European Journal of Human Genetics (1999) 7, 12–19
© 1999 Stockton Press All rights reserved 1018–4813/99 $12.00
http://www.stockton-press.co.uk/ejhg

ARTICLES

An integrated map of chromosome 18 CAG trinucleotide repeat loci

Andy J Grierson[1,5], Marjon van Groenigen[1], Nancy PB Groot[2], Kerstin Lindblad[3], Jan MN Hoovers[2], Martin Schalling[3], Jackie de Belleroche[4] and Frank Baas[1]

[1]Department of Neurology, University of Amsterdam, Academic Medical Center, Amsterdam
[2]Department of Human Genetics, University of Amsterdam, Academic Medical Center, Amsterdam, The Netherlands
[3]Neurogenetics Unit, Center for Molecular Medicine, Karolinska Hospital, Stockholm, Sweden
[4]Molecular Neurobiology Research Group, Department of Biochemistry, Charing Cross and Westminster Medical School, London, UK
[5]Department of Neuroscience, Institute of Psychiatry, De Crespigny Park, Denmark Hill, London SE5 8AF, UK (present address)

Expansions of trinucleotide CAG repeats have been demonstrated in at least eight neurodegenerative disorders, and suggested to occur in several others, including bipolar disorder and schizophrenia. Chromosome 18 loci have been implicated in bipolar disorder pedigrees by linkage analysis. To address this putative link between chromosome 18 CAG trinucleotide repeats and neuropsychiatric illness, we have screened a chromosome 18 cosmid library (LL18NCO2□AD□) and identified 14 novel candidate loci. Characterisation of these loci involved repeat flank sequencing, estimation of polymorphism frequency and mapping using FISH as well as radiation hybrid panels. These mapped trinucleotide loci will be useful in the investigation of chromosome 18 in neurodegenerative or psychiatric conditions, and will serve to integrate physical and radiation hybrid maps of chromosome 18.

Keywords: chromosome 18; trinucleotide repeats; FISH; integrated mapping; neuropsychiatric disorders

Introduction

A growing number of human diseases are caused by expansions of trinucleotide repeats in their genes. The CAG, CTG, AAG and CCG repeats have so far been implicated in this way. CAG repeats represent the most numerous class of expansions, at present known to account for eight distinct diseases: Kennedy's disease, Huntington's disease, dentatorubralpallidoluysian atrophy and spinocerebellar ataxia types 1, 2, 3, 6 and 7 (reviewed by Ross[1]). Expansion of CAG repeats may also be involved in other spinocerebellar ataxias, including spinocerebellar ataxia type 5.[2] These diseases are each characterised by specific neuronal cell death, but share genetic and phenotypic similarities, including progressive neurodegeneration, genetic anticipation, and the translation of the expanded trinucleotide into a polyglutamine stretch in the encoded protein.

A number of other diseases have been suggested to be caused by expansion of as yet unidentified trinucleotide loci, including spinocerebellar ataxias and neuropsychiatric disorders. Evidence for the involvement of expanded CAG/CTG repeats in psychiatric disease came from the repeat expansion detection (RED) technique,[3] which demonstrated an increased frequency

Correspondence: Frank Baas, University of Amsterdam, Academic Medical Center, PO Box 22700, 1100DE, Amsterdam, The Netherlands. Tel: 31 20 5665998; Fax: 31 20 5664440; E-mail: f.baas@amc.uva.nl
Received 5 February 1998; revised 15 May 1998; accepted 11 June 1998

of longer CAG/CTG trinucleotides in bipolar disorder and schizophrenia patient DNA compared with control populations.[4–6] Consistent with this, previous studies suggested this anticipation in psychotic illness.[7–10]

Linkage analysis in the field of psychiatric genetics has a chequered history,[11] accounted for by factors such as divergent phenotypes and the existence of multiple disease genes. Recently a number of reports have suggested linkage of bipolar disorder to markers on chromosome 18.[12–14] It could be hypothesised that an expanded CAG trinucleotide from this chromosome is involved in the pathogenesis of the disease. Previously Schalling *et al*[3] demonstrated an expanded CAG locus (RED1) on chromosome 18 in CEPH pedigree 1334. Another expanded chromosome 18 allele recently cloned by Breschel,[15] was first identified using fluorescence *in situ* hybridisation (FISH) on metaphases obtained from a schizophrenic individual.[16] This repeat was subsequently excluded from expansion in CEPH pedigree 1334, suggesting the presence of at least two non-allelic CAG repeats which can show expansions on chromosome 18.

