

SHORT REPORT

Ashkenazi Jewish mtDNA haplogroup distribution varies among distinct subpopulations: lessons of population substructure in a closed group

Jeanette Feder¹, Ofer Ovadia¹, Benjamin Glaser² and Dan Mishmar^{*,1}

¹Department of Life Sciences, Ben-Gurion University of the Negev, Beer-Sheva, Israel; ²Department of Endocrinology, Hadassah-Hebrew University Medical School, Jerusalem, Israel

The quest for genes associated with diseases is widely recognized as an essential task in the effort to investigate the genetic basis of complex human disorders and traits. A basic stage in association studies is the careful choice of the model population, with preference to closed groups having little population substructure. Here, we show evidence for significant geographic substructure ($P = 0.017$) of the maternal lineage represented by mitochondrial DNA variation in one of the most commonly studied populations, the Ashkenazi Jews. Most of the substructure effect stems from differential representation of haplogroups K and H. Our results underline the essentiality of adjusting data of population genetic variation for substructure during the design of association studies, even in apparently closed populations.

European Journal of Human Genetics (2007) 15, 498–500. doi:10.1038/sj.ejhg.5201764; published online 24 January 2007

Keywords: ashkenazi jews; population sub-structure; mtDNA

Ashkenazi Jews, considered to be an isolated population that has undergone a recent bottleneck,^{1–4} constitute a model population for the search of disease-causing mutations and disease-susceptibility genes. One of the most commonly used arguments for choosing Ashkenazi Jews for such studies is the lack of population substructure.⁵ Nevertheless, differential distribution of disease-causing mutations among Ashkenazi Jews of Eastern European versus Central European ancestry suggests genetic drift and may thus imply possible population substructures among Ashkenazi Jews.⁶ This consideration is particularly crucial when association studies are performed with the common disease variant approach, as the prevalence of the disease-associated variant may vary among populations owing to genetic drift.⁴ Indeed, during an association study on type

II diabetes mellitus and its complications recently conducted by our group, we found a significant difference in the distribution of linked sets of mitochondrial DNA (mtDNA) common variants (haplogroups) in Ashkenazi Jews of different geographic origins (Feder *et al.*, 2007, submitted). This observation, along with the frequent use of mtDNA as a target for association studies, led us to assess, using mtDNA, population substructure in the maternal lineage of Ashkenazi Jews.

A cohort of 300 healthy unrelated subjects of Ashkenazi Jewish origin, who previously served as control subjects in association studies at the Hebrew University, was analyzed in a hierarchical manner for mtDNA haplogroups, starting from haplogroups K and H, reported to be the most prevalent in Ashkenazi Jews, followed by the less prevalent types.^{1,7} To increase the sample size, we added to the analysis previously published data for mtDNA haplogroup distribution in 565 unrelated subjects of the Ashkenazi Jewish origin.⁷ Among these 865 subjects, 704 were with known maternal geographic origin. Classification of these subjects according to maternal geographic origin gave three subpopulations that were considered sufficiently

*Correspondence: Dr D Mishmar, Department of Life Sciences, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel.

Tel.: +972 8 6461355; Fax: +972 8 6461356;

E-mail: dmishmar@bgu.ac.il

Received 9 June 2006; revised 24 October 2006; accepted 21 November 2006; published online 24 January 2007

Table 1 Haplogroup distribution in Ashkenazi Jews (Ashj) according to the maternal geographic origin

Haplogroup	Un-planned test	Poland Ashj		Romanian Ashj		RU Ashj	
		N = 192	%	N = 104	%	N = 150	%
U (non-K)	a	12	6.3	5	4.8	11	7.3
K	c	72	37.5	32	30.8	25	16.7
HV*	a	20	10.4	8	7.7	19	12.7
H	b	27	14.1	27	26	41	27.3
J	a	18	9.4	9	8.7	15	10
T	a	9	4.7	3	2.9	11	7.3
N1	a	12	6.3	11	10.6	11	7.3
I		3	1.6	0	0	0	0
W	a	6	3.1	1	1	4	2.7
X		1	0.5	2	1.9	0	0
L1 and L2	a	6	3.1	2	1.9	2	1.3
M	a	2	1	2	1.9	0	0
Pre-HV		2	1	1	1	6	4
Others		2	1	1	1	5	3.3

HV* harbors the entire HV lineage, excluding the H haplogroup. To avoid small sample size, haplogroups were either grouped on the basis of phylogeny (ie, N1 and I, W and X) or added to 'others' (haplogroups M and Pre-HV). Different letters in the unplanned test column indicate significant contribution of a specific haplogroup to the overall heterogeneity. RU, Russia and the Ukraine.

