

ARTICLE

# Highly discrepant proportions of female and male Scandinavian and British Isles ancestry within the isolated population of the Faroe Islands

Thomas D Als<sup>\*,1</sup>, Tove H Jorgensen<sup>1</sup>, Anders D Børghlum<sup>2</sup>, Peter A Petersen<sup>3</sup>,  
Ole Mors<sup>1</sup> and August G Wang<sup>4,5</sup>

<sup>1</sup>Centre for Basic Psychiatric Research, Aarhus University Hospital, Risskov, Denmark; <sup>2</sup>Institute of Human Genetics, University of Aarhus, Denmark; <sup>3</sup>The Faroese Church, Copenhagen, Denmark; <sup>4</sup>Department of Psychiatry, Landssjúkrahúsið (National Hospital), Torshavn, Faroe Islands and <sup>5</sup>Department of Psychiatry, Amager Hospital, Copenhagen University Hospital, Copenhagen S, Denmark

The Faroe Islands in the North Atlantic Ocean are inhabited by a small population, whose origin is thought to date back to the Viking Age. Historical, archaeological and linguistic evidence indicates that the present population of the Faroe Islands may have a mixture of Scandinavian and British Isles ancestry. In the present study we used 122 new and 19 previously published hypervariable region I sequences of the mitochondrial control region to analyse the genetic diversity of the Faroese population and compare it with other populations in the North Atlantic region. The analyses suggested that the Faroese mtDNA pool has been affected by genetic drift, and is among the most homogenous and isolated in the North Atlantic region. This will have implications for attempts to locate genes for complex disorders. To obtain estimates of Scandinavian vs British Isles ancestry proportions, we applied a frequency-based admixture approach taking private haplotypes into account by the use of phylogenetic information. While previous studies have suggested an excess of Scandinavian ancestry among the male settlers of the Faroe Islands, the current study indicates an excess of British Isles ancestry among the female settlers of the Faroe Islands. Compared to other admixed populations of the North Atlantic region, the population of the Faroe Islands appears to have the highest level of asymmetry in Scandinavian vs British Isles ancestry proportions among female and male settlers of the archipelago.

*European Journal of Human Genetics* (2006) 14, 497–504. doi:10.1038/sj.ejhg.5201578; published online 25 January 2006

**Keywords:** mtDNA; genetic diversity; ancestry proportions; Faroe Islands; North Atlantic region

## Introduction

The North Atlantic region was dominated primarily by Norwegian Vikings for several centuries starting in the late 700 AD. Scandinavian Vikings travelled the coastal regions stretching from Norway, Sweden, and Denmark to the

Shetlands, Orkney, Scotland, the Hebrides, and Ireland where they settled and intermarried with the original populations.<sup>1–3</sup> The Faroese saga<sup>4</sup> dates the first settlement of the archipelago to approximately 825–875 AD. Irish monks may, however, have lived on the islands as early as 650 AD and later deserted them due to the appearance of the Vikings.<sup>5–7</sup> Historical and archaeological records date the discovery of Iceland slightly later, that is, just before 870 AD. The oldest archaeological findings on the Faroe Islands are of Norse origin and date back to 1100 or later, but items of Celtic origin from this period have also been

\*Correspondence: Dr TD Als, Centre for Basic Psychiatric Research, Aarhus University Hospital, Skovagervej 2, Risskov DK-8240, Denmark, Tel: +45 77893505; Fax: +45 77893599; E-mail: tda@psykiatri.aaa.dk  
Received 9 September 2005; revised 13 December 2005; accepted 14 December 2005; published online 25 January 2006

found.<sup>8,9</sup> The term 'Celtic' is in this context defined as Scotland and Ireland. The Faroese language has its closest affinity to Icelandic and old Norwegian languages, but a Celtic influence has also been identified.<sup>5,10</sup> Throughout the remaining text Scottish/Irish ancestry will be referred to as British Isles ancestry.

Due to the relatively isolated geographic position of the archipelago, migration to the Faroe Islands has presumably been sparse after the initial settlement period. After the reign of Norwegian kings, the Faroe Islands became part of the Danish kingdom in 1380 AD. In the centuries after, Danish priests and officials settled on the islands, and sailors and traders from other European countries are very likely to have left descendants on the islands.<sup>6</sup> The Faroese population was presumably started by only a few founders and experienced early slow growth to a population size of only 4000 for the entire Faroese archipelago in the late 1300s, increasing to approx. 5000 around 1800 AD and then growing rapidly during recent years to the present 48 000 inhabitants.

In recent years, haploid, non-recombining and uniparentally inherited parts of the human genome (the Y chromosome and the mitochondrial genome) have become standard tools for genetic analyses of human populations.<sup>11–13</sup> Specific markers or regions of these units have been successfully used in conjunction with archaeological, historical and linguistic sources to trace the origin of human populations in a number of studies.<sup>14,15</sup> Genetic markers of the Y chromosome reveal patterns of population history through the paternal line, whereas the mitochondrial genome reveals population history through the maternal line.

