

Control region mtDNA variants: Longevity, climatic adaptation, and a forensic conundrum

Pinar E. Coskun, Eduardo Ruiz-Pesini, and Douglas C. Wallace[†]

Center for Molecular and Mitochondrial Medicine and Genetics, University of California, Irvine, CA 92692-3940

Mitochondrial oxidative phosphorylation (OXPHOS) has been observed to decline with age in association with the accumulation of somatic mtDNA rearrangement (1) and base substitution (2–4) mutations. These observations have led to the theory that the mtDNA is the aging clock. To this aging story was added a new twist in a recent issue of PNAS, an mtDNA mutation that promotes longevity (5).

MtDNA rearrangement mutations accumulate exponentially with age in postmitotic tissues, with the inflection point generally occurring about age 45. Although any one mtDNA rearrangement is rare in a particular tissue, surveys of the array of mtDNA rearrangement mutations using long extension-PCR have revealed that a broad spectrum of different mtDNA rearrangements accumulate in each tissue with age (1).

Certain control region (CR) mtDNA mutants have also been reported to accumulate with age in specific tissues (4), and evidence of heteroplasmic (mixtures of mutant and normal mtDNAs) CR mutants has been reported in a variety of contexts including human brain tissue (6), different hairs of single individuals (7), tissues of the Grand Duke of the Romanov family (8), and cancer cells (9, 10).

A surprising aspect of these age-related CR mtDNA mutants is their tissue specificity. One somatic CR mtDNA mutation, a T to G transversion occurring at nucleotide position (np) 414 (T414G), is generally found in skin fibroblast cultures from individuals over 60 years of age with the level of mutant mtDNAs exceeding 60% in some cases (Table 1). Two additional mtDNA CR mutations, A189G and T408A, have been reported to accumulate with age in skeletal muscle (3, 5), but unlike T414G these mutations have also been reported to be germ-line-transmitted polymorphisms associated with specific mtDNA haplogroups (groups of related mtDNA haplotypes) (Table 1). Thus, it appears that each tissue may accumulate its own unique somatic mtDNA CR mutations with age, but some of these same variants might also be inherited.

This story is further complicated in a recent issue of PNAS (5) by the report of a C150T variant that accumulates with age in skin fibroblasts, but is also present in the blood cell lymphocytes of centenar-

Table 1. Age-related heteroplasmic mtDNA CR variant reported as somatic mutations and germ-line variations in various haplogroups

Sequence variation	Germ line		Somatic tissue specificity
	Clusters	No. of cases (total 444)	
C150T	Lo (26, 27)	2	Leukocyte, fibroblast (5)
	L2 (24, 26–28)	9	
	L3 (26, †)	8	
	G (26)	1	
	M* (†)	2	
	D (29, 30, †)	3	
	N no R (26, 30, 31)	4	
	N1b (24)	1	
	U* (†)	1	
	U2 (28)	1	
	U3 (28)	2	
	U5 (26, 28, 32, †)	11	
	J (24, 32–34, †)	9	
	T (26, 28, †)	3	
	HV* (28)	1	
H (26, †)	2		
A189G	Lo (24, 26–28, 32, 35)	7	Skeletal muscle (2)
	U5 (†)	1	
	L3 (26)	1	
	W (24, 28, 32)	39	
	B (36)	2	
	J (32)	2	
	V (32)	1	
T408A	L3 (26)	1	Skeletal muscle (3)
	M* (29)	1	
T414G	—	—	Fibroblast (4), skeletal muscle (2)

References are given in parentheses.
[†]Unpublished data.

ians and twins (Table 1). The association of the C150T mutation with centenarians and twins suggested to the authors that the C150T imparts resistance to stress and thus promotes longevity. Yet, this mutation has also been reported to be an inherited polymorphism in some instances (Table 1).

