

ARTICLE

Linkage disequilibrium and demographic history of the isolated population of the Faroe Islands

Tove H Jorgensen^{*1}, Birte Degn², August G Wang^{3,4}, Maria Vang³, Hugh Gurling⁵, Gursharan Kalsi⁵, Andrew McQuillin⁵, Torben A Kruse⁶, Ole Mors¹ and Henrik Ewald^{1,2}

¹Institute for Basic Psychiatric Research, Department of Psychiatric Demography, Psychiatric Hospital in Aarhus, Aarhus University Hospital, Denmark; ²Department of Biological Psychiatry, Psychiatric Hospital in Aarhus, Aarhus University Hospital, Denmark; ³Department of Psychiatry, National Hospital, Torshavn, Faroe Islands; ⁴Copenhagen University Hospital, Denmark; ⁵Molecular Psychiatry Laboratory, Department of Psychiatry and Behavioural Sciences, Windeyer Institute for Medical Sciences, Royal Free and University College London Medical School, University of London, 46 Cleveland Street, London W1T 4JF, UK; ⁶Department of Clinical Biochemistry and Genetics, Odense University Hospital, Denmark

The isolated population of the Faroe Islands has a history of recent expansion after being limited to a small size for centuries. Such an isolated population may be ideal for linkage disequilibrium mapping of disease genes if linkage disequilibrium (LD) extends over large regions. Analyses of 18 markers on 12q24.3, spanning a region of 4.3 Mb (16 cM), revealed extensive LD in the Faroese population. Maximum LD was found between marker pairs separated by more than 3.8 Mb. The same region had a maximum LD of only 1.2 and 1.4 Mb respectively in two outbred Danish and British populations analysed here for comparison. The analyses of gene diversity excess at 15 unlinked microsatellite markers did not reveal any sign of a severe bottleneck to have occurred within approximately 1200 years' history of the Faroese population. The extensive LD in this population may, therefore, have arisen primarily by random genetic drift. The implications for future gene mapping studies are discussed.

European Journal of Human Genetics (2002) 10, 381–387. doi:10.1038/sj.ejhg.5200816

Keywords: association mapping; complex diseases; bottleneck; random genetic drift; British–Danish

Introduction

Recent interest in association mapping has put focus on factors generating linkage disequilibrium (LD) in human populations. The success of association mapping, or linkage disequilibrium mapping, in locating candidate genes for a disease depends on the extent of LD between typed markers and the disease gene in the population from where the affected individuals are sampled. Various factors can influence the extent of LD in human populations. The age of the disease or marker alleles, recombination and gene conversion, natural selection and population admixture will influence LD in a population.¹ Other important factors are the size of the

population and its demographic history. Established population genetic theory predicts LD to be low in large populations under mutation-drift equilibrium whereas it will increase in small populations due to genetic drift.^{2,3} At the extreme is the situation in which a population has been through a population reduction or is founded by few individuals. This situation decreases the allelic diversity and enhances the effect of genetic drift.⁴ In populations where bottlenecks have been followed by rapid expansion, genetic drift has been suggested to have little effect on the amount of linkage disequilibrium.⁵ Predictions about the role of genetic drift in growing human populations are still under debate, though, and may vary with the fecundity distribution of the population in question.⁶ If a rapidly expanding population has been little influenced by genetic drift, a disease gene that was present in one or a few founders may be surrounded by a single or limited

*Correspondence: Tove Hedegaard Jorgensen; Department of Psychiatric Demography, Aarhus University Hospital, Skovagervej 2, 8240 Aarhus, Denmark. Tel: +45 77892805, Fax +45 77892899;
E-mail: demograf@psykiatri.aaa.dk

Received 8 November 2001; revised 1 March 2002; accepted 9 April 2002

number of marker haplotypes if the LD between these markers is maintained in the expanding population.⁷ In small populations of constant size, though, it is not necessarily a single marker haplotype at a locus that has increased to a high frequency through genetic drift. In these populations several marker haplotypes may be found associated with a disease gene.⁷

Based on the predictions from population genetic theory we expect isolated populations that have gone through severe bottlenecks or which have remained at a constant small size for a long time to have extensive linkage disequilibrium compared to outbred populations. Such inbred and selected populations tend to be ideal for the linkage disequilibrium mapping of disease genes. Several monogenic diseases have been mapped successfully by this method in isolated populations.^{8,9} Expectations are that linkage disequilibrium mapping can be equally useful in genetic studies of complex diseases,^{10,11} and that the same advantages of using population isolates may apply to diseases with more than one gene involved, depending on the underlying genetic model.^{12,13} So far the method has been used to locate a number of candidate regions that may harbour risk genes for complex diseases.^{14,15}