Several populations of the North Atlantic region (Figure 3) have recently been studied with respect to male and female ancestry.<sup>16–20</sup> These studies suggest that areas remote from Scandinavia have excess Scandinavian ancestry among males and excess British Isles ancestry among females, whereas areas close to Scandinavia have a higher degree of symmetric ancestry proportions among females and males. In this study we add an important piece of missing information to the population history of the North Atlantic region by presenting an analysis of HVR I sequences from the mitochondrial control region in the Faroe Islanders. The two main focuses of this paper are to (i) assess the effect of genetic drift on shaping genetic diversity within the Faroese population, and (ii) to evaluate the genetic impact of Scandinavian *versus* British Isles female ancestors on the Faroese population and compare it with previous estimates of male ancestry based on Faroese Y-chromosomal data<sup>20</sup> applying identical methods.<sup>19</sup>

## Methods

DNA samples were provided by 122 unrelated males whose grandmothers originated from various islands, thereby

representing all major regions of the archipelago. These males were originally collected for a study of male ancestry proportions using Y-chromosomal markers on the Faroe Islands.<sup>20</sup> Sixty-one unrelated Danish males were added to an existing data set of 54 Danish males to increase the Danish contribution to the Scandinavian data set. All individuals had voluntarily consented to participate in the genetic survey. In addition to the mitochondrial haplotypes sequenced by the authors, the following previously published HVR I sequences were obtained for 12 European populations: 394,<sup>17</sup> 551,<sup>21</sup> 14,<sup>22</sup> and 39<sup>23</sup> Icelanders, 78 sequences from the Orkney Islands,<sup>18</sup> 49 from Isle of Skye,<sup>18</sup> 181 sequences from the Western Scottish isles,<sup>18</sup> 219 sequences from the Scottish North West coast,<sup>18</sup> 502 sequences from the Shetlands,<sup>19</sup> 672 sequences from Inland Scotland,<sup>18</sup> 100<sup>24</sup> and 200<sup>15</sup> Irish sequences, 19 sequences from the Faroe Islands,<sup>25</sup> 323,<sup>18</sup> 215,<sup>26</sup> 74<sup>11</sup> and 16<sup>24</sup> sequences from Norway, 28<sup>27</sup> and 32<sup>28</sup> Swedish sequences, 33,<sup>22</sup> 5<sup>24</sup> and 16<sup>25</sup> Danish sequences. The past Scandinavian source population is represented in our study by contemporary samples from Norway, Sweden and Denmark. These populations were therefore grouped in the subsequent analyses. The past source population of the Northern and Western parts of the British Isles is represented by contemporary samples from Ireland and inland Scotland,<sup>19</sup> which were grouped for analytical purposes. Sequences from the Western Scottish isles and the neighbouring Isle of Skye were likewise grouped, and a total of eight populations were used in the subsequent analyses (Figure 3).

DNA was prepared from blood samples, using a standard sucrose/Triton-lysis protocol with sodium chloride/isopropanol precipitation. All mtDNA site numbers referred to are in accordance with the reference sequence by Anderson *et al.*<sup>29</sup>

Fragments of 540 bp of the mitochondrial HVR I region were obtained by standard PCR amplification using primer pairs L15 999/H16 498.<sup>17</sup> Amplification reactions were performed on 18–36 ng of DNA in a 6- $\mu$  volume by use of taq polymerase (Roche). The typical thermal cycling profile was 94°C for 5 min, 94°C for 30 s, 50°C for 30 s, 72°C for 60 s, for 35 cycles, and 72°C for 7 min. After amplification, the double-stranded DNA was purified using MicroClean PCR purification kit (Microzone Ltd) prior to direct sequencing. Cycle sequencing reactions (8  $\mu$ l) were performed using ABI Prism™ Big Dye v.3.0 terminator cycle sequencing kits. Both strands of the PCR product were sequenced using fluorescence-labeled primers on a 3100 Genetic Analyzer (Applied Biosystems, CA, USA). Approximately 500 bp of forward and reverse sequences were aligned and manually checked and edited using SeqMan™ II version 5.0 (DNASTAR Inc.). To allow comparison with previously published data, only the sequence starting with position 16090 and ending with position 16366 were considered in the subsequent comparative analyses.

**Table 1** Summary statistics

(a)										
	N	k	k/n	Private haplotypes	S	S/n	Gene diversity (S.E.)	Nucleotide diversity (S.E.)		
Iceland	998	170	0.1703	0.5000	89	0.0892	0.9706	0.0022	0.0155	0.0085
Faroe Islands	141	28	0.1986	0.3929	29	0.2057	0.8648	0.0214	0.0073	0.0046
Shetland	502	167	0.3327	0.4970	91	0.1813	0.9362	0.0087	0.0122	0.0069
Orkney	78	37	0.4744	0.2432	35	0.4487	0.9507	0.0142	0.0135	0.0076
Western Isles & Skye	230	94	0.4087	0.3404	65	0.2826	0.9781	0.0035	0.0153	0.0084
Scottish NW coast	219	96	0.4384	0.2500	69	0.3151	0.9704	0.0065	0.0150	0.0083
Norway, Sweden & Denmark	803	308	0.3836	0.6136	127	0.1582	0.9531	0.0055	0.0138	0.0077
Inland Scotland & Ireland	972	317	0.3261	0.5363	120	0.1235	0.9626	0.0039	0.0140	0.0078

