Annals of human genetics

doi: 10.1111/j.1469-1809.2011.00682.x

mtDNA Lineages Reveal Coronary Artery Disease-Associated Structures in the Lebanese Population

Marc Haber[1,†], Sonia C. Youhanna[1,†], Oleg Balanovsky[2,3], Stephanie Saade[1], Begoña Martínez-Cruz[4], Michella Ghassibe-Sabbagh[1], Nabil Shasha[5], Raed Osman[5], Hamid el Bayeh[1], Sergey Koshel[6], Valery Zaporozhchenko[2], Elena Balanovska[2], David F. Soria-Hernanz[4], Daniel E. Platt[7] and Pierre A. Zalloua[1,8*]

[1] *The Lebanese American University, Chouran, Beirut 1102 2801, Lebanon*
[2] *Research Centre for Medical Genetics, Russian Academy of Medical Sciences, Moscow, Russia*
[3] *Vavilov Institute for General Genetics, Russian Academy of Sciences, Moscow, Russia*
[4] *Evolutionary Biology Institute, Pompeu Fabra University -CSIC -PRBB, Barcelona, Spain*
[5] *Rafic Hariri University Hospital, Cardiology Department, Beirut, Lebanon*
[6] *Moscow State University, Faculty of Geography, Moscow, Russia*
[7] *Bioinformatics and Pattern Discovery, IBM T. J. Watson Research Centre, Yorktown Hgts, NY 10598, USA*
[8] *Harvard School of Public Health, Boston, MA*

Summary

Population origins and ancestry have previously been found to be important determinants of coronary artery disease (CAD). This study investigates associations of Lebanese mitochondrial DNA lineages with CAD and studies their correlation with other populations, exploring population structures that may infer mitochondria functional associations and reveal population movements and origins. Sequencing the mitochondrial hypervariable sequence 1 (HVS-1) of 363 controls and 448 cases revealed that haplogroup W was more frequent ($P = 0.013$) in cases compared to controls, and was associated with increased risk of CAD (OR = 5.50, 95% CI = 1.50–35.30, $P = 0.026$) among Lebanese samples. Haplogroup A was only found in controls ($P = 0.029$). We have detected stronger geographic correlation between haplogroup W and CAD (Pearson's $r = 0.316$, $P < 0.001$) than between haplogroup A and CAD ($r = 0.149$, $P < 0.001$). HVS-1 phylogenetic network of haplogroup W shows controls are restricted to European clusters while cases belong mostly to Middle Eastern natives. The network of haplogroup A shows that the controls belong to a cluster dominated by Central Asians. Our results show evidence of a gene flow into Lebanon, creating CAD-associated population structures that are similar to those in the source populations, maintained by limited admixture, and probably encompassing variations on the nuclear and/or the mitochondrial genome that are correlated with the disease.

Keywords: Coronary artery disease, HVS-1, Lebanon, mtDNA haplotype, population structure

Introduction

The maternally inherited human mitochondrial DNA (mtDNA) has around 16,500 base pairs. It contains ribosomal RNA, tRNA, and 37 genes supporting metabolic oxidative phosphorylation (Anderson et al., 1981). The very high mitochondrial evolutionary rate compared to the nuclear genome has provided for the accumulation of diverse mtDNA variations, marking a number of haplotypes, thus providing an excellent tool for studying human evolution, migration, and population histories (Pakendorf & Stoneking, 2005). Over the last two decades, analysis of modern human mtDNA variations have revealed modern human origins in

*Corresponding author: Pierre Zalloua, The Lebanese American University, Chouran, Beirut 1102 2801, Lebanon. Tel: +961-1-784408 Ext. 2855; Fax: +961-9-546090; E-mail: pierre.zalloua@lau.edu.lb
†Marc Haber and Sonia C. Youhanna contributed equally to the paper.

Africa and subsequent migrations to Asia and Europe (Cann et al., 1987).

It has been found that some mtDNA haplotypes not only elucidate population structures, but may also predispose to, or protect against, certain diseases (Wallace, 2005). In fact, mtDNA haplotype studies have reported associations with several diseases such as Parkinson's disease (Takasaki, 2009), Alzheimer's disease (Santoro et al., 2010), hepatocellular carcinoma (Zhang et al., 2010), breast cancer (Bai et al., 2007), multiple sclerosis (Kalman et al., 1999), Leber's hereditary optic neuropathy (Koilkonda & Guy, 2011), and type 2 diabetes (Feder et al., 2009). Mitochondrial functional differences are thought to be among the most important risk factors of coronary artery diseases (CAD), including myocardial infarction (MI) (Nishigaki et al., 2007).

Mitochondria are the primary site of superoxide production in vascular endothelial cells. Mitochondrion-derived reactive oxygen species play an important role in the pathogenesis of atherosclerosis and CAD (Guzik et al., 2006). Consequently, polymorphisms in the mtDNA are expected to associate with CAD.