The unusual behavior of the CR mtDNA mutations raises a myriad of questions. Why is it that mtDNA rearrangement mutations are ubiquitous in aging postmitotic tissues, but each individual rearrangement is rare; whereas the age-related CR mtDNA base substitution mutations are tissue-specific, but each mutation can be present at high levels

throughout a tissue? How can most somatic CR mtDNA mutations be deleterious and contribute to aging in their tissue of origin, whereas the C150T mutation may be advantageous when occurring in lymphocytes? Why is it that certain CR somatic mutations are thought to contribute to aging, and yet the same variants have been reported to be inherited in certain mtDNA haplogroups? Finally, how can one distinguish between germ-line and somatic mutations in forensic specimens, and what does this ambiguity imply about the reliability of CR sequences in individual identification?

Cell-Specific Amplification of Rearrangement Mutations. The low level of each mtDNA rearrangement in a tissue has been interpreted as each individual mutation arising *de novo* within a single cell or a region of skeletal muscle fiber. Once the mutation appears, it is selectively amplified until it comes to be the predominate mtDNA in the cell or fiber domain of origin. This idea has been substantiated by *in situ* hybridization to mtDNAs within individual muscle fibers from the elderly which revealed the clonal amplification of different mtDNA rearrangements in different regions of the same muscle fiber (11), and by studying cultured cardiomyocytes which revealed that each cardiomyocyte harbored only one mtDNA rearrangement that was clonally expanded within that cell (12, 13). Thus, each somatic mtDNA rearrangement reaches high levels of heteroplasmy within its cell of origin, however, when a tissue is homogenized, all of the rearranged mtDNA are mixed together diluting each of the molecules to low average levels.

But why would the deleterious mtDNA rearrangement mutations be selectively amplified? The human mtDNA encodes 13 polypeptides that are essential for the mitochondrial energy generating system, OXPHOS, plus the 12S and 16S rRNA genes and 22 tRNA genes necessary for their expression. Thus, any mutation in the mtDNA coding region will alter mitochondrial energy production. The mitochondria use OXPHOS to generate most of the cellular ATP, but as a toxic by-

See companion article on page 1116 in issue 3 of volume 100.

[†]To whom correspondence should be addressed. E-mail: dwallace@uci.edu.

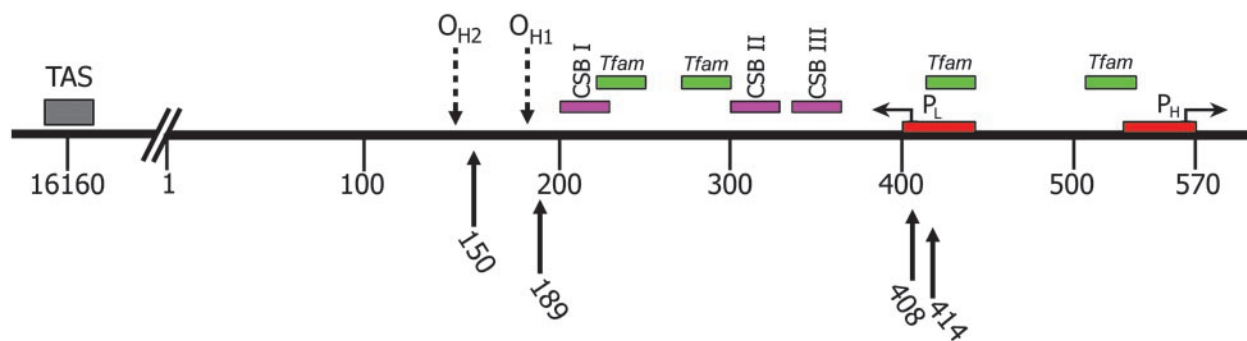


Fig. 1. Schematic diagram of the mtDNA-CR. The functional sequences are indicated with boxes. Dashed arrows point to H-strand origin of replication 1 and 2 (O_{H1} and O_{H2}). The arrows pointing left and right show transcription orientation of H- and L-strand promoters (P_H and P_L). Arrows below the line indicate the heteroplasmic nucleotide positions associated with aging. TAS, termination association sequence; CSB I, II, III, conserved sequence block I, II, and III; *Tfam*, *Tfam* binding sites.

product they produce most of the endogenous reactive oxygen species (ROS). The ROS, in turn, can damage the OXPHOS enzymes and the mtDNA, eroding mitochondrial function. When mitochondrial energy production gets too low and/or mitochondrial ROS damage becomes too high, the mitochondrial permeability transition pore (mtPTP) is activated and the cell with its damaged mitochondria is removed by apoptosis (14).

