

# Hardy-Weinberg Equilibrium in different mitochondrial haplogroups of four genes associated with neuroprotection and neurodegeneration

Equilíbrio de Hardy-Weinberg em diferentes haplogrupos mitocondriais de quatro genes associados à neuroproteção e neurodegeneração

Sheida MEHRPOUR<sup>1</sup>, Camila Ronqui RODRIGUES<sup>2</sup>, Renata Carmona FERREIRA<sup>1,3</sup>, Marcelo Ribeiro da Silva BRIONES<sup>4</sup>, Acary Souza Bulle OLIVEIRA<sup>1</sup>

## ABSTRACT

**Background:** Malfunctioning or damaged mitochondria result in altered energy metabolism, redox equilibrium, and cellular dynamics and is a central point in the pathogenesis of neurological disorders such as Alzheimer's disease, Parkinson's disease, Huntington's disease and Amyotrophic Lateral Sclerosis. Therefore, it is of utmost importance to identify mitochondrial genetic susceptibility markers for neurodegenerative diseases. Potential markers include the respiratory chain enzymes Riboflavin kinase (*RFK*), Flavin adenine dinucleotide synthetase (*FAD*), Succinate dehydrogenase B subunit (*SDHB*), and Cytochrome C1 (*CYC1*). These enzymes are associated with neuroprotection and neurodegeneration. **Objective:** To test if variants in genes *RFK*, *FAD*, *SDHB* and *CYC1* deviate from Hardy-Weinberg Equilibrium (HWE) in different human mitochondrial haplogroups. **Methods:** Sequence variants in genes *RFK*, *FAD*, *SDHB* and *CYC1* of 2,504 non-affected individuals of the 1,000 genomes project were used for mitochondrial haplogroup assessment and HWE calculations in different mitochondrial haplogroups. **Results:** We show that *RFK* variants deviate from HWE in haplogroups G, H, L, V and W, variants of *FAD* in haplogroups B, J, L, U, and C, variants of *SDHB* in relation to the C, W, and A and *CYC1* variants in B, L, U, D, and T. HWE deviation indicates action of selective pressures and genetic drift. **Conclusions:** HWE deviation of particular variants in relation to global populational HWE, could be, at least in part, associated with the differential susceptibility of specific populations and ethnicities to neurodegenerative diseases. Our data might contribute to the epidemiology and diagnostic/prognostic methods for neurodegenerative diseases.

**Keywords:** Alzheimer Disease; Parkinson Disease; Huntington Disease; Amyotrophic Lateral Sclerosis.

## RESUMO






**Introdução:** Mitocôndrias defeituosas ou danificadas resultam em alterações do metabolismo energético, equilíbrio redox e dinâmica celular e são, portanto, identificadas como o ponto central da patogênese em muitos distúrbios neurológicos, como a doença de Alzheimer, a doença de Parkinson, a doença de Huntington e a Esclerose Lateral Amiotrófica. Portanto, é de fundamental importância identificar marcadores de susceptibilidade genética mitocondrial para doenças neurodegenerativas. Entre os potenciais marcadores relevantes estão as enzimas da cadeia respiratória riboflavina quinase (*RFK*), flavina adenina dinucleotídeo sintetase (*FAD*), succinato desidrogenase subunidade B (*SDHB*) e citocromo C1 (*CYC1*). Estas enzimas estão associadas à neuroproteção e à neurodegeneração. **Objetivo:** Testar se variantes nas sequências dos genes *RFK*, *FAD*, *SDHB* e *CYC1* desviam do Equilíbrio de Hardy-Weinberg (HWE) em diferentes haplogrupos mitocondriais humanos. **Métodos:** Neste trabalho utilizamos os variantes nos genes *RFK*, *FAD*, *SDHB* e *CYC1* de sequências de 2.504 indivíduos não afetados do projeto de 1.000 genomas para o cálculo dos valores de HWE em diferentes haplogrupos mitocondriais. **Resultados:** As variantes de *RFK* desviam de HWE nos haplogrupos G, H, L, V e W, variantes de *FAD* nos haplogrupos B, J, L, U e C, variantes de *SDHB* em relação às variantes C, W e A e *CYC1* em B, L, U, D e T. O desvio de HWE indica a ação de pressões seletivas e

<sup>1</sup>Universidade Federal de São Paulo, Departamento de Neurologia e Neurocirurgia, São Paulo SP, Brazil.

<sup>2</sup>Universidade Federal de São Paulo, Departamento de Microbiologia, Imunologia e Parasitologia, São Paulo SP, Brazil.

<sup>3</sup>Bridges Genomics, São Paulo SP, Brazil.

<sup>4</sup>Universidade Federal de São Paulo, Escola Paulista de Medicina, Divisão de Bioinformática, São Paulo SP, Brazil.

Sheida MEHRPOUR  <https://orcid.org/0000-0002-1217-8071>; Camila Ronqui RODRIGUES  <https://orcid.org/0000-0001-5250-9712>; Renata Carmona FERREIRA  <https://orcid.org/0000-0003-1481-2926>; Marcelo Ribeiro da Silva BRIONES  <https://orcid.org/0000-0001-8045-2477>; Acary Souza Bulle OLIVEIRA  <https://orcid.org/0000-0002-6986-4937>

**Correspondence:** Acary Souza Bulle Oliveira; E-mail: acary.bulle@unifesp.br

**Conflict of interest:** There is no conflict of interest to declare.

**Support:** Coordination for the Improvement of Higher Education Personnel (Capes), Foundation for Research Support of the State of São Paulo (Fapesp), and National Council for Scientific and Technological Development (CNPq).