Empirical assessments of the levels of LD in population isolates do not always conform to expectations based on the historic record. Modest levels of LD in the Finnish and Sardinian isolates suggest introduction of common variants by multiple founders.^{16–18} Similar low levels of LD have been obtained in isolates such as the Afrikaners and Ashkenazi,¹⁹ whereas predicted high levels of LD have been confirmed in populations such as an expanding Costa Rican population²⁰ and in the constantly small Saami population.²¹ The purpose of the present study was to evaluate the amount of linkage disequilibrium in the population of the Faroe Islands and to compare this with its demographic history as revealed from microsatellite data. The historic record suggests that the Faroe Islands is an ideal population isolate in which to search for candidate genes of various genetic diseases²² and the interpretation of the amount of LD found in this population is important for such future studies. Linkage disequilibrium between microsatellite markers on 12q was compared to that of two larger and less isolated populations, the British and the Danish. Our results confirm our expectations of high levels of LD in the Faroese compared to the two mixed outbred populations, but do not indicate the occurrence of a severe population bottleneck.

Materials and methods

History of the Faroese

Historic records suggest that the isolated Faroese archipelago in the North Atlantic Ocean (Figure 1) was settled around 825 AD by emigration from the western part of Norway.²³ Paleo-botanical findings of cultivated species and associated weeds does, however, indicate the presence

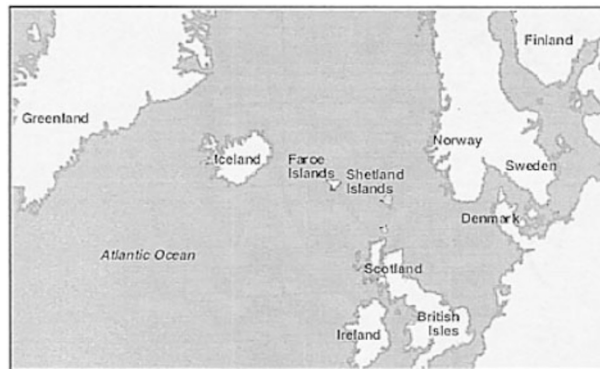


Figure 1 The Faroe Islands in the North Atlantic Ocean.

of people on the islands two centuries earlier.²⁴ These people had possibly left the archipelago before the Norse settlement.^{23,24} Population size in the archipelago as a whole may have been as small as 4000 in the late 1300s increasing to 9000 in the 1800s and further to around 48000 inhabitants at present.²⁵ A count of 4773 inhabitants was made in 1769.²⁵ Migration to and from the Faroe Islands has been sparse as a consequence of their geographic position, the commercial monopoly prevalent from 1380 until 1856 and legislation aimed at avoiding depopulation. Epidemics, such as a smallpox epidemic in 1709, and periods of famine may have reduced the population severely.²⁵ Genealogical investigations of Faroese families show that considerable movements occurred between villages despite the remote positions that many of them have (AG Wang, pers. comm.).

Sampling of individuals

Forty-three unrelated Faroese individuals, 120 unrelated Danish individuals and 316 unrelated British individuals were sampled from various regions of the three countries. The Faroese are part of a control population in an ongoing study of the genetic aetiology of complex diseases. British individuals were mainly of English origin although a small number had Irish, Welsh or Scottish parents.

Genotyping

DNA was extracted from whole blood using a standard Triton lysis protocol with sodium chloride/isopropanol precipitation or a phenol/chloroform method.²⁶ Single or multiplex PCR was performed with amplification of a maximum of three markers simultaneously and analysed on an ABI Prism 310 Genetic Analyser or a Li-Cor IR2 automated sequencer.

Fourteen markers on 12q24.3, spanning a region of 2.2 Mb (5 cM), were genotyped for the analysis of linkage disequilibrium (Table 1). Three of these were typed only for the British and Danish individuals. An additional four markers positioned distal to the 2.2 Mb region were typed in the Faroese so that a total of 4.3 Mb (16 cM) were investigated in this population.

