



OPEN

## Artificial neural network applied to fragile X-associated tremor/ataxia syndrome stage diagnosis based on peripheral mitochondrial bioenergetics and brain imaging outcomes

Cecilia Giulivi<sup>1,2✉</sup>, Jun Yi Wang<sup>2,3</sup> & Randi J. Hagerman<sup>2,4</sup>

No proven prognosis is available for the neurodegenerative disorder fragile X-associated tremor/ataxia syndrome (FXTAS). Artificial neural network analyses (ANN) were used to predict FXTAS progression using data from 127 adults (noncarriers and *FMR1* premutation carriers with and without FXTAS) with five outcomes from brain MRI imaging and 22 peripheral bioenergetic outcomes from two cell types. Diagnosis accuracy by ANN predictions ranged from 41.7 to 86.3% (depending on the algorithm used), and those misclassified usually presented a higher FXTAS stage. ANN prediction of FXTAS stages was based on a combination of two imaging findings (white matter hyperintensity and whole-brain volumes adjusted for intracranial volume) and four bioenergetic outcomes. Those at Stage 3 vs. 0–2 showed lower mitochondrial mass, higher oxidative stress, and an altered electron transfer consistent with mitochondrial unfolded protein response activation. Those at Stages 4–5 vs. 3 had higher oxidative stress and glycerol-3-phosphate-linked ATP production, suggesting that targeting mGPDH activity may prevent a worse prognosis. This was confirmed by the bioenergetic improvement of inhibiting mGPDH with metformin in affected fibroblasts. ANN supports the prospect of an unbiased molecular definition in diagnosing FXTAS stages while identifying potential targets for personalized medicine.

Carriers of fragile X messenger ribonucleoprotein 1 (*FMR1*) premutation alleles are those with 55–200 CGG repeats in the 5'UTR gene region, whereas those with > 200 CGG repeats have the full mutation. Premutation carriers may present with neurodevelopmental, emotional, and psychiatric problems, all with lower severity than those with the full mutation, which is associated with fragile X syndrome. Many adult male premutation carriers develop an age-dependent neurological disorder known as fragile X-associated tremor/ataxia syndrome (FXTAS)<sup>1–3</sup>. This disease is characterized by parkinsonism, intention tremor, cognitive-behavioral changes, and cerebellar ataxia, among other features. It may also present with lower limb proximal muscle weakness, peripheral neuropathy, and autonomic dysfunction<sup>4</sup>. Neurocognition may also be impacted, ranging from mild memory deficits to clear dementia, attention deficits, and impairments in declarative and procedural learning, executive functioning, and working memory<sup>5–10</sup>. Other symptoms may include anxiety, depression, irritability, reclusive behavior, and socially inappropriate behavior<sup>7,8</sup>.

Neuropathologically, FXTAS is an inclusion body disease because of the detection of eosinophilic ubiquitin-positive intranuclear inclusions in both neurons and astrocytes, especially in the hippocampus, frontal cortex, and cerebellar nuclei<sup>11,12</sup>. Consistent with the frontal cortex and hippocampal inclusions<sup>11</sup>, FXTAS involves frontal cognitive, learning, and recall deficits. Only confirmatory diagnoses of FXTAS are obtained by detecting inclusions in post-mortem brain samples.

<sup>1</sup>Department of Molecular Biosciences, School of Veterinary Medicine, University of California Davis, Davis, CA, USA. <sup>2</sup>MIND Institute, University of California at Davis Medical Center, Sacramento, CA, USA. <sup>3</sup>Center for Mind and Brain, University of California Davis, Davis, CA, USA. <sup>4</sup>Department of Pediatrics, University of California at Davis Medical Center, Sacramento, CA, USA. ✉email: cgiulivi@ucdavis.edu

In terms of progression, FXTAS, like other neurodegenerative diseases, proceeds in six successive stages of physical disability (tremor and balance problems, falls) and level of interference with activities of daily living<sup>7</sup>. The characteristics of each stage include the following: Stage 1—subtle or uncertain signs; Stage 2—minor, but clear, tremor and/or balance issues with minimal daily living activities' interference; Stage 3—moderate issues with tremor and balance and at least occasional falls with significant interference of daily activities; Stage 4—severe tremor and balance problems requiring the use of a cane or walker aid; Stage 5—daily use of a wheelchair, and Stage 6—bedridden<sup>7</sup>. The diagnosis of FXTAS (definite, possible, probable) is established by neurological examinations and a combination of major and minor findings from clinical and MRI findings. Major FXTAS symptoms include (1) intention tremor of the hands, (2) gait ataxia, (3) detection of white matter hyperintensities (WMHs) by MRI in the middle cerebellar peduncles, and (4) FXTAS intranuclear inclusions in neuropathological studies. Minor FXTAS indications include (1) parkinsonism, (2) short-term memory and executive function deficits, (3) neuropathy, and (4) brain WMHs on T2-weighted MRI in the splenium of the corpus callosum, pons, and periventricular regions, along with generalized brain atrophy. Those with a definite FXTAS diagnosis are individuals with one major clinical symptom (1 or 2) and one major radiological finding (3 under major). Those with a probable FXTAS diagnosis have two major clinical signs or one minor clinical (from 1–3) and one major radiological finding (3). Those with a possible FXTAS diagnosis are individuals with one major clinical symptom and one minor radiological result<sup>13</sup>.

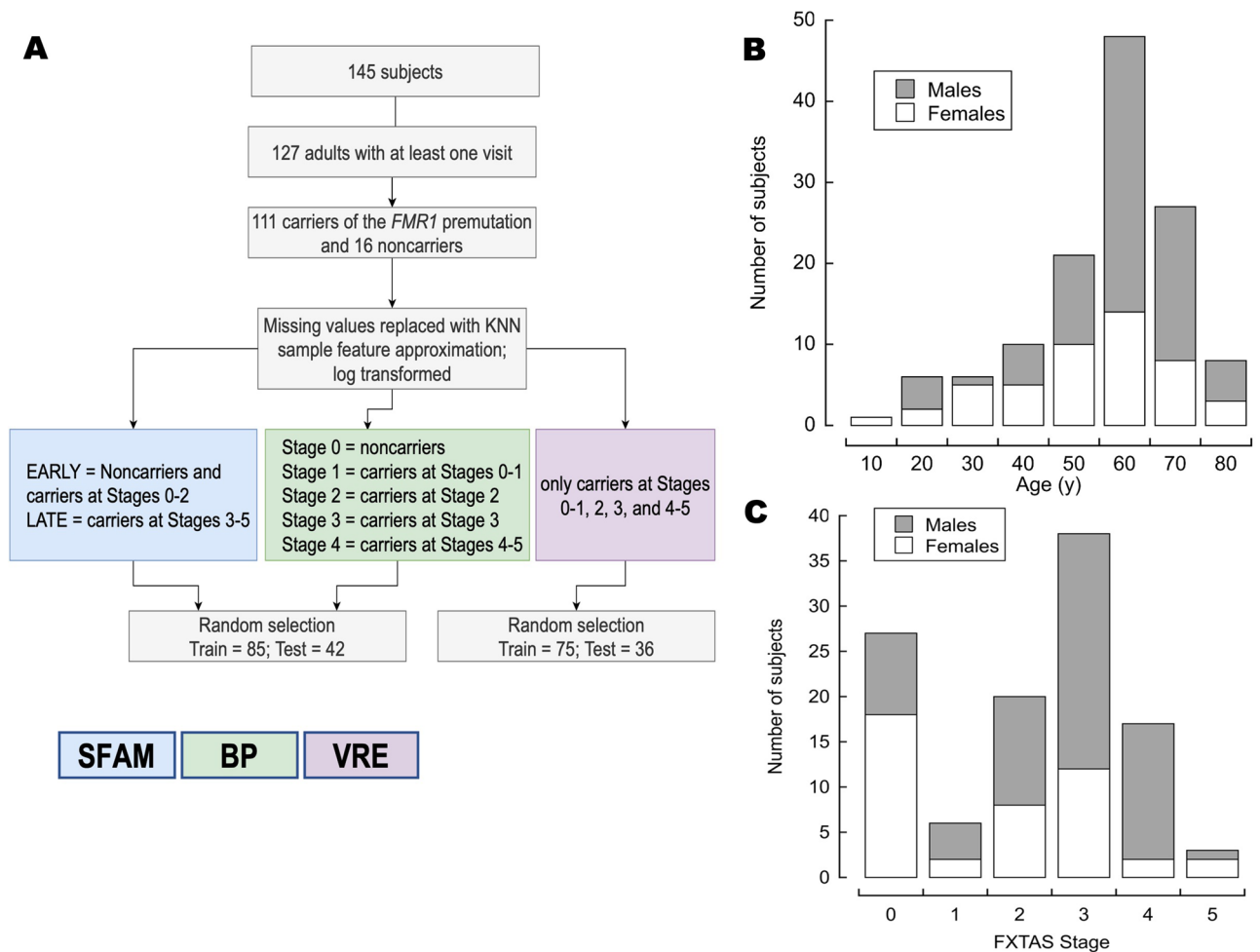
As indicated above, FXTAS diagnosis is based on neurological and MRI findings<sup>14</sup>. Some of these brain findings correlate with peripheral mitochondrial dysfunction<sup>15</sup>. The volume of white matter disease increases with FXTAS stages, following the trajectory or morbidity of the disease<sup>15–17</sup>; however, it has a low predictive value similar to the confirmatory diagnosis reached by the presence of brain inclusions.

Diagnosing FXTAS is challenging for clinicians because its phenotype overlaps with aging and other neurodegenerative diseases<sup>18</sup>. Furthermore, both onset and FXTAS disease progression are highly variable, and as of today, the progression of FXTAS trajectory cannot accurately be predicted. There is a great demand for early detection and more effective interventions to delay the onset of FXTAS or slow the course of its progression. Understanding the disease progression is a prerequisite for developing outcomes for predicting the course of FXTAS and determining effective intervention strategies.

Notably, no studies have used the power of combining molecular outcomes with clinical and imaging findings to develop a predictive model of FXTAS, even though this has been reported for Alzheimer's disease<sup>19</sup>. At the molecular level and consistent with numerous reports on placing mitochondrial dysfunction at the center of initiation and propagation of cognitive and behavioral abnormalities in neurological/neurodegenerative diseases<sup>20</sup>, our research group was the first to report mitochondrial dysfunction (reflected as a shift from mitochondrial bioenergetics to glycolysis, a.k.a. the Warburg effect) in carriers of the premutation even before developing overt FXTAS symptoms<sup>21–24</sup>, results expanded by our team and others in a variety of biospecimens<sup>21–23,25–29</sup> and mouse *Fmr1* premutation models<sup>26,30,31</sup>. Energy failure precedes the detection of the FXTAS hallmark, i.e., brain ubiquitin-positive intranuclear inclusions<sup>11,12</sup>, and some of the mitochondrial outcomes correlate with CGG repeats and severity of the phenotype<sup>21,22,25</sup>, highlighting the use of peripheral markers to mirror global mitochondrial dysfunction in disease-susceptible, post-mitotic tissues<sup>32–37</sup>. The brain is a highly aerobic organ, and within the cells that constitute the CNS, neurons rely heavily on mitochondrial ATP. Furthermore, considering the critical role mitochondria play in neurotransmitter metabolism<sup>28,29,38</sup>, it is not surprising that mitochondrial deficits likely compound the proteo- and RNA-triggered toxicities in CNS, contributing to FXTAS morbidity progression. This is well exemplified by our study of sulforaphane-treated fibroblasts from carriers at late stages whose bioenergetics were recovered by controlling their unfolded protein response and antioxidant capacities<sup>39</sup>. Although in vitro overexpression of the repeat-associated non-AUG (RAN)-mediated FMRpolyG product resulted in mitochondrial dysfunction<sup>40</sup>, no significant FMRpolyG levels in biospecimens from carriers have been found<sup>41,42</sup> even with evident mitochondrial dysfunction.

Artificial Neural Network (ANN), a branch of artificial intelligence, allows the detection of pre-disease states based on early warning signals (for instance, mitochondrial dysfunction) prognosticating complex or multifactorial diseases of unexpected and large-scale critical passages in biosystems<sup>43</sup>. ANN's detection of the “pre-disease state,” or in our case, progressing to more advanced FXTAS stages, is highly relevant because reversals are difficult or impossible to accomplish once the transition occurs, in contrast to relatively more direct intervention and prevention before the change occurs.

To this end, we turned to ANN to aid with diagnosing FXTAS stages. The ANN must first be “trained” by processing various input patterns and presenting each output result. The next step utilizes the trained ANN to recognize similarities under new input patterns and predict output patterns. In this way, ANN can detect early warning signs of crucial changes, defined as unexpected and large-scale state changes occurring in highly complex biological systems<sup>43,44</sup>. Based on our previous studies on early signs of mitochondrial dysfunction associated with executive dysfunction in carriers, we used ANN to extract similarities among a breadth of molecular and imaging outcomes and build a predictive model for FXTAS diagnosis and stage progression. In our model, disease development was considered as a critical change or transition<sup>19,44–46</sup> even with noisy input data such as those originated from a diverse population of human participants (even if all are carriers) as opposed to those emerging from more homogenous populations (e.g., specific murine breed or strains; artificially controlled living conditions). The hypothesis of this study was to test whether the diagnosis of FXTAS stages can be aided by ANN analysis of mitochondrial outcomes, MRI findings, sex, and age and if this predicted diagnosis would have better predictive value than the use of more traditional diagnosis protocols.



**Figure 1.** Experimental design and participants' sex, age, and FXTAS stage distribution. (A) Scheme of the experimental design utilized in this study. From 145 eligible participants, we filtered those aged 18 years old and older with at least one visit to the MIND. The remaining 127 subjects were either carriers of the *FMR1* premutation ( $n = 111$ ) or noncarriers ( $n = 16$ ). Missing values were replaced by using the sample feature k-nearest neighbor algorithm, and the final values were log-transformed. The values were analyzed by using three different algorithms represented with colored boxes: blue, green, and purple boxes indicate that the data were tested with the simplified fuzzy adaptive resonance theory map (SFAM), backpropagation (BP), and visual rule extraction (VRE) algorithms, respectively. SFAM was used to predict EARLY vs. LATE FXTAS stages, a backpropagation algorithm to test numerical stage assignment, and VRE to unveil the combination of critical factors that predict stage assignment. More details were included in the text. Age and sex (B), and sex and FXTAS stage (C) distribution of the participants.

## Results

**Study design and subject characteristics.** This study included 127 participants, 18 years and older, of both sexes, carriers ( $n = 111$ ) and noncarriers ( $n = 16$ ) of the *FMR1* premutation that had at least one visit to the MIND, with bioenergetic data from PBMC and fibroblasts, and MRI findings (Fig. 1). If more than one visit were available for a patient, only the first was used in the analysis. To maximize the statistical power, the small number of carriers at FXTAS Stage 1 ( $n = 6$ ) were combined with those at Stage 0 ( $n = 27$ ) and grouped under "Stage 1" ( $n = 33$ ), whereas those at FXTAS Stage 5 ( $n = 3$ ) were combined with those at Stage 4 ( $n = 17$ ) and labeled as "Stage 4" ( $n = 20$ ; Fig. 1A). The ratio of male to female carriers was 1.65 (79 males, 48 females) with a similar median age (median, range for males: 64, 18–86; for females 65, 25–86; Fig. 1B). As expected for an X-linked genetic disorder and consistent with other studies in the field<sup>47,48</sup>, women are usually less affected than men, with a median FXTAS stage lower than men (51% of male carriers at Stage 3 vs. 43% of females at Stage 0; Fisher exact test  $p$ -value (2-sided) = 0.004; Fig. 1C).

