

LETTERS

On the origin of Y-chromosome haplogroup N1b

European Journal of Human Genetics (2009) 17, 1540–1541;
doi:10.1038/ejhg.2009.100; published online 17 June 2009

Recently, Mirabal *et al*¹ have noted that their results of Y-chromosome haplogroup (hg) N1b analysis challenge earlier findings^{2,3} and suggest that N1b-P43 mutation may have occurred in the Uralic range rather than in Siberia and much earlier than has been proposed (12.9 ± 4.1 instead of 5.2 ± 2.7 kya). Mirabal *et al* performed high-resolution haplotyping of Y chromosomes from populations in the East European region of Russia and the Uralic mountains and compared them with relevant previously published data. Meanwhile, they did not find it possible to compare the data obtained with South Siberian data set, so we are intending here to make up this deficiency. In addition to the data obtained by Mirabal *et al*, we have used the data on hg N1b diversity of Northwest Siberian Khants and Mansi⁴ and Russian Pomors⁵ as well as our published data on N1b diversity in South Siberians (Altaians, Shors, Tuvinians, Tofalars, Sojots, Khakassians, Buryats, Mongolians and Kalmyks) and Central and Northeast Siberians (Yakuts, Evenks, Evens and Koryaks).³ In the latter study, 83 individuals were found as belonging to hg N1b. Median-joining analysis of 53 N1b haplotypes revealed in 168 individuals from Siberia, Ural mountains and Eastern Europe shows (as in the previous studies^{2,3}) a well-resolved bipartite distribution of haplotypes with separate Siberian and European branches, N1b-A and N1b-E, respectively (Figure 1). Previous studies have shown that South Siberian subcluster, N1b-A, may have originated first, later giving rise to the derived Northwest Siberian/Uralic/Northeast European subcluster, N1b-E.^{2,3} However, new data presented here do not allow resolving the problem of geographic origin of hg N1b ancestor because the putative ancestral N1b haplotype (indicated by a star in Figure 1) appears to be equally widespread both in South Siberia (among Khakassians, Altaians, Mongolians, Tofalars, Evenks and Shors) and in Northwest Siberia (among Khants and Mansi). It is noteworthy, however, that earlier haplotype preceding subcluster N1b-E in the network is the haplotype revealed in Khants, pointing to the possibility of Northwest Siberian origin of the N1b-E subcluster.

According to our data, the age of STR variation within hg N1b as a whole is about 12 ky (12.1 ± 4.4 ky). This age was estimated as the average squared difference in the number of repeats between all the current chromosomes and the founder haplotype, averaged over microsatellite loci and divided by the effective mutation rate of 0.00069 per locus per 25 years.⁷ We used the median haplotypes (formed by the median values of the repeat scores at each microsatellite locus within each hg) as the founding ones.⁸ The age of subcluster N1b-A in Siberian and Uralic populations is 4.2 ± 1.1 ky, with the age prevailing in South Siberia in comparison with Northwest Siberia and Urals (4.8 ± 1.7 vs 3.1 ± 0.9 ky, respectively). However, the age of subcluster N1b-E in Northwest Siberian, Uralic and Northeast European Russian populations is much higher, being equal to 6.8 ± 2.6 ky. Subcluster N1b-E differs sharply from N1b-A in its median repeat scores at four loci (DYS19, DYS385a,b, DYS389II and DYS392), thus suggesting that N1b-E chromosomes descended from single founders with multiple jumps at several loci, thus causing a possible shift in statistical estimates that assumes a step-wise mutation model.²



Figure 1 Median-joining network of hg N1b based on 11 STR loci (DYS19, DYS385a,b, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438 and DYS439) constructed using the Network 4.5.1.0. (www.fluxus-engineering.com)⁶ Presumed founder (and median) haplotype with allelic structure 14-12,13-13-16-23-10-14-13-14-10-10 for the aforementioned loci is indicated by a star. Circles represent STR haplotypes with area proportional to frequencies, and colors indicate the region of origin. South Siberian haplotypes (Tuviniian, Buryat, Mongolian, Altaian, Tofalar, Khakassian, Shor, Kalmyk and Evenk) are shown in black, Northwest Siberian haplotypes (Khants and Mansi) are shown in gray, whereas Uralic and Eastern European (Komi, Tatar, Mari, Vepsa and Russian) haplotypes are shown in white.

