Available online at www.sciencedirect.com
SCIENCE DIRECT®
Forensic Science International xxx (2005) xxx–xxx
Forensic Science International
www.elsevier.com/locate/forsciint

Geographical heterogeneity of Y-chromosomal lineages in Norway

Berit Myhre Dupuy [a,*], Margurethe Stenersen [a], Tim T. Lu [b], Bjørnar Olaisen [a]

[a] *Institute of Forensic Medicine, University of Oslo, Rikshospitalet, Oslo, Norway*
[b] *Institute of Medical Informatics and Statistics, Christian-Albrechts-University, Kiel, Germany*

Received 10 February 2005; received in revised form 2 November 2005; accepted 2 November 2005

Abstract

Y-chromosomal variation at five biallelic markers (Tat, YAP, 12f2, SRY_{10831} and 92R7) and nine multiallelic short tandem repeat (STR) loci (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS385I/II and DYS388) in a Norwegian population sample are presented. The material consists of 1766 unrelated males of Norwegian origin. The geographical distribution of the population sample reflects fairly well the population distribution around the year 1942, which is the median birth year of the index persons. Seven hundred and twenty-one different Y-STR haplotypes but 726 different lineages (Y-STRs plus biallelic markers) were encountered. We observed six known (P*(xR1a), BR(xDE, J, N3, P), R1a, N3, DE, J), and one previously undescribed haplogroup (probably a subgroup within haplogroup P*(xR1a)). Four of the haplogroups (P*(xR1a), BR(xDE, J, N3, P), R1a and N3) represented about 98% of the population sample. The analysis of population pairwise differences indicates that the Norwegian Y-chromosome distribution most closely resembles those observed in Iceland, Germany, the Netherlands and Denmark. Within Norway, geographical substructuring was observed between regions and counties. The substructuring reflects to some extent the European Y-chromosome gradients, with higher frequency of P*(xR1a) in the south-west and of R1a in the east. Heterogeneity in major founder groups, geographical isolation, severe epidemics, historical trading links and population movements may have led to population stratification and have most probably contributed to the observed regional differences in distribution of haplotypes within two of the major haplogroups.

Keywords: Short tandem repeat; Y-chromosomal lineages; Heterogeneity

1. Introduction

The first traces of human life in Norway, based on archeological data, were found in bays and inlets dating from the earliest period after the coast was free of ice approximately 9000–10,500 years ago [1–6]. The people were hunters, fishermen and gatherers until animal domestication and grain cultivation became more widespread some 4500 years ago [7]. During the period of the Vikings (750–1050 A.D.), Norwegians settled in northern Scotland, Shetland, the Orkneys, Isle of Man, Central England, Ireland, Normandy and Iceland. The Norwegian kingdom, previously ruled by several local petty kings, was united by Harald Hårfagre around 872. Until the end of the 12th century, the country was ruled by consecutive Norwegian kings. Several periods of shared kingdom with Denmark and Sweden then followed, until the termination of the last union in 1905.

The population density has always been low and the population widespread, the latter partly because of the mountainous geography. Before the Black Death in 1349

* Corresponding author. Tel.: +47 230 71312.
E-mail address: b.m.dupuy@labmed.uio.no (B.M. Dupuy).

0379-0738/$ – see front matter © 2005 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.forsciint.2005.11.009

and other epidemics that followed, the population size has been estimated to have been about 300,000 individuals. Only one-third to one-half of the population survived the epidemics. Numerous farms were laid waste and it took about 150 years for the population to regain its previous size. In the middle of the 14th century, 900–1800 German traders called the Hanseatics, lived permanently in Bergen (on the west coast). The number increased to 2000–3000 during the summers, quite significant for a city of about 7000 inhabitants. Scots and Dutchmen were also traders on the west coast and the Scots comprised 8% of the inhabitants in Bergen around 1640. At the first complete census in 1801, the Norwegian population registered 883,603 inhabitants of whom 20,071 had foreign names. At that time, Bergen still had 18% German named inhabitants. In most cities along the coast, Danes, Germans and Swedes represented, on average, 15% of the total population. From 1900 to 1946, 2.3% of the total population had been born abroad, 2/3 of these in Sweden. In the 16th and 17th centuries, people of Finnish extraction (known as Kvener) migrated from Finland via Sweden to northern Norway because of war and famine. In 1855, 6300 Kvener were registered in Troms and Finnmark. From 1620, Finns also migrated from central Finland via Sweden to eastern Norway. In 1686, 1200 Finns were registered in this region. The Saami are considered the aboriginal inhabitants of Fennoscandia. At present, approximately 50,000 Saami are inhabitants of Norway. Most information given in this section is taken from Kjeldstadli et al. [8].

The distributions of Y-chromosome STRs and biallelic markers in European populations have been extensively studied [9–13] and knowledge about population stratification is currently increasing [14–18]. Y-chromosome variation in forensic genetics has been validated and found to be a valuable tool in case work [19–26]. A European reference database, the Y-STR Haplotype Reference Database (http://www.yhrd.org/index.html) [27]) created by Sascha Willuweit and Lutz Roewer, comprised in December 2004 a European population sample of 17,373 haplotypes (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS385I/II) from a set of 126 populations (release 15).

Several studies of Y-chromosome variation in Norway have been published (Supplementary Data Online, Table 1), but this is by far the largest and most comprehensive study, analyzing a total of 1766 Norwegian samples at five biallelic markers and nine Y-STR loci. The aim is to map the haplotype and haplogroup distribution at the national level, to compare it with other European populations, and to study the regional distribution within the country and relate the findings to other genetical and historical data.

Table 1
Population sample: number of samples from the different counties

County code	County	Region	No. in population sample
1	Østfold	East	85
2	Akershus	East	58
3	Oslo	Capital (east)	109
4	Hedemark	East	78
5	Oppland	East	56
6	Buskerud	East	80
7	Vestfold	East	75
8	Telemark	East	61
9	Aust-Agder	South	34
10	Vest-Agder	South	42
11	Rogaland[a]	West	153
12	Hordaland	West	91
13	Bergen	City (west)	93
14	Sogn og Fjordane	West	57
15	Møre og Romsdal[a]	Middle	123
16	Sør-Trøndelag	Middle	132
17	Nord-Trøndelag	Middle	62
18	Nordland	North	191
19	Troms	North	105
20	Finnmark	North	81

County code and geographical region are given.