**ANN to diagnose FXTAS stages.** The goal of this study was to utilize a neural network to discover patterns or a combination of imaging and bioenergetic outcomes that will allow categorizing carriers in FXTAS stages. A neural network is a software (or hardware) simulation of a biological brain (also named as Artificial Neural Network or "ANN"). The purpose of a neural network is to learn to recognize patterns in the data that

**Figure 2.** Artificial neural network prediction of FXTAS stages. The ANN utilized age, sex, imaging, and bioenergetic outcomes from PBMCs and fibroblasts from carriers and noncarriers to predict EARLY or LATE FXTAS stages by applying an SFAM algorithm (A), predicted FXTAS stage by utilizing the backprop algorithm (B), and VRE to identify key parameters that predict stage classification (C,D). Panel (A): Confusion matrix indicating the actual (diagnosed) stage and the predicted one (via SFAM). The number in the boxes are subjects in that category. In grey, subjects for which the diagnosed EARLY or LATE classification matched the predicted one. The leftmost column shows predicted errors, and the bottom row shows actual errors. Overall error is 16.67%. Panel (B): Data were analyzed by using the backprop algorithm. The results (diagnosed vs. predicted FXTAS stage) were visualized in a scatter plot, and a linear regression fitting was used. The correlation equation is shown; below it, the Pearson's correlation is shown [R (degrees of freedom) = value; p-value]. The dotted line marks the 95% confidence interval for the regression; M and F, male and female, respectively. Panel (C): Decision tree obtained by analyzing the carrier data with the VRE algorithm. Abbreviations for imaging outcomes: *mv\_WMH* normalized white matter hyperintensities and *mv\_WB* normalized whole brain volume. For bioenergetic outcomes: GP-sustained ATP production = glycerol-3-phosphate-sustained oxygen-linked ATP production; CCO activity (or Complex IV activity) was expressed as nmol oxygen consumed  $\times$  (min  $\times$  million cells)<sup>-1</sup>, and the CCO/CS ratio = CCO activity normalized to that of citrate synthase used as a proxy for mitochondrial mass. SRC or spare respiratory capacity is calculated as the ratio of oxygen uptake under State 3u to that of State 3. It indicates the excess or reserve capacity of mitochondria to respond to stress (see more details under text). Those bioenergetic outcomes with the abbreviation FB were obtained with fibroblasts (e.g., FB CCO/CS), while all others refer to PBMCs. Colored boxes indicate those with stage prediction, Stages 0–1 in grey, Stage 2 in yellow, Stage 3 in orange, and Stages 4–5 in red. Boxes indicated with A, B, C, and D are used to discuss the differences in predicting the diagnosis of Stages 0–1 vs. 2–3 (A vs. B) and Stage 3 vs. 4–5 (C vs. D). Panel (D): Confusion matrix of diagnosed and predicted FXTAS stage resulting from VRE data analysis. Diagnosed and predicted stages ranged from 0–1, 2, 3, and 4–5. Box numbers indicate the number of subjects in that category. Boxes in grey indicate those matching the predicted and diagnosed stage. The leftmost column shows the prediction errors, whereas the bottom row shows the diagnosed errors. The overall error of the VRE analysis is 41.67%, with the most mismatches biased towards higher predicted stages.

are not apparent through a simple analysis. Consequently, software that learns through this process is named "Artificial Intelligence." The neural network must be "trained" by processing many input patterns resulting in outputs from each input pattern. Once trained, the neural network will make predictions by detecting similar patterns when presented with a new situation. The training of the ANN is a critical step in applying the ANN, as the predictions will rely on this training. To select a training set size, we performed a learning curve analysis<sup>49</sup> by utilizing carrier status, FXTAS diagnosis, stage, sex, age, five MRI findings, and 22 units of mitochondrial outcomes for each cell type (PBMC and fibroblasts). Basically, a relatively small number of samples from our data was randomly chosen to train the ANN as detailed before<sup>50</sup>. The trained ANN was then used to predict the pattern of a randomly sampled test set not included in the training set. Then, iterative steps were taken to increase the size of our training sample while keeping that of the test set. By tracking the degree to which predictive accuracy on the fixed test set increased with the training set size, we obtained the needed size of training data until the differences between the network classifications and the clinical diagnoses became acceptable. The learning curve analysis was performed multiple times with different test sets to reduce variation in predictive accuracy. The resulting sample sizes were 85 subjects for training and 42 for testing.

The experimental outcomes included in our study were constituted by MRI imaging and bioenergetic findings. Of the five MRI imaging outcomes, two were binary variables constituted by the presence or absence of WMHs in the middle cerebellar peduncles and the cerebellum. The other three were the volumes of the WMHs, whole brain, and ventricles normalized by intracranial volume. These outcomes are supported by previous studies on brain MRI imaging changes associated with FXTAS progression, which include white matter lesions, generalized brain atrophy, ventricular enlargement, and accumulation of iron deposits in the subcortical nuclei<sup>16,51–54</sup>. While some of these changes can be detected before overt development of FXTAS symptoms, including loss of cerebellar and brain stem volume<sup>55,56</sup>, changes in white matter tracts<sup>57</sup>, and modifications in functional MRI recordings<sup>58</sup>, microstructural white matter disease in the middle cerebellar peduncles and the genu of the corpus callosum correlates with the severity of executive dysfunction and impairments in processing speed in carriers<sup>59</sup>.

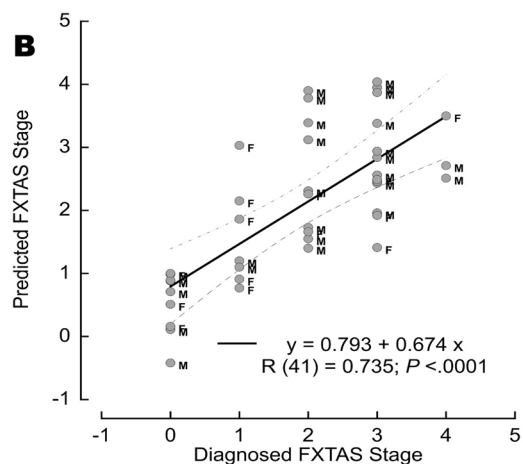
The 22 bioenergetic outcomes included 11 from each cell type tested (PBMC and primary fibroblasts). Since many parameters can assess different aspects of mitochondrial dysfunction, here we provide a brief justification for the 11 chosen outcomes. Four of them were constituted by functional assays on ATP production sustained by various substrates and specific inhibitors allowing the evaluation of different segments of the electron transport chain. Among them, oxygen-linked ATP production was evaluated with glucose–glutamine in intact cells, and malate–glutamate, succinate, and glycerophosphate in permeabilized cells allowing mitochondrial access to these cell-impermeable substrates. Mitochondrial oxidative stress was estimated by evaluating the mitochondrial oxygen consumption under non-phosphorylating conditions (when ADP is limiting and its phosphorylation is almost null) normalized or not by the respiration with glucose–glutamine. Essentially these parameters reflect electrons leaking through the electron transport chain that, when combined with oxygen, generate reactive oxygen species (ROS)<sup>60</sup>. As inner membrane leakiness may impair ATP production, membrane integrity of mitochondria<sup>61</sup> was assessed by evaluating the coupling between electron transfer and ATP production<sup>62</sup>. Under conditions of maximum oxygen uptake, the uncoupled respiratory control ratio (RCRu) is the ratio between maximal respiratory capacity, or State 3u, and the residual respiration after inhibition of ATP synthesis with oligomycin, or State 4o. In contrast, the ratio of phosphorylating to nonphosphorylating mitochondria is RCR. A critical parameter evaluated here was the spare respiratory capacity (SRC) This parameter is a key indicator



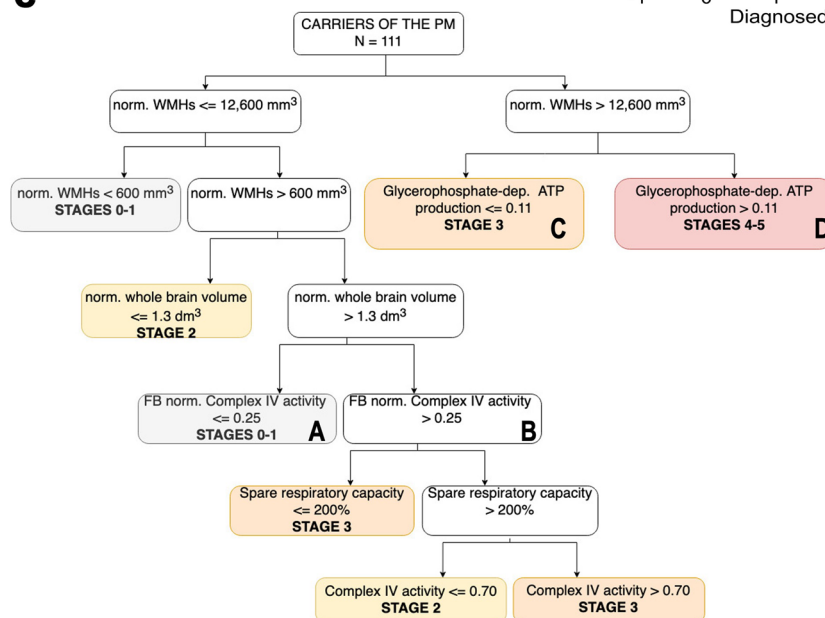
**A**

Prediction	Actual		Prediction totals (n)	Prediction error (%)
	EARLY (n)	LATE (n)		
EARLY (n)	22	3	25	12
LATE (n)	4	13	17	23.53
Actual totals (n)	26	16	42	16.67
Actual errors (%)	15.38	18.75	16.67	

**B**



**C**



**D**

Predicted Stage (n)	Diagnosed stage (n)				Predicted totals (n)	Prediction error (%)
	0-1	2	3	4-5		
0-1	6	3	1	0	10	40.00
2	2	3	2	0	7	57.14
3	2	0	6	0	8	25.00
4-5	0	1	4	6	11	45.45
Diagnosed totals (n)	10	7	13	6	36	41.47
Diagnosed errors (%)	40.00	57.14	53.85	0.00	41.67	

of neuronal susceptibility to cellular stress<sup>63</sup> and estimates the capacity of mitochondria to adapt to increasing energy demands. Even small decreases in spare or reserve respiratory capacity enhance neuronal vulnerability, predisposing the tissue to energy deficits and neurodegeneration<sup>27</sup>. This index was calculated as the ratio (in percentage) between the oxygen uptake under uncoupling conditions and that under glucose–glutamine. Finally, the activities of the terminal oxidase of the electron transport chain Complex IV and citrate synthase were evaluated. From all five Complexes, Complex IV is implicated in ataxia and other neurodegenerative diseases of all five complexes<sup>64–66</sup>. Citrate synthase is a Krebs' cycle enzyme whose activity does not normally fluctuate with physiological or pathological conditions. As such, it can be used as a surrogate marker for mitochondrial mass<sup>27,36,36,67</sup>. Conversely, mtDNA copy number is an unlikely mitochondrial mass marker because it changes under several conditions and diseases resulting in either depletion (e.g., premature aging, failure to antagonize oxidative stress)<sup>68</sup> or over-replication, as observed in FXTAS<sup>23,27</sup> and others<sup>23,27,69–77</sup>.

To utilize ANN to analyze the data, we first applied a simplified fuzzy adaptive resonance theory map (SFAM) to predict a class, i.e., in our case, EARLY or LATE stages. SFAM uses numeric inputs (age, imaging and bioenergetic outcomes, sex, and carrier status as binary ones) to predict a class or category (i.e., EARLY or LATE). In lay terms, it provides a “fuzzy lookup” into the training database based on the similarity of coordinates within the multidimensional input space. For instance, if SFAM analyzes the parameters of a new FXTAS patient, it will return the diagnosis of the most similar pattern in the training data (early or late stage). For this algorithm, the ANN is constituted by a network of artificial neurons (also known as “nodes”). The strength of their connections is assigned a value (weight). Training begins with just one hidden node whose weights are set equal to the first record, and prediction is set equal to the class (e.g., Early) of the first record. Similarly, a new node is created whenever a new class is encountered (e.g., Late). The node whose weights best match the current input determines the Early or Late stage prediction if the degree of match exceeds the vigilance threshold value. If this prediction is correct, the weights of this winning node are adjusted according to this input. If the prediction is wrong or the vigilance threshold is not achieved, a new node is created with weights and predictions equal to this record. The algorithm learns through this iterative process justifying the label “Artificial Intelligence.”

In our dataset, *Early* stages included 9 noncarriers and 17 carriers at FXTAS Stages 0–2 (median age [95% CI] = 62.3 years [56–69]; 10 females and 16 males; 7 carriers at Stage 1 and 10 carriers at Stage 2) whereas *late* stages included 16 carriers at Stages  $\geq 3$  (median age [95% CI] = 64 [62–66]; 5 females and 11 males; 13 at Stage 3 and 3 at Stages 4–5; Fig. 1A blue box and Fig. 2A). This algorithm accurately predicted 84.6 and 81.3% of those at early and late stages, respectively. Overall, this classification resulted in an 83.3% sensitivity (95% CI 68.64–93.03), 100% specificity (95% CI 66.37–100), and an overall 86.27% accuracy (95% CI 73.74–94.30). The small number of mismatches between predicted and actual stages ( $n = 7$ , F/M = 2/5; 3 at late and four at early) resulted from the reclassification of 4 carriers diagnosed at early stages to late-predicted ones and the remaining three in the opposite direction.

Second, we analyzed the same test set of 42 participants (median age [95% CI] = 63 years [58–69]; 9 noncarriers, 7, 10, 13, and 3 at Stages 1, 2, 3, and 4; 15 females and 27 males) by using a backpropagation neural network (Fig. 1A, green box). This is the most common and robust type of neural network because solves classification problems. In our case, it assigned a numeric value to each diagnosed class (in our case stage), predicting a stage<sup>78–80</sup>, not a combined class (Early or Late) as it was when we used SFAM (vide supra). Training begins with the weights of all input parameter set to random numbers. For each input parameter, the predicted stage is compared to the actual stage, and the algorithm adjusts the weights to shift the predicted stage closer to the diagnosed one. Many cycles are made through the entire training data set, with the weights continually adjusted to produce more accurate stage predictions.

The analysis done by the backpropagation algorithm showed a strong and statistically significant association between the predicted and actual diagnoses of FXTAS stages (Fig. 2B). To calculate the error of this approach, we considered a mismatch between predicted and diagnosed stages if the absolute numerical difference was  $\geq 2$ . Of the 42 participants, four carriers were identified as mismatches between the predicted and diagnosed stages (9.5% error). Most misclassified ones were the noncarriers (6 of 9 classified at Stage 1). The three carrier mismatches resulted from their reclassification to higher predicted stages ( $n = 3$ ; M/F = 2/1). The sensitivity of this algorithm to detect the diagnosed stage was 87.88% (95% CI 71.80–96.60), with a specificity of 33.33% (95% CI 7.49–70.07), and an accuracy of 76.19% (95% CI 60.55–87.95).

Finally, a Visual Rule Extraction (VRE) algorithm<sup>81,82</sup> was utilized to ascertain those factors (or their combination) important for diagnosing a stage that did not rely on clinical or neuropsychiatric evaluations. The algorithm behind this analysis (inductive rule extraction) is related to Machine Learning, Knowledge Discovery, Expert Systems, and Artificial Intelligence and is also called “Decision Tree Classification.” This approach is based on the concept of entropy, a physicochemical term used to estimate randomness, disorder, or uncertainty in a population. Its use in artificial intelligence dates back more than 50 years ago<sup>82,83</sup>. Numeric and text input parameters (in our case, age, sex, stage, imaging and mitochondrial outcomes) are loaded to discover rules contained within the data to predict a class or category (in our case, stage). The VRE algorithm (based on an optimized version of the C4.5 algorithm<sup>84</sup>) builds the branches of the “decision tree” based on the lowest entropy (or randomness) or the most reliable information.