(b)											
	$\theta_k$	95% CI		$\theta$ estimators		$\theta_\pi$ SD		$D_{obs}$	Tajima's $D$		P-value
				$\theta_s$	SD				$D_{sim}$	SD	
Iceland	58.6	48.9	70.0	11.8	2.4	4.28	2.35	-1.7744	-0.1211	0.8968	0.007
Faroe Islands	10.2	6.6	15.6	5.2	1.5	2.02	1.27	-1.796	-0.0573	0.9457	0.008
Shetland	87.2	71.4	106.2	13.5	2.9	3.37	1.91	-2.1616	-0.0555	0.9105	<0.001
Orkney	26.9	17.1	42.2	6.9	2.1	3.73	2.11	-1.4812	-0.0898	0.9091	0.045
Western Isles & Skye	58.8	44.6	77.4	10.9	2.7	4.21	2.32	-1.8270	-0.0796	0.8821	0.002
Scottish NW coast	64.7	48.8	85.4	11.6	2.8	4.15	2.29	-1.9404	-0.0756	0.8942	0.001
Norway, Sweden & Denmark	182.3	154.2	211.5	17.5	3.5	3.81	2.12	-2.2230	-0.0179	0.9477	<0.001
Inland Scotland & Ireland	163.2	139.5	—	16.0	3.3	3.85	2.14	-2.1429	-0.1045	0.9108	<0.001

N = sample size, k = number of distinct haplotypes, S = number of polymorphic sites.

Several summary statistics were calculated as implemented in the Arlequin 2001 software package<sup>30</sup> (Table 1). Gene diversity measures were estimated using the formula by Nei,<sup>31</sup> defined as the probability that two randomly chosen haplotypes in a population are different. Three methods were applied to estimate the population parameter  $\theta$  as implemented in Arlequin: (i)  $\theta_k$  was estimated from the infinite-allele-mutation model equilibrium relationship between the expected number of haplotypes ( $k$ ), the sample size ( $n$ ) and  $\theta$  using Ewens' sampling formula.<sup>32</sup> (ii) Watterson's estimator<sup>33</sup> of  $\theta$ ,  $\theta_s$  is based on the infinite-site mutation model relationship between the number of segregating sites ( $S$ ), the sample size ( $n$ ) and  $\theta$ . (iii) Tajima's estimator<sup>34</sup> of  $\theta$ ,  $\theta_\pi$  is also based on the infinite-site mutation model, but on the relationship between the mean number of pairwise differences ( $\pi$ ) and  $\theta$ . Tajima's  $D$  is the normalised difference between  $\theta_s$  and  $\theta_\pi$ .<sup>35</sup>  $D$  was estimated for each population, and its significance was tested by generating random samples under the hypothesis of selective neutrality and population equilibrium.

The following measures of population differentiation were obtained using Arlequin.<sup>30</sup> Analysis of molecular variance (AMOVA) was applied and an empirical  $P$ -value was obtained by performing 3000 permutations of haplotypes among populations. In addition, exact tests of population differentiation were performed and pairwise significances were evaluated using 10 000 Markov chain steps. Pairwise genetic distances between populations were

calculated in the form of Slatkins linearised  $F_{st}$ <sup>36</sup> and presented in two-dimensional space by multidimensional scaling using the ASCAL procedure implemented in SPSS 13.0. Haplotypes were permuted 3000 times between populations to obtain the null distribution of  $F_{st}$  values and evaluate significance.

The phylogenetic relationship of the Faroese mtDNA haplotypes was analysed using a median-joining network approach as implemented in Network 4.1.1.1.<sup>37</sup> All sites and substitutions were weighted equally, assuming mutation rate homogeneity.

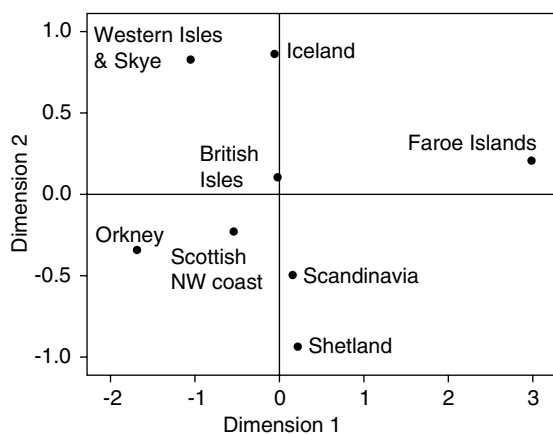
We examined the female contribution of the Scandinavian and the British Isles source populations to the putative admixed population of the Faroe Islands by applying the  $m\rho$  approach<sup>18,19</sup> on maternally inherited mitochondrial DNA sequences. The best fitting admixture model was found by searching through all possible admixture proportions from  $\eta_N=0$  to 1 in intervals of 0.0001, where  $\eta_N$  represents the proportion of either Scandinavian or British Isles founders. Randomly selected Faroese haplotypes were assigned a probability<sup>18,19</sup> of Scandinavian or British Isles ancestry, determined by their relative frequency in the two source populations. For specific 'private haplotypes' not present in the source populations, this probability was derived from the relative frequency of the presumed founder haplotype, that is, the phylogenetically closest haplotypes in terms of mutational differences.<sup>18,19</sup> An average proportion of haplotypes assigned a Scandina-

vian or British Isles origin ( $m\rho_N$ ) was obtained by performing 10 000 Monte Carlo simulations for each admixture model.