Table 1 Gene diversity (*D*) and allelic diversity (*n*) with corrections according to the rarefaction method in brackets

	British		Danish		Faroes	
	<i>n</i>	<i>D</i>	<i>n</i>	<i>D</i>	<i>n</i>	<i>D</i>
D12S1614	7 (6)	0.72	6 (6)	0.69	6	0.51
D12S342	13 (10)	0.76	13 (10)	0.75	8	0.75
AFMB332ZC1	6 (5)	0.61	6 (5)	0.65	–	–
D12S340	7 (5)	0.66	5 (4)	0.67	5	0.66
D12S1639	14 (11)	0.76	14 (12)	0.83	9	0.76
D12S324	7 (5)	0.65	5 (5)	0.65	5	0.65
D12S866	8 (6)	0.72	10 (7)	0.66	–	–
D12S386	17 (9)	0.75	10 (8)	0.75	6	0.76
AFM337ZD5	5 (4)	0.62	5 (5)	0.63	–	–
DS12S1634	11 (8)	0.53	9 (8)	0.62	6	0.42
D12S307	9 (6)	0.34	7 (5)	0.39	6	0.36
D12S1658	9 (7)	0.76	8 (7)	0.77	6	0.76
GATA41E1	8 (7)	0.80	7 (7)	0.80	6	0.75
D12S2075	8 (7)	0.82	8 (7)	0.78	7	0.77
D12S1675	–	–	–	–	7	0.76
D12S1659	–	–	–	–	8	0.79
D12S1714	–	–	–	–	8	0.79
D12S97	–	–	–	–	7	0.66
Mean <i>D</i>		0.69		0.70		0.65
2 <i>N</i>		632		240		86

2*N* is the total number of chromosomes in the investigation. Mean *D* is calculated from the 11 loci shared by all three populations.

Table 2 Test for the presence of a historic population bottleneck in the Faroese population

<i>Locus</i>	<i>n</i>	<i>H_o</i>	<i>H_{eq}</i>	<i>SD_{eq}</i>	<i>DH/sd</i>	<i>P</i>
D1S1631	8	0.831	0.804	0.043	0.640	0.311
D1S518	8	0.830	0.804	0.042	0.618	0.302
D2S1329	5	0.380	0.675	0.081	–3.638	0.011
D2S1328	6	0.791	0.731	0.069	0.855	0.166
D3S1744	8	0.765	0.804	0.044	–0.885	0.163
D7S2201	4	0.625	0.594	0.110	0.278	0.488
D7S1824	11	0.888	0.863	0.027	0.920	0.163
D8S1119	7	0.766	0.772	0.054	–0.116	0.357
D8S373	6	0.741	0.730	0.065	0.174	0.498
D11S2001	6	0.788	0.729	0.066	0.887	0.166
D12S1052	5	0.722	0.674	0.084	0.576	0.312
D12S2075	7	0.779	0.773	0.053	0.101	0.445
D13S787	6	0.643	0.730	0.063	–1.370	0.096
D13S317	7	0.740	0.772	0.055	–0.574	0.215
D22S1169	7	0.761	0.771	0.055	–0.191	0.335

Expected number of loci with heterozygosity excess: 8.85
Probability of observed: 0.579

Listed is the allelic diversity (*n*) and observed (*H_o*) and expected (*H_{eq}*) heterozygosities at 15 unlinked loci. *H_{eq}* is generated by simulations under the assumption of mutation drift equilibrium. *DH/sd* is the standardised difference in observed and expected heterozygosity (*H_o*–*H_{eq}*)/*SD* and *P* the probability of the observed heterozygosity at a given locus.

Fifteen unlinked microsatellite markers were typed in 41 individuals from the same sample of Faroese for the demographic analyses (Table 2). All were tri- or tetra nucleotide markers assumed to evolve under the stepwise mutation model.²⁷

Allele and gene diversity

Allelic diversity is the number of alleles at a locus in a population. Differences in sample sizes were corrected for by the rarefaction method^{28,29} using the programme CONTRIB (www.pierroton.inra.fr/genetics/labo/Software/). This method adjusts for the fact that microsatellites will have rare alleles because of their relatively high mutation rate and that these rare alleles are less likely to be detected with smaller sample sizes. A minimum of 84 chromosomes was successfully typed at every locus in all three populations and allelic diversity in the three populations was therefore standardised to a sample size of 84 chromosomes using the rarefaction method.

Nei's average heterozygosity or *gene diversity* was calculated as a measure of variation at a locus *l* with *u* alleles: $D_l = 1 - \sum p_{li}^2$ and as a population mean over *m* loci: $D = 1 - 1/m \sum \sum p_{li}^2$.^{4,30} Therefore *D* is the probability that two alleles sampled at random are different. Only the loci shared between all three populations were used in the calculation of *D*.