It was previously reported that the mitochondrial haplogroup D, represented by an mtSNP 5178C>A (ND2: Leu237Met), was associated with longevity and with resistance against MI (Takagi et al., 2004). In another study, mtDNA haplogroup T was found to be associated with CAD and diabetic retinopathy in a Caucasian population from Austria (Kofler et al., 2009). The mtDNA variant 16189T>C increased risk of CAD and MI in Saudi Arabs (Abu-Amero et al., 2010) and of CAD in Austrians (Mueller et al., 2011), while the mitochondrial haplogroup N9b was reported to be protective against MI in Japanese males (Nishigaki et al., 2007). However, in a large study on the Danish population, no association was found between mitochondrial haplogroups and risk of ischemic cardiovascular disease (Benn et al., 2008).

Since CAD risk is both genetically and environmentally determined, and family history (Youhanna et al., 2010) along with ethnicity (Kain et al., 2003, De Lima Santos et al., 2011, Fernandes et al., 2011) play a major role in disease occurrence, the strong association of mtDNA with population structure may explain the inconsistent implication of CAD among mtDNA haplogroups, both in terms of correlation with inherited autosomal mutations, as well as correlation with cultural and environmental risks. Studies exploring population structures associated with CAD and incorporated into admixed populations at a resolution finer than ethnicity are lacking. In this study, our aim is to investigate such fine structures in a population of CAD cases and controls from Lebanon using mtDNA lineages that may infer mitochondria functional associations and/or reveal population structures, resolving population movements and origins. Lebanon possesses a unique genetic diversity; population structures among current Lebanese have been established since at least the Bronze Age (Haber et al., 2011). Post-establishment expansions to this region, such as the Islamic expansion in the 7th century and the Crusades in the 11th and 13th century, were marked by gene flow still detectable today (Zalloua et al., 2008). Lebanon therefore, provides an opportunity to observe genetic signatures of recent migrations, potentially giving a view on which lineages are disease correlated and how they are related to other populations.

Table 1 Summary statistic of study population.

	CAD status		
Variables	Cases	Controls	Combined
	$n = 448$	$n = 363$	$n = 811$
Age, in years (mean ± SD[a])	53.9 ± 9.4	54.9 ± 11.4	54.4 ± 10.4
Sex (male%)	86.4	48.5	69.4
Hypertension (%)	48.9	59.8	53.8
Hyperlipidemia (%)	50.7	37.5	44.7
Diabetes(%)	29.0	24.0	26.8
Smoking (%)	78.3	55.4	68.1
Family history of CAD[b] (%)	64.7	54.8	60.3

[a]Standard deviation.
[b]Parents or siblings.

Materials and Methods

Study Subjects and Comparative Data

The study consists of 811 unrelated Lebanese subjects randomly chosen from a previously established CAD cohort and selected from subjects who were catheterised for MI, unstable angina, or for CAD workup (Youhanna et al., 2010). The subjects were classified as 363 controls and 448 cases (Table 1). Controls have a normal angiogram defined by the absence of any atherosclerosis and/or any lesions in all coronary arteries. Cases were diagnosed with >50% stenosis in any of the coronary arteries. The study was approved by the IRB of the Lebanese American University.

Comparative data were used from the MURKA database and integrated software (Zaporozhchenko et al., 2010), which contains hypervariable sequence 1 (HVS-1) records from published sources.

Sequencing of Human Mitochondrial HVS-1

DNA was extracted from blood samples. Briefly, extraction from blood was done by lysing erythrocytes at 37°C in

0.144 M NH_4Cl, 1 mM $KHCO_3$. Leukocytes were lysed at 55°C in 10 mM Tris, 5 mM EDTA, 400 mM NaCl, 0.15% SDS and digested with 0.2 mg/ml Proteinase K. DNA was isolated using phenol/chloroform/isoamyl alcohol and subsequently precipitated in ethanol. Samples were sequenced from mtDNA positions 16024 to 16569 covering the HVS-1. Sequencing was performed on an Applied Biosystems 3130xl Genetic Analyzer. The primary amplification was achieved by primers 15876F (5'-TCAAATGGGCCTGTCCTTGTAG-3') and 639R (5'-GGGTGATGTGAGCCCGTCTA-3') (Applied Biosystems, Foster City, CA, USA). PCR products were cleaned using Exonuclease I and Shrimp Alkaline Phosphatase (Bio Basic Inc., Markham, ON, Canada). A sequencing amplification was achieved using BigDye Terminator v3.1 with primers 15946F (5'-CAAGGACAAATCAGAGAAAA-3') and 132R (5'-GACAGATACTGCGACATAGG-3') and cleaned using BigDye Xterminator, all supplied by Applied Biosystems.

Polymorphisms, Nomenclature, and Haplogroup Assignment

The HVS-1 sequences were aligned against the revised Cambridge Reference Sequence (Andrews et al., 1999) using SeqScape2.5 (Applied Biosystems) and the position of the polymorphism (16024–16569) was reported. Transitions are reported by their position on the mtDNA genome and transversions are indicated by capital letters (16147G). Deletions are marked by the letter d (16262d), heteroplasmies by the letter N (16260N), and insertions were reported as follows (16120.1A).

mtDNA haplogroups were determined using a nearest neighbour-based methodology developed for haplogroup assignment from HVS-1 sequence data (Behar et al., 2007). The analytical tool is available online at http://nnhgtool.nationalgeographic.com/classify/.