## Results

An Analysis of molecular variance approach based on the mutational differences between haplotypes<sup>38</sup> indicated that 99.2% of the genetic variation was due to variance within populations, whereas only 0.80% was due to differences among populations. This small AMOVA  $F_{st}$  value (0.00802) was, however, significantly different from 0 ( $P < 0.00001$ ), suggesting that the distribution of haplotypes among populations deviates from random expectations. The exact test was based solely on haplotype frequencies<sup>39</sup> and showed a significant difference ( $P < 0.0001$ ) between the haplotype distribution of all pairs of populations except: Scotland/Ireland vs Orkney ( $P = 0.07$ ), Scotland/Ireland vs Scottish NW coast ( $P = 0.27$ ), Scandinavia vs Orkney ( $P = 0.073$ ). In terms of haplotype frequencies, the Faroe Islands thus appeared to be significantly differentiated from all of the remaining seven neighbouring populations in the North Atlantic region. Furthermore, the Faroese population appeared as an outlier with regard to pairwise genetic distance (Figure 1) indicating its genetic isolation. All pairwise measures of Slatkin's  $F_{st}$  were significantly different from 0 except that between the British Isles sample and the Scottish NW-coast.

As opposed to the AMOVA approach and Slatkin's  $F_{st}$ , the exact tests do not take mutational differences between haplotypes into account when testing for population differentiation. The significance of almost all pairwise comparisons of the exact test therefore suggests the detected population structure to be due to the differential distribution of individual haplotypes rather than to



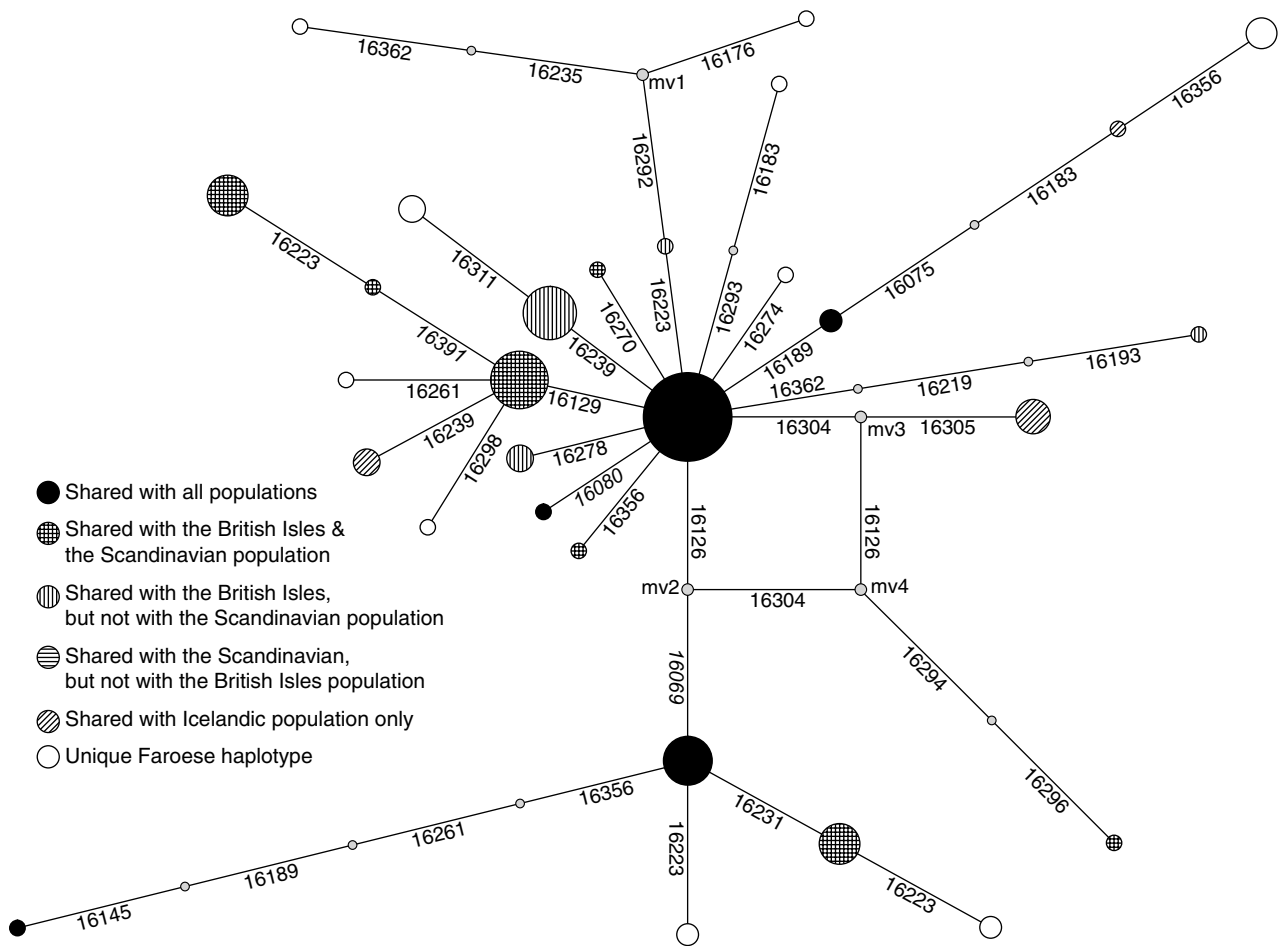
**Figure 1** Multidimensional scaling of the genetic distance between populations, measured as Slatkin's linearized  $F_{st}$ , calculated on the basis of coalescence times for mitochondrial haplotypes within and between populations.

differences in the distribution of deep-rooted phylogenetic clades among populations.<sup>19</sup>

The diversity statistics in Table 1 indicate that smaller, geographically isolated island populations (Faroe Islands and Iceland) have fewer haplotypes per number of individuals sampled than mainland populations, which are likely to have a larger effective population size. Small island populations (Western Isles and Skye, Orkney, Shetlands) relatively close to the mainland have a higher number of haplotypes per number of individuals, probably attributable to a higher level of gene flow from the mainland compared to the more remote island populations (Iceland and the Faroe Islands). Values of Tajima's  $D$  close to zero are generally interpreted to reflect selective neutrality and constant population size, whereas negative values are interpreted to indicate either a selective sweep or a past population expansion affecting a single locus under the first scenario or all loci under the latter scenario.<sup>35,40</sup> In the context of European HVR I sequences, the low values of the population parameter  $\theta$  ( $\theta_k$ ,  $\theta_S$  or  $\theta_\pi$ ) and the relatively high values of Tajima's  $D$  indicate that the Faroe Islands and Iceland have the smallest effective female population size.<sup>21</sup> The Faroe Islands also appeared to have a substantially lower level of gene diversity than the other populations (Table 1).

