ATTENTION! THIS ARTICLE IS OUTDATED; SEE AN UPDATED 67 MARKER ARTICLE IN http://www.mv.helsinki.fi/home/jphakkin/N1c1.xps

Stratification of Y-haplogroup N1c

(Jaakko Häkkinen, August 5th, 2010)

Explanations:
> = direction of a mutation
10 = original ancestral value
11 = result of one-step mutation (10 > 11)
12 = result of two-step mutation (10 > 12) or double mutation (10 > 11 > 12)
10 = result of (assumed) back-mutation (10 > 11 > 10)
[1] = two similar haplotypes which for areal reason are probably of different origin
* = Haplogroup is N but not tested further
(REC.) = reconstructed haplotype (not yet found)

1. Rooting of N1c

In this analysis I use the name N1c (Tat-C), as it seems to be identical with N1c1 (M178): even the most remote Yakut haplotypes have both of these mutations (Derenko et al. 2007).

The status of different N1c (former N3) subgroups differs radically, depending on how the N1c is rooted (connected) to the phylogenetic Y-chromosomal haplogroup tree. The descending relations (which haplotype is ancestral and which has developed from it) may even turn totally upside down.

For example, a thorough presentation ”Suomen miehet” (”The Men of Finland”) by Kalevi Wiik follows the rooting of N1c presented in the study of Lappalainen et al. (2006), where the N1c (former N3) founder haplotype is:

14 11 14 14 16 24 11 14 13 (Lappalainen et al. 2006: ht 85)
(DYS19, 385a, 385b, 389-I, 389b [II–I], 390, 391, 392, 393)

This haplotype differs only in two *loci* from the reference haplotype presenting haplogroup K:

14 **12** 14 14 16 24 11 **13** 13 (Lappalainen et al. 2006: ht 205)

Based on this rooting, all the other Finnish N1c haplotypes are derived from the founder haplotype 85 by Wiik. However, after the above mentioned study there have occurred a plenty of data from the other branches of haplogroup N (Rootsi et al. 2007; Derenko et al. 2007; Balanovsky et al. 2008; Mirabal et al. 2009), and there seems to be a need for a serious reconsideration:

14 11 12 14 16 23 10 13 13 (Rootsi et al. 2007: Chinese N*)
13 12 12 13 17 22 10 14 14 (Rootsi et al. 2007: Fiji N*)
14 12 13 13 16 22 10 15 13 (Rootsi et al. 2007: Kazah N1a [former N1])
14 12 13 13 16 23 10 14 13 (Derenko et al. 2007: Khakass N1b [former N2])

By comparing these N haplotypes from different branches, we may reconstruct the N1 founder haplotype, which happens to be identical with the Khakass N1b haplotype:

14 12 13 13 16 23 10 14 13 (N1 FHT)

Now the before mentioned ht 85 differs from this most probable N1 founder haplotype in even 5 *loci* out of the 9 in common!

14 **11** **14** **14** 16 **24** **11** 14 13 (Lappalainen et al. 2006: ht 85)
So it seems to be a pure coincidence that in the widespread and numerous N1c haplogroup there has been born a haplotype quite similar to the ancestral haplogroup K. Consequently, the evolution of N1c haplotypes of Finland has unfortunately been derived reversedly from the tip of the branch to the root. This particular Eastern Finnish (Southern Karelian) haplotype 85 is actually very high in the tree, as we may see by creating the following path to it, using already found haplotypes and starting from the reconstructed N1 founder haplotype:

14 12 13 13 16 23 10 14 13 (N1 FHT)
14 **11** 13 13 16 23 10 14 13 (Khakass)
14 **11** **14** 14 16 23 10 14 13 (Udmurt, Finn.W)
14 **11** **14** **14** 16 23 10 14 13 (Komi.Pri, Komi.Izh, Altaian, Mansi, Khanty, Kar, Rus.Iwa, Rus.Pen, Mari, Chuvash)
14 **11** **14** **14** 16 23 **11** 14 13 (Finn.E)
14 **11** **14** **14** 16 **24** **11** 14 13 (Finn.E, Saami.Swe)

It is also crucial to take into consideration as large (geographically, not necessarily numerically) haplotype corpus as possible. There is no point in trying to derive one Finnish haplotype from another, if there are similar haplotypes elsewhere in the world. It is of course more credible that the particular mutation has occurred once and then the ancestral form and the descendant form have together spread to the wider area, than assume that the same mutations have occurred several times in several places.

Below is my attempt to stratificate the large corpus of N1c haplotypes, collected from the studies listed at the end of this writing. It may well be, that if any new SNP's within the haplogroup N1c are found (so far there are not enough results), this classification may turn out to be incorrect. Or if there some day are enough eastern N1c haplotypes with 67 marker resolution, this classification may similarly turn out to be incorrect.

But while the ultimate descent of a certain group may change, the areal notifications (in parenthesis) still stay. And by the areal distribution it is possible to tentatively connect certain groups to certain historical occasions: if your N1c haplotype belongs to the "Gulf of Finland" -group, this probably does not change even if there were corrections made in the descent of the lineage. So your haplotype still has closest relatives among the Estonians and Karelians, and your lineage still can be connected to the (Proto-)Finnic people. And this kind of things are what people want to know, when they get their DNA test results. N1c is so wide-spread haplogroup, that this kind of stratification is the only way to give people more accurate knowledge about their paternal lineage.

