ARTICLE

Founder effect in spinal and bulbar muscular atrophy (SBMA) in Scandinavia

A Lund[1], B Udd[2,3], V Juvonen[1], PM Andersen[4], K Cederquist[5], L-O Ronnevi[6], P Sistonen[7], SA Sörensen[8], L Tranebjærg[9], C Wallgren-Pettersson[10,11] and M-L Savontaus[1,12]

[1]Department of Medical Genetics, University of Turku; [2]Neurological Department, Vasa Central Hospital, Vasa; [3]Department of Neurology, University of Tampere, Finland; Departments of [4]Neurology and [5]Medical Genetics, University of Umeå; [6]Department of Neurology, Karolinska Sjukhuset, Stockholm, Sweden; [7]Finnish Red Cross, Blood Transfusion Service, Helsinki, Finland; [8]Department of Medical Genetics, University of Copenhagen, Denmark; [9]Department of Medical Genetics, University Hospital of Tromsö, Norway; [10]The Folkhälsan Department of Medical Genetics, Helsinki; [11]Department of Medical Genetics, University of Helsinki; [12]Department of Biology, University of Turku, Finland

We haplotyped 13 Finnish, 10 Swedish, 12 Danish and 2 Norwegian SBMA (spinal and bulbar muscular atrophy, Kennedy disease) families with a total of 45 patients and 7 carriers for 17 microsatellite markers spanning a 25.2 cM region around the androgen receptor gene on chromosome Xq11–q12 in search of a genetic founder effect. In addition, the haplotypes of 50 Finnish, 20 Danish and 22 Swedish control males were examined. All the Scandinavian SBMA families shared the same 18 repeat allele for the intragenic GGC repeat, which was present in only 24% of the controls. Linkage disequilibrium was also seen for the closest microsatellite markers. In addition, extended haplotypes of the Finnish, Swedish and Danish SBMA families revealed country-specific common founder haplotypes, which over time became gradually shortened by recombinations. No common haplotype was found among the controls. The data suggest that the SBMA mutation was introduced into western Finland 20 generations ago. Haplotype analysis implies a common ancestor for the majority of Scandinavian SBMA patients. *European Journal of Human Genetics* (2000) **8**, 631–636.

Introduction

Spinal and bulbar muscular atrophy (SBMA), or Kennedy disease (MIM 313200), is an X-linked recessive late-onset motor neurone disorder. Patients show slowly progressive spinal and bulbar muscular atrophy with fasciculations and generalised muscle weakness, frequently sensory symptoms and signs of partial androgen insensitivity, such as gynecomastia and impotence.[1,2] SBMA is caused by a CAG repeat expansion in the first exon of the androgen receptor (*AR*) gene.[3] The gene normally has 10–36 CAG repeats,[4] but in SBMA patients the number of repeats is increased to 38–72.[5,6] Several other neurodegenerative disorders caused by trinucleotide repeat expansions have also been reported.[7–9] It has

been shown that proteins with expanded polyglutamine stretches encoded by the CAG repeats form fibrillary aggregates in neurones, leading to apoptosis.[10] Such aggregates have also been reported in SBMA.[4,6,11]

The human *AR* gene is located in Xq11–q12. The gene contains 8 exons.[3] In the 3' end of the first exon of the *AR* gene, 1.2 kb downstream from the CAG repeat there is a polymorphic GGC repeat sequence.[12] In Japanese SBMA patients linkage disequilibrium between the CAG mutation and the intragenic GGC repeat has been found, indicating a founder effect.[13] The high prevalence of SBMA among Swedish-speaking people in the district of Vaasa in Western Finland[2] and in Northern Sweden (6 in 930 000) motivated us to search for linkage disequilibrium between the mutation and 17 microsatellite markers around the *AR* gene, and also to estimate the time when the SBMA mutation was introduced into Finland. Our initial studies showed that all

13 Finnish families shared identical alleles for at least five adjacent markers. This finding prompted us to extend the study to patients from the linguistically related neighbouring Scandinavian countries in the search for a genetic founder effect.