A possible scenario explaining the phylogeographic pattern of hg N1b is that the subclusters N1b-A and N1b-E might have diverged from a common ancestor anywhere in Siberia approximately 15 ky ago (with upper bound of divergence time, T_D , between these subclusters estimated (according to Zhivotovsky⁹) as 14.7 ± 5.9 ky). According to STR variation, expansion of Northwest Siberian/Uralic/Northeast European branch, N1b-E, may have occurred approximately 7 kya and the expansion of South Siberian branch N1b-A about 5 kya. Divergence time between South Siberian and Northwest Siberian/Uralic N1b-A haplotypes is 4 ky ($T_D = 4.0 \pm 1.0$ ky), so this suggests that expansion of N1b-A lineages in Northwest Siberia and Urals (about 3 kya) might have occurred right after their migrations from South Siberia. Future studies of Y-chromosome variation using a combined STR-SNP approach may enable a better definition of geographic origin of N1b-P43 mutation.

Acknowledgements

This study was supported by the Program of Presidium of Russian Academy of Sciences 'Biodiversity and Gene Pools' and the Russian Foundation for Basic Research (07-04-00445).

Boris Malyarchuk¹ and Miroslava Derenko¹

¹Genetics Laboratory, Institute of Biological Problems of the North, Russian Academy of Sciences, Magadan, Russia
E-mail: malyarchuk@ibpn.ru

References

- Mirabal S, Regueiro M, Cadenas AM *et al*: Y-Chromosome distribution within the geo-linguistic landscape of northwestern Russia. *Eur J Hum Genet* 2009; **17**: 1260–1273.
- Rootsi S, Zhivotovsky LA, Baldovic M *et al*: A counter-clockwise northern route of the Y-chromosome haplogroup N from Southeast Asia towards Europe. *Eur J Hum Genet* 2007; **15**: 211–405.
- Derenko M, Malyarchuk B, Denisova GA *et al*: Y-chromosome haplogroup N dispersals from south Siberia to Europe. *J Hum Genet* 2007; **52**: 763–770.
- Pimenoff VN, Comas P, Palo JU, Vershubsky G, Kozlov A, Sajantila A: Northwest Siberian Khanty and Mansi in the junction of West and East Eurasian gene pools as revealed by uniparental markers. *Eur J Hum Genet* 2008; **16**: 1254–1264.
- Balanovsky O, Rootsi S, Pshenichnov A *et al*: Two sources of the Russian patrilineal heritage in their Eurasian context. *Am J Hum Genet* 2008; **82**: 236–250.
- Bandelt H-J, Forster P, Roehl A: Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 1999; **16**: 37–48.
- Zhivotovsky LA, Underhill PA, Cinnioglu C *et al*: The effective mutation rate at Y chromosome short tandem repeats, with application to human population divergence time. *Am J Hum Genet* 2004; **74**: 50–61.
- Sengupta S, Zhivotovsky LA, King R *et al*: Polarity and temporality of high-resolution Y-chromosome distributions in India identify both indigenous and exogenous expansions and reveals minor genetic influence of Central Asian pastoralists. *Am J Hum Genet* 2006; **78**: 202–221.
- Zhivotovsky LA: Estimating divergence time with the use of microsatellite genetic distances: impacts of population growth and gene flow. *Mol Biol Evol* 2001; **18**: 700–709.

Reply to B Malyarchuk and M Derenko: a need for further investigation of Uralic and Siberian populations in the search for haplogroup N1b's origins

European Journal of Human Genetics (2009) **17**, 1541–1543;
doi:10.1038/ejhg.2009.101; published online 17 June 2009

Malyarchuk and Derenko (see accompanying letter) have expanded on our previous work¹ by including northwest² and south³ Siberian samples in Time to Most Recent Common Ancestor (TMRCA) calculations under haplogroup N1b (P43) (The Russian Pomor populations⁴ the authors claim to have newly included in their analyses were already included in our TMRCA calculations¹). We have taken the liberty of reanalyzing the data published in the aforementioned articles along with the results from the publications already investigated in our previously published work¹ using a population based as opposed to a regionally based approach to compliment Malyarchuk and Derenko's findings.

