EXTENDED ABSTRACT

Journal of the World Mitochondria Society
Issue n°2, Vol. 2
DOI 10.18143/JWMS_v2i2_1911

Mitochondrial genome wide association analysis in several complex diseases and lifestyle factors using the KORA population

FLAQUER, Antònia[1,2], HEIER, Hanna-Teresa[1,2], LADWIG, Karl-Heinz[3], WALDENBERGER, Melanie[3,4], GRALLERT, Harald[3,4,5], MEISINGER, Christa[3,5], MEITINGER, Thomas[6,7], PETERS Annette[3,5], STRAUCH, Konstantin[1,2].

[1] Institute of Medical Informatics, Biometry and Epidemiology, Chair of Genetic Epidemiology, Ludwig-Maximilians-Universität, Munich, 81377, Germany . [2] Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, 85764, Germany. [3] Institute of Epidemiology II, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, 85764, Germany.[4] Research Unit of Molecular Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, 85764, Germany. [5] German Center for Diabetes Research (DZD e.V.), Neuherberg, 85764, Germany. [6] Institute of Human Genetics, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, 85764, Germany. [7] Institute of Human Genetics, Technical University, Munich, 81675, Germany.

Corresponding author:
Antònia Flaquer
antonia.flaquer@lmu.de

Abstract

Our main goal is to identify mitochondrial susceptibility genes for human complex diseases. Using different genotyping platforms we performed mitochondrial genome wide association analysis (mtGWAs) in the KORA population (3,080 individuals) with several complex diseases (post-traumatic stress disorder, metabolic syndrome, depression, and asthma), physiological parameters (BMI, cholesterol profile, and triglycerides) as well as with life style factors (smoking, alcohol consumption, and physical activity). In this study we report several mitochondrial genetic variants associated with post-traumatic stress disorder and metabolic syndrome. BMI, HDL cholesterol and triglyceride (TG) levels were also associated to mitochondrial variants, indicating that the presence of heteroplasmy in these variants may influence the balance of HDL cholesterol and TG levels. We also observed that lack of physical activity, alcohol abuse, and smoking may increase the heteroplasmy levels of mitochondrial genetic variants. As fluctuations in mtDNA heteroplasmy may signify alterations in cellular activity, our findings highlight the important role of the mtDNA as a biomarker for aging-related diseases, metabolic syndrome, and lifestyle factors.

Introduction

The study of human genetic diseases has often centered on the human nuclear genome - the sprawling linear code of about 3.2 billion nucleotides and more than 20,000 genes spread across 23 pairs of chromosomes (nDNA). In contrast, the "other" genome, the mitochondrial genome (mtDNA), has received less attention. Although the mtDNA dwarfs it in size – a short, circular genome with just 16,569 nucleotides and exactly 37 genes (2 rRNA genes, 22 tRNA genes, and 13 structural genes) – it is arguably just as important as the nDNA. The mitochondrial genes are essential for encoding subunits of the mitochondrial respiratory chain responsible for oxidative metabolism and adenosine triphosphate (ATP)

Received: 26 October 2016; Accepted: 24 February 2017
© Copyright JWMS 2016-2017 – All rights reserved.

generation. In addition mitochondria are major producers of reactive oxygen species (ROS). ROS levels can increase drastically during times of environmental stress and cause significant damage to cellular structures including the mitochondria themselves which, in turn, can lead to mitochondrial dysfunction, possibly even to apoptosis.

There can be tens, hundreds or thousands of mitochondria per cell. Since each mitochondrion carries on average five mitochondrial genomes, there can, theoretically, be tens, hundreds or thousands of mitochondrial genomes per cell. MtDNA is characterized by a high rate of mutation, much higher than nDNA. When there is a mutation in a mitochondrial genome, it may occur in only one mitochondrion in a cell, or the mitochondrion may have many exact copies, so the same mutation is seen either in a few, some, or most of the mitochondria in the cell or the person as a whole. Consequently mutant and wild-type mtDNA can co-exist (Wallace and Chalkia, 2013). This is known as heteroplasmy effect which was originally believed to be a rare phenomenon. New studies found that heteroplasmy is quite common (Ye et al., 2014) – and tends to increase with age. Currently, mutations of mtDNA are under a growing scientific spotlight and there is increasing evidence that these mutations play a central role in many, if not most, human diseases.