So why might somatic mtDNA rearrangements be selectively amplified? Apparently, the only way that a cell can cope with chronic energy deficiency is to make more mitochondria; but only those cells or muscle regions that have mutant mtDNAs have energy deficiencies. Hence, only the nuclei associated with the defective mitochondria are induced to make more mitochondria. Thus, the mutant mtDNAs are selectively amplified.

But what tells the nuclei to make more mitochondria? One possibility is that the nuclei monitor the redox state of the cell and thus the mitochondria. The mitochondria oxidize NADH derived from carbohydrates and fats with O_2 to generate H_2O and trap the energy released in ATP. Thus, a defect in OXPHOS will reduce ATP production and increase the NADH/NAD⁺ ratio making the cell more reduced. The more reduced environment of the cell could be monitored by a ubiquitous transcription factor, such as the REBOX-binding factor. This factor only binds to its 5' *cis*-elements in nDNA bioenergetic genes when in a reducing environment (15). Thus, when the defective mitochondria inadequately oxidize the NADH the increasingly reduced environment would activate the REBOX-binding factor and this in turn would induce the transcription of nuclear DNA (nDNA) proteins involved in the bioenergetics and biogenesis of mitochondria. Only the nuclei associated with mutant mtDNA would be activated and hence the mutant mtDNAs would be preferentially amplified.

So why don't rearranged mtDNAs come to predominate in elderly tissues? Because once the rearranged mtDNA level becomes too high within a cell, the reduced energy output and increased oxidative stress activates the mtPPT, and the cell and its mutant mtDNAs are eliminated (14).

Tissue-Specific Amplification of CR Mutations. Unlike individual mtDNA rearrangements, specific CR mtDNA variants can come to represent the majority of the mtDNAs in a tissue. This means that most of the cells must contain high levels of the same mutant. How could this be? One possibility is that the CR base substitution mutation could arise by chance at various times during development. The earlier the mutation, the greater the percentage of cells in the tissue that would inherit the mutant mtDNA. If this particular CR mutant imparted a replicative advantage in a particular tissue, then the mutant mtDNAs would be amplified in all of the cells of that tissue. Thus, the more cells that inherited the mutant mtDNA, the greater the percentage of the mtDNAs in the tissue that would ultimately be mutant.

But why might specific CR mutants be amplified in specific cells? The CR is known to contain regulatory elements. Its 1,122 nps encompasses the H- and L-strand promoters (P_L , P_H), the origin of H-strand replication (O_H), three conserved sequence blocks (CSB I, II, III), and the termination associated sequences (TAS). Yet these elements represents only for a small proportion of the CR sequence in a genome notorious for its coding sequence economy (Fig. 1).

What then is the function of the rest of the CR sequence? This might be deduced from two observations. First, both P_L and P_H are associated with binding sites for the nDNA coded transcription factor *Tfam*. The T414G mutation occurs in close proximity to the *Tfam* binding site of the P_L and thus could perturb both

L-strand transcription and, because the L-strand acts as its primer, H-strand replication (4). Thus, CR sequences may bind nDNA-encoded regulatory factors. Second, two novel *cis*-regulatory elements, the OXBOX and the REBOX, are found 5' to multiple nDNA energy genes (15, 16). Moreover, homologous *cis*-elements are also observed in the CR of the mtDNA, and these mtDNA sequences can gel-shift the nDNA-encoded transcription factors. Whereas the REBOX-binding factor is ubiquitous, the OXBOX-binding factor is heart and muscle specific (15). Such *trans*-acting factors might coordinate nDNA and mtDNA transcription and replication. If so, there might be many such factors, some of which are tissue-specific. In this case, a chance CR mtDNA mutation in the *cis*-binding site of a tissue-specific regulatory factor would remove the tissue-specific regulation. If this factor imposed a negative control on mtDNA replication, then the mtDNA harboring the CR mutant would be selectively amplified in every cell of a tissue expressing that *trans*-factor.

