J Bioenerg Biomembr (2009) 41:407–410
DOI 10.1007/s10863-009-9240-8

Mitochondrial haplogroups associated with Japanese Alzheimer's patients

Shigeru Takasaki

Published online: 1 October 2009

Abstract The relationships between Japanese Alzheimer's disease (AD) patients and their mitochondrial single nucleotide polymorphism (mtSNP) frequencies at individual mtDNA positions of the entire mitochondrial genome are described using the radial basis function (RBF) network and the modified method. Japanese AD patients are associated with the haplogroups G2a, B4c1, and N9b1. In addition, to compare mitochondrial haplogroups of the AD patients with those of other classes of Japanese people, the relationships between four classes of Japanese people (i.e., Japanese centenarians, Parkinson's disease (PD) patients, type 2 diabetic (T2D) patients, and non-obese young males) and their mtSNPs are also described. The four classes of people are associated with following haplogroups: Japanese centenarians—M7b2, D4b2a, and B5b; Japanese PD patients—M7b2, B4e, and B5b; Japanese T2D patients—B5b, M8a1, G, D4, and F1; and Japanese healthy non-obese young males—D4g and D4b1b. The haplogroups of the AD patients are therefore different from those of the other four classes of Japanese people. As the analysis method described in this article can predict a person's mtSNP constitution and the probabilities of becoming an AD patient, centenarian, PD patient, or T2D patient, it may be useful in initial diagnosis of various diseases.

Keywords mtSNPs · Mitochondrial haplogroups · Alzheimer's disease · Centenarians · Parkinson's disease · Radial basis function (RBF)

S. Takasaki (✉)
Toyo University,
Izumino 1-1-1, Ora-gun Itakuracho,
Gunma 374-0193, Japan
e-mail: s_takasaki@toyonet.toyo.ac.jp

Introduction

Mitochondria are essential cytoplasmic organelles, generating cellular energy in the form of adenosine triphosphate by oxidative phosphorylation. Most cells contain hundreds of mitochondria and individual mitochondria have a few mitochondrial DNA (mtDNA) copies, so each cell contains several thousands of mtDNA copies. mtDNA has a very high mutation rate. When a mutation occurs, a cell initially contains a mixture of wild-type and mutant mtDNAs, a situation known as heteroplasmy. As the percentage of mutant mtDNAs increases, the cellar energy capacity declines until it falls below the bioenergetic threshold, the minimum energy output necessary for a cell or tissue to function normally. Beyond this point, disease symptoms appear and become progressively worse. Mitochondrial defects have been implicated in a wide variety of diseases, such as Alzheimer's disease (AD), Parkinson's disease (PD), cancer, etc. It is clear that mitochondrial diseases encompass an extraordinary assemblage of clinical problems, usually involving tissues that require large amounts of energy, such as heart, muscle, and renal and endocrine tissues (Wallace 1999; Vila and Przedborski 2003; Taylor and Turnbull 2005).

mtDNA mutations have been reported to be related to aging and a wide variety of diseases (e.g., AD, PD, type 2 diabetic disease, cancer, etc.) (Lin et al. 1992; Schoffner et al. 1993; Kosel et al. 1994; Mayr-Wohlfart et al. 1996; Schnopp et al. 1996; Simon et al. 2000; Tanaka et al. 2002; Dawson and Dawson 2003; Ross et al. 2003; Lustbader et al. 2004; Niemi et al. 2005; Alexe et al. 2007; Fuku et al. 2007; Chinnery et al. 2008; Kim et al. 2008; Maruszak et al. 2008; Feder et al. 2008). There are, however, few reports regarding the relationships between all mtDNA mutations and disease patients or centenarians. In addition, previous

works focused on amino acid replacements by mtDNA mutations. Although mitochondrial functions can be affected directly by amino acid replacements, they can also be affected indirectly by mutations in the mtDNA control regions. It is therefore important to examine the relationships between all mtDNA mutations and disease patients or centenarians.