Received on August 12, 2019; Received in final its form on November 22, 2019; Accepted on December 09, 2019.

desvio genético. **Conclusões:** O desvio do HWE de variantes particulares em relação ao HWE populacional global poderia estar, pelo menos em parte, associado à suscetibilidade diferencial de populações e etnias específicas a doenças neurodegenerativas. Nossos dados podem contribuir para a epidemiologia e métodos diagnósticos/prognósticos para doenças neurodegenerativas.

**Palavras-chave:** Doença de Alzheimer; Doença de Parkinson; Doença de Huntington; Esclerose Amiotrófica Lateral.

Several studies have shown that mitochondria play an important role in the pathogenesis of neurodegenerative diseases<sup>1</sup>. Robust evidence suggests that structural changes in mitochondria, including increased mitochondrial fragmentation and decreased mitochondrial fusion, are critical factors associated with mitochondrial dysfunction and with several types of neurons' death in ALS, including motor neurons<sup>2</sup>.

Human populations can be classified in mitochondrial haplogroups based on specific single nucleotide variants (SNVs) scattered along the mitochondrial genome. Mutations in mitochondrial DNA are very common, so the distribution of mutant mitochondria during mitosis affects mitochondrial function in the daughter cells, generating variability in mitochondrial function<sup>3</sup>.

Common complex diseases have been associated with specific mitochondrial haplogroups. There is strong evidence that mitochondrial DNA variation plays a role in the development and progression of complex human diseases. Mitochondrial genetic variations have been widely recognized in the context of genome-wide association studies (GWAS), confirming the need for studies that investigate the function of mitochondrial DNA variation in human health and disease<sup>4</sup>.

Studies on mitochondrial haplogroups allow the identification of adequate pathways in the discovery of susceptibility of a specific disease, alteration in response of a drug or environmental factors. Characterization of the haplogroups is done according to the mitochondrial mutation rate, mainly the mutations located in the control region of mitochondria. Therefore, if some diseases are affected by mitochondrial dysfunction, they behave differently depending on the patient's haplogroup<sup>5</sup>.

Another factor that can make variations in the genetic characteristics and consequently affect the function of a gene in relation to the incidence of a disease is the microevolution process involving nuclear genes and/or rare mitochondrial DNA<sup>6</sup>.

Most mitochondrial diseases are generally identified in relation to mitochondrial respiratory chain compromised due to quantitative and qualitative defects of both mitochondrial DNA and nuclear DNA, which are inherited according to the mitochondrial genetics rules and Mendelian inheritance, respectively<sup>7</sup>. These diseases are classified genetically by three types: sporadic appearance (by mitochondrial DNA mutation — duplications or deletions); maternal inheritance (typically by point mutations in mitochondrial DNA); mendelian inheritance (typically due to nuclear DNA defects)<sup>8</sup>.

Previous studies have shown that enzymes of the mitochondrial respiratory chain complexes such as Riboflavin kinase (*RFK*), Flavin adenine synthase (*FAD* synthase), Succinate dehydrogenase B subunit (*SDHB*), Cytochrome C1 (*CYCI*), are involved in neuroprotection mechanisms ensuring neuron survival. Quantitative or qualitative defects in genes of these enzymes lead to neurodegeneration<sup>9,10,11,12,13</sup>. In addition, in a study with patients with Amyotrophic Lateral Sclerosis (ALS), Lin and collaborators, using quantitative real-time polymerase chain reaction (PCR) method showed that mRNA expression levels of *RFK*, *FAD*, *SDHB*, *CYCI* are significantly lower in comparison to controls, confirming quantitative effects of these enzymes, especially in ALS patients<sup>14</sup>.

Considering that these enzymes may be associated with neurodegenerative diseases, we studied the combination of genotypic information of mitochondrial haplogroups and variants of these enzymes in normal, non-affected individuals. After studying the combination between the genotypic information of the individuals with their own mitochondrial haplogroups, we were able to identify the differential distribution of these variants in the different mitochondrial haplogroups.

## METHODS

### Genome sequence data

The 2,504 genome sequences used in this project were from the 1,000 genomes project (<https://www.internationalgenome.org/>)<sup>15,16,17,18</sup>. This sample consists of 26 worldwide distributed populations grouped in five super-populations (African, American, East Asian, European and South Asian)<sup>15</sup>. Although the level of sample kinship is small due to its global scale, some authors identified that some level of inbreeding is detectable with the FSuite program<sup>19</sup>. The 1,000 genomes project consisted in three phases, phase 1 being the pilot. This phase contained sequences derived from trios (Father, Mother, Proband). The 1,000 genomes project guidelines ensured that a significant proportion of samples should be from unrelated individuals: [https://www.internationalgenome.org/sample\\_collection\\_principles](https://www.internationalgenome.org/sample_collection_principles).