Linkage disequilibrium

Tests for Hardy–Weinberg equilibrium were performed on genotypic data as a prerequisite for further analyses of linkage disequilibrium using the method of Guo and Thompson.³¹

The 43 Faroese consisted of 22 parental couples that had their haplotypes reconstructed from one offspring. To be able to reconstruct the phase of the two parents' genotypes from the genotype of the offspring we need to assume that no recombination in this one generation has occurred at the segment on 12q investigated. Exact tests for linkage disequilibrium between all pairwise markers were then performed using the haplotype data. The test is an extension of Fisher's exact probability test on contingency tables, which evaluates the probability of tables with the same marginal totals and a probability equal or less than the observed table.⁵ It uses a Markov Chain to explore the space of all possible tables.³¹

Gametic phase was unknown in the British and Danish populations and pairwise tests for linkage disequilibrium were therefore performed using a likelihood ratio test whose empirical distribution was obtained by a permutation procedure.³² In this test the haplotype frequencies were estimated using the EM algorithm to obtain the likelihood of the data not assuming linkage equilibrium. The EM algorithm has a very high ability of inferring correct haplotypes when sample sizes are large (>100 individuals) as is the case in this investigation.³³

All three tests are implemented in Arlequin version 2.00 and performed with 1000 permutations.³⁴

Population bottleneck and demography

The number of alleles at a locus is reduced faster than the gene diversity at that locus when a population is reduced

in size since rare alleles are more readily affected by drift than more frequent ones.³⁵ As population size is restored the average number of alleles increases faster than the gene diversity until reaching mutation-drift equilibrium.³⁵ After going through a bottleneck, a population will have a transient excess in gene diversity compared to that expected for the allele number. A test for the presence of a recent genetic bottleneck can be based on the differences in the observed and expected gene diversities in a population across a large set of unlinked loci.³⁶ For such a test, the program Bottleneck³⁷ computes the gene diversity H_{eq} expected at each locus from the observed number of alleles given the sample size N under the assumption of mutation-drift equilibrium. This is done through simulation of the coalescent process from N genes under the stepwise mutation model. Power analyses and theoretical models suggest that a 100-fold reduction in population size can be detected with a 60–80% probability within a range of $0.25 \times 2N_e$ to $2N_e$ generations after the reduction occurred.³⁶ N_e is the post-change effective population size. The power for the detection of gene diversity excess increases as the relative reduction in population size becomes larger. Based on these power analyses we predict that a founder event on the Faroe Islands approximately 1200 years ago with a post-bottleneck size (N_e) of up to a few hundred will be detected with a probability of at least 60% in our analysis.³⁶ This is assuming that the founder population is at least 100 times smaller than the source population. The method has proven successful in detecting past changes in population size in a series of populations including humans.³⁸

Results

Allelic and gene diversity

All markers for the linkage disequilibrium analyses were in Hardy–Weinberg equilibrium in all the three populations (results not shown). Eight of the 11 loci typed in all three populations had lower allele counts in the Faroese population and in six of these this persisted after correcting for differences in sample sizes (Table 1). Marker D12S1614 had relatively low gene diversity in the Faroese in relation to its high allele number. All loci contained rare alleles that were present in just one population. Twenty unique alleles were found in the British, whereas the Danish had five and the Faroese population only one. All unique alleles had low frequencies not exceeding 1% in the British and Danish populations. The unique allele in the Faroese was present in only two copies. Very few of the remaining alleles present in all three populations had frequencies below 1%. These frequency patterns may reflect a sampling problem and not differences in the demographic history of the populations. The numbers obtained using the rarefaction method may be a better measure of allelic diversity. Also, it should be kept in mind that this sampling bias may lead to higher estimates of LD in the British and Danish samples compared to the Faroese due to the sensi-

tivity of this estimate to the number and frequencies of alleles. Overall, the mean gene diversity in the Faroese population is slightly smaller than in the other two populations. The results suggest different demographic histories of the three populations and may reflect a founder effect in the Faroese population. No formal test of differences in allelic and gene diversities are performed due to the lack of independence of the loci.

Linkage disequilibrium

There was extensive LD between markers on 12q in the Faroese compared to the other two populations (Figure 2). Linkage disequilibrium extended as far as 3.8 Mb in the Faroese population (marker D12S1639 and D12S97) while marker pairs in LD were separated by a maximum of 1.4 Mb in the British (D12S866 – D12S2075) and 1.2 Mb in the Danish populations (D12S342 – AFMB337ZD5). No corrections for multiple tests were performed due to the lack of independence of marker pairs. A summary of the number of different haplotypes at each segment in a population reveal no decreased variability in the haplotypes found in the Faroese when taking the lower sample size of this population into account (Figure 3). LD correlated with physical distance in the Danish populations (Mantel test with 10 000 permutations, $r=0.377$, $P=0.002$), had a non-significant correlation coefficient in the British ($r=0.180$, $P=0.099$) and a low and non-significant correlation coefficient in the Faroese population ($r=0.020$, $P=0.420$).^{34,39}

