Mitochondrial genetics of exceptional longevity in multigeneration matrilineages

Richard A. Kerber[1], Elizabeth O'Brien[1], Ron Munger[2], Ken R. Smith[3], and Richard M. Cawthon[4]

1 Department of Epidemiology & Population Health, School of Public Health and Information Sciences , University of Louisville , Louisville , KY

2 Department of Nutrition, Dietetics, and Food Sciences, Utah State University, Logan, UT

3 Department of Family and Consumer Studies, University of Utah, Salt Lake City, UT

4 Dept. of Human Genetics, University of Utah, Salt Lake City, UT

Correspondence: rcawthon@genetics.utah.edu

Abstract

Some heritable mitochondrial DNA (mtDNA) sequence variants may slow the rate of aging. The European mitochondrial haplogroup K has previously been reported to be increased in frequency in centenarians and nonagenarians relative to its frequency in younger individuals, by standard case/control study designs. To select for mitochondrial genomes likely to carry beneficial genetic variants, we screened a large genealogical database (the Utah Population Database, UPDB) for mitochondrial lineages in which the frequency of survival past 90 years was significantly higher than in the general population, and also significantly higher than in close non-matrilineal relatives. We ranked 14,900 distinct matrilineages by the strength of their association with longevity. Full sequencing of the mtDNAs from a single individual from each of 53 matrilineages in the top longevity ranks and each of 374 control matrilineages from the general Utah population, followed by analyses of the mtDNA haplogroup frequencies, identified haplogroup K2 as the haplogroup most enriched in frequency by the longevity selection (Odds Ratio = 23.05). We then analyzed overall survival and cause-specific mortality in the several thousand individuals aged 40 years or older whose mtDNA genotypes could be imputed from the 374 fully sequenced control mtDNAs. In these control matrilineages Haplogroup K2 individuals (n=332) enjoyed a significantly lower all-cause mortality risk than the general population (HR=0.81), attributable in part to a significantly lower risk of dying from heart disease (HR=0.50), as well as lower (though not significantly lower) risks of dying from cancer (HR=0.72) and diabetes (HR=0.74). Furthermore, K2 was the only haplogroup in which mortality was reduced for all three of these common causes of death.

Keywords: genetics of longevity, mitochondrial DNA, Utah Population Database

Introduction

The heritable component to a long and healthy life is likely to involve the actions and interactions of both nuclear and mitochondrial genetic variants. Mitochondrial haplogroup K has been reported to be increased in frequency in exceptionally long-lived individuals as compared to younger controls, in French [1], Irish [2], and Finnish [3] populations. Other studies, also with a case-control design, have reported an association of haplogroup K with decreased risk of Parkinson Disease [4], Alzheimer Disease in APOE 4 allele carriers [5], pancreatic cancer [6], thyroid cancer [7], and both transient ischemic attack and ischemic stroke [8].

Since the mtDNA is always maternally transmitted, all members of a multigeneration matrilineage share the same mtDNA sequence; therefore, one need only sequence the mtDNA of one individual in a matrilineage in order to know the mtDNA sequence of all members of the matrilineage. Using the Utah Population Database (https://healthcare.utah.edu/huntsmancancerinstitute/research/updb/), a large genealogical database with records on more than 8 million individuals, we identified the matrilineages with the highest frequency of exceptionally long-lived individuals, and collected blood samples from 53 of them for full mtDNA sequencing. These longevity-selected mtDNAs were compared to mtDNAs from 101 different matrilineages from the Utah CEPH (Centre d'Etudes du Polymorphisme Humain) families [9], and from an additional 273 distinct matrilineages from the Cache County Memory Study [10] provided by Dr. Munger, to identify mitochondrial haplogroups and mtDNA sequence variants enriched in the longevity-selected matrilineages. Haplogroups and sequence variants found in the control mtDNAs were also tested for association with all-cause, heart disease, cancer, and diabetes mortality among all of the known control matrilineage members who survived to 40 years or older, several thousand individuals.

