ANNALS of Clinical and Translational Neurology
Open Access
ANA American Neurological Association since 1875

BRIEF COMMUNICATION

Contribution of non-reference alleles in mtDNA of Alzheimer's disease patients

Tiziana Casoli[1], Giuseppina Di Stefano[1,a,*], Liana Spazzafumo[2,a], Marta Balietti[3], Belinda Giorgetti[1], Cinzia Giuli[4], Demetrio Postacchini[4], Patrizia Fattoretti[3] & Fiorenzo Conti[1,5]

[1]Center for Neurobiology of Aging, INRCA IRCCS, Ancona 60121, Italy
[2]Center of Biostatistics, INRCA IRCCS, Ancona 60124, Italy
[3]Cellular Bioenergetics Laboratory, Center for Neurobiology of Aging, INRCA IRCCS, Ancona 60121, Italy
[4]Unit of Geriatrics, INRCA IRCCS Hospital, Fermo 63023, Italy
[5]Section of Neuroscience and Cell Biology, Department of Experimental and Clinical Medicine, Università Politecnica delle Marche, Ancona 60126, Italy

Correspondence
Tiziana Casoli, Center for Neurobiology of Aging, INRCA IRCCS, Via Birarelli 8, Ancona 60121, Italy. Tel: +39 071 8004203; Fax: +39 071 206791; E-mail: t.casoli@inrca.it

*Present address: Scientific Direction, INRCA IRCCS, Ancona 60124, Italy.

Funding information
Supported by grant RF-INR-2006-346774 from Italian Ministry of Health and revenues of the "5 per Mille" donations to INRCA.

Received: 29 August 2013; Revised: 20 December 2013; Accepted: 20 January 2014

***Annals of Clinical and Translational Neurology* 2014; 1(4): 284–289**

doi: 10.1002/acn3.42

[a]Giuseppina Di Stefano and Liana Spazzafumo contributed equally to this work.

Abstract

Many observations suggest that mutations of mitochondrial DNA (mtDNA) could be responsible for the neurodegenerative changes of Alzheimer's disease (AD). Here we examined the signal intensity of the four alleles of each mtDNA nucleotide position (np) in whole blood of AD patients and age-matched controls using MitoChip v2.0 array. Our analysis identified 270 significantly different nps which, with one exception, showed an increased contribution of non-reference alleles in AD patients. Principal component analysis (PCA) and cluster analysis showed that five of these nps could discriminate AD from control subjects with 80% of cases correctly classified. Our data support the hypothesis of mtDNA alterations as an important factor in the etiology of AD.

Introduction

An interesting hypothesis concerning Alzheimer's disease (AD) development posits a cause–effect relationship between accumulation of mitochondrial DNA (mtDNA) mutations and neurodegenerative changes associated with this pathology, for example, defective oxidative phosphorylation, increased oxidative stress, accumulation of amyloid-ß (Aß) and apoptosis.[1,2] Human cells contain thousands of mtDNA copies and the most common mutations are insertion, deletions, and point mutations that are classified as homoplasmic or heteroplasmic if they regard all mtDNA copies or only a fraction of them, respectively.[3] As regards homoplasmic point mutations, studies on association between mtDNA haplogroups and AD have not provided to date definite conclusions.[4,5] In brains of patients with AD, heteroplasmic mutations are increased by 63% compared to controls,[6] while AD and elderly subjects show a higher aggregate burden of low-level heteroplasmic point mutations than younger subjects.[7] To further our understanding of allelic variations, we investigated in whole blood of AD patients and age-matched controls the quantitative levels of the four bases for each nucleotide position (np) on both strands of the entire sequence of mtDNA by the resequencing array MitoChip v2.0 which allows analysis of mtDNA sequences with high reproducibility and sensitivity.[8,9]

Table 1. Characteristics of the study subjects.