Population bottleneck and demography

All 15 markers for the demographic analysis were polymorphic (Table 2) and all loci conformed to Hardy–Weinberg expectations (results not shown). Nine loci had gene diversity excess compared to the 8.85 loci expected. The probability of obtaining this result is 0.58. There is therefore no significant gene diversity excess and no indication of a recent genetic bottleneck in the Faroese population. There is also no significant gene diversity deficiency across loci (Wilcoxon one-tailed test, $P=0.68$). A consistent gene diversity deficiency across loci would be indicative of a recent population expansion.³⁶

Discussion

Increased levels of linkage disequilibrium on 12q in the Faroese relative to the British and Danish populations reveals prominent differences in the demographic history of the three populations. Levels of allelic diversity at micro-satellite loci on chromosome 12q are somewhat lower than found in outbred British and Danish populations. However, there is no sign of a general gene diversity excess at other unlinked loci neither are the haplotypes longer or less variable in the Faroese population compared to the other two populations. There is, therefore, no evidence of a strong population bottleneck in the Faroese population. It is possi-

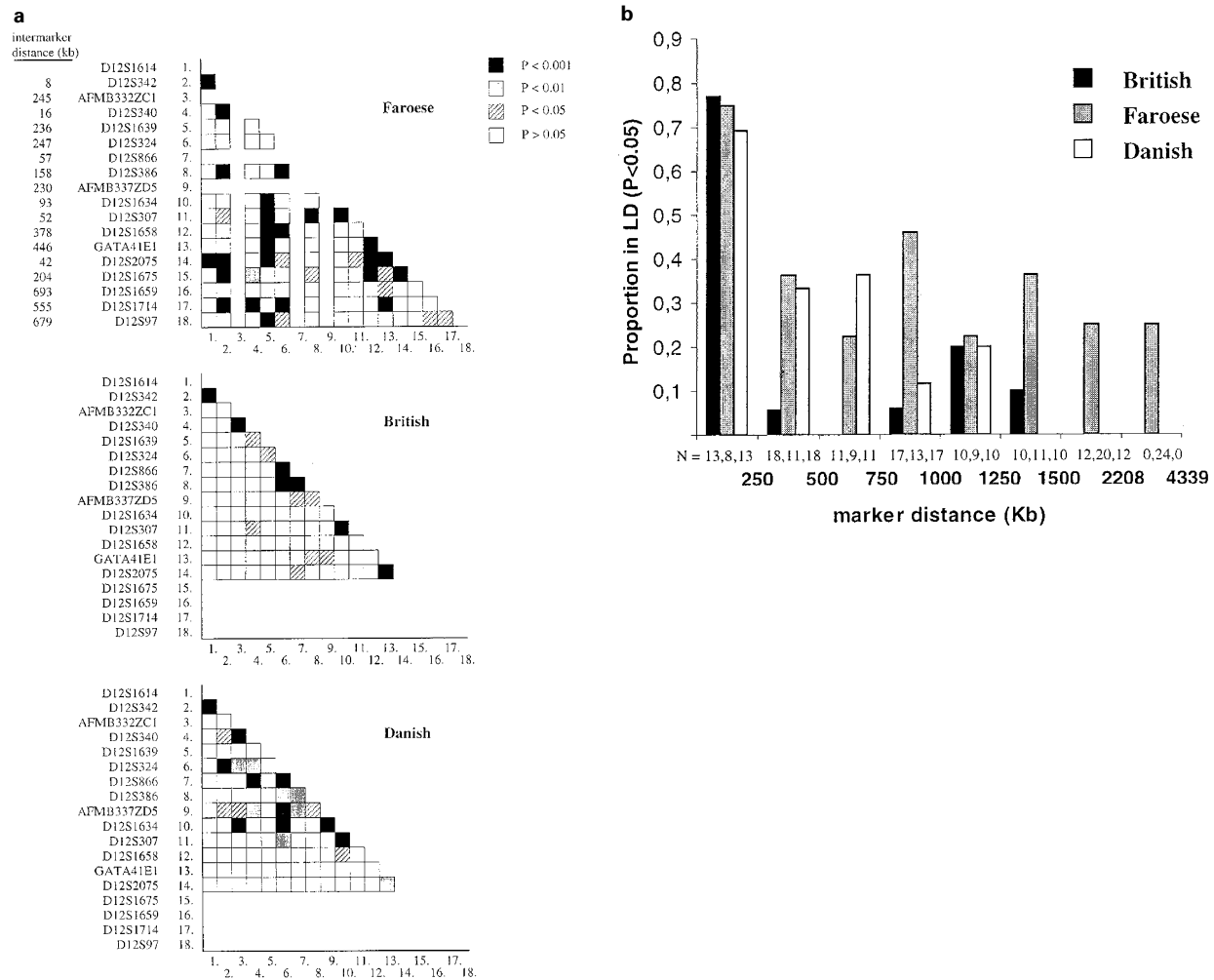


Figure 2 (A) Fisher's exact test of pairwise linkage disequilibrium between microsatellites. No correction for multiple tests were performed. Distances between neighbouring markers are indicated in Kb. (B) Summary of the frequency of marker pairs in linkage disequilibrium. N is the total number of marker pairs tested within the distance interval. Notice the non-linear scale above 1500 Kb.