I shall not list all the haplotypes, because there are several hundreds of them, and my collected data table is ten pages long; I give only the founder haplotype of each group and a few haplotypes differing from it mainly by one mutation. The rest of the N1c haplotypes are descendants of those presented below, and mostly quite easily placeable into the given groups.

2. Stratification of N1c haplotypes

1. group: DYS385a = 12 > 11

14 **11** 13 13 16 23 10 14 13 (Khakass)

1a. DYS393 = 13 > 14 (FENNOSCANDIA)
14 **11** 13 13 16 23 10 14 **14** (Swe.Västerb, Finn, Est, Kar, Rus.Arch)
14 **11** 13 13 16 **24** 10 14 **14** (Rus.Arch, Rus.Smol, Finn.W)
14 **11** 13 13 16 ***25*** 10 14 **14** (Finn.E)
14 **11** 13 13 16 **22** 10 14 **14** (Swe.Upp)
14 **11** 13 13 ***18*** 23 10 14 **14** (Altaian)

1a1. DYS19 = 14 > 13 (GULF OF BOTHNIA)
13 11 13 13 16 23 10 14 **14** (Swe.Skara, Finn.W)
13 11 13 13 **17** 23 10 14 **14** (Swe.GötE)

1a2. DYS19 = 14 > 15 (EASTERN BALTIC SEA)
15 11 13 13 16 23 10 14 **14** (Est, Finn.W, Lit, Swe.Västerb, Rus.Vol)
15 11 13 13 16 **24** 10 14 **14** (Finn.W)
15 ***10*** 13 13 16 23 10 14 **14** (Lat, Lit, Est)
15 11 14 13 16 23 10 14 **14** (Rus.Bri)

1a2a. DYS391 = 10 > 11
15 11 13 13 16 23 **11** 14 **14** (Est, Finn, Lat, Lit, Komi.Izh, Rus)
15 11 13 13 16 **22 11** 14 **14** (Finn.W)

1a2a1. DYS385b = 13 > 14 (BALTIA)
15 11 14 13 16 23 **11** 14 **14** (Lit, Lat, Est, Rus.Now, Slovak [1])
15 *12* **14** 13 16 23 **11** 14 **14** (Lat, Lit)
15 11 14 14 16 23 **11** 14 **14** (Lit, Lat, Est, Rus.Now)

1a2b. DYS389-I = 13 > 14
15 11 13 **14** 16 23 10 14 **14** (Est, Finn.W, Lit, Lat [2])
15 11 13 **14** 15 **24** 10 14 **14** (Finn.E)

1a2b1. DYS391 = 10 > 11
15 11 13 **14** 16 23 **11** 14 **14** (Lit, Lat, Est, Finn, Kar)
15 11 13 **14** 16 23 **11 15 14** (*Gediminid*)
16 **11** 13 **14** 16 23 **11** 14 **14** (Lit, Lat, Est, Finn, Kar)
15 11 12 14 16 23 **11** 14 **14** (Est, Lat, Lith)
15 *12* 13 **14** 16 23 **11** 14 **14** (Lit, Lat)
15 11 13 **14 17** 23 **11** 14 **14** (Lit)
15 11 13 **14 17 24 11** 14 **14** (Lit, Finn.E)
15 11 13 **14** 16 **24 11** 14 **14** (Lit, Finn, Swe)
15 11 13 **14** 16 **24** ***12*** 14 **14** (FinnE)

1a2c. DYS385b = 13 > 14
15 11 14 13 16 23 10 14 **14** (Finn.Österb, Rus.Bri)

1a3. DYS389-I = 13 > 14
14 **11** 13 **14** 16 23 10 14 **14** (Tuvan, Buryat, Russian, Swe.Upp, Finn, Saami.FT)
14 **11** 13 **14** 16 **24** 10 14 **14** (Rus.Bri, Finn, Saami.FT, Swe.Got)

1a3a. DYS391 = 10 > 11 (MOST EXPANSIVE)
14 **11** 13 **14** 16 23 **11** 14 **14** (Komi.pri, Komi.Izh, Russ, Ukr, Est, Kar, Mong, Chukchi, Esk, Turk, Madjar [Kazakhstan]*, Swe, Saa, Finn, Lit, *Rurikid*)
14 **11** 13 **14 15** 23 **11** 14 **14** (Kar, Finn.E)
13 11 13 **14** 16 23 **11** 14 **14** (Finn.E, Kar)
14 **11** 13 **14 17** 23 **11** 14 **14** (Finn, Rus.Arch)
15 11 13 **14** 16 23 **11** 14 **14** (Even [2])

1a3a1. DYS385a = 11 > 10
14 ***10*** 13 **14** 16 23 **11** 14 **14** (Est, Swe)
14 ***10*** 13 **14 15** 23 **11** 14 **14** (Kar)