The correlation between  $\theta_k$  and the proportion of private haplotypes<sup>19</sup> was  $r = 0.749$  ( $P = 0.032$ ), with Iceland above and the Scottish NW-coast below the predicted 95% confidence limit, as the only two populations found outside the 95% confidence region. This indicates that the Icelandic population has a relative excess of private haplotypes, and the Scottish NW-coast has a deficit of private haplotypes. The Faroe Islands and the Shetlands were placed above the regression line at the 85% confidence limit, suggesting that the Faroe Islands and the Shetlands have a modest excess of private haplotypes. Leaving out Iceland produces a correlation of  $r = 0.816$  ( $P = 0.025$ ) and places the Faroese population at the upper 92% confidence limit, suggesting that the inclusion of the larger Icelandic gene pool makes it appear less isolated. Genetically isolated populations are expected to have an excess of private haplotypes,<sup>19</sup> which is consistent with Iceland, the Faroe Islands, and the Shetlands being the most geographically remote populations.

Median-joining network analysis was performed using all 141 Faroese mtDNA sequences based on positions 16 090–16 365 (corresponds to the positions used in the comparative analysis and revealed 28 distinct haplotypes, network not shown) and using only the 122 sequences obtained in this study based on positions 16 026–16 485 (revealed 29 distinct haplotypes). The overall pattern did not differ between the two networks, but the latter provided a slightly better phylogenetic resolution (Figure 2).

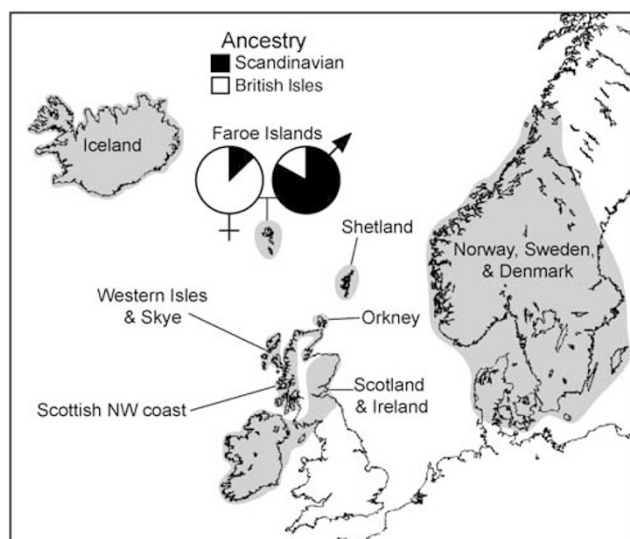


**Figure 2** Median joining haplotype network representing the phylogenetic relationship between the 29 observed Faroese haplotypes. This analysis was based on the 122 sequences obtained in the current study and position 16026 to 16485. In order to obtain more phylogenetic resolution and make use of the full-length of sequences obtained in the current study, the 19 Faroese sequences by Miller<sup>25</sup> were excluded. Numbers on branches refer to the positions undergoing mutational changes from one distinct haplotype to another, positions are numbered according to Anderson *et al.*<sup>27</sup> Numbers in italics refer to the numbers outside position 16090 to 16356 not used in the haplotype sharing analysis to allow comparison with previously published data. The haplotype sharing is thus based on a collapsed network (not shown) with 28 distinct haplotypes based on position 16090 to 16356 and all 141 Faroese sequences. Each circle represents a single distinct haplotype and their sizes are proportional to the frequency of that particular haplotype in the Faroese sample. According to the legend to the left in the figure, different shading of circles indicates with which populations the Faroese population shares that particular haplotype. Light grey circles, denoted mv (median vectors), represent ancestral nodes not present in the sample.

Figure 3 shows a map of the NW Europe and matrilineal and patrilineal admixture proportions in terms of  $m\rho$  estimated from mtDNA (current study) and Y-chromosomal data (reanalysed using  $m\rho$  by Goodacre *et al.*<sup>19</sup>), whereas Table 2 presents the admixture proportions of the Faroese Islands in relation to previously estimated  $m\rho$  admixture proportions for the remaining North Atlantic populations.<sup>19</sup> The proportion of Scandinavian matrilineal ancestry of the Faroese population was only 16.7% ( $CI_{0.95}$ : 0.114–0.227), whereas the previous estimate of Scandinavian patrilineal ancestry was 87%<sup>19</sup> ( $CI_{0.95}$ : 0.809–0.921, Helgason personal communication). British Isles matrilineal ancestry of the Faroese populations was 83.3% ( $CI_{0.95}$ : 0.773–0.886).