ble that the high levels of LD are instead due to the general effect of genetic drift in a small population. The total population on the Faroe Islands remained at a small size through centuries until the rapid expansion in the 1800s. Simulation studies on human populations supplement predictions from population genetic theory that such slow early growth can be one of the key factors responsible for increasing the extent of LD in an isolated population as compared to an early rapid expansion.¹³ Similar simulation studies can be made on the Faroese population provided that relevant demographic information can be obtained.

Our results confirm that some population isolates have elevated levels of LD and may therefore be valuable in linkage disequilibrium mapping of genetic diseases even without the use of dense marker maps. The estimate by Kruglyak¹³ of an average (biallelic) marker distance of

3 Kb needed to detect useful levels of LD in an outbred population is much exceeded by this, and some of the other rare empirical studies of LD in isolated populations.^{20,21,40} Estimating marker densities to be used in linkage disequilibrium mapping in the Faroese population may not be appropriate at the present state due to the non-uniform distribution of LD known to occur on the human genome.⁴¹

The demographic history revealed by the genetic data may have some important implications for gene mapping studies. If no strong bottleneck has been present in the Faroese population we should not expect a single haplotype to be shared by the affected disease gene. This is particularly true if the disease has a complex background and is attributable to a number of common alleles that could have entered the population through the multiple founders.⁷

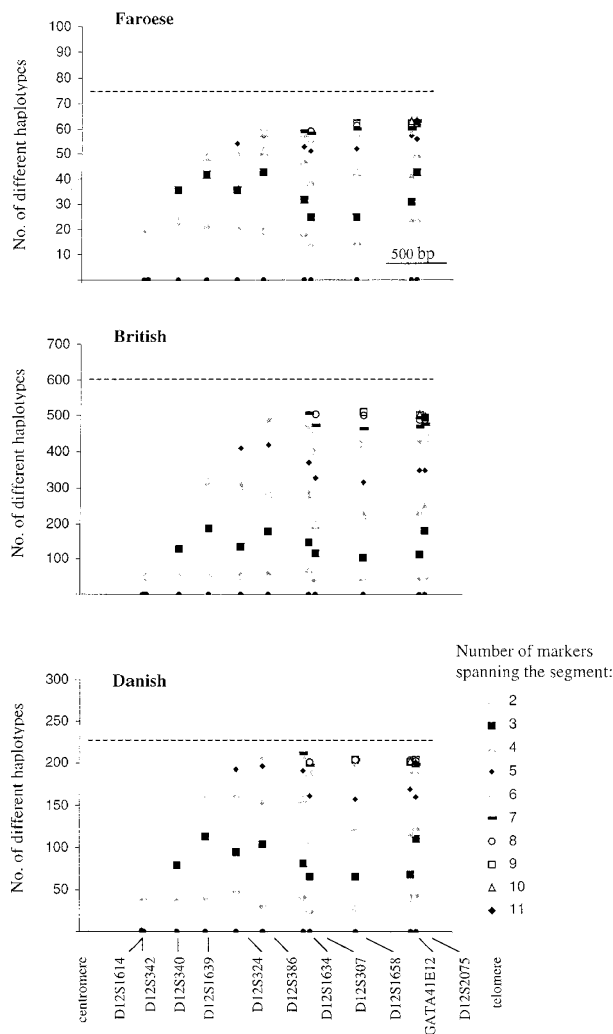


Figure 3 The number of different haplotypes found in a population. A total number of 86,632 and 240 chromosomes were typed in the Faroese, British and Danish populations respectively. A few missing data lowered the total number of haplotypes that could be determined in each population from the maximum possible. Maximum values were 84 in the Faroese, 628 in the British and 238 in the Danish population. Minimum values were 70, 574 and 216. The stippled lines indicate the average number of haplotypes determined in each population. Each point in the graph refers to a segment with the marker on the x-axis at the right end of that segment. Haplotypes in the British and Danish populations were estimated from phase-unknown genotypes using the EM algorithm.