The mitochondrial haplogroups are groups of types of mitochondrial genomes with a genealogical relatedness and defining certain sets of mutations (<http://www.phylotree.org/>)<sup>20</sup>. The consent for subjects included in the sampled populations followed the project guidelines: ([https://www.internationalgenome.org/sample\\_collection\\_principles/](https://www.internationalgenome.org/sample_collection_principles/))

and are available in: <https://www.internationalgenome.org/sites/1000genomes.org/files/docs/Informed%20Consent%20Form%20Template.pdf>

Ensemble Genome Browser was used to access the sequencing information and single nucleotide variants of 2,504 normal, non-affected, individuals from the 1000 genomes project<sup>21</sup>. Genes selected for the analysis were Riboflavin kinase (*RFK*), Flavin adenine dinucleotide synthetase (*FAD*), Succinate dehydrogenase B subunit (*SDHB*), and Cytochrome C1 (*CYCI*). Variants were filtered by minimum allele frequency (MAF)>0.1. The SNVs with highest global allele frequency for each individual gene were selected for HWE analysis. The SNVs analyzed are depicted with the dbSNP database ID<sup>22</sup>. The information on each individual SNV is available using the “rs” ID number for search in the dbSNP (<https://www.ncbi.nlm.nih.gov/snp/>).

### Hardy-Weinberg Equilibrium Calculation

For the analysis of the data from the Hardy-Weinberg calculation, chi-square test was used, with one degree of freedom and significance level set at 0.05 ( $\alpha=0.05$ ), (The Pearson’s chi-square goodness-of-fit)<sup>23,24</sup>. For more accuracy, we used the program’s in-built Hardy-Weinberg calculator to compute the chi-square index: <https://www.easycalculation.com/health/hardy-weinberg-equilibrium-calculator.php>. Critical value is 3.841 for one degree of freedom and  $\alpha=0.05$ . If chi-square values for each haplogroup are greater than the genotype, frequencies of the variant deviate from equilibrium.

The null hypothesis in this test states that there is no difference between observed and expected values, that is, variables are independent, and according to the alternative hypothesis, there is a significant difference between observed and expected values and the two variables are dependent. According to the test, when the chi-square index is greater than the table value or critical value, the null hypothesis is rejected, determining the existence of the correlation between the two variables and, evidently, as the chi-square index was higher than the critical value, we had stronger evidence to reject the null hypothesis in favor of the existence of stronger association between the two variables.

Therefore, for a particular haplogroup, deviation from the HWE suggests the relationship between variants and mitochondrial haplogroups, such that variants deviate in relation to a haplogroup are in higher distributions in haplogroup populations and, consequently, populations are more susceptible to neurodegeneration. We describe these haplogroups as “haplogroups at risk”.

## RESULTS

The results here presented bear on variants in genes *RFK*, *FAD*, *SDHB* and *CYCI* that were selected for highest global allele frequency. The analysis showed 36 variants that are not in HWE in the genes *RFK*, *FAD*, *SDHB* and *CYCI*. These four genes are associated with neuroprotection mechanisms. Here we tested if differences in allele dynamics, as revealed by HWE, might suggest a differential distribution of neuroprotection in different mitochondrial haplogroups. The chi-square critical value in HWE was 3.841 (one gene, two alleles, one degree of freedom for  $\alpha=0.05$ ). Variants with chi-square value higher than the threshold were considered a risk factor for neurodegeneration.

The mitochondrial haplogroups were defined for the 1,000 genomes project database from the mitochondrial genomes using Haplogrep<sup>25</sup> and were fully consistent with a parallel published inference<sup>26</sup>.

There are nine variants in disequilibrium in the *RFK* gene in seven mitochondrial haplogroups, as described in Table 1. Haplogroups G, V and W showed to be more susceptible to neurodegeneration than others haplogroups at risk.

In the *FAD* gene, there are eight variants in disequilibrium in five mitochondrial haplogroups (Table 2).

Haplogroups B and L show chi-square values significantly higher than the other haplogroups. Halogroup J showed risk for neurodegeneration for all variants in disequilibrium. The *SDHB* gene showed eight variants in disequilibrium in seven mitochondrial haplogroups (Table 3).

Haplogroups A and W showed to be more susceptible to neurodegeneration than other haplogroups. The *CYCI* gene

**Table 1.** The chi-square statistics of seven variants with highest global allele frequency in the *RFK* gene. The variants with “-” are in Hardy-Weinberg Equilibrium. Variant ID=number in dbSNP (<https://www.ncbi.nlm.nih.gov/snp/>).

Variant ID	Haplogroups						
	C	D	G	H	L	V	W
rs2501925	-	-	4.3556	4.6172	5.5946	7.4903	-
rs2501928	-	-	4.3556	-	-	4.3900	4.5918
rs2490579	-	-	4.3556	-	-	7.4903	4.5918
rs2501923	-	-	4.3556	-	-	7.4903	4.5918
rs10716702	-	-	4.3556	-	-	5.7600	4.5918
rs11447646	-	-	-	-	-	4.0000	6.6345
rs11144870	3.8649	4.6646	-	-	10.5977	-	-

showed nine variants in disequilibrium. The deletion (A/-) in variant rs60547285 showed a greater disequilibrium in 20 mitochondrial haplogroups (Table 4).