	AD patients	Controls	*P* value
Age	76.0 ± 1.24	74.1 ± 1.46	0.328
Gender (M/F)	3/15	4/14	0.674
MMSE	18.2 ± 1.01	28.0 ± 0.22	<0.001
ADL	5.4 ± 0.26	6.0 ± 0.00	0.039
IADL	3.7 ± 0.43	8.0 ± 0.00	<0.001

Values are expressed as mean ± standard error of the mean (SEM). *P* values have been calculated by *t*-test except for the comparison of gender where χ^2 has been applied.

Materials and Methods

Subjects

This study included a total of 18 AD patients and 18 age-matched controls enrolled from October 2009 to September 2010 at the INRCA Geriatrics Unit of Fermo. A priori power analysis, performed by one-sided *t*-test with 5% significance level, indicated that a sample size of 36 subjects (18 AD and 18 controls) provides ~80% of power to detect differences among groups. The research was approved by the Institutional Ethical Committee and each participant or caregiver provided informed consent to participate. Patients were diagnosed according to National Institute of Neurological and Communicative Diseases and Stroke/Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria for probable AD by an extended neuropsychological and functional evaluation, neuroimaging, and laboratory tests. Controls consisted of volunteers and spouses of the patients and underwent the same assessment as the AD group (Table 1).

MitoChip v2.0 resequencing array protocol

Total DNA was extracted from whole blood using QIAamp DNA Blood Mini kit (Qiagen, Hilden, Germany) and amplified by REPLI-g mitochondrial DNA kit (Qiagen, Hilden, Germany). After purification, DNA was quantified spectrophotometrically and Genechip Resequencing array kit (Affymetrix, Santa Clara, CA) was employed for fragmentation and labeling. After loading, MitoChip arrays were washed and stained in Fluidics Station 450 and scanned in Affymetrix GeneChip Scanner 3000 7G.

MitoChip v2.0 array analysis

Affymetrix MitoChip v2.0 is a mtDNA sequencing array with eight 25-mer probes/base position (four oligonucleotide probes/strand) corresponding to the whole revised Cambridge Reference Sequence (rCRS). Each 25-mer probe is varied at the central position to incorporate each possible nucleotide (A, G, C, or T). Data acquisition was performed using the Affymetrix Genechip Command Console (AGCC) software and analysis was carried out with GSEQ 4.1. Output files utilized for this study were the Single Nucleotide Polymorphism (SNP) View and Probe Intensity Files that provide, for each np, the base call classified as homoplasmic or heteroplasmic and value of fluorescence intensity of the four bases for both sense and antisense filaments. The quantitative estimates of allelic contribution were used to calculate REA (Ratio of Expected Allele) index, which is the log ratio of the signal intensity of the reference allele at any site, as indicated in the rCRS, to the average signal intensity of the other three alleles from the sense and antisense strand.[10] A high value of the index suggests that the reference allele is prevailing, while a low value indicates that the contribution of the other alleles is significant.

Statistical analysis

Statistical analysis was performed as follows. First, one-sided *t*-test ($P \leq 0.02$) was used to identify differences among REA values in the two groups for every np. According to correction for multiple testing, we calculated the *q* values with FDR (false discovery rate) method by QVALUE software with a significant threshold of 0.05.[11] Second, principal component analysis (PCA) was applied on a restricted set of nps ($n = 34$) with the lowest *P*-values ($P < 0.01$) between AD and controls.[12] The eigenvalues >1.0 were retained in the analysis and a Varimax orthogonal rotation was used to obtain the factors. To interpret the results from factor analysis, the pattern of the factor loading was examined to determine which original variables represented primary constituents of each factor. An absolute loading value ≥0.85 was used to select REAs useful for cluster analysis. Third, receiver operating characteristic (ROC) curves were calculated to estimate the area under the curve (AUC) for selected REAs with $P < 0.01$ in relation to disease. AUC values ≥0.80 were taken into account for cluster analysis. Finally, hierarchical clustering was performed based on Euclidean distance between AD and control samples with complete linkage method. The linear discriminant analysis was adopted to verify the "goodness" of the classification. Contingency analyses were applied to examine frequency distributions.

