Immunity*, *42*(3), 406–417. http://doi.org/10.1016/j.immuni.2015.02.002

Wu, M., Neilson, A., Swift, A. L., Moran, R., Tamagnine, J., Parslow, D., … Chomicz, S. (2007). Multiparameter metabolic analysis reveals a close link between attenuated mitochondrial bioenergetic function and enhanced glycolysis dependency in human tumor cells. *The American Journal of Physiology - Cell Physiology*, *292*, 125–136. http://doi.org/10.1152/ajpcell.00247.2006.

Yao, J., Irwin, R. W., Zhao, L., Nilsen, J., Hamilton, R. T., & Brinton, R. D. (2009). Mitochondrial bioenergetic deficit precedes Alzheimer's pathology in female mouse model of Alzheimer's disease. *PNAS*, *106*(34), 14670–14675.

Zhang, J., Khvorostov, I., Hong, J. S., Oktay, Y., Vergnes, L., Nuebel, E., … Teitell, M. A. (2011). UCP2 regulates energy metabolism and differentiation potential of human pluripotent stem cells. *The EMBO Journal*, *30*(24), 4860–4873. http://doi.org/10.1038/emboj.2011.401

Zhang, J., Nuebel, E., Wisidagama, D. R. R., Setoguchi, K., & Hong, J. S. (2012). Measuring energy metabolism in cultured cells, including human pluripotent stem cells and differentiated cells. *Nature Protocols*, *7*(6). http://doi.org/10.1038/nprot.2012.048.Measuring

Zhou, W., Choi, M., Margineantu, D., Margaretha, L., Hesson, J., Cavanaugh, C., … Ruohola-Baker, H. (2012). HIF1α induced switch from bivalent to exclusively glycolytic metabolism during ESC-to-EpiSC/hESC transition. *The EMBO Journal*, *31*(9), 2103–2116. http://doi.org/10.1038/emboj.2012.71

Zong, W.-X., Rabinowitz, J. D., & White, E. (2016). Mitochondria and Cancer. *Molecular Cell*, *166*(3), 555–566. http://doi.org/10.1016/j.cell.2016.07.002