It should be noted that the results presented by Malyarchuk and Derenko should be tempered given that all claims made are based solely on TMRCA estimates. Although TMRCA estimates provide a signal of demographic expansion dating to the last recognizable bottleneck, determination of high frequency and/or diversity areas are particularly pertinent. In addition, these estimates (TMRCA) are subject to large standard errors, can be affected by founder population size and may also reflect gene flow within the same lineage. Insights from archaeology, linguistics and other lines of evidence besides molecular clocks may be worth considering. It should also be noted that STR haplotypes generally coalesce to more recent times than those ascertained by TMRCA as a result of lineage extinctions that are not always possible to detect.⁵

Overall, age of accumulated Y-chromosome short tandem repeat (Y-STR) variation of 12.1 ± 4.4 kilo years ago (kya) as per Malyarchuk and Derenko's analyses is in close agreement

with our estimate yielding 11.5 ± 3.1 kya, when all individuals from the regions analyzed (European, Uralic and Siberian) are pooled (Table 1). Malyarchuk and Derenko explain that data attained in their analyses suggests that the subset of Y-chromosomes belonging to haplogroup N1b-E (European branch) may have arisen earlier than N1b-A (Asian branch), as the former attains older age estimates (6.8 ± 2.6 kya) than the latter (4.8 ± 1.7 kya). Our analyses yield similar ages to the ones reported by Malyarchuk and Derenko for Europeans and Siberians/Mongolians (Table 1), and also indicate that Y-STR variance and accumulated age estimates in a population-by-population basis are higher in European collections, ranging from 0.083 in the Mezen (3.0 ± 2.1 kya) to 0.598 in the Other Slavs (exhibiting a bipartite structure including both Asian (6.0 ± 3.7 kya) and European (7.0 ± 3.6 kya) core haplotypes) than in Asian groups, which fluctuate from 0.046 in the South Siberian Tuvinians (1.7 ± 0.7 kya) to 0.292 in the northwest Siberian Mansi (9.7 ± 4.4 kya) (Table 2). Therefore, our results are in support of Malyarchuk and Derenko's findings, but in disagreement with the traditional view that branch N1b-A is older than N1b-E.⁶

The highest overall accumulated Y-STR variance and age estimates are achieved in the Komi from Priluzski (0.611 and 12.9 ± 4.1 kya, respectively), as previously indicated.¹ Meanwhile, with the exemption of the Mansi from northwest Siberia, all Asian populations exhibit accumulated

ages of 6.3 Kya or less (Table 2). These results, along with the overall homogeneity (low Y-STR variance values as compared with the European or Uralic groups included in this study) throughout northwest and south Siberia and Mongolia are not suggestive of these regions as the cradle for haplogroup N1b. Generally, populations within Siberia are highly stratified, exhibiting pronounced levels of differentiation between them^{8,9} and had haplogroup N1b arisen in Siberia, it is likely that high variance estimates would be apparent; however, the Siberian collections included in this work do not possess this said variation (Table 2) despite the large number of Asian Y-chromosomes sampled ($N = 140$).

Malyarchuk and Derenko also suggest that N1b-E may have arisen in the Khants, as the proclaimed ancestral modal haplotype for this branch is preceded by a profile present in this group in the figure they have presented. The accumulated variance in the Khants (taking both data from our previous work¹ and from Pimenoff *et al*²) yields an age estimate of 4.3 ± 2.3 kya for the haplogroup's expansion in the population, a value that is in accordance with the rest of the collections in the region. The conclusion that northwest Siberia may be the birthplace for this branch of haplogroup N1b is undermined by these relatively young TMRCA results. In addition, the Khants and the Mansi are known to have arrived in northwest Siberia only recently during the Soviet era, and represent the

Table 1 Time to most recent common ancestor (TMRCA) for haplogroup N1b (P43)

Population (region)	N	Network shape	TMRCA ^a (kya)
Europeans	29	Bipartite star	11.9 ± 4.6 (6.0 ± 3.7 and 7.7 ± 3.8) ^b
Uralic	16	Bipartite star	12.8 ± 5.0 (5.0 ± 3.6 and 6.6 ± 4.4) ^b
Siberians/Mongolians	140	Bipartite star	4.9 ± 1.7 (6.9 ± 3.3 and 3.5 ± 1.6) ^b
All	185	Star-like	11.5 ± 3.1

Abbreviation: kya, kilo years ago.

^aA mutation rate of 0.00069 mutations per locus per generation was used to estimate time.

^bSubclustering of haplotypes led to separate time estimates for each branch.