Results

We compared REA values for each site between AD patients and controls and identified 270 nps with $P \leq 0.02$ by *t*-test (Table S1). All the 270 nps had a *q* value <0.05 and therefore were considered truly significant. These sites, numbered according to GenBank posi-

tion and ID number, were not associated with homo- or heteroplasmic mutations, with the exception of three cases, each of which regarded only one subject: nps 4762 (ID 49, heteroplasmic mutation in controls), 5489 (ID 93, heteroplasmic mutation in AD), and 13035 (ID 109, homoplasmic mutation in AD). It should be evidenced that GSEQ 4.1 software assigns heteroplasmy calls when there are two different bases with fluorescence intensity values close to each other, resembling the diploid condition with about 50% heteroplasmy.[9,13] The most remarkable result was that all but one (np 4762, ID 49) of the 270 nps displayed the highest value for controls (Figure S1). This means that controls had a notably higher fidelity to reference base than AD patients, which consequently showed a significant degree of allelic shift. Table S1 shows the amino acid changes when the reference base is substituted by one of the other bases and substitutions resulting in nonsynonymous mutations were significantly more abundant than the synonymous ones (χ^2 test; $P < 0.001$). The 270 nps did not belong to a particular gene, suggesting random nucleotide variation (Table 2). Table S2 reports the six factors identified by PCA applied to the 34 most significant nps ($P < 0.01$). Factorial weights ≥0.85 were selected in order to identify REAs that had a very high correlation with factors. Thus, nps 2620, 1414, 6226, and 13016 were considered important for discriminating healthy subjects and AD patients. Diagnostic accuracy of REA values was evaluated with ROC curve analyses (Table S3). When comparing AD versus controls, np 16195 showed the highest value of AUC (0.804). Hierarchical cluster analysis including nps 2620 (16S), 1414 (12S), 6226 (COI), 13016 (ND5), and 16195 (D-Loop) showed that groups were related to the selected REAs and the percentage of cases correctly classified was 80% (Fig. 1). There was no prevalence of a specific gene or a functional gene class in the selected nps (Monte Carlo analysis; $P = 1.000$). We calculated the percentage of heteroplasmy for these nps according to Coon et al.,[10] and found that these values were between 6.32% and 15.9%, corresponding to low-level heteroplasmy. Higher heteroplasmy percentage values were found in AD subjects compared to controls, for each discriminant site, but the comparisons resulted significant only for 16195 (one-sided *t*-test; $P < 0.05$). Data discussed here were deposited in NCBI's Gene Expression Omnibus,[14] and are accessible through GEO Series accession number GSE49160 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE49160).

Table 2. Nps for each gene counted in the 270 REAs with $P < 0.02$ (*t*-test).

Genes	nps	*P* value
D-loop	26	0.854
tRNA phenylalanine	1	1.000
12S	17	1.000
16S	39	0.152
tRNA leucine 1	6	0.128
ND1	14	0.579
tRNA methionine	1	1.000
ND2	25	0.346
tRNA asparagine	3	0.616
tRNA cysteine	1	1.000
COI	26	1.000
tRNA serine 1	1	1.000
tRNA aspartic acid	3	0.616
COII	5	0.115
Non coding	3	0.247
tRNA lysine	1	1.000
ATPase8	3	1.000
ATPase6	4	0.073
COIII	11	0.539
ND3	5	1.000
ND4	18	0.420
tRNA histidine	1	1.000
ND5	31	0.893
ND6	7	0.805
tRNA glutamic acid	1	1.000
Cytb	17	0.610

Statistical analysis shows that these nps are randomly distributed among genes (χ^2 test).