Material and methods

Thirteen Finnish families (FI-FXIII) with a total of 17 SBMA patients and four carriers, 10 Swedish families (SI-SX) with 13 patients, 12 Danish patients from different families (DI-DXII) and two Norwegian families (NI and NII) with three patients and three carriers were haplotyped for 17 microsatellite markers spanning a 25.2 cM region around the androgen receptor gene (Figure 1). The locations and distances of the markers were obtained from The Genome Database (GDB) (http://gdbwww.gdb.org/). The haplotypes of 50 Finnish, 20 Danish and 22 Swedish control males were also examined, as well as the CAG repeat numbers in 25 Finnish control women.

After obtaining informed consent, blood samples were taken of patients with SBMA. DNA was extracted from EDTA

Figure 1 Genetic distances of the microsatellite markers and their locations on the X chromosome according to GDB.

blood. The primers for the microsatellites were from the GDB Polymorphism Detail, except for the CAG[3] and the GGC[13] repeats in the *AR* gene. The microsatellites were amplified by a standard PCR method with 50 ng of template DNA, 10 pmol of each primer, 200 μM of dNTP, 1 $\times$ PCR buffer (Finnzymes), 0.1 μCi of α-^{32}P-dCTP and 1 U of either *Taq* DNA polymerase (Promega) (CAG, GGC, ALAS2, DXS8032, DXS8062, DXS339 and DXS1213) or Dynazyme DNA polymerase (Finnzymes). In addition, the reactions for CAG and DXS8032 contained 10% glycerol, and CAG contained 2% formamide. For the GGC repeat, 7-deaza-dGTP was included in the dNTP mix, and 10% DMSO was used. The reaction volumes were 25 μl, except for GGC, in which volume was 15 μl. The initial denaturation was at 95°C for 5 min, denaturation at 95°C, annealing at 50–64°C, elongation at 72°C and a final elongation at 72°C for 4 min. Denaturation, annealing and elongation were 45 s each, except for CAG 1 min each and for GGC 1, 2 and 2 min, respectively. Thirty-five cycles were used for DXS1690, while the others went through 30 cycles. The sizes of the PCR products were analysed by comparison with an M13 DNA sequencing ladder in 4–6% denaturing PAGE gels. The CAG and GGC repeat numbers were verified by comparison with sequenced references.

The allele frequencies of the SBMA patients and the controls were compared using the Mann-Whitney test, and the CAG repeat ranges of the Finnish, Swedish and Danish patients were compared by variance analysis.

A tentative age estimate in the Finnish sample based on the number of SBMA chromosomes, linkage disequilibrium observed, recombination fractions obtained from GDB and assumptions on ascertainment was made using the DMLE program (version 0.25).[14] We assumed complete isolation of the Finnish Swedish-speaking coastal population and a population growth equal to that generally estimated for Finland from mediaeval times (population base of about 100 000) from around year 1500 until the present (population base of about 5 million). As the algorithm carrying out the Monte Carlo replicates in DMLE only applies to marker loci with two allelic variants, only two loci were considered in the analysis, ALAS2 and DXS1125, the closest recombinant flanking loci to AR, both showing biallelic segregation in the Finnish SBMA families. For each log likelihood, a million Monte Carlo replicates were computed using a generation interval of 2. X-chromosomal inheritance was allowed for by simply halving the recombination fractions.

