

SHORT COMMUNICATION

On the Y-chromosome haplogroup C3c classification

Boris A Malyarchuk, Miroslava Derenko and Galina Denisova

As there are ambiguities in classification of the Y-chromosome haplogroup C3c, relatively frequent in populations of Northern Asia, we analyzed all three haplogroup-defining markers M48, M77 and M86 in C3-M217-individuals from Siberia, Eastern Asia and Eastern Europe. We have found that haplogroup C3c is characterized by the derived state at M48, whereas mutations at both M77 and M86 define subhaplogroup C3c1. The branch defined by M48 alone would belong to subhaplogroup C3c*, characteristic for some populations of Central and Eastern Siberia, such as Koryaks, Evens, Evenks and Yukaghirs. Subhaplogroup C3c* individuals could be considered as remnants of the Neolithic population of Siberia, based on the age of C3c*-short tandem repeat variation amounting to 4.5 ± 2.4 thousand years.

Journal of Human Genetics (2012) 57, 685–686; doi:10.1038/jhg.2012.93; published online 19 July 2012

Keywords: haplogroup; phylogeography; Siberia; STR; Y chromosome

According to a modified version of the Y-chromosome Consortium tree,¹ haplogroup C3c is defined by three parallel mutations: M48, M77 and M86. However, Pakendorf *et al.*^{2,3} have demonstrated that a few Siberian individuals from Evenk and Yukaghir populations have the derived state at M48, but the ancestral state at M86. They suggested that the derived state at M48 and the ancestral state at M86 defines subhaplogroup C3c*, whereas the derived state at M86 further defines subhaplogroup C3c1. Unfortunately, this suggestion had not been taken into account and, in general, any one of three markers is analyzed in studies of Eurasian populations (for example, M48^{4–8}, M86^{9–12} and M77^{13–16} in studies). Moreover, according to the structure of the Y-chromosome haplogroup tree proposed by the International Society of Genetic Genealogy (<http://www.isogg.org>; Version: 7.28; Date: 18 June 2012), haplogroup C3c is determined by parallel mutations at M48 and M77, whereas the derived state at M86 defines its C3c1 branch.

We have recently analyzed haplogroup C structure in populations of Northern Eurasia in a total sample of 1449 males from 18 ethnic groups of Siberia, Eastern Asia and Eastern Europe.¹⁶ We have found that, besides haplogroup C3c defined in our study by marker M77, some Siberian populations are characterized by high frequency of C3* Y chromosomes (for instance, Koryaks have 38.5% of C3* haplotypes). Because there are aforementioned ambiguities in C3c classification and a possibility of incomplete identification of haplogroup C3c using the only M77 marker, we have analyzed all three single-nucleotide polymorphisms (SNPs)—M48, M77 and M86—in two sets of samples in the present study, previously typed as belonging to haplogroups C3c-M77 (141 samples) and C3* (x C3a, C3b, C3c, C3d, C3e and C3f) (90 samples).

M48 and M86 typing of C3c-M77 and C3* Y chromosomes from populations of Northern Asia and Eastern Europe has shown that haplogroup C3c is strongly defined by marker M48 and consists of

two subhaplogroups—C3c1 characterizing by parallel mutations at loci M77 and M86 and C3c* with ancestral states at M77 and M86 (Table 1, Supplementary Table S1). Therefore, all previously published C3c-M77 and C3c-M86 haplotypes would belong, in fact, to subhaplogroup C3c1, and it might be that some of the individuals who were classified previously as belonging to haplogroup C3c, based on analysis of the M48 alone, belong to subhaplogroup C3c*.

In our study, the highest frequencies of subhaplogroup C3c1-(M77, M86) were observed in Tungusic-speaking people of North-Eastern Asia, such as Evens and Evenks, as well as in Turkic-speaking Altaian Kazakhs and Mongolic-speaking Kalmyks. These results are in agreement with previous observations based on separate or joint genotyping of M77 and M86 markers.^{3,9,12,16}

C3c* haplotypes were detected in aboriginal populations of North-Eastern Asia—Koryaks (28.2%) and Evens (1.6%) from the Sea of Okhotsk coast (Magadan region) and West Evenks (2.4%) from Central Siberia (Evenki Autonomous District) (Table 1). Earlier, two Evenk individuals from southern part of Yakutia, one Yakut-speaking Evenk and one Yukaghir were found to belong to C3c*.^{2,3} Therefore, the geographic distribution of subhaplogroup C3c* is limited to the eastern part of Siberia.