As with the previous ANN analyses, the VRE was trained with a stratified random sampling of 75 carriers of the *FMRI* premutation (14, 10, 38, and 13 at Stages 1, 2, 3, and 4; F/M = 28/47; median age [95% CI] = 66.2 years [63.5–68.9]), which represented roughly 1/3 of all carriers (Fig. 1A, purple box). The VRE algorithm analysis resulted in a decision tree with 15 nodes with an overall confidence of 41.67% (Fig. 2C,D) that relied on the following parameters: two outcomes from imaging, i.e., WMHs normalized by intracranial volume (nv\_WMH) and normalized whole brain volume (nv\_WB), and four bioenergetic outcomes, one from fibroblasts (Complex IV activity normalized to that of citrate synthase) and three from PBMcs (spare respiratory capacity, Complex IV activity, and glycerol-3-phosphate-sustained ATP production; Fig. 2C).

The decision tree indicated that WMHs normalized by intracranial volume are critical to distinguish carriers at early vs. late stages with those with a value  $\leq 600 \text{ mm}^3$  at Stages 0–1. Those with values between 600 and  $12,600 \text{ mm}^3$  may end up at Stage 2 if the normalized whole brain volume is  $\leq 1.3 \text{ dm}^3$ . If higher and Complex IV activity normalized to that of citrate synthase in fibroblasts is  $\leq 0.25$ , the carriers are classified in Stages 0–1. If this last ratio is higher than 0.25, the spare respiratory capacity in PBMCs up to 200% classifies carriers at Stage 3. If SRC  $> 200\%$ , then the defining factor is Complex IV activity in PBMCs. If the lymphocytic activity of Complex IV is  $\leq 0.70 \text{ nmol oxygen consumed} \times (\text{min} \times 10^6 \text{ cells})^{-1}$ , carriers are at Stage 2, whereas if above this threshold, carriers are at Stage 3. Carriers with WMHs normalized by intracranial volume above  $12,600 \text{ mm}^3$  may end up at Stages 3 or 4–5 depending on the mitochondrial ATP production with the substrate glycerol-3-phosphate in PBMCs ( $> 0.11 \text{ nmol oxygen consumed} \times (\text{min} \times 10^6 \text{ cells})^{-1}$  at Stage 4 or  $\leq 0.11$  at Stage 3).

**Validation of ANN-mediated FXTAS stage diagnosis.** We validated the above results from the decision tree using principal component analysis by utilizing all carriers ( $n = 111$ ) and all outcomes (22 bioenergetic outcomes for each cell type, 5 MRI findings, sex, and age) vs. those six parameters selected by the VRE algorithm. To this end, we used principal component analysis because it allows analyzing large datasets with large numbers of parameters per subject by reducing the dataset complexity to increase the interpretability of data while preserving the maximum information. The data are linearly transformed into a new coordinate system where most of the data variation is observed with fewer dimensions than the original data (usually in a 2-dimensional plot) to visually identify clusters of closely related data points (in our case, stage clusters). The strategy was to use a statistical technique to visualize whether the VRE algorithm produced a better cluster separation of stages when using all vs. VRE selected parameters and utilizing all carriers, not just the test set.

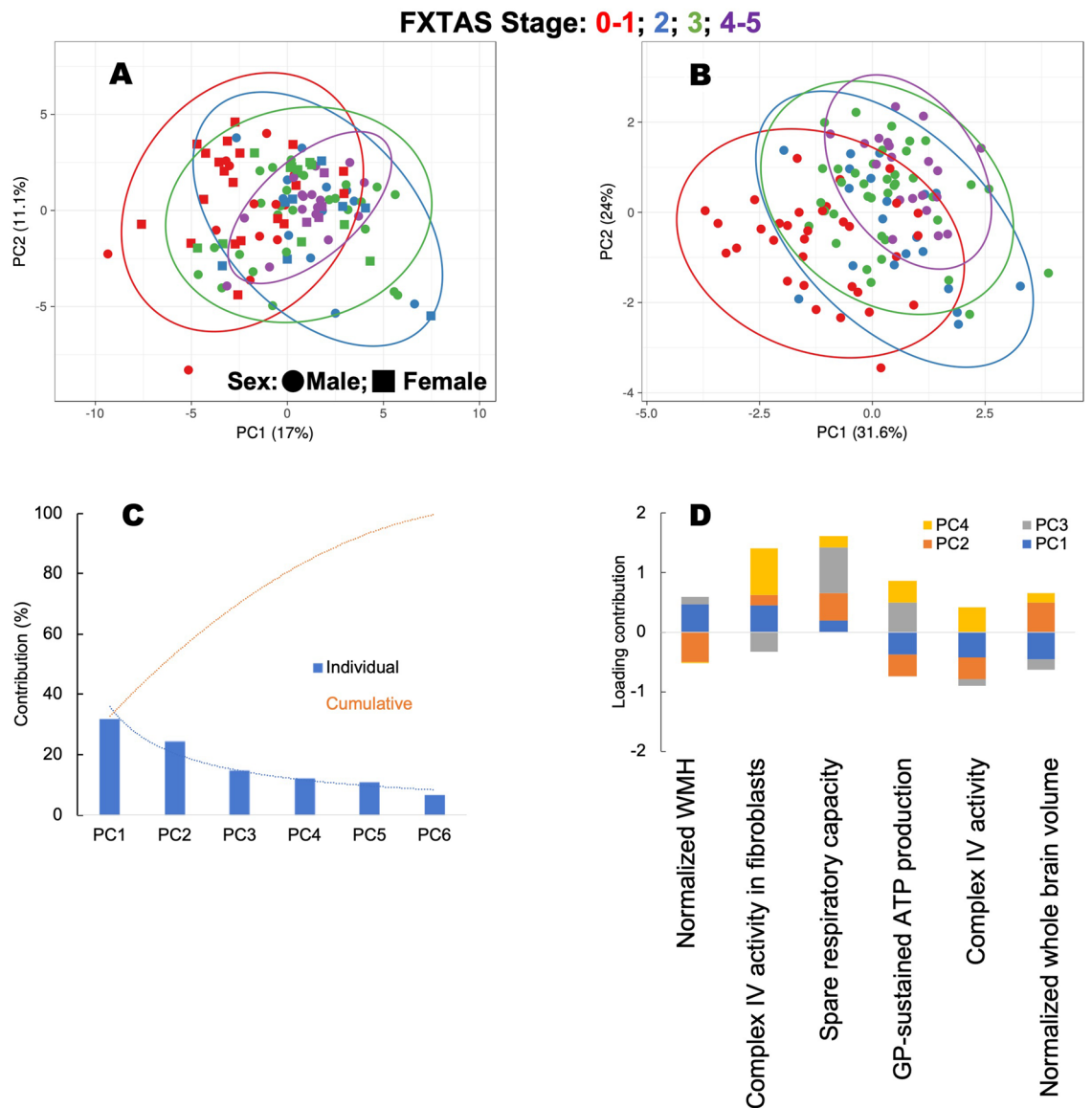
When all carriers and parameters were utilized, no clear separation of the groups was noted in a principal component analysis score plot for the first two principal components. As indicated above, this plot is an orthogonal linear transformation that visualizes data to a new coordinate system. The greatest variance of the data lies on the first coordinate (X-axis) or first principal component (17%), with the second greatest variance (second principal component; 11.1%) on the second coordinate (Y-axis; Fig. 3A). Contrary to the analysis performed with all carriers and all parameters, when the six parameters identified by the VRE algorithm were used, better separation of the groups was achieved with larger variances in principal components 1 (31.6%) and 2 (24%; Fig. 3B). A scree plot was used to determine which parameters from those six VRE-selected ones needed to be retained as principal components to obtain the best separation of stages. This analysis indicated that all six were needed because considering the first two principal components, the cumulative variance contribution was close to 60% (Fig. 3C). Imaging and bioenergetic outcomes contributed mostly to principal components 1 and 2, but the bioenergetic ones also contributed to principal components 3 and 4 (Fig. 3D). These results indicated that the six parameters identified by the VRE algorithm were adequate to identify clusters of carriers at different stages even when using the larger set of carriers ( $n = 111$ ).

**Biological implications of the VRE predictive value.** Notably, the decision tree classification predicted that Complex IV activity in fibroblasts—normalized to that of citrate synthase—from participants at Stages 0–1 (Fig. 2C indicated as box A) was lower than in those at Stages 2–3 (Fig. 2C marked as box B). However, these results should not be interpreted as an “increase in mitochondrial function” at higher stages. A comprehensive analysis of mitochondrial bioenergetic data revealed that although Complex IV activity was higher in those at Stages 2–3 vs. Stages 0–1, citrate synthase activity was lower (Fig. 4). More importantly, malate-glutamate-fueled ATP production and the ROS and proton leak were higher in those at higher stages. Consequently, those at higher stages presented a lower coupling (RCR) and higher I/III and lower (I/IV and III/IV) Complex ratios (Fig. 4). The lower citrate synthase activity in fibroblasts from subjects at Stages 2–3 (by 30%) suggests lower mitochondrial mass. The other possibility (shrinking of Krebs’ cycle without changing the mitochondrial mass) does not seem viable as no changes in another Krebs’ cycle enzyme, succinate dehydrogenase or Complex II, were noted (Fig. 4 bottom panel; oxygen-linked ATP production sustained by succinate;  $p = 0.251$ ).

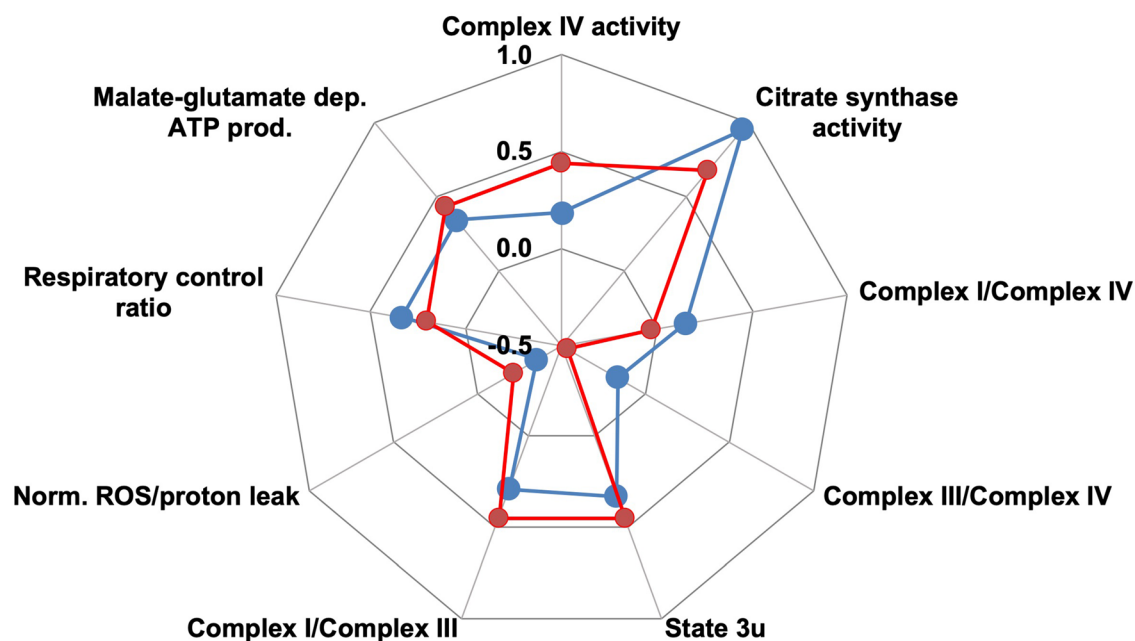
This indicates that fibroblasts from carriers at Stages 2–3 vs. those at Stages 0–1 show lower mitochondrial mass, increased mitochondrial ROS, and an altered electron transfer across Complexes. This last observation is relevant because an altered mitochondrial and nuclear-encoded proteome (mitochondria-nuclear balance) reflects not only an inefficient ATP production because OXPHOS Complexes are not synthesized proportionally to each other, but also a protein imbalance that could lead to activation of the mitochondrial unfolded protein response (UPRmt)<sup>85,86</sup>.

Mitochondrial bioenergetic differences were observed between carriers at Stages 0–1 vs. Stages 2–3 (Fig. 2C, boxes labeled as A and B) and between those at Stage 3 vs. Stages 4–5 (Fig. 2C, boxes labeled as C and D). In this regard, a higher glycerophosphate-fueled ATP production was observed in PBMCs from carriers at Stages 4–5 than those at Stage 3. Confirming the decision tree classification, those PBMCs from carriers at higher stages showed higher glycerol-3-phosphate-fueled ATP production (to a lesser extent with succinate) than those at Stage 3 (Fig. 5A). Similar to bioenergetic data from patients from Stages 0–1 vs. those at Stages 2–3, differences between Stages 3 vs. 4–5 included a higher mtROS production and altered Complex ratios, but with no changes in either Complex IV or citrate synthase activities.

The glycerol-3-phosphate-fueled ATP production is mediated by the coordinated action of the mitochondrial glycerol-3-phosphate dehydrogenase (mGPDH) with a distinct, cytosolic cGPDH to transfer reducing equivalents from NADH—generated at the glycolytic step catalyzed by glyceraldehyde dehydrogenase—into the mitochondrial respiratory chain through the glycerophosphate shuttle. The reoxidation of NADH, critical for sustaining glycolysis, is achieved not only by this shuttle but also by the malate-aspartate shuttle and lactate dehydrogenase<sup>87</sup>. To estimate the contribution of the shuttles glycerophosphate and malate-aspartate (both



**Figure 3.** Principal component analysis of FXTAS stages based on sex, age, MRI findings, and bioenergetic outcomes. PCA (performed with ClustVis 2.0<sup>119</sup>) was applied to the test dataset by utilizing all 111 carriers and either all outcomes as input (panel A) or only those selected by VRE algorithm<sup>81,82</sup> (panel B). Outcomes selected for panel B were taken from the decision tree in Fig. 2C. The data for constructing each PCA plot were scaled using a unit variance scaling applied to rows, and a singular value decomposition with imputation was used to calculate the principal component analysis method (for both panels). Other options were set as follows: no data transformation was used as the data were in the log form, no collapse of columns with similar annotations was performed, and no removal of constant columns was applied. (A) PCA of all subjects classified by stages by considering all parameters and outcomes (sex, age, diagnosed stage, and all bioenergetic and imaging outcomes). The X-axis and Y-axis show principal component 1 and principal component 2, which explain 17% and 11.1% of the total variance, respectively. The prediction ellipses demarcate the 0.99 probability that a new observation from the same group will fall inside the ellipse. (B) The sample analysis indicated under A was then performed only with the six (two imaging and four bioenergetic ones) outcomes selected by the VRE algorithm (from Fig. 2C). Using these parameters allows more precise separation of stages, as indicated by the larger variance contributions (principal component 1 and principal component 2 contributed to 31.6% and 24% of the total variance, respectively). Panels (A,B): Circles and square data points represent males and females, respectively; colors represented FXTAS stages: Stages 0–1 (red), blue (Stage 2), Stage 3 (green), and Stages 4–5 (purple). (C) A scree plot shows the individual contribution of the principal components (PC, blue bars) for explaining the variance in the data whose cumulative contribution (shown as a dotted orange line) accounts for 100% of the variance. PC1 and PC2 are the major contributors accounting for 57% of the variance. (D) A stacked column plot shows the individual contribution of the six outcomes selected by the VRE algorithm on principal components 1–4 accounts for >80% of the total variance. Imaging outcomes contribute mainly to PC1 (blue) and PC2 (orange); similarly, bioenergetic outcomes contribute to PC1–PC2 but also PC3 (grey) and PC4 (yellow). Bioenergetic outcomes were obtained with PBMCs except one indicated with “fibroblasts.” PC principal component.