This same logic could explain the excess of mtDNA CR mutants recently reported in cancer cells (9, 10). Loss of *cis*-binding sites that negatively regulated mtDNA replication could impart a significant growth advantage to a cancer cell.

CR mtDNA mutations do have the potential of being functionally important. Recently, a polymorphism in the TAS region at np 16,189 has been reported to correlate with an increased risk for type II diabetes (17).

So why don't the amplified CR mtDNA mutations induce apoptosis and remove the cells with the mutant mtDNAs? The answer probably lies with the fact that the CR mutants would not generate defective mtDNA gene products. Rather, they would alter the ratio of the mtDNA and nDNA gene products. This might erode the efficiency of a tissue without actually inhibiting OXPHOS.

Advantages of the C150T Mutation

How then could the C150T mutation have both a replicative advantage in skin fibroblasts and be protective to centenarians and twins in lymphocytes? The authors suggest that the C150T mutation alters the location of H-strand replication origin and this imparts a replicative advantage to the mtDNA.

But why might C150T be advantageous? One possibility is that it enhances the immune system in some way, possibly by slowing the turnover of memory T-cells. Alternatively, because most of the centenarians with the C150T mutation appear to have inherited it through the germ line, the C150T mutation might simply have been linked to other variants in the mtDNA that promote longevity.

Somatic Versus Germ-Line CR mtDNA Mutations

Previously, the authors of the C150T paper reported that in this same group of centenarians there was enrichment for the European haplogroup J (18). Since their report, haplogroup J has been associated with longevity in three other European studies (19–21). Furthermore, a subset of haplogroup M mtDNAs has also been associated with longevity in a Japanese study (22).

We have surveyed the mtDNA sequence of over 400 mtDNAs for the presence of the C150T mutation, and found that it was present in certain African-specific haplogroups (L0, L2, L3), European-specific haplogroups (NnoR, N1b, U*, U2, U3, U5, J, T, HV*, and H), and Asian-specific haplogroups (M* and D). The presence of C150T in European haplogroup J is particularly interesting. Possibly the centenarians with haplogroup J

mtDNAs are also the ones with the germ-line C150T variant (Table 1).

How might C150T and/or haplogroup J contribute to longevity? Human mtDNAs radiated out of Africa and around the world through successive transmission from mother to daughter, only changing by the accumulation of sequential mutations along radiating maternal lineage (23). Africans encompass the African-specific mtDNA haplogroups L0-L3, and the M and N lineages derived from L3. However, only M and N left Africa to colonize Eurasia. In Europe, lineage N radiated to give the plethora of European mtDNA haplogroups including J. In Asia, both M and N radiated into numerous mtDNA lineages, but only four of the southern Asian lineages (A, C, D, and G) came to predominate in northern Siberia. Only three of these (A, C, and D) crossed the Bering land bridge to populate the Americas.

The striking transitions in mtDNA types between Africa and Eurasia and between southern Asia and Siberia correlate with latitude and indicate that mtDNA diversity has been subjected to climatic selection (24). The mitochondria burn calories to make ATP to do work, but also to generate heat to maintain our body temperature. The balance between these two mutually exclusive processes is determined by the OXPHOS efficiency, which is regulated in large measure by the mtDNA ATP6 gene protein. Highly efficient OXPHOS generates primarily ATP with little waste heat. Less efficient OXPHOS generates more heat for the same amount of ATP. The former is preferable in the tropics, whereas the later would be critical for survival in the arctic. Interestingly, the amino acid sequence of the human mtDNA ATP6 protein is hy-

pervariable in the arctic, implying that mutations in this mtDNA gene have been important for human adaptation to extreme cold. A similar argument can be made for variation in the mtDNA cytochrome gene among the European mtDNA lineages (24).