Fig. 1

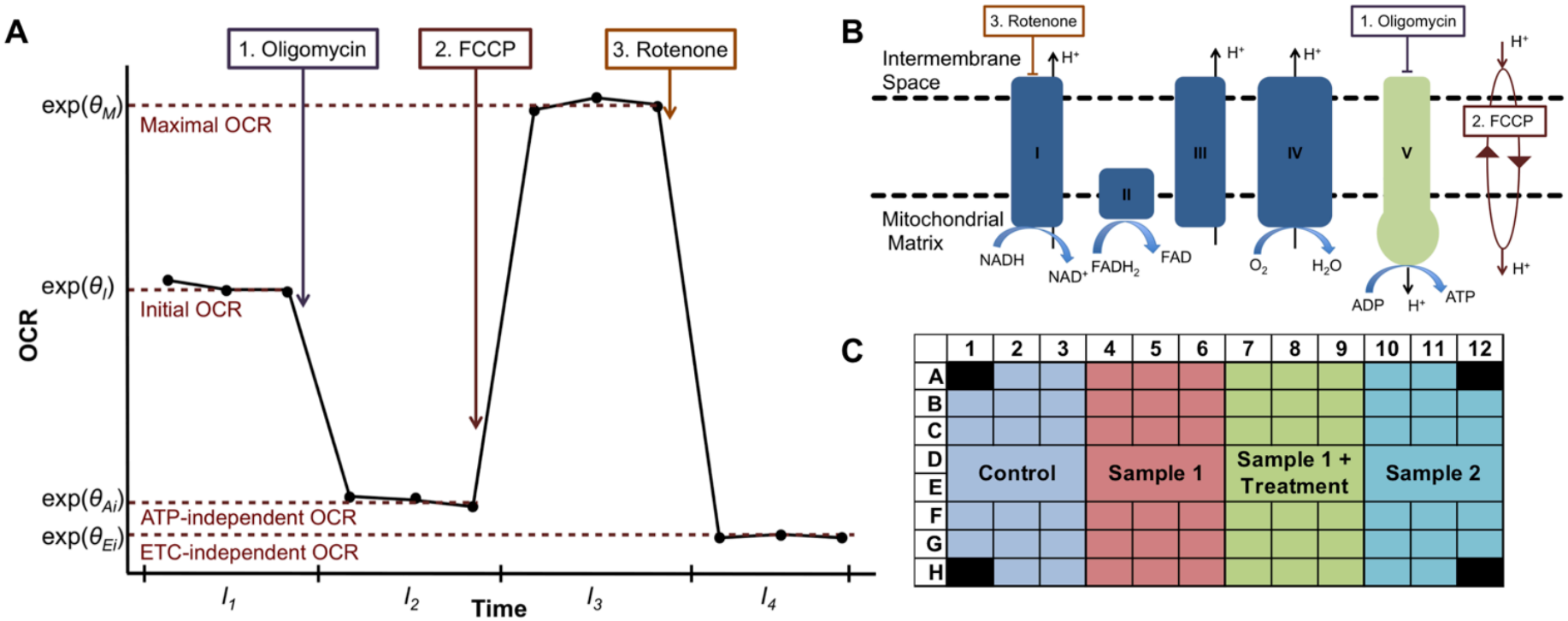

A
1. Oligomycin
2. FCCP
3. Rotenone
exp(θ_M)
Maximal OCR
exp(θ_I)
Initial OCR
exp(θ_{Ai})
ATP-independent OCR
exp(θ_{Ei})
ETC-independent OCR
OCR
I_1
I_2
Time
I_3
I_4
B
3. Rotenone
1. Oligomycin
2. FCCP
Intermembrane Space
Mitochondrial Matrix
H+
I
II
III
IV
V
NADH
NAD+
FADH2
FAD
O2
H2O
ADP
ATP
C
1 2 3 4 5 6 7 8 9 10 11 12
A B C D E F G H
Control
Sample 1
Sample 1 + Treatment
Sample 2