Haplogroups that were significantly different between cases and controls were genotyped using the Applied Biosystems' 7900HT Fast Real-Time PCR System with custom marker assays (Applied Biosystems) from the mtDNA coding region, confirming the haplogroup assignment. SNP at nucleotide position 11947 defined haplogroup W; SNP at nucleotide position 8794 defined haplogroup A.

Data Analysis

Haplogroup frequency analysis

mtDNA haplogroup frequencies in cases and controls were compared using the R statistical package (R Development Core Team, 2010). Pearson's χ^2 was used to test association of CAD cases or CAD controls to a certain mtDNA haplogroup; a P-value of <0.05 was considered significant. Disease risk was evaluated by calculating Odds ratios and the corresponding 95% confidence intervals. Proportional odds logistic regression (LR) analysis, from the MASS package of R, was used to assess the association between CAD and the mtDNA haplogroups, adjusting for smoking, family history of CAD (CAD in siblings or parents), hypertension (blood pressure >140/90 mmHg), hyperlipidemia (LDL >130 mg/dl and HDL <40 mg/dl), diabetes (fasting plasma glucose ≥126 mg/dl.), and age (overall mean age = 54 years).

CAD and Associated Haplogroup Mapping

In order to characterise correlations between CAD, geography, and specific populations, a frequency map of ischaemic heart disease burden was created using the WHO global infobase (WHO, 2004). In addition, frequency maps of mtDNA haplogroups associated with CAD were also constructed using data from the MURKA database (Zaporozhchenko et al., 2010). The frequency maps were computed with the software GeneGeo using algorithms described previously (Balanovsky et al., 2008). To assess the correlation between the haplogroups' maps and the CAD map, Pearson product-moment correlation coefficient (r) was calculated using the GeneGeo software implementing $r = \sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})/[(n-1)S_X S_Y]$ where $\bar{X}$ and $\bar{Y}$ are the frequency means of the mtDNA haplogroups and CAD burden, S_X and S_Y are their standard deviations, and $n = 39{,}258$ is the total number of pairs scanned on the map. Pearson's r will measure the extent to which, as one haplogroup increases, the CAD burden tends to increase.

Reduced Median (RM) Networks

RM Networks (Bandelt et al., 1995) of the haplogroups that were associated with a CAD phenotype were calculated using a reduction threshold of 1 and with equal weight on all polymorphisms. The phylogenetic relationships between haplotypes within a haplogroup can be estimated using the RM networks, which can also ascertain haplotype characteristics among the cases and controls when compared to the literature.

Results

Sequencing the HVS-1 of 811 subjects from the CAD database resulted in 428 haplotypes (Table S1) and 21 mtDNA haplogroups (Table 2). A significant difference existed between CAD cases and CAD controls; haplogroup W was found to be more frequent ($P = 0.013$) in cases (2.90%)

Table 2 mtDNA haplogroup distribution in patients with CAD and controls.

Haplogroup	Case	Control	χ^2 *P*-value	OR	95% CI	LR *P*-value	OR-Adjusted*	95% CI	LR *P*-value*
A	0	4	0.029†	–	–	–	–	–	–
D	1	2	–	–	–	–	–	–	–
H	157	134	0.581	0.91	0.68–1.21	0.497	0.91	0.68–1.21	0.497
HV	20	19	0.610	0.86	0.44–1.64	0.634	1.03	0.51–2.07	0.934
I	6	7	0.506	0.69	0.22–2.12	0.519	0.70	0.21–2.26	0.551
J	36	28	0.866	1.06	0.83–1.78	0.829	1.01	0.59–1.76	0.962
K	38	30	0.911	1.04	0.63–1.72	0.873	1.03	0.60–1.76	0.923
L0	1	0	–	–	–	–	–	–	–
L2	5	2	0.387	2.07	0.44–14.49	0.387	1.92	0.35–15.00	0.476
L3	6	2	0.259	2.49	0.49–12.21	0.266	3.19	0.64–23.83	0.187
M	9	5	0.492	1.49	0.51–4.88	0.480	1.18	0.39–3.98	0.777
N	0	1	–	–	–	–	–	–	–
N1	15	12	0.973	1.02	0.47–2.26	0.950	1.03	0.45–2.39	0.951
R	2	0	–	–	–	–	–	–	–
R0	11	14	0.251	0.59	0.26–1.29	0.190	0.59	0.25–1.35	0.215
R9	1	1	–	–	–	–	–	–	–
T	49	39	0.930	1.03	0.66–1.62	0.886	1.17	0.73–1.89	0.518
U	58	48	0.908	0.99	0.65–1.49	0.955	0.96	0.62–1.48	0.837
V	10	8	0.978	1.02	0.40–2.71	0.959	1.22	0.46–3.38	0.696
W	13	2	0.013†	5.50	1.50–35.3†	0.026	4.81	1.28–31.4†	0.043
X	10	5	0.369	1.66	0.58–5.36	0.361	1.74	0.57–5.94	0.344

*Proportional odds logistic regression adjusted for smoking, family history of CAD, hypertension, hyperlipidemia, diabetes, and age.
†Significant.

compared with controls (0.55%) by χ^2. Given 15 haplogroups, the probability of observing at least one of those 15 at a significance level of $P = 0.013$ is 0.178. Haplogroup W was found to be associated with increased risk for CAD by proportional odds LR (OR = 5.50, 95% CI = 1.50–35.30, $P = 0.026$). The probability of observing one or more of the 15 haplogroups is 0.326. Haplogroup A was only found in controls (1.1%) ($P = 0.029$) suggesting a significant decreased risk for CAD.