A reduction in OXPHOS efficiency would also burn more calories. As a result, fewer reducing equivalents (electrons) from the diet would be available to make ROS. A life-long reduction in mitochondrial ROS stress would, in turn, decrease apoptosis and increase longevity. Hence, the C150T variant or a linked polymorphism in haplogroup J might have changed OXPHOS efficiency and thus ROS production, reducing oxidation stress and increasing longevity.

CR mtDNA Variants and Forensics

It is clear that the same CR mutations can occur either as tissue-specific somatic mutations or systemic germ-line mutations (Table 1). At the same time the mtDNA CR sequence has become increasingly popular as a forensic tool for individual identification, primarily because of its hypervariability. However, this application is predicated on the assumption that the mtDNA CR sequence variants of a person will be the same in all tissues of his or her body. But it is now clear that the CR sequences from two different tissues of the same person could harbor different somatic mutations, erroneously excluding a match. For example, the C150T mutation has already been observed in the heteroplasmic state in the blood samples of one Spanish subject (25). Clearly, more research is still required on tissue-specific somatic mtDNA mutations before CR sequences can be fully reliably for individual identification in forensics.

- Wallace, D. C. (1999) *Science* **283**, 1482–1488.
- Murdoch, D. G., Christacos, N. C. & Wallace, D. C. (2000) *Nucleic Acids Res.* **28**, 4350–4355.
- Wang, Y., Michikawa, Y., Mallidis, C., Bai, Y., Woodhouse, L., Yarasheski, K. E., Miller, C. A., Askanas, V., Engel, W. K., Bhasin, S. & Attardi, G. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 4022–4027.
- Michikawa, Y., Mazzuchelli, F., Bresolin, N., Scarfato, G. & Attardi, G. (1999) *Science* **286**, 774–779.
- Zhang, J., Asin-Cayuela, J., Fish, J., Michikawa, Y., Bonafé, M., Olivieri, F., Passarino, G., De Benedictis, G., Franceschi, C. & Attardi, G. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 1116–1121.
- Jazin, E. E., Cavellier, L., Eriksson, I., Orelund, L. & Gyllensten, U. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 12382–12387.
- Bendall, K. E., Macaulay, V. A. & Sykes, B. C. (1997) *Am. J. Hum. Genet.* **61**, 1303–1308.
- Ivanov, P. L., Wadhams, M. J., Roby, R. K., Holland, M. M., Weeden, V. W. & Parsons, T. J. (1996) *Nat. Genet.* **12**, 417–420.
- Polyak, K., Li, Y., Zhu, H., Lengauer, C., Willson, J. K., Markowitz, S. D., Trush, M. A., Kinzler, K. W. & Vogelstein, B. (1998) *Nat. Genet.* **20**, 291–293.
- Fliiss, M. S., Usadel, H., Caballero, O. L., Wu, L., Buta, M. R., Eleff, S. M., Jen, J. & Sidransky, D. (2000) *Science* **287**, 2017–2019.
- Muller-Hocker, J., Schneiderbanger, K., Stefani, F. H. & Kadenbach, B. (1992) *Mutat. Res.* **275**, 115–124.
- Muller-Hocker, J. (1992) *Brain Pathol.* **2**, 149–158.
- Muller-Hocker, J., Seibel, P., Schneiderbanger, K., Zietz, C., Obermaier-Kusser, B., Gerbitz, K. D. & Kadenbach, B. (1992) *Hum. Pathol.* **23**, 1431–1437.