**Discussion**

The current study suggests that only about 17% of the female settlers of the Faroe Islands were of Scandinavian descent, whereas a much larger 83% had British Isles ancestry. Previous studies<sup>19,20</sup> suggested that 87% of the male settlers were of Scandinavian descent, with only 13% having British Isles ancestry. These results may, however, have been affected by the high level of genetic drift occurring in the Faroese population. The applied admixture approach should, however, take into account uncertainty in estimates of haplotype frequencies in the admixed and the ancestral populations due to the effects of genetic drift subsequent to admixture.<sup>18</sup> The results of the admixture analysis are inevitably influenced by the



**Figure 3** A map of the North Atlantic region showing the geographical position of the eight populations. The pie chart represents female admixture proportions ( $m_p$ ) for the Faroe Islands estimated using the mtDNA data set (left pie) in relation to male admixture proportions ( $m_p$ )<sup>21,39</sup> for the Y-chromosomal data set (right pie).

**Table 2** Admixture proportions

Admixed population	Scandinavian		British	
	Y-chr	mtDNA	Y-chr	mtDNA
Shetland	0.445	0.430	0.555	0.570
Orkney	0.310	0.305	0.690	0.695
Scottish NW coast	0.150	0.145	0.850	0.855
Western Isles & Skye	0.225	0.110	0.775	0.890
Faroe Islands	0.870	0.167 <sup>1</sup>	0.130	0.833 <sup>1</sup>
Iceland	0.750	0.340	0.250	0.660

Scandinavian and British Isles patrilineal and matrilineal ancestry based on the  $m_p$  procedure for estimating admixture for Y-chromosome and mitochondrial haplotypes. All estimates, except 1, were obtained from Goodacre *et al* (2005).<sup>19</sup>

choice of source populations. The British Isles sample included 169 (~21%) mtDNA sequences more than the Scandinavian sample, and this may have biased the estimate towards increased British Isles ancestry. For the Y-chromosomal data set, 279 (~90%) more individuals were used for the British Isles data set than for the Scandinavian data set,<sup>19</sup> suggesting that the observed 87% Scandinavian ancestry may represent an underestimation.

The original study of the Faroese male admixture proportions<sup>20</sup> suggested that 50% of the Faroese Y chromosomes had Scandinavian ancestry, whereas 23% of haplotypes were assigned to the British Isles and 27% to the Icelandic population. This study used a different approach for assigning y-chromosomal haplotypes, but the data set was reanalysed by Goodacre *et al*<sup>19</sup> using the  $m_p$  approach, and is therefore directly comparable with the present study on Faroese mtDNA sequences. The method used by

Jorgensen *et al* does not rely on the assumptions of linkage equilibrium, but the exact performance of this method using completely linked markers remains unexplored.<sup>41</sup> The  $m_p$  approach<sup>16,18</sup> takes into account the nonrecombining nature of the Y chromosome. Both analyses leave little doubt that a majority of Scandinavian ancestry existed among the male settlers of the Faroe Islands. The study by Jorgensen *et al*<sup>20</sup> consider potential ancestral populations separately instead of pooling contemporary samples likely to represent the same ancestral source population.<sup>19</sup> In the current study we have adopted the latter approach for two main reasons. (i) The exact constitution of the original source populations remain unknown and the original source populations are therefore best represented by the combined samples of their descendants. These contemporary populations are closely related which justifies the pooling into a Scandinavian and a British Isles source population. (ii) The  $m_p$  admixture approach,<sup>18</sup> which seems superior to other methods in the present context,<sup>19</sup> is currently not extended to deal with more than two ancestral populations. Standard frequency-based methods for estimating admixture proportions do not take into account private haplotypes of the admixed population<sup>42</sup> and are therefore unlikely to be informative for the current data set as the Faroese sample has a high proportion of private haplotypes (Table 1). Admixture estimates ( $m_y$ ) based on the mean coalescence times of haplotype pairs do take private haplotypes into account, but estimates are only accurate if the ancestral populations are substantially differentiated,<sup>43</sup> which is not the case for the current data set (Figure 1).

Leaving out Iceland, Shetland, Orkney, the Western Isles & Skye and the Scottish NW-coast as potential source populations is reasonable as they have probably been heavily affected by the same migration processes that gave rise to the Faroese population, according to historical and genetic evidence. Of the 28 distinct Faroese haplotypes, 10% were shared only with the Icelandic population and an additional nine haplotypes (32%) with the Icelandic and other populations. This high degree of similarity with Iceland suggests that the female settlers originated from the same regions<sup>18,19</sup> and settled the archipelago and Iceland at about the same time. Alternatively, the high degree of similarity could be explained by an extensive immigration from the much larger population of Iceland to the Faroe Islands during later periods. However, no support for the latter scenario exists in the historical records.

The present study supports the historical, archaeological and linguistic record by suggesting that a considerable number of the first female settlers in the archipelago originated from the British Isles whereas the majority of the males originated from Scandinavia.<sup>20</sup> Goodacre *et al*<sup>19</sup> conclude that Scandinavian settlement of the North Atlantic region during the Viking Age was primarily family-based in the closer and more secure areas (Orkney,

Shetland, Scottish NW coast), whereas pronounced male-biased settlement occurred at the 'frontier' (the Western Isles, Skye, Iceland). Lone Viking males, who later established families with British Isles females, were thus more prominent in the remote and less secure areas. The asymmetry in female and male ancestry proportions observed in the Faroe Islands thus fit well into the pattern of male and female admixture proportions of the North Atlantic region. The population of the Faroe Islands appears to have even more discrepant admixture proportions among male and female settlers than the putative 'sister' population in Iceland.<sup>16,18,19</sup> It thereby exhibits the greatest discrepancy in male vs female ancestry proportions of all populations in the North Atlantic region (Table 2).