There are essentially two distinct strategies available to identify an expanded CAG repeat locus. Random genome, cDNA or chromosome-wide searches based on screening cosmid or cDNA libraries have been described.[17–19] In addition novel cloning approaches which are locus specific have been used, including use of the antibody 1C2 which recognises expanded polyglutamine tracts in proteins,[20–22] as well as the Direct principle[23] which utilises selective hybridisation or the RAPID screening procedure which uses stringent selection procedures prior to cloning steps.[24]

We set out to identify CAG repeats on chromosome 18, so as to facilitate a systematic screen for possible expansions in bipolar disorder, as has been previously conducted for schizophrenia.[25] We used a PCR-based cloning strategy to identify CAG repeat loci in a chromosome 18-specific cosmid library derived from a flow-sorted monochromosomal somatic cell hybrid. This method avoids time-consuming subcloning procedures, and since only the repeat flanks are cloned, any cloning bias against longer repeat sequences will be avoided.[26,27] The new loci were used for a small-scale screen of individuals who showed expanded RED products and were affected with bipolar disorder or other neuropsychiatric conditions.

In addition to identifying novel CAG loci on chromosome 18, we aimed to characterise their polymorphism frequency, and to define precisely their map locations on both genetic and physical maps, which will provide a link between these two types of map. The CAG trinucleotide-containing repeats reported here may help in the mapping of genes for hereditary psychiatric or degenerative disorders, including bipolar disorder and schizophrenia, and will contribute to the development of human STS-based maps.

Materials and Methods

Hybridisation Conditions

The chromosome cosmid library LL18NCO2[illegible]AD[illegible] was replicated in high-density arrays on to Hybond N filters using a Biomek 1000 robot (Beckman Instruments Inc., Fullerton, CA, USA). The filters were hybridised to a ^{32}P labelled $(CAG)_{10}$ oligonucleotide as previously described.[3] Filters were stringency washed with either $1 \times$ SSC for normal stringency, or $0.1 \times$ SSC for high stringency, both at 65°C for 15 min. To determine coverage of the library, the forward PCR primer of the GCT5D07 STS was ^{32}P labelled as above, and hybridised to the filters.

Vectorette PCR

Positive clones were isolated and cosmid DNA was prepared using Qiagen columns according to the manufacturer's instructions (Qiagen, GmBH, Hilden, Germany). Cosmid DNA (0.5 μg) was digested with one or more restriction enzymes; AluI, HaeIII, EcoRV, RsaI. Vectorette PCR was performed essentially as previously described,[28] except that repeat complementary primers $(CAG)_6$ or $(CTG)_6$ were used at an annealing temperature of 63°C for amplification of repeat-flanking DNA. The repeat flanks were cloned into the pGEM5 vector (Promega Corp., Madison, WI, USA) by virtue of 5' terminal NotI and NcoI restriction sites on the PCR primers. Plasmid DNA was prepared using the Wizard Plus kit (Promega Corp., Madison, WI, USA), and sequenced using the FS+ dye terminator kit (Applied Biosystems Inc., Perkin Elmer Corporation, Norwalk, Conn., USA).

Development of CAG Repeat Containing STSs

Primers were chosen from each of the two amplified flanks of a single cosmid using the GCG Wisconsin package (Genetics Computer Group, Madison, WI, USA). In the event that only a single flank was amplified from a cosmid clone, a single sequencing primer was designed, and direct sequencing of 1 μg of PEG-precipitated cosmid DNA was performed. PCR amplification was carried out on 25 ng of genomic DNA in a total volume of 10 μl consisting of 20 mM dNTPs, 10 ng of each primer, 0.01% NP40, 0.05% gelatin, 50 mM KCl, 10 mM Tris (pH 8.3–9.9), 1.5–3.5 mM $MgCl_2$ and 2 units Taq polymerase for the following cycling times: 94°C for 3 min followed by 30 cycles of 94°C for 30 s, 53–58°C (see Table 1) for 30 s, and 72°C for 1 min followed by 72°C for 5 min. $MgCl_2$ concentration (1.5–3.5 mM) and Tris pH (8.3–9.9) were determined empirically for each primer pair. Polymorphism frequency was determined by amplification using the above conditions in the presence of 1 ng ^{33}P dATP end-labelled primer. Products were separated on 6% denaturing polyacrylamide gels. In some instances unlabelled PCR products were visualised by hybridisation of electro-blotted denaturing