Fig. 2

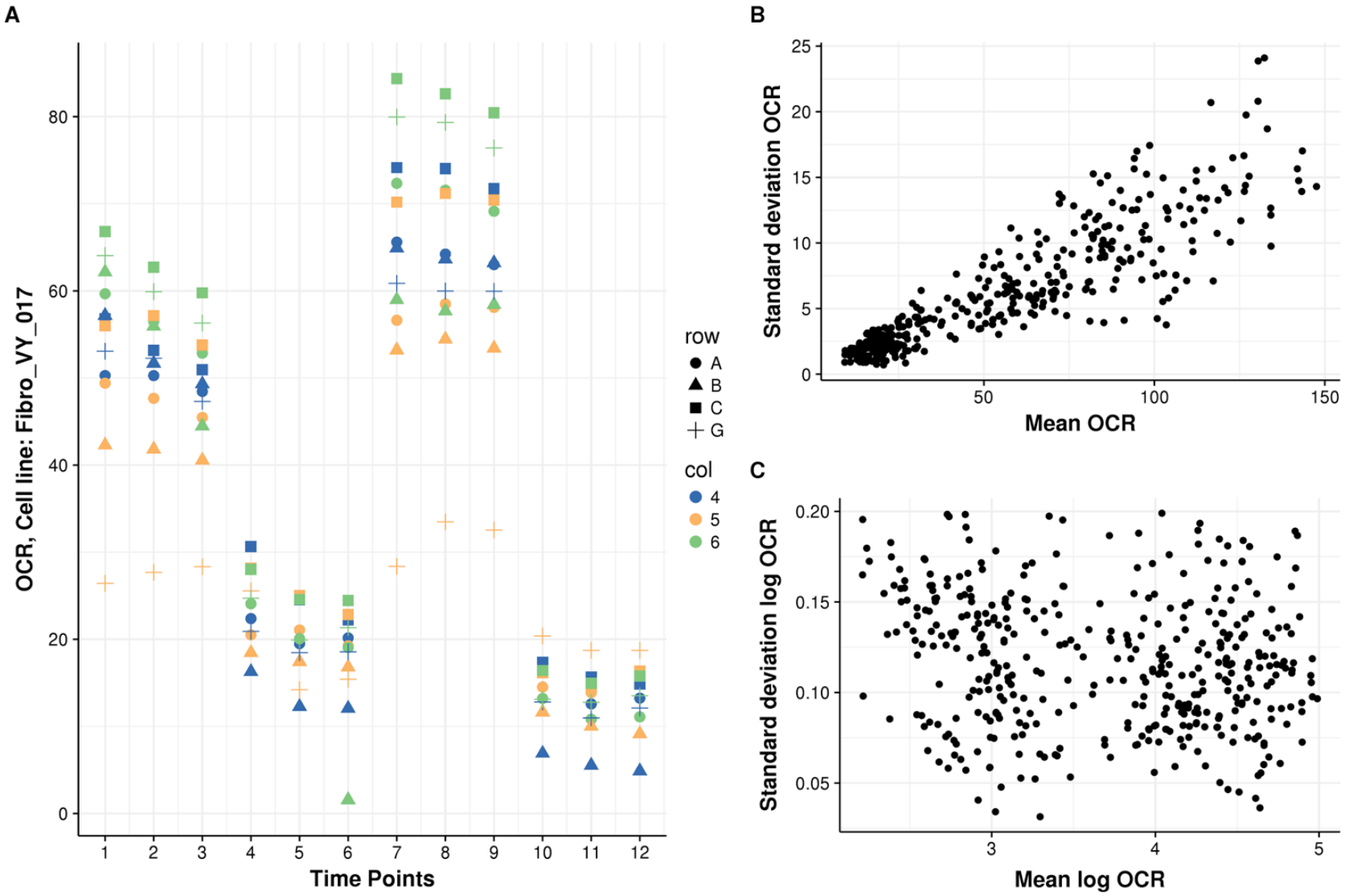

A
OCR, Cell line: Fibro_VY_017
Time Points
row
A
B
C
G
col
4
5
6
B
Standard deviation OCR
Mean OCR
C
Standard deviation log OCR
Mean log OCR