1a3a2. DYS392 = 14 > 15 (SIBERIAN TURKIC)
14 **11** 13 **14** 16 23 **11 15 14** (Tofalar)
14 **11** 13 **14** 16 23 **11 15 14** (Evenk)
14 **11** 13 **14 17** 23 **11 15 14** (Yakut)
14 **11** 13 **14 *18*** 23 **11 *16* 14** (Yakut)

1a3a3. DYS385b = 13 > 11 (SOUTHERN)
14 **11 *11* 14** 16 23 **11** 14 **14** (Komi.Pri, Khanty, Slovak, Tuva, Altai, Yakut)
14 **11 *11* 14** 16 **24 11** 14 ***15*** (Finn.E)

1a3a4. DYS390 = 23 > 24 (GULF OF FINLAND)
14 **11** 13 **14** 16 **24 11** 14 **14** (Finn, Est, Kar, Swe, Saami.FT)
14 **11** 13 ***15*** 16 **24 11** 14 **14** (Est, Kar, Finn.E, Saami.FT)

1a3a4a. DYS389b = 16 > 17
14 **11** 13 **14 17 24 11** 14 **14** (Est, Kar, Finn.E)
14 **11** 13 **14 17 24 11 *16* 14** (Rus.Lip)
14 **11** 13 ***15* 17 24 11** 14 **14** (Rus.Iwa, Est)

1a3a4b. DYS390 = 24 > 25 (SAVONIAN)
14 **11** 13 **14** 16 ***25* 11** 14 **14** (Finn.E, Swe.Värml, Kar)
14 **11** 13 **14** 16 ***25 12*** 14 ***15*** (Finn.E)

1a4. DYS389b = 16 > 17 (WESTERN FINNISH)
14 **11** 13 13 **17** 23 10 14 **14** (Finn.W)
14 **11** 13 13 **17 24** 10 14 **14** (Finn.W)

1b. DYS391 = 10 > 11 (FINNO-SAAMIC)
14 **11** 13 13 16 23 **11** 14 13 (Est, Finn, Rus, Swe.Värm)

ATTENTION! THIS ARTICLE IS OUTDATED; SEE AN UPDATED 67 MARKER ARTICLE IN http://www.mv.helsinki.fi/home/jphakkin/N1c1.xps

14 **11** 13 13 16 23 ***12*** 14 13 (Finn.E)

1b1. DYS389b = 16 > 17
14 **11** 13 13 **17** 23 **11** 14 13 (Finn.E, Chukchi)

1b1a. DYS390 = 23 > 24 (ANCIENT KARELIAN)
14 **11** 13 13 **17** **24** **11** 14 13 (Finn.E, Kar)
14 **11** 13 **14** **17** **24** **11** 14 13 (Finn.E)
14 **11** 13 13 **17** **24** **11** 14 **14** (Finn.E, Kar)
14 **11** **14** 13 **17** **24** **11** 14 13 (Finn.E)

1b2. DYS393 = 13 > 14 (BALTIC SEA)
14 **11** 13 13 16 23 **11** 14 **14** (Est, Finn, Kar, Finn.Österb)
14 ***10*** 13 13 16 23 **11** 14 **14** (Finn.Österb)
14 **11** 13 13 **15** 23 **11** 14 **14** (Swe, Finn.W)
14 **11** 13 13 16 **24** **11** 14 **14** (Finn, Kar, Swe, Est, Komi.Izh)
14 **11** **12** 13 16 23 **11** 14 **14** (Finn, Kar, Est)
14 **11** 13 13 **17** 23 **11** 14 **14** (Kar, Finn.W, Saami.FT)

1b3. DYS389-I = 13 > 14
14 **11** 13 **14** 16 23 **11** 14 13 (Mansi, Swe.Värm, Saami.Swe)
14 **11** 13 **14** 16 23 *10* 14 13 (Rus.Now)
14 **11** 13 **14** 16 **24** **11** 14 12 (Finn.E)

1b4. DYS19 = 14 > 15
15 **11** 13 13 16 23 **11** 14 13 (Saami.swe, Saami.FT, Finn.E)
15 **11** 13 13 16 **24** **11** 14 13 (Finn.E)

1b5. DYS391 = 10 > 11 (corrected: former 1c1a)
14 **11** **14** 13 16 23 **11** 14 13 (REC.)
14 **11** **14** 13 16 23 **11** **15** 13 (Estonian)
14 **11** **14** 13 16 23 **11** 14 **14** (Finn.W, Est, Lat, Swe)
13 **11** **14** 13 17 23 **11** 14 **14** (Saami.Swe)
15 **11** **14** 13 16 23 **11** 14 **14** (Saami.FT [1])
14 **11** **14** 13 **17** 23 **11** 14 **14** (Swe)

1c. DYS385b = 13 > 14 (VOLGA-URAL)
14 **11** **14** 14 16 23 10 14 13 (Udmurt, Finn.W)

1c1. DYS393 = 13 > 14
14 **11** **14** 14 16 23 10 14 **14** (Khanty, Buryat, Swe)
14 **11** ***15*** 14 16 23 10 14 **14** (Finn)