Results

The two tables below present mtDNA haplogroup and mtDNA sequence variant associations with membership in a set of 53 Utah matrilineages selected for the frequent occurrence of extreme longevity (LSL, or Longevity Selected Lineages); overall survival beyond age 40 years in 374 control matrilineages from the general Utah population (273 from the NIA-funded Cache County Study on Memory Health and Aging, and 101 from the Utah CEPH families); and mortality from heart disease, cancer, and diabetes in those 374 matrilineages. Haplogroup K subclade K2, and the T9716C sequence variant shared by all haplogroup K2 members, showed significantly lower all-cause and heart disease mortality, as well as reduced (though non-significantly) risks of dying from cancer and diabetes.

	Association with LSL membership			**Association with UPDB Cohort Survival**					**Heart**			**Cancer**			**Diabetes**		
Haplogroup	**OR***	**p-value**	**FDR**	**HR****	**95% CI**		**z**	**p-value**	**HR**	**95% CI**		**HR**	**95% CI**		**HR**	**95% CI**	
H	1.13	0.822	1.000	**0.98**	**0.95**	**1.00**	**-1.74**	**0.082**	1.00	0.94	1.06	0.99	0.91	1.08	1.00	0.82	1.22
J	0.00	0.248	0.870	1.02	0.98	1.07	0.88	0.379	0.94	0.85	1.04	1.12	0.98	1.27	0.68	0.46	1.00
K	**6.46**	**0.002**	**0.010**	**0.95**	**0.91**	**1.00**	**-1.82**	**0.068**	1.08	0.97	1.20	1.02	0.88	1.18	0.86	0.59	1.25
K1	3.06	0.105	0.883	0.98	0.93	1.03	-0.70	0.483	**1.17**	**1.05**	**1.30**	1.07	0.92	1.25	0.88	0.59	1.30
K2	**23.05**	**0.002**	**0.018**	**0.81**	**0.70**	**0.93**	**-3.00**	**0.003**	**0.50**	**0.34**	**0.76**	0.72	0.45	1.16	0.74	0.24	2.31
T	0.48	0.708	0.999	1.01	0.97	1.06	0.48	0.633	1.02	0.93	1.13	0.95	0.83	1.08	1.11	0.82	1.50
U	0.55	0.551	0.993	**1.04**	**1.00**	**1.07**	**1.98**	**0.048**	0.96	0.89	1.04	0.93	0.83	1.03	1.08	0.85	1.37
W	2.31	0.386	0.965	0.96	0.86	1.07	-0.69	0.488	1.03	0.82	1.30	0.78	0.54	1.11	0.68	0.25	1.83

***OR > 1 indicates increased probability of LSL selection, OR < 1 indicates decreased probability of selection.**

****HR > 1 indicates increased probability of death, HR < 1 indicates decreased probability of death.**

Overrepresentation among the 53 LSL (longevity-selected) matrilineages and association with reduced mortality in the UPDB cohort (controls) are independent tests, and represent a replication of results (albeit within the same population) for the mtDNA nucleotide changes at 9716, 16224, and 146, and is highly suggestive for 10550, 3480, 9055, 14167, 11299, 9698, and 523.1/523.2. Some variants found in the longevity-selected matrilineages did not occur in the control matrilineages, and so could not be tested against survival in the control lineages.