We employed proportional odds LR to measure the impact of smoking, family history of CAD, hypertension, hyperlipidemia, and diabetes on the association of the haplogroups with CAD. Haplogroup W correlation with increased risk of CAD persisted after adjustment (OR = 5.31, 95% CI = 1.41–34.7, $P = 0.031$) suggesting an association that is independent from the traditional CAD risk factors. The chances of seeing at least one of the 15 haplogroups at a significance of $P = 0.031$ is 0.376. The impact of gender on LR adjustment yielded large error bars since only one female was found within haplogroup W. An adjusted OR for haplogroup A could not be computed due to the small number of individuals belonging to this group.

A frequency map of world CAD burden (Fig. 1A) shows that populations at high risk are located in a region that extends from Central Asia to Eastern Europe passing through Western Russia and the Caucasus. Lebanon appears to be in a relatively moderate risk region surrounded by higher risk regions in Egypt and Iraq. Figure 1B shows that haplogroup W is found at a very low frequency in Asia, the Middle East, and Europe with sporadic high occurrences in Northern India, Kurdistan, European Russia, and Finland. Haplogroup A, on the other hand, is almost exclusive to Central and East Asia (Fig. 1C). Pearson's r shows that geographic correlation of haplogroup W with CAD by region ($r = 0.316$, $P < 0.001$) was two times higher than correlation of haplogroup A with CAD ($r = 0.149$, $P < 0.001$).

Specific regional affinity of haplogroups W and A was further investigated with the RM networks. Haplotypes within haplogroup W (Fig. 2A) show significant differentiation between cases and controls. Controls belong to a branch restricted to Europeans, (Fig. 2A, cluster α) and specific to Germans, English, and Greeks. Cases belong to branches that are restricted to Middle Easterners (Emiraties, Turkish, Saudis) (Fig. 2A, cluster β) or shared between Middle Easterners (Armenians, Turkish, Saudis, Egyptians, Kurds) and Europeans (French, Greeks, Germans, Italians) (Fig. 2A, cluster γ). In addition, cases formed a separate branch of haplotypes not found in European or Middle Eastern populations (Fig. 2A, cluster δ). The probability that random region assignments have isolated CAD controls on a branch formed only by West Europeans by chance has a Fisher *P*-value of 0.002. Moreover, removing West Europeans from the

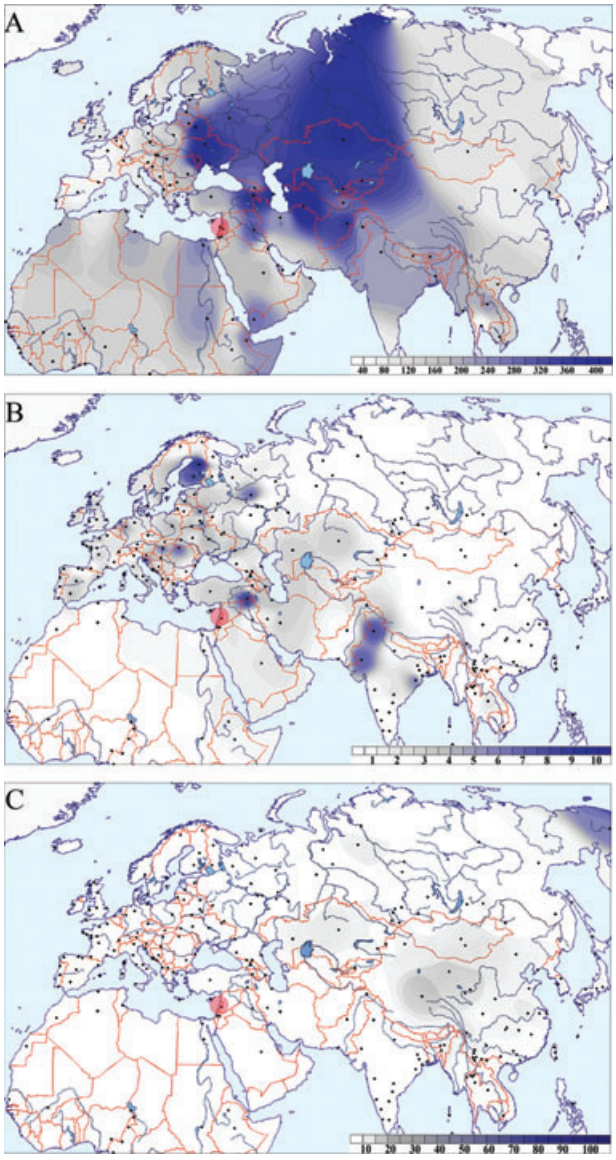

Figure 1 Frequency map of CAD burden and mtDNA haplogroups W and A.
Black dots refer to the location of the populations used in the analysis. Lebanon is highlighted by a red disk. **(**A) Map shows ischaemic heart disease burden using the WHO global infobase. The scale indicates the mortality rate per 100,000. (B) and (C) Distribution of haplogroup W and A, respectively. The scale indicates the frequency in percentage.

correlation analysis increases Pearson's (r) by 13% ($r = 0.356$, $P < 0.001$). The RM network of haplogroup A shows that the only four samples found in this study are CAD controls, which belongs to a branch dominated by Central Asians (Kazakhs, Nogays, Turkmens, Uzbecks, Kyrghyz) (Fig. 2B, cluster ε).