## DISCUSSION

The accumulation of mutations in mitochondrial DNA leads to oxidative damage, energy reduction, and increased Reactive Oxygen Species (ROS) production. This can cause disease or negatively affect longevity in individuals or in a population that shares the same genotype of mitochondrial DNA. Perhaps the opposite may also be true for different polymorphisms, which may be protective<sup>27</sup>.

Several studies correlate mitochondrial haplogroups with disease susceptibility or protective roles. In 114 healthy Spanish men, Marcuello et al.<sup>28</sup> found that J Haplogroup was associated with reduced Electron Transport Chain (ETC) efficiency, decreased Adenosine Triphosphate (ATP) and ROS production, possibly explaining the accumulation of this variant in elderly people.

**Table 2.** The chi-square statistics of eight variants with highest global allele frequency in the *FAD* gene. Variants with “-” are in Hardy-Weinberg Equilibrium. Variant ID=number in dbSNP (<https://www.ncbi.nlm.nih.gov/snp/>).

Variant ID	Haplogroups				
	B	C	J	L	U
rs10908449	14.0890	-	6.7867	7.0089	5.2518
rs11587981	14.0890	-	6.7867	7.9965	5.2518
rs34756997	14.0890	-	6.7867	7.5335	5.2518
rs7528060	19.7781	6.4471	6.7867	4.3409	5.2518
rs7525580	19.7781	6.4471	6.7867	-	5.2518
rs12096640	19.7781	6.4471	6.7867	-	5.2518
rs7535144	-	6.4471	6.7867	-	8.7015
rs6427111	-	-	-	49.8291	-

The data here used for HWE calculations are obtained from a global sample and therefore small local changes might not have the necessary significance in the HWE. The 1,000 genomes project dataset contains samples that can be considered as close as possible to a random population, although in phase 1 (pilot) some trios were included. Because of the global nature of the sample collected from 26 worldwide locations in five different continents, it is likely that minor local population effects are not interfering with global HWE here measured.

Van der Walt et al.<sup>29</sup> observed that haplogroups J and K reduced the incidence of Parkinson’s disease (PD) by 50%. Another analysis revealed that the SNP 9055A in *ATP6* (which defines haplogroup K) reduced the risk in women, and the SNP 13708A in *ND5* was protective in individuals over 70 years (haplogroup J)<sup>29</sup>. The UKJT group was associated with a 22% reduction in risk for PD but not with the risk of Alzheimer’s disease (AD), confirming that the association with PD was disease specific and not a general effect observed in all neurodegenerative diseases<sup>30</sup>. In a large cohort of 620 Italian patients with PD, it was found that haplogroup K was associated with a lower risk of PD<sup>31</sup>.

In relation to Friedreich’s ataxia (FA), Giacchetti et al.<sup>32</sup> studied 99 patients with FA and 48 control subjects, all from southern Italy, showing that patients with haplogroup U had a delay of five years in the onset of the disease and a lower rate of cardiomyopathy. However, no significant difference was found in the frequency distribution of haplogroups between patients and controls<sup>32</sup>.

In a study conducted by Mancuso and collaborators<sup>33</sup>, fifty-one patients with Huntington’s disease (HD), a trinucleotide expansion disorder caused by a CAG expansion in the IT15 gene, were compared with 181 controls. The frequency of haplogroups and clusters of haplogroups did not differ between the two groups, and they did not observe a correlation with sex, age of onset, and status of the disease<sup>33</sup>. However, in a recent study, Arning et al. demonstrated a significantly lower initial age of patients with H haplogroup<sup>34</sup>.

**Table 3.** The chi-square statistics of eight variants with highest global allele frequency in the *SDHB* gene. The variants with “-” are in Hardy-Weinberg Equilibrium. Variant ID=number in dbSNP (<https://www.ncbi.nlm.nih.gov/snp/>).

Variant ID	Haplogroups						
	A	C	F	H	L	W	X
rs59090025	-	4.6656	-	-	-	4.5217	-
rs9435739	6.0196	-	-	-	-	-	-
rs3754508	-	5.7248	-	-	-	6.5789	-
rs2871775	7.6271	-	-	-	3.8495	-	4.4444
rs2295059	5.1739	-	-	-	-	6.74	-
rs10887990	5.1739	-	-	-	-	6.74	-
rs11203287	5.1739	-	-	-	-	6.74	-
rs978528	-	-	5.8302	4.8024	-	-	-

In order to better clarify the involvement of mitochondria in the pathogenesis of AD, studies showed the possible association of mitochondrial haplogroups and susceptibility to the disease. According to studies, European mitochondrial haplogroups may be related to longevity<sup>35,36</sup> as well as to neurodegeneration, the risk of AD, and therefore death in Caucasians. In an article by Chagnon et al.<sup>37</sup>, haplogroup T is described as under-represented in AD, whereas haplogroup J is overrepresented. Van der Walt et al.<sup>38</sup> showed that men classified as U-haplogroup had a significant increase in the risk of AD, whereas women showed a significant decrease in risk with the same U-haplogroup. To assess the relationship between haplogroup and AD in an Iranian population, the two hypervariable segments of mitochondrial DNA in 30 patients with AD and 100 control subjects were sequenced<sup>39</sup>. It was found that H and U haplogroups are significantly more abundant in AD patients, presuming that these two haplogroups may act synergistically to increase the penetrance of AD<sup>39</sup>. When studying an Italian sample, Carrieri et al.<sup>40</sup> hypothesized that haplogroups K and U may act by neutralizing the effect of the major risk factor allele of AD, apolipoprotein E. However, this association was not confirmed in another study<sup>41</sup>. An association between AD and haplogroups G2a, B4c1 and N9b1 was described in 96 Japanese AD patients<sup>42</sup>. A study conducted in a Polish population found that the HV cluster is significantly associated with the risk for AD, although no evidence was reported for the involvement of haplogroups U, K, J or T in the risk for AD<sup>43</sup>.