Methods

Study design and population: Data are based on the Cooperative Health Research in the Region of Augsburg (KORA) study. The KORA study is a series of independent population-based epidemiological surveys and follow-up studies of unrelated participants living in the region of Augsburg, in southern Germany (Wichmann et al., 2005). The most recent KORA studies are KORA S4 (1999-2001) with its follow-up KORA F4 (2006-2008) and KORA S3 (1994-1995) with its follow-up KORA F3 (2004-2005). The present study includes data of the studies F3, F4, and additionally from those individuals of S4 that did not participate in F4, including a total number of 6,528 unrelated persons. Testing for population stratification using EIGENSOFT and principal components analysis found no evidence of population stratification in the KORA genotype data.

Genotyping: DNA was extracted from full blood after the blood draw and then stored at -80°C. Genotyping was performed using different platforms such as the Affymetrix 6.0 GeneChip (465 mtSNPs), Affymetrix Axiom chip (252 mtSNPs), Illumina Human Exome Beadchip (226 mtSNPs), and Illumina MetaboChip 200K (135 mtSNPs). Only 9 positions are common to all four chips. In total, 645 independent mtSNPs have been genotyped. However, when all chips are considered together, good overall coverage of the mitochondrial genome is obtained (Flaquer et al., 2014).

Phenotypes investigated within this project: Table 1 summarizes the investigated disease phenotypes, physiological parameters, and lifestyle phenotypes.

Table 1: Summary of the phenotypes investigated within this project.	**Type of variable**	**KORA survey**	**Sample size cases/controls**
Disease phenotypes			
Posttraumatic stress disorder (PTSD): No PTSD* / Partial PTSD / Full PTSD	Categorical	F4	1,238 (total) 875 / 312 / 51
Metabolic syndrome	Binary	F4	1,076 / 1,937
Asthma	Binary	S4,F4	269 / 2768
Depression	Binary	F4	581 / 1970
Physiological parameters			
Body mass index (BMI)	Quantitative	S4,F4,F3	6,528
Cholesterol profiles: Total cholesterol (TC) / HDL / LDL	Quantitative	F4	2,804
Triglycerides (TG)	Quantitative	F4	2,804
Lifestyle factors			
Physical activity (hours/week): frequently (2) / regularly (1) / unregularly (1) / no physical activity (0)	Categorical	F4	3,021 (total) 732 / 915 / 398 / 976
Alcohol intake (g/day): no alcohol (0) / light (1–20) / moderate (21–60) / heavy intake (>60)	Categorical	F4	3,021 (total) 912 / 1,223 / 766 / 120
Smoking: non-smoker, former smoker, occasional smoker, regular smoker	Categorical	F4	3021 (total) 1332 / 1145 / 79 / 465

Sample size: total number of individuals with phenotype and genotype information after quality control.
**In the category no-PTSD (controls) are considered individuals that had a trauma but having no symptoms of PTSD.*

Statistical methods: After stringent quality control we approached the heteroplasmy present in 978 mtSNPs using raw signals of luminous intensity (Flaquer et al., 2014), where every measurement is associated with a specific mtSNP and represents one of its alleles. The number of measurements *n* per mtSNP depends on the vendor-specific technology employed on the genotyping chip. At the very least there have to be two signals, one for each of the two alleles. Often,

Received: 26 October 2016; Accepted: 24 February 2017
© Copyright JISANH 2016-2017 – All rights reserved.