Results

All the Finnish SBMA families shared an identical haplotype 2–3–18–2–4 for the loci DXS1213–DXS1194–AR GGC–DXS1111–DXS135, spanning 4.8 cM calculated from GDB, and all except FV had the same DXS1213–DXS981 haplotype 2–3–18–2–4–6–7 +, spanning 5.3 cM (GDB) (Table 1). Nine out of ten Swedish SBMA families had the same DXS1213–DXS135 haplotype as most of the Finnish families. Ten out of
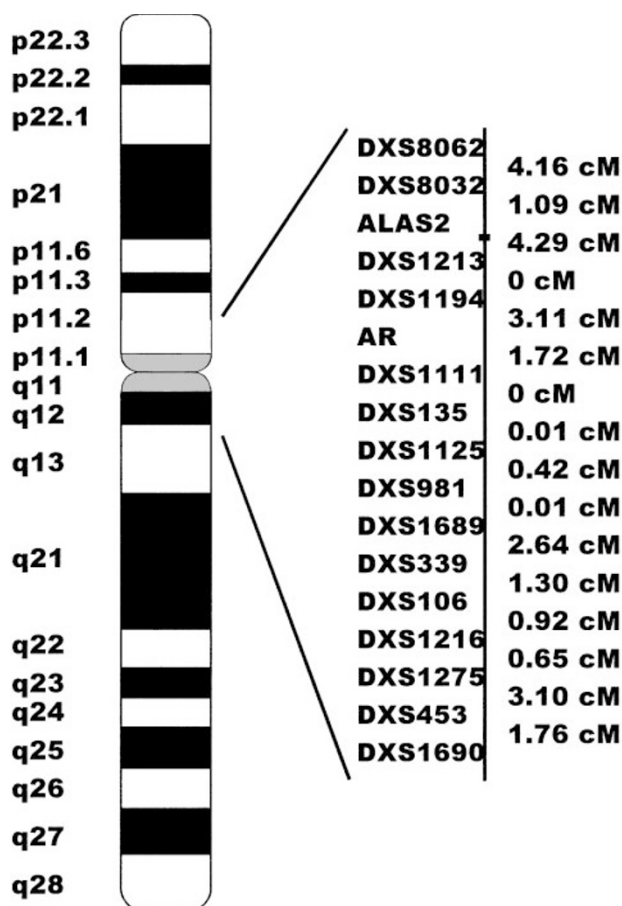
Table 1 The haplotypes of the Finnish, Swedish, Danish and Norwegian SBMA families and the base-pair sizes of the alleles. The common haplotypes shared by SBMA families in each country are shaded. The number behind the family symbol stands for the individual in the family. When there is no number after the symbol, all the examined family members shared the same haplotype. Different CAG numbers in families FIII, FV and NII are shown by multiple CAG numbers. *DXS981 is a tetranucleotide combined with deletions. + and – stand for alleles that are longer or shorter by one nucleotide, eg 7+ = 188 bp and 6–=189 bp

	8062	8032	ALAS2	1213	1194	AR CAG	AR GGC	1111	135	1125	*981	1689	339	106	1216	1275	453	1690
FII	4	3	5	2	3	45	18	2	4	6	7+	11	8	10	1	2	3	3
FXI1	4	3	5	2	3	41	18	2	4	6	7+	11	8	10	1	2	3	3
FXII1	4	3	5	2	3	42	18	2	4	6	7+	11	8	10	1	2	3	3
FVI1	4	2	2	2	3	42	18	2	4	6	7+	11	8	10	1	2	3	3
FVIII1	4	2	2	2	3	45	18	2	4	6	7+	11	8	10	1	2	3	3
FIX1	4	2	2	2	3	47	18	2	4	6	7+	11	8	10	1	2	3	3
FX1	4	3	5	2	3	47	18	2	4	6	7+	11	7	7	4	7	2	13
FXIII1	4	3	5	2	3	45	18	2	4	6	7+	11	7	7	4	7	2	13
FIV1	4	3	5	2	3	44	18	2	4	6	7+	11	7	7	4	7	2	13
FVII1	4	2	5	2	3	44	18	2	4	6	7+	11	7	7	4	7	2	13
FI2	4	3	5	2	3	43	18	2	4	6	7+	11	7	7	4	7	2	13
FI1+3	4	3	5	2	3	44	18	2	4	6	7+	2	3	7	4	1	2	1
FIII	4	3	5	2	3	45/46	18	2	4	6	7+	5	8	9	1	1	1	7
FV	4	2	2	2	3	44/45	18	2	4	5	3+	11	7	9	2	4	2	3
SII1	2	2	3	2	3	42	18	2	4	10	6–	11	7	8	2	6	1	13
SIII1	2	2	3	2	3	44	18	2	4	10	6–	11	7	8	2	6	1	13
SVI1	2	2	3	2	3	43	18	2	4	10	6–	11	7	8	2	6	1	13
SVII1	4	2	3	2	3	47	18	2	4	10	6–	11	7	8	2	6	1	13
SX1	2	2	3	2	3	46	18	2	4	10	6–	11	7	8	2	6	1	13
SVIII1	4	2	2	2	3	44	18	2	4	10	6–	11	8	7	2	4	1	2
SIX2	2	2	4	2	3	45	18	2	4	10	6–	11	7	8	2	6	1	13
SIX1	2	2	3	2	3	45	18	2	4	10	6–	2	3	7	3	5	1	13
SV1	2	2	3	2	3	41	18	2	4	3	4–	11	7	10	3	4	3	3
SI1	4	2	4	2	3	38	18	2	4	6	7+	11	8	10	1	2	3	3
SIV	4	2	2	2	3	45	18	1	4	5	4–	5	3	9	3	1	2	8
DIV1	4	7	5	3	3	44	18	2	6	5	7+	5	3	7	3	5	3	4
DVI1	4	7	5	3	3	44	18	2	6	5	7+	5	3	7	3	5	3	4
DVII1	4	7	5	3	3	49	18	2	6	5	7+	5	3	7	3	5	3	4
DVIII1	4	7	5	3	3	50	18	2	6	5	7+	5	3	7	3	5	3	4
DXI1	4	7	5	3	3	49	18	2	6	5	7+	5	3	7	3	5	3	4
DI1	4	3	7	3	3	48	18	2	6	5	7+	5	3	7	3	5	2	4
DX1	4	3	7	3	3	49	18	2	6	5	7+	5	3	7	3	5	2	4
DIX1	4	2	4	6	7	47	18	2	6	5	5+	5	1	8	1	1	2	12
DII1	4	7	5	3	3	51	18	2	6	6	9–	5	3	7	3	5	2	13
DIII1	4	7	5	3	3	41	18	2	6	4	7+	5	3	7	3	5	3	4
DV1	4	7	5	3	3	47	18	2	6	4	6–	5	7	8	2	7	2	12
DXII1	4	3	3	3	3	45	18	2	4	8	5	5	2	4	3	5	2	13
NI	4	7	6	2	3	44	18	2	4	6	7+	11	8	10	1	2	3	3
NII	4	2	2	2	3	43/44/45	18	5	4	5	2–	11	3	7	2	7	2	4