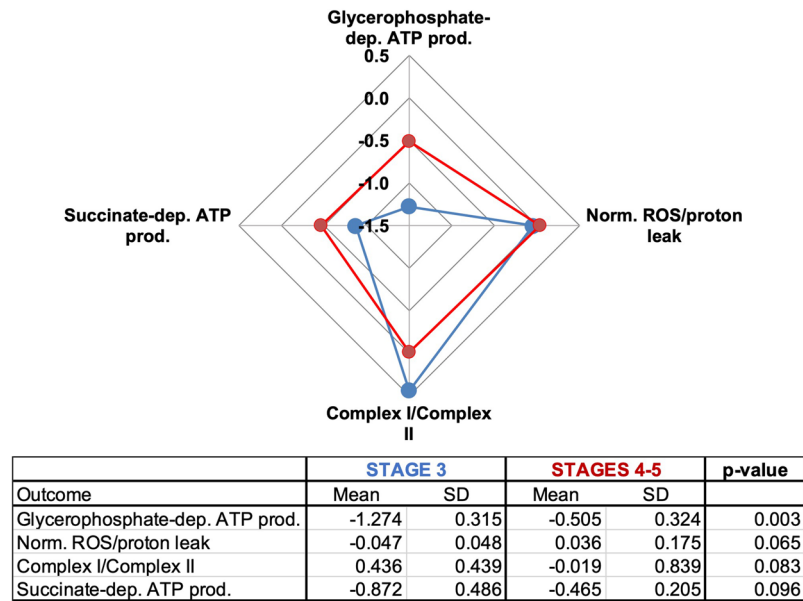
**Figure 4.** Bioenergetics of fibroblasts from patients at Stages 0–1 vs. 2–3. Top panel: Bioenergetic outcomes from primary fibroblasts obtained from carriers and assigned by the VRE algorithm as predicted Stages 0–1 (box A in Fig. 2C) or 2–3 (box B in Fig. 2C). The outcomes were plotted using a radar chart or spider web chart. For simplicity, only those with  $p \leq 0.1$  were shown—bottom panel: Statistical analyses performed using an unpaired two-tailed Student's *t* test with unequal variances. Values were expressed as LOG FC.

contributing to transferring reducing equivalents to the mitochondrial electron transport chain), the oxygen-linked ATP production rates sustained by glycerol-3-phosphate or basal (glucose–glutamine) were normalized to the maximum respiratory rate (Fig. 5B). While the activity of the malate–aspartate shuttle was similar between Stages 3 and 4–5 when PBMCs utilize glucose–glutamine, that of the glycerol-3-phosphate one was higher in those at higher stages. Taken together, at higher stages, there is higher lipolysis (with an increased flux of glycerol) at the expense of higher mtROS as mGPDH participates significantly in the mitochondrial ROS production<sup>87–89</sup>.

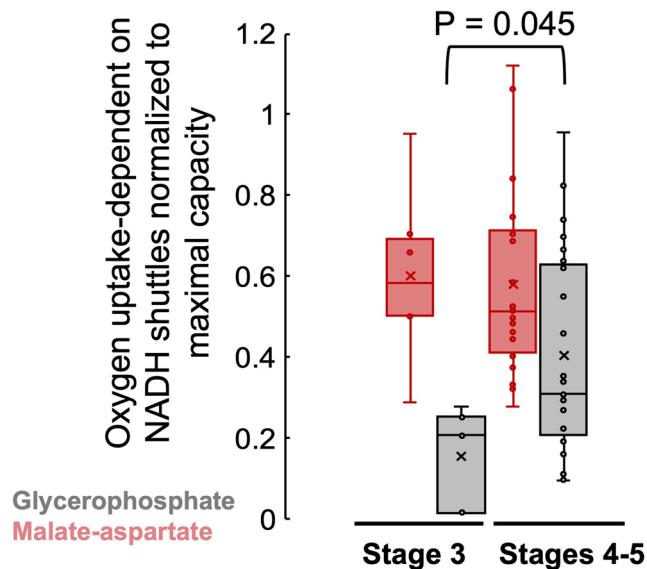
To test whether the bioenergetics of carriers with higher glycerol-3-phosphate-fueled ATP production would benefit from lowering the activity of the glycerol-3-phosphate shuttle, we treated fibroblasts from four carriers (3 males, one female, all within 67–68 years of age) selected from the decision tree classification (Fig. 2C labeled D as they had relatively higher glycerol-3-phosphate activity) with either metformin (50  $\mu$ M) or amino-oxyacetate. The rationale of this strategy was that bioenergetic improvements would be obtained upon inhibiting directly (metformin-dependent inhibition of mGDPH<sup>90</sup>) or indirectly (inhibitor of the malate–aspartate shuttle<sup>91,92</sup> that disrupts the NADH/NAD<sup>+</sup> ratio) the glycerophosphate shuttle (Fig. 6A). In this experimental setting, we used biologically relevant metformin concentrations within the 10–40  $\mu$ M plasma levels from patients treated with 1 g metformin twice a day<sup>93</sup> and within the 74–100  $\mu$ M detected in plasma and livers from rats treated acutely with 50 mg/kg body weight<sup>90</sup>. While it is known that metformin may exhibit a plethora of effects<sup>94,95</sup>, at these relatively low concentrations, and similar to other studies, no changes in Complex I activity were observed, and the incubation period did not allow for following putative changes in mitochondrial biogenesis (assessed by citrate synthase activity). As expected, both treatments increased levels of glycerol-3-phosphate compared to vehicle-treated ones (Fig. 6B), while only amino-oxyacetate disrupted metabolites within the malate–aspartate shuttle



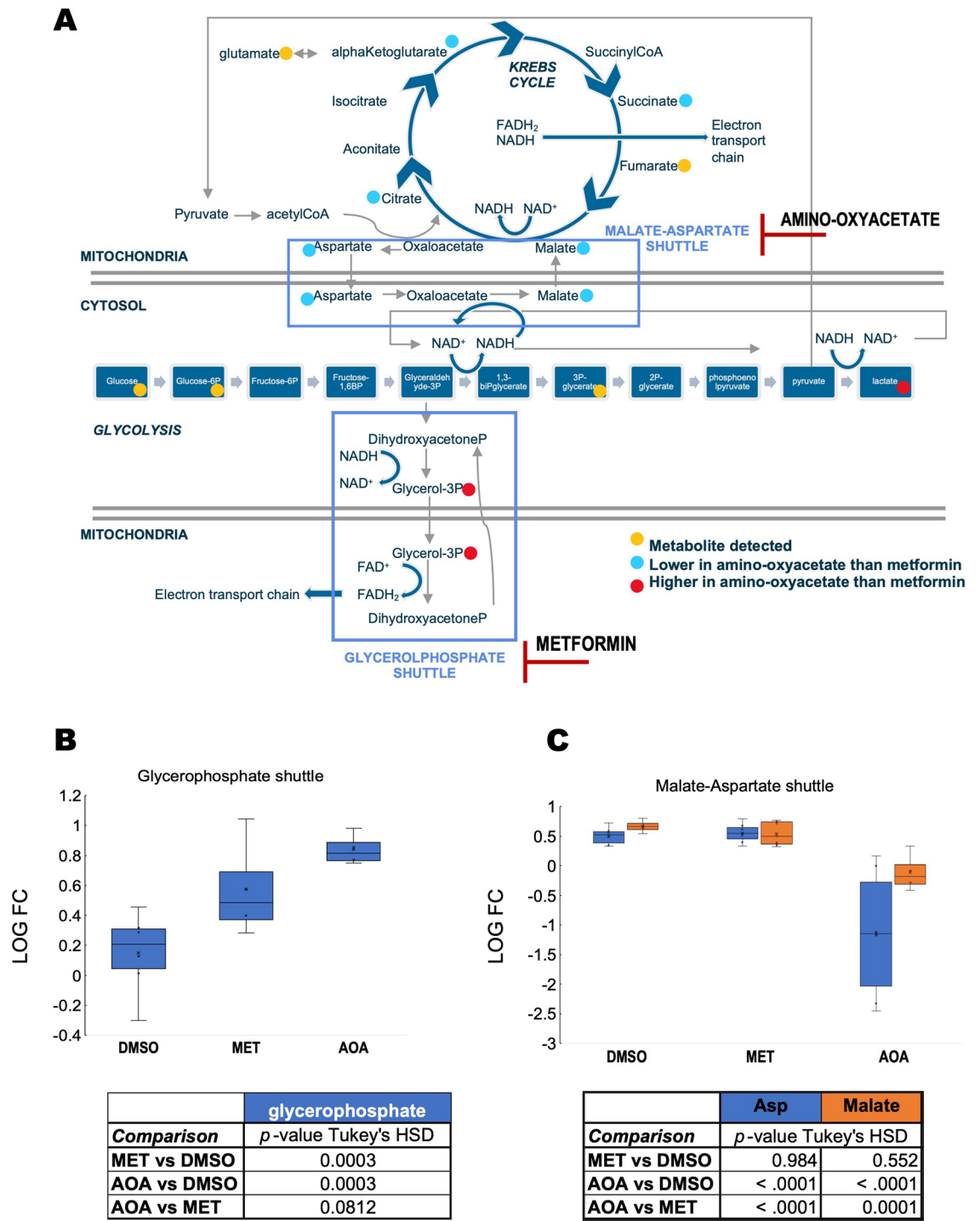
## A. Mitochondrial bioenergetics



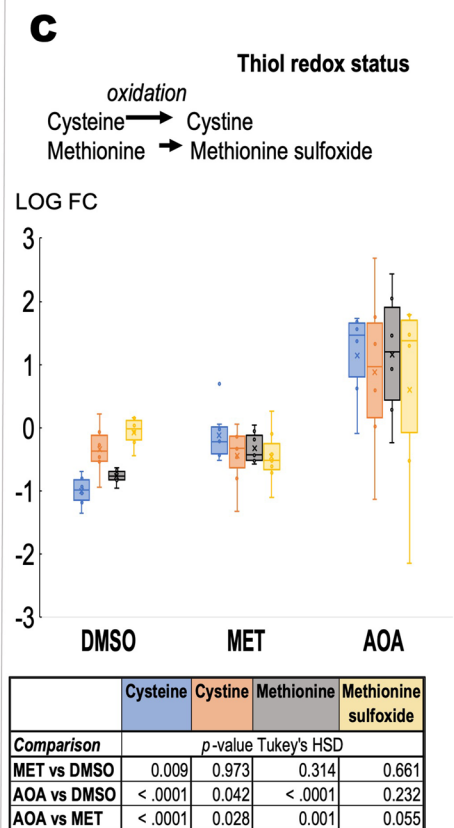
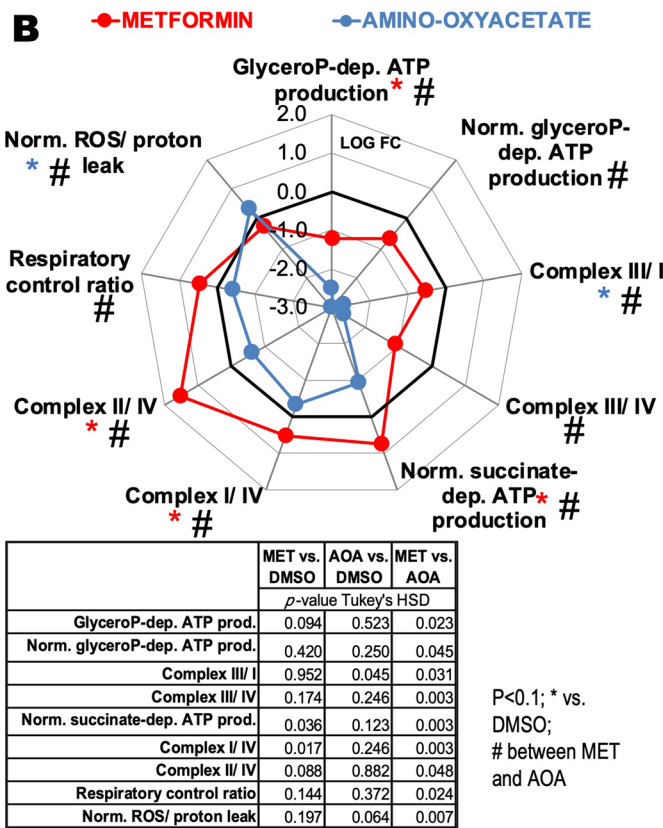
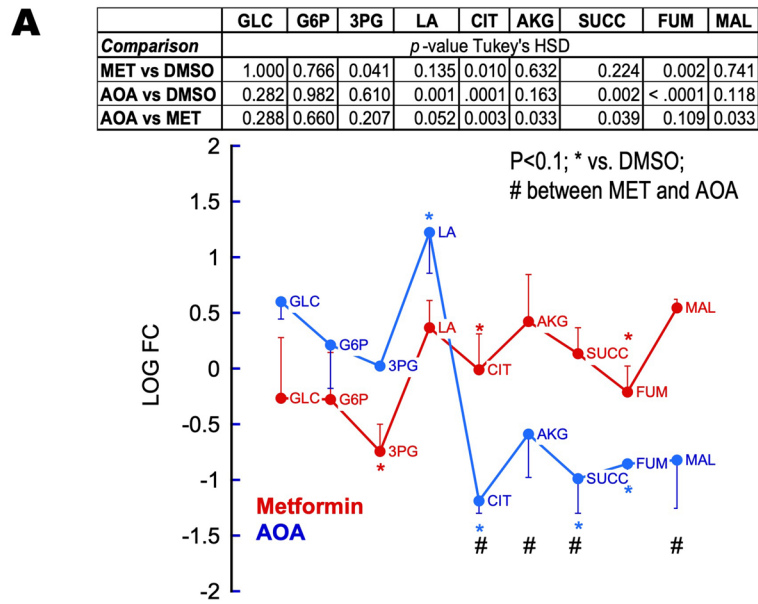
## B. Contribution of NADH shuttles



**Figure 5.** Bioenergetics of PBMCs from patients at Stages 3 vs. 4–5. **(A)** Mitochondrial bioenergetic outcomes from PBMCs obtained from carriers and assigned by the VRE algorithm at Stage 3 (Fig. 2C, box C) and Stages 4–5 (Fig. 2C, box D) were plotted using a spiderweb chart. For simplicity, only those with  $p \leq 0.1$  were shown. Bottom panel: Statistical analyses performed by using an unpaired two-tailed Student's  $t$  test with unequal variances. *dep.* Dependent, *prod.* Production, *Norm.* normalized to glucose–glutamine respiration. **(B)** Contribution of NADH shuttles: The oxygen-linked ATP production fueled by glucose–glutamine of primary fibroblasts from carriers (Malate-aspartate in red) was taken as a surrogate for the contribution of the malate-aspartate shuttle. The oxygen-linked ATP production fueled by glycerol-3-phosphate (GP in grey) was taken as the contribution of the glycerophosphate shuttle. Both rates were normalized to the maximum respiratory rate<sup>120</sup>. A two-tailed Student  $t$  test for unpaired data with unequal variance was run for GP with no statistical difference [ $t(29) = 0.203$ ;  $p = 0.840$ ] but with statistical significance for Malate-Aspartate [ $t(26) = -2.102$ ;  $p = 0.045$ ]. Data were visualized by using a box plot. Each box enclosed 50% of the data (closed or open circles) with the median value of the variable displayed as a line and the mean as an x. The top and bottom of the box mark the limits of  $\pm 25\%$  of the variable population. The lines extending from the top and bottom of each box mark the minimum and maximum values within the dataset that fall within an acceptable range. Any value outside this range, called an outlier, is displayed as an individual point whose value is outside the upper quartile +  $1.5 \times$  interquartile range and lower quartile –  $1.5 \times$  interquartile range. The height of each box represents the interquartile range.