In this article, the relationships between Japanese AD patients and their mitochondrial single nucleotide polymorphism (mtSNP) frequencies at individual mtDNA positions of the entire mitochondrial genome are described using a radial basis function (RBF) network (Poggio and Girosi 1990; Wu and McLarty 2000) and a modified version of its classification method (Takasaki 2009). In addition, the relationships between the haplogroups of the AD patients and those of the other four classes of people are also described using the same analysis method. The results described here are quite different from those reported previously (Saxena et al. 2006; Alexe et al. 2007; Fuku et al. 2007; Bilal et al. 2008).

mtSNPs of Japanese people

The following was used in this article. Website: http://mtsnp.tmig.or.jp/mtsnp; mtSNPs used: those in 96 Japanese AD patients, 96 Japanese centenarians, 96 Japanese PD patients, 96 Japanese type 2 diabetic (T2D) patients, and 96 Japanese healthy non-obese young males; tissue: blood; sex: male and female except for healthy non-obese young males; origin: Asian (Tanaka et al. 2004).

Method of mtSNP classification for Japanese AD patients

The RBF network is a type of artificial network for applications to problems of supervised learning, such as regression, classification and time series prediction. In the supervised learning, the function is learned from the examples (training set) that a teacher supplies. The training set contains elements, which consist of paired