Marker Allele	8062	8032	ALAS2	1213	1194			1111	135	1125	981	1689	339	106	1216	1275	453	1690
1	98	198	167	244	267			129	238	138	199	161	214	107	246	212	170	144
2	94	192	165	240	283			127	236	136	198	159	212	105	244	210	168	142
3	90	200	163	236	269			125	234	134	195	157	210	103	248	216	166	140
4	92	190	161	242	275			123	232	132	194	155	208	101	250	214	164	138
5	96	196	159	238	263			121	230	130	191	153	206	99		206	160	136
6	88	188	157	234	261			119	228	128	190	151	204	97		220	172	134
7		194	155	232	277			131	226	126	187	149	202	95		218	150	132
8		202	153	230	265					124	186	147	200	93		208		130
9		204	149		271					122	183	145		91		222		128
10			169							144	182	143		89				126
11											202	167		87				124
12												169						146
13												171						148
14												165						150
15																		152

12 Danish families shared the same DXS1213–DXS135 haplotype. However they harboured allele 3 for the locus DXS1213 instead of 2, and allele 6 for the locus DXS135 instead of 4. In the Finnish, Swedish and Danish SBMA families a country-specific common DXS8062–DXS1690 founder haplotype was detected (Table 1, shaded areas). Moreover, all the Finnish, Swedish, Danish and Norwegian families shared the same 18 GGC repeats allele, and 92% of the families had the DXS1194–DXS1111 haplotype 3–18–2 (Table 1). One Swedish patient, SI1, and one Norwegian family, NI, had the DXS1213–DXS1690 haplotype 2–3–18–2–4–6–7 + –11–8–10–1–2–3–3 (15.6 cM), which is common among Finnish SBMA families (Table 1). All the controls showed different haplotypes (data not shown).

The 142 control chromosomes displayed a CAG repeat number ranging from 15 to 29 (Figure 2). The number of CAG repeats in the patients ranged from 38 to 51 (Figure 3). The range of CAG repeat numbers in Finnish patients was 41–47, with the mean of 44.3 ± 1.7. In Swedish patients the range of repeat numbers was 38–47, mean 43.6 ± 2.5. The Danish patients displayed 41–51 CAG repeats, mean 47 ± 3 (Figure 3). The two Norwegian families had 43–45 repeats (Table 1). The difference between the CAG ranges of the Finnish and the Swedish patients was not statistically significant, but the Danish patients differed significantly from the others ($P = 0.003$).