		Association with LSL membership			Association with UPDB Cohort Survival					Heart			Cancer			Diabetes		
Sequence variant	**Haplogroup**	**OR***	**p-value**	**FDR**	**HR****	**95% CI**		**z**	**p-value**	**HR**	**95% CI**		**HR**	**95% CI**		**HR**	**95% CI**	
523.1+523.2 (CA ins)		3.87	0.0224	0.988	**0.95**	**0.91**	**1.00**	**-1.92**	**0.055**	1.03	1.03	1.03	1.06	0.92	1.22	0.77	0.52	1.13
A3480G	**K**	6.46	0.0017	0.192	**0.95**	**0.91**	**1.00**	**-1.82**	**0.068**	1.08	0.97	1.20	1.02	0.88	1.18	0.86	0.59	1.25
G9055A	**K**	6.46	0.0017	0.192	**0.95**	**0.91**	**1.00**	**-1.82**	**0.068**	1.08	0.97	1.20	1.02	0.88	1.18	0.86	0.59	1.25
T9698C	**K + U8a1 (2)**	5.95	0.0024	0.494	**0.96**	**0.91**	**1.00**	**-1.76**	**0.079**	1.07	0.96	1.18	1.02	0.88	1.17	0.79	0.54	1.16
T9716C	**K2**	**23.05**	**0.0016**	**0.182**	**0.81**	**0.70**	**0.93**	**-3.00**	**0.003**	**0.50**	**0.33**	**0.76**	0.72	0.45	1.16	0.74	0.24	2.31
A10550G	**K + H1e (singleton)**	8.10	0.0003	0.026	**0.95**	**0.91**	**1.00**	**-1.82**	**0.068**	1.08	0.97	1.20	1.02	0.88	1.18	0.86	0.59	1.25
T11299C	**K**	6.20	0.0021	0.216	**0.95**	**0.91**	**1.00**	**-1.82**	**0.068**	1.08	0.97	1.20	1.02	0.88	1.18	0.86	0.59	1.25
C14167T	**K**	6.46	0.0017	0.192	**0.95**	**0.91**	**1.00**	**-1.82**	**0.068**	1.08	0.97	1.20	1.02	0.88	1.18	0.86	0.59	1.25
T16224C		**5.73**	**0.0029**	**0.525**	**0.94**	**0.90**	**0.99**	**-2.59**	**0.010**	1.04	0.95	1.15	1.00	0.87	1.15	1.01	0.73	1.40
T146C		**4.51**	**0.0048**	**0.594**	**0.95**	**0.90**	**0.99**	**-2.41**	**0.016**	0.94	0.85	1.04	0.97	0.85	1.12	0.94	0.67	1.32

***OR > 1 indicates increased probability of LSL selection, OR < 1 indicates decreased probability of selection.**

****HR > 1 indicates increased probability of death, HR < 1 indicates decreased probability of death.**

Discussion

To our knowledge this report is the first to use survival and cause-specific mortality data in multigeneration matrilineages to identify mitochondrial genetic variants that may slow aging and contribute to longevity by lowering the risk of not just one, but several common causes of death. The previously reported published case-control studies [1-8] attributing protection from several aging-related diseases to haplogroup K, combined with the evidence presented here of haplogroup K2's broad health benefits, based on morbidity and mortality data from multiple multigeneration matrilineages, strongly support the conclusion that haplogroup K, and especially the K2 subclade, contributes to better survival by reducing the risk of dying from not just one, but multiple common diseases. How haplogroup K2 accomplishes this has yet to be determined, and may involve multiple mechanisms.

Reactive Oxygen Species (ROS) are generated when electrons fall off the Electron Transport Chain (ETC) prematurely (i.e. at any point prior to their donation to oxygen by Complex IV to make water). The steeper the hydrogen ion gradient across the inner mitochondrial membrane, the harder it is for electron transport to be coupled to the pumping of protons up their gradient, and the more likely it becomes for electrons to fall off the ETC prematurely. The free radical theory of aging argues that ROS damage to DNA, proteins, and lipids accumulates with age, eventually causing cell and tissue dysfunction, disease, and death.