**Figure 6.** Metformin-dependent inhibition of glycerophosphate shuttle and amino-oxacetate inhibits the malate-aspartate shuttle in fibroblasts from FXTAS carriers with high mitochondrial glycerol-3-phosphate dehydrogenase activity. Four primary fibroblasts (50  $\mu$ M metformin or MET, 5 mM amino-oxacetate or AOA) were treated for 24 h. Their metabolome and mitochondrial bioenergetic capacity were compared to DMSO-treated as indicated in the “Methods”. Panel (A): Inhibition sites for metformin and amino-oxacetate. The inhibition exerted by these compounds is shown in the context of glycolysis, Krebs’ cycle, and malate-aspartate and glycerophosphate shuttles. Panel (B): Levels of glycerol-3-phosphate were evaluated by metabolomics under each condition (vehicle or dimethylsulfoxide, DMSO; MET metformin; AOA amino-oxacetate) as a proxy for glycerophosphate shuttle. Panel (C): Malate-aspartate shuttle components (Aspartate or Asp in blue; malate in orange). For panels (B,C): Data were visualized by using a box plot (box height is the interquartile range; top and bottom of boxes are the upper and lower quartiles; top and bottom lines represent the maximum and minimum values; median is shown as a line; average as an x). Tables at the bottom of each panel show the *p*-values obtained from Tukey’s honestly significant difference (HSD) for each comparison and show only outcomes with *p*-values  $\leq 0.1$ . Matched-samples ANOVA was run to test for statistical significance across treatments followed; if  $p \leq 0.05$ , a Tukey’s HSD post-test was run. All analyses were performed with SPSS.



◀ **Figure 7.** Metformin-dependent inhibition of glycerophosphate shuttle improves bioenergetic outcomes and decreases oxidative stress of fibroblasts from FXTAS carriers with high mitochondrial glycerol-3-phosphate dehydrogenase activity. Panel (A): Glucose catabolism, including glycolysis and Krebs' cycle, was evaluated by assessing key metabolites by using metabolomics. The metabolites were obtained from four primary fibroblasts (50  $\mu$ M metformin or MET, 5 mM amino-oxyacetate or AOA) treated for 24 h. For simplicity, only one-side error bars are shown, representing the standard error of the mean. P-values obtained using Tukey's HSD are shown with an asterisk (if the comparison was against DMSO) or # (if the comparison was between MET and AOA). *GLC* glucose, *G6P* glucose-6-phosphate, *3PG* 3-phosphoglycerate, *LA* lactic acid, *CIT* citrate, *AKG* alpha-ketoglutarate, *SUCC* succinate, *FUM* fumarate, *MAL* malate. Panel (B): Bioenergetic outcomes-Spider plot showing bioenergetic outcomes of fibroblasts from four carriers treated with metformin or amino-oxyacetate as indicated above. Metformin data are shown in red; amino-oxyacetate in blue. *glyceroP* glycerophosphate, *dep.* dependent, *prod.* Production, *Norm.* normalized by the activity of citrate synthase except ROS/proton leak normalized to oxygen-linked ATP production. P-values obtained using Tukey's HSD for each comparison; shown only outcomes with  $p$ -values  $\leq 0.1$ . Matched-samples ANOVA was run to test for statistical significance across treatments followed; if  $p \leq 0.05$ , a Tukey's HSD post-test was run. All analyses were performed with SPSS.

(Fig. 6C). Both inhibitors showed similar patterns of glucose metabolism, i.e., higher lactic acid and decreases in Krebs' cycle intermediates, with more profound and significant changes observed for amino-oxyacetate than metformin (Fig. 7A). The lower levels of Krebs' cycle intermediates may result from disrupted anaplerotic reactions (i.e., reactions that serve to replenish Krebs' cycle intermediates) affected by the imbalances in cytosolic NADH/NAD<sup>+</sup> status. These results indicated that by inhibiting either shuttle (but quantitatively more the malate-aspartate one), the reduced equivalents produced during glycolysis are kept in the cytosol in the form of lactic acid via the activity of lactate dehydrogenase.

As predicted, the mitochondrial bioenergetic analysis indicated that both treatments (albeit metformin at higher degrees) decreased the ATP production sustained by glycerophosphate (normalized or not to citrate synthase); as such, those Complex ratios in which Complex III was either a denominator or numerator changed accordingly. Changes specific to metformin were the increase in the succinate-dependent ATP production, Complex I/IV and II/IV ratios, respiratory control ratio, and a decrease in ROS/proton leak production. Taken together, metformin vs. amino-oxyacetate treatment was beneficial in those fibroblasts with relatively high glycerol-3-phosphate-fueled ATP production because it decreased the rate of ROS, increased the rate of ATP production fueled by succinate, and increased the coupling between electron transfer and ATP production. Conversely, amino-oxyacetate vs. vehicle treatment increased the rate of ROS production normalized to basal respiration and dramatically decreased the Complex III/Complex I ratio (Fig. 7B). As such, differences between metformin and amino-oxyacetate treatments were essentially based on proton leak and mtROS production and the segment of the electron transport chain between the entry points of succinate and glycerophosphate.

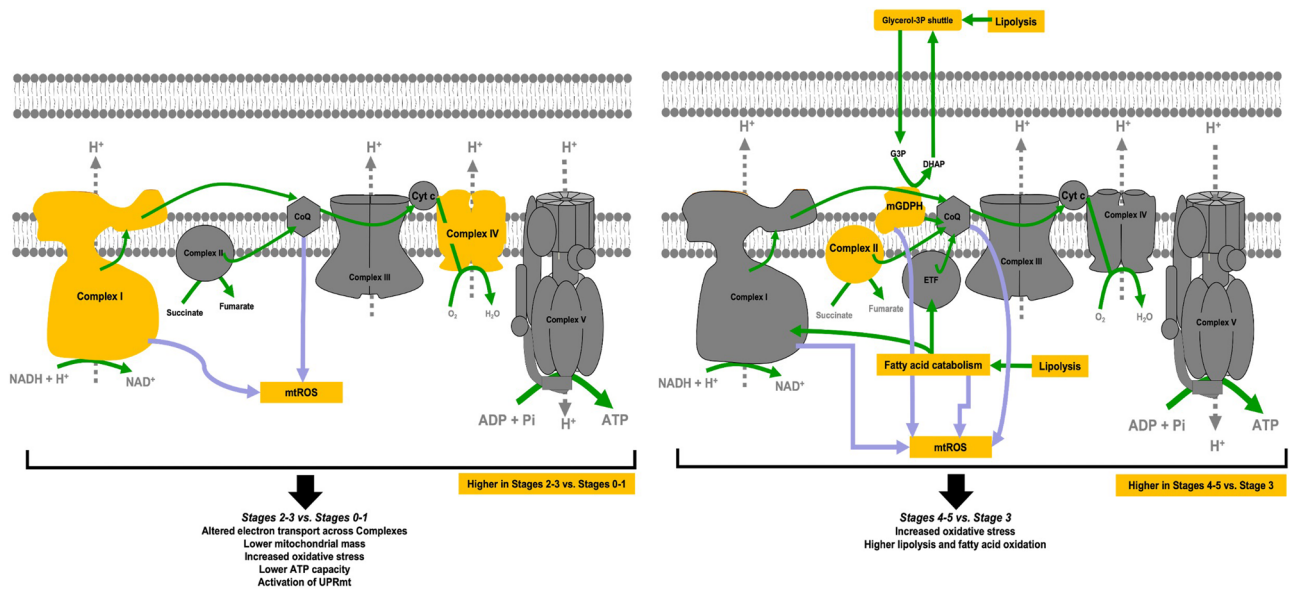
To test the impact of oxidative stress by inhibiting either the malate-aspartate or the glycerol-3-phosphate shuttle, we evaluated the levels of cysteine, methionine, and their oxidized products (cystine and methionine sulfoxide, respectively) as sulfur-containing amino acids are strongly affected by oxidative stress (Fig. 7C). Metformin significantly increased cysteine levels with no changes in methionine or their respective oxidized products, suggesting that by decreasing mGPDH-mediated ROS production, reduced cysteine levels were spared. These results are consistent with the lower mtROS production indicated above achieved by metformin treatment. As expected for a condition resembling nonphosphorylating mitochondria due to the shortage of NADH supply via the malate-aspartate shuttle, amino-oxyacetate treatment increased the mtROS production compared to vehicle-treated cells. However, despite the increase in oxidative stress, the levels of cysteine and methionine were higher than those with the vehicle (Fig. 7C). This can be understood by the aminoxyacetate-dependent inhibition of cystathionine gamma-lyase<sup>96</sup>, a key enzyme in the catabolism of these amino acids.

## Discussion

In this study, we tested whether ANN approaches could aid in diagnosing FXTAS stages based on age, sex, MRI findings, and bioenergetic outcomes from two cell types, PBMC and fibroblasts. Our results indicated that the 3 ANN approaches utilized here were able to make a suitable stage classification with an accuracy of 86.27% (SFAM), 76% (backpropagation), and 41.67% (VRE) based only on two MRI and four bioenergetic findings—both experimental, unbiased outcomes—of the many outcomes included in the study, all of which were objective measures of brain imaging and cellular bioenergetic assessments.

Although the progression of FXTAS increases in older male carriers, age was not among the critical factors expected for predicting the progression of FXTAS because, in our study, we stratified the sampling for the training and test sets, in which age and sex were proportionally represented. Notably, although sex was also not among the main discriminators for predicting FXTAS stages as expected for an X-linked disorder, on average, the ratio of male-to-female mismatch ratio was 1.45 across the three methods tested here, supporting the notion that women may be protected by the second X chromosome<sup>97</sup> and estrogen<sup>98,99</sup>.

The decision tree classification obtained by using VRE algorithm utilized two of the five brain imaging outcomes, i.e., WMHs and whole brain volumes normalized by intracranial volume, for predicting FXTAS stages.



**Figure 8.** Bioenergetic differences across stages. Changes in bioenergetics (in yellow) that match the assignment of stages according to the decision tree classification (Fig. 2C). Altered Complex ratio and increased mitochondrial ROS characterized Stages 2–3 vs. Stages 0–1 (left panel), accompanied by lower mitochondrial mass (citrate synthase activity; see text). These changes result in an overall lower ATP production (altered Complex ratio and lower mitochondrial mass), and may activate the mitochondrial unfolded protein response (UPRmt). Higher fatty acid oxidation (Complex II, mGDPH) and mitochondrial ROS characterized Stages 4–5 vs. Stage 3. While these changes may increase ATP derived from mitochondrial fatty acid beta-oxidation, they may increase the susceptibility of neurons to hypoxic events and to ATP deficits as glucose catabolism is faster than that from fatty acids. Arrows in green indicate the electron flow and direction of biological processes; purple arrows indicate the ROS sources. CoQ coenzyme Q, Cyt c cytochrome c, Pi phosphate, mGDPH mitochondrial glycerophosphate dehydrogenase, ETF electron transfer flavoprotein, G3P glycerol-3-phosphate, DHAP dihydroxyacetone-phosphate.

As indicated before, white matter lesions presented as WMHs on T2 and FLAIR images in the middle cerebellar peduncles, brainstem, deep white matter, and corpus callosum are important radiologic biomarkers for FXTAS diagnosis. Neuropathological studies revealed that white matter disease is characterized by marked spongiosis, loss of axons, and demyelination in the cerebrum and cerebellum in most brains of carriers with FXTAS<sup>11,12</sup>. However, WMHs are not unique to FXTAS as they have also been observed in aging and other neurodegenerative diseases<sup>100–103</sup>. Conversely, bilateral WMHs in the middle cerebellar peduncles are much rarer and present in FXTAS-affected patients but also in some patients with Leigh syndrome, brain stroke, demyelination, multiple system atrophy-cerebellar variants, spinocerebellar ataxia, and leukoencephalopathy<sup>51,104–106</sup>.

In the decision tree, the values of WMHs normalized by intracranial volume helped classify FXTAS Stages 0–3 ( $\leq 12,600 \text{ mm}^3$ ) versus 3–5 ( $> 12,600 \text{ mm}^3$ ). In contrast, the classification of stages 3 versus 1, 2, and 4 relied on both imaging (normalized whole brain volume) and bioenergetic outcomes (Fig. 2C). This result is consistent with our recent publication reporting a significant correlation between WMH volume and FXTAS stage in 74 premutation carriers (54% males; aged  $\geq 50$  years) after adjusting for age, sex, and intracranial volume, although substantial overlapping in WMHs volume across all FXTAS stages can be observed<sup>15</sup>. Values of WMHs normalized by intracranial volume  $< 600 \text{ mm}^3$  help classify carriers at Stages 0–1, whereas those with values between  $600 \text{ mm}^3$  and  $12,600 \text{ mm}^3$  and normalized whole brain volume  $\leq 1.3 \text{ dm}^3$  predicted a Stage 2 diagnosis. Bioenergetic outcomes were necessary for distinguishing among Stages 0–1, 2, and 3 for those with normalized whole brain volume  $> 1.3 \text{ dm}^3$ . This is also consistent with our previous finding of age-related accelerated brain volume loss in carriers without FXTAS than noncarriers, which may become more evident as FXTAS develops<sup>16</sup>.

Intriguingly, the two binary variables, the presence of WMHs in the middle cerebellar peduncles and the cerebellum-brainstem, were not selected by ANN for FXTAS stage classification. In our recent study of 175 carriers and 82 controls aged  $\geq 45$  years<sup>54</sup>, the WMHs presence in the middle cerebellar peduncles was detected only in the premutation carriers (52%) with significant correlations with impaired motor and executive function in those carriers with  $> 75$  CGG repeats. In the current study of 111 carriers, the prevalence of WMHs in the middle cerebellar peduncles and WMHs in the cerebellum and brainstem increased with FXTAS stage (middle cerebellar peduncles/cerebellum and brainstem: Stage 0–1, 6.1%/57.6%; Stage 2, 25%/80%; Stage 3, 42.1%/94.7%; Stages 4–5, 80%/100%). It is conceivable that WMH volume, providing a fine-grained measurement for white matter damage, is more useful for stage classification than the information provided by two binary variables. Similarly, ventricular volume was not selected by ANN as it may carry overlapping information with the WMH volume, further evidenced by their significant positive correlation [Pearson's  $r(109) = 0.598$ ,  $p < 0.0001$ ].

Regarding bioenergetic outcomes selected by ANN to discriminate between early and late stages, it is remarkable that one of the outcomes (Complex IV) was affected in both cell types (bioenergetic outcomes summarized



in Fig. 8). The consistency in the outcome Complex IV activity between fibroblasts and PBMCs indicates a conserved mechanism despite the apparent differences between the cell types, namely, proliferative vs. quiescent, grown under optimal cell culture conditions vs. reflecting the carrier's homeostasis within days to several weeks before blood extraction, germ layer of origin ectoderm vs. endoderm. The results indicate that cells from carriers at Stages 0–3 may have a mitochondrial-nuclear imbalance evidenced by changes in Complex IV activity because this Complex is encoded by the coordinated action of both the nuclear and the mitochondrial genomes. As indicated before, an altered ratio of mitochondrial Complexes ensues in not only deficits in ATP production but also in the activation of UPRmt<sup>85</sup>. This is supported by the activation of UPRmt in worms and mammals in response to the stoichiometric imbalance between nuclear DNA- and mitochondrial DNA-encoded oxidative phosphorylation proteins, also known as mitochondrial-nuclear protein imbalance<sup>85</sup>. This scenario extends our previous reports on a disrupted proteostasis in the *FMRI* premutation<sup>24,39,107</sup> with an accumulation of oxidatively modified proteins and deficits in mitochondrial protein import, processing, and folding<sup>21–23</sup>. Notably, the apparent functional connection between mitochondrial-nuclear protein imbalance and UPRmt, and their link to aging<sup>85</sup> may prompt future research to explore whether targeting UPRmt can prevent or slow down age-dependent neurodegeneration.