The intragenic GGC repeat sizes ranged from 11 to 21 in controls, with 17 being the most common allele, present in 48% of the Finnish, 55% of the Swedish and 70% of the Danish controls (Figure 4). All the Finnish, Swedish, Danish and Norwegian patients had 18 GGC repeats, also present in 26% of the Finnish, 27% of the Swedish and 15% of the Danish controls ($P = 0.0001$), suggesting linkage disequilibrium between the CAG repeat expansion and the GGC microsatellite.

Linkage disequilibrium was also seen between the mutation and the locus DXS1111, since allele 2 was present in 94.5% of the SBMA families and in only 40% of the Finnish, 18% of the Swedish and 25% of the Danish controls ($P = 0.0001$). All the Finnish, Swedish and Norwegian families shared the same allele 4 for the locus DXS135, which was present in 41% ($P = 0.004$) and 50% ($P = 0.002$) of the respective controls. However, allele 4 was seen in only one Danish family, the remaining 11 (92%) carrying allele 6 present in 20% of the Danish controls ($P = 0.0003$). All the Finnish, Swedish and Norwegian families shared allele 2 in DXS1213, present in 62% ($P = 0.05$) and 50% ($P = 0.03$) of the respective controls, and 11 (92%) of the Danish families had allele 3, present in 10% of the Danish controls ($P = 0.08$). Allele 3 of the locus DXS1194 was present in all except one (97%) patients, and also in 66% of the controls ($P = 0.0002$).

The maximum likelihood peak in the age estimation using the DMLE program (see Materials and methods for assumptions) for the closest recombinant flanking loci, ALAS2 and
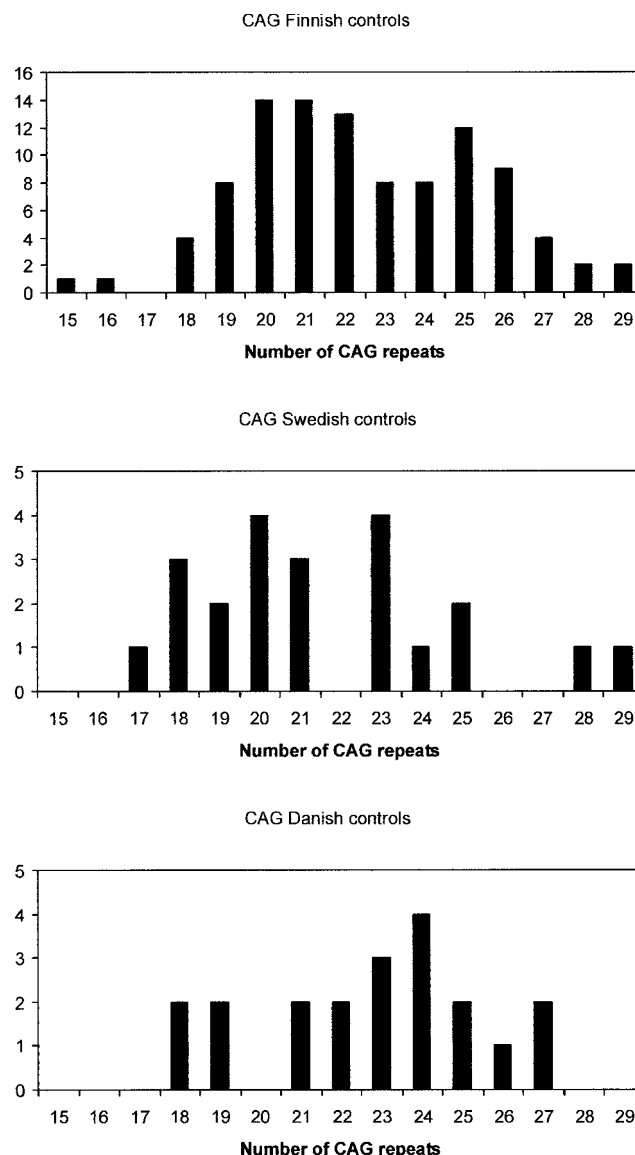

Figure 2 Numbers of CAG repeats in the controls.