To investigate whether specific genetic polymorphisms within mitochondrial DNA could act as susceptibility factors and contribute to the clinical expression of sporadic amyotrophic lateral sclerosis (ALS), the mitochondrial haplogroups in 222 patients of clear Italian origin were analyzed for sporadic ALS and 151 paired controls<sup>44</sup>. The frequency of haplogroups was lower in cases of ALS than in controls. Age of onset, severity and neurological system involved in the disease were not associated with haplogroups. In a comparison developed to test what makes haplogroup I different from the other haplogroups tested, a highly significant difference was found in alleles 16391A and 10034C. They concluded that mitochondrial DNA polymorphisms may contribute to motor neuron degeneration, possibly interacting with unknown genetic or environmental factors. However, this finding was not confirmed by a study with a large British cohort of 504 patients with ALS and 493 controls, and reported no evidence that mitochondrial DNA haplogroups contribute to the risk of developing ALS<sup>45</sup>.

The conflicting findings in relation to several neurodegenerative diseases may be related to analysis only in the mitochondrial DNA. Other studies have shown the reciprocal interaction between mitochondrial DNA and nuclear genome. Gene expression uses a high proportion of cellular energy and protein content of mitochondria is responsible for ATP production through the mechanism of oxidative phosphorylation. Mitochondria, depending on genome content, may alter the expression profile of the nuclear genes of cells differently due to a high amount of the mitochondrial

**Table 4.** The chi-square results of nine variants with highest global allele frequency in the *CYC1* gene. The variants with “-” are in Hardy-Weinberg Equilibrium. Variant ID=number in dbSNP (<https://www.ncbi.nlm.nih.gov/snp/>).

Haplogroup	rs60547285	rs144925641	rs11433813	rs11780874	rs13255347	rs13254954	rs7822232	rs12550729	rs11541475
A	79.3915	-	-	-	-	-	-	-	-
B	160.5336	6.0456	12.0432	-	-	6.0456	6.0456	6.0456	6.0456
C	46.8595	-	-	-	-	-	-	-	-
D	109.9769	-	-	8.7873	8.7873	-	-	-	-
F	88.0856	-	-	-	-	-	-	-	-
G	20.0000	-	-	-	-	-	-	-	-
H	138.585	-	-	-	-	-	-	-	-
HV	16.3543	-	-	-	-	-	-	-	-
J	28.4549	-	-	-	-	-	-	-	-
K	18.7995	-	-	-	-	-	-	-	-
L	118.2736	-	7.6375	10.3141	9.4369	-	-	-	-
M	264.4310	-	-	-	-	-	-	-	-
N	37.0000	-	-	-	-	-	-	-	-
R	93.6729	-	4.4856	-	-	-	-	-	-
T	39.1443	-	4.2287	-	-	-	-	-	-
U	102.1655	-	-	4.0536	4.0536	-	-	-	-
V	16.0000	-	-	-	-	-	-	-	-
W	21.3018	-	-	-	-	-	-	-	-
X	6.6942	-	-	-	-	-	-	-	-
Z	4.8395	-	-	-	-	-	-	-	-



DNA sequence that vary between human populations<sup>46</sup> and over time evolution, mitochondrial-nuclear interactions come highly specific<sup>47</sup>. On the other hand, more than 99% of mitochondrial proteins are encoded through nuclear genes and deleterious or protective effect of mitochondria compromises the importance of cytosolic proteins<sup>48</sup>. Therefore, the expression of nuclear genes may vary depending on the mitochondrial haplogroup, also if the nuclear gene is affected by some genetic alterations in its variants due to the process of microevolution, it can affect mitochondrial function differently. Through these interactions and considering that, in general, point mutations in enzyme variants may affect the stabilization of the enzyme transition state and consequently reduce their total catalytic activity<sup>49</sup>, one can conclude that the genetic expression of the components of the mitochondrial respiratory chain complexes encoded through the nuclear genes may vary depending on the mitochondrial haplogroup. Moreover, if the nuclear gene is affected by some mutations due to the microevolution process, in a different way, it can affect the function of mitochondria in final ATP production in population haplogroups.

In our work, we found variants that were out of equilibrium in relation to specific haplogroups. These were considered as “Haplogroups at risk”, which in most cases were clusters of identical haplogroups.

According to the analysis of outcome measures, variants of *RFK* encoded enzyme, including rs2501928, rs2490579, rs2501923, rs10716702, deviate from HWE relative to haplogroups G, V, W. Other variants of this enzyme had at least some of these groups in common, for example, rs2501925 (G, H, L, V) and rs11447646 (V, W). Some variants deviate from the HWE relative to the totally different haplogroups, such as rs11144870 (C, D, L), while still others did not deviate from any haplogroups, including rs60002266 and rs7210880.