While increases in mtROS (oxidative stress) seem to underlie all transitions from early to late stages, the increases in glycerol-3-phosphate shuttle in Stages 4–5 may represent an attempt to overcome energy failure by increasing glycerol and fatty acid oxidation at the expense of increased oxidative stress (Fig. 8). This is a crucial point as it indicates that controlling mGPDH and oxidative stress (especially cumulative damage over time) may be critical to preventing the transition of FXTAS carriers to higher stages. As indicated above, while the upregulation of glycerophosphate-mediated ATP production in cells from carriers at higher FXTAS stages may be interpreted as an attempt to supply ATP via mitochondrial fatty acid beta-oxidation, several factors need to be considered to infer that this attempt might be detrimental to carriers at higher stages. Among them were (i) the inherent higher mtROS production of both mGPDH activity and fatty acid beta-oxidation, (ii) the higher glycolytic capacity of cells with high mGPDH expression<sup>108</sup>, which may limit the indispensable mitochondria-derived ATP production in neurons, and (iii) the slower ATP production and higher oxygen demand to oxidize fatty acids vs. glucose which may not be matched by the high demand of rapid and sustained neuronal firing, putting neurons at risk of hypoxic events<sup>109</sup>. In support of these concepts, in this study, metformin-dependent inhibition of mGPDH decreased mtROS production, improved the coupling and increased the ATP production sustained by succinate (Fig. 7B). Treatment of fibroblasts from carriers at Stages 3–4 with the potent NRF2 activator sulforaphane successfully recovered all bioenergetic outcomes highlighting the need for enhancing UPR and antioxidant responses to protect mitochondria from further damage<sup>39</sup>.

Limitations of this study include the lack of testing for the specificity of these ANN analyses in other diseases displaying a psychiatric and cognitive performance pattern overlapping with FXTAS (e.g., several spinocerebellar ataxias and frontotemporal dementia). Due to sample availability and cell number, some bioenergetic data were missing from a few samples and were estimated with suitable algorithms. All participants were recruited from one facility (MIND Institute), leading to potential sample bias. Finally, most male carriers included in this study were at Stage 3, which may limit the applicability to other stages.

Despite the aforementioned limitations, this study delineates a strong molecular signature in detecting FXTAS across stages, offering the best therapeutic intervention window. If validated in prospective larger cohorts, the predicted trajectories through the implementation of a neural network will represent a breakthrough in clinical screening and swift detection of FXTAS morbidity in carriers of the *FMRI* premutation without overt FXTAS symptoms or to prevent the transition to stages with more detrimental symptoms and co-morbidities.

## Methods

**Study design and participants.** Written informed consent was obtained from all participants before participation in line with the Declaration of Helsinki. The study was reviewed and approved by the Institutional Review Board of the University of California Davis Medical Center, and all methods were performed by their guidelines and regulations.

The present study consisted of a retrospective chart review of adult patients with FXTAS. Participants were carriers with the premutation enrolled in a large research study at the UC Davis Medical Investigation of Neurodevelopmental Disorders (MIND) Institute. We included individuals who participated in the Genotype-Phenotype Study of Families with Fragile X (NICHD HD036071) from 2013 through 2019 and with the availabilities of MRI, primary skin fibroblasts, and blood samples. The inclusion criteria required that participants were noncarriers (n = 9) or carriers of the premutation (n = 89), all adults (≥ 18 years), and both sexes included. Among the carriers, they would have a definite, probable, or possible FXTAS diagnosis in clinical Stages 1–5. We considered FXTAS Stage 0 for a carrier with no FXTAS diagnosis. FXTAS Stage<sup>7</sup> was scored by a trained physician (R.J.H.) based on the severity of movement and gait impairments (1: subtle or questionable signs, 2–6: evident tremor/balance problems with minor to severe interference of daily living). Only those with the most outcomes were considered if multiple visits were available. Only the oldest was included if two or more visits had the most outcomes. As such, no longitudinal data were taken into consideration. Exclusion criteria were other neurodegenerative dementias, a history of cerebrovascular accidents, traumatic brain injury, and alcohol abuse or dependence. A total of 127 participants (both genders) were included in the final data analysis, and their demographics are summarized in Fig. 1.

**Chemicals and biochemicals.** EDTA, EGTA, KH<sub>2</sub>PO<sub>4</sub>, sodium succinate, digitonin, rotenone, antimycin A, oligomycin, malonate, ascorbic acid, N,N,N',N'-tetramethyl-p-phenylenediamine, KCN, HEPES, metformin and AOA were purchased from Sigma (St Louis, MO, USA). Tris-HCl, glycine, sodium chloride, and KCl were

purchased from Fisher (Pittsburg, PA, USA). Bovine serum albumin (fatty acid-free) was obtained from MP Biomedicals. All reagents were of analytical grade or higher.

**PBMCs preparation.** PBMCs were isolated from blood samples (5–7 ml) collected in BD Vacutainer™ glass mononuclear cell preparation tubes (BD Biosciences, Franklin Lakes, NJ, USA) as previously described<sup>25</sup>. Lymphocyte suspensions were pelleted by centrifugation for 1 min at 2000 rpm in a microfuge at 4 °C. The supernatants were discarded, and cell pellets were used immediately to evaluate mitochondrial outcomes.

**Culture of primary skin fibroblasts.** Skin biopsies were obtained from participants from the same skin area (punch biopsy on left upper back) to minimize variances due to body site or different exposure to the external milieu. To minimize confounding factors, such as significant telomere shortening across passages and corresponding features of replicative senescence, we used fibroblasts at their earliest passage available (on average 4–5). This approach also prevented the putative instability of cultured cells from subjects with respiratory chain disorders, which tend to return to normal values across passages because cells less affected may proliferate faster than those more severely affected<sup>110</sup>. Fibroblasts were obtained from Dr. P. Hagerman's laboratory with the help of Ms. Glenda Jackson. Fibroblasts were grown in high glucose Minimum Essential Medium supplemented with 15% fetal bovine serum, 2 mM glutamine, and 1 mM sodium pyruvate, as previously described<sup>22</sup>.

Metformin and amino-oxyacetate treatments were performed on primary fibroblasts from four carriers with FXTAS, three males (2 at Stage 3 and one at Stage 4) and one female (Stage 3) with a median (range) age of 67 (67–68). This subset was selected randomly from patients within box D as they presented relatively higher activity of the glycerol-3-phosphate shuttle (Fig. 2C). For the treatment with metformin and AOA, cells were grown at ~50% confluence and treated with 50 μM metformin, 5 mM AOA, or with the equivalent volume of DMSO for 24 h, detached by trypsinization (ThermoFisher Scientific), and assessed for viability by trypan blue with a TC20 cell counter (BioRad). Bioenergetic outcomes were determined as it is detailed below.

**Mitochondrial outcomes of PBMCs and fibroblasts.** Mitochondria-dependent oxygen consumption was evaluated in either fibroblasts<sup>27</sup> or freshly isolated lymphocytes permeabilized with digitonin as previously described<sup>36</sup>. Briefly, lymphocytes ( $0.5\text{--}1.0 \times 10^6$ ) or fibroblasts ( $2.0 \times 10^6$ ) were added to the Clark-type oxygen electrode chamber containing 0.3 ml of a buffer (0.22 M sucrose, 50 mM KCl, 1 mM EDTA, 10 mM  $\text{KH}_2\text{PO}_4$ , and 10 mM HEPES, pH 7.4). After establishing the baseline, the oxygen consumption rates were evaluated with (i) 1 mM ADP, 1 mM malate-10 mM glutamate for 5–8 min until reaching a steady rate, followed by the addition of 5 μM rotenone to inhibit Complex I; (ii) 10 mM succinate for 5–8 min until reaching a steady rate, followed by adding 1 mM malonate and 3.6 μM antimycin A; and (iii) 10 mM ascorbate and 0.2 mM  $N,N,N',N'$ -tetramethyl-*p*-phenylenediamine followed by the addition of 1 mM KCN to inhibit Complex IV. The succinate-fueled oxygen uptake and cytochrome *c* oxidase activity were evaluated as the difference in oxygen consumption rate recorded before and after the addition of the respective inhibitors, malonate, and KCN. The respiratory control ratio (RCRu) was calculated as the quotient of oxygen consumption rates of intact cells in 10 mM glucose (present in RPMI-1640) in State 3u (with 2 μM carbonylcyanide-*p*-trifluoromethoxyphenylhydrazone; FCCP) and that of State 4 (with 0.2 μM oligomycin). SRC (spare respiratory capacity) was calculated as the ratio (in percentage) between State 3u and basal respiration.

In contrast, ROS/proton leak was calculated as the ratio (in percentage) between State 4 and basal respiration. Protein concentration was determined by the Lowry method and calculated according to the bovine serum albumin standard curve. Citrate synthase activity was evaluated at 412 nm in both cell types with a Tecan Infinite M200 microplate reader equipped with the Magellan software (Austria)<sup>36</sup>. All carrier parameters were normalized to the pooled noncarrier values. For some samples (<5), bioenergetic data from either fibroblasts or PBMC were missing (sample loss, not enough cells for reliable detection, among others). In those cases, missing data were estimated using the feature *k*-nearest-neighbor imputation<sup>111</sup>.

**MRI acquisition and processing.** All methods had been described in detail before<sup>15</sup>. MRI scans were acquired on a Siemens Trio 3 T MRI scanner equipped with a 32-channel head coil (Siemens Medical Solutions, Erlangen, Germany). One-millimeter isotropic T1-weighted scans were collected covering the whole brain using the magnetization-prepared rapid gradient-echo sequence in 192 sagittal slices with a repetition time of 2170 ms, echo time of 4.82 ms, and 7° flip angle. Fluid attenuated inversion recovery (FLAIR) images for quantifying WMHs were acquired in 104 sagittal slices of 1.9-mm thickness with an in-plane resolution of 0.47 mm<sup>2</sup>, repetition time of 5000 ms, echo time of 456 ms, and inversion time 1700 ms. Both T1 and FLAIR scans were corrected for intensity inhomogeneities due to the MRI bias field using N4<sup>112</sup>. Whole-brain volume, ventricular volume, and brain scaling factor (for correcting individual differences in cranial size) were obtained on magnetization-prepared rapid gradient-echo scans using the SIENAX program<sup>113</sup> from FSL. Optimal values were obtained by adjusting the parameter used by the BET program for brain extraction. Whole-brain WMH volume was quantified on FLAIR images using a lesion prediction algorithm from SPM12<sup>114</sup>. Lesion masks were generated by setting appropriate thresholds on lesion probability maps using the FSL command, `fslmaths`, followed by manual correction for errors using `ITK-Snap`<sup>115</sup>. Volume calculation of WMH was performed using the FSL command, `fslstats`. Normalized volumes of the whole brain, ventricles, and WMHs normalized by intracranial volume were computed by their multiplication with the brain scaling factor. Binary variables were coded according to the presence or absence of WMHs in the middle cerebellar peduncles and cerebellar white matter, respectively, based on the generated WMHs lesion mask. Cerebellar white matter included the middle cerebellar peduncles, transverse pontine fiber in the pons, superior cerebellar peduncle, and its decussation in the midbrain.

**Artificial neural network (ANN) for FXTAS stage diagnosis.** The ANN approach applied to this study followed that described before for Wilson's disease<sup>50</sup>. The ANN design consisted of a three-layer network: an input layer with either 2 or 5 units for defining FXTAS Stages (Fig. 1A), sex, age, five units of brain imaging (as indicated above), and 22 units of mitochondrial outcomes for each cell type (PBMC and fibroblasts); a hidden layer, with four units; and an output layer for the FXTAS stage. The simplified fuzzy adaptive resonance theory map (SFAM) algorithm was used for the prediction of "early" (noncarriers and carriers at Stages 0–2) and "late" (carriers at Stage  $\geq 3$ ), whereas the backpropagation algorithm was used to predict numerical values of FXTAS stages (Fig. 1A) (NeuNetPro software; CorMac Technologies Inc., Canada). The unsupervised training variables included all of those indicated above. To select the training set size for each analysis, we performed a learning curve analysis<sup>49</sup>. We withdrew a relatively small random sample from our data to train an ANN and used this training set to predict a randomly sampled test set. Then, the size of our training sample was increased while keeping the same test set to track the degree to which predictive accuracy on the fixed test set increased with the training set size. This iterative process resulted in a training data size that minimized the differences between the network classifications and the clinical diagnoses. We repeated the learning curve analysis multiple times with different test set samples to reduce variation in predictive accuracy (85 samples were used for training, and 42 were used for testing with SFAM and backprop). The last algorithm (Visual Rule Extraction or VRE)<sup>81,82</sup> utilized the patterns of input factors associated with stages. It was trained with a randomly selected sample of 75 carriers from the 111 premutation carriers, all with known clinical status. This set of 75 carriers was selected using stratified random sampling (Python package; SciKitLearn) to ensure that each subgroup within the population received proper representation within the validation sample. The stratified random sampling provides better population coverage since it allows control over the subgroups to ensure that all of them are represented in the sample. Once the network was trained, the remaining 36 carriers were "tested" using the trained network. The neural network classifications were then compared with the known clinical diagnoses to determine whether the network could verify disease status reliably. The output was visualized with a decision tree with adjustable tree pruning (55% pruning with a minimum of five participants).

**Statistical analysis.** The authors were blinded to the diagnosis of the subjects until all data were collected. If data for a subject were available from multiple visits, the oldest or the one with the most outcomes was taken. As such, no longitudinal data were taken into consideration for this study. Bioenergetic outcomes in fibroblasts and PBMCs were evaluated in at least triplicates. The statistical analyses involved comparisons between boxes or before and after treatments. The treatments (DMSO, metformin, and AOA) were compared using matched-samples 1-way ANOVA with Tukey's multiple comparisons post-test. Differences between boxes were analyzed via an unpaired, two-tailed Student's *t* test with unequal variances. The variance equality was confirmed by using Levene's test<sup>116–118</sup>. For every comparison, a power analysis was run in which the alpha and power of the experiment were imputed, given the beta/alpha ratio, sample size, and effect size. For a two-tailed *t* test (difference between two independent means), an effect size of 2, beta/alpha ratio = 1, the computed alpha values ranged from 0.087 to 0.176 with a power of 0.823 to 0.913, highlighting that the *p*-values of 0.1 are within the alpha range. Fisher exact test and Pearson correlation were performed with SPSS. For boxes' analyses, points whose values were either greater than upper quartile  $2 \times$  interquartile distance or less than lower quartile  $-2 \times$  interquartile distance were considered outliers and excluded from further analysis. The data statistical software SPSS 27.0.0 and the data analysis software GraphPad Prism 9 were used to analyze the experimental data.

**Institutional review board statement.** The University of California Davis Institutional Review Board and the State of California Committee for the Protection of Human Subjects approved this study.

**Informed consent.** Neither data nor specimens were collected until written informed consent was obtained from the patients. Written informed consent has been obtained from the patient(s) to publish this paper.