Comparison of short tandem repeat (STR) haplotypes belonging to subhaplogroup C3c* demonstrates that haplotypes found in Koryaks, Evenks, Yukaghirs and Even are very similar to each other and probably belong to the same branch (Supplementary Table S2). A search of similar nine-marker haplotypes (loci DYS19, DYS385a, DYS385b, DYS389I, DYS389II, DYS390, DYS391, DYS392 and DYS393) in the YHRD 3.0 database¹⁷ (<http://www.yhrd.org>; release 37 built at 17 February 2012; 99 258 haplotypes within 732 world populations) has shown (Supplementary Table S2) that analogous haplotypes are

Table 1 Haplogroups C3c and C3* distribution (number of individuals and % values in parentheses) in populations studied

Population	N	C3c-M48		
		C3c*-M48 (xM77 and M86)	C3c1-M77, M86	C3*-M217 (xC3a, C3b, C3c, C3d, C3e and C3f)
Altaians	89	0	5 (5.6)	9 (10.1)
Teleuts	44	0	0	4 (9.1)
Altaiian Kazakhs	36	0	15 (41.7)	6 (16.7)
Khakassians	64	0	0	1 (1.6)
Shors	38	0	0	1 (2.6)
Todjins	26	0	2 (7.7)	0
Tuvinians	108	0	6 (5.6)	3 (2.8)
Yakuts	10	0	2 (20.0)	0
Sojots	28	0	0	0
Mongols	46	0	5 (10.9)	18 (39.1)
Buryats	217	0	13 (6.0)	18 (8.3)
Khamnigans	51	0	0	1 (0.5)
Kalmyks	91	0	41 (45.1)	5 (5.5)
Evenks	41	1 (2.4)	18 (43.9)	0
Evens	63	1 (1.6)	33 (52.4)	0
Koryaks	39	11 (28.2)	0	4 (10.3)
Koreans	52	0	0	5 (9.6)
Russians	406	0	1 (0.2)	2 (0.5)

Table 2 The age of STR variation within haplogroup C3c and its subgroups based on STR diversity

Haplogroups and sample size	Time estimates in thousand years (ky) ^a
C3c (154)	9.9 ± 5.2
C3c1 (141)	6.7 ± 3.2
C3c* (13)	4.5 ± 2.4

Abbreviation: STR, short tandem repeat.

For tri- and tetranucleotide loci DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437 and DYS439.

^aAges of haplogroups calculated by the method of Zhivotovsky *et al.*¹⁹ with the evolutionary effective mutation rate equal to 6.9×10^{-4} per 25 years. STR loci DYS19, DYS385a, DYS385b and DYS438 were excluded from analysis.

present only in Yakuts from western part of Yakutia (Vilui region).¹⁸ Unfortunately, these haplotypes were not typed using SNP markers, so their exact phylogenetic affiliation remains unclear.

The ages of STR variation of haplogroups C3c and its branches are <10 ky (Table 2), reflecting relatively recent population expansions, connected probably with the Holocene warming. Meanwhile, the divergence time (T_D)¹⁹ between subhaplogroups C3c1 and C3c* is very high, being estimated as 30.2 ± 25.9 ky, pointing to the Upper Pleistocene origin and divergence of the C3c lineages. We should note, however, that the calculated divergence time is so high due to very high repeat variance at locus DYS389I (5.7 versus 0.00025–0.1 at the remaining seven loci), which is characterized by the average means of the repeat number of 10 and 13.4 in subhaplogroups C3c* and C3c1, respectively. However, if we assume that the deletion of three repeats at locus DYS389I in the founder C3c* haplotype was the result of a single mutation event, then the divergence time of C3c1 and C3c* haplotypes is estimated as 8.6 ± 4.6 ky ago, started in the Holocene epoch. Unfortunately, results of molecular dating in this case are controversial, and further studies are required to clarify the haplogroup C3c diversification patterns.

To sum up, by analyzing three SNPs (M48, M77 and M86), we found that haplogroup C3c is defined by marker M48, and both M77 and M86 mutations occurred on the background of derived M48 chromosomes, yielding the subhaplogroup C3c1. Subhaplogroup C3c* individuals have been detected, albeit at low frequencies, in populations of Central and Eastern Siberia (Koryaks, Evenks, Evens and Yukaghirs) up to the Okhotsk Sea coast, and based on the age of STR variation, they could be considered as remnants of the Neolithic population of Siberia.

ACKNOWLEDGEMENTS

This study was supported by the Program of Presidium of Russian Academy of Sciences 'Biodiversity' (subprogram 'Gene pools and genetic diversity') and the Russian Foundation for Basic Research (11-04-00620 and 11-04-98508).