The test for gene diversity excess used here only detects bottlenecks of relatively small size. It does, however, include the range of 10–100 N_e suggested to be required for the population reduction to result in increased levels of LD in post-bottlenecked populations.¹³

It is well known that LD may not necessarily be associated with the physical distance between markers.⁴² Although an

overall correlation between LD and physical distance is present in the Danish populations and a positive but non-significant correlation is present in the British population, all three populations have several examples of close neighbouring markers that are in linkage equilibrium. It has been suggested that such a pattern can be due to gene conversions that break up LD between closely linked markers but have negligible effect on more distant sites.^{42,43,44} LD extends over long regions in the Faroese population and after identification of a candidate region by association analyses additional typing and haplotype analyses will be needed to narrow down the gene of interest. That LD is not completely predictable from small physical distances has obvious consequences for such fine-mapping. The test for linkage equilibrium at a marker closely linked to the disease locus may be significant and haplotypes around a disease locus may be broken up, which will hamper the localisation of the disease locus. Analyses of multiple close markers in a candidate region will therefore be necessary for gene localisation.

Acknowledgements

Financial support was given by The Lundbeck Foundation (grant no. 39/98 and 59/99); Ingeborg and Leo Dannin; Ministry of Education, Culture and Research, The Faroese Government and The Research Fund of the Faroese Savings Bank, Torshavn, Faroese Islands

References

- 1 Wright AF, Carothers AD, Pirastu M: Population choice in mapping genes for complex diseases. *Nat Genet* 1999; **23**: 397–404.
- 2 Hill WG, Robertson A: Linkage disequilibrium in finite populations. *Theor Appl Genet* 1968; **38**: 226–231.
- 3 Ohta T, Kimura M: Linkage disequilibrium at steady state determined by random genetic drift and recurrent mutation. *Genetics* 1969; **63**: 229–238.
- 4 Nei M: *Molecular population genetics and evolution*. American Elsevier: New York, 1975.
- 5 Slatkin M: Linkage disequilibrium in growing and stable populations. *Genetics* 1994; **137**: 331–336.
- 6 te Meerman GJ, Van der Meulen MA: Genomic sharing surrounding alleles identical by descent: effects of genetic drift and population growth. *Genet Epidemiol* 1997; **14**: 1125–1130.
- 7 Terwilliger JD, Zollner S, Laan M, Paabo S: Mapping genes through the use of linkage disequilibrium generated by genetic drift: 'drift mapping' in small populations with no demographic expansion. *Hum Hered* 1998; **48**: 138–154.
- 8 Houwen RH, Baharloo S, Blankenship K, Raeymaekers P, Juyn J, Sandkuijl LA, Freimer NB: Genome screening by searching for shared segments: mapping a gene for benign recurrent intrahepatic cholestasis. *Nat Genet* 1994; **8**: 380–386.
- 9 Peltonen L, Jalanko A, Varilo T: Molecular genetics of the Finnish disease heritage. *Hum Mol Genet* 1999; **8**: 1913–1923.
- 10 Risch N, Merikangas K: The future of genetic studies of complex human diseases. *Science* 1996; **273**: 1516–1517.
- 11 Martin ER, Lai EH, Gilbert JR *et al*: SNPing away at complex diseases: analysis of single-nucleotide polymorphisms around APOE in Alzheimer disease. *Am J Hum Genet* 2000; **67**: 383–394.