[a] In population frequency studies of classical markers, these two counties were divided in two. Rogaland in south/west and Møre and Romsdal in west/middle. The geographical division differs slightly from the topographical division defined by [29] later used with minor modifications in population frequency studies of classical markers e.g. [69] and [70]. A map of Norway showing the geographical localization of all counties is available online in Supplementary Data, Fig. 2.

2. Material and methods

2.1. Material

The material consists of 1766 unrelated Norwegian males involved in consecutive paternity cases during the years 1993–1999. Males with obvious non-Norwegian surnames were omitted as well as males whose oldest known patrilineal relatives (index persons), as reported in the National Register, were born abroad. The latter represented 3% of the population sample. Males with surnames of Scandinavian or German origin whose index person was born in Norway were included. The material includes males from confirmed father–son pairs only and is part of a large mutation study of Y-STRs [28]. The geographical origin of each sample is based on the place of birth of the index person whose median year of birth is 1942 (Supplementary Data Online, Fig. 1). The geographical distribution of these index persons reflects fairly well that of the population at that time (Table 1). For statistical analyses, the population sample was grouped according to geographical principles as described by Hartmann et al. [29]. Initially, the material was divided into regions: East, South, West, Middle and North. The capital, Oslo and the main city on the west coast, Bergen, were also separated as independent regions for a total of seven regions. Alternatively, the population was divided at the county level into 20 regions (Supplementary Data Online, Fig. 2) with each county representing from one to four ''fogderier'', the main administrative centre until 1891 (e.g. [30]).

2.2. Extraction, Y-STR amplification conditions and detection

Blood samples (EDTA) were extracted by the salting out method [31]. PCR conditions for Multiplex I (DYS19, DYS389I, DYS389II and DYS390) and Multiplex II (DYS391, DYS392 and DYS393) were as described by Kloosterman et al. [32]. PCR conditions for DYS385I/II and DYS388 have been described earlier, as well as electrophoretic methods, software and fragment length analysis [33]. Fragment length analysis was performed with the GS500 internal standard (Applied Biosystems) and sequenced allelic ladders for all loci. Published primer sequences [34] were used for PCR.

2.3. Amplification and detection of biallelic markers

The Tat locus was amplified as described in Zerjal et al. [35]. The resulting 112-bp product was screened by digestion with Hsp92II (Promega) and the C-alleles were confirmed by MaeII (Boehringer) digestion. The SRY_{10831} locus was amplified according to Whitfield et al. [36], and the resulting 432-bp product was screened by digestion with MaeIII. The 92R7 locus was amplified as described in Mathias et al. [37] and Hurles et al. [38]. The resulting 730-bp product was screened by digestion with *Hin*dIII. The 12f2 locus was amplified according to Casanova et al. [39] and Rosser et al. [14]. The resulting 88-bp product amplifies in non-haplogroup J chromosomes only. A control amplicon was amplified using the Tat primers. The YAP locus was typed according to the method of Hammer and Horai [40]. The recurrent 12f2 deletion was verified by M9 typing according to Underhill et al. [41]. The biallelic markers were separated by electrophoresis in a 2.5% agarose gel containing ethidium bromide in a 1× TBE-buffer and visualized by UV-radiation.

2.4. Nomenclature and definitions

The repeat number nomenclature follows the guidelines of the International Society of Forensic Genetics (ISFG) [22]. For DYS385I/II, the nomenclature of Schneider et al. [42], also used in Kayser et al. [34], was applied. The nomenclature for the biallelic markers is according to the recommendations of the Y-Chromosome Consortium [43]. Distinct Y-chromosomes were grouped into different levels of resolution; haplogroups on the basis of biallelic markers, haplotypes on the basis of STR loci and lineages based on both biallelic markers and STR loci [44]. The haplotype composition is always given as follows: DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS385I/II and DYS388, unless otherwise indicated.

2.5. Phylogenetic and statistical analysis

We used the Arlequin package [45] for analysis of molecular variance (AMOVA) and for calculation of population pairwise F_{ST}s and R_{ST}s (F_{ST} analogue taking into account the (mostly) stepwise mutation mechanism of STRs). AMOVA was performed on haplotype frequencies. The duplicated locus DYS385I/II contains two repeat tracts [46–48], which remain undistinguishable by the primers applied. The locus was therefore excluded in all statistical analysis because the designation of the haplotype is arbitrary at this locus. The phylogenetic relationships between the microsatellite haplotypes within haplogroups were determined by using the program Network 4.1.0.9. Networks were calculated by the median-joining method ($\varepsilon = 0$) [49]. The most variable loci were removed by giving them weight zero. DYS385I/II, DYS389I and DYS389II were omitted, the latter two because the lengths do not reflect underlying mutational events in this complex locus (Peter Forster, personal communication). In order to better visualize the genetic landscape in Norway and Europe, principal coordinates were identified from the F_{ST} and R_{ST} estimates between regions (7 samples), counties (20 samples) and countries (23 samples) by multidimensional scaling analysis (MDS) using the *R* statistical computing platform [50] with the packages MASS [51] and VEGAN [52]. For comparison of Norwegian regional and county data sets, dissimilarity matrices required for MDS analysis were simply symme-