Table 1 CAG-repeat loci identified in this investigation

STS	*Cosmid ID*[1]	*PCR primers*	*Repeat structure*	*Tm °C*	*Size (bp)*	*No. of alleles*	*Heterozygote frequency*
LL18CAG10	AD11H12 AD23E8 AD112A6 AD63G8	GGCAAAACACCCATCTTCTCTT AGCTGCTATGTATTCCAGGCAT	$(CAG)_8$	55	149	2	25%
LL18CAG12	AD12F4 AD31B4 AD86A7	GCATTCTTTATTACCCGGCAC AAAACAAGGACACGCAGAGG	$(CAG)_7$CAA	55	211	2	67%
LL18CAG17	AD16G4 AD82H2 AD57G8	GCATTTGTTCTCTGGCTCCTTG GGCACCCACTAGCACCCT	$(CAG)_4$TAG$(CAG)_5$CAT	55	225	2	15%
LL18CAG19	AD18H6	GCTAATAGGATTAGGATGCTGCTGA GCACCCTGAAGGTTTCCGT	$(CAG)_6$	56	125	1	–
LL18CAG20	AD19C2 AD75B9	CTTCGCCCCACAAATCCT CGCCGACATGCCTCAC	CAGCATCAC$(CAG)_6$	55	214	1	–
LL18CAG25	AD29H4 AD31G7	TGTGAATGGGGAAGGAAGAAG GGACTGCCAAGAGAATGGAA	$(CAG)_8$CAACAC	58	203	3	65%
LL18CAG30	AD31A12 AD63D8 AD66G3 AD823G3 AD89A7	TTTGGGGAGCACTGAGGG TTTCCACAGAGAAAGGAAGGG	$(CAG)_8$CAT	55	135	4	60%
LL18CAG31	AD41F12 AD129G2 AD63D8	ACCAATTTCAGGCCCAA CAAAGGGTTTCCTGCTTTGA	$(CAG)_2$CAA$(CAG)_8$CAA $(CAG)_2$CAA$(CAG)_7$	55	196	3	26%
LL18CAG39	AD65A12 AD116A5	CCTTTTCATTCATGCCAAACT GGCTTCTGGCTCTGTGTTTCT	$(CAG)_7$	55	220	1	–
LL18CAG48	AD80E3	GCCACTTGCCCTTCTCCT GTTCCCTTCCTTCATTTTCGTT	$(CAG)_7$	53	320	1	–
LL18CAG50	AD38D8 AD81H6	CCTTTTCTGAAGGCCCCTG CCTTACCTGGAGGTGAAATGG	$(CAG)_5$	58	155	1	–
LL18CAG61	AD102H10 AD109D9	ATCTTATTTTCTCCAATTCCCCAG TCCCTATGATTTCTTTTTGATTGC	$(CAG)_6$	52	121	1	–
LL18CAG73	AD132B2	CGGCCCCTTCAGGCAG GAGAACTAAAACCCAGCAGCTAAGA	$(CAG)_6$	55	175	1	–
LL18CAG218	AD95F3	CCCCGCAACAAATCTTTGA GGGAGATGGGGATAATGGTAATAG	TAG$(CAG)_5$CAT	54	172	1	–

[1]Cosmids were derived from the library LL18NCO2"AD" (Lawrence Livermore National Laboratory). Multiple IDs represent independent cosmids containing the same repeat flank sequences.

polyacrylamide gels with a $(CAG)_{10}$ ^{32}P-labelled probe as described above. DNA from 30 unrelated individuals was analysed to determine the polymorphism frequency of the CAG repeat.

Radiation Hybrid Mapping

The Stanford G3 (Research Genetics Inc., Huntsville, AL, USA) and Genebridge 4[29] radiation hybrid panels were used to place the CAG-containing STSs on the framework radiation hybrid maps. The Stanford Human Genome Center and Whitehead Institute Radiation Hybrid WWW servers were used for this analysis.

Physical Mapping

STSs were physically mapped and localised to chromosomal sub-bands by FISH, using the appropriate cosmid DNA. Localisation was performed by measuring the relative position of the probe on the chromosome. Map positions were expressed as the fractional length of the whole chromosome relative to the tip of the p arm (FLpter).[30] For each cosmid, FLpter values were calculated for 10–15 prometaphase chromosomes and a 95% confidence interval was calculated. When compared with ideograms in a similar state of contraction, this method gives a very accurate chromosomal assignment.[31]

Screening for Expansions

At least three unrelated bipolar affective disorder patients and between nine and 15 individuals with other neuropsychiatric disorders who all showed expanded RED products were screened by PCR using the primers and conditions shown in Table 1. The products were separated on denaturing 6% polyacrylamide gels, transferred to nylon membranes and hybridised to a $(CAG)_{10}$ probe as described above. Membranes were overexposed to permit the detection of poorly amplified expanded alleles.