Table 2 Haplotype variance and time to most recent common ancestor (TMRCA) for haplogroup N1b (P43)

Population (region)	N	Haplotype variance	Network shape	TMRCA ^a (kya)	References
Komi Izhemski (Uralic)	9	0.181	Star-like	6.7 ± 4.2	1
Komi Priluzski (Uralic)	7	0.611	Star-like	12.9 ± 4.1	1
Mezen (Slavic)	4	0.083	Star-like	3.0 ± 2.1	4
Pinega (Slavic)	15	0.163	Star-like	5.6 ± 2.9	4
Other Slavs ^b	10	0.598	Bipartite Star	16.3 ± 5.8 (6.0 ± 3.7 and 7.0 ± 3.6) ^c	3,4
Hezhen (Siberian)	8	0.077	Star-like	3.0 ± 1.8	7
Khakassian (Siberian)	21	0.056	Star-like	6.3 ± 5.2	3
Tuviniian (Siberian)	25	0.046	Star-like	1.7 ± 0.7	3
Mansi (Siberian)	15	0.292	Star-like	9.7 ± 4.4	2
Khanty (Siberian)	28	0.129	Star-like	4.3 ± 2.3	1,2
Other Siberians/Mongolians ^d	43	0.059	Star-like	2.2 ± 1.2	3,7

Abbreviation: kya, kilo years ago.

^aA mutation rate of 0.00069 mutations per locus per generation was used to estimate time.

^bKrasnoborsk, Vologda, Belgorod, Cossacs, Livni, Porhov.

^cSubclustering of haplotypes led to separate time estimates for each branch.

^dUyghur, Oroqen, Outer Mongolian, Shors, Kalmyks, Tofalars, Altaian, Buryats, Evenks.

expansion from the Uralic range¹⁰ (they both speak Finno-Ugric languages^{11,12}).

The higher variance and age estimates (Tables 1 and 2) for both European and Uralic populations as compared to the Siberian groups lend support to our previous statement suggesting that N1b may have arisen in the Uralic range or in an area/region proximal to it; however, a much more exhaustive investigative effort must be conducted to trace the origins of N1b and its branches N1b-A and N1b-E.

Acknowledgements

We thank Drs Boris A Malyarchuk and Miroslava Derenko for their interest in our work and for taking the time to closely examine our data.

Sheyla Mirabal^{1,2}, Peter A Underhill³ and Rene J Herrera^{1,2}

¹Department of Molecular and Human Genetics, College of Medicine, Florida International University, Miami, FL, USA

²Department of Biological Sciences, Florida International University, Miami, FL, USA; ;

³Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA, USA
E-mail: herrerar@fiu.edu

References

- 1 Mirabal S, Regueiro MM, Cadenas AM *et al*: Y-chromosome distribution within the geo-linguistic landscape of northwestern Russia. *Eur J Hum Genet* 2009. Published On-line.
- 2 Pimenoff VN, Comas D, Palo JU *et al*: Northwest Siberian Khanty and Mansi in the junction of West and Eurasian gene pools as revealed by uniparental markers. *Eur J Hum Genet* 2008; **82**: 236–250.
- 3 Derenko M, Malyarchuk B, Denisova GA *et al*: Y-chromosome haplogroup N dispersals from south Siberia to Europe. *J Hum Genet* 2007; **52**: 763–770.
- 4 Balanovsky O, Rootsi S, Pshenichnov A *et al*: Two sources of the Russian patrilineal heritage in their Eurasian context. *Am J Hum Genet* 2008; **82**: 236–250.
- 5 Bosch E, Calafell F, Santos FR *et al*: Variation in short tandem repeats is deeply structured by genetic background on the human Y chromosome. *Am J Hum Genet* 1999; **65**: 1623–1638.
- 6 Rootsi S, Zhivotovsky LA, Baldovic M *et al*: A counter-clockwise northern route of the Y chromosome haplogroup N from Southeast Asia towards Europe. *Eur J Hum Genet* 2007; **15**: 204–211.
- 7 Xue Y, Zerjal T, Ban W *et al*: Male demography in East Asia: a north-south contrast in human population expansion times. *Genetics* 2006; **172**: 2431–2439.
- 8 Karafet TM, Xu L, Du R *et al*: Paternal population history of East Asia: sources, patterns, and microevolutionary processes. *Am J Hum Genet* 2001; **69**: 615–628.
- 9 Karafet TM, Osipova LP, Gubina MA *et al*: High levels of Y-chromosome differentiation among native Siberian populations and the genetic signature of a boreal hunter-gatherer way of life. *Hum Biol* 2002; **74**: 761–789.
- 10 Osherenko G: Indigenous Political and Property rights and Economic/Environmental Reform in Northwest Siberia. *Post-Sov Geogr* 1995; **36**: 225–237.
- 11 Abondolo DM: Khnty; In: Abondolo DM (ed): *The Uralic Languages*. Routledge: London and New York, 1998, pp 358–386.
- 12 Kersztes L: Mansi; In: Abondolo DM (ed): *The Uralic Languages*. Routledge: London and New York, 1998, pp 387–427.