Traditional statistics, such as gene diversity and mean pairwise mutational differences, have recently been suggested to be unreliable indicators of the degree of homogeneity and varying effective population size among closely related, recently diverged populations.<sup>21</sup> This is considered a problem, especially for loci with many haplotypes (such as mtDNA), where drift primarily acts to reduce the number of rare haplotypes.<sup>21</sup> Statistics based on the number of distinct haplotypes ( $\theta_k$ ) or the number of segregating sites ( $\theta_s$ , Tajima's  $D$ ) are, however, suggested to be much more sensitive to such differences in the effect of events (bottlenecks and founder events) in recent population histories, and therefore more reliable as indicators of differences in effective population size.<sup>21</sup> The Faroese population had a small number of founders, slow population growth over centuries and putative occasional reductions in population size due to epidemics, followed by recent population expansion.<sup>44</sup> Such a demographic history makes genetic drift likely to have played a major role in shaping the genetic diversity of the population. Although no severe bottleneck ( $N_e = 50-100$ ) could be detected using autosomal markers,<sup>45</sup> the above-mentioned factors are still likely to affect the genetic diversity and distribution of Faroese mtDNA haplotypes. The putative effects of increased genetic drift during founder events and bottlenecks of the Faroese population are reflected by the low number of haplotypes per sampled mtDNA sequence, the low level of gene diversity, and the small effective population size (Table 1).<sup>21</sup>

We conclude that the relative Faroese homogeneity provides strong evidence that genetic drift has played an important role in shaping Faroese genetic diversity compared to most other European populations. This will have implications for attempts to locate genes of complex disorders. In populations with increased genetic homogeneity individuals sharing a phenotype are more likely to do so because they also share genetic material than would be the case in more heterogeneous populations. This may facilitate the use of linkage and association mapping methods by reducing the problem of allelic- and locus-heterogeneity. If genetic drift has played a major role in

shaping the patterns of genetic diversity, it will also have affected the patterns of linkage disequilibrium.<sup>21</sup> Randomly selected chromosomes from the Faroe Islands appeared to contain larger shared segments (ie segments that have not been reshuffled by recombination) than chromosomes from more heterogeneous European populations.<sup>45</sup> Consequently, certain aspects of association mapping are expected to be easier on the Faroe Islands, whereas others are expected to be more difficult.<sup>46</sup>

Both of the two main focuses of this paper (i) assessing the effects of genetic drift on shaping genetic diversity and (ii) inferring population history of the Faroe Islands will be of importance when interpreting results of gene mapping studies and in choosing populations for replication. It also adds important missing information to the otherwise well-examined population history of the North Atlantic Europe.

### Acknowledgements

We are grateful to Dr A Helgason, deCODE Genetics, for kindly providing the software for the  $m\mu$  analysis and for comments on an earlier version of the manuscript. We are grateful to the Faroe House Copenhagen, Faroese community and sport organisations, and Faroese students for assistance. Three anonymous referees made valuable comments on the manuscript. Finally, we thank MM Hansen, A El Daoud & T Flint for assistance in the laboratory. Financial support was provided by 'Puljen til styrkelse af Psykiatrisk forskning i Aarhus Amt' to TDA.

### References

- 1 Jones G: *A history of the Vikings*. Oxford: Oxford University Press, 1984.
- 2 Davies N: *The Isles: a history*. London: Macmillan, 1999.
- 3 Wilson JF, Weiss DA, Richards M *et al*: Genetic evidence for different male and female roles during cultural transitions in the British Isles. *Proc Natl Acad Sci USA* 2001; **98**: 5078-5083.
- 4 Johnston G: *The Faroe Islanders' saga*. Ottawa: Oberon, 1975.
- 5 Debes HJ: *Føroya søga 1*. Tórshavn: Føroya Skúlabókagrunnur, 1990.
- 6 Young GVC: *From the Vikings to the reformation: a chronicle of the Faroe Islands up to 1538*. Onchan, Isles of Man: Shearwater Press, 2005.
- 7 Jóhansen J: *Studies in the vegetational history of the Faroe and Shetland Islands*. Thorshavn: Føroya Fróðskaparfelag, 1985.
- 8 Larsen AC, Stummann Hansen S: Viking Ireland and the Scandinavian communities in the North Atlantic; in: Larsen AC (ed): *The Vikings in Ireland*. Roskilde: The Viking Ship Museum, 2001.
- 9 Stummann Hansen S: The early settlement of the Faroe Islands; in: Barrett J (ed): *Culture contact, continuity and collapse: the Norse colonisation of the North Atlantic. Studies in the early middle ages 5*. York: University of York, 2002.
- 10 Matras C: *Greinaval-máaelfróðigreindir*. Torshavn: Føroya Fróðskaparfelag, 2000.
- 11 Passarino G, Cavalleri GL, Lin AA *et al*: Different genetic components in the Norwegian population revealed by the analysis of mtDNA and Y chromosome polymorphisms. *Eur J Hum Genet* 2002; **10**: 521-529.
- 12 Gonzalez AM, Brehm A, Perez JA *et al*: Mitochondrial DNA affinities at the Atlantic fringe of Europe. *Am J Phys Anthropol* 2003; **120**: 391-404.