Database Analysis of STSs

Blast similarity searches were performed at the NCBI using the BlastN server and NR, STS and EST databases.

Results

Library Screening and STS Development

The strategy for the identification and analysis of CAG repeat loci is outlined in Figure 1. To determine the coverage of the LL18NCO2AD cosmid library we hybridised the filters to the forward primer of the GCT5D07 STS, a CAG trinucleotide repeat locus, which has been mapped to chromosome 18. This identified 7 cosmids containing the STS, predicting a seven-fold coverage of the chromosome in the cosmid library. The filters were then screened with a $(CAG)_{10}$ oligonucleotide probe, washed to normal stringency ($1 \times$ SSC) and 200 positive clones were identified. Of these, 73 cosmids were randomly chosen for further analysis. A second high-stringency screen using $0.1 \times$ SSC yielded 66 positive signals, and the 20 strongest clones were chosen. Of these 20 clones 5 had already been identified in the normal stringency screen, and 11 turned out to be duplicates of clones already selected in the normal stringency screen. Initially triple digests with EcoRV, HaeIII and RsaI were performed prior to bubble-anchor ligation. Subsequently multiple single digests were used, to enable a suitably sized flank to be selected for cloning. In total, from the 93 cosmids, data were generated from 66 independent clones. Multiple clones representing 14 novel CAG repeat loci were identified using this approach. We also identified multiple cosmids representing four previously identified chromosome 18 repeat loci (GCT3E06, GCT3G01, GCT5D07, GCT6G01), but we did not isolate clones corresponding to four other CAG loci which had previously been mapped to chromosome 18 (GCT3A09, GCT7G01, GCT13D05, CTG18.1).[26,30] However we subsequently reassigned the GCT3A09 locus to chromosome 15 (see below). Fourteen additional clones were isolated which corresponded to hamster DNA sequences (data not shown). This was expected as the library contains 10% hamster clones. Seven uninformative clones gave no flank-PCR products with the $(CAG/CTG)_{10}$ primers, suggesting a false-positive hybridisation result. Two repeat flanks were derived from the inappropriate amplification of *E. coli* sequences. These were present in the PCR reaction as a result of contamination of the cosmid preparation with genomic DNA. These misamplifications were readily identified by database analysis. If restriction sites had been too close (less than 20 bp) to the repeat we may have had difficulty in detecting a 40–60 bp PCR product. This problem was usually overcome by choosing another restriction enzyme prior to bubble-anchor ligation. As expected, multiple clones were identified for several repeats, and this data is summarised in Table 1 along with PCR primers, conditions for PCR, repeat structure and polymorphism frequencies. Briefly, of 14 novel clones, 6 are polymorphic based on the amplification of 30 independent DNA samples (43%).

Screening for Expansions in a Set of Individuals with Neuropsychiatric Disorders and Long RED Products

To investigate if any of the novel CAG repeat loci described here are pathologically expanded, a small-scale screen was undertaken. A set of individuals with bipolar affective disorder or other neuropsychiatric disorders where unidentified expansions had previously been detected using RED were used for this screen. A

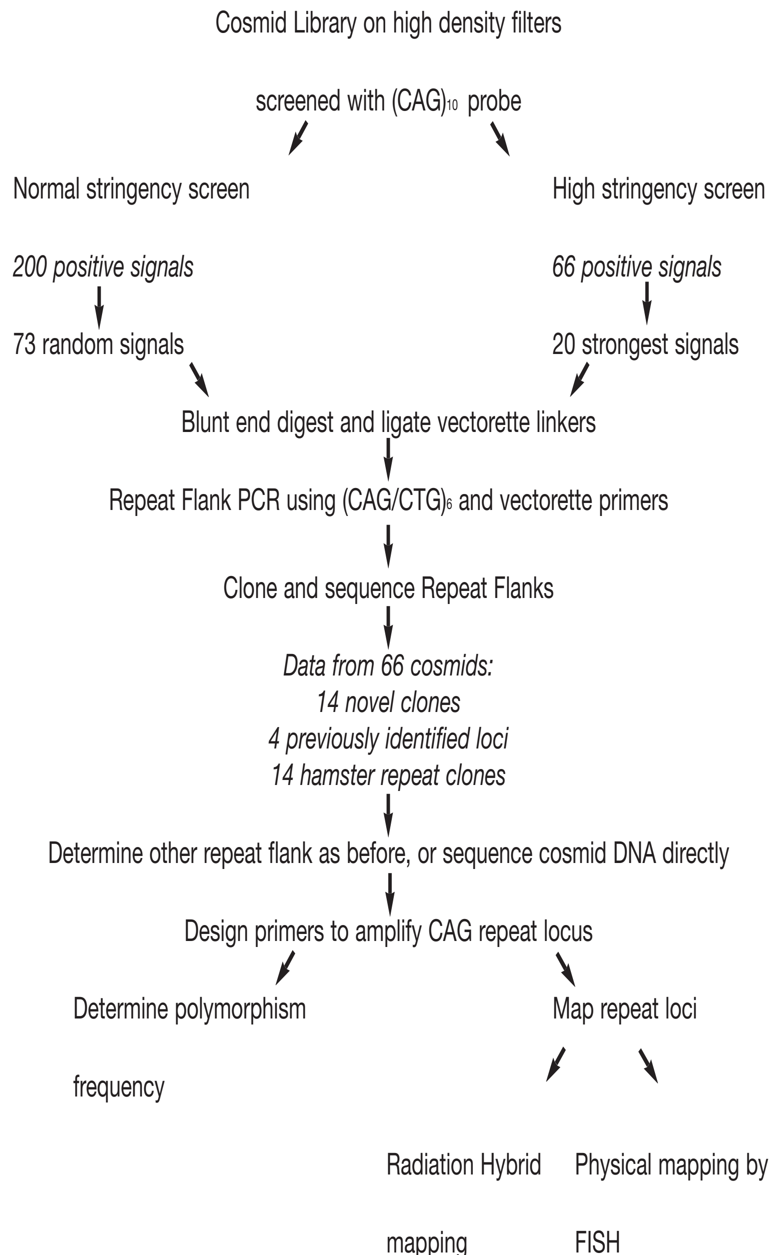

Figure 1 *Schematic outline of the strategy used to identify and map CAG repeat containing cosmids.*

PCR strategy optimised for detection of poorly amplified expanded alleles was used. All 14 loci identified in this study were screened in 15 unrelated individuals with either bipolar affective disorder or other neuropsychiatric conditions. The bipolar cases were familial, but not specifically linked to chromosome 18. No expansions at the novel CAG repeat loci were detected in this small set of neuropsychiatric disorder individuals with this screen.

Radiation Hybrid and Physical Mapping of the STSs

All clones were mapped on the Stanford G3 or Genebridge 4 radiation hybrid panels. In order to link the radiation hybrid and physical maps, we also performed FISH with whole cosmids. The results of mapping the STSs are presented in Table 2. The results of both approaches are in good accord, and thus these STSs serve as reference markers linking the two maps.

Table 2 Database homology and map location of chromosome 18 CAG repeat loci

STS	*Database homology*	*Physical position (FLpter)*	*Sub-band localisation*	*Linked marker on radiation hybrid map*[1]	*Distance from linked marker (centiRays)*
LL18CAG10	T06032 Foetal Brain EST	pter	18p11.32	D18S1132	6.43
LL18CAG12	U55976, H09960, R25492, R59809, H15991, R19422, R42000 ESTs	0.18–0.23	18p11.1–p11.21	D18S1228 D18S852	11.69 14.80
LL18CAG17	–	0.15–0.18	18p11.21	D18S1226	63.70
LL18CAG19	–	0.57–0.72	18q12.3–q21.2	D18S1311	33.52
LL18CAG20	Z78337 Foetal Brain EST	0.65–0.74	18q21.2–q21.31	D18S1311	31.40
LL18CAG25	–	0.55–0.66	18q12.3–q21.2	D18S1212	14.84
LL18CAG30	–	0.91–0.95	18q22.3	D18S823	17.35
LL18CAG31	–	failed	–	D18S1282	4.11
LL18CAG39	–	0.16–0.19	18p11.21	WI-5607	6.83†
LL18CAG48	–	0.35–0.52	18q11.2–q12.1	D18S869	15.55†
LL18CAG50	–	0.9–1.0	18q22.3–q23	D18S1161	41.36
LL18CAG61	L34155 LAMA3	failed	18q11.2*	D18S869 D18S1067	27.10 41.10
LL18CAG73	–	pter	18p11.32	D18S1231	3.1
LL18CAG218	–	failed	–	D18S60	25.97
GCT3E06	–	–	–	D18S1161	21.20
GCT3G01	–	–	–	D18S828	0††
GCT6F12		–	–	D18S73	0††
GCT6G01	–	–	–	D18S58	0††
GCT7G01	–	–	–	D18S542	24.09
GCT5D07	U55976	–	–	D18S71	31.58
GCT13D05	–	–	–	D18S869	35.89
GCT3A09	–	–	–	D15S148	33.45

[1]Markers were typed on the StanfordG3 Radiation Hybrid panel unless indicated otherwise.
*FISH failed, sequence identity to LAMA3 which is already mapped to 18q11.2.
†These markers were typed on the Genebridge4 hybrid panel.
††These markers are present on the Genebridge4 hybrid map.

Three of the novel loci, LL18CAG31, LL18CAG61 and LL18CAG218, failed to give interpretable FISH data. Clone LL18CAG61 showed sequence identity to the *LAMA3* gene, which allows us to map it to 18p11.2. We also used radiation hybrid panels to map the previously identified markers GCT3E06, GCT3G01, GCT5D07, GCT6G01, GCT3A09, GCT7G01 and GCT13D05. It is noteworthy that we found GCT3A09 mapped to chromosome 15 using the Genebridge 4 radiation hybrid panel (Table 2). This STS had previously been mapped to chromosome 18 using the NIGMS panel.[18]

Database Homology Analysis

The rapidly expanded database service at the NCBI enables the rapid cross-referencing of new and existing sequence data from diverse sequencing projects. The EST and Non-Redundant (NR) databases enable us to predict which of the CAG-repeat sequences are expressed. Expression of the CAG repeat is particularly relevant to the involvement of these repeats in human genetic diseases, and this data is presented in Table 2. Of the novel genomic clones presented here, four had perfect homologies with expressed Genbank sequences. In one instance the repeat was within the homologous region, LL18CAG20, which was also identified by Neri *et al.*[19] In the remaining three cases the flanking sequence rather than the CAG repeat was homologous to an EST sequence (LL18CAG10, 12 and 61), suggesting that the repeat sequence is located in the proximity of an EST, but is not necessarily exonic.

Discussion

We have identified 18 chromosome 18 CAG-repeat containing STSs of which 14 are novel. These clones,

together with the previously reported chromosome 18 loci may be used for screening diseases associated with the expansion of CAG repeats on this chromosome. The vectorette PCR techniques used here[28] demonstrate the efficacy of 'cloning' CAG repeats without actually cloning the potentially unstable CAG repeat itself. This may be useful to investigators interested in trying to identify novel expanded repeats in human genomic DNA.

Our extensive analysis of the repeat loci identified here enables comparisons to be drawn with other studies which have identified CAG repeats. Gastier *et al*[18] reported the largest set of CAG-based STSs to date, including 8 loci on chromosome 18. We have shown that one of these markers is actually located on chromosome 15, and confirmed the data for four of the others. Taken together with another locus identified by Breschel *et al*,[15] we now know the location of 22 different CAG-repeat loci on this chromosome. On the basis of hybridisation with a single copy probe, we originally estimated there were 22–27 CAG loci represented in the library based on 200 positives in a library comprising 14 000 clones (of which 1400 are likely to be hamster clones) with 7-fold coverage. In fact we isolated 14 hamster CAG repeat loci (data not shown) which is in keeping with these calculations. Since we failed to identify clones representing four other loci in our analysis of 93/200 clones it is possible that portions of the chromosome are not equally or adequately represented in the cosmid library. Of the clones for which we generated human CAG repeat sequence data, we usually identified at least two cosmids representing each repeat locus, with a mean of 1.6 cosmids, and a range of 1 to 5 cosmids per locus. This data is based on the identification of multiple copies within the 93 clones chosen from 200 hybridisation signals, and is not based upon rescreening the library with each novel clone. Therefore, we propose that it is unlikely that we have failed to identify a significant number of CAG loci present in this library, and suggest that any chromosome 18 loci which we failed to identify are likely to be absent from, or under-represented in, this library.

Our polymorphism analysis demonstrated that 6/14 (43%) of the novel repeats were polymorphic. Inclusion of polymorphism data for the other chromosome 18 loci,[32,33] and our own analysis of six of these loci (data not shown), revealed that 5/8 (63%) loci were polymorphic. Thus in total, of 22 loci on chromosome 18, 11 were polymorphic (50%). It has been suggested that loci containing greater than nine repeat units are the most likely to be polymorphic, and this likelihood is increased further for perfect repeats containing only CAG motifs, compared with complex repeats which are interspersed with CAA or CAC or AAG trinucleotides.[19] However, the SCA6 repeat in the 3' end of the $\alpha 1_A$-voltage dependent calcium channel may contain as few as four CAG repeats on normal chromosomes,[34] so it seems to be the case that shorter and less polymorphic repeats may be candidates for disease-associated expansion. It is interesting to note that our longest CAG clone, LL18CAG31, containing 22 imperfect repeats, also has a low polymorphism frequency, with only three alleles, one of which has an allele frequency of 0.94. The sequence of the repeat at this locus is $(CAG)_2CAA(CAG)_8CAA(CAG)_2CAA(CAG)_7$.

The identification, cloning, and integrated mapping of 14 novel and eight previously identified CAG repeat-containing loci on human chromosome 18 will be of use to investigations of neuropsychiatric disorders. For example, bipolar disorder demonstrates anticipation and has been associated with CAG/CTG repeat expansions. It is also possible that a locus on chromosome 18 contributes to this disorder. We therefore analysed a number of unrelated bipolar samples where repeat expansions had previously been identified using the RED technique. These were familial cases, but not specifically families linked to chromosome 18. Although no expansions were found in this preliminary screen, firm conclusions regarding the involvement of any of these new CAG repeat loci would have to await a more comprehensive analysis of chromosome 18 linked families. A small sample of unmapped neuropsychiatric disorders were also tested because of the presence of verified RED expansions. Although no expansions were detected, it would be premature to rule out the involvement of these chromosome 18 loci in neuropsychiatric disorders based only on this limited patient sample.

Acknowledgements

We thank the YAC Screening Service at Leiden for use of their Biomek robot when replicating the cosmid library. We acknowledge the HGMP for providing the Genebridge 4 radiation hybrid panel. AJG, MG and KL were funded by an EU Biomed award. We thank ALMA ten Asbroek, Drs W Scheper, LJ Valentijn, R Versteeg and A Westerveld for critical comments on this manuscript.

References

1 Ross CA: Intranuclear neuronal inclusions: A common pathogenic mechanism for glutamine-repeat neurodegenerative diseases? *Neuron* 1998; **19**: 1147–1150.
2 Ranum LP, Schut LJ, Lundgren JK, Orr HT, Livingston DM: Spinocerebellar ataxia type 5 in a family descended from the grandparents of President Lincoln maps to chromosome 11. *Nat Genet* 1994; **8**: 280–284.
3 Schalling M, Hudson TJ, Buetow KH, Housman DE: Direct detection of novel expanded trinucleotide repeats in the human genome. *Nat Genet* 1993; **4**: 135–139.
4 O'Donovan MC, Guy C, Craddock N *et al*: Expanded CAG repeats in schizophrenia and bipolar disorder. *Nat Genet* 1995; **10**: 380–381.
5 Morris AG, Gaitonde E, McKenna PJ, Mollon JD, Hunt DM: CAG repeat expansions and schizophrenia: association with disease in female and with early age-at-onset. *Hum Mol Genet* 1996; **4**: 1957–1961.
6 O'Donovan MC, Craddock GN, Bowen T *et al*: Confirmation of association between expanded CAG/CTG repeats and both schizophrenia and bipolar disorder. *Psychol Med* 1996; **26**: 1145–1153.
7 McInnis MG, McMahon FJ, Chase GA, Simpson SG, Ross CA, DePaulo JR jr: Anticipation in bipolar affective disorder. *Am J Hum Genet* 1993; **53**: 385–390.
8 Asherson P, Walsh C, Williams J *et al*: Imprinting and anticipation. Are they relevant to genetic studies of schizophrenia? *Br J Psychiatry* 1994; **16**: 619–624.
9 Bassett AS, Honer WG: Evidence for anticipation in schizophrenia. *Am J Hum Genet* 1994; **54**: 864–870.
10 Nylander PO, Engstrom C, Chotai J, Wahlstrom J, Adolfsson R: Anticipation in Swedish families with bipolar affective disorder. *J Med Genet* 1994; **31**: 686–689.
11 Risch N, Botstein D: A manic depressive history. *Nat Genet* 1996; **12**: 351–353.
12 Berrettini WH, Feraro TN, Goldin LR *et al*: Chromosome 18 DNA markers and manic-depressive illness: Evidence for a susceptibility gene. *Proc Natl Acad Sci USA* 1994; **91**: 5918– 5921.
13 Stine OC, Xu J, Koskela R *et al*: Evidence for linkage of Bipolar Disorder to chromosome 18 with parent-of-origin effect. *Am J Hum Genet* 1995; **57**: 1384–1394.
14 Fremier 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 Breschel TS, McInnis MG, Margolis RL *et al*: A novel, heritable, expanding CTG repeat in an intron of the *SEF2-1* gene on chromosome 18q21. *Hum Mol Genet* 1997; **6**(11): 1855–1863.
16 Haaf T, Sirugo G, Kidd KK, Ward DC: Chromosomal localization of long trinucleotide repeats in the human genome by fluorescence in situ hybridization. *Nat Genet* 1996; **12**: 183–185.
17 Gastier JM, Pulido JC, Sunden S *et al*: Survey of trinucleotide repeats in the human genome: assessment of their utility as genetic markers. *Hum Mol Genet* 1995; **4**: 1829–1836.
18 Gastier JM, Brody T, Pulido JC *et al*: Development of a screening set for new (CAG/CTG)n dynamic mutations. *Genomics* 1996; **32**: 75–85.
19 Neri C, Albanese V, Lebre AS *et al*: Survey of CAG/CTG repeats in human cDNAs representing new genes: candidates for inherited disorders. *Hum Mol Genet* 1996; **5**: 1001–1009.
20 Trottier Y, Lutz Y, Stevanin G *et al*: Polyglutamine expansion as a pathological epitope in Huntington's disease and four dominant cerebellar ataxias. *Nature* 1995; **378**: 403–406.
21 Imbert G, Saudou F, Yvert G *et al*: Cloning of the gene for spinocerebellar ataxia 2 reveals a locus with high sensitivity to expanded CAG/glutamine repeats. *Nat Genet* 1996; **14**: 285–291.
22 Stevanin G, Trottier Y, Cancel G *et al*: Screening for proteins with polyglutamine expansions in autosomal dominant cerebellar ataxias. *Hum Mol Genet* 1996; **5**: 1887–1892.
23 Sanpei K, Takano H, Igarashi S *et al*: Identification of the spinocerebellar ataxia type 2 gene using a direct identification of repeat expansion and cloning technique, DIRECT. *Nat Genet* 1996; **14**: 277–284.
24 Koob MD, Benzow KA, Bird TD, Day JW, Moseley ML, Ranum LPW: Rapid cloning of expanded trinucleotide repeat sequences from genomic DNA. *Nat Genet* 1998; **18**: 72–76.
25 Bowen T, Guy C, Speight G *et al*: Expansion of 50 CAG/CTG repeats excluded in schizophrenia by application of a highly efficient approach using repeat expansion detection and a PCR screening set. *Am J Hum Genet* 1996; **59**: 912–917.
26 Gibbs M, Collick A, Kelly RG, Jeffreys AJ: A tetranucleotide repeat mouse minisatellite displaying substantial instability during early preimplantation development. *Genomics* 1993; **17**: 121–128.
27 Kang S, Jaworski A, Ohshima K, Wells RD: Expansion and deletion of CTG repeats from human disease genes are determined by the direction of replication in *E. coli*. *Nat Genet* 1995; **10**: 213–218.
28 Munroe DJ, Haas M, Whitton T *et al*: IRE-Bubble PCR: A rapid method for efficient and representative amplification of human genome DNA sequences from complex sources. *Genomics* 1994; **19**: 506–514.
29 Gyapay G, Schmitt K, Fizames C *et al*: A radiation hybrid map of the human genome. *Hum Mol Genet* 1996; **5**: 339–346.
30 Lichter P, Tang CJC, Call K *et al*: High-resolution mapping of human chromosome 11 by *in situ* hybridization with cosmid clones. *Science* 1990; **247**: 64–69.
31 Hoovers JMH, Mannens M, John R *et al*: High-resolution localisation of 69 potential human zinc finger protein genes: A number are clustered. *Genomics* 1992; **12**: 254–263.
32 Giacalone J, Li X, Lehrach H, Francke U: High-density radiation hybrid map of human chromosome 18 and contig of 18p. *Genomics* 1996; **37**: 9–18.
33 Schuler GD, Boguski MS, Stewart EA *et al*: A gene map of the human genome. *Science* 1996; **274**: 540–546.
34 Zhuchenko O, Bailey J, Bonnen P *et al*: Autosomal dominant cerebellar ataxia (SCA6) associated with small polyglutamine expansions in the $\alpha 1_A$-voltage-dependent calcium channel. *Nat Genet* 1997; **15**: 62–69.