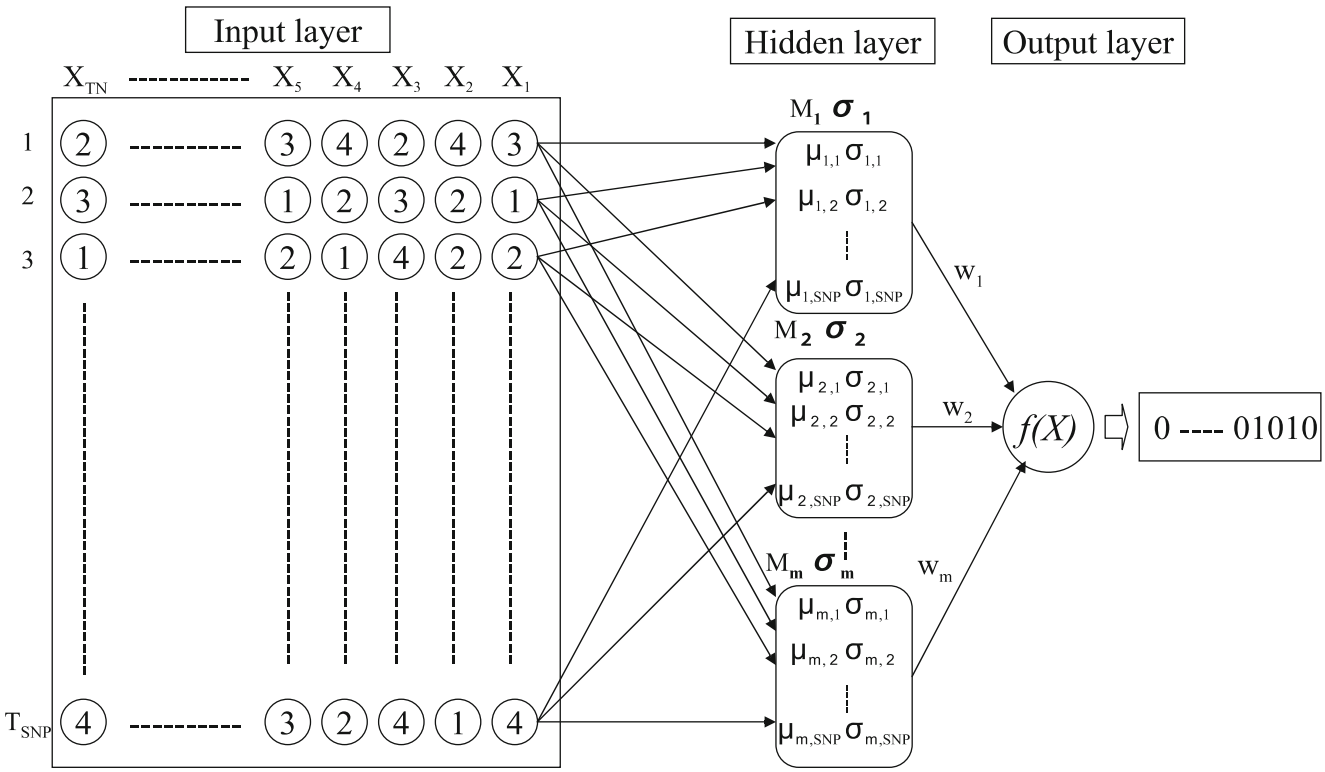

Fig. 1 RBF network representation of the relationships between individual mtSNPs and the AD patients. The input layer is the set of mtSNP sequences represented numerically (A, G, C, and T are converted to 1, 2, 3, and 4). The hidden layer classifies the input vectors into several clusters depending on the similarities of individual input vectors. The output layer is determined depending on which analysis is carried out. In the case of AD patients, 1 corresponds to AD patients and 0 corresponds to the other four classes of people. In the case of centenarians, 1 corresponds to centenarians and 0 corresponds to the other four classes of people. The PD patients, T2D patients, and healthy non-obese young males are also carried out in a similar way. X_i : i-th input vector, TN : maximum number of vectors (in this example, TN=320 (64×5)), T_{SNP} : maximum number of mtSNPs (in this example, T_{SNP} =562), M_m: the location vector, m: the number of basis functions, μ: basis function, σ: standard deviation, w_i: i-th weighting variable, $f(X)$: weighted sum function

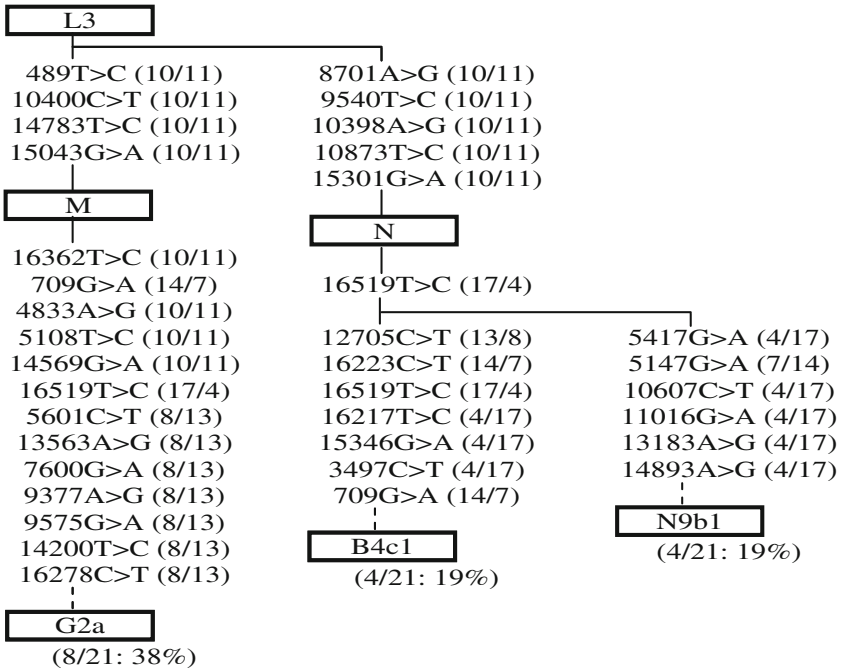

Fig. 2 Associations between Asian/Japanese haplogroups and mtSNPs of the AD patients. Associations between Asian/Japanese haplogroups and mtSNPs are described based on the phylogenetic tree for macrohaplogroups M and N described in Tanaka et al. (2004). The locus of mtDNA polymorphism (*mmm*), normal nucleotide (rCRS) at the position *mmm* (N_N), mtDNA mutation at the same position (N_M), the number of the mtDNA mutations at *mmm* in individual highest clusters (*Y*), and the number of the normal nucleotides at *mmm* in individual highest clusters (*X*) are expressed as $mmmN_N>N_M(Y/X)$. For example, 16362T>C (10/11) indicates "16362"—the locus of mtDNA, "T"—the normal nucleotide at the position 16362, "C"—the mtDNA mutation at the same position, "10"—the number of the mtDNA mutations, and "11"—the number of the normal nucleotides in the highest cluster

values of the independent (input) variable and dependent (output) variable.