- 13 Jobling MA, Tyler-Smith C: The human Y chromosome: an evolutionary marker comes of age. *Nat Rev Genet* 2003; **4**: 598–612.
- 14 Hurles ME, Nicholson J, Bosch E *et al*: Y chromosomal evidence for the origins of oceanic-speaking peoples. *Genetics* 2002; **160**: 289–303.
- 15 McEvoy B, Richards M, Forster P, Bradley DG: The longue duree of genetic ancestry: multiple genetic marker systems and celtic origins on the atlantic facade of europe. *Am J Hum Genet* 2004; **75**: 693–702.
- 16 Helgason A, Sigureth aS, Nicholson J *et al*: Estimating Scandinavian and Gaelic ancestry in the male settlers of Iceland. *Am J Hum Genet* 2000; **67**: 697–717.
- 17 Helgason A, Sigureth aS, Gulcher JR, Ward R, Stefansson K: mtDNA and the origin of the Icelanders: deciphering signals of recent population history. *Am J Hum Genet* 2000; **66**: 999–1016.
- 18 Helgason A, Hickey E, Goodacre S *et al*: mtDna and the islands of the North Atlantic: estimating the proportions of Norse and Gaelic ancestry. *Am J Hum Genet* 2001; **68**: 723–737.
- 19 Goodacre S, Helgason A, Nicholson J *et al*: Genetic evidence for a family-based Scandinavian settlement of Shetland and Orkney during the Viking periods. *Heredity* 2005; **95**: 129–135.
- 20 Jorgensen TH, Buttenschon HN, Wang AG *et al*: The origin of the isolated population of the Faroe Islands investigated using Y chromosomal markers. *Hum Genet* 2004; **115**: 19–28.
- 21 Helgason A, Nicholson G, Stefansson K, Donnelly P: A reassessment of genetic diversity in Icelanders: strong evidence from multiple loci for relative homogeneity caused by genetic drift. *Ann Hum Genet* 2003; **67**: 281–297.
- 22 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.
- 23 Sajantila A, Lahermo P, Anttinen T *et al*: Genes and languages in Europe: an analysis of mitochondrial lineages. *Genome Res* 1995; **5**: 42–52.
- 24 Richards M, Macaulay V, Hickey E *et al*: Tracing European founder lineages in the Near Eastern mtDNA pool. *Am J Hum Genet* 2000; **67**: 1251–1276.
- 25 Miller KW: *Molecular Genetic Analysis of Human Populations in Orkney and the North Atlantic Region*. Cambridge: University of Cambridge, 1996.
- 26 Opdal SH, Rognum TO, Vege A *et al*: Increased number of substitutions in the D-loop of mitochondrial DNA in the sudden infant death syndrome. *Acta Paediatr* 1998; **87**: 1039–1044.
- 27 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.
- 28 Sajantila A, Salem A, Savolainen P *et al*: Paternal and maternal DNA lineages reveal a bottleneck in the founding of the Finnish population. *PNAS* 1996; **93**: 12035–12039.
- 29 Anderson S, Bankier AT, Barrell BG *et al*: Sequence and organization of the human mitochondrial genome. *Nature* 1981; **290**: 457–465.
- 30 Schneider S, Roessli D, Excoffier L: *Arlequin: A software for population genetics data analysis (2001)*. Genetics and Biometry Lab, Department of Anthropology, University of Geneva 2001.
- 31 Nei M: *Molecular evolutionary genetics*. New York: Columbia University Press, 1987.
- 32 Ewens WJ: Sampling Theory of Selectively Neutral Alleles. *Theoretical Population Biology* 1972; **3**: 87–112.
- 33 Watterson GA: Number of segregating sites in genetic models without recombination. *Theoret Popul Biol* 1975; **7**: 256–276.
- 34 Tajima F: Evolutionary relationship of Dna-sequences in finite populations. *Genetics* 1983; **105**: 437–460.
- 35 Tajima F: Statistical-method for testing the neutral mutation hypothesis by Dna polymorphism. *Genetics* 1989; **123**: 585–595.
- 36 Slatkin M: A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 1995; **139**: 457–462.
- 37 Bandelt HJ, Forster P, Rohl A: Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 1999; **16**: 37–48.
- 38 Excoffier L, Smouse PE, Quattro JM: Analysis of molecular variance inferred from metric distances among Dna haplotypes – application to human mitochondrial-Dna restriction data. *Genetics* 1992; **131**: 479–491.
- 39 Raymond M, Rousset F: An exact test for population differentiation. *Evolution* 1995; **49**: 1280–1283.
- 40 Tajima F: The effect of change in population-size on Dna polymorphism. *Genetics* 1989; **123**: 597–601.
- 41 Cornuet JM, Piry S, Luikart G, Estoup A, Solignac M: New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics* 1999; **153**: 1989–2000.
- 42 Long JC: The genetic-structure of admixed populations. *Genetics* 1991; **127**: 417–428.
- 43 Dupanloup I, Bertorelle G: Inferring admixture proportions from molecular data: extension to any number of parental populations. *Mol Biol Evol* 2001; **18**: 672–675.
- 44 West J: *Faroe: the emergence of a nation*. London: Hurst, 1972.
- 45 Jørgensen TH, Degn B, Wang AG *et al*: Linkage disequilibrium and demographic history of the isolated population of the Faroe Islands. *Eur J Hum Genet* 2002; **10**: 381–387.
- 46 Bourgain C, Genin E: Complex trait mapping in isolated populations: are specific statistical methods required? *Eur J Hum Genet* 2005; **13**: 698–706.