### Data availability

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

Received: 4 May 2022; Accepted: 1 December 2022

Published online: 10 December 2022

### References

- Jacquemont, S. *et al.* Aging in individuals with the *FMR1* mutation. *Am. J. Ment. Retard.* **109**, 154–164. [https://doi.org/10.1352/0895-8017\(2004\)109%3c154:AIWTF%3e2.0.CO;2](https://doi.org/10.1352/0895-8017(2004)109%3c154:AIWTF%3e2.0.CO;2) (2004).
- Jacquemont, S. *et al.* Penetrance of the fragile X-associated tremor/ataxia syndrome in a premutation carrier population. *JAMA* **291**, 460–469. <https://doi.org/10.1001/jama.291.4.460> (2004).
- Leehey, M. A. *et al.* Progression of tremor and ataxia in male carriers of the *FMR1* premutation. *Mov. Disord.* **22**, 203–206. <https://doi.org/10.1002/mds.21252> (2007).
- Jacquemont, S., Hagerman, R. J., Hagerman, P. J. & Leehey, M. A. Fragile-X syndrome and fragile X-associated tremor/ataxia syndrome: Two faces of *FMR1*. *Lancet Neurol.* **6**, 45–55. [https://doi.org/10.1016/S1474-4422\(06\)70676-7](https://doi.org/10.1016/S1474-4422(06)70676-7) (2007).
- Grigsby, J. *et al.* Impairment in the cognitive functioning of men with fragile X-associated tremor/ataxia syndrome (FXTAS). *J. Neurol. Sci.* **248**, 227–233. <https://doi.org/10.1016/j.jns.2006.05.016> (2006).
- Grigsby, J. *et al.* Cognitive impairment in a 65-year-old male with the fragile X-associated tremor-ataxia syndrome (FXTAS). *Cogn. Behav. Neurol.* **19**, 165–171. <https://doi.org/10.1097/01.wnn.0000213906.57148.01> (2006).
- Bacalman, S. *et al.* Psychiatric phenotype of the fragile X-associated tremor/ataxia syndrome (FXTAS) in males: Newly described fronto-subcortical dementia. *J. Clin. Psychiatry* **67**, 87–94. <https://doi.org/10.4088/jcp.v67n0112> (2006).

8. Bourgeois, J. A. *et al.* Cognitive, anxiety and mood disorders in the fragile X-associated tremor/ataxia syndrome. *Gen. Hosp. Psychiatry* **29**, 349–356. <https://doi.org/10.1016/j.genhosppsy.2007.03.003> (2007).
9. Bourgeois, J. A. *et al.* Dementia with mood symptoms in a fragile X premutation carrier with the fragile X-associated tremor/ataxia syndrome: Clinical intervention with donepezil and venlafaxine. *J. Neuropsychiatry Clin. Neurosci.* **18**, 171–177. <https://doi.org/10.1176/jnp.2006.18.2.171> (2006).
10. Grigsby, J. *et al.* The cognitive neuropsychological phenotype of carriers of the *FMR1* premutation. *J. Neurodev. Disord.* **6**, 28. <https://doi.org/10.1186/1866-1955-6-28> (2014).
11. Greco, C. M. *et al.* Neuropathology of fragile X-associated tremor/ataxia syndrome (FXTAS). *Brain* **129**, 243–255. <https://doi.org/10.1093/brain/awh683> (2006).
12. Greco, C. M. *et al.* Neuronal intranuclear inclusions in a new cerebellar tremor/ataxia syndrome among fragile X carriers. *Brain* **125**, 1760–1771. <https://doi.org/10.1093/brain/awf184> (2002).
13. Hall, D. A. *et al.* Emerging topics in FXTAS. *J. Neurodev. Disord.* **6**, 31. <https://doi.org/10.1186/1866-1955-6-31> (2014).
14. Hagerman, R. J. & Hagerman, P. Fragile X-associated tremor/ataxia syndrome—Features, mechanisms and management. *Nat. Rev. Neurol.* **12**, 403–412. <https://doi.org/10.1038/nrneurol.2016.82> (2016).
15. Wang, J. *et al.* Brain atrophy and white matter damage linked to peripheral bioenergetic deficits in the neurodegenerative disease FXTAS. *Int. J. Mol. Sci.* <https://doi.org/10.3390/ijms22179171> (2021).
16. Wang, J. Y. *et al.* Abnormal trajectories in cerebellum and brainstem volumes in carriers of the fragile X premutation. *Neurobiol. Aging* **55**, 11–19. <https://doi.org/10.1016/j.neurobiolaging.2017.03.018> (2017).
17. Wang, J. Y. *et al.* Open-label allopregnanolone treatment of men with fragile X-associated tremor/ataxia syndrome. *Neurotherapeutics* **14**, 1073–1083. <https://doi.org/10.1007/s13311-017-0555-6> (2017).
18. Hagerman, P. J. & Hagerman, R. J. Fragile X-associated tremor/ataxia syndrome (FXTAS). *Ment. Retard. Dev. Disabil. Res. Rev.* **10**, 25–30. <https://doi.org/10.1002/mrdd.20005> (2004).
19. Jiang, L. *et al.* Impaired functional criticality of human brain during Alzheimer's disease progression. *Sci. Rep.* **8**, 1324. <https://doi.org/10.1038/s41598-018-19674-7> (2018).
20. Gibson, G. E., Starkov, A., Blass, J. P., Ratan, R. R. & Beal, M. F. Cause and consequence: Mitochondrial dysfunction initiates and propagates neuronal dysfunction, neuronal death and behavioral abnormalities in age-associated neurodegenerative diseases. *Biochim. Biophys. Acta* **1802**, 122–134. <https://doi.org/10.1016/j.bbadis.2009.08.010> (2010).
21. Ross-Inta, C. *et al.* Evidence of mitochondrial dysfunction in fragile X-associated tremor/ataxia syndrome. *Biochem. J.* **429**, 545–552. <https://doi.org/10.1042/BJ20091960> (2010).
22. Napoli, E. *et al.* Altered zinc transport disrupts mitochondrial protein processing/import in fragile X-associated tremor/ataxia syndrome. *Hum. Mol. Genet.* **20**, 3079–3092. <https://doi.org/10.1093/hmg/ddr211> (2011).
23. Song, G. *et al.* Altered redox mitochondrial biology in the neurodegenerative disorder fragile X-tremor/ataxia syndrome: Use of antioxidants in precision medicine. *Mol. Med.* **22**, 548–559. <https://doi.org/10.2119/molmed.2016.00122> (2016).
24. Napoli, E. *et al.* Impact of *FMR1* premutation on neurobehavior and bioenergetics in young monozygotic twins. *Front. Genet.* **9**, 338. <https://doi.org/10.3389/fgene.2018.00338> (2018).
25. Napoli, E. *et al.* Warburg effect linked to cognitive-executive deficits in *FMR1* premutation. *FASEB J.* **30**, 3334–3351. <https://doi.org/10.1096/fj.201600315R> (2016).
26. Napoli, E. *et al.* Premutation in the fragile X mental retardation 1 (*FMR1*) gene affects maternal Zn-milk and perinatal brain bioenergetics and scaffolding. *Front. Neurosci.* **10**, 159. <https://doi.org/10.3389/fnins.2016.00159> (2016).
27. Napoli, E., Song, G., Wong, S., Hagerman, R. & Giulivi, C. Altered bioenergetics in primary dermal fibroblasts from adult carriers of the *FMR1* premutation before the onset of the neurodegenerative disease fragile X-associated tremor/ataxia syndrome. *Cerebellum* **15**, 552–564. <https://doi.org/10.1007/s12311-016-0779-8> (2016).
28. Giulivi, C., Napoli, E., Tassone, F., Halmaj, J. & Hagerman, R. Plasma metabolic profile delineates roles for neurodegeneration, pro-inflammatory damage and mitochondrial dysfunction in the *FMR1* premutation. *Biochem. J.* **473**, 3871–3888. <https://doi.org/10.1042/BCJ20160585> (2016).
29. Giulivi, C., Napoli, E., Tassone, F., Halmaj, J. & Hagerman, R. Plasma biomarkers for monitoring brain pathophysiology in *FMR1* premutation carriers. *Front. Mol. Neurosci.* **9**, 71. <https://doi.org/10.3389/fnmol.2016.00071> (2016).
30. Kaplan, E. S. *et al.* Early mitochondrial abnormalities in hippocampal neurons cultured from *Fmr1* pre-mutation mouse model. *J. Neurochem.* **123**, 613–621. <https://doi.org/10.1111/j.1471-4159.2012.07936.x> (2012).
31. Conca Dioguardi, C. *et al.* Granulosa cell and oocyte mitochondrial abnormalities in a mouse model of fragile X primary ovarian insufficiency. *Mol. Hum. Reprod.* **22**, 384–396. <https://doi.org/10.1093/molehr/gaw023> (2016).
32. Washizuka, S., Iwamoto, K., Kakiuchi, C., Bundo, M. & Kato, T. Expression of mitochondrial complex I subunit gene *NDUFV2* in the lymphoblastoid cells derived from patients with bipolar disorder and schizophrenia. *Neurosci Res.* **63**, 199–204. <https://doi.org/10.1016/j.neures.2008.12.004> (2009).
33. Scherzer, C. R. *et al.* Molecular markers of early Parkinson's disease based on gene expression in blood. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 955–960. <https://doi.org/10.1073/pnas.0610204104> (2007).
34. Sullivan, P. F., Fan, C. & Perou, C. M. Evaluating the comparability of gene expression in blood and brain. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **141B**, 261–268. <https://doi.org/10.1002/ajmg.b.30272> (2006).
35. Jasinska, A. J. *et al.* Identification of brain transcriptional variation reproduced in peripheral blood: An approach for mapping brain expression traits. *Hum. Mol. Genet.* **18**, 4415–4427. <https://doi.org/10.1093/hmg/ddp397> (2009).
36. Giulivi, C. *et al.* Mitochondrial dysfunction in autism. *JAMA* **304**, 2389–2396. <https://doi.org/10.1001/jama.2010.1706> (2010).
37. Chou, Y. F. & Huang, R. F. Mitochondrial DNA deletions of blood lymphocytes as genetic markers of low folate-related mitochondrial genotoxicity in peripheral tissues. *Eur. J. Nutr.* **48**, 429–436. <https://doi.org/10.1007/s00394-009-0031-0> (2009).
38. Attwell, D. & Laughlin, S. B. An energy budget for signaling in the grey matter of the brain. *J. Cereb. Blood Flow Metab.* **21**, 1133–1145. <https://doi.org/10.1097/00004647-200110000-00001> (2001).
39. Napoli, E., Flores, A., Mansuri, Y., Hagerman, R. J. & Giulivi, C. Sulforaphane improves mitochondrial metabolism in fibroblasts from patients with fragile X-associated tremor and ataxia syndrome. *Neurobiol. Dis.* **157**, 105427. <https://doi.org/10.1016/j.nbd.2021.105427> (2021).
40. Gohel, D. *et al.* *FMR1* polyG alters mitochondrial transcripts level and respiratory chain complex assembly in Fragile X associated tremor/ataxia syndrome [FXTAS]. *Biochim. Biophys. Acta Mol. Basis Dis.* **1865**, 1379–1388. <https://doi.org/10.1016/j.bbadis.2019.02.010> (2019).
41. Holm, K. N. *et al.* Human cerebral cortex proteome of fragile X-associated tremor/ataxia syndrome. *Front. Mol. Biosci.* **7**, 600840. <https://doi.org/10.3389/fmolb.2020.600840> (2020).
42. Ma, L. *et al.* Composition of the intranuclear inclusions of fragile X-associated tremor/ataxia syndrome. *Acta Neuropathol. Commun.* **7**, 143. <https://doi.org/10.1186/s40478-019-0796-1> (2019).
43. Scheffer, M. *et al.* Early-warning signals for critical transitions. *Nature* **461**, 53–59. <https://doi.org/10.1038/nature08227> (2009).
44. Chen, L., Liu, R., Liu, Z. P., Li, M. & Aihara, K. Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers. *Sci. Rep.* **2**, 342. <https://doi.org/10.1038/srep00342> (2012).
45. Liu, X. *et al.* Quantifying critical states of complex diseases using single-sample dynamic network biomarkers. *PLoS Comput. Biol.* **13**, e1005633. <https://doi.org/10.1371/journal.pcbi.1005633> (2017).