- Karafet, T. M., Mendez, F. L., Meilerman, M. B., Underhill, P. A., Zegura, S. L. & Hammer, M. F. New binary polymorphisms reshape and increase resolution of the human Y-chromosomal haplogroup tree. *Genome Res.* **18**, 830–838 (2008).
- Pakendorf, B., Novgorodov, I. N., Osakovskij, V. L., Danilova, A. P., Protod'jakonov, A. P. & Stoneking, M. Investigating the effects of prehistoric migrations in Siberia: genetic variation and the origins of Yakuts. *Hum. Genet.* **120**, 334–353 (2006).
- Pakendorf, B., Novgorodov, I. N., Osakovskij, V. L. & Stoneking, M. Mating patterns amongst Siberian reindeer herders: inferences from mtDNA and Y-chromosomal analyses. *Am. J. Phys. Anthropol.* **133**, 1013–1027 (2007).
- Wells, R. S., Yuldashaeva, N., Ruzibakiev, R., Underhill, P. A., Evseeva, I., Blue-Smith, J. *et al.* The Eurasian heartland: a continental perspective on Y-chromosome diversity. *Proc. Natl Acad. Sci. USA* **98**, 10244–10249 (2001).
- Lell, J. T., Sukernik, R. I., Starikovskaya, Y. B., Su, B., Jin, L., Schurr, T. G. *et al.* The dual origin and Siberian affinities of Native American Y chromosomes. *Am. J. Hum. Genet.* **70**, 192–206 (2002).
- Zerjal, T., Wells, R. S., Yuldashaeva, N., Ruzibakiev, R. & Tyler-Smith, C. A genetic landscape reshaped by recent events: Y-chromosomal insights into central Asia. *Am. J. Hum. Genet.* **71**, 466–482 (2002).
- Xue, Y., Zerjal, T., Bao, W., Zhu, S., Shu, Q., Xu, J. *et al.* Male demography in East Asia: a north-south contrast in human population expansion times. *Genetics* **172**, 2431–2439 (2006).
- Zhong, H., Shi, H., Qi, X. B., Xiao, C. J., Jin, L., Ma, R. Z. *et al.* Global distribution of Y-chromosome haplogroup C reveals the prehistoric migration routes of African exodus and early settlement in East Asia. *J. Hum. Genet.* **55**, 428–435 (2010).
- Karafet, T. M., Osipova, L. P., Gubina, M. A., Posukh, O. L., Zegura, S. L. & Hammer, M. F. High levels of Y-chromosome differentiation among native Siberian populations and the genetic signature of a boreal hunter-gatherer way of life. *Hum. Biol.* **74**, 761–789 (2002).
- Zegura, S. L., Karafet, T. M., Zhivotovsky, L. A. & Hammer, M. F. High-resolution SNPs and microsatellite haplotypes point to a single, recent entry of Native American Y chromosomes into the Americas. *Mol. Biol. Evol.* **21**, 164–175 (2004).
- Hammer, M. F., Karafet, T. M., Park, H., Omoto, K., Harihara, S., Stoneking, M. *et al.* Dual origins of the Japanese: common ground for hunter-gatherer and farmer Y chromosomes. *J. Hum. Genet.* **51**, 47–58 (2006).
- Dulik, M. C., Osipova, L. P. & Schurr, T. G. Y-chromosome variation in Altaian Kazakhs reveals a common paternal gene pool for Kazakhs and the influence of Mongolian expansions. *PLoS One* **6**, e17548 (2011).
- Regueiro, M., Cadenas, A. M., Gayden, T., Underhill, P. A. & Herrera, R. J. Iran: tricontinental nexus for Y-chromosome driven migration. *Hum. Hered.* **61**, 132–143, 2006.
- Gayden, T., Cadenas, A. M., Regueiro, M., Singh, N. B., Zhivotovsky, L. A., Underhill, P. A. *et al.* The Himalayas as a directional barrier to gene flow. *Am. J. Hum. Genet.* **80**, 884–894 (2007).
- Kharkov, V. N., Stepanov, V. A., Medvedev, O. F., Spiridonova, M. G., Maksimova, N. R., Nogovitsyna, A. N. *et al.* The origin of Yakuts: analysis of Y-chromosome haplotypes. *Mol. Biol. (Mosk)* **42**, 226–237 (2008).
- Malyarchuk, B., Derenko, M., Denisova, G., Wozniak, M., Grzybowski, T., Dambueva, I. *et al.* Phylogeography of the Y-chromosome haplogroup C in northern Eurasia. *Ann. Hum. Genet.* **74**, 539–546 (2010).
- Willuweit, S. & Roewer, L. Y chromosome haplotype reference database (YHRD): update. *Forensic Sci. Int. Genet.* **1**, 83–87 (2007).
- Thèves, C., Balaresque, P., Evdokimova, L. E., Timofeev, I. V., Alekseev, A. N., Sevina, A. *et al.* Population genetics of 17 Y-chromosomal STR loci in Yakutia. *Forensic Sci. Int. Genet.* **4**, e129–e130 (2010).
- Zhivotovsky, L. A., Underhill, P. A., Cinnioglu, C., Kayser, M., Morar, B., Kivisild, T. *et al.* The effective mutation rate at Y chromosome short tandem repeats, with application to human population-divergence time. *Am. J. Hum. Genet.* **74**, 50–61 (2004).

Supplementary Information accompanies the paper on Journal of Human Genetics website (<http://www.nature.com/jhg>)