

ARTICLE

# Y-chromosome diversity in Sweden – A long-time perspective

Andreas O Karlsson<sup>\*1</sup>, Thomas Wallerström<sup>2</sup>, Anders Götherström<sup>3</sup> and Gunilla Holmlund<sup>1</sup>

<sup>1</sup>The National Board of Forensic Medicine, Department of Forensic Genetics and Forensic Toxicology, Linköping, Sweden; <sup>2</sup>Department of Archaeology, University of Tromsø, Norway; <sup>3</sup>Department of Evolutionary Biology, Uppsala University, Sweden

Sixteen Y-chromosomal binary markers and nine Y-chromosome short tandem repeats were analyzed in a total of 383 unrelated males from seven different Swedish regions, one Finnish region and a Swedish Saami population in order to address questions about the origin and genetic structure of the present day population in Sweden. Haplogroup I1a\* was found to be the most common haplogroup in Sweden and accounted, together with haplogroups R1b3, R1a1 and N3, for over 80% of the male lineages. Within Sweden, a minor stratification was found in which the northern region Västerbotten differed significantly ( $P < 0.05$ ) from the other Swedish regions. A flow of N3 chromosomes into Västerbotten mainly from Saami and Finnish populations could be one explanation for this stratification. However, the demographic history of Västerbotten involving a significant male absence during the 17th Century may also have had a large impact. Immigration of young men from elsewhere to Värmland at the same time, can be responsible for a similar deviation with I1a\* haplotypes. Y chromosomes within haplogroup R1b3 were found to have the highest STR variation among all haplogroups and could thus be considered to be one of the earliest major male lineages present in Sweden. Regional haplotype variation, within R1b3, also showed a difference between two regions in the south of Sweden. This can also be traced from historical time and is visible in archaeological material. Overall this Y chromosome study provides interesting information about the genetic patterns and demographic events in the Swedish population.

*European Journal of Human Genetics* (2006) 14, 963–970. doi:10.1038/sj.ejhg.5201651; published online 24 May 2006

**Keywords:** Y chromosome; Sweden; demography; neolithisation

## Introduction

Fennoscandia may have been populated as early as 130 000 years ago<sup>1</sup> but anatomically modern humans appeared in the area after the last glacial maximum. For example, there are human remains from the Swedish west coast that are at least 10 200 years old<sup>2</sup> and there may have been humans in Scandinavia even earlier.<sup>3</sup> Several recent radiocarbon dates

also indicate human presence in northern Fennoscandia more than 9000 calendar years ago.<sup>4</sup> Thus, northern and southern Scandinavia have settlement continuity since the hunter-gatherers colonized the peninsula.

Prehistoric and historic demographic events are visible in archaeological and historic source material, as well as in current population genetic patterns. Y-chromosome DNA, for example, has proven to be a useful tool in phylogeography.<sup>5</sup> Once a mutation occurs on the Y chromosome, it is a slow process for it to spread in the population. This is due to the fact the Y chromosome is, for the most part, non-recombining and only passes through paternal lineages, leading to less effective population size. By combining the slow mutating binary markers with the

\*Correspondence: AO Karlsson, The National Board of Forensic Medicine, Department of Forensic Genetics and Forensic Toxicology, Artillerigatan 12, SE-581 33 Linköping, Sweden.

Tel: +46 13 25 21 43; Fax: +46 13 13 60 05;

E-mail: andreas.karlsson@rmv.se

Received 15 December 2005; revised 27 March 2006; accepted 13 April 2006; published online 24 May 2006

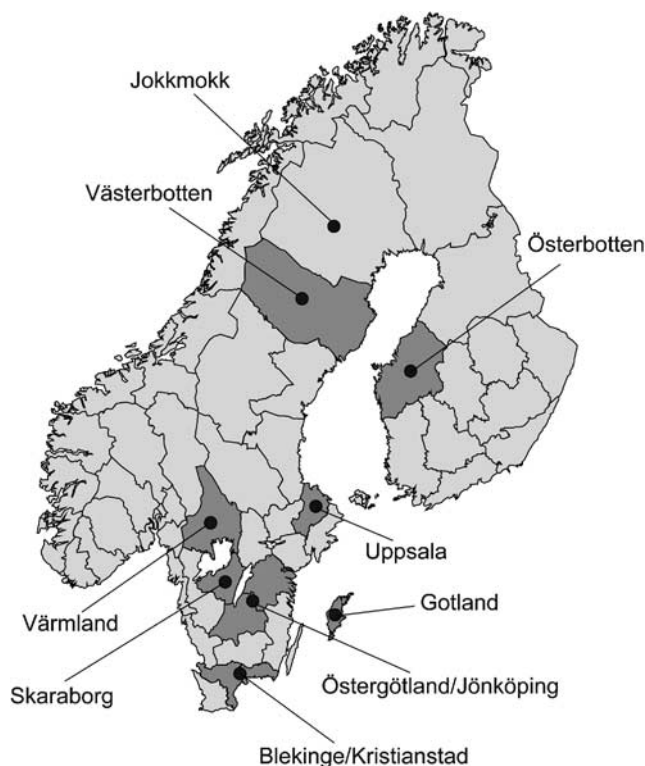
faster mutating short tandem repeats (STRs) information about major events in the past can thus be obtained.