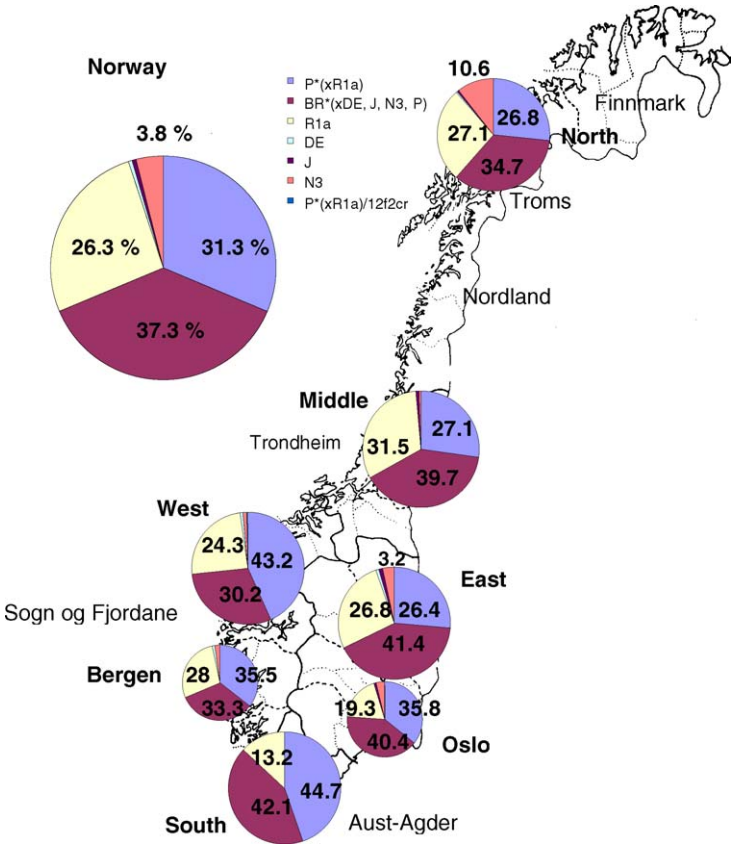

Fig. 1. Distribution of Y-haplogroups in Norway and Norwegian regions.

trical pairwise F_{st} and R_{st} tables. For the comparison of Norway to other European countries, a dissimilarity matrix was from computed with the Bray–Curtis distance using the comprehensive Norwegian F_{st} values.

Because pairwise distances in the dissimilarity matrix must be positive values (except for self-distance), F_{st} and R_{st} values that were zero or negative were converted to a comparatively very small number (0.00001). Kruskal's stress (S) and the coefficient of determination (R^2) from the MDS were examined for each data set, and the number of dimensions accounting for more than 90% of the variation in the data were chosen. To generate raster plots for the geographical representation, surface interpolation using inverse distance squared weighting was done using Geographical Resources Analysis Support System (GRASS) software [53]. Masking of the European shoreline was provided by the GTOPO30 (30-arcsec topographical) data set, publicly available at the US Land Processes Distribute Active Archive Center [54]. Genealogical depths or time to the most recent common ancestor, TMRCA, were calculated as the average across loci of the square differences, ASD, in STR repeat numbers between two haplotypes [55,56], one being defined as the ancestral haplotype [57].

3. Results

3.1. Haplogroup distribution in Norway and Norwegian regions

Among the 1766 individuals typed, 7 different haplogroups of Y-chromosomes were identified (Fig. 1). Six haplogroups have been described previously; P*(xR1a), BR(xDE, J, N3, P), R1a, DE, J and N3. Four of the haplogroups (P*(xR1a), BR(xDE, J, N3, P), R1a and N3) represented ~98% of the population sample. A new haplogroup, not described earlier, was found in a single sample. Deduced from its biallelic type, it might represent a new 12f2 deletion within haplogroup P*(xR1a). The haplogroup it defines has been given the temporary name P*(xR1a)/12f2c (M. Jobling personal communication). Its haplotype composition is 15-10-17-24-10-13-14-11,14-12. There are already two known 12f2 deletions within hgJ and hgD2. The recurrence is a non-allelic homologous double recombination event between directly repeated HERV elements that flank the AZFa region on Yq [58].

Haplogroup frequency distributions in the different Norwegian regions are presented (Fig. 1). The frequency of P*(xR1a) varied from 26% in the east to 45% in the south,

BR*(xDE, J, N3, P) from 30% in the west to 42% in the south and R1a from 13% in the south to 32% in the middle. N3 was most frequent in the north (11%; 18.6% in the northernmost county Finnmark) and totally absent in the south. Haplogroup DE and J were rare in all regions. We observe a relatively high frequency of P*(xR1a) and R1a in the population sample from south-west and east, respectively.

Y-STR allele frequency distribution in each haplogroup is available online (Supplementary Data Online, Table 2). Allele frequency distributions by regions have been described earlier [59].

3.2. Haplotype and lineage distribution in Norway and Norwegian regions

A total of 721 different haplotypes (based on Y-STRs) and 726 different lineages (based on Y-STRs and biallelic markers) were observed among 1766 individuals (Supplementary Data Online, Table 3). Five haplotypes were thus identical by state (chance) but not by descent (common ancestor). Of these, four haplotypes were present in haplogroup P*(xR1a) and R1a and one was present in haplogroup BR*(xDE, J, N3, P) and R1a (for details see Supplementary Data Online, Table 3). Five hundred lineages (28%) were only observed once; furthermore 35% of the total material consists of lineages less frequent than 0.5%.

The frequency distribution of the six most common haplotypes within haplogroups P*(xR1a), BR(xDE, J, N3, P), R1a and N3 in Norway is presented in Fig. 2. The six haplotypes represent 24, 37, 28 and 52%, of each haplogroup, respectively. Regional distribution of the most common haplotypes, within haplogroups P*(xR1a), BR*(xDE, J, N3, P) and R1a is available online (Supplementary Data Online, Fig. 3). Within P*(xR1a), the distribution of haplotype 14-10-16-24-11-13-13-11,14-12 ranged from 3% in Bergen to 15% in the south. The most frequent haplotype of all (14-9-16-23-10-11-13-14,15-14) is within haplogroup BR*(xDE, J, N3, P), and represented 4.4% of the total population sample. The distribution ranged from 3% in the middle to 19% in Bergen. Within R1a, the haplotype distribution of 15-11-17-25-11-11-13-11,14-12 ranged from 5% in Oslo to 40% in the south. Haplotype networks for P*(xR1a), BR*(xDE, J, N3, P) and R1a are presented in Supplementary Data Online, Fig. 4.