Table 2 Haplogroup distribution in non-Jewish local populations and Jews

Haplogroup	Polish Jews		Polish non-Jews		Russian and Ukrainian (RU) Jews		Russian non-Jews	
	N = 192	%	N = 436	%	N = 150	%	N = 201	%
U (non K)	12	6.3	70	16.1	11	7.3	36	17.9
K	72	37.5	15	3.4	25	16.7	6	3
HV*	20	10.4	25	5.7	19	12.7	15	7.5
H	27	14.6	197	45.2	41	27.3	85	42.3
J	18	9.4	34	7.8	15	10	16	8
T	9	4.7	50	11.5	11	7.3	22	10.9
N1	12	6.3	2	0.5	11	7.3	0	0
I	3	1.6	8	1.8	0	0	5	2.5
W	6	3.1	16	3.7	4	2.7	4	2
X	1	0.5	8	1.8	0	0	7	3.5
L1 and L2	6	3.1	0	0	2	1.3	0	0
L3	0	0	1	0.2	0	0	0	0
R*	0	0	2	0.5	0	0	1	0.5
M	2	1	8	1.8	0	0	3	1.5
Pre-HV	2	1	0	0	6	4	1	0.5
Others	2	1	0	0	5	3.3	0	0

Numbers for the non-Jewish populations are taken from Malyarchuk *et al.*¹⁰ To avoid small sample size, haplogroups were either grouped based on phylogeny (ie, N1 and I, W and X) or added to 'others' (haplogroups L, M and Pre-HV).

large for further analysis: that originating from Russia and the Ukraine (combined into one group due to geographical proximity, 'RU'), that from Poland, and that from Romania. This division into subpopulations resulted in a total of 446 subjects included in the study (Table 1). To avoid comparisons of haplogroups for which sample sizes were small, we grouped some haplogroups together according to phylogenetic considerations, that is, N1 with I, and W with X. Haplogroups M and Pre-HV (as well as haplogroup L in Table 2) were considered together with 'others' (Table 1). We used the $R \times C$ test of independence to compare the haplogroup distribution among the three different Jewish communities.⁸ To determine which of the haplogroups

made the largest contribution to the overall heterogeneity, we performed an unplanned test.⁸ The statistical analyses were performed using Systat 9.0 (Systat Software, Inc., CA, USA), and results were considered statistically significant if $\alpha < 0.05$.

Comparison of overall haplogroup distribution ($R \times C$ test of independence) revealed significant differences among the three Jewish populations (RU, Poland and Romania, $P = 0.017$). Using an unplanned test, we found that haplogroups K and H made the largest contribution to the observed overall heterogeneity (Table 1). Accordingly, a close examination of the data (Table 1) revealed that the frequency of haplogroup K in the RU sample was

considerably lower than that in the other two populations; similarly, the frequency of haplogroup H was lower in the Polish group. Since the Ashkenazi Jewish population originated ~1000 years ago, we speculate that these differences may have been caused by genetic drift rather than non-random processes such as natural selection.

Alternatively, our findings may imply varying degrees of admixture of Jewish with non-Jewish populations in the above-mentioned communities. Thomas *et al.*⁹ compared mtDNA variation of Jewish with local non-Jewish populations in a number of geographic locations throughout the world and did not find support for such an explanation. However, these researchers considered Ashkenazi Jews as a single entity, assuming that all Ashkenazi Jewish communities from Central and Eastern Europe originated exclusively from the Rhine basin and thus compared Ashkenazi Jewish mtDNA diversity only with that of non-Jewish Germans. To test for possible admixture between Ashkenazi communities and local non-Jewish populations, we used a log-linear model⁸ to compare mtDNA haplogroup distribution of the Jewish RU and Polish groups to the available data on the haplogroup distribution of non-Jewish RU and Polish populations, respectively¹⁰ (Table 2). This analysis revealed a significant divergence in total haplogroup distribution between the Ashkenazi Jewish and the local populations (ethnic background \times haplogroup interaction term, $G = 173$, $df = 8$, $P < 0.001$). A nonsignificant three-way interaction term ($G = 7.7$, $df = 8$, $P = 0.463$) reflects that the differences between Jews and non-Jews was consistent both in the RU and Polish populations. These findings, taken together with HVR1 sequences for some of the haplogroups, such as N1b, that contain motifs restricted and common to all Ashkenazi Jewish populations,¹ may further support the interpretation of little or no gene flow of the local non-Jewish communities in Poland and Russia to the Jewish communities in these countries. This conclusion may also suggest founder event(s), resulting in the Jewish settlement in Poland and Russia, consistent with the view that the ancestry of the Ashkenazi Jewish population is a result of at least four different founder events.¹ Clearly, the differences between Jews and non-Jews (Table 2) are far larger than those observed among the Jewish communities (Table 1). Hence, differences between the Jewish communities can be overlooked when non-Jews are included in the comparisons.