DXS1125, in the Finnish SBMA families was obtained at 20 generations for each (data not shown).

Discussion

All the Finnish and 9 of the 10 Swedish SBMA chromosomes harboured an identical haplotype spanning a 4.8 cM region (GDB) (Table 1) on chromosome Xq11–q12. This suggests a common founder haplotype for these patients. The different alleles of the loci DXS1213 and DXS135 in the Danish SBMA patients compared to those carried by Finnish and Swedish patients indicate a more ancient separation. The DXS1194–DXS1111 haplotype shared by 92% of the families (Table 1) composes a potential Scandinavian founder haplotype. There was also evidence of migration between countries, since SI,
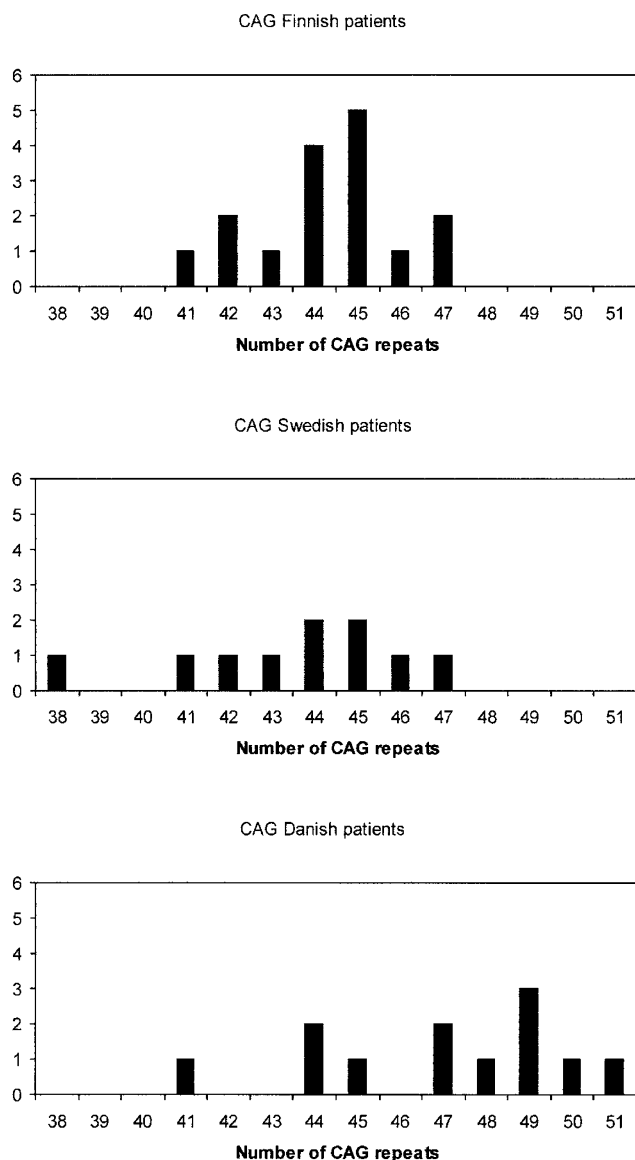
CAG Finnish patients

CAG Swedish patients

CAG Danish patients

Figure 3 Numbers of CAG repeats in the patients.

NI and six Finnish SBMA families shared the same haplotype for DXS1213–DXS1690.

The ranges of the CAG repeat sizes of the Finnish, Swedish and Norwegian patients (Figure 3, Table 1) were quite similar, 44 and 45 repeats being the most common. The Danish patients had a somewhat wider range with slightly longer repeats, 49 repeats being the most common. If the Danish mutation had a more ancient origin, it would have had more time to develop a wider range. The comparable ranges shown here in the Finnish, Swedish and Norwegian patients suggest a more recent separation between these.