3.3. AMOVA and genetic structure

The analysis of AMOVA shows that the main source of variation is within the population. The percentage of variation found between regions is only 0.47 while it is 99.53 within the population, however the *P*-value was significant (0.00347 ± 0.00055, 10,000 permutations), indicating substructuring between regions in the Norwegian population sample. Signs of regional substructuring were also observed when using population pairwise F_{ST}s. The matrix of pairwise F_{ST} values for regions and their significance level, based on haplogroup frequencies is presented online (Supplementary Data Online, Table 4). Most regions, apart from the main cities, differ from one another. A geographical presentation is given in Fig. 3. Population pairwise R_{ST}s were calculated based and Y-

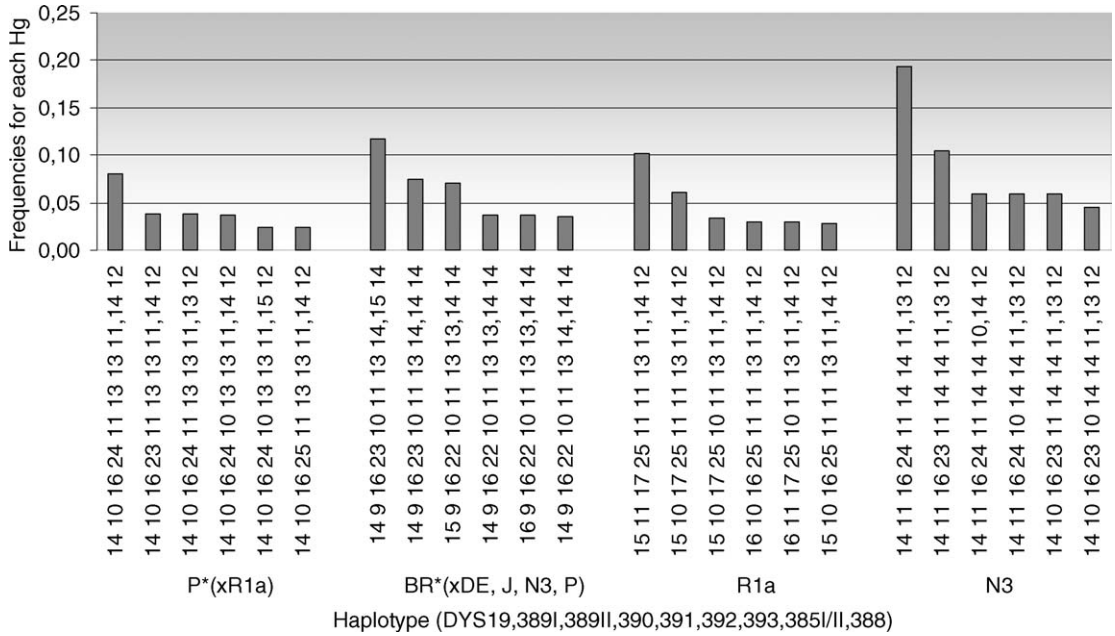

Fig. 2. Distribution of the six most common lineages within haplogroups (P*(xR1a), BR(xDE, J, N3, P)), R1a and N3.

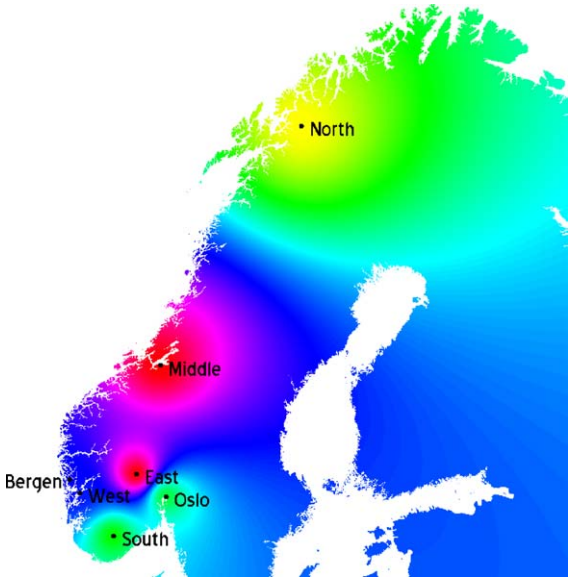

Fig. 3. Multidimensional scaling analysis of pairwise Y-SNP based F_{ST} between seven Norwegian regions. Data were gathered only from the geographical coordinates indicated. Map coloration is the result of algorithmic interpolation and must be interpreted appropriate skepticism.

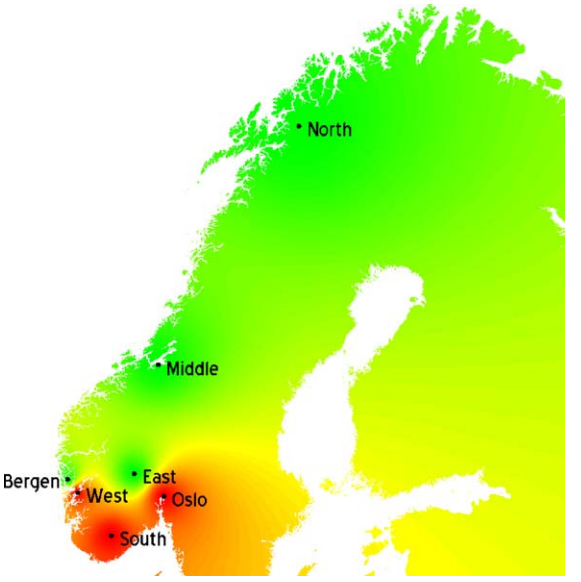

Fig. 4. Multidimensional scaling analysis of pairwise Y-STR based R_{ST} between seven Norwegian regions. Data were gathered only from the geographical coordinates indicated. Map coloration is the result of algorithmic interpolation and must be interpreted appropriate skepticism.

STR haplotypes (Supplementary Data Online, Table 5). Most of the population differences previously detected at the regional level were confirmed; however it seems that separation between regions is better with haplogroup, rather than haplotype, distributions. A geographical presentation is given in Fig. 4.

The differentiation test (AMOVA) between counties gave a P-value of 0.00119 ± 0.00032 (10,000 permutations), indicating population substructuring at the county level as well. These finding initiated the search for a more precise geographical location of substructuring. Low interpopulation variance was observed between most counties. $R_{ST} < 0$, between 0 and 0.01 and >0.01 were observed in 39, 31 and 30% of the comparisons respectively. For most of the pairwise population comparisons, the interpopulation differences were non-significant ($P > 0.05$). The counties Sogn og Fjordane in the west, Aust-Agder in the south and Finnmark in the north were significantly different from 19, 12 and 9 other counties, respectively (Supplementary Data Online, Table 6). A geographical presentation is given in Fig. 5.