Our observation that haplogroup K is one of the main contributors to differences in mtDNA haplogroup distribution in Ashkenazi Jews is of special interest. Haplogroup K was previously associated with protection against Parkinson's disease in Americans of European ancestry.¹¹ Hence, since Ashkenazi Jews are also of European origin and since this haplogroup could associate with Parkinson's disease or with other complex disorders in Ashkenazi Jews, we stress that geographic origins should be taken into

account during the design of association studies, as was found for the apparently homogenous population in Iceland.¹²

It is notable that haplogroups K and H can be further divided into subgroups, of which only a few are present in Ashkenazi Jews;¹ such subgroups may further contribute to our observed population differences and thus their distribution should be studied in Ashkenazi populations of different maternal geographic origins.

In conclusion, although Ashkenazi Jews may be regarded as a defined ethnic group, it is necessary to take population substructure into account when designing association studies, that is, patients and controls of Ashkenazi Jewish origin should be also matched for geographic origin.

Acknowledgements

This research project was supported by research grants from the Israel Science Foundation and the Israel Ministry of Health (D.M.), a fellowship from the Israel Ministry of Absorption (J.F.) and a grant from the Russell Berrie Foundation and D-Cure, Diabetes Care in Israel (B.G.).

References

- Behar DM, Metspalu E, Kivisild T *et al*: The matrilineal ancestry of Ashkenazi Jewry: portrait of a recent founder event. *Am J Hum Genet* 2006; **78**: 487–497.
- Charrow J: Ashkenazi Jewish genetic disorders. *Fam Cancer* 2004; **3**: 201–206.
- Ostrer H: A genetic profile of contemporary Jewish populations. *Nat Rev Genet* 2001; **2**: 891–898.
- Service S, DeYoung J, Karayiorgou M *et al*: Magnitude and distribution of linkage disequilibrium in population isolates and implications for genome-wide association studies. *Nat Genet* 2006; **38**: 556–560.
- Shifman S, Bronstein M, Sternfeld M *et al*: A highly significant association between a COMT haplotype and schizophrenia. *Am J Hum Genet* 2002; **71**: 1296–1302.
- Risch N, Tang H, Katzenstein H, Ekstein J: Geographic distribution of disease mutations in the Ashkenazi Jewish population supports genetic drift over selection. *Am J Hum Genet* 2003; **72**: 812–822.
- Behar DM, Hammer MF, Garrigan D *et al*: MtDNA evidence for a genetic bottleneck in the early history of the Ashkenazi Jewish population. *Eur J Hum Genet* 2004; **12**: 355–364.
- Sokal RR, Rohlf FJ: *Biometry: the Principles and Practice of Statistics in Biological Research*. W.H. Freeman and Company: New York, 1995.
- Thomas MG, Weale ME, Jones AL *et al*: Founding mothers of Jewish communities: geographically separated Jewish groups were independently founded by very few female ancestors. *Am J Hum Genet* 2002; **70**: 1411–1420.
- Malyarchuk BA, Grzybowski T, Derenko MV, Czarny J, Wozniak M, Miscicka-Sliwka D: Mitochondrial DNA variability in Poles and Russians. *Ann Hum Genet* 2002; **66**: 261–283.
- van der Walt JM, Nicodemus KK, Martin ER *et al*: Mitochondrial polymorphisms significantly reduce the risk of Parkinson disease. *Am J Hum Genet* 2003; **72**: 804–811.
- Helgason A, Yngvadottir B, Hrafnkelsson B, Gulcher J, Stefansson K: An Icelandic example of the impact of population structure on association studies. *Nat Genet* 2005; **37**: 90–95.