In the case of the *CYCI* encoded enzyme, variants rs11780874 and rs13255347 deviate from the equilibrium relative to the set (D, L, U) and variants rs13254954, rs78222232, rs11541475, rs12550729, rs144925641 deviated in relation to the single haplogroup (B), while the insertion of rs11433813 deviated from equilibrium in relation to the haplogroups B,

L, R, T. Interestingly, variant rs60547285 deviated from the equilibrium in almost all haplogroups except I and Y, probably due to the very small sample size.

Therefore, according to the results here described, our hypothesis is that departure from the equilibrium in relation to a specific haplogroup results in a reciprocal interaction between mitochondrial genome and nuclear genes. This is related to the modification of the expression of variants through the mitochondrial genome and the presence of point mutations in the deviant variants, which leads to reduction in the catalytic activity of enzymes. All these interactions cause the deviant variants carriers in different haplogroup populations to be susceptible to neurodegeneration in different levels.

The main objective of this study was to identify variants of enzymes of *RFK*, *FAD*, *SDHB*, and *CYCI*, which, theoretically and according to analysis of calculations and statistical studies, may have greater distribution in haplogroup populations and increased susceptibility to neurodegenerative diseases. Because the study was based on the probability of occurrence of events, the results obtained should be compared to those seen in clinical trials. Another critical point of the study is the small size of some populations of haplogroups, as it may probably affect the accuracy of results in HWE calculation, because the larger the study population is, the more reliable results will be.

In this study, we identified differential distribution of variants in four genes encoding mitochondrial respiratory chain enzymes. These four genes (*RFK*, *FAD*, *SDHB*, *CYCI*) differ significantly in HWE statistics in different haplogroups. Departure from HWE in specific mitochondrial haplogroups can be explained by: (a) differential natural selection, (b) genetic drift due to small population size, (c) extensive migration and (e) significant non-random mating. Departure from HWE in some of the haplogroups and not others suggests that the specific haplogroups within a population could be more susceptible to neurodegeneration associated with AD, PD, HD and ALS, and that this susceptibility could increase the higher the chi-square index for each specific haplogroup is.

## References

1. Lin MT, Beal MF. Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases. *Nature*. 2006 Oct;443(7113):787-95. <https://doi.org/10.1038/nature05292>
2. Smith EF, Shaw PJ, De Vos KJ. The role of mitochondria in amyotrophic lateral sclerosis. *Neurosci Lett*. 2017 Jun;710:132933. <https://doi.org/10.1016/j.neulet.2017.06.052>
3. Samuels DC, Carothers AD, Horton R, Chinnery PF. The Power to Detect Disease Associations with Mitochondrial DNA Haplogroups. *Am J Hum Genet*. 2006 Apr;78(4):713-20. <https://doi.org/10.1086/502682>
4. Mitchell SL, Goodloe R, Brown-Gentry K, Pendergrass SA, Murdock DG, Crawford DC. Characterization of mitochondrial haplogroups in a large population-based sample from the United States. *Hum Genet*. 2014 Feb;133(7):861-8. <https://doi.org/10.1007/s00439-014-1421-9>
5. Urzua-Traslavina CG. Relationship of Mitochondrial DNA Haplogroups with Complex Diseases. *J Genet Genome Res*. 2014;1(2). <https://doi.org/10.23937/2378-3648/1410011>
6. Tatarenkov A, Avise JC. Rapid concerted evolution in animal mitochondrial DNA. *Proc R Soc B Biol Sci*. 2007 May;274(1619):1795-8. <https://doi.org/10.1098/rspb.2007.0169>
7. DiMauro S, Schon EA. Mitochondrial respiratory-chain diseases. *N Engl J Med*. 2003 Jun;348(26):2656-68. <https://doi.org/10.1056/NEJMra022567>