Mitochondrially encoded amino acid substitutions in the proteins of oxidative phosphorylation could lower the rate of ROS production per unit oxygen metabolized in at least three different ways: 1) making the hydrogen ion gradient less steep by increasing the leak of protons, without ATP production, across Complex V; 2) making the hydrogen ion gradient less steep by decreasing the number of protons translocated per electron transported down the full ETC; or 3) for a given level of hydrogen ion gradient, increasing the efficiency of electon transport through the ETC, so that fewer electrons fall off prematurely.

Because haplogroup K was absent in endurance athletes (present in 0 of 52 endurance athletes vs. 48 of 1060 controls) and associated with longevity, Niemi et al. [11] suggested that it may be an "uncoupling genome." Uncoupling could be due to the G9055A-encoded amino acid substitution (Ala->Thr) in ATP synthase subunit 6 promoting proton leakage through Complex V [12]. This would be expected to reduce the efficiency of oxidative phosphorylation (limiting athletic performance) and decrease the production of reactive oxygen species (extending lifespan). Moreover, an uncoupling mitochondrial genome may prevent or reduce the rise in ROS levels associated with excessive caloric intake, thereby reducing the risk of developing insulin resistance and diabetes. Whether cells from haplogroup K2 individuals actually have a decreased mitochondrial membrane potential and lowered rates of mitochondrial ROS production remains to be tested.

While the G9055A variant is shared by all haplogroup K carriers, the K2 subclade appears to have health benefits above and beyond those shared by K haplogroup members generally. Therefore, one or more sequence variants found in K2 but missing from other K subclades may further promote health. The T9716C sequence variant is specific to K2, shared by all K2 individuals, and not found in any other mitochondrial haplogroups. Interestingly, it is a synonymous nucleotide substitution, i.e. it does not alter the amino acid sequence of its encoded protein.

It has been reported that species with higher maximum lifespans tend to have higher GC-content in their mtDNA sequences [13]. Higher GC-content may protect mitochondrial genomes, since single-stranded DNA is more prone to damage of various types than double-stranded DNA, and stretches of DNA with higher GC-content will spend less time open and single-stranded than otherwise similar sequences with lower GC-content, especially at high temperatures. Slippage of DNA polymerases during DNA replication, resulting in

small deletions and insertions, also occurs more frequently at higher temperatures, and more frequently in AT-rich sequences than in GC-rich sequences. A recent study showed that in human cells mitochondrial physiology maintains mitochondria at a temperature near 50°C [14], much higher than the typical core body temperature of around 37°C, making the hypothesis of differences in mtDNA damage rates being sensitive to haplogroup differences in GC-content even more attractive. Among European mitochondrial haplogroups, haplogroup K is among the highest for GC content. In haplogroup K2 the bolded red "t" at position 9716 in the revised Cambridge Reference Sequence, 5'-attttactgggtctctattttac-3', is mutated to a "c". Perhaps the AT-rich 7-base direct repeat (attttac) flanking 9716 makes this region prone to replication errors, and the increased GC-content provided by T9716C provides increased stability, lowering error rates. The just upstream T9698C base change, which is specific to, and shared by all, K haplogroup members, further raises GC-content, perhaps further stabilizing the region during replication. Further work will be needed to investigate these and other possible explanations for haplogroup K2's health benefits.

Research subjects & methods

Longevity-selected mitochondrial lineages

The Utah Population Database (UPDB) is a genealogical database with more than 8 million family history, demographic, and medical records of Utah residents. This is primarily a Caucasian population of Northern European origin.

We identified 14,900 human mitochondrial lineages in the UPDB with female founders born from 1800-1860, and selected those matrilineages in which the proportion of members over 65 who survived to age 90 was significantly greater than the proportion of the general population over 65 who survived to 90. We further selected for matrilineages whose members showed significantly better survival than their non-matrilineal relatives. 53 matrilineages resulting from this filtering process were selected for further study. One blood sample from an individual age 65 years or older was obtained from each of the 53 matrilineages.