Discussion

Array-based analysis of mtDNA showed that AD whole blood is associated with an increased allelic shift compared to age-matched controls, and that some nps discriminate between the groups with 80% of cases correctly classified. Affymetrix MitoChip v2.0 provides the advantage of performing quantitative estimation for each of the four possible alleles in every np while avoiding interference of background nonspecific signal by setting specific conditions in GSEQ 4.1 algorithm. The sites with increased allelic shift in AD were not associated with homo- or heteroplasmic mutations, and conceivably their allelic imbalance has a limited phenotypic effect. However, many nucleotide substitutions alter protein structure and possibly function, and the aggregate burden of all variations may impact on mitochondria and cell oxidative capacity until exceeding the minimum threshold for a specific cell to function.[15] Indeed, results from AD cybrid model experiments indicate that, by introducing mitochondria from an AD donor into a clonal host cell depleted of endogenous mtDNA, an AD cybrid is obtained that exhibits reduced cytochrome c oxidase (COX) activity, increased oxidative stress, increased intracellular and extracellular ß-amyloid levels, and cell death activation compared to control cybrids.[16,17] The nps included in hierarchical cluster analysis did not belong to

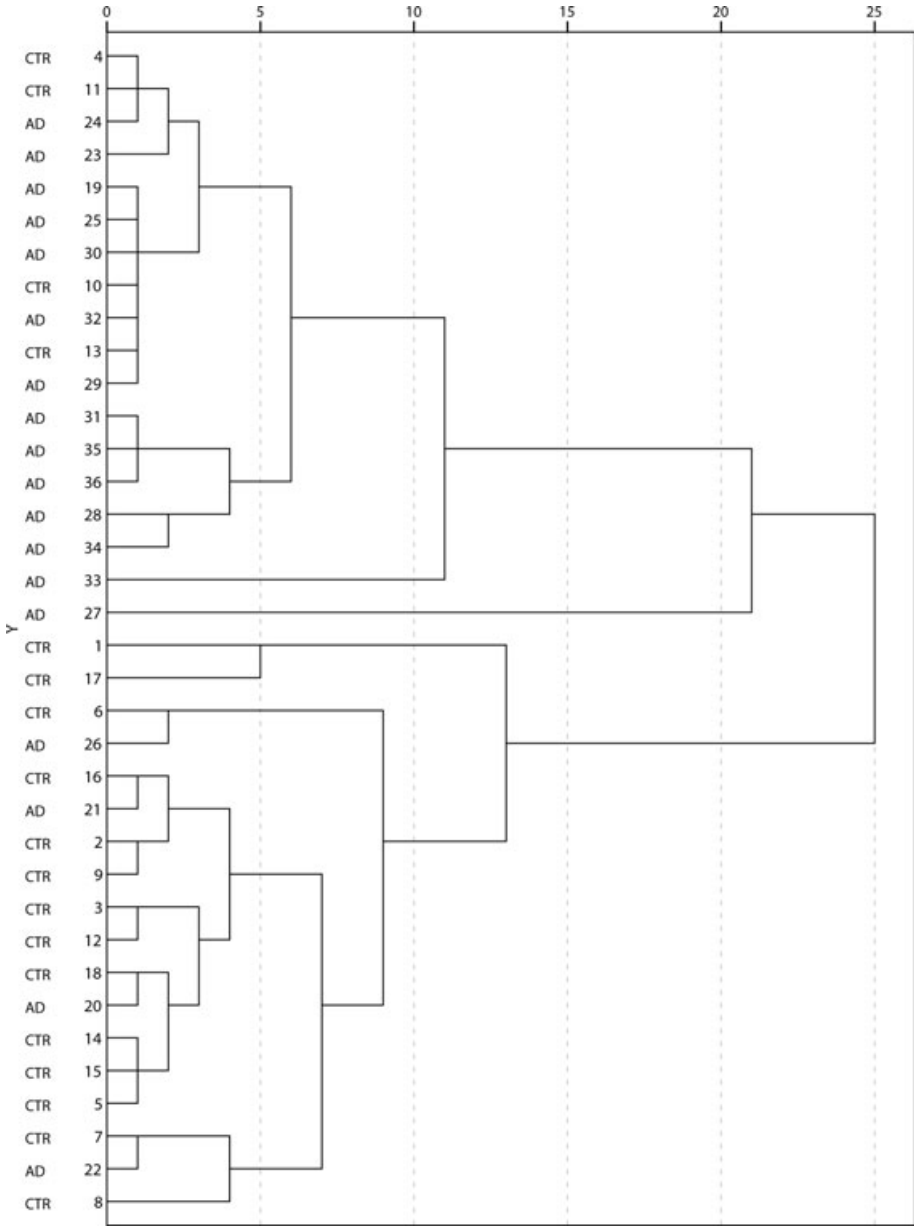

Figure 1. Hierarchical clustering of REA values for the selected nps (2620, 1414, 6226, 13016, and 16195) well divided the AD patients from control subjects. The discriminant analysis showed that the percentage of cases correctly classified was 80% (n = 36 individuals). Applying bootstrap method with 10,000 resampling we found that the 99% CI of percentage of correct classification was between 72% and 94%. We found in the dendrogram a percentage of subjects correctly classified corresponding to 78% that is within the CI defined by bootstrap. CTR, control group; AD, Alzheimer's disease patients.