8. DiMauro S, Bonilla E, Davidson M, Hirano M, Schon EA. Mitochondria in neuromuscular disorders. *Biochim Biophys Acta*. 1998 Aug;1366(1-2):199-210. [https://doi.org/10.1016/s0005-2728\(98\)00113-3](https://doi.org/10.1016/s0005-2728(98)00113-3)
9. Marashly ET, Bohlega SA. Riboflavin has neuroprotective potential: focus on Parkinson's disease and migraine. *Front Neurol*. 2017 Jul;8. <https://doi.org/10.3389/fneur.2017.00333>
10. Udhayabanu T, Manole A, Rajeshwari M, Varalakshmi P, Houlden H, Ashokkumar B. Riboflavin responsive mitochondrial dysfunction in neurodegenerative diseases. *J Clin Med*. 2017;6(5):52. <https://doi.org/10.3390/jcm6050052>
11. Mao P, Ardeshiri A, Jacks R, Yang S, Hurn PD, Alkayed NJ. Mitochondrial mechanism of neuroprotection by CART. *Eur J Neurosci*. 2007 Aug;26(3):624-32. <https://doi.org/10.1111/j.1460-9568.2007.05691.x>
12. Yu Z, Zhang Y, Liu N, et al. Roles of neuroglobin binding to mitochondrial complex III subunit cytochrome c1 in oxygen-glucose deprivation-induced neurotoxicity in primary neurons. *Mol Neurobiol*. 2016;53(5):3249-57. <https://doi.org/10.1007/s12035-015-9273-4>
13. Weng T-Y, Tsai S-YA, Su T-P. Roles of sigma-1 receptors on mitochondrial functions relevant to neurodegenerative diseases. *J Biomed Sci*. 2017 Sep;24(1):74. <https://doi.org/10.1186/s12929-017-0380-6>
14. Lin J, Diamanduros A, Chowdhury SA, Scelsa S, Latov N, Sadiq SA. Specific electron transport chain abnormalities in amyotrophic lateral sclerosis. *J Neurol*. 2009 Feb;256(5):774-82. <https://doi.org/10.1007/s00415-009-5015-8>
15. Sudmant PH, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, Huddleston J, et al. An integrated map of structural variation in 2,504 human genomes. *Nature*. 2015 Sep;526(7571):75-81. <https://doi.org/10.1038/nature15394>
16. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature*. 2015 Sep;526(7571):68-74. <https://doi.org/10.1038/nature15393>
17. The 1000 Genomes Project Consortium, Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, et al. A map of human genome variation from population-scale sequencing. *Nature*. 2010 Oct;467(7319):1061-73. <https://doi.org/10.1038/nature09534>
18. The 1000 Genomes Project Consortium, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, et al. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012 Nov;491(7422):56-65. <https://doi.org/10.1038/nature11632>
19. Gazal S, Sahbatou M, Babron M-C, Génin E, Leutenegger A-L. High level of inbreeding in final phase of 1000 Genomes Project. *Sci Rep*. 2015;5:17453. <https://doi.org/10.1038/srep17453>
20. van Oven M, Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat*. 2009 Feb;30(2):E386-E394. <https://doi.org/10.1002/humu.20921>
21. Hunt SE, McLaren W, Gil L, Thormann A, Schuilenburg H, Sheppard D, et al. Ensembl variation resources. *Database*. 2018 Nov;2018. <https://doi.org/10.1093/database/bay119>
22. Kitts A, Sherry S. The Single Nucleotide Polymorphism Database (DbSNP) of nucleotide sequence variation. Bethesda (MD): National Center for Biotechnology Information (US); 2011. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK21088/>
23. Hardy GH. Mendelian proportions in a mixed population. *Science*. 1908 Jul;28(706):49-50. <https://doi.org/10.1126/science.28.706.49>
24. Weinberg W. Über den Nachweis der Vererbung beim Menschen [On the demonstration of heredity in man]. *Naturkunde Württemberg*. 1908;64:368-82.
25. Weissensteiner H, Pacher D, Kloss-Brandstätter A, Forer L, Specht G, Bandelt HJ, et al. HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res*. 2016 Jul;44(W1):W58-W63. <https://doi.org/10.1093/nar/gkw233>
26. Rishishwar L, Jordan IK. Implications of human evolution and admixture for mitochondrial replacement therapy. *BMC Genomics*. 2017 Feb;18(1):140. <https://doi.org/10.1186/s12864-017-3539-3>
27. Ienco EC, Simoncini C, Orsucci D, Petrucci L, Filosto M, Mancuso M, et al. May "Mitochondrial Eve" and mitochondrial haplogroups play a role in neurodegeneration and Alzheimer's disease? *Int J Alzheimers Dis*. <https://doi.org/10.4061/2011/709061>
28. Marcuello A, Martínez-Redondo D, Dahmani Y, Casajús JÁ, Ruiz-Pesiniab E, Montoya J, et al. Human mitochondrial variants influence on oxygen consumption. *Mitochondrion*. 2009 Feb;9(1):27-30. <https://doi.org/10.1016/j.mito.2008.10.002>
29. van der Walt JM, Nicodemus KK, Martin ER, Schmechel DE, Pericak-Vance MA, Vance JM, et al. Mitochondrial Polymorphisms Significantly Reduce the Risk of Parkinson disease. *Am J Hum Genet*. 2003 Apr;72(4):804-11. <https://doi.org/10.1086/373937>
30. Pyle A, Foltynie T, Tiangyou W, Lambert C, Keers SM, Allcock LM, et al. Mitochondrial DNA haplogroup cluster UKJT reduces the risk of PD. *Ann Neurol*. 2005 Apr;57(4):564-67. <https://doi.org/10.1002/ana.20417>
31. Ghezzi D, Marelli C, Achilli A, Goldwurm S, Pezzoli G, Barone P, et al. Mitochondrial DNA haplogroup K is associated with a lower risk of Parkinson's disease in Italians. *Eur J Hum Genet*. 2005;13(6):748-52. <https://doi.org/10.1038/sj.ejhg.5201425>
32. Giacchetti M, Monticelli A, De Biase I, Pianese L, Turano M, Filla A, et al. Mitochondrial DNA haplogroups influence the Friedreich's ataxia phenotype. *J Med Genet*. 2004;41(4):293-5. <https://doi.org/10.1136/jmg.2003.015289>
33. Mancuso M, Kiferle L, Petrozzi L, Nestia C, Rocchia A, Ceravolo R, et al. Mitochondrial DNA haplogroups do not influence the Huntington's disease phenotype. *Neurosci Lett*. 2008 Oct;444(1):83-6. <https://doi.org/10.1016/j.neulet.2008.08.013>
34. Arning L, Haghikia A, Taherzadeh-Fard E, Saft C, Andrich J, Pula B, et al. Mitochondrial haplogroup H correlates with ATP levels and age at onset in Huntington disease. *J Mol Med*. 2010 Jan;88(4):431-6. <https://doi.org/10.1007/s00109-010-0589-2>
35. Niemi A-K, Moilanen JS, Tanaka M, Hervoonen A, Hurme M, Lehtimäki T, et al. A combination of three common inherited mitochondrial DNA polymorphisms promotes longevity in Finnish and Japanese subjects. *Eur J Hum Genet*. 2005 Feb;13(2):166-70. <https://doi.org/10.1038/sj.ejhg.5201308>
36. Zhang J, Asin-Cayuela J, Fish J, Michikawa Y, Bonafé M, Olivieri F, et al. Strikingly higher frequency in centenarians and twins of mtDNA mutation causing remodeling of replication origin in leukocytes. *Proc Natl Acad Sci U S A*. 2003 Feb;100(3):1116-21. <https://doi.org/10.1073/pnas.242719399>
37. Chagnon P, Gee M, Filion M, Robitaille Y, Belouchi M, Gauvreau D. Phylogenetic analysis of the mitochondrial genome indicates significant differences between patients with Alzheimer disease and controls in a French-Canadian founder population. *Am J Med Genet*. 1999 Jul;85(1):20-30. [https://doi.org/10.1002/\(sici\)1096-8628\(19990702\)85:1%3C20::aid-ajmg6%3E3.0.co;2-k](https://doi.org/10.1002/(sici)1096-8628(19990702)85:1%3C20::aid-ajmg6%3E3.0.co;2-k)
38. van der Walt JM, Dementieva YA, Martin ER, Scott WK, Nicodemus KK, Kronera CC, et al. Analysis of European mitochondrial haplogroups with Alzheimer disease risk. *Neurosci Lett*. 2004 Jul;365(1):28-32. <https://doi.org/10.1016/j.neulet.2004.04.051>
39. Fesahat F, Houshmand M, Panahi MSS, Gharagozli K, Mirzajani F. Do haplogroups H and U act to increase the penetrance of Alzheimer's disease? *Cell Mol Neurobiol*. 2007 May;27(3):329-34. <https://doi.org/10.1007/s10571-006-9126-9>
40. Carrieri G, Bonafé M, De Luca M, Rose G, Varcasia O, Bruni A, et al. Mitochondrial DNA haplogroups and APOE4 allele are non-independent variables in sporadic Alzheimer's disease. *Hum Genet*. 2001 Mar;108(3):194-8. <https://doi.org/10.1007/s004390100463>
41. Elson JL, Herrnstadt C, Preston G, Thal L, Morris CM, Edvardson JA, et al. Does the mitochondrial genome play a role in the etiology of Alzheimer's disease? *Hum Genet*. 2006 Jan;119(3):241-54. <https://doi.org/10.1007/s00439-005-0123-8>