Control mitochondrial lineages

One hundred one individuals from different matrilineages (based on UPDB identification of the female founders) were chosen from the 47 three-generation Utah CEPH (Centre d'etude du polymorphisme humain) families [12] for full mtDNA sequencing. The full mtDNA sequences of 273 individuals from the Cache County Memory Study [13], who were determined to be from an additional distinct set of UPDB matrilineages, were provided by Dr. Munger.

DNA sequencing

The longevity-selected mtDNAs and the Cache County mtDNAs were fully sequenced by Family Tree DNA (https://www.familytreedna.com/), Houston, TX. Overlapping PCR products covering the entire mtDNA sequence were prepared, using primer sets that have been proven not to amplify from mitochondrial-like pseudogenes in the nuclear DNA (a.k.a. numts), and sequenced by standard Sanger sequencing. Some of the Utah CEPH mtDNAs were sequenced by Family Tree DNA, and others were sequenced by the DNA sequencing core laboratory at the University of Utah also using mtDNA-specific primers.

Genetic association, survival, and cause-specific mortality analyses

Survival was analyzed in Utah Population Database cohorts who were alive and followed from the time they turned 40 years old, until death, loss to follow-up, or present time. Overall survival and deaths from specific causes were compared to mitochondrial genetic variation measured as haplogroup, subclade, and specific variant. Full mitochondrial genotypes were extrapolated to all members of each genotyped person's matrilineage. Analyses were adjusted for sex, year of birth, connectedness to pedigree structures, nuclear familial excess longevity, and familial standardized mortality ratios for specific causes of death.

Acknowledgements

We thank Dr. Mark Leppert for providing DNA samples from 101 Utah CEPH individuals for full mtDNA sequencing, Missy Dixon for Utah CEPH grandparent survival data, and Alison Fraser and the UPDB staff for survival and cause-specific mortality data for all Utah CEPH and Cache County matrilineage members. We thank the Pedigree and Population Resource (funded by the Huntsman Cancer Foundation) for its valuable role in the ongoing collection, maintenance, and support of the Utah Population Database.

Funding

This work was supported by the National Institutes of Health / National Institute on Aging, grants to RM

Cawthon AG000767, AG014495, and AG038797. The Cache County Study on Memory, Health and Aging was supported by the National Institutes of Health grant R01-1138. USTAR, the Utah Science Technology and Research initiative at the University of Utah, funded the mtDNA sequencing of the Cache County research subjects. Partial support for all datasets within the Utah Population Database was provided by the University of Utah Huntsman Cancer Institute and the Huntsman Cancer Institute Cancer Center Support grant, P30 CA2014 from the National Cancer Institute. Support for the Utah Cancer Registry is provided by Contract No HHSN2612013000171 from the National Cancer Institute with additional support from the Utah Department of Health.

References

1. Ivanova, R., V. Lepage, D. Charron, and F. Schachter. Mitochondrial genotype associated with French Caucasian centenarians. Gerontology, 1998. **44**(6): p. 349.

2. Ross, O.A., R. McCormack, M.D. Curran, R.A. Duguid, Y.A. Barnett, I.M. Rea, and D. Middleton. Mitochondrial DNA polymorphism: its role in longevity of the Irish population. Exp Gerontol, 2001. **36**(7): p. 1161-78.

3. Niemi, A.K., A. Hervonen, M. Hurme, P.J. Karhunen, M. Jylha, and K. Majamaa. Mitochondrial DNA polymorphisms associated with longevity in a Finnish population. Hum Genet, 2003. **112**(1): p. 29-33.

4. van der Walt, J.M., K.K. Nicodemus, E.R. Martin, W.K. Scott, M.A. Nance, R.L. Watts, J.P. Hubble, J.L. Haines, W.C. Koller, K. Lyons, R. Pahwa, M.B. Stern, A. Colcher, B.C. Hiner, J. Jankovic, W.G. Ondo, F.H. Allen, Jr., C.G. Goetz, G.W. Small, F. Mastaglia, J.M. Stajich, A.C. McLaurin, L.T. Middleton, B.L. Scott, D.E. Schmechel, M.A. Pericak-Vance, and J.M. Vance. Mitochondrial polymorphisms significantly reduce the risk of Parkinson disease. Am J Hum Genet 2003. **72**(4): p. 804-11.