a specific gene or a functional class of genes, they were selected for the best separation of the two groups of subjects. This apparent lack of spatial and functional linkage suggests that the allelic shift could be completely random and unspecific. Unlike homoplasmic and many heteroplasmic mutations that are inherited or acquired at early stages of development, these allelic alterations increases probably throughout life[18] and could eventually increase heteroplasmy levels. Engineered mice, heteroplasmic for NZB and 129S6 mtDNAs, showed impaired spatial memory retention and cognitive impairment, even at low-level heteroplasmy of NZB and 129S6 mtDNA genomes.[19] Our results support the hypothesis that mtDNA could accumulate mutations as a result of damage from reactive oxygen species, determining an impairment of oxidative phosphorylation and neurodegeneration. Actually,

changes of oxidative markers such as 8-hydroxyguanosine, lipid peroxidation, nitration, and nucleic acid oxidation generally precede occurrence of Aß aggregates and hyperphosphorylated tau protein, considered the hallmarks of AD.[20,21] These observations strengthen the notion that mitochondria are a potential target for preventive intervention in AD.

Acknowledgments

Supported by grant RF-INR-2006-346774 from Italian Ministry of Health, and revenues of the "5 per Mille" donations to INRCA. We thank Moreno Solazzi for technical support, and Dr Marco Catalano (Università Vita e Salute, Ospedale San Raffaele, Milan, Italy) for useful suggestions on a previous version of the paper.

Conflict of Interest

None declared.