1c2. DYS 389-I = 13 > 14
14 **11** **14** **14** 16 23 10 14 13 (Komi.Pri, Komi.Izh, Altaian, Khanty, Kar, Rus.Iwa, Rus.Pen)
14 **11** **14** **14** 16 23 **11** 14 13 (Finn.E, Mansi)
14 **11** **14** **14** 16 **24** **11** 14 13 (Finn.E, Saami.Swe)
14 **11** ***15*** **14** 16 23 **11** 14 13 (Mari)
14 **11** **14** **14** **15** 23 **11** 14 13 (Chuvash)

1c2a. DYS393 = 13 > 14
14 **11** **14** **14** 16 23 10 14 **14** (Est)
14 **11** **14** **14** 16 **22** 10 14 **14** (Rus.Lip)
14 **11** **14** **14** 16 23 **11** 14 **14** (Est, Kar, Finn.E)

1c2a1. DYS390 = 23 > 24
14 **11** **14** **14** 16 **24** **11** 14 **14** (Est, Finn, Swe, Kar)
14 **11** **14** **14** **15** **24** **11** 14 **14** (Finn)

2. group: DYS393 = 13 > 14

14 12 13 13 16 23 10 14 **14** (REC.)
14 12 13 13 16 **24** 10 14 **14** (Rus.Vol)
13 12 13 13 **17** 23 10 14 **14** (Lit)

2a. DYS389-I = 13 > 11 (ALTAIC)
xx xx xx ***11*** 16 23 10 14 **14** (Mongol, Uyghur, Xibe)
xx xx xx ***11*** 16 23 10 14 *13* (Han Chinese)
xx xx xx ***11*** 16 23 10 14 **12** (Daur)

2a1. DYS391 = 10 > 11 (YUKAGHIR)
14 12 13 ***11*** 16 23 **11** 14 **14** (Yukaghir)
14 **11** 13 ***11*** **17** 23 **11** ***16*** **14** (Yakut)

2b. DYS391 = 10 > 11
14 12 13 13 16 23 **11** 14 **14** (Bashkir)
14 12 13 13 16 **24** **11** 14 **14** (Kar)
15 12 13 13 **17** 23 **11** 14 **14** (Hung)

2c. DYS389-I = 13 > 14
14 12 13 **14** 16 23 10 14 **14** (Vepsian, Turkish ?)
14 **13** 13 **14** 16 23 10 14 **14** (Finn.W)

2b/c1. DYS 391 = 10 > 11 / DYS389I = 13 > 14 (UGRIAN)
14 12 13 **14** 16 23 **11** 14 **14** (Khanty, Rus.Vol, Finn.W)
14 12 13 **14** **17** 23 **11** 14 **14** (Khanty)
15 xx xx **14** 16 23 **11** 14 **14** (Turkish) ?

2b/c1a. DYS390 = 23 > 24
14 12 13 **14** 16 **24** **11** 14 **14** (Finn, Swe, Kar)
14 12 13 **14** 16 **24** ***12*** 14 **14** (Finn.E)
14 12 13 **14** **17** **24** **11** 14 **14** (Finn.W)

2b/c1b. DYS19 = 14 > 13, DYS285a = 12 > 13
13 **13** 13 **14** 16 23 **11** 14 **14** (Finn.E)

3. group: DYS19 = 14 > 15, DYS389I = 13 > 14

15 12 13 **14** 16 23 10 14 13 (REC.)

3a. DYS385b = 13 > 14 (SAYAN 1)
15 12 **14** **14** 16 23 10 14 13 (REC.)
15 12 **14** **14** 16 23 10 14 **14** (Tuva)
15 12 **14** **14** 16 23 **11** 14 13 (Tuva)
15 12 **14** **14** 16 23 ***12*** 14 13 (Finn.E)

3b. DYS391 = 10 > 11
15 12 13 **14** 16 23 **11** 14 13 (Rus.Kursk)
15 **13** 13 **14** 16 23 **11** 14 13 (Lit, Finn.W)

3c. DYS385b = 13 > 12 (SAYAN 2)
15 12 **12** **14** 16 23 10 14 13 (Tuva)
15 12 **12** **14** 16 23 **11** 14 13 (Tuva, Khakass, Shor)
15 12 **12** *13* 16 23 **11** 14 13 (Tuva)
13 12 **12** **14** **17** 23 10 **12** 13 (Khanty) ?

N1c can be divided to three different groups (**1, 2, 3**), and all of them can again be divided to three different subgroups (**a, b, c**). Groups **2** and **3** can be separated clearly to the European (**2b, 2c, 3b**) and Asian subgroups (**2a, 3a, 3c**), while in the group **1** the situation is more complicated. Group **3** is somewhat mysterious, as its founder haplotype has not (yet) been found anywhere, even though there have occurred two different mutations.

Subgroup **1a** is mostly European, but there is one lineage including haplotypes found also in Asia: **1a, 1a3, 1a3a, 1a3a2** and **1a3a3**. It is possible that **1a, 1a3, 1a3a** and **1a3a3** have spread together from Europe to Asia or *vice versa*, and **1a3a2** has later been born in Southern Siberia from **1a3a**. Group **1a+** (= including subgroups) seems to present the earliest N1c expansion to the Baltic Sea region, due to its wide distribution: it is found in Uralic, Scandinavian, Baltic and some Slavic peoples.