46. Liu, X., Liu, R., Zhao, X. M. & Chen, L. Detecting early-warning signals of type 1 diabetes and its leading biomolecular networks by dynamical network biomarkers. *BMC Med. Genom.* **6**(Suppl 2), S8. <https://doi.org/10.1186/1755-8794-6-S2-S8> (2013).
47. Seritan, A. L. *et al.* Dementia in fragile X-associated tremor/ataxia syndrome (FXTAS): Comparison with Alzheimer's disease. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **147B**, 1138–1144. <https://doi.org/10.1002/ajmg.b.30732> (2008).
48. Hagerman, R. J. *et al.* Fragile-X-associated tremor/ataxia syndrome (FXTAS) in females with the *FMR1* premutation. *Am. J. Hum. Genet.* **74**, 1051–1056. <https://doi.org/10.1086/420700> (2004).
49. Perlich, C., Provost, F. & Simonoff, J. S. Tree induction vs. logistic regression: A learning-curve analysis. *J. Mach. Learn. Res.* **4**, 211–255. <https://doi.org/10.1162/153244304322972694> (2004).
50. Medici, V., Czlonkowska, A., Litwin, T. & Giulivi, C. Diagnosis of Wilson disease and its phenotypes by using artificial intelligence. *Biomolecules* **11**, 1243. <https://doi.org/10.3390/biom11081243> (2021).
51. Brunberg, J. A. *et al.* Fragile X premutation carriers: Characteristic MR imaging findings of adult male patients with progressive cerebellar and cognitive dysfunction. *AJNR Am. J. Neuroradiol.* **23**, 1757–1766 (2002).
52. Wang, J. Y., Hagerman, R. J. & Rivera, S. M. A multimodal imaging analysis of subcortical gray matter in fragile X premutation carriers. *Mov. Disord.* **28**, 1278–1284. <https://doi.org/10.1002/mds.25473> (2013).
53. Wang, J. Y. *et al.* Interaction between ventricular expansion and structural changes in the corpus callosum and putamen in males with *FMR1* normal and premutation alleles. *Neurobiol. Aging* **86**, 27–38. <https://doi.org/10.1016/j.neurobiolaging.2019.09.009> (2020).
54. Wang, J. Y. *et al.* Clinical and molecular correlates of abnormal changes in the cerebellum and globus pallidus in fragile X premutation. *Front. Neurol.* **13**, 797649. <https://doi.org/10.3389/fneur.2022.797649> (2022).
55. Hashimoto, R., Javan, A. K., Tassone, F., Hagerman, R. J. & Rivera, S. M. A voxel-based morphometry study of grey matter loss in fragile X-associated tremor/ataxia syndrome. *Brain* **134**, 863–878. <https://doi.org/10.1093/brain/awq368> (2011).
56. Hashimoto, R., Srivastava, S., Tassone, F., Hagerman, R. J. & Rivera, S. M. Diffusion tensor imaging in male premutation carriers of the fragile X mental retardation gene. *Mov. Disord.* **26**, 1329–1336. <https://doi.org/10.1002/mds.23646> (2011).
57. Wang, J. Y. *et al.* Fragile X-associated tremor/ataxia syndrome: Influence of the *FMR1* gene on motor fiber tracts in males with normal and premutation alleles. *JAMA Neurol.* **70**, 1022–1029. <https://doi.org/10.1001/jamaneurol.2013.2934> (2013).
58. Wang, J. M. *et al.* Male carriers of the *FMR1* premutation show altered hippocampal-prefrontal function during memory encoding. *Front. Hum. Neurosci.* **6**, 297. <https://doi.org/10.3389/fnhum.2012.00297> (2012).
59. Filley, C. M. *et al.* White matter disease and cognitive impairment in *FMR1* premutation carriers. *Neurology* **84**, 2146–2152. <https://doi.org/10.1212/WNL.0000000000001612> (2015).
60. Brand, M. D., Hafner, R. P. & Brown, G. C. Control of respiration in non-phosphorylating mitochondria is shared between the proton leak and the respiratory chain. *Biochem. J.* **255**, 535–539 (1988).
61. Sherratt, H. S., Watmough, N. J., Johnson, M. A. & Turnbull, D. M. Methods for study of normal and abnormal skeletal muscle mitochondria. *Methods Biochem. Anal.* **33**, 243–335. <https://doi.org/10.1002/9780470110546.ch6> (1988).
62. Chance, B. & Williams, G. R. The respiratory chain and oxidative phosphorylation. *Adv. Enzymol. Relat. Subj. Biochem.* **17**, 65–134. <https://doi.org/10.1002/9780470122624.ch2> (1956).
63. Murphy, A. N., Fiskum, G. & Beal, M. F. Mitochondria in neurodegeneration: Bioenergetic function in cell life and death. *J. Cereb. Blood Flow Metab.* **19**, 231–245. <https://doi.org/10.1097/00004647-199903000-00001> (1999).
64. Di Bella, D. *et al.* Mutations in the mitochondrial protease gene AFG3L2 cause dominant hereditary ataxia SCA28. *Nat. Genet.* **42**, 313–321. <https://doi.org/10.1038/ng.544> (2010).
65. Kish, S. J. *et al.* Decreased brain protein levels of cytochrome oxidase subunits in Alzheimer's disease and in hereditary spinocerebellar ataxia disorders: A nonspecific change?. *J. Neurochem.* **72**, 700–707. <https://doi.org/10.1046/j.1471-4159.1999.0720700.x> (1999).
66. Schapira, A. H. Mitochondrial involvement in Parkinson's disease, Huntington's disease, hereditary spastic paraplegia and Friedreich's ataxia. *Biochim. Biophys. Acta* **1410**, 159–170. [https://doi.org/10.1016/s0005-2728\(98\)00164-9](https://doi.org/10.1016/s0005-2728(98)00164-9) (1999).
67. Capkova, M. *et al.* Activities of cytochrome c oxidase and citrate synthase in lymphocytes of obese and normal-weight subjects. *Int. J. Obes. Relat. Metab. Disord.* **26**, 1110–1117. <https://doi.org/10.1038/sj.sjo.0802055> (2002).
68. Liu, C. S. *et al.* Oxidative stress-related alteration of the copy number of mitochondrial DNA in human leukocytes. *Free Radic. Res.* **37**, 1307–1317. <https://doi.org/10.1080/10715760310001621342> (2003).
69. Naviaux, R. K. & Nguyen, K. V. POLG mutations associated with Alpers' syndrome and mitochondrial DNA depletion. *Ann. Neurol.* **55**, 706–712. <https://doi.org/10.1002/ana.20079> (2004).
70. Saada, A. *et al.* Mutant mitochondrial thymidine kinase in mitochondrial DNA depletion myopathy. *Nat. Genet.* **29**, 342–344. <https://doi.org/10.1038/ng751> (2001).
71. Sarzi, E. *et al.* Mitochondrial DNA depletion is a prevalent cause of multiple respiratory chain deficiency in childhood. *J. Pediatr.* **150**, 531–534 e531–536. <https://doi.org/10.1016/j.jpeds.2007.01.044> (2007).
72. Vu, T. H. *et al.* Clinical manifestations of mitochondrial DNA depletion. *Neurology* **50**, 1783–1790. <https://doi.org/10.1212/wnl.50.6.1783> (1998).
73. Barrientos, A. *et al.* Reduced steady-state levels of mitochondrial RNA and increased mitochondrial DNA amount in human brain with aging. *Brain Res. Mol. Brain Res.* **52**, 284–289. [https://doi.org/10.1016/s0169-328x\(97\)00278-7](https://doi.org/10.1016/s0169-328x(97)00278-7) (1997).
74. Lee, H. C., Yin, P. H., Lu, C. Y., Chi, C. W. & Wei, Y. H. Increase of mitochondria and mitochondrial DNA in response to oxidative stress in human cells. *Biochem. J.* **348**(Pt 2), 425–432 (2000).
75. Masayeva, B. G. *et al.* Mitochondrial DNA content increase in response to cigarette smoking. *Cancer Epidemiol. Biomark. Prev.* **15**, 19–24. <https://doi.org/10.1158/1055-9965.EPI-05-0210> (2006).
76. May-Panloup, P. *et al.* Increased sperm mitochondrial DNA content in male infertility. *Hum. Reprod.* **18**, 550–556. <https://doi.org/10.1093/humrep/deg096> (2003).
77. Noack, H. *et al.* TFAM-dependent and independent dynamics of mtDNA levels in C2C12 myoblasts caused by redox stress. *Biochim. Biophys. Acta* **1760**, 141–150. <https://doi.org/10.1016/j.bbagen.2005.12.007> (2006).
78. LeCun, Y., Bottou, L., Orr, G. B. & Muller, K. R. Efficient backprop. *Neural Netw. Tricks Trade* **1524**, 9–50. [https://doi.org/10.1007/3-540-49430-8\\_2](https://doi.org/10.1007/3-540-49430-8_2) (1998).
79. Parker, K. L. & Thornbrugh, A. L. parallelized back-propagation training and its effectiveness. *IJCNN-90-Wash DC: International Joint Conference on Neural Networks, Vols 1 and 2*, B179–B182 (1990).
80. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. *Nature* **323**, 533–536. <https://doi.org/10.1038/323533a0> (1986).
81. Shannon, C. E. The mathematical theory of communication. 1963. *MD Comput.* **14**, 306–317 (1997).
82. Quinlan, J. R. *Readings in Knowledge Acquisition and Learning: Automating the Construction and Improvement of Expert Systems* 349–361 (Morgan Kaufmann Publishers Inc., 1993).
83. Shannon, C. E. *A Mathematical Theory of Communication* (American Telephone and Telegraph Company, 1948).
84. Quinlan, J. R., Morris, G., Jackson, J. & O'Connell, P. *C4.5: Programs for Machine Learning* (Morgan Kaufmann Publishers, 1993).
85. Houtkooper, R. H. *et al.* Mitonuclear protein imbalance as a conserved longevity mechanism. *Nature* **497**, 451–457. <https://doi.org/10.1038/nature12188> (2013).
86. Soto, I. *et al.* Balanced mitochondrial and cytosolic translomes underlie the biogenesis of human respiratory complexes. *Genome Biol.* **23**, 170. <https://doi.org/10.1186/s13059-022-02732-9> (2022).



87. Orr, A. L. *et al.* Inhibitors of ROS production by the ubiquinone-binding site of mitochondrial complex I identified by chemical screening. *Free Radic. Biol. Med.* **65**, 1047–1059. <https://doi.org/10.1016/j.freeradbiomed.2013.08.170> (2013).
88. Chowdhury, S. K., Gemin, A. & Singh, G. High activity of mitochondrial glycerophosphate dehydrogenase and glycerophosphate-dependent ROS production in prostate cancer cell lines. *Biochem. Biophys. Res. Commun.* **333**, 1139–1145. <https://doi.org/10.1016/j.bbrc.2005.06.017> (2005).
89. Honzik, T. *et al.* Specific properties of heavy fraction of mitochondria from human-term placenta—Glycerophosphate-dependent hydrogen peroxide production. *Placenta* **27**, 348–356. <https://doi.org/10.1016/j.placenta.2005.03.010> (2006).
90. Madiraju, A. K. *et al.* Metformin suppresses gluconeogenesis by inhibiting mitochondrial glycerophosphate dehydrogenase. *Nature* **510**, 542–546. <https://doi.org/10.1038/nature13270> (2014).
91. Rognstad, R. & Clark, D. G. Effects of aminoxyacetate on the metabolism of isolated liver cells. *Arch. Biochem. Biophys.* **161**, 638–646. [https://doi.org/10.1016/0003-9861\(74\)90348-8](https://doi.org/10.1016/0003-9861(74)90348-8) (1974).
92. Cornell, N. W., Zuurendonk, P. F., Kerich, M. J. & Straight, C. B. Selective inhibition of alanine aminotransferase and aspartate aminotransferase in rat hepatocytes. *Biochem. J.* **220**, 707–716. <https://doi.org/10.1042/bj2200707> (1984).
93. Foretz, M. *et al.* Metformin inhibits hepatic gluconeogenesis in mice independently of the LKB1/AMPK pathway via a decrease in hepatic energy state. *J. Clin. Investig.* **120**, 2355–2369. <https://doi.org/10.1172/JCI40671> (2010).
94. Froment, P. *et al.* At the crossroads of fertility and metabolism: The importance of AMPK-dependent signaling in female infertility associated with hyperandrogenism. *Hum. Reprod.* **37**, 1207–1228. <https://doi.org/10.1093/humrep/deac067> (2022).
95. Faure, M. *et al.* Metformin in reproductive biology. *Front. Endocrinol. (Lausanne)* **9**, 675. <https://doi.org/10.3389/fendo.2018.00675> (2018).
96. John, R. A. & Charteris, A. The reaction of amino-oxycetate with pyridoxal phosphate-dependent enzymes. *Biochem. J.* **171**, 771–779. <https://doi.org/10.1042/bj1710771> (1978).
97. Berry-Kravis, E., Potanos, K., Weinberg, D., Zhou, L. & Goetz, C. G. Fragile X-associated tremor/ataxia syndrome in sisters related to X-inactivation. *Ann. Neurol.* **57**, 144–147. <https://doi.org/10.1002/ana.20360> (2005).
98. Hundscheid, R. D., Smits, A. P., Thomas, C. M., Kiemeny, L. A. & Braat, D. D. Female carriers of fragile X premutations have no increased risk for additional diseases other than premature ovarian failure. *Am. J. Med. Genet. A* **117A**, 6–9. <https://doi.org/10.1002/ajmg.a.10862> (2003).
99. Hara, Y. *et al.* Presynaptic mitochondrial morphology in monkey prefrontal cortex correlates with working memory and is improved with estrogen treatment. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 486–491. <https://doi.org/10.1073/pnas.1311310110> (2014).
100. Wardlaw, J. M., Valdes Hernandez, M. C. & Munoz-Maniega, S. What are white matter hyperintensities made of? Relevance to vascular cognitive impairment. *J. Am. Heart Assoc.* **4**, 001140. <https://doi.org/10.1161/JAHA.114.001140> (2015).
101. Prins, N. D. & Scheltens, P. White matter hyperintensities, cognitive impairment and dementia: An update. *Nat. Rev. Neurol.* **11**, 157–165. <https://doi.org/10.1038/nrneurol.2015.10> (2015).
102. Haller, S. *et al.* Do brain T2/FLAIR white matter hyperintensities correspond to myelin loss in normal aging? A radiologic-neuropathologic correlation study. *Acta Neuropathol. Commun.* **1**, 14. <https://doi.org/10.1186/2051-5960-1-14> (2013).
103. Shim, Y. S. *et al.* Pathological correlates of white matter hyperintensities on magnetic resonance imaging. *Dement. Geriatr. Cogn. Disord.* **39**, 92–104. <https://doi.org/10.1159/000366411> (2015).
104. Morales, H. & Tomsick, T. Middle cerebellar peduncles: Magnetic resonance imaging and pathophysiologic correlate. *World J. Radiol.* **7**, 438–447. <https://doi.org/10.4329/wjr.v7.i12.438> (2015).
105. Jiang, J. *et al.* Bilateral middle cerebellar peduncle lesions: Neuroimaging features and differential diagnoses. *Brain Behav.* **10**, e01778. <https://doi.org/10.1002/brb3.1778> (2020).
106. Aulbert, W. *et al.* Long survival in Leigh syndrome: New cases and review of literature. *Neuropediatrics* **45**, 346–353. <https://doi.org/10.1055/s-0034-1383823> (2014).
107. Napoli, E. *et al.* Characterization of the metabolic, clinical and neuropsychological phenotype of female carriers of the premutation in the X-linked FMR1 gene. *Front. Mol. Biosci.* **7**, 578640. <https://doi.org/10.3389/fmolb.2020.578640> (2020).
108. Mracek, T., Drahotka, Z. & Houstek, J. The function and the role of the mitochondrial glycerol-3-phosphate dehydrogenase in mammalian tissues. *Biochim. Biophys. Acta* **1827**, 401–410. <https://doi.org/10.1016/j.bbabo.2012.11.014> (2013).
109. Schonfeld, P. & Reiser, G. Why does brain metabolism not favor burning of fatty acids to provide energy? Reflections on disadvantages of the use of free fatty acids as fuel for brain. *J. Cereb. Blood Flow Metab.* **33**, 1493–1499. <https://doi.org/10.1038/jcbfm.2013.128> (2013).
110. van den Heuvel, L. P., Smeitink, J. A. & Rodenburg, R. J. Biochemical examination of fibroblasts in the diagnosis and research of oxidative phosphorylation (OXPHOS) defects. *Mitochondrion* **4**, 395–401. <https://doi.org/10.1016/j.mito.2004.07.005> (2004).
111. Petrazzini, B. O., Naya, H., Lopez-Bello, F., Vazquez, G. & Spangenberg, L. Evaluation of different approaches for missing data imputation on features associated to genomic data. *BioData Min.* **14**, 44. <https://doi.org/10.1186/s13040-021-00274-7> (2021).
112. Tustison, N. J. *et al.* N4ITK: Improved N3 bias correction. *IEEE Trans. Med. Imaging* **29**, 1310–1320. <https://doi.org/10.1109/TMI.2010.2046908> (2010).
113. Smith, S. M. *et al.* Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. *Neuroimage* **17**, 479–489. <https://doi.org/10.1006/nimg.2002.1040> (2002).
114. Schmidt, P. *et al.* An automated tool for detection of FLAIR-hyperintense white-matter lesions in Multiple Sclerosis. *Neuroimage* **59**, 3774–3783. <https://doi.org/10.1016/j.neuroimage.2011.11.032> (2012).
115. Yushkevich, P. A. *et al.* User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage* **31**, 1116–1128. <https://doi.org/10.1016/j.neuroimage.2006.01.015> (2006).
116. Levene, H. In *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling* (eds Olkin, I. & Hotteling, H.) 278–292 (Stanford University Press, 1960).
117. Martin, W. E. & Bridgmon, K. D. *Quantitative and Statistical Research Methods: From Hypothesis to Results* 1st edn. (Jossey-Bass, 2012).
118. Nordstokke, D. W. & Zumbo, B. D. A new nonparametric Levene test for equal variances. *Psicologica* **31**, 401–430 (2010).
119. Metsalu, T. & Vilo, J. ClustVis: A web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res.* **43**, W566–W570. <https://doi.org/10.1093/nar/gkv468> (2015).
120. Orr, A. L. *et al.* Novel inhibitors of mitochondrial sn-glycerol 3-phosphate dehydrogenase. *PLoS One* **9**, e89938. <https://doi.org/10.1371/journal.pone.0089938> (2014).

## Acknowledgements

We thank the participants and their families for their dedication and support for this research. We thank the technical assistance of Drs. Eleonora Napoli and Julian Halmaj, Gyu Song, and Jennifer William.

## Author contributions

Conceptualization, C.G.; methodology, R.J.H., J.W., C.G.; formal analysis, C.G.; data resources, all; data curation, R.J.H.; writing—original draft preparation, C.G.; writing—review and editing, all; funding acquisition, R.J.H., C.G. All authors have agreed on the final version of this study.

## Funding

This research was supported by NICHD (HD 036071), the MIND Institute Intellectual and Developmental Disabilities Research Center (IDDR) (P50 HD103526), and NINDS (NS 128751 and NS 128751). The content is solely the authors' responsibility and does not necessarily represent the official views of the National Institutes of Health or any other funding agency.

## Competing interests

The funders (NICHD, NINDS) had no role in the study's design, in the collection, analysis, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results. R.J.H. has received funding from Zynherba and the Azrieli Foundation for carrying out treatment studies in patients with fragile X syndrome. She has also consulted with Zynherba regarding treatment studies in individuals with FXS. C.G. is an Editorial Board Member of *Scientific Reports* and received compensation as a Specialty Chief Editor for *Frontiers in Cell and Developmental Biology*. R.J.H. and C.G. received honoraria from participating in NIH peer review meetings. J.Y.W. has no conflict to declare. No other potential financial and non-financial competing interests that could directly or indirectly undermine the objectivity, integrity, and value of this publication through a potential influence on the judgments and actions of authors regarding objective data presentation, analysis, and interpretation, are found.

## Additional information

**Correspondence** and requests for materials should be addressed to C.G.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022