As the RBF network shown in Fig. 1 is learned from the training set (correct and incorrect data), mtSNPs of the AD patients are regarded as correct and the other four classes of people, i.e., Japanese centenarians, PD patients, T2D patients, and healthy non-obese young males, are regarded as incorrect. Similarly, in the case of mtSNP classification for the centenarians, mtSNPs of the centenarians are regarded as correct (right) and the other four classes of people are regarded as incorrect. The mtSNP classifications for the PD patients, T2D patients and healthy non-obese young males are also carried out in the same way as those for the AD patients and centenarians (Fig. 1).

The mitochondrial genome sequences of the AD patients are partitioned into two sets, one of training data and the other of validation data. The training data are the sequences for individual 64 of the AD patients, and the validation data are the sequences for the other individual 32 of the AD patients. The processes of the classifications are carried out in two phases: training and validation. The detailed steps are described by Takasaki et al. (2006).

Modified classification method based on probabilities predicted by the RBF network

Since a RBF network can predict the probabilities that persons with certain mtSNPs belong to certain classes (e.g., AD patients, centenarians, PD patients, T2D patients, or healthy non-obese young males), these predicted probabilities are used to identify mtSNP features. By examining the relationships between individual mtSNPs and the persons with high predicted probabilities of belonging to one of these classes, we are able to identify other mtSNPs useful for distinguishing between the members in different classes. The modified classification method based on the probabilities predicted by the RBF network is carried out in the following way (Takasaki 2009).

1) Select the target of analysis (i.e., AD patients, centenarians, PD patients, T2D patients, or healthy non-obese young males).
2) Rank individuals according to their predicted probabilities of belonging to the target class.
3) Select individuals whose probabilities are greater than a certain value or the desired number of individuals from the top and set them as a modified cluster.

Associations between Asian/Japanese haplogroups and mtSNPs of the AD patients

The mtSNP classifications for the AD patients were executed by the above described method. As a result, seven mtSNP clusters were obtained. The average predicted probabilities of these clusters for becoming the AD patients were respectively 62.1%, 59.4%, 35.1%, 28.6%, 24%, 10.5% and 0%. Then individuals whose probabilities were greater than 70% for the AD patients were selected using the modified classification method and nucleotide distributions of the selected 21 AD patients were examined at individual mtDNA positions. After that, the relationships between Asian/Japanese haplogroups and mtSNPs for the AD patients were examined (Herrnstad et al. 2002; Kong et al. 2003; Tanaka et al. 2004). The associations between the haplogroups and mtSNPs for the AD patients are shown in Fig. 2. The features of associations for the AD patients were L3-M-G2a (38%), L3-N-B4c1 (19%), and N9b1 (19%).

To compare the mitochondrial haplogroups of the AD patients with those of other classes of Japanese people, the relationships between four classes of Japanese people (i.e., Japanese centenarians, Parkinson's disease (PD) patients, type 2 diabetic (T2D) patients, and non-obese young males) and their mtSNPs were also examined using the same modified method. The four classes of people were associated with the following haplogroups: Japanese centenarians—L3-M-M7b2 (40%), L3-M-D-D4b2a (27%), and L3-N-B5b (20%); the PD patients—L3-M-M7b2 (50%), L3-N-B4e (20%), and B5b (20%); the T2D patients—L3-M-D-D4 (10%), L3-M-M8a1 (10%), G (10%), L3-N-B5b (30%), and F1 (10%); and the healthy non-obese young males—L3-M-D-D4g (38%), and D4b1b (38%).