Subgroup **1b** is mainly restricted to Finnic and Saamic peoples, and subgroup **1c** is a bit more eastern, concentrated in Volga-Ural region, although it has scarcely spread to Scandinavia. All the subgroups of **1** are mostly European, but the founder haplotype is only found within the Khakass of Southern Siberia. Judging by distribution, haplogroup N1c may have started to expand somewhere near the Ural mountains: groups **2** and **3** seem to have split there, and part of both groups spread to the east, part to the west. Group **1** spread mainly to the west, but **1a3+** has later spread from Europe to Asia or from Asia to Europe (in which case it first spread to the east).

ATTENTION! THIS ARTICLE IS OUTDATED; SEE AN UPDATED 67 MARKER ARTICLE IN http://www.mv.helsinki.fi/home/jphakkin/N1c1.xps

3. Finnish N1c

	FinnE	SKar	NKar	NSavo	NOB	**FinnW**	SOB	SSOB	Satak	Häme	FinPr
1a	**5** 2,3 %	1 3 %		1 1,2 %	3 3,6 %	**13** 13,7 %		2 20 %	3 23,1 %	1 3,7 %	6 20,7 %
1a1	**0**					**1** 1,1 %		1 10 %			
1a2	**0**					**6** 6,3 %			4 30,8 %		2 6,9 %
1a2a	**1** 0,5 %			1 1,2 %		**1** 1,1 %				1 3,7 %	
1a2b	**4** 1,9 %	1 3 %		3 3,6 %		**2** 2,1 %			1 7,7 %		1 3,4 %
1a2b1	**7** 3,2 %	1 3 %	2 13,3 %	2 2,4 %	2 2,4 %	**2** 2,1 %	1 6,7 %			1 3,7 %	
1a3	**17** 7,9 %		2 13,3 %	12 14,3 %	3 3,6 %	**8** 8,4 %	2 13,3 %		2 15,4 %	2 7,4 %	2 6,9 %
1a3a	**11** 5,1 %	2 6,1 %	1 6,7 %	5 6 %	3 3,6 %	**9** 9,5 %	1 6,7 %	2 20 %		3 11,1 %	3 10,3 %
1a3a4	**122** 56,5 %	21 63,6 %	3 20 %	47 56 %	51 60,7 %	**16** 16,8 %	4 26,7 %	3 30 %		7 25,9 %	2 6,9 %
1a4	**0**					**2** 2,1 %	1 6,7 %				1 3,4 %
1a+ TOTAL	**167** 77,3 %	**26** 78,8 %	**8** 53,3 %	**71** 84,5 %	**62** 73,8 %	**60** 63,2 %	**10** 66,7 %	**7** 70 %	**10** 76,9 %	**15** 55,6 %	**17** 58,6 %
1b	**2** 0,9 %	1 3 %			1 1,2 %	**3** 3,2 %			2 15,4 %	1 3,7 %	
1b1	**1** 0,5 %			1 1,2 %		**0**					
1b1a	**4** 1,9 %	2 6,1 %	1 6,7 %		1 1,2 %	**0**					
1b2	**19** 8,8 %	1 3 %	6 40 %	5 6 %	7 8,3 %	**17** 17,9 %	3 20 %	2 20 %	1 7,7 %	7 25,9 %	4 13,8 %
1b3	**5** 2,3 %	1 3 %		1 1,2 %	3 3,6 %	**0**					
1b4	**3** 1,4 %			2 2,4 %	1 1,2 %	**1** 1,1 %					1 3,4 %
1c1a	**0**					**2** 2,1 %					2 6,9 %
1b+ TOTAL	**34** 15,7 %	**5** 15,2 %	**7** 46,7 %	**9** 10,7 %	**13** 15,5 %	**23** 24,2 %	**3** 20 %	**2** 20 %	**3** 23,1 %	**8** 29,6 %	**7** 24,1 %
1c	**0**					**1** 1,1 %					1 3,4 %
1c1	**1** 0,5 %				1 1,2 %	**0**					
1c2	**4** 1,9 %	1 3 %			3 3,6 %	**0**					
1c2a	**2** 0,9 %			2 2,4 %		**0**					
1c2a1	**4** 1,9 %				4 4,8 %	**5** 5,3 %	3 20 %			1 3,7 %	1 3,4 %
1c+ TOTAL	**11** 5,1 %	**1** 3 %	**0**	**2** 2,4 %	**8** 9,5 %	**6** 6,3 %	**3** 20 %	**0**	**0**	**1** 3,7 %	**2** 6,9 %
2a1	**0**					**1** 1,1 %				1 3,7 %	
2b/c1	**4** 1,9 %	1 3 %		2 2,4 %	1 1,2 %	**1** 1,1 %				1 3,7 %	2 6,9 %
2c	**0**					**2** 2,1 %				1 3,7 %	1 3,4 %
TOTAL	**216 =**	**33**	**15**	**84**	**84**	**95 =**	**15**	**10**	**13**	**27**	**29**

3.1. Greater groups

Next survey of the internal relations of Finland is based on Lappalainen et al. 2006, where there were studied four eastern (SKar = Southern Karelia, NKar = Northern Karelia, NSavo = Northern Savolax, NOB = Northern Ostrobothnia) and five western regions (SOB = Southern Ostrobothnia, SSOB = Swedish-speaking Ostrobothnia, Satak = Satakunta, Häme = Häme, FinPr = Finland Proper or South-western Finland).