Discussion

We have shown that in the Lebanese population mtDNA haplogroups W and A appear to be respectively associated with increased and decreased risk to CAD. While some previous studies found different mtDNA haplogroups and haplotypes associated with CAD in various populations (Nishigaki et al., 2007, Kofler et al., 2009, Abu-Amero et al., 2010), other studies did not correlate mtDNA haplogroups to ischaemic heart disease (Benn et al., 2008) or other clinical disorders (Herrnstadt & Howell, 2004, Saxena et al., 2006). This dissimilarity in the results has been attributed to differences in study design, including the number and choice of cases and controls along with the ability to adjust for confounders (Benn et al., 2008). However, the wide confidence intervals observed in our samples suggest relatively low frequencies, suggesting the possibility of immigration. Furthermore, the P-values for association of these haplogroups with region are stronger than for disease, while CAD is also noted to be more common in those regions than surrounding regions. This suggests that the associations observed in our Lebanese samples reflect population structure and stratification marked by Hardy–Weinberg disequilibrium, reflected in correlations of somatic mutations with mtDNA haplogroups.

Population structures that are formed by the mtDNA haplogroups were not explored in previous association studies, although mtDNA is an excellent tool to study population structure and origin (Underhill & Kivisild, 2007), which are known to be independent predictors of many diseases including CAD (Kain et al., 2003, De Lima Santos et al., 2011, Fernandes et al., 2011). Interestingly, as noted in the previous paragraph, populations with the highest frequencies of haplogroups W and A were historically associated, respectively, with exceptionally high and low CAD rates, but recent environmental changes have significantly affected the disease outcome in those populations. Finland has the highest frequency of haplogroup W (Zaporozhchenko et al., 2010), and Finland had been previously reported to be the country with the highest rate of coronary events (Tunstall-Pedoe et al., 1999). However, improved treatments and control for risk factors resulted in decline of CAD mortality by 63% in 15 years (Laatikainen et al., 2005). Conversely, American Indians and Alaska Natives have the highest frequency of haplogroup A (Zaporozhchenko et al., 2010), and their rate of CAD has been historically extremely low. Nevertheless, adoption of new diets and life styles have made CAD the leading cause of death in those populations (Galloway, 2005).

mtDNA haplogroups designate lineages that were established thousands of years ago (tya). Specifically, haplogroup W is believed to have emerged 21 tya and haplogroup A appeared 30 tya (Soares et al., 2009). European W lineages have possibly originated in the Near East during the Late Upper

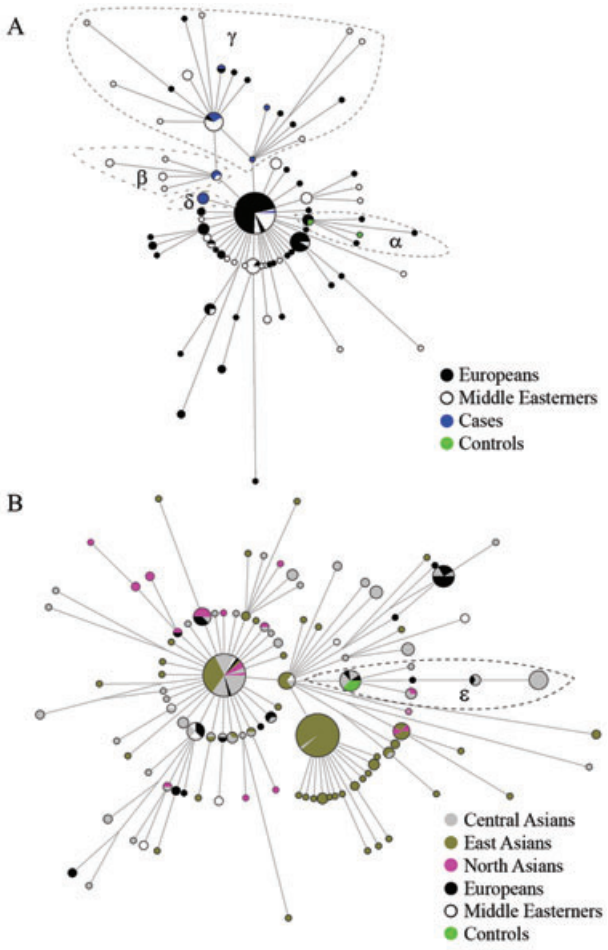

Figure 2 RM network of mtDNA haplotypes belonging to haplogroup W (A) and A (B).
Circles represent mtDNA haplotypes; area is proportional to frequency, and colour indicates the region of origin. Lines represent the mutational differences between haplotypes.