By calculating population pairwise R_{ST}s based on Y-STR haplotype data within each of the three major haplogroups, we observed regional differences within haplogroups BR*(xDE, J, N3, P) and R1a (Supplementary Data Online, Table 7 and 8). The analysis of AMOVA gave P-values of 0.03099 ± 0.00160 and 0.01020 ± 0.00096 for BR*(xDE, J, N3, P) and R1a, respectively (10,000 permutations). Within haplogroup BR*(xDE, J, N3, P) the middle region showed significant differences to the north, east and west. Regional differences were also observed between the south and the north. Within haplogroup R1a, Bergen showed significant differences to all regions but Oslo.

3.4. Comparison to other European populations

The frequency distribution of haplogroups in Norway has been compared to those of other European countries (Supplementary Data Online, Table 9 and Fig. 5). P*(xR1a) representing 32% in Norway, varies from 1% in Finland to 83% in Wales. BR(xDE, J, N3, P) representing 36% in Norway, varies from 4% in Basques to 59% in Gotland (Sweden) and R1a representing 26% in Norway varies from 1% in Ireland to 55% in Poland. Haplogroups P*(xR1a) and R1a are interpreted as signatures of expansion from isolated populations in the Iberian Peninsula and the present Ukraine, respectively, whereas haplogroup BR(xDE, J, N3, P) probably originated in Europe 20,000–25,000 years ago [15]. Haplogroup N3 has been interpreted as a signature of Uralic Finno-Ugric speaking males migrating to northern Scandinavia about 4000–5000 years ago [9,17,35,60]. In the present study, N3 is observed at 4% in the overall population and at 11% in the northern region corresponding to 150,000 and 50,000 inhabitants, respectively. These numbers exceed the total number of Saami inhabitants, which is officially recognized as about 50,000 (http://www.sametinget.se). In northern Norway, the N3 percentage is 18.6% in Finnmark, 8.6% in Troms and 8.4% in Nordland (which are the three northernmost counties—Nordland being located to the south of the other two (Supplementary Data Online, Fig. 2)). There is thus a

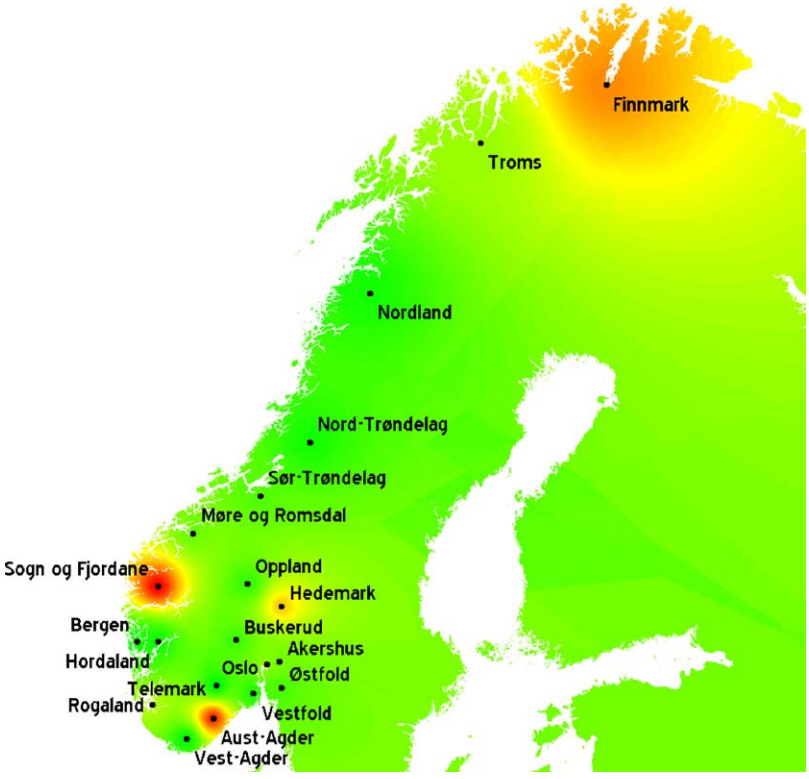

Fig. 5. Multidimensional scaling analysis of pairwise Y-STR based R_{ST} between 20 Norwegian counties. Data were gathered only from the geographical coordinates indicated. Map coloration is the result of algorithmic interpolation and must be interpreted appropriate skepticism.

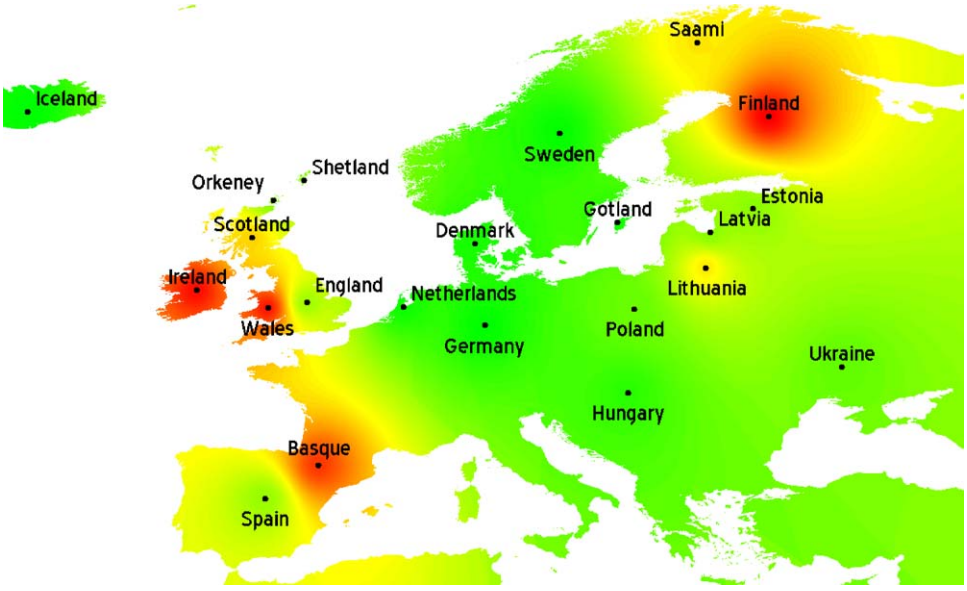

Fig. 6. Multidimensional scaling analysis of pairwise Y-SNP based F_{ST} between 23 European countries. Data were gathered only from the geographical coordinates indicated. Map coloration is the result of algorithmic interpolation and must be interpreted appropriate skepticism.