however, the chip design includes more than one measurement per SNP and allele. That is, for every individual and SNP we have intensity measurements (A_1, B_1), ..., (A_n, B_n) with n >= 1 where A_i and B_i represent the intensities of the two alleles A and B. To assess association between phenotype-mtSNP intensities we applied regression analysis using the phenotype of interest as response variable (*Y*). The mtSNP enters the model as a covariate via the $\log_2$-transformed intensity ratio, $z=\log_2(\overline{A}/\overline{B})$, where $\overline{A}$ and $\overline{B}$ denote the mean intensity, or single measure in case of n=1, for the A allele and B allele, respectively. Sex and age are introduced in the model as covariates. The type of the regression model we used depends on the distribution of Y; if it is continuous and approximately normally distributed we used a linear regression model; if it is dichotomous we used logistic regression. For all phenotypes P-values were obtained from the Wald test, and adjusted for multiple comparisons applying the Bonferroni correction or the Benjamini-Hochberg false discovery rate (FDR) in the case of BMI. The adjusted significance threshold was set at 0.05. All the analyses were performed with the statistical software R (R Core Team., 2013). For more details about the statistical method we refer to (Flaquer et al., 2014).

Results

Disease phenotypes and physiological parameters

The association results that remained significant after adjustment for multiple testing ($P_{adjusted} \leq 0.05$) are plotted in Fig. 1 for each disease phenotype.

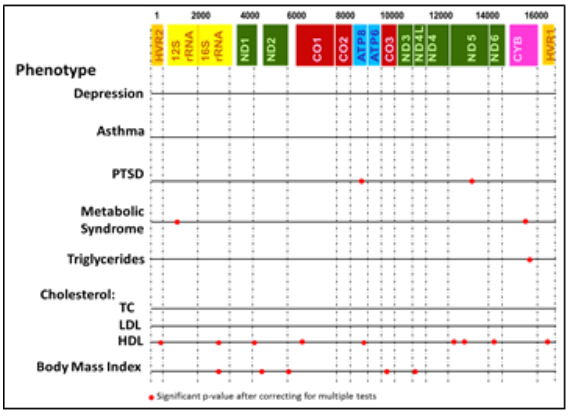

Figure 1. *illustrates the significant mtSNPs for each of the disease phenotypes.*

Depression: No significant mtSNPs were obtained after analysing the depression phenotype.

Asthma: Nominally significant association was found for two mtSNPs, but they were not strongly supportive after correcting for multiple testing.

PTSD: Significant associations were obtained between full versus no PTSD in two mtSNPs, mt8414 located in ATP synthase subunit 8 (*MT-ATP8*) and mt12501 located in the NADH dehydrogenase subunit 5 (*MT-ND5*). Heteroplasmy for the two variants towards a larger number of the respective mutated variant increases the risk of having PTSD (Flaquer et al., 2015a).

Metabolic syndrome: Two mtSNPs were significantly associated with metabolic syndrome, mt14793 located in the subunit of ubiquinol-cytochrome c reductase of complex III (MT-CYB) and mt927 located in the region coding for 12S rRNA (MT-RNR1). These findings suggest that participants with metabolic syndrome had a significantly higher level of heteroplasmy for the *MT-RN1*$_{mt927}$ and the *MT-CYB*$_{mt14793}$ mtSNPs than participants without metabolic syndrome. The effect estimates for sex and age were significant, indicating that men are at higher risk of developing metabolic syndrome with increasing age than women.

Triglyceride (TG) levels: For TG levels the only significant mtSNP, mt15074, was found in the *MT-CYB* gene. Those individuals with presence of mt15074 heteroplasmy show a higher level of TG than those individuals being homoplasmic for the wild type allele at this mtSNP. In the regression model the estimates of sex and age were both significant, indicating significantly higher TG levels in males than in females and with older age (Flaquer et al., 2015b).