- Kokoszka, J. E., Coskun, P., Esposito, L. A. & Wallace, D. C. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 2278–2283.
- Haraguchi, Y., Chung, A. B., Neill, S. & Wallace, D. C. (1994) *J. Biol. Chem.* **269**, 9330–9334.
- Chung, A. B., Stepien, G., Haraguchi, Y., Li, K. & Wallace, D. C. (1992) *J. Biol. Chem.* **267**, 21154–21161.
- Poulton, J., Luan, J., Macaulay, V., Hennings, S., Mitchell, J. & Wareham, N. J. (2002) *Hum. Mol. Genet.* **11**, 1581–1583.
- De Benedictis, G., Rose, G., Carrieri, G., De Luca, M., Falcone, E., Passarino, G., Bonafé, M., Monti, D., Baggio, G., Bertolini, S., et al. (1999) *FASEB J.* **13**, 1532–1536.
- Niemi, A. K., Hervonen, A., Hurme, M., Karhunen, P. J., Jylha, M. & Majamaa, K. (2003) *Hum. Genet.* **112**, 29–33.
- Rose, G., Passarino, G., Carrieri, G., Altomare, K., Greco, V., Bertolini, S., Bonafé, M., Franceschi, C. & De Benedictis, G. (2001) *Eur. J. Hum. Genet.* **9**, 701–707.
- Ross, O. A., McCormack, R., Curran, M. D., Duguid, R. A., Barnett, Y. A., Rea, I. M. & Middleton, D. (2001) *Exp. Gerontol.* **36**, 1161–1178.
- Tanaka, M., Gong, J. S., Zhang, J., Yoneda, M. & Yagi, K. (1998) *Lancet* **351**, 185–186.
- Wallace, D. C., Brown, M. D. & Lott, M. T. (1999) *Gene* **238**, 211–230.
- Mishmar, D., Ruiz-Pesini, E., Golik, P., Macaulay, V., Clark, A. G., Hosseini, S., Brandon, M., Easley, K., Chen, E., Brown, M. D., et al. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 171–176.
- Crespillo, M., Luque, J. A., Paredes, M., Fernandez, R., Ramirez, E. & Valverde, J. L. (2000) *Int. J. Legal Med.* **114**, 130–132.
- Ingman, M., Kaessmann, H., Paabo, S. & Gyllensten, U. (2000) *Nature* **408**, 708–713.
- Torrioni, A., Rengo, C., Guida, V., Cruciani, F., Sellitto, D., Coppa, A., Calderon, F. L., Simionati, B., Valle, G., Richards, M., Macaulay, V. & Scozzari, R. (2001) *Am. J. Hum. Genet.* **69**, 1348–1356.
- Maca-Meyer, N., Gonzalez, A. M., Larruga, J. M., Flores, C. & Cabrera, V. M. (2001) *BMC Genet.* **2**, 13.
- Kivisild, T., Tolk, H. V., Parik, J., Wang, Y., Papiha, S. S., Bandelt, H. J. & Villems, R. (2002) *Mol. Biol. Evol.* **19**, 1737–1751.
- Ozawa, T., Tanaka, M., Ino, H., Ohno, K., Sano, T., Wada, Y., Yoneda, M., Tanno, Y., Miyatake, T., Tanaka, T., et al. (1991) *Biochem. Biophys. Res. Commun.* **176**, 938–946.
- Ozawa, T. (1995) *Exp. Gerontol.* **30**, 269–290.
- Finnila, S., Lehtonen, M. S. & Majamaa, K. (2001) *Am. J. Hum. Genet.* **68**, 1475–1484.
- Brown, M. D., Zhadanov, S., Allen, J. C., Hosseini, S., Newman, N. J., Atamonov, V. V., Mikhailovskaya, I. E., Sukernik, R. I. & Wallace, D. C. (2001) *Hum. Genet.* **109**, 33–39.
- Brown, M. D., Starikovskaya, E., Derbeneva, O., Hosseini, S., Allen, J. C., Mikhailovskaya, I. E., Sukernik, R. I. & Wallace, D. C. (2002) *Hum. Genet.* **110**, 130–138.
- Horai, S., Hayasaka, K., Kondo, R., Tsugane, K. & Takahata, N. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 532–536.
- Yamasoba, T., Goto, Y., Oka, Y., Nishino, I., Tsukuda, K. & Nonaka, I. (2002) *Neuromuscul. Disord.* **12**, 506–512.