In Japanese SBMA patients linkage disequilibrium was found between the CAG repeat expansion and the GGC repeat, indicating a founder effect.[13] Nevertheless, in the Japanese patients the linkage disequilibrium was not as

complete as in our study, where all Nordic patients shared the same GGC repeat number. A founder effect was also suggested to be responsible for the clustering of SCA6 families in the Nordrhein–Westfalia area in Germany, where the majority of the families shared allelic characteristics of three markers around the associated gene.[15] A founder effect has also been reported in SCA1,[16] SCA2[17,18] and SCA3 (Machado-Joseph disease).[19]

Twelve of the 13 Finnish SBMA families originated from the Vaasa region on the western coast of Finland. One patient (FVIII1) lived in Southern Finland. The fairly restricted area from which the SBMA patients originated and the fact that 11 of these 13 families were Swedish-speaking supports the theory of a common ancestor. Nevertheless, genealogical
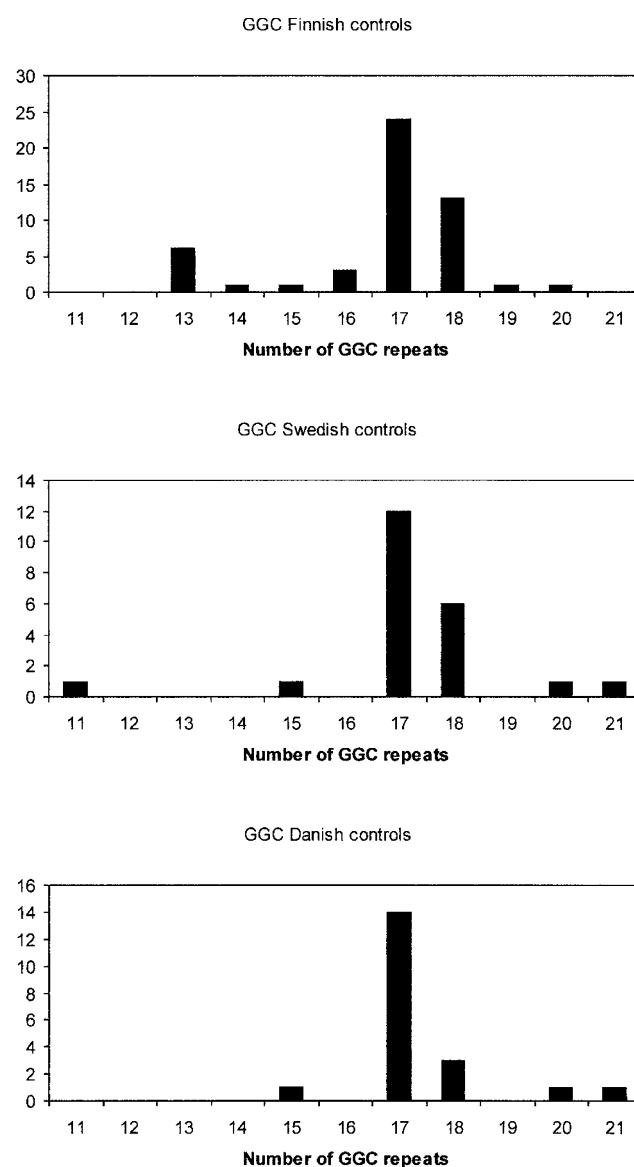
GGC Finnish controls

GGC Swedish controls

GGC Danish controls

Figure 4 Numbers of GGC repeats in the controls.

studies have not revealed any definite genetic links between these families over the past 200 years.

The Vaasa region in western Finland was inhabited during the Stone, Bronze and Iron Ages. Some later resettlement from Sweden is documented from mediaeval times, along with the expansion of Christianity, and later until the 18th century. Since many of the Swedish patients come from the areas located across the Gulf of Bothnia from the Vaasa region, and because of the common language, a common genetic background for the Finnish and the Swedish SBMA patients is indicated.

The maximum likelihood peak for ALAS2 and DXS1125 of the Finnish SBMA patients obtained by DMLE, 20 generations, would indicate the introduction of the SBMA mutation into the Finnish Swedish-speaking coastal population around 1400–1500. However, whilst the estimate seems close to the truth, it has very wide confidence limits and, on the other hand, is based for simplicity on assumptions that almost certainly are not the whole truth.

The fact that all the Scandinavian patients with SBMA carry the same intragenic GGC allele 18 in addition to other similarities in allele distribution suggests an ancient Scandinavian founder effect. There also seems to be a specific more extended sub-founder haplotype in each country. Later on, we plan to elucidate the possible similarities between patients from around the world, and to find out whether multiple founder effects exist.