42. Takasaki S. Mitochondrial haplogroups associated with Japanese Alzheimer's patients. *J Bioenerg Biomembr*. 2009;41(5):407-10. <https://doi.org/10.1007/s10863-009-9240-8>
43. Maruszak A, Canter JA, Styczyńska M, Żekanowski C, Barcikowska M. Mitochondrial haplogroup H and Alzheimer's disease--is there a connection? *Neurobiol Aging*. 2009 Nov;30(11):1749-55. <https://doi.org/10.1016/j.neurobiolaging.2008.01.004>
44. Mancuso M, Conforti FL, Rocchi A, Tessitore A, Muglia M, Tedeschi G, et al. Could mitochondrial haplogroups play a role in sporadic amyotrophic lateral sclerosis? *Neurosci Lett*. 2004 Nov;371(2-3):158-62. <https://doi.org/10.1016/j.neulet.2004.08.060>
45. Chinnery PF, Mowbray C, Elliot H, Elson JL, Nixon H, Hartley J, et al. Mitochondrial DNA haplogroups and amyotrophic lateral sclerosis. *Neurogenetics*. 2007 Jan;8(1):65-7. <https://doi.org/10.1007/s10048-006-0066-9>
46. Muir R, Diot A, Poulton J. Mitochondrial content is central to nuclear gene expression: Profound implications for human health. *BioEssays News Rev Mol Cell Dev Biol*. 2016 Feb;38(2):150-6. <https://doi.org/10.1002/bies.201500105>
47. Reinhardt K, Dowling DK, Morrow EH. Mitochondrial replacement, evolution, and the clinic. *Science*. 2013 Sep;341(6152):1345-6. <https://doi.org/10.1126/science.1237146>
48. Boengler K, Heusch G, Schulz R. Mitochondria in postconditioning. *Antioxid Redox Signal*. 2011 Jan;14(5):863-80. <https://doi.org/10.1089/ars.2010.3309>
49. Ishida T. Effects of point mutation on enzymatic activity: correlation between protein electronic structure and motion in chorismate mutase reaction. *J Am Chem Soc*. 2010 Apr;132(20):7104-8. <https://doi.org/10.1021/ja100744h>