Cholesterol levels: No significant mtSNPs were obtained for TC. However, when analysing cholesterol subtypes, ten mtSNPs for HDL cholesterol reached significance. Six of the ten significant mtSNPs for HDL are

Received: 26 October 2016; Accepted: 24 February 2017
© Copyright JISANH 2016-2017 – All rights reserved.

located in the NADH subunit dehydrogenase genes of complex I (*MT-ND1*$_{mt3336}$, *MT-ND2*$_{mt5285}$, *MT-ND5*$_{mt13855,\ mt13598,\ mt14000}$, and *MT-ND6*$_{mt14580}$), the others are located in the ATP synthase subunit 6 (*MT-ATP6*$_{mt9163}$), cytochrome c oxidase subunit 1 (*MT-CO1*$_{mt6671,\ mt6591}$), and tRNA (*MT-TL1*$_{mt3285}$). Taking the most strongly associated variant mt3285T>A, in the *MT-TL1* gene, based on our estimates HDL cholesterol increases with a higher proportion of T (wild type variant) at this locus, making the T variant favourable to HDL. Similar arguments can be applied to the other mtSNPs.The regression models for all significant mtSNPs regarding HDL showed no significant effect for age; however, the estimate of sex was significant with $\beta_{sex}<0$ indicating significantly higher levels of HDL cholesterol in females than in males (Flaquer et al., 2015b).

BMI: Five mtSNPs resulted to be significant when analysing BMI. The two most significant mtSNPs are located in the *Cytochrome c oxidase* 1 and 3 (*MT-CO1*$_{mt6663}$ and *MT-CO3*$_{mt9698}$) genes of complex IV. Another three locations of *NADH subunit dehydrogenase* genes of complex I were also significant (*MT-ND1*$_{mt3336}$, *MT-ND2*$_{mt4851}$, *MT-ND4L*$_{mt10550}$). Heterogeneity in these mtSNPs towards the mutated variant is associated with a higher BMI. In all significant mtSNPs the estimates of sex and age were also significant, reflecting the fact that on average the BMI is higher in males and with older age (Flaquer et al., 2014).

Lifestyle factors

For all lifestyle factors significant effects are considered when Bonferroni corrected p-values are <0.05.

Physical activity: Significant results were obtained in two mtSNPs of the MT-RNR2 gene (MT-RNR2$_{mt2789,\ mt3022}$) when comparing the group of individuals with no physical activity versus those groups of individuals with at least one hour a week of physical activity. No difference in heteroplasmy was observed between the group performing physical activity 1h/week and the group performing physical activity 2h/week. These results show that, in particular, no physical activity increases the level of heteroplasmy at both mtSNPs.

Alcohol intake: When comparing the group of people that does not consume alcohol with the heavy consumers group, two variants showed a higher heteroplasmy level in the latter group (MT-ND5$_{mt12633}$ and MT-CYB$_{mt15833}$). When comparing the group of moderate alcohol intake with the heavy alcohol intake group, one of the former mtSNPs (MT-ND5$_{mt12633}$) and another one (MT-CYBmt$_{15311}$) showed a higher level of heteroplasmy for the latter group. Anyway, no differences were observed when comparing individuals who do not drink alcohol at all with those individuals that consume a moderate amount of alcohol. Significant mtSNPs with higher heteroplasmy were only detected in the group of individuals with a heavy amount of alcohol consumption.

Smoking: When considering the group of ex-smokers with the group of regular smokers significant differences were observed in three mtSNPs (MT-RNR1$_{mt1018,\ mt1048,}$ and MT-CYB$_{mt15074}$). When comparing ex-smokers with non-smokers significant differences were obtained in 6 mtSNPs (MT-RNR1$_{mt2706}$, MT-CO1$_{mt7028}$, MT-ND4$_{mt11719}$, MT-ND5$_{mt13708}$, MT-CYB$_{mt14769,\ mt15074}$). Interestingly, significant results were only obtained in those analyses in which the group of ex-smokers was considered, indicating that ex-smokers are the individuals with higher heteroplasmy levels, even higher than regular smokers. Anyway, interpretation based on these results needs to be done very carefully. When interpreting these results directly the conclusion might be: if a person is smoker, she/he should not stop smoking. Evidently, this conclusion is not true. This kind of controversy result has been observed in several epidemiological studies when analyzing the variable "smoking". It needs to be considered that most of the ex-smokers stopped smoking possibly due to health problems. So, generally ex-smokers may have some defects on the mitochondria as a result of the time they smoked and due to other health related issues.