Fig. 3

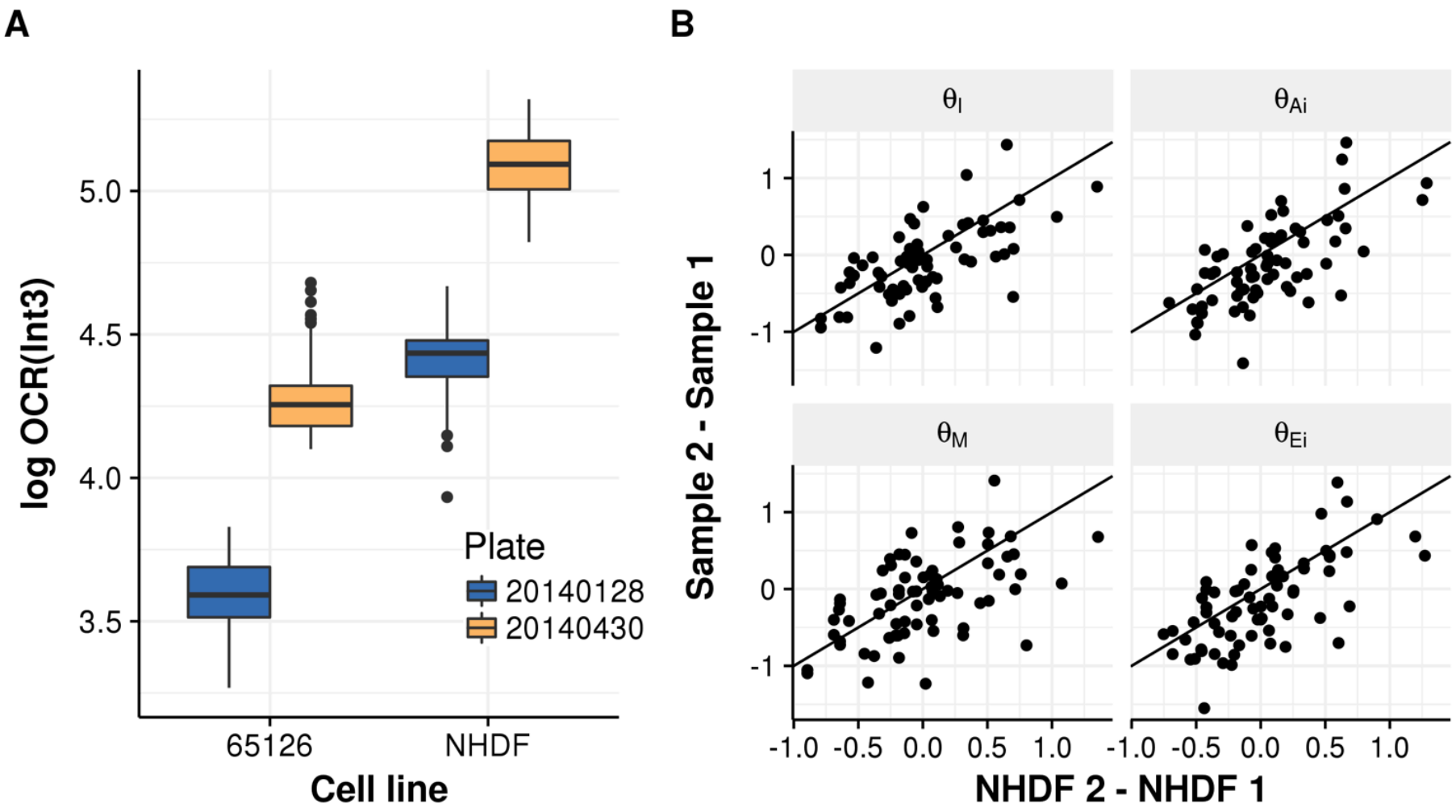

A
log OCR(Int3)
5.0
4.5
4.0
3.5
Plate
20140128
20140430
65126
NHDF
Cell line
B
θI
θAi
θM
θEi
Sample 2 - Sample 1
1
0
-1
-1.0
-0.5
0.0
0.5
1.0
NHDF 2 - NHDF 1