Paleolithic, establishing founder populations that may have later expanded inside Europe (Richards et al., 2000). Haplogroup A is frequent in Asia and among American Indians (Zaporozhchenko et al., 2010) but is found with very low frequency and diversity in the Near East (Zaporozhchenko et al., 2010), indicating a recent gene flow into the Levant from Asia. Since today's populations are formed by different sublineages that were founded through diverse old migratory patterns of the major haplogroups, it is plausible that the previously reported associations of haplogroups and diseases, concern specific sublineages. The RM network of haplogroup W shows that cases and controls are separated into distinct lineages. The presence of two sublineages with contrasting association to CAD could explain the high confidence interval depicting an uncertainty of the risk correlation level in the presence of a minor W lineage protective against the disease. This could also explain the non-consistency observed in the various studies exploring associations between mtDNA haplogroups and CAD, where different sublineages of a certain haplogroup are analyzed in different populations.

The haplotype diversity of the W cases appears to be similar to the surrounding populations, however, the controls' haplotypes are exclusive to a branch restricted to West Europeans. A previous study has detected West European paternal lineages in the Lebanese population attributed to Crusader activities in the 11th–13th centuries CE (Zalloua et al., 2008). The crusades brought West European men and women into established colonies in Lebanon (Hitti, 1957). Modern West Europeans have a lower CAD burden than Middle Easterners and East Europeans, among whom most of the cases are clustered on the RM network. In fact, removing the West Europeans from the correlation analysis increased association of haplogroup W and CAD, suggesting that some W lineages are negatively correlated with CAD. The gene flow from Western Europe was restricted to specific communities (Zalloua et al., 2008), with limited admixture to surrounding populations (Haber et al., 2011), enhancing the continued existence of a CAD-associated structure in the Lebanese population. It is unlikely that the presence of a minority of mtDNA haplogroup W in a population would account for enhanced CAD burden within the population, but it may suggest correlation that persists in the stratified population, as is observed within Lebanon. It appears that environmental changes after migrations of these lineages had a limited effect on CAD risk, suggesting a genetically determined protection sustained by an environmentally favourable milieu. This also applies to haplogroup A lineages that are uncommon in Lebanon but frequent in populations with a low CAD burden. The RM network shows clustering with Central Asian populations, one of which are the Nogays that were displaced during the Ottoman era to the Levant region (Hitti, 1957).

In conclusion, we have demonstrated using mtDNA data that specific matrilineal lineages in our samples were associated with CAD phenotypes. It appears that structured gene flow into Lebanon created CAD associations retained from the source populations, subsequently maintained by limited admixture, and probably encompassing variations on the nuclear and/or the mitochondrial genome that are correlated with the disease.

Acknowledgements

We would like to thank Alexey Romanov for his help in creating the frequency distribution maps. OB is funded by the Presidium of the Russian Academy of Sciences (Program "Molecular and

Cell Biology"). EB is funded by the Russian Foundation of Basic Research (grant 10–06-00451).

References

Abu-Amero, K. K., Al-Boudari, O. M., Mousa, A., Gonzalez, A. M., Larruga, J. M., Cabrera, V. M., & Dzimiri, N. (2010) The mitochondrial DNA variant 16189T>C is associated with coronary artery disease and myocardial infarction in Saudi Arabs. *Genet Test Mol Biomarkers* **14**, 43–47.

Anderson, S., Bankier, A. T., Barrell, B. G., De Bruijn, M. H., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J., Staden, R., & Young, I. G. (1981) Sequence and organization of the human mitochondrial genome. *Nature* **290**, 457–465.

Andrews, R. M., Kubacka, I., Chinnery, P. F., Lightowlers, R. N., Turnbull, D. M., & Howell, N. (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* **23**, 147.

Bai, R. K., Leal, S. M., Covarrubias, D., Liu, A., & Wong, L. J. (2007) Mitochondrial genetic background modifies breast cancer risk. *Cancer Res* **67**, 4687–4694.

Balanovsky, O., Rootsi, S., Pshenichnov, A., Kivisild, T., Churnosov, M., Evseeva, I., Pocheshkhova, E., Boldyreva, M., Yankovsky, N., Balanovska, E., & Villems, R. (2008) Two sources of the Russian patrilineal heritage in their Eurasian context. *Am J Hum Genet* **82**, 236–250.

Bandelt, H. J., Forster, P., Sykes, B. C., & Richards, M. B. (1995) Mitochondrial portraits of human populations using median networks. *Genetics* **141**, 743–753.

Behar, D. M., Rosset, S., Blue-Smith, J., Balanovsky, O., Tzur, S., Comas, D., Mitchell, R. J., Quintana-Murci, L., Tyler-Smith, C., & Wells, R. S. (2007) The Genographic Project public participation mitochondrial DNA database. *PLoS Genet* **3**, e104, doi:10.1371/journal.pgen.0030104.

Benn, M., Schwartz, M., Nordestgaard, B. G., & Tybjaerg-Hansen, A. (2008) Mitochondrial haplogroups: Ischemic cardiovascular disease, other diseases, mortality, and longevity in the general population. *Circulation* **117**, 2492–2501.

Cann, R. L., Stoneking, M., & Wilson, A. C. (1987) Mitochondrial DNA and human evolution. *Nature* **325**, 31–36.