References

1. Smigrodzki RM, Khan SM. Mitochondrial microheteroplasmy and a theory of aging and age-related disease. Rejuvenation Res 2005;8:172–198.
2. Coskun P, Wyrembak J, Schriner SE, et al. A mitochondrial etiology of Alzheimer and Parkinson disease. Biochim Biophys Acta 2012;1820:553–564.
3. Payne BA, Wilson IJ, Yu-Wai-Man P, et al. Universal heteroplasmy of human mitochondrial DNA. Hum Mol Genet 2013;22:384–390.
4. Hudson G, Sims R, Harold D, et al. No consistent evidence for association between mtDNA variants and Alzheimer disease. Neurology 2012;78:1038–1042.
5. Lakatos A, Derbeneva O, Younes D, et al. Association between mitochondrial DNA variations and Alzheimer's disease in the ADNI cohort. Neurobiol Aging 2010;31:1355–1363.
6. Coskun PE, Beal MF, Wallace DC. Alzheimer's brains harbor somatic mtDNA control-region mutations that suppress mitochondrial transcription and replication. Proc Natl Acad Sci USA 2004;101:10726–10731.
7. Lin MT, Simon DK, Ahn CH, et al. High aggregate burden of somatic mtDNA point mutations in aging and Alzheimer's disease brain. Hum Mol Genet 2002;11:133–145.
8. Maitra A, Cohen Y, Gillespie SE, et al. The Human MitoChip: a high-throughput sequencing microarray for mitochondrial mutation detection. Genome Res 2004;14:812–819.
9. Xie HM, Perin JC, Schurr TG, et al. Mitochondrial genome sequence analysis: a custom bioinformatics pipeline substantially improves Affymetrix MitoChip v2.0 call rate and accuracy. BMC Bioinformatics 2011;12:402.
10. Coon KD, Valla J, Szelinger S, et al. Quantitation of heteroplasmy of mtDNA sequence variants identified in a population of AD patients and controls by array-based resequencing. Mitochondrion 2006;6:194–210.
11. Storey JD. A direct approach to false discovery rate. J R Stat Soc Series B 2002;64:479–498.
12. Tollervey JR, Wang Z, Hortobágyi T, et al. Analysis of alternative splicing associated with aging and neurodegeneration in the human brain. Genome Res 2011;21:1572–1582.
13. Van Eijsden RG, Gerards M, Eijssen LM, et al. Chip-based mtDNA mutation screening enables fast and reliable genetic diagnosis of OXPHOS patients. Genet Med 2006;8:620–627.
14. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 2002;30:207–210.
15. Coskun PE, Wyrembak J, Derbereva O, et al. Systemic mitochondrial dysfunction and the etiology of Alzheimer's disease and down syndrome dementia. J Alzheimers Dis 2010;20(Suppl. 2):S293–S310.
16. Khan SM, Cassarino DS, Abramova NN, et al. Alzheimer's disease cybrids replicate ß-amyloid abnormalities through cell death pathways. Ann Neurol 2000;48:148–155.
17. Onyango I, Khan S, Miller B, Swerdlow R, et al. Mitochondrial genomic contribution to mitochondrial dysfunction in Alzheimer's disease. J Alzheimers Dis 2006;9:183–193.
18. Schon EA, DiMauro S, Hirano M. Human mitochondrial DNA: roles of inherited and somatic mutations. Nat Rev Genet 2012;13:878–890.
19. Sharpley MS, Marciniak C, Eckel-Mahan K, et al. Heteroplasmy of mouse mtDNA is genetically unstable and results in altered behavior and cognition. Cell 2012;151:333–343.
20. Nunomura A, Perry G, Aliev G, et al. Oxidative damage is the earliest event in Alzheimer disease. J Neuropathol Exp Neurol 2001;60:759–767.
21. Praticò D, Uryu K, Leight S, et al. Increased lipid peroxidation precedes amyloid plaque formation in an animal model of Alzheimer amyloidosis. J Neurosci 2001;21:4183–4187.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Line graphs of Alzheimer's disease patient (AD) and control (CTR) REA values for the nps reported in Table S1. Nps are represented with their corresponding ID. (A) Nps from 1 to 135 are represented. REAs for controls have the highest value except for np 4762 corre-

sponding to ID 49. (B) Nps from 136 to 270 show the highest REA value for controls.

Table S1. Nps corresponding to significant comparisons of REA values between AD and controls with $P \leq 0.02$ (*t*-test) sorted in ascending order of *P*. The reference base, the gene involved, as well as the possible amino acid change, are also indicated. Amino acid substitutions have been defined by MitoTool software.

Table S2. Factor loadings, identified by the Principal Component Analysis (PCA), of REA values corresponding to the 34 most significant nps ($P < 0.01$). The percentage of explained variance was 32.6 for factor 1, 15.6 for factor 2, 12.5 for factor 3, 8.1 for factor 4, 7.6 for factor 5 and 6.3 for factor 6. In bold the nps selected for cluster analysis.

Table S3. ROC curve analysis of REA values corresponding to the 34 most significant nps ($P < 0.01$). In bold the nps selected for cluster analysis.