Fig. 4

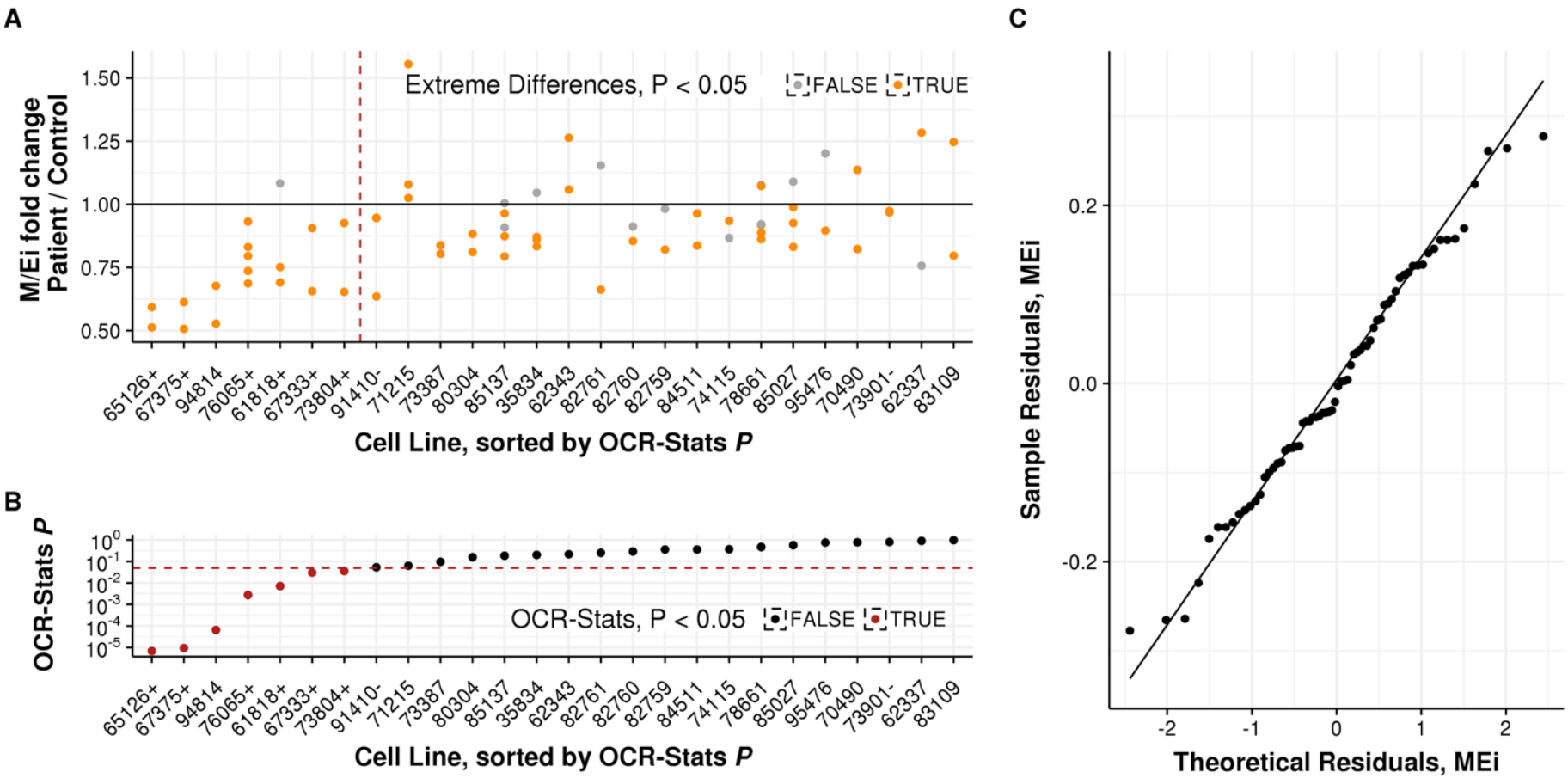
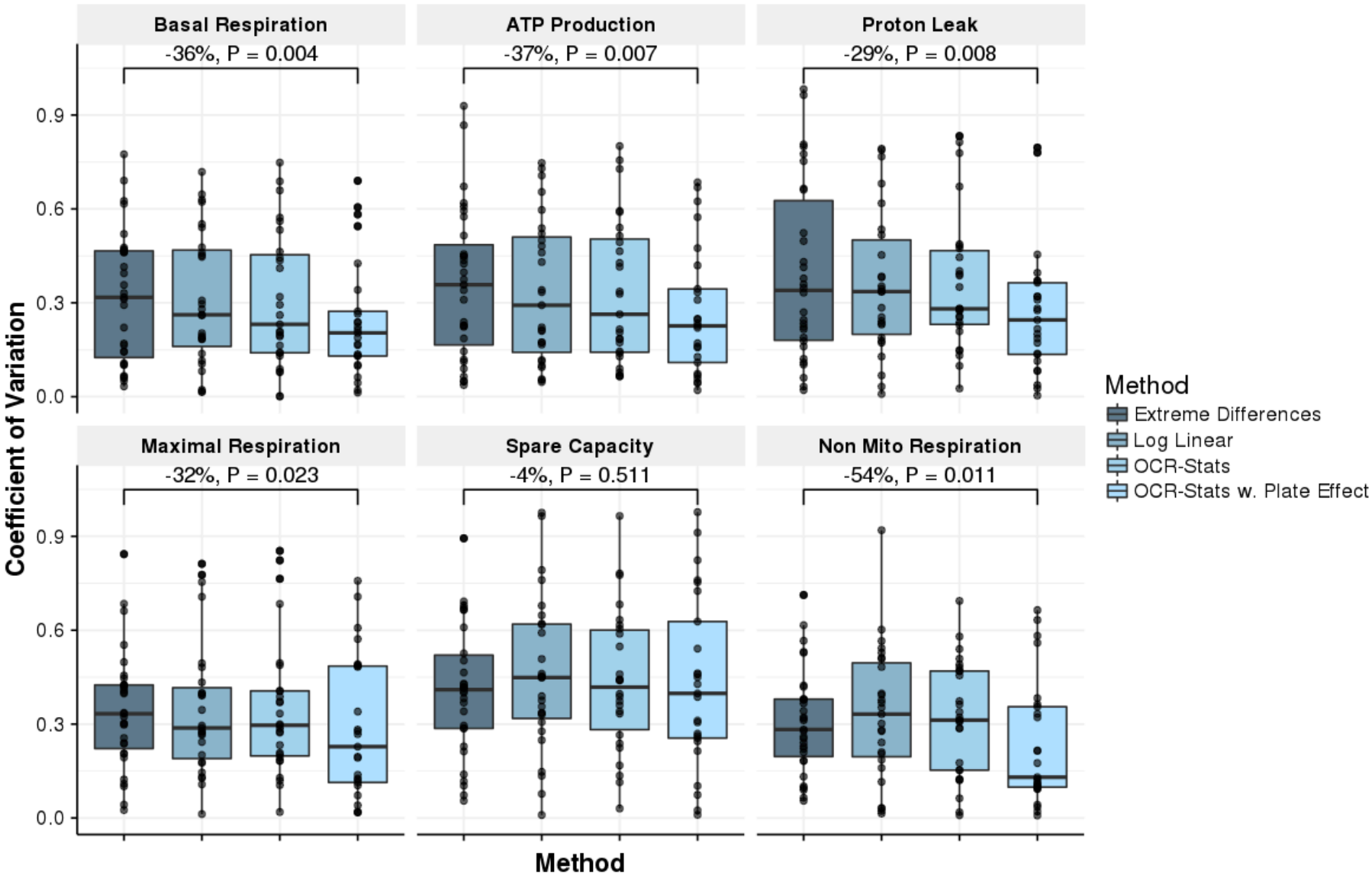

A
Extreme Differences, P < 0.05
FALSE
TRUE
M/Ei fold change
Patient / Control
1.50
1.25
1.00
0.75
0.50
65126+
67375+
94814
76065+
61818+
67333+
73804+
91410-
71215
73387
80304
85137
35834
62343
82761
82760
82759
84511
74115
78661
85027
95476
70490
73901-
62337
83109
Cell Line, sorted by OCR-Stats P
B
OCR-Stats P
OCR-Stats, P < 0.05
FALSE
TRUE
Cell Line, sorted by OCR-Stats P
C
Sample Residuals, MEi
Theoretical Residuals, MEi
0.2
0.0
-0.2
-2
-1
0
1
2

Fig. 5


Basal Respiration
-36%, P = 0.004
ATP Production
-37%, P = 0.007
Proton Leak
-29%, P = 0.008
Maximal Respiration
-32%, P = 0.023
Spare Capacity
-4%, P = 0.511
Non Mito Respiration
-54%, P = 0.011
Coefficient of Variation
0.0
0.3
0.6
0.9
Method
Method
Extreme Differences
Log Linear
OCR-Stats
OCR-Stats w. Plate Effect