considerable pool of Saami and/or Finnish Y-chromosomes in the Norwegian population and particularly in the north. F_{ST} data demonstrated significant differences of haplogroup frequencies in Norway to all other European countries (Supplementary Data Online, Table 10). Nevertheless, F_{ST} values were low when the Norwegian population sample was compared with samples from Iceland, Germany and Sweden (Supplementary Data Online, Table 10, Column 2). A geographical presentation is given in Fig. 6. The large Norwegian sample size implies that smaller F_{ST} values would become significant. At the regional level, F_{ST} values were low when Oslo, Bergen, West and South, were compared to Iceland and Germany. The southern region showed in addition low F_{ST} values when compared to haplogroup frequencies in the Netherlands and Denmark.

The most common Norwegian haplotypes have been matched against the Y-STR database (also comprising 300 previously analyzed Norwegian samples [33] (Supplementary Data Online, Table 11). Two haplotypes correspond to the "Nordic haplotype" with alleles 14-23-10-11-13, defined by DYS19-390-391-392-393, respectively [60]. They belong to subgroup I1a which accounts for most of BR*(xDE, J, N3, P) in Scandinavia [61]. This is also consistent with results of subtyping of BR*(xDE, J, N3, P) with M170 and M9 demonstrating that almost all Norwegian BR*(xDE, J, N3, P) chromosomes are within hgI [13,62].

4. Discussion

The Norwegian Y-chromosomes are mainly (98%) composed of haplogroups P*(xR1a), BR*(xDE, J, N3, P), R1a and N3. This finding is in agreement with previous reports. However, there are significant differences between some of the different Norwegian population samples based on haplogroup frequency distribution in the materials (Supplementary Data Online, Table 1). Such differences may be caused by the geographical substructuring observed here. As the male population is heterogeneous, ample precautions must be taken in order to collect representative population samples for specific genetic investigations and for forensic genetic purposes. The new knowledge regarding Y-chromosome substructuring among Norwegian males presented here might have implications for the interpretation of specific Y-STR haplotypes in forensic genetics. The most common haplotype within BR*(xDE, J, N3, P) represents almost 12% in the overall Norwegian population sample. In Bergen, this particular haplotype represent nearly 20% of all BR*(xDE, J, N3, P) samples (Supplementary Data Online, Fig. 3). Similarly, the most common haplotype within R1a represents 10% in the overall Norwegian population sample but 40% in samples from the south. These observations could be partly explained by relatively small population sizes (Bergen: $n = 93$, South: $n = 76$, Table 1) but regional clustering remains the most probable explanation. These findings should therefore be taken into consideration when applying Y-STR analysis in forensic casework.

From a historical and genealogical view the observed geographical substructuring might also throw further light on the Norwegian Y-chromosome contribution to, e.g., Iceland [9] and the British Isles [13].

The observed similarities at the haplogroup and haplotype level between the Nordic countries are most probably signs of our common ancestors. It is common opinion that different cultures (e.g. Ahrensburgian and Swiderian) directly or indirectly emerged from the three ice age refugia located in Iberia, Balkan and Ukraine, and that they migrated, little by little, to the north as the ice melted. The people of these cultures are thus thought to be the pro-Germans that later populated Norway from the south [63–65]. The observed similarity to Germany may have been strengthened by the important trade between the countries in the 12th century. This is probably reflected in the observed regional differences between Bergen and the rest of the regions in the pairwise comparisons of haplotypes within haplogroup R1a (Supplementary Data Online, Table 8). The similarity to Iceland is probably explained by the fact that the country was partly populated by Norwegians in the eighth century as many petty kings and their men were forced to leave the country after the union of the Norwegian kingdom. It is also obvious that the severe epidemics that raged in the country in the 12th and 13th centuries must have led to important bottlenecks and influenced the distribution of Y-chromosomes.

Genealogical depths of haplogroups in Norway have been calculated using lineage specific mutation rates (DYS385 omitted) [28] and a generation time of 20 years. We estimated a date for Norwegian P*(xR1a) coalescence of 4000 (95% CI 19,900–2200 BP). For BR(xDE, J, N3, P), R1a and N3 coalescences of 5000 (95% CI 15,000–3000 BP), 2500 (95% CI 5700–1600 BP) and 2100 (95% CI 0–700 BP) were estimated, respectively. All confidence intervals overlap and do not support different waves of settlers. However, postglacial population expansion from Franco-Cantabria could have contributed mainly BR(xDE, J, N3, P). The spread of Ahrensburgian and Swederian Mesolithic technologies associated with the recolonization after the last glacial maximum could have brought P*(xR1a) to the population, while R1a might represent the spread of the Corded Ware and Battle-Axe cultures from central and east Europe. Finally, N3 is interpreted as a signature of Finno-Ugric speaking males migrating to the north.

In this study population, substructuring is demonstrated in Aust-Agder (south), Sogn og Fjordane (west) and Finnmark (north). This is consistent with other genetic studies reporting regional differences. The frequency of K+ in the Kell blood group is 4.61% in Aust-Agder and 10.36% in Sogn og Fjordane. Both counties differ from 8%, which is the Norwegian average [66]. It is interesting to note that two alleles for recessive diseases, where the "disease" alleles are particularly frequent in Norway (R272X in PAH on chromo-

some 12 and A2526G in TGM1 on chromosome 14), both seem to be absent in the Aust-Agder population [67,68].

We have thus demonstrated geographical heterogeneity of Y-chromosomal lineages in Norway, an observation that might have implications for the interpretation of Y-STR haplotypes in forensic and population genetics. The substructuring reflects to some extent the European Y-chromosome gradients. Further, geographical isolation, severe epidemics, historical trading links and population movements may have contributed to the geographical heterogeneity observed in this study.