5. Carrieri, G., M. Bonafe, M. De Luca, G. Rose, O. Varcasia, A. Bruni, R. Maletta, B. Nacmias, S. Sorbi, F. Corsonello, E. Feraco, K.F. Andreev, A.I. Yashin, C. Franceschi, and G. De Benedictis. Mitochondrial DNA haplogroups and APOE4 allele are non-independent variables in sporadic Alzheimer's disease. Hum Genet, 2001. **108**(3): p. 194-8.

6. Lam ET, Bracci PM, Holly EA, Chu C, Poon A, Wan E, White K, Kwok PY, Pawlikowska L, Tranah GJ. Mitochondrial DNA sequence variation and risk of pancreatic cancer. Cancer Res. 2012 Feb 1;72(3):686-95. doi: 10.1158/0008-5472.CAN-11-1682. Epub 2011 Dec 15. PubMed PMID: 22174369; PubMed Central PMCID: PMC3271167.

7. Cocoş R, Schipor S, Badiu C, Raicu F. Mitochondrial DNA haplogroup K as a contributor to protection against thyroid cancer in a population from southeast Europe. Mitochondrion. 2018 Mar;39:43-50. doi: 10.1016/j.mito.2017.08.012. Epub 2017 Aug 26. PubMed PMID: 28851673.

8. Chinnery PF, Elliott HR, Syed A, Rothwell PM. Oxford Vascular Study. Mitochondrial DNA haplogroups and risk of transient ischaemic attack and ischaemic stroke: a genetic association study. Lancet Neurol. 2010 May;9(5):498-503. doi: 10.1016/S1474-4422(10)70083-1. Epub 2010 Mar 31. PubMed PMID: 20362514; PubMed Central PMCID: PMC2855429.

9. Prescott SM, Lalouel JM, Leppert M. From linkage maps to quantitative trait loci: the history and science of the Utah genetic reference project. Annu Rev Genomics Hum Genet. 2008;9:347-58. doi: 10.1146/annurev.genom.9.081307.164441. Review. PubMed PMID: 18544033.

10. Welsh-Bohmer KA, Breitner JC, Hayden KM, Lyketsos C, Zandi PP, Tschanz JT, Norton MC, Munger R. Modifying dementia risk and trajectories of cognitive decline in aging: the Cache County Memory Study. Alzheimers Dement. 2006 Jul;2(3):257-60. doi: 10.1016/j.jalz.2006.04.011. PubMed PMID: 19595895.

11. Niemi, A.K. and K. Majamaa, Mitochondrial DNA and ACTN3 genotypes in Finnish elite endurance and sprint athletes. Eur J Hum Genet, 2005. **13**(8): p. 965-9.

12. Kadenbach, B. Intrinsic and extrinsic uncoupling of oxidative phosphorylation. Biochim Biophys Acta, 2003. **1604**(2): p. 77-94.

13. Lehmann, G., E. Segal, K.K. Muradian, and V.E. Fraifeld. Do mitochondrial DNA and metabolic rate complement each other in determination of the mammalian maximum longevity? Rejuvenation Res, 2008. **11**(2): p. 409-17.

14. Chrétien D, Bénit P, Ha HH, Keipert S, El-Khoury R, Chang YT, Jastroch M, Jacobs HT, Rustin P, Rak M. Mitochondria are physiologically maintained at close to 50 °C. PLoS Biol. 2018 Jan 25;16(1):e2003992. doi: 10.1371/journal.pbio.2003992. eCollection 2018 Jan. PubMed PMID: 29370167; PubMed Central PMCID: PMC5784887.