De Lima Santos, P. C., De Oliveira Alvim, R., Ferreira, N. E., De Sa Cunha, R., Krieger, J. E., Mill, J. G., & Pereira, A. C. (2011) Ethnicity and arterial stiffness in Brazil. *Am J Hypertens* **24**, 278–284.

Feder, J., Ovadia, O., Blech, I., Cohen, J., Wainstein, J., Harman-Boehm, I., Glaser, B., & Mishmar, D. (2009) Parental diabetes status reveals association of mitochondrial DNA haplogroup J1 with type 2 diabetes. *BMC Med Genet* **10**, 60, doi:10.1186/1471-2350-10-60.

Fernandes, V. R., Cheng, S., Cheng, Y.J., Rosen, B., Agarwal, S., Mcclelland, R. L., Bluemke, D. A., & Lima, J. A. (2011) Racial and ethnic differences in subclinical myocardial function: the Multi-Ethnic Study of Atherosclerosis. *Heart* **97**, 405–410.

Galloway, J. M. (2005) Cardiovascular health among American Indians and Alaska natives – Successes, challenges, and potentials. *Am J Prev Med* **29**, 11–17.

Guzik, T. J., Sadowski, J., Guzik, B., Jopek, A., Kapelak, B., Przybylowski, P., Wierzbicki, K., Korbut, R., Harrison, D. G., & Channon, K. M. (2006) Coronary artery superoxide production and nox isoform expression in human coronary artery disease. *Arterioscler Thromb Vasc Biol* **26**, 333–339.

Haber, M., Platt, D. E., Badro, D. A., Xue, Y., El-Sibai, M., Bonab, M. A., Youhanna, S. C., Saade, S., Soria-Hernanz, D. F., Royyuru, A., Wells, R. S., Tyler-Smith, C., Zalloua, P. A., & The Genograhic Consortium (2011) Influences of history, geography, and religion on genetic structure: The Maronites in Lebanon. *Eur J Hum Genet* **19**, 334–340.

Herrnstadt, C. & Howell, N. (2004) An evolutionary perspective on pathogenic mtDNA mutations: Haplogroup associations of clinical disorders. *Mitochondrion* **4**, 791–798.

Hitti, P. (1957) *Lebanon in history from the earliest times to the present*. London: Macmillan.

Kain, K., Catto, A. J., & Grant, P. J. (2003) Associations between insulin resistance and thrombotic risk factors in high-risk South Asian subjects. *Diabet Med* **20**, 651–655.

Kalman, B., Li, S., Chatterjee, D., O'connor, J., Voehl, M. R., Brown, M. D., & Alder, H. (1999) Large scale screening of the mitochondrial DNA reveals no pathogenic mutations but a haplotype associated with multiple sclerosis in Caucasians. *Acta Neurol Scand* **99**, 16–25.

Kofler, B., Mueller, E. E., Eder, W., Stanger, O., Maier, R., Weger, M., Haas, A., Winker, R., Schmut, O., Paulweber, B., Iglseder, B., Renner, W., Wiesbauer, M., Aigner, I., Santic, D., Zimmermann, F. A., Mayr, J. A., & Sperl, W. (2009) Mitochondrial DNA haplogroup T is associated with coronary artery disease and diabetic retinopathy: A case control study. *BMC Med Genet* **10**, 35, doi:10.1186/1471-2350-10-35.

Koilkonda, R. D. & Guy, J. (2011) Leber's hereditary optic neuropathy-gene therapy: From benchtop to bedside. *J Ophthalmol* **2011**, 179412, doi:10.1155/2011/179412.

Laatikainen, T., Critchley, J., Vartiainen, E., Salomaa, V., Ketonen, M., & Capewell, S. (2005) Explaining the decline in coronary heart disease mortality in Finland between 1982 and 1997. *Am J Epidemiol* **162**, 764–773.

Mueller, E. E., Eder, W., Ebner, S., Schwaiger, E., Santic, D., Kreindl, T., Stanger, O., Paulweber, B., Iglseder, B., Oberkofler, H., Maier, R., Mayr, J. A., Krempler, F., Weitgasser, R., Patsch, W., Sperl, W., & Kofler, B. (2011) The mitochondrial T16189C polymorphism is associated with coronary artery disease in Middle European populations. *PLoS One* **6**, e16455, doi:10.1371/journal.pone.0016455.

Nishigaki, Y., Yamada, Y., Fuku, N., Matsuo, H., Segawa, T., Watanabe, S., Kato, K., Yokoi, K., Yamaguchi, S., Nozawa, Y., & Tanaka, M. (2007) Mitochondrial haplogroup N9b is protective against myocardial infarction in Japanese males. *Hum Genet* **120**, 827–836.

Pakendorf, B. & Stoneking, M. (2005) Mitochondrial DNA and human evolution. *Annu Rev Genomics Hum Genet* **6**, 165–183.