Acknowledgments

We would like to thank Anne Grete Flønes, Karianne Tomassen, Rikke Wallum and Uzma Shaffi for technical assistance, Thore Egeland for valuable discussions and Ian Frame for introducing us to the Arlequin software. The research was approved by the Regional Research Ethics Committee (Helseregion Sør) and supported by The Institute of Forensic Medicine, Department of Family Genetics, University of Oslo.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at 10.1016/j.forsciint.2005.11.009.

References

[1] H.B. Bjerck, The Fosna-Nøstvedt Problem. A consideration of archaeological units and chronozones in the South Norwegian mesolithic period, Norwegian Archaeol. Rev. 19 (1986).
[2] B. Skar, The early mesolithic site Rørmyr II. A re-examination of one of the Høgnipen sites. SE Norway, Acta Archaeol. 56 (1987).
[3] S.E. Nygaard, The stone age of Northern Scandinavia: a review, J. World Prehistory 3 (1989) 71–116.
[4] S. Bang-Andersen, The Myrvatn group, a preboreal find-complex in south-west Norway, in: P.M.V.P. Vermeersch (Ed.), Contributions to the Mesolithic in Europe, Leuven, 1990, pp. 215–226.
[5] L. Prøsch-Danielsen, M. Høgestøl, A coastal Ahrensburgian site found at Galta, Rennesøy, Southwest Norway, in: A. Fischer (Ed.), Man and Sea in the Mesolithic, Oxford, 1995.
[6] K. Amundsen, The physical conditions for earliest settlement during the last deglaciation in Norway, in: L. Larsson (Ed.), The Earliest Settlement of Scandinavia, Almquist & Wiksell International, Stockholm, 1996, pp. 207–218.
[7] E.H. Hofseth, From Ice Age to Nordic Christianity, ed. U.M.o.C. Heritage, Oslo.
[8] K. Kjeldstadli, et al. in: K. Kjeldstadli (Ed.), Norsk innvandringshistorie (“The history of immigration in Norway”), vols. 1–3, Pax Forlag A/S, Oslo, 2003.
[9] A. Helgason, et al., Estimating Scandinavian and Gaelic ancestry in the male settlers of Iceland, Am. J. Hum. Genet. 67 (2000) 697–717.
[10] E.W. Hill, M.A. Jobling, D.G. Bradley, Y-chromosome variation and Irish origins, Nature 404 (2000) 351–352.
[11] M. Raitio, et al., Y-chromosomal SNPs in Finno-Ugric-speaking populations analyzed by minisequencing on microarrays, Genome Res. 11 (2001) 471–482.
[12] M. Brion, et al., Insights into Iberian population origins through the construction of highly informative Y-chromosome haplotypes using biallelic markers, STRs, and the MSY1 minisatellite, Am. J. Phys. Anthropol. 122 (2003) 147–161.
[13] C. Capelli, et al., A Y chromosome census of the British Isles, Curr. Biol. 13 (2003) 979–984.
[14] Z.H. Rosser, et al., Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language, Am. J. Hum. Genet. 67 (2000) 1526–1543.
[15] O. Semino, et al., The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective, Science 290 (2000) 1155–1159.
[16] R. Scozzari, et al., Human Y-chromosome variation in the western Mediterranean area: implications for the peopling of the region, Hum. Immunol. 62 (2001) 871–884.
[17] T. Zerjal, et al., Geographical, linguistic, and cultural influences on genetic diversity: Y-chromosomal distribution in Northern European populations, Mol. Biol. Evol. 18 (2001) 1077–1087.
[18] M.E. Weale, et al., Y chromosome evidence for Anglo-Saxon mass migration, Mol. Biol. Evol. 19 (2002) 1008–1021.
[19] V.L. Pascali, M. Dobosz, B. Brinkmann, Coordinating Y-chromosomal STR research for the Courts, Int. J. Legal Med. 112 (1999) 1.
[20] L. Roewer, et al., A new method for the evaluation of matches in non-recombining genomes: application to Y-chromosomal short tandem repeat (STR) haplotypes in European males, Forensic Sci. Int. 114 (2000) 31–43.
[21] D. Corach, et al., Routine Y-STR typing in forensic casework, Forensic Sci. Int. 118 (2001) 131–135.
[22] P. Gill, et al., DNA commission of the International Society of Forensic Genetics: recommendations on forensic analysis using Y-chromosome STRs, Int. J. Legal Med. 114 (2001) 305–309.
[23] M.A. Jobling, Y-chromosomal SNP haplotype diversity in forensic analysis, Forensic Sci. Int. 118 (2001) 158–162.
[24] M. Krawczak, Forensic evaluation of Y-STR haplotype matches: a comment, Forensic Sci. Int. 118 (2001) 114–115.
[25] M. Prinz, et al., Validation and casework application of a Y chromosome specific STR multiplex, Forensic Sci. Int. 120 (2001) 177–188.
[26] Y. Bilge, et al., The identification of a dismembered human body: a multidisciplinary approach, Forensic Sci. Int. 137 (2003) 141–146.
[27] L. Roewer, et al., Online reference database of European Y-chromosomal short tandem repeat (STR) haplotypes, Forensic Sci. Int. 118 (2001) 106–113.
[28] B.M. Dupuy, et al., Y-chromosomal microsatellite mutation rates: differences in mutation rate between and within loci, Hum. Mut. 23 (2004) 117–124.
[29] O. Hartmann, K. Hadland, C.J. Bjerklund, Blood group distribution in Norway, Norsk Vidensk.-Akad. i Oslo, Mat-Naturvit. Klasse 5 (1941) 1–41.