By far the most frequent N1c group in Finland is **1a3a4** ”Gulf of Finland”, which is present at 25–64 % elsewhere, but only 7 % in Finland Proper and totally lacking in Satakunta. Frequency in Northern Karelia (only 20 %) reminds those of Western Finland. This group may – on the distributional basis – be connected to the Proto-Finnic expansion. Subgroup **1a3a4b** with the mutation DYS390 = 24 > 25 is only found within Eastern Finnish (only in Northern Savolax and Northern Ostrobothnia), plus one Karelian and one Swedish (Värmland, most probably descendant of a forest Finn from Savolax) occurrence. This subgroup can be called ”Savonian” (though it has not been counted separately in the table above) as it can be connected to the Savonian expansion from the late 15th century onwards.

The second group is **1b2** (”Baltic Sea”), which reaches the frequency of 19 % in Western Finland and 9 % in Eastern Finland, but the highest frequence is 40 % in Northern Karelia, probably reflecting the ancient Karelian colonisation or genetic drift. Satakunta is closer to the lower frequencies of Eastern Finland. This group has probably spread to Finland quite early, because it is the only group present in all the regions studied.

The third group in Western Finland is **1a** ”Fennoscandia” (14 %). **1a** is present only at the frequency of few percents in Eastern Finland, as well as in the inland of Western Finland: this is clearly a coastal group, reaching its highest frequency (over 20 % of all N1c) in Satakunta, Swedish-speaking Ostrobothnia and Finland Proper. This group is present in coastal Sweden (Uppsala, Västerbotten) too, and thus probably reflects the contacts over the Gulf of Bothnia. This group has a wider total distribution – being present only in low frequencies in Estonia and Northern Russia and lacking in Latvia and Lithuania – but the subgroup **1a1** ”Gulf of Bothnia” (DYS19 = 14 > 13) is found only in the Gulf of Bothnia region (Sweden: Skaraborg, Eastern Götland; Finland: Swedish-speaking Ostrobothnia) and is clearly of Swedish origin.

The third group in Eastern Finland is **1a3** (8 %), shared with Saami and some eastern populations (Russian, Buryat, Tuvan). It has similar frequencies in Western Finland (lacking only in swedish-speaking Ostrobothnia), where it is the fifth biggest group. Subgroup **1a3a** ”Most expansive” is the fourth biggest group in Western Finland (10 %) and Eastern Finland (5 %). This group can – for most part at least – be connected to the Swedish expansion, as the frequency in swedish-speaking Ostrobothnia is even 20 %. This group also includes the Rurikid haplotype and is present in Sweden, Karelia, Russia and Ukraine, where the Varangians were active, which also points to its Scandinavian origin here in the west.

There might be a separate Southern Siberian lineage (group **1** within Khakass, **1a** within Altaian, **1a3** within Tuvan and Buryat), or perhaps all these ancestral groups (**1**, **1a** and **1a3**) have spread widely along with the expansion of **1a3a** ”Most expansive”. At this point it is impossible to determine the original homeland of **1a3a**, because the above mentioned ancestral groups are also wide-spread.

3.2. Lesser groups

Group **1a2** ”Eastern Baltic Sea” has a frequency of 31 % in Satakunta while it is absent almost everywhere else in Finland (except only 7 % in Finland Proper). There is one match in Northern

Sweden (Västerbotten), but this distinct occurrence in the South-Western coast of Finland most probably originates in Baltia: this group is present in Estonia, Latvia and Lithuania. It is interesting, that this group complements the frequences of **1a3a** ”Most expansive” and **1a3a4** ”Gulf of Finland”, which were lacking in Satakunta. Subgroup **1a2b1** is clearly of Baltic origin.

Group **1b** ”Finno-Saamic” founder haplotype with the offspring has again clearly highest frequency in Satakunta (15 %).

Group **1c2a1** has the frequency of 19 % in Southern Ostrobothnia while it is only scarcely observed elsewhere. This group descends from **1c** ”Volga-Ural”.

Group **1b1a** ”Ancient Karelian” has the frequency of over 6 % in Southern and Northern Karelia and 1 % in Northern Ostrobothnia but is lacking elsewhere in Finland. As this group is also found in Karelians outside of Finland, it can be connected to ancient Karelians, whose northern expansion from the Ladoga area took place ca. 11th–12th centuries.

Group **1c1a** is present only in Finland Proper, and additionally within the Swedish and Saami, scarcely also within Estonians and Latvians.