- 12 Peltonen L, Palotie A, Lange K: Use of population isolates for mapping complex traits. *Nat Rev Genet* 2000; **1**: 182–190.
- 13 Kruglyak L: Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat Genet* 1999; **22**: 139–144.
- 14 Freimer NB, Reus VI, Escamilla MA *et al*: Genetic mapping using haplotype, association and linkage methods suggests a locus for severe bipolar disorder (BPI) at 18q22-q23. *Nat Genet* 1996; **12**: 436–441.
- 15 Degn B, Lundorf M, Wang AG *et al*: Further evidence for a bipolar risk gene on chromosome 12q24 suggested by investigation of haplotype sharing and allelic association in patients from the Faroe Islands. *Mol Psychiatry* 2001; **6**: 450–455.
- 16 Eaves IA, Merriman TR, Barber RA *et al*: The genetically isolated populations of Finland and Sardinia may not be a panacea for linkage disequilibrium mapping of common disease genes. *Nat Genet* 2000; **25**: 320–323.
- 17 Varilo T, Laan M, Hovatta I, Wiebe V, Terwilliger JD, Peltonen L: Linkage disequilibrium in isolated populations: Finland and a young sub-population of Kuusamo. *Eur J Hum Genet* 2000; **8**: 604–612.
- 18 Taillon-Miller P, Bauer-Sardina I, Saccone NL *et al*: Juxtaposed regions of extensive and minimal linkage disequilibrium in human Xq25 and Xq28. *Nat Genet* 2000; **25**: 324–328.
- 19 Dunning AM, Durocher F, Healey CS *et al*: The extent of linkage disequilibrium in four populations with distinct demographic histories. *Am J Hum Genet* 2000; **67**: 1544–1554.
- 20 Service SK, Ophoff RA, Freimer NB: The genome-wide distribution of background linkage disequilibrium in a population isolate. *Hum Mol Genet* 2001; **10**: 545–551.
- 21 Laan M, Paabo S: Demographic history and linkage disequilibrium in human populations. *Nat Genet* 1997; **17**: 435–438.
- 22 Ewald H, Wang AG, Vang M, Mors O, Nyegaard M, Kruse TA: A haplotype-based study of lithium responding patients with bipolar affective disorder on the Faroe Islands. *Psychiatr Genet* 1999; **9**: 23–34.
- 23 Young GVC: *From the vikings to the reformation: a chronicle of the Faroe Islands up to 1538*. Onchan, Isle of Man: Shearwater Press Limited, 1979.
- 24 Jóhansen J: *Studies in the vegetational history of the Faroe and Shetland Islands*. Thorshavn: Føroya Fróðskaparfelag, 1985.
- 25 West JF: *Faroe. The emergence of a nation*. London: Hurst, 1972.
- 26 Kunkel LM, Smith KD, Boyer SH *et al*: Analysis of human Y-chromosome-specific reiterated DNA in chromosome variants. *Proc Natl Acad Sci USA* 1977; **74**: 1245–1249.
- 27 Shriver MD, Jin L, Chakraborty R, Boerwinkle E: VNTR allele frequency distributions under the stepwise mutation model: a computer simulation approach. *Genetics* 1993; **134**: 983–993.
- 28 Hurlbert SH: The nonconcept of species diversity: A critique and alternative parameters. *Ecology* 1971; **52**: 577–585.
- 29 Petit RJ, El Mousadik A, Pons O: Identifying populations for conservation on the basis of genetic markers. *Conserv Biol* 1998; **12**: 844–855.
- 30 Weir B: *Genetic Data Analysis II*. Massachusetts, Sunderland: Sinauer Associates, 1996.
- 31 Guo SW, Thompson EA: Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* 1992; **48**: 361–372.
- 32 Slatkin M, Excoffier L: Testing for linkage disequilibrium in genotypic data using the Expectation-Maximization algorithm. *Heredity* 1996; **76**: 377–383.
- 33 Excoffier L, Slatkin M: Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol Biol Evol* 1995; **12**: 921–927.
- 34 Schneider S, Roessli D, Excoffier L: *Arlequin ver. 2.00: A software for population genetic data analysis*. Genetic and Biometry Laboratory, University of Geneva, Switzerland, 2001.
- 35 Nei M, Maruyama T, Chakraborty R: The bottleneck effect and genetic variability in populations. *Evolution* 1975; **29**: 1–10.
- 36 Cornuet JM, Luikart G: Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 1996; **144**: 2001–2014.
- 37 Piry S, Luikart G, Cornuet JM: *Bottleneck: A program for detecting recent effective population size reductions from allele frequency data*. INRA, UR1B, Laboratoire de Modélisation et Biologie Evolutive, 1998.
- 38 Luikart G, Cornuet JM: Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. *Conserv Biol* 1998; **12**: 228–237.
- 39 Mantel N: The detection of disease clustering and a generalized regression approach. *Cancer Res* 1967; **27**: 209–220.
- 40 Zavattari P, Deidda E, Whalen M *et al*: Major factors influencing linkage disequilibrium by analysis of different chromosome regions in distinct populations: demography, chromosome recombination frequency and selection. *Hum Mol Genet* 2000; **9**: 2947–2957.
- 41 Huttley GA, Smith MW, Carrington M, O'Brien SJ: A scan for linkage disequilibrium across the human genome. *Genetics* 1999; **152**: 1711–1722.
- 42 Pritchard JK, Przeworski M: Linkage disequilibrium in humans: models and data. *Am J Hum Genet* 2001; **69**: 1–14.
- 43 Ardlie K, Liu-Cordero SN, Eberle MA *et al*: Lower-than-expected linkage disequilibrium between tightly linked markers in humans suggests a role for gene conversion. *Am J Hum Genet* 2001; **69**: 582–589.
- 44 Andolfatto P, Nordborg M: The effect of gene conversion on intra-locus associations. *Genetics* 1998; **148**: 1397–1399.