References
1 Kennedy WR, Alter M, Sung J: Progressive proximal spinal and bulbar muscular atrophy of late onset: A sex-linked recessive trait. *Neurology* 1968; **18**: 671–680.

2 Udd B, Juvonen V, Hakamies L *et al*: High prevalence of Kennedy's disease in Western Finland – is the syndrome underdiagnosed? *Acta Neurol Scand* 1998; **98**: 128–133.

3 La Spada AR, Wilson EM, Lubahn DB, Harding AE, Fischbeck KH: Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* 1991; **352**: 77–79.

4 Merry DE, Kobayashi Y, Bailey CK, Taye AA, Fischbeck KH: Cleavage, aggregation and toxicity of the expanded androgen receptor in spinal and bulbar muscular atrophy. *Hum Mol Genet* 1998; **7**: 693–701.

5 Burright EN, Orr HT, Clark HB: Mouse models of human CAG repeat disorders. *Brain Pathol* 1997; **7**: 965–977.

6 Butler R, Leigh PN, McPhaul MJ, Gallo J-M: Truncated forms of the androgen receptor are associated with polyglutamine expansion in X-linked spinal and bulbar muscular atrophy. *Hum Mol Genet* 1998; **7**: 121–127.

7 Pearson CE, Sinden RR: Trinucleotide repeat DNA structures: dynamic mutations from dynamic DNA. *Curr Opin Struct Biol* 1998; **8**: 321–330.

8 Koob MD, Moseley ML, Schut LJ *et al*: An untranslated CTG expansion causes a novel form of spinocerebellar ataxia (SCA8). *Nat Genet* 1999; **21**: 379–384.

9 Holmes SE, O'Hearn EE, McInnis MG *et al*: Expansion of a novel CAG trinucleotide repeat in the 5' region of *PPP2R2B* is associated with SCA12. *Nat Genet* 1999; **23**: 391–392.

10 Igarashi S, Koide R, Shimohata T *et al*: Suppression of aggregate formation and apoptosis by transglutaminase inhibitors in cells expressing truncated DRPLA protein with an expanded polyglutamine stretch. *Nat Genet* 1998; **18**: 111–117.

11 Abdullah AAR, Trifiro MA, Panet-Raymond V *et al*: Spinobulbar muscular atrophy: polyglutamine-expanded androgen receptor is proteolytically resistant *in vitro* and processed abnormally in transfected cells. *Hum Mol Genet* 1998; **7**: 379–384.

12 Sleddens HFBM, Oostra BA, Brinkmann AO, Trapman J: Trinucleotide (GGN) repeat polymorphism in the human androgen receptor (AR) gene. *Hum Mol Genet* 1993; **2**: 493.

13 Tanaka F, Doyu M, Ito Y *et al*: Founder effect in spinal and bulbar muscular atrophy (SBMA). *Hum Mol Genet* 1996; **5**: 1253–1257.

14 Rannala B, Slatkin M: Likelihood analysis of disequilibrium mapping, and related problems. *Am J Hum Genet* 1998; **62**: 459–473.

15 Dichgans M, Schöls L, Herzog J *et al*: Spinocerebellar ataxia type 6: Evidence for a strong founder effect among German families. *Neurology* 1999; **52**: 849–851.

16 Wakisaka A, Sasaki H, Takada A *et al*: Spinocerebellar ataxia 1 (SCA1) in the Japanese in Hokkaido may derive from a single common ancestry. *J Med Genet* 1995; **32**: 590–592.

17 Hernandez A, Magarino C, Gispert S *et al*: Genetic mapping of the spinocerebellar ataxia 2 (SCA2) locus on chromosome 12q23–q24.1. *Genomics* 1995; **25**: 433–435.

18 Mizushima K, Watanabe M, Kondo I *et al*: Analysis of spinocerebellar ataxia type 2 gene and haplotype analysis: (CCG)1-2 polymorphism and contribution to founder effect. *J Med Genet* 1999; **36**: 112–114.

19 Endo K, Sasaki H, Wakisaka A *et al*: Strong linkage disequilibrium and haplotype analysis in Japanese pedigrees with Machado-Joseph disease. *Am J Med Genet* 1996; **67**: 437–444.