Conclusion and future research

Our study reports several mtSNPs significantly associated with various complex diseases, physiological parameters, and lifestyle factors. In both cases some mitochondrial genomic regions seem to be altered and are likely to

Received: 26 October 2016; Accepted: 24 February 2017
© Copyright JISANH 2016-2017 – All rights reserved.

influence common human diseases such as posttraumatic stress disorder, metabolic diseases, and cholesterol levels among others. Although further analyses are needed to follow up on the present results, these findings highlight the important role of the mitochondrial genome among the factors that contribute to the risk of complex diseases and suggest that expression of this genome may be more important than has previously been suspected. The exact mechanisms linking mitochondrial dysfunction and complex diseases have not been unraveled so far. Anyway, it is known that a coordinated action between the nuclear and mitochondrial genomes is essential for sustaining life. How the two genomes communicate and interact has not been clarified yet. Human diseases are often genetically complex and, and the penetrance of underlying genetic variants is variable. The understanding of mitochondrial-nuclear interaction will provide a plausible mechanism that contributes to the explanation of this complexity. However, a plausible hypothesis is that mitochondrial dysfunction (possibly due to environmental factors, somatic mutations, and inherited mutations) increases the production of ROS. This excess might induce alteration in the nuclear genome that can alter miRNA expression, and this, in turn, could impact other aspects of epigenetic regulation and ultimately affect the disease phenotype. If such a pathway were to be discovered and understood in sufficient detail, it would help to understand physiological mechanisms leading to complex diseases. Furthermore, it would open up entirely new perspectives for drug development and therapy.

Acknowledgments

We thank all members of the field staff involved in the KORA study and we express our appreciation to all study participants. The KORA research platform (KORA, Cooperative Research in the Region of Augsburg) was initiated and financed by the Helmholtz Zentrum München - German Research Center for Environmental Health which is funded by the German Federal Ministry of Education and Research and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ.

References

Flaquer, A., C. Baumbach, J. Kriebel, T. Meitinger, A. Peters, M. Waldenberger, H. Grallert, and K. Strauch. 2014. Mitochondrial Genetic Variants Identified to Be Associated with BMI in Adults. *PloS one*. 9:e105116.

Flaquer, A., C. Baumbach, K.H. Ladwig, J. Kriebel, M. Waldenberger, H. Grallert, J. Baumert, T. Meitinger, J. Kruse, A. Peters, R. Emeny, and K. Strauch. 2015a. Mitochondrial genetic variants identified to be associated with posttraumatic stress disorder. *Transl Psychiatry*. 5:e524.

Flaquer, A., S. Rospleszcz, E. Reischl, S. Zeilinger, H. Prokisch, T. Meitinger, C. Meisinger, A. Peters, M. Waldenberger, H. Grallert, and K. Strauch. 2015b. Mitochondrial GWA Analysis of Lipid Profile Identifies Genetic Variants to Be Associated with HDL Cholesterol and Triglyceride Levels. *PloS one*. 10:e0126294.

R Core Team. 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Wallace, D.C., and D. Chalkia. 2013. Mitochondrial DNA genetics and the heteroplasmy conundrum in evolution and disease. *Cold Spring Harbor perspectives in medicine*. 3:a021220.

Wichmann, H.E., C. Gieger, and T. Illig. 2005. KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen*. 67 Suppl 1:S26-30.

Ye, K., J. Lu, F. Ma, A. Keinan, and Z. Gu. 2014. Extensive pathogenicity of mitochondrial heteroplasmy in healthy human individuals. *Proceedings of the National Academy of Sciences of the United States of America*. 111:10654-10659.

Received: 26 October 2016; Accepted: 24 February 2017
© Copyright JISANH 2016-2017 – All rights reserved.