From the points of common haplogroups among individual classes of people, the haplogroup M7b2 was common in the centenarians and PD patients, and B5b was common in the centenarians, PD patients and T2D patients. The haplogroups of the AD patients are therefore different from those of other four classes of Japanese people. Furthermore, although the haplogroups of the centenarians and T2D patients have been reported (Alexe et al. 2007; Fuku et al. 2007; Bilal et al. 2008), there are few reports for the AD patients (Takasaki 2009). The results are therefore considered as new findings.

Differences between statistical technique and the modified RBF method

Although the haplogroups of the AD patients were obtained by the modified RBF method, there are clear differences between the previously reported statistical technique and the method described here. In the statistical technique, odds ratios or relative risks are analyzed based on the relative relationships between target and control data at each polymorphic mtDNA locus. On the other hand, in the modified RBF method, clusters indicating predicted probabilities are examined on the basis of the RBF using correct and incorrect data for the entire polymorphic mtDNA loci. The statistical technique determines characteristics of haplogroups using independent mtDNA polymorphisms that indicate high odds ratios, whereas the modified method determines them by checking individuals with high predicted probabilities. This means that the statistical technique uses the results of independent mutation positions, whereas the modified RBF method uses the results of entire mutation positions. As there are the differences between both methods, which method is better depends on future research. Furthermore, the method described here may have possibilities for use in the initial diagnosis of various diseases on the basis of the individual predicted probabilities.

Acknowledgements I thank Dr. Y. Kawamura for his encouragement and technical support.

References

Alexe G, Fuku N, Bilal E et al (2007) Hum Genet 121:347–356
Bilal E, Rabadan R, Alexe G et al (2008) PLoS ONE 3(6):e2421
Chinnery PF, Mowbray C, Patel SK et al (2008) J Med Genet 44(6):e80
Dawson TM, Dawson VL (2003) Science 302:819–822
Feder J, Blech I, Ovadia O et al (2008) BMC Genomics 9:198
Fuku N, Park KS, Yamada Y et al (2007) Am J Hum Genet 80:407–415
Herrnstad C, Elson JL, Fahy E et al (2002) Am J Hum Genet 70:1152–1171
Kim W, Yoo TK, Shin DJ et al (2008) PLoS ONE 3(5):e2211
Kong QP, Yao YG, Sun C et al (2003) Am J Hum Genet 73:671–676
Kosel S, Egensperger R et al (1994) Biochem Biophys Res Commun 203:745–749
Lin F, Lin R et al (1992) Biochem Biophys Res Commun 182:238–246
Lustbader JW, Cirilli M, Lin C et al (2004) Science 304:448–452
Maruszak A, Canter JA et al (2008) Neurobiol Aging doi:10.1016/j.neurobiolaging.2008.01.004
Mayr-Wohlfart U, Paulus C, Rodel G (1996) Acta Neurol Scand 94:167–171
Niemi AK, Moilanen JS et al (2005) Eur J Hum Genet 13:166–170
Poggio T, Girosi F (1990) Proc IEEE 78:1481–1497
Ross OA, MaCormack R et al (2003) Exp Gerontol 38:397–405
Saxena R, de Bakker PI et al (2006) Am J Hum Genet 79:54–61
Schnopp NM, Kosel S et al (1996) Clin Neuropathol 15:348–352
Schoffner JM, Brown MD, Torroni A et al (1993) Genomics 17:171–184
Simon DK, Mayeux R et al (2000) Neurology 54:703–709
Takasaki S (2009) J Genetics Genomics 36:425–434
Takasaki S, Kawamura Y, Konagaya A (2006) BMC Bioinformatics 7 (Suppl 5):S22
Tanaka M, Fuku N, Takeyasu T et al (2002) J Neurosci Res 70:347–355
Tanaka M, Cabrera VM, Gonzalez AM et al (2004) Genome Res 14:1832–1850
Taylor RW, Turnbull DM (2005) Nat Rev Genet 6:389–402
Vila M, Przedborski S (2003) Nat. Rev. Neurosci. 4:1–11
Wallace DC (1999) Science 283:1482–1488
Wu CH, McLarty JW (2000) Elsevier Science, New York