R Development Core Team (2010) *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Richards, M., Macaulay, V., Hickey, E., Vega, E., Sykes, B., Guida, V., Rengo, C., Sellitto, D., Cruciani, F., Kivisild, T., Villems, R., Thomas, M., Rychkov, S., Rychkov, O., Rychkov, Y., Golge, M., Dimitrov, D., Hill, E., Bradley, D., Romano, V., Cali, F., Vona, G., Demaine, A., Papiha, S., Triantaphyllidis, C., Stefanescu, G., Hatina, J., Belledi, M., Di Rienzo, A., Novelletto, A., Oppenheim, A., Norby, S., Al-Zaheri, N., Santachiara-Benerecetti, S., Scozari, R., Torroni, A., & Bandelt, H. J. (2000) Tracing European founder lineages in the Near Eastern mtDNA pool. *Am J Hum Genet* **67**, 1251–1276.

Santoro, A., Balbi, V., Balducci, E., Pirazzini, C., Rosini, F., Tavano, F., Achilli, A., Siviero, P., Minicuci, N., Bellavista, E., Mishto, M.,

Salvioli, S., Marchegiani, F., Cardelli, M., Olivieri, F., Nacmias, B., Chiamenti, A. M., Benussi, L., Ghidoni, R., Rose, G., Gabelli, C., Binetti, G., Sorbi, S., Crepaldi, G., Passarino, G., Torroni, A. & Franceschi, C. (2010) Evidence for sub-haplogroup h5 of mitochondrial DNA as a risk factor for late onset Alzheimer's disease. *PLoS One*, **5**, e12037, doi:10.1371/journal.pone.0012037.

Saxena, R., De Bakker, P.I., Singer, K., Mootha, V., Burtt, N., Hirschhorn, J. N., Gaudet, D., Isomaa, B., Daly, M. J., Groop, L., Ardlie, K. G., & Altshuler, D. (2006) Comprehensive association testing of common mitochondrial DNA variation in metabolic disease. *Am J Hum Genet* **79**, 54–61.

Soares, P., Ermini, L., Thomson, N., Mormina, M., Rito, T., Rohl, A., Salas, A., Oppenheimer, S., Macaulay, V., & Richards, M. B. (2009) Correcting for purifying selection: An improved human mitochondrial molecular clock. *Am J Hum Genet* **84**, 740–759.

Takagi, K., Yamada, Y., Gong, J. S., Sone, T., Yokota, M., & Tanaka, M. (2004) Association of a 5178C–>A (Leu237Met) polymorphism in the mitochondrial DNA with a low prevalence of myocardial infarction in Japanese individuals. *Atherosclerosis* **175**, 281–286.

Takasaki, S. (2009) Mitochondrial haplogroups associated with Japanese centenarians, Alzheimer's patients, Parkinson's patients, type 2 diabetic patients and healthy non-obese young males. *J Genet Genomics* **36**, 425–434.

Tunstall-Pedoe, H., Kuulasmaa, K., Mahonen, M., Tolonen, H., Ruokokoski, E., & Amouyel, P. (1999) Contribution of trends in survival and coronary-event rates to changes in coronary heart disease mortality: 10-year results from 37 WHO MONICA project populations. Monitoring trends and determinants in cardiovascular disease. *Lancet* **353**, 1547–1557.

Underhill, P. A. & Kivisild, T. (2007) Use of y chromosome and mitochondrial DNA population structure in tracing human migrations. *Annu Rev Genet* **41**, 539–564.

Wallace, D. C. (2005) A mitochondrial paradigm of metabolic and degenerative diseases, aging, and cancer: A dawn for evolutionary medicine. *Annu Rev Genet* **39**, 359–407.

World Health Organization (2004) WHO global infobase: International comparisons. In: *Ischaemic Heart Disease*, (https://apps.who.int/infobase/Comparisons.aspx).

Youhanna, S., Platt, D. E., Rebeiz, A., Lauridsen, M., Deeb, M. E., Nasrallah, A., Alam, S., Puzantian, H., Kabbani, S., Ghoul, M., Zreik, T. G., El Bayeh, H., Abchee, A., & Zalloua, P. (2010) Parental consanguinity and family history of coronary artery disease strongly predict early stenosis. *Atherosclerosis*, **212**, 559–563.

Zalloua, P. A., Xue, Y., Khalife, J., Makhoul, N., Debiane, L., Platt, D. E., Royyuru, A. K., Herrera, R. J., Hernanz, D. F., Blue-Smith, J., Wells, R. S., Comas, D., Bertranpetit, J., & Tyler-Smith, C. (2008) Y-chromosomal diversity in Lebanon is structured by recent historical events. *Am J Hum Genet* **82**, 873–882.

Zaporozhchenko, V., Balanovsky, O., Pshenichnov, A., & Balanovska, E. (2010) *MURKA: Global mitochondrial database and intergated software*. Moscow: Research Centre for Medical Genetics, Russian Academy of Medical Sciences.

Zhang, R., Zhang, F., Wang, C., Wang, S., Shiao, Y. H., & Guo, Z. (2010) Identification of sequence polymorphism in the D-Loop region of mitochondrial DNA as a risk factor for hepatocellular carcinoma with distinct etiology. *J Exp Clin Cancer Res* **29,** 130, doi:10.1186/1756-9966-29-130.

Supporting Information

Additional supporting information may be found in the online version of this article:

Table S1 mtDNA haplogroups and haplotypes in 811 unrelated subjects from the Lebanese CAD database.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organised for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

Received: 22 June 2011
Accepted: 1 September 2011