Data available on Y chromosomes in Sweden, based on a limited number of markers, indicate little or no variation between different regions.<sup>6</sup> Other studies suggest that there are regional differences in Sweden concerning Y-STR variation.<sup>7</sup> In order to further investigate the Y-chromosomal variation in Sweden and to apply it to Scandinavian demographic history, we have used additional markers and methods well suited for comparing the evolutionary information on Y chromosomes from different geographical regions. This information was related to available historical and archaeological data.

## Materials and methods

### Population samples

DNA samples were collected from 305 unrelated Swedish males from seven regions in Sweden of various geographical locations (Figure 1). The regions were represented by administrative provinces (Swedish län) and were as follows: Gotland ( $n=40$ ), Uppsala (55), Östergötland/Jönköping (41), Blekinge/Kristianstad (41), Skaraborg (45), Värmland (42) and Västerbotten (41). In this way, northern, southern, coastal and inland regions were accounted for, as well as an island in the Baltic Sea.



**Figure 1** Map showing the regions in which the males included in this study were born.

The sampled areas are representative for contemporary demography as well as major regions indicated in the archaeological material.<sup>8</sup> Most of them are situated far apart from each other; hence the DNA samples can be expected to give information about seven contrasting demographic histories. Some examples: In 1571, Västerbotten län is estimated to have had about 0.1 inhabitants/km<sup>2</sup>, a striking contrast to the situation in the Kristianstad län with its 7.0, Blekinge's 4.0 or Uppsala's 5.6 (also estimations). The population increased in Västerbotten from 0.1 to 0.3 up to 1751. Two- or threefold growth is regarded as normal for this period but the Värmland case is exceptional; from 0.6 to 4.8.<sup>9</sup> Also notable is a documented absence of males in different parts of Västerbotten, due to the great wars at the end of the 16th which continued during the first decades of the 18th C.<sup>10–14</sup>

Forty samples from the Finnish region Österbotten and 38 samples from a Swedish Saami population (Jokkmokk nomads) were also analyzed only as reference populations. The Österbotten samples represent Finnish-speaking Finns as well as the Swedish-speaking minority.

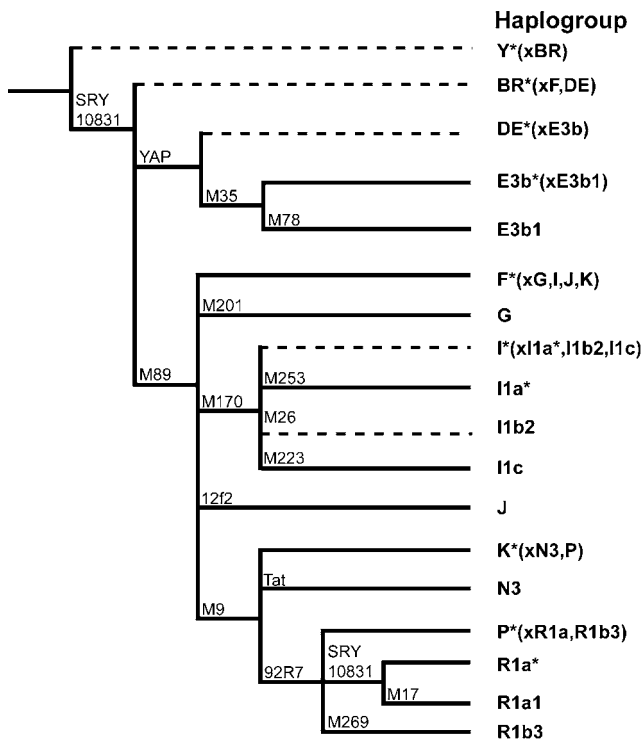
### Y-chromosomal markers

Fourteen Y-chromosome single nucleotide polymorphisms (Y-SNPs), (M9, Tat, 92R7, M17, M35, M78, M89, M201, M170, M26, M223, SRY10831, M253 and M269), an Alu insertion polymorphism (YAP) and an insertion/deletion polymorphism (12f2) were selected for the haplogroup determination (Figure 2). The Y-SNPs were typed using Pyrosequencing technique, while the two length polymorphisms were analyzed using agarose gel electrophoresis (Supplementary Table 1). All samples were typed for the markers SRY10831, M89, M9 and 92R7. Depending on the results, samples were further analyzed for additional markers in a hierarchical approach. The haplogroup nomenclature used is according to the recommendation of the Y Chromosome Consortium.<sup>15</sup>

Information on the Y-STR markers DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393 and DYS385 was already available for the 305 Swedish males.<sup>16</sup> However, in the present study we also typed the Saami and Österbotten samples and separated DYS385 into DYS385a and DYS385b for all 383 chromosomes. This was performed according to Kittler *et al*,<sup>17</sup> with the minor modification that 10 ng template DNA was used.

### Statistical analysis

An exact test for population differentiation<sup>18</sup> was performed using haplogroup frequencies for the different regions. One hundred thousand Markov chain steps were used for the test of significance. The haplogroup frequencies from one region were compared with frequencies of all the other regions combined. Principal component analysis (PCA) was carried out on haplogroup frequencies and



**Figure 2** Y-chromosomal haplogroups defined by the 16 binary markers used. The solid lines represent haplogroups found in the study while the dashed lines are haplogroups not detected in the sample.

Y-STR based  $R_{ST}$  values using the software MATLAB 7.0 (MathWorks).

Arlequin package version 2.000,<sup>19</sup> was used to calculate haplotype diversities and also  $R_{ST}$  values based on all nine Y-STRs.  $P$  values for  $R_{ST}$  were calculated using 10000 permutations.  $R_{ST}$  was calculated between all regions, as well as between one region and all the other regions together. For the calculation of  $R_{ST}$ , a weighting procedure<sup>20</sup> was used to compensate for the different mutation rates at the individual Y-STR loci.<sup>21,22</sup>

Median Joining networks were constructed by NETWORK 4.0.<sup>23</sup> Networks were created from haplogroup specific haplotypes with a five fold range weighting scheme due to different mutation rates among the markers. The scheme was based on allele variation at individual loci.

Estimation of time to most recent common ancestor (TMRCA) was carried out by means of BATWING.<sup>24</sup> This software uses a Bayesian coalescent approach for calculation of TMRCA. The mutation priors were constructed as gamma distribution based on mutations rates.<sup>21,22</sup> The demographic model was chosen to be an exponential growth from a constant ancestral population size. A generation time of 30 years<sup>25</sup> was assumed and the 95% confidence intervals take uncertainties such as mutation rates and population growth into account.

For all statistical analyses, the repeat length of DYS389I was subtracted from DYS389II, into DYS389AB, to give the true length of the individual loci.<sup>26</sup>

## Results Haplogroups

The set of the 16 different Y binary markers identifies 18 haplogroups in theory (Figure 2). In the Swedish population, however, only 13 different haplogroups were found (Table 1). Four major haplogroups (I1a\*, R1b3, R1a1 and N3) accounted for 80% of the Swedish male lineages. The most common haplogroup was I1a\*, to which 37% of the male lineages belonged. I1a\* was the most frequent haplogroup in five (Gotland, Uppsala, Blekinge/Kristianstad, Värmland and Västerbotten) of the seven Swedish regions, while R1b3 was most common in the remaining two (Skaraborg and Östergötland/Jönköping). In both Östergötland and among the Saamis, N3 was found to be the haplogroup with the highest frequency.

A Principal component analysis (PCA), based on haplogroup frequencies, revealed that Västerbotten differed from the other Swedish regions (Figure 3). This deviation was further supported by an exact test for population differentiation (Table 2), in which Västerbotten was compared with all the other Swedish subpopulations combined ( $P < 0.01$ ). Haplogroup diversities also showed that Västerbotten (0.86) differed from the other regions (0.73–0.80).

## Haplotypes

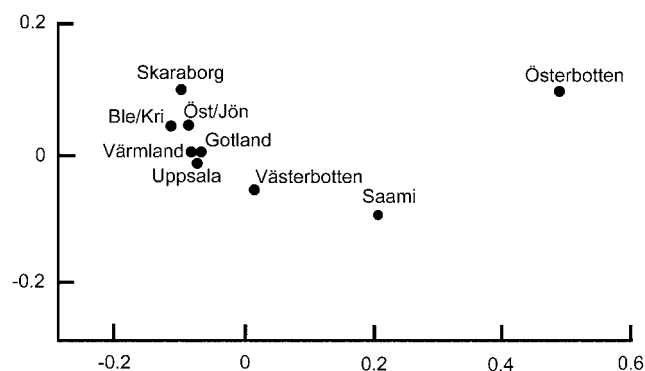
The data from the Y-STR markers were used to resolve the dataset further (Supplementary Table 2). Nine Y-STR loci, including the separation of DYS385a and DYS385b, revealed 228 distinctive haplotypes among the 305 Swedish Y chromosomes, which represents a haplotype diversity of 0.995.

Pairwise  $R_{ST}$  values with all nine STRs were compared (Figure 4) and showed, some significant differences between a few of the Swedish regions, mostly involving Skaraborg or Västerbotten (data not shown). Since only a few subpopulations showed such differences, new  $R_{ST}$  values were calculated in which one region was compared with the haplotypes of all the other Swedish regions. In this new test, both Skaraborg and Västerbotten differed ( $P < 0.05$ ) from the other Swedish regions (Table 2). Furthermore, of all the Swedish regions, Västerbotten showed the shortest distance to the Saami population,  $R_{ST} = 0.03$ ,  $P < 0.05$ , while  $R_{ST}$  was 0.04–0.08 ( $P < 0.05$  in all cases) for the other six Swedish regions. This pattern was also the same for the distance to the Östergötland sample,  $R_{ST} = 0.15$  for Västerbotten and  $R_{ST} = 0.17–0.23$  for the remaining Swedish regions ( $P < 0.01$  in all cases).

**Table 1** Haplogroups found among 305 Swedish males, 40 Österbotten males and 38 Saami males

Population (N <sup>a</sup> )	Number of individuals in each haplogroup												
	E3b1	E3b*(xE3b1)	F*(xG,I,J,K)	G	I1a*	I1c	J	K*(xN3,P)	N3	P*(xR1a,R1b3)	R1a1	R1a*	R1b3
Gotland (40)	1	0	0	0	18	2	0	2	4	2	5	0	6
Uppsala (55)	0	0	0	0	21	0	3	1	8	1	6	1	14
Östergötland/ Jönköping (41)	0	1	0	0	12	1	4	0	3	1	4	2	13
Blekinge/Kristianstad (41)	0	0	1	2	17	2	0	1	1	1	7	0	9
Skaraborg (45)	1	0	0	0	16	1	2	0	1	3	4	0	17
Värmland (42)	0	0	0	1	19	2	2	0	4	1	6	0	7
Västerbotten (41)	1	0	1	2	10	7	0	1	8	1	4	0	6
<b>Sweden as a whole (305)</b>	<b>3</b>	<b>1</b>	<b>2</b>	<b>5</b>	<b>113</b>	<b>15</b>	<b>11</b>	<b>5</b>	<b>29</b>	<b>10</b>	<b>36</b>	<b>3</b>	<b>72</b>
Swedish Saami (38)	0	0	0	0	12	0	0	0	17	0	6	0	3
Österbotten (40)	0	0	0	2	8	0	0	0	26	0	3	0	1

<sup>a</sup>Number of Y-chromosomes analysed.



**Figure 3** Principal component analysis on  $F_{ST}$  distances based on haplogroup frequencies in the Swedish subpopulations. The first two axes account for 91% of the total variance. Öst/Jön stands for Östergötland/Jönköping and Ble/Kri for Blekinge/Kristianstad.

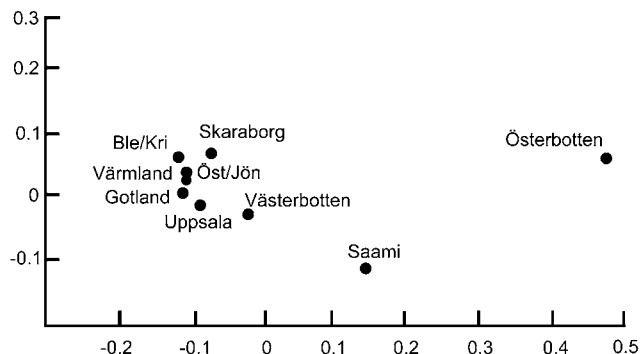
**Table 2** Tests for population differentiation based on haplogroup frequencies and haplotype data

Population <sup>a</sup>	Exact test for population differentiation <sup>b</sup>	$R_{ST}$ value ( $P$ -value) <sup>c</sup>
Gotland	0.78	-0.008 (0.86)
Uppsala	0.67	-0.003 (0.67)
Östergötland/ Jönköping	0.13	-0.002 (0.55)
Blekinge/ Kristianstad	0.35	-0.001 (0.50)
Skaraborg	0.30	0.016 (0.04)
Värmland	0.97	0.0003 (0.43)
Västerbotten	0.003	0.017 (0.03)

<sup>a</sup>The Y-chromosomes from one region were compared with all chromosomes from the other Swedish regions combined.

<sup>b</sup>Exact test for population differentiation was performed on haplogroup frequencies.  $P$ -values are given in the table.

<sup>c</sup> $R_{ST}$  values based on Y-STRs DYS19, DYS385a, DYS385b, DYS389I, DYS389AB, DYS390, DYS391, DYS392 and DYS393.



**Figure 4** Principal component analysis on pairwise  $R_{ST}$  values based on STR haplotypes. The first two axes account for 90% of the total variance. Öst/Jön stands for Östergötland/Jönköping and Ble/Kri for Blekinge/Kristianstad.

### Haplogroup specific haplotype diversities

$R_{ST}$  analysis of STR haplotypes in I1a\* chromosomes provided significant distances between Värmland and two other Swedish regions (Värmland-Gotland,  $R_{ST}=0.07$ ,  $P<0.05$ , Värmland-Skaraborg,  $R_{ST}=0.1$ ,  $P<0.05$ ). For STR haplotypes within I1a\*, Värmland was also the only Swedish region that showed significant distances to the Saami ( $R_{ST}=0.10$ ,  $P<0.05$ ) and Österbotten ( $R_{ST}=0.13$ ,  $P<0.05$ ). The I1a\* haplotypes from Värmland were also compared with 83 I1a\* haplotypes from three neighbouring Norwegian regions (Östfold, Akershus and Hedmark; BM Dupuy, personal communication.). No closer relationship was detected between Värmland and the three Norwegian regions than between the Norwegian regions and any other of the six Swedish regions.

The twelve I1a\* chromosomes in the Swedish Saami population formed a close cluster around a modal haplotype (14-12-28-23-10-11-13-14-14, DYS19-DYS389I-DYS389II-DYS390-DYS391-DYS392-DYS393-DYS385a-DYS385b), which comprises 66% of the Saami lineages within this

**Table 3** Estimation of TMRCA based on Y-STR data for the four most common haplogroups in Sweden

Haplogroup	Number of Y chromosomes	Mode TMRCA <sup>a</sup> (95% CI) (years)
I1a*	113	6200 (2500–15 700)
N3	29	3300 (1500–8700)
R1b3	72	9100 (3300–25 000)
R1a1	36	4000 (1800–9900)

<sup>a</sup>TMRCA was calculated by means of BATWING. A generation time of 30 years was assumed.

haplogroup. The same haplotype is also the most common in Sweden.<sup>16</sup>

Eleven different haplotypes were found among the 15 samples belonging to haplogroup I1c. This haplogroup was especially common in Västerbotten and in order to deduce its origin, a search in the YHRD (Y-STR haplotype reference database, release 17) was performed. The search for the most common haplotype belonging to I1c in Västerbotten, (15-14-32-23-10-12-14-15,15) yielded hits in low frequencies around Europe but was most frequent in populations from Germany and the Netherlands (excluding the Swedish haplotypes, eight out of fifteen European haplotypes were found in Germany or the Netherlands).

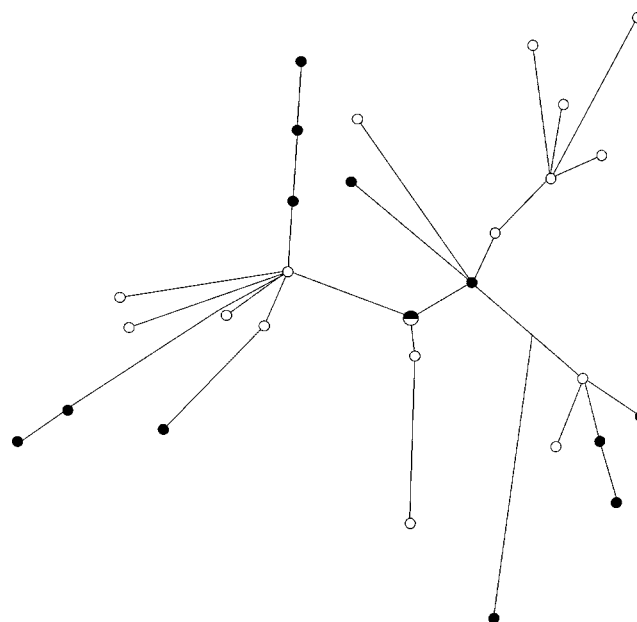
N3 chromosomes showed low haplotype diversity in Österbotten (0.75) compared with the Swedish regions (0.96–1.0) and the Saami population (0.91). Pairwise  $R_{ST}$  values showed a closer relationship between Västerbotten and Saami ( $R_{ST}=0.02$ ,  $P=0.3$ ) than between Västerbotten and Österbotten ( $R_{ST}=0.11$ ,  $P=0.03$ ).

Haplotypes belonging to R1b3 were shown to have the highest variance among the haplogroups found in Sweden. This was revealed by TMRCA analyses (Table 3), which show that the R1b3 haplotypes in Sweden have a common ancestor from around 9000 (3300–25 000) years ago. Further analysis revealed differences between the eastern and western parts of the south of Sweden (Östergötland/Jönköping and Skaraborg, respectively). This deviation was shown both by a significant  $R_{ST}$  value ( $R_{ST}=0.1$ ,  $P=0.02$ ) and a median-joining network, illustrating a divided cluster (Figure 5).

## Discussion

### Haplogroups found in Sweden

Using PCA analysis, the haplogroups found in the Swedish population sample agreed with haplogroup distributions in other European countries (data not shown). The most common haplogroup in Sweden was I1a\*, which is also present in the same frequency in Norway.<sup>27</sup> I1a\* is thought to have a decreasing gradient from Scandinavia towards both the east (Ural) and the west (Atlantic) and Western Europe has been suggested as the source of the Scandinavian I1a\*.<sup>27</sup>



**Figure 5** Median-joining network of R1b3 haplotypes from Skaraborg and Östergötland/Jönköping. The sizes of the circles are proportional to the haplotype frequencies. The open circles represent haplotypes from Skaraborg and the closed circles are haplotypes found in Östergötland/Jönköping.

Semino *et al*<sup>28</sup> concluded that haplogroups with the M170 mutation (defining haplogroup I) and the M173 mutation (defining haplogroup R1\*) have been present in Europe since the palaeolithic period, while the other haplogroups entered, independently, from the Middle East and the Urals. R1b3 and R1a1 (24 and 12% in Sweden respectively) are today two common haplogroups in the rest of Europe. It has been suggested that the mutation determining R1\*(xR1a1) originated 35 000–40 000 years ago in western Europe and that the mutation defining R1a1 arose later on in eastern Europe. Semino *et al* also suggested that these two haplogroups expanded into the central parts of the continent after the last glacial maximum. Studies of today exhibit that these two haplogroups show a gradient pattern in Europe. R1\*(xR1a1) is frequent in western Europe, with a decreasing frequency towards the east while R1a1, on the other hand, has the opposite pattern with a high frequency in the east and a decreasing occurrence in the west.<sup>28,29</sup> Sweden, which is considered to be a northern European population has frequencies similar to other Scandinavian populations (Danes, Norwegians) and has a higher frequency of the 'western' haplogroup R1\*(xR1a1) compared with haplogroup R1a1.

The fourth most frequent haplogroup in Sweden was N3 (10%). This haplogroup is mostly present in the northern Swedish regions, indicating a closer relationship with Saami and Finnish populations, in which N3<sup>30,31</sup> is very common (this is discussed further below). Tambets *et al*<sup>31</sup>

suggested that the higher diversity found in eastern Europe (compared to Siberia) would make eastern Europe a possible origin for this haplogroup. The time for its expansion remains, however, unclear.

### Regional Y-chromosome variation

There are differences in haplogroup frequencies among all the regions studied. Some may be due to the limited sample size. Anyhow, all analyses in this study provided information suggesting that Västerbotten differed from the rest of the Swedish regions. This deviation is mostly due to the high frequencies of N3 and I1c haplotypes compared to the other regions. N3 and I1c together account for 37% of the Y chromosomes in Västerbotten, but only 4–15% in the other Swedish regions and, at the same time, the lower frequency of I1a\* and R1b3, which are very common in the other regions (40% together in Västerbotten but 60–73% in other Swedish regions). Furthermore, when N3 and I1c were excluded from the exact test for population differentiation the earlier significant difference disappeared.

The high frequency of N3 in Västerbotten can be explained by the short geographical distance to both the Saami and the Finnish populations. However, our study suggests a closer relationship (represented by pairwise  $R_{ST}$  values) between Saami and Västerbotten N3 haplotypes than between Västerbotten and Österbotten haplotypes. Among the Österbotten N3 Y chromosomes, 13 of 26 haplotypes were identical (14-14-30-24-11-14-14-11-13) which makes the Österbotten haplotype diversity extremely low, suggesting a high grade of endogamy or founder effects. This observation has been made in many previous studies.<sup>30,32</sup> The same haplotype is found three times among 17 Saami N3 Y chromosomes and in only one of the Västerbotten Y chromosomes. The absence of the common Finnish haplotype in Västerbotten may explain a more distant relationship to Österbotten, relative to the Swedish Saami population.

The frequency of haplogroup I1c also caused deviation of Västerbotten from the other Swedish regions. The presence of this lineage in Västerbotten is not as easy to explain. Rootsi and Magri *et al*<sup>27</sup> suggested that I1c was not present in the north of Sweden (Lapland). However, it was relatively common in German (12.5%) and Dutch (10%) population samples and it can be expected that it is present in most parts of Sweden, since there has been a more or less constant immigration from Germany to Sweden since the 13th C (at the latest) and from Holland, perhaps from the 15th C onwards.<sup>33–36</sup> Many Germans, Dutch and Scots settled in Sweden, particularly during its 'Age of Greatness' (1611–1718).

One possible reason why haplogroup I1c is present at such a high frequency in Västerbotten and not in any of the remaining six Swedish samples could be due to the local historical demography, that is the absence of men in the 17th C, as mentioned earlier. Thus, the geneflow from

Germany and Dutch populations may have had a greater impact. People with German or Dutch names were abundant in vicarages and officers residences in Västerbotten since that time onwards and among the officials or royal commissioners in northern Sweden.<sup>37</sup> The 'Age of Greatness' is considered a 'mortal but fertile era' with families having large numbers of children and the population thus being able to override the mortality rates caused by wars, epidemics and famines.<sup>38</sup> Thus, I1c, brought in by immigrants, has persisted in the area, probably due to a founding effect in Västerbotten during the 17th C.

I1a\* is the most common haplogroup in nearly all regions in Sweden. Within this haplogroup, the regions did not show any deviation among themselves except for the I1a\* haplotypes found in Värmland. This region differed significantly from two Swedish regions and both the Saami and Österbotten I1a\* lineages. No other Swedish region differed from the Saami or the Österbotten samples. No doubt Värmland's population growth and rate of colonisation, which was outstanding between 1571 and 1751 and remaining considerable until the 1930s<sup>9,38–40</sup> compared to Sweden in general, could be a part of the explanation. In contrast to other parts of Sweden, as in the case of Västerbotten, Värmland was not affected by military conscription during the great wars. The mines and iron works, important for the war industry, attracted young workers from other parts of Sweden as well as foreigners (Germans, Wallons, Danes, Norwegians)<sup>39</sup> and the landscape was colonized (partly by Finns). Still in the High Middle Ages, Värmland was something 'between' the emerging Swedish and Norwegian kingdoms.<sup>41,42</sup>

I1a\* haplotypes in the Jokkmokk-Saami sample indicate a 'Swedish' origin (of I1a\* in Saami), since the most common Saami I1a\* STR haplotype is also a typical Swedish haplotype. This affiliation is easy to explain as a result of connections across a cultural and linguistic border between Saamis and settling Swedes at the coast around Luleå starting during the 14th C.<sup>43</sup>

Haplogroup R1b3 was shown to have the highest variation in Y-STRs among all haplogroups found in Sweden (Table 3). It could indicate that this was the first, or one of the first, major haplogroups in Sweden after the latest glacial maximum. As this haplogroup is pre-Neolithic in Europe, it may actually reveal some of the oldest demographic events in Scandinavian prehistory. R1b3 also suggests a differentiation in southern Scandinavia between east and west. Both  $R_{ST}$  values and a median joining network provided information suggesting a difference, probably as a sign of prehistoric demographic events. It is also visible in archaeological material, from the megalithic architecture onwards, to the Medieval period<sup>8</sup> and further on into recent time, in accordance with observed differences between the east and the west of Sweden that have long since been long discussed in terms of economy, religion, settlement, social structure, politics etc. It is all

about differences in degree, not in kind.<sup>44</sup> The geneflow between the eastern and western parts of southern Scandinavia, across Lake Vättern, has not been strong enough to completely erase the founder effects of the earliest settlement in these areas.

The high frequency of Paleolithic haplogroups and the low frequency of suggested Neolithic haplogroups<sup>45</sup> is indicative. The origin of a farming economy has been discussed for 80 years,<sup>46</sup> and genetic data have been used to argue for a massive migration, spreading agriculture from the Near East,<sup>47,48</sup> as well as for a more moderate migration.<sup>28,49</sup>

Anyhow, given that E3b, F\*, and G are Neolithic migrant haplogroups (4% in Sweden), and I1a\*, R1b3 and R1a1 Paleolithic haplogroups (73% in Sweden), our data do not correlate with the agricultural spreading to northern Europe with migrating Anatolian farmers.<sup>28</sup> Furthermore, in Scandinavia there is no evidence for a swift replacement of the hunter-gatherer economy,<sup>2</sup> and there are also indications that the local wild fauna was used in the northern European domestic stock to a greater extent than in southern Europe.<sup>50</sup> Thus, we believe that our data indicate population continuity, acculturation and acceptance of new ideas rather than migration and population replacement in the Mesolithic-Neolithic transition. They also underline that migration is not a necessary prerequisite of cultural change, often presumed.<sup>51</sup> Anyhow, we have to conclude that emerging agriculture was an introduction of ideas and/or a result of immigration not visible in our Y-chromosome data, i.e. as for example female immigration.

Although not all our questions were answered by our Y-chromosome data, they represent an important tool for use in population genetics. This study of the Y-chromosome variation in several Swedish regions provides interesting information about genetic patterns in existence today but derived from demographic events in the past.

### Acknowledgements

We thank Dr Pertti Sistonen at the Finnish Red Cross Blood Transfusion Service for the collection of the Österbotten samples, Dr Minttu Hedman and Dr Antti Sajantila at the Department of Forensic Medicine, University of Helsinki for general help with the Österbotten samples. We also thank the late Professor Lars Beckman at the University of Gotland and Dr Berit Myhre Dupuy at the Institute of Forensic Medicine, University of Oslo for the collection of the Saami samples and the Norwegian data, respectively. This study was partly funded by The Board for Research and Development, the National Board of Forensic Medicine, Sweden (grant no 55217), the Swedish National Laboratory of Forensic Science and the Forensic Science Centre at Linköping University.

### Electronic databases

Y-STR haplotype reference database: <http://www.yhrd.org>

### References

1 Schulz HP: The lithic industry from layers IV–V, Susiluola Cave, Western Finland, dated to the Eemian interglacial. *Préhistoire Européenne* 2002; **16–17**: 7–23.

- 2 Nordqvist B: Coastal adaptations in the Mesolithic. A study of coastal sites with organic remains from the Boreal and Atlantic periods in Western Sweden. Appendix: Eva-Lena Larsson. GO-TARC, Series B 13 Göteborg, 2000.
- 3 Larsson L: *Segebro: En tidigatlantisk boplats vid Sege ås mynning*. Malmö: Malmö mus, 1982.
- 4 Bergman I, Olofsson A, Hörnberg G, Zackrisson O, Hellberg E: Deglaciation and colonization: Pioneer settlements in northern Fennoscandia. *J World Prehist* 2004; **18**: 155–177.
- 5 Jobling MA, Tyler-Smith C: The human Y chromosome: an evolutionary marker comes of age. *Nat Rev Genet* 2003; **4**: 598–612.
- 6 Holmlund G, Nilsson H, Engdahl L, Götherström A, Lindblom B: No obvious geographical Y-chromosome gradient in the Swedish population; in Douthett C, Morling N (eds): *Progress in Forensic Genetics 10*, International Congress Series, Elsevier: Amsterdam, The Netherlands, 2004, vol 1261: 357–359.
- 7 Roewer L, Croucher PJ, Willuweit S *et al*: Signature of recent historical events in the European Y-chromosomal STR haplotype distribution. *Hum Genet* 2005; **116**: 279–291.
- 8 Rysenstrand Å: *Arkeologisk regionindelning av Sverige*. Stockholm: Hysantikvarieämbetet, 1979.
- 9 Palm LA: *Folkmängden i Sveriges socknar och kommuner 1571–1997. Med särskild hänsyn till perioden 1571–1751*. Göteborg: LA Palm, 2000.
- 10 Westin JG: Bygden växer; in Fahlgren K (eds): *Skellefteå sockens historia I*. Uppsala: Almqvist & Wiksell, 1953; Vol 1.
- 11 Bylund E: *Koloniseringen av Pite lappmark t.o.m. 1867*. Uppsala: Geographica, 1956; Vol 30.
- 12 Lindegren J: *Utskrivning och utsugning. Produktion och reproduktion i Bygdeå 1620–1640*. Uppsala: Studia Historica Upsaliensia, 1980; Vol 117.
- 13 Eneqvist G: *Nedre Luledalens byar*. Uppsala: Geographica, 1937; Vol 4.
- 14 Göthe G: *Om Umeå lappmarks svenska kolonisation från mitten av 1500-talet till omkr. 1750*. Uppsala: Almqvist & Wiksell, 1929.
- 15 Y Chromosome Consortium: A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res* 2002; **12**: 339–348.
- 16 Holmlund G, Nilsson H, Karlsson A, Lindblom B: Y-chromosome STR haplotypes in Sweden. *Forensic Sci Int*, (in press).
- 17 Kittler R, Erler A, Brauer S, Stoneking M, Kayser M: Apparent intrachromosomal exchange on the human Y chromosome explained by population history. *Eur J Hum Genet* 2003; **11**: 304–314.
- 18 Raymond M, Rousset F: An exact test for population differentiation. *Evolution* 1995; **49**: 1280–1283.
- 19 Schneider P, Roesseli D, Excoffier L: *ARLEQUIN, version 2\*000: a software for population genetics data analysis*. Geneva: Genetics and Biometry Laboratory, University of Geneva, 2000.
- 20 Goodman SJ:  $R_{ST}$  Calc: a collection of computer programs for calculating estimates of genetic differentiation from microsatellite data determining their significance. *Mol Ecol* 1997; **6**: 881–885.
- 21 Dupuy BM, Stenersen M, Egeland T, Olaisen B: Y-chromosomal microsatellite mutation rates: differences in mutation rate between and within loci. *Hum Mutat* 2004; **23**: 117–124.
- 22 Kayser M, Roewer L, Hedman M *et al*: Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome, as revealed by direct observation in father/son pairs. *Am J Hum Genet* 2000; **66**: 1580–1588.
- 23 Bandelt HJ, Forster P, Sykes BC, Richards MB: Mitochondrial portraits of human populations using median networks. *Genetics* 1995; **141**: 743–753.
- 24 Wilson IJ, Balding DJ: Genealogical inference from microsatellite data. *Genetics* 1998; **150**: 499–510.
- 25 Tremblay M, Vezina H: New estimates of intergenerational time intervals for the calculation of age and origins of mutations. *Am J Hum Genet* 2000; **66**: 651–658.

- 26 Rolf B, Meyer E, Brinkmann B, de Knijff P: Polymorphism at the tetranucleotide repeat locus DYS389 in 10 populations reveals strong geographic clustering. *Eur J Hum Genet* 1998; **6**: 583–588.
- 27 Rootsi S, Magri C, Kivisild T *et al*: Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe. *Am J Hum Genet* 2004; **75**: 128–137.
- 28 Semino O, Passarino G, Oefner PJ *et al*: The genetic legacy of Paleolithic Homo sapiens in extant Europeans: A Y chromosome perspective. *Science* 2000; **290**: 1155–1159.
- 29 Rosser ZH, Zerjal T, Hurler ME *et al*: Y-chromosomal diversity in Europe is clinal influenced primarily by geographic, rather than by language. *Am J Hum Genet* 2000; **67**: 1526–1543.
- 30 Hedman M, Pimenoff V, Lukka M, Sistonen P, Sajantila A: Analysis of 16 Y STR loci in the Finnish population reveals a local reduction in the diversity of male lineages. *Forensic Sci Int* 2004; **142**: 37–43.
- 31 Tambets K, Rootsi S, Kivisild T *et al*: The western and eastern roots of the Saami- the story of genetic 'outliers' told by mitochondrial DNA and Y chromosomes. *Am J Hum Genet* 2004; **74**: 661–682.
- 32 Kittles RA, Bergen AW, Urbanek M *et al*: Autosomal, mitochondrial, and Y chromosome DNA variation in Finland: evidence for a male-specific bottleneck. *Am J Phys Anthropol* 1999; **108**: 381–399.
- 33 Kumlien K: *Tyskar. Kulturhistoriskt lexikon för nordisk medeltid*. Malmö: Allhem, 1975; Vol 19.
- 34 Yrwing H: *Hollandshandel. Kulturhistoriskt lexikon för nordisk medeltid*. Malmö: Allhem, 1961; Vol 6.
- 35 Svanberg I, Tydén M: *Tusen år av invandring. En svensk kulturhistoria*. Stockholm: Gidlund, 1992.
- 36 Sandström Å: *Mellan Torneå och Amsterdam: en undersökning av Stockholms roll som förmedlare av varor i regional- och utrikeshandel 1600–1650*. Stockholm: Monografier utgivna av Stockholms stad, 1990; Vol 102.
- 37 Steckzén B: *Birkarlar och lappar. En studie i birkarleväsendets, lappbefolkningens och skimmhandels historia*. Stockholm: Kungliga Vitterhets-, historie- och antikvitetsakademiens handlingar. Historiska serien, 1964; Vol 9.
- 38 Palm LA: *Livet, kärleken och döden: fyra uppsatser om svensk befolkningsutveckling 1300–1800*. Göteborg: LA Palm, 2001.
- 39 Palm LA: *Människor och skördar. Studier kring agrarhistoriska metodproblem 1540–1770*. Göteborg: Institutet för lokalhistorisk forskning, 1993; Vol 4.
- 40 Olausson P: *Värmlänningar i rörelse. Några anteckningar om migrationsrörelser, migrationsforskning och källor till migrationsforskningen avseende Värmland från 1500-talet fram till våra dagar, med tonvikt på emigrationen till Amerika*. Karlstad: Högsolan i Karlstad, 1997.
- 41 Gahrn L: *Värmland mellan Sveariket och Norge (I)*; in: *Werm-landica*. Karlstad: Föreningen för Värmlandslitteratur, 1994; pp 7–22.
- 42 Gahrn L: *Värmland mellan Sveariket och Norge (II)*; in: *Werm-landica*. Karlstad: Föreningen för Värmlandslitteratur, 1995; pp 93–114.
- 43 Wallerström T: *Norrbottnen, Sverige och medeltiden. Problem kring makt och bosättning i en europeisk periferi I*. Stockholm: Almqvist & Wiksell International, 1995.
- 44 Winberg C: *Hur Västsverige blev västsvenskt*. Göteborg: Humanistiska fakulteten Göteborgs Universitet, 2000.
- 45 King R, Underhill PA: Congruent distribution of neolithic painted pottery and ceramic figurines with Y-chromosome lineages. *Antiquity* 2002; **76**: 704–714.
- 46 Childe VG: *The dawn of European civilization*. London, Routledge & Kegan Paul, 1925.
- 47 Renfrew C: *Archaeology and language: the puzzle of Indo-European origins*. London: Jonathan Cape, 1987.
- 48 Cavalli-Sforza LL, Piazza A: Human genomic diversity in Europe: a summary of recent research and prospects for the future. *Eur J Hum Genet* 1994; **1**: 3–8.
- 49 Richards M, Corte-Real H, Forster P *et al*: Paleolithic and neolithic lineages in the European mitochondrial gene pool. *Am J Hum Genet* 1996; **59**: 185–203.
- 50 Götherström A, Anderung C, Hellborg L *et al*: Cattle domestication in the Near East was followed by hybridization with aurochs bulls in Europe. *Proc Biol Sci* 2005; **22**: 2345–2350.
- 51 Cassel K: *Where are we going? Attitudes toward migrations in archaeological thought*. *Current Swedish Archaeology*, 2000; **8**: 33–49.

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