[30] T. Gedde-Dahl Jr., Population structure in Norway. Inbreeding, distance and kinship, Hereditas 73 (1973) 211–232.
[31] S.A. Miller, D.D. Dykes, H.F. Polesky, A simple salting out procedure for extracting DNA from human nucleated cells, Nucleic Acids Res. 16 (1988) 1215.
[32] A.D. Kloosterman, P.M. Daselaar, H.J.T. Janssen, Population genetic study of Y-chromosome specific STR-loci in Dutch Caucasians, in: 17th Congress of the International Society for Forensic Haemogenetics, Springer, Oslo, 1997.
[33] B.M. Dupuy, et al., Y-chromosome variation in a Norwegian population sample, Forensic Sci. Int. 117 (2001) 163–173.
[34] M. Kayser, et al., Evaluation of Y-chromosomal STRs: a multicenter study, Int. J. Legal Med. 110 (1997) 125–133.
[35] T. Zerjal, et al., Genetic relationships of Asians and Northern Europeans, revealed by Y-chromosomal DNA analysis, Am. J. Hum. Genet. 60 (1997) 1174–1183.
[36] L.S. Whitfield, J.E. Sulston, P.N. Goodfellow, Sequence variation of the human Y chromosome, Nature 378 (6555) (1995) 379–380.
[37] N. Mathias, M. Bayes, C. Tyler-Smith, Highly informative compound haplotypes for the human Y chromosome, Hum. Mol. Genet. 3 (1994) 115–123.
[38] M.E. Hurles, et al., Recent male-mediated gene flow over a linguistic barrier in Iberia, suggested by analysis of a Y-chromosomal DNA polymorphism, Am. J. Hum. Genet. 65 (1999) 1437–1448.
[39] M. Casanova, et al., A human Y-linked DNA polymorphism and its potential for estimating genetic and evolutionary distance, Science 230 (4732) (1985) 1403–1406.
[40] M.F. Hammer, S. Horai, Y chromosomal DNA variation and the peopling of Japan, Am. J. Hum. Genet. 56 (1995) 951–962.
[41] P.A. Underhill, et al., Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography, Genome Res. 7 (1997) 996–1005.
[42] P.M. Schneider, et al., Tandem repeat structure of the duplicated Y-chromosomal STR locus DYS385 and frequency studies in the German and three Asian populations, Forensic Sci. Int. 97 (1998) 61–70.
[43] YCC, A nomenclature system for the tree of human Y-chromosomal binary haplogroups, Genome Res. 12 (2002) 339–348.
[44] P. de Knijff, Messages through bottlenecks: on the combined use of slow and fast evolving polymorphic markers on the human Y chromosome, Am. J. Hum. Genet. 67 (2000) 1055–1061.
[45] S. Schneider, K.J. Roessli D., L. Excoffier, Arlequin Ver. 1.1: A Software for Population Genetic Data Analysis, Genetics and Biometry Laboratory, University of Geneva, Switzerland, 1997
[46] M. Kayser, et al., Applications of microsatellite-based Y chromosome haplotyping, Electrophoresis 18 (1997) 1602–1607.
[47] P.M. Schneider, et al., Tandem repeat structure of the duplicated Y-chromosomal STR locus DYS385 and frequency studies in the German and three Asian populations, Forensic Sci. Int. 97 (1998) 61–70.
[48] R. Kittler, et al., Apparent intrachromosomal exchange on the human Y chromosome explained by population history, Eur. J. Hum. Genet. 11 (2003) 304–314.
[49] H.J. Bandelt, et al., Mitochondrial portraits of human populations using median networks, Genetics 141 (1995) 743–753.
[50] R. Ihaka, R. Gentleman, A language for data analysis and graphics, J. Comput. Graphical Stat. 5 (1996) 299–314.
[51] W.N. Venables, B.D. Ripley, Modern Applied Statistics with S, fourth ed., Springer, New York, 2002.
[52] J. Oksanenen, R. Kindt, R.B. O'Hara, Community Ecol. Package Version 1.6–9 (2005).
[53] GRASS, Geographical Resources Analysis Support System.
[54] Gtop30, US Land Processes Distribution Active Archive Center.
[55] D.B. Goldstein, et al., An evaluation of genetic distances for use with microsatellite loci, Genetics 139 (1995) 463–471.
[56] M. Slatkin, A measure of population subdivision based on microsatellite allele frequencies, Genetics 139 (1995) 457–462.
[57] N.O. Bianchi, et al., Characterization of ancestral and derived Y-chromosome haplotypes of New World native populations, Am. J. Hum. Genet. 63 (1998) 1862–1871.
[58] P. Blanco, et al., Divergent outcomes of intrachromosomal recombination on the human Y chromosome: male infertility and recurrent polymorphism, J. Med. Genet. 37 (2000) 752–758.
[59] B.M. Dupuy, et al., A study of Y-chromosomal microsatellites and biallelic markers in Norway, in: 20th International ISFG Congress, Arcachon, France, 2003.
[60] K. Tambets, 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. 74 (2004) 661–682.
[61] S. Rootsi, et al., Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe, Am. J. Hum. Genet. 75 (2004) 128–137.
[62] G. Passarino, et al., Different genetic components in the Norwegian population revealed by the analysis of mtDNA and Y chromosome polymorphisms, Eur. J. Hum. Genet. 10 (2002) 521–529.
[63] M. Otte, in: C.G.O. Soffer (Ed.), The World at 18000 BP, Unwin Hyman, London, 1990, pp. 54–68.
[64] P.M. Dolukhanov, in: K.J.K. Wiik (Ed.), The Most Ancient North Europeans: Consensus in Sight? The Roots of Peoples and Languages of Northern Eurasia, Societas Historiae Fenno-Ugricae, Turku, 1998, pp. 9–27.
[65] K. Wiik (Ed.), The Roots of the Europeans, in press.
[66] L. Kornstad, Frequency of the blood group antigen K and the A1A2BO groups in the Norwegian counties, Gene Geogr. 11 (1997) 37–46.
[67] H.G. Eiken, et al., Relative frequency, heterogeneity and geographic clustering of PKU mutations in Norway, Eur. J. Hum. Genet. 4 (1996) 205–213.
[68] M. Pigg, et al., Strong founder effect for a transglutaminase 1 gene mutation in lamellar ichthyosis and congenital ichthyosiform erythroderma from Norway, Eur. J. Hum. Genet. 6 (1998) 589–596.
[69] P. Teisberg, The distribution of C 3 types in Norway, Hum. Hered. 21 (2) (1971) 154–161.
[70] T. Reinskou, Distribution of the Gc types in Norway, Acta Genet. Stat. Med. 15 (1965) 33–44.