There are only scarce occurrences of groups **2** and **3** in Finland, and these point to the rare Asian or easternmost European (Ural region) influence. Group **2b/c1** ”Ugrian” is interesting, as there are a few descendants of it both in Eastern and Western Finland. A descendant of the most remote immigrant in Finland is a solitary Tavastian (from Häme) representing group **2a1** ”Yukaghir”, otherwise found only in Eastern Siberia. There are also single haplotypes from the groups **3a** and **3b**, at least the former being clearly of Asian origin from the Altay-Sayan area.

3.3. Different Satakunta

Of all the Finnish regions, Satakunta seems to be the most peculiar judging from N1c groups. 39 % of the N1c of Satakunta originates in Baltia (31 % **1a2** ”Eastern Baltic Sea” + 8 % **1a2b**), 38 % probably in Scandinavia or coastal Finland (23 % **1a** ”Fennoscandia” + 15 % **1a3**), 23 % represents quite early eastern migration (15 % **1b** ”Finno-Saamic” + 8 % **1b2** ”Baltic Sea”).

So far Satakunta totally lacks the ”Swedish” **1a3a** and ”Finnic” **1a3a4**, but of course the situation might change when more N1c men from Satakunta are tested.

4. Uralic and non-Uralic N1c

It has often been (even among the geneticists) a tendency to predict language from the genes: haplogroup N1c has been interpreted as an evidence of Finno-Ugrian presence in the area. However, this is only circular argumentation and cannot withstand the scientific criteria. N1c has high frequencies within some Finno-Ugrian peoples, but also within some other peoples. There are no arguments to support the claim that Finno-Ugrian languages would have ever been spoken in Eastern Siberia, Mongolia, Southern Baltia or Scandinavia (before the Saamic expansion in the last 2 000 years).

The early waves of N1c to the Baltic Sea area may well have been spread along with the Combed Ware culture, because it too has a distribution from Southern Baltia to Finland and Lapland, and possibly has some similarities with the Scandinavian Pitted Ware culture. However, it has recently been argued that Uralic expansion is much later than the Combed Ware culture, which thus could not have been carried by Uralic speakers (Kallio 2006; Häkkinen 2009). At the present, it seems that both Combed Ware culture and haplogroup N1c are mostly independent of the Uralic languages.

Of course some subgroups of N1c may – for the distributive reasons – be connected to the Uralic expansion – in my analysis particularly groups **1b+** ”Finno-Saamic”, **1c+** ”Volga-Ural”, and subgroups **1a3a4** ”Gulf of Finland”, **1a3a4b** ”Savonian”, **1a4** ”Western Finnish”, **2b/c1** ”Ugrian”.

Similarly, it now seems obvious that some subgroups may be connected to some other language families: there are clearly Baltic-speaker dominated groups like **1a2a1** and **1a2b1**, Siberian Turkic like **1a3a2**, Sayan-Turkic like **3a** and **3c**, Altaic like **2a** and Yukaghir like **2a1** (Turkic-speaking Yakuts are later newcomers in the former Yukaghir area). There are even Swedish groups like **1a1** and part of **1a3a**, so it is no more possible to claim that N1c in Sweden points to the ancient presence of Uralic-speakers.

4.1. Saamic N1c

There were only few Saami N1c haplotypes available: 9 Swedish Saami from Karlsson et al. 2006, and 11 Saami from FamilyTreeDNA project. The total frequencies are: **1b2** 25 %, **1a3** 20 %, **1b4** 15 %, **1a3a4** 15 %, **1a2a1** 5 %, **1b3** 5 %, **1c2** 5 %, **1a2b1** 5 %, **1c1a** 5 %.

The greatest difference compared to the Finns is the high frequency of group **1b4**, found only within the Saami and eastern Finns (and a single occurrence in Finland Proper). Groups **1b3** and **1c2** are also present only within the Saami and eastern Finns, and these groups happen to be those to which the only two available Mansi haplotypes also belong. This may reflect eastern influence to Karelia, where the Saamic homeland is located. The frequencies of the other groups are similar to those of the Finns.

5. Age and place of origin

Recently it has been proposed that N1c is divided into two sub-branches, separated only by DYS391, its value being either 11 (in the bigger and older sub-branch) or 10 (Rootsi et al. 2007; Derenko et al. 2007). However, through the comparison of different N-subgroups it seems evident that the value 10 in this *locus* must be original and older of the two. Even if the borderline was real (which my analysis does not support) and it was really located in that particular spot, the younger branch may well still have got older age estimations.

Namely, the age of the branch (or cluster) is based on the variation observed within the branch: the more there are different allele values, the deeper is the time depth. But actually this is the case *only* if there are similar number of carriers within each branch. The occurrence of a mutation is dependent not only of the time spent, but also of the size of the population. In a population twice as big, the number of mutations may be reached in the time half of that of the other population.

Consequently, the younger branch may be more variative, if it has expanded and grown earlier than the ancestral branch. The variation of any kind (diversity, variance etc.) cannot automatically prove that the branch would be older: we must root the branches to the phylogenetic tree before we can tell which is the older one.

Mirabal et al. 2009 propose that in Europe the direction of N1c expansion could have been from west to east, based on the fact that the western populations possess higher variance levels. However, the variance (or any other kind of variation) cannot truthfully tell the relative age of the branches (see above). Secondly, it is totally misleading to even calculate any kind of variation *population-wise*: the only thing that matters is the descending of the lineage – to read more about the subject, see chapter 6 in my article on the haplogroup N1b (”Analyzing the inner structure of the Y-chromosomal haplogroup N1b” http://www.mv.helsinki.fi/home/jphakkin/N1b.pdf)

By the help of the reconstructed N1 founder haplotype we can locate the area of the deepest diversity of the haplogroup N1c – that is, the area where the most ancient mutations occurred. For the smaller groups **2** and **3** this area seems to be near the Ural mountains, as these groups have

clearly divided to European and Asian subgroups, with very little crossing over the Ural. But for the largest group **1** this area may be in Southern Siberia, as the founder haplotype of group **1** is only observed within the Khakass, as well as the very similar N1b founder haplotype. Thus Khakassia is so far the strongest candidate for the original homeland of the whole N1c. Furthermore, the haplotypes of China and Mongolia (subgroup **2a**) are not very close to the N1 founder haplotype, so there is no need to derive N1c from the area so eastern.

SOURCES

Bíró A. Z, Zalán A, Völgyi A, Pamjav H 2009: A Y-Chromosomal Comparison of the Madjars (Kazakhstan) and the Magyars (Hungary). – *American Journal of Physical Anthropology* 2009.

Cinnioğlu C, King R, Kivisild T, Kalfoğlu E, Atasoy S, Cavalleri G.L, Lillie A.S, Roseman C.C, Lin A.A, Prince K, Oefner P.J, Shen P, Semino O, Cavalli-Sforza L.L, Underhill P. 2004: Excavating Y-chromosome haplotype strata in Anatolia. – *Human Genetics* 114.

Derenko M, Malyarchuk B, Denisova G, Wozniak M, Grzybowski T, Dambueva I, Zakharov I 2007: Y-chromosome haplogroup N dispersals from south Siberia to Europe. – *Journal of Human Genetics* 52.

Häkkinen, Jaakko 2009: Kantauralin ajoitus ja paikannus: perustelut puntarissa. [The timing and locating of Proto-Uralic: weighing the arguments.] – *Journal de la Société Finno-Ougrienne* 92. http://www.sgr.fi/susa/92/hakkinen.pdf

Kallio, Petri 2006: Suomen kantakielten absoluuttista kronologiaa. [Absolute chronology of the proto-languages of Finnish.] – *Virittäjä* 1/2006. http://www.kotikielenseura.fi/virittaja/hakemistot/jutut/2006_2.pdf

Karlsson A. O, Wallerström T, Götherström A, Holmlund G 2006: Y-chromosome diversity in Sweden – A long-time perspective. – *European Journal of Human Genetics* 14.

Lappalainen T, Koivumäki S, Salmela E, Huoponen K, Sistonen P, Savontaus M-L, Lahermo P. 2006: Regional differences among the Finns: A Y-chromosomal perspective. – *Gene* 376.

Lappalainen T, Laitinen V, Salmela E, Andersen P, Huoponen K, Savontaus M-L, Lahermo P. 2008: Migration Waves to the Baltic Sea Region. – *Annals of Human Genetics* 72.

Mirabal S, Regueiro M, Cadenas A.M, Cavalli-Sforza L.L, Underhill P.A, Verbenko D.A, Limborska A.A. and Herrera R.J. 2009: Y-Chromosome distribution within the geo-linguistic landscape of northwestern Russia. – *European Journal of Human Genetics* 17.

Pakendorf B, Novgorodov I.N, Osakovskij V.L, Danilova A.P, Protod'jakonov A.P, Stoneking M 2006: Investigating the effects of prehistoric migrations in Siberia: genetic varioation and the origins of Yakuts. – *Human Genetics* 120.

Pimenoff V.N, Comas D, Palo J.U, Vershubsky G, Kozlov A and Sajantila A. 2008: Northwest Siberian Khanty and Mansi in the junction of West and East Eurasian gene pools as revealed by uniparental markers. – *European Journal of Human Genetics* 16.

Rootsi S, Zhivotovsky L.A, Baldovic M, Kayser M, Kutuev I.A, Khusainova R, Bermisheva M.A, Gubina M, Fedorova S.A, Ilumäe A-M, Khusnutdinova E.K, Voevoda M.I, Osipova L.P, Stoneking M, Lin A.A, Ferak V, Parik J, Kivisild T, Underhill P.A, Villems R. 2007: A counter-clockwise northern root of the Y-chromosome haplogroup N from Southeast Asia towards Europe. – *European Journal of Human Genetics* 15.

Völgyi A, Zalán A, Szvetnik E, Pamjav H 2008: Hungarian population data for 11 Y-STR and 49 Y-SNP markers. (Announcement of Population data). Genetics 3.

Xue Y, Zerjal T, Bao W, Zhu S, Shu Q, Xu J, Du R, Fu S, Li P, Hurles M.E, Yang H, Tyler-Smith C. 2006: Male Demography in East Asia: A North–South Contrast in Human Population Expansion Times. – *Genetics* 172.