J Hum Genet (2003) 48:111–118

ORIGINAL ARTICLE

Mingshun Li · Kinya Ishikawa · Shuta Toru
Hiroyuki Tomimitsu · Minoru Takashima · Jun Goto
Yoshihisa Takiyama · Hidenao Sasaki · Issei Imoto
Johji Inazawa · Tatsushi Toda · Ichiro Kanazawa
Hidehiro Mizusawa

Physical map and haplotype analysis of 16q-linked autosomal dominant cerebellar ataxia (ADCA) type III in Japan

Received: June 25, 2002 / Accepted: November 22, 2002

Abstract Autosomal dominant cerebellar ataxia (ADCA) is a group of heterogeneous neurodegenerative disorders. We previously mapped a gene locus for ADCA with pure cerebellar syndrome (ADCA type III) to a 3-cM region in chromosome 16q, and found a common haplotype among affected individuals. This region was exactly within the locus for another ADCA, spinocerebellar ataxia type 4 (SCA4). To identify the gene causing 16q-linked ADCA type III, we constructed a contig with 38 bacterial artificial chromosome clones between D16S3043 and D16S3095. The size of this contig was estimated to be 4.8 Mb. We found more than 500 nucleotide tandem repeats, including 9 CAG/CTG repeats in this candidate region, although none of the 94 tandem repeats analyzed were expanded in affected individuals. However, we found 11 new polymorphic markers, giving 22 markers spanning the candidate region. By typing these markers on eight Japanese families with ADCA type III, including two new families, we found that a common "founder" haplotype is seen in a more restricted 3.8-Mb region, spanning markers GGAA05 and D16S3095. We present here a newly refined critical interval of 16q-ADCA type III/SCA4. Data of 11 new DNA markers on 16q22.1 would also be useful for other research of genes mapped to this region.

Key words Autosomal dominant cerebellar ataxia type III (ADCA type III) · Spinocerebellar ataxia type 4 (SCA4) · 16q22.1 · Physical map · Bacterial artificial chromosome (BAC) contig · Haplotype analysis

M. Li · K. Ishikawa · S. Toru · H. Tomimitsu · M. Takashima · H. Mizusawa (✉)
Department of Neurology and Neurological Science, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8519, Japan
Tel. +81-3-5803-5233; Fax +81-3-5803-0134
e-mail: h-mizusawa.nuro@tmd.ac.jp

S. Toru
Japan Health Sciences Foundation, Tokyo, Japan

J. Goto · I. Kanazawa
Department of Neurology, Division of Neuroscience, Graduate School of Medicine, University of Tokyo, Tokyo, Japan

Y. Takiyama
Department of Neurology, Jichi Medical School, Tochigi, Japan

H. Sasaki
Department of Neurology, Hokkaido University School of Medicine, Sapporo, Japan

I. Imoto · J. Inazawa
Department of Molecular Cytogenetics, Medical Research Institute, Tokyo Medical and Dental University, Tokyo, Japan

T. Toda
Division of Functional Genomics, Department of Post-Genomics and Diseases, Course of Advanced Medicine, Osaka University Graduate School of Medicine, Suita, Japan

Introduction

Autosomal dominant cerebellar ataxia (ADCA) is a group of heterogeneous neurodegenerative disorders (Harding 1982). It is clinically characterized by progressive cerebellar ataxia, although other extracerebellar signs, such as pyramidal or extrapyramidal signs, ophthalmoparesis, amyotrophy, and peripheral neuropathy may be variably present. ADCA is classified into three types on the basis of its clinical features (Harding 1982). ADCA types I and II are clinically characterized by prominent extracerebellar signs and macular dystrophy, respectively. In contrast, ADCA type III is characterized by purely cerebellar ataxia throughout the course of illness (Harding 1982; Ishikawa et al. 1996).

At present, 18 loci have been reported to be responsible for ADCA, although there are still other families that do not map to any of these loci. Among these disorders, causative genes have been further identified as expansions of a trinucleotide (CAG) repeat for SCA1 (Orr et al. 1993), SCA2 (Imbert et al. 1996; Sanpei et al. 1996), Machado-Joseph disease/SCA3 (Kawaguchi et al. 1994), SCA6 (Zhuchenko et al. 1997), SCA7 (David et al. 1997), SCA12 (Holmes et al. 1999), SCA17 (Koide et al. 1999; Nakamura et al. 2001) and dentatorubral and pallidoluysian atrophy (DRPLA) (Nagafuchi et al. 1994). Spinocerebellar ataxia type 8 has been proposed to be caused by the expansion of

a CTG repeat in the 3′ noncoding region of the SCA8 gene (Koob et al. 1999), and the SCA10 mutation has been demonstrated to be the expansion of the pentanucleotide ATTCT repeat in an intron of this gene (Matsuura et al. 2000). Only the gene loci have been identified for the remaining 8 SCAs: SCA4 on 16q22.1 (Gardner et al. 1994; Flanigan et al. 1996), SCA5 on 11 cen (Ranum et al. 1994), SCA11 on 15q14–q21.3 (Worth et al. 1999), SCA13 on 19q13.3–q13.4 (Herman-Bert et al. 2000), SCA14 on 19q13.4–q ter (Yamashita et al. 2000), SCA16 on 8q22.1–q24.1 (Miyoshi et al. 2001), SCA19 (Verbeek et al. 2002), and SCA21 (Vuillaume et al. 2002).

We previously mapped a gene responsible for ADCA type III to chromosome 16q (Nagaoka et al. 2000). The region spans a 10.9-cM region flanked by D16S3089 and D16S515, the locus identical to the SCA4 locus (Gardner et al. 1994; Flanigan et al. 1996). We further identified that our "16q-linked ADCA type III" has a common founder haplotype for markers between D16S3043 and D16S3095 in this region (Takashima et al. 2001).

To identify the causative gene mutation for 16q-linked ADCA type III, we first constructed a physical map spanning D16S3043 and D16S3095 because no complete physical maps had been available. We found 11 new polymorphic markers in this region. Using these markers, we analyzed the haplotype of eight families and narrowed the critical region.

We searched tandem repeats in this candidate region for any abnormal expansion in our patients, because expansions of certain repeat sequences, such as CAG/CTG or ATTCT repeats, are the only causative mutations for SCAs.

Here we show a complete bacterial artificial chromosome (BAC) contig of the locus for 16q-linked ADCA type III, information of newly developed markers in this region, and results of haplotype analysis.

Patients and methods

Patients and families

Eight families were studied, six of which have been previously described (Ishikawa et al. 1996; Nagaoka et al. 2000; Takashima et al. 2001). Two other families have been included for the first time in this study because they also showed later-onset pure cerebellar ataxia with autosomal dominant inheritance (family trees are available on request). From our interviews with family members, we determined that these two families originated from the Tokyo Metropolitan area (T1) and the Gunma Prefecture (T2), indicating that none of our eight families are related.

Thirty-eight Japanese individuals whose family members have no history of neurodegenerative diseases were chosen as normal controls.

A peripheral blood sample was obtained with informed consent from each family member and control individual, and genomic DNA was extracted as previously described (Ishikawa et al. 1996).

Construction of the BAC contig

BAC clones were initially searched from National Center for Biotechnology Information (NCBI) genome sequencing contigs (http://www.ncbi.nlm.nih.gov/genome/seq) by microsatellite DNA markers linked to the critical interval. By this method, we obtained eight BAC clones, Rp11-123C5, Rp11-112G1, Rp11-361L15, Rp11-167P11, Rp11-292B23, Rp11-521L9, Rp11-502K10, and Rp11-311C24, from Research Genetics (Houtsville, AL, USA).

DNA from each BAC clone was extracted by a modified QIAGEN Plasmid Mini Purification Protocol (QIAGEN, Hilden, Germany). BAC clones were first confirmed to contain microsatellite DNA markers or Sequence targged site (STS) markers by polymerase chain reaction (PCR). These clones were then "end-sequenced" by T7 (5′-ATACGAC TCACTATA-3′) and Sp6 (5′-ATTTAGGTGACACTA TA-3′) primers.

We then browsed for sequences of BAC clones in BLASTN (http://www.ncbi.nlm.nih.gov/BLAST/) after eliminating repeat sequences using the RepeatMasker program (http://ftp.genome.washington.edu/cgi-bin/RepeatMasker), and, by repeating this method, we finally obtained five partial BAC contigs that nearly covered the entire candidate region. Four small remaining gaps were filled with six BAC clones by browsing databases: Ensembl Map view (http://www.ensembl.org/perl/mapview?chr=16), HGREP (http://hgrep.ims.u-tokyo.ac.jp/cgi-bin/HTG_tool/view.cgi), and TIGR (http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_intro.htm/).

Sequencing

Sequencing reactions were performed using the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction kit (PE Applied Biosystems, Foster City, CA, USA). We cycle-sequenced 2–3 μg BAC DNA for 50 cycles with 6.4 pmol of either T7 or Sp6 primer on an ABI GeneAmp PCR system 9700 (PE Applied Biosystems).

For genomic DNA, 10 ng PCR product was cycle-sequenced for 25 cycles with 3.2 pmol of each primer. Nucleotide sequences were analyzed on an ABI PRISM 377 DNA Sequencer (PE Applied Biosystems).

Searching nucleotide-repeat variation

We derived the BAC clone sequence from NCBI (http://www.ncbi.nlm.nih.gov/entrez). To identify tandem repeats, namely, 2 (di-), 3 (tri-), 4 (tetra-), 5 (penta-), and 6 (hexa-) tandem repeats, we used the RepeatMasker program. Oligonucleotide primers were then designed to amplify the genomic DNA segment flanking each repeat. PCR was performed in a final volume of 20 μl, containing 10 ng of genomic DNA, 3.4 pmol of each primer, 2.5 mM each of deoxyribonucleoside triphosphates, and 0.75 unit of Taq polymerase (Takara, Tokyo, Japan) or Gold Taq polymerase (Applied Biosystems Ampli-Taq Gold with Gene Amp).

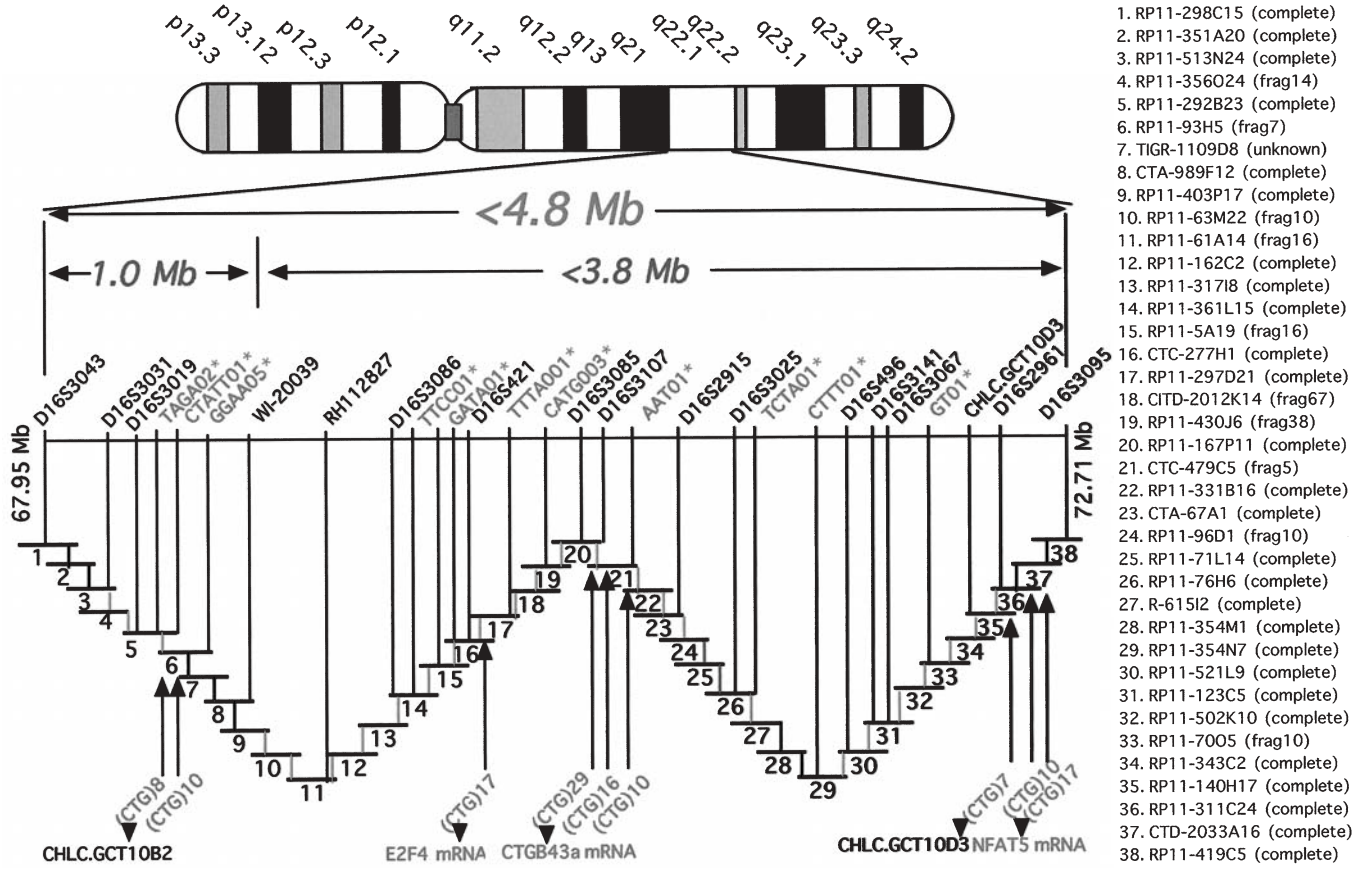

Fig. 1. The physical map of the candidate region of 16q-linked autosomal dominant cerebellar ataxia (ADCA) type III. The size of this bacterial artificial chromosome (BAC) contig was estimated to be no more than 4.8 Mb, with 38 BAC clones spanning the centromeric marker D16S3043 (shown on the *left*) and the telomeric marker D16S3095 (shown on the *right*). The identification of each BAC clone and the most up-to-date sequence status of each clone [i.e., completely sequenced or still in numbers of fragments (*frag*)] are shown on the *right*. *Shaded lines* between BAC clones indicate that the connection of the two BAC clones was confirmed by end sequences, and *bold lines* indicate that the connection was confirmed from database searches. The sequence of BAC clone TIGR-1109D8 is unknown. BAC Rp11-93H5, CTC277H1, CTC479C5, Rp11-331B16, Rp11-140H17, and Rp11-311C24 contain (CAG/CTG)n repeats. The locations of 11 newly identified polymorphic markers are indicated with *asterisks*. The new founder haplotype region is no more than 3.8 Mb between GGAA05 and D16S3095

The PCR products were separated on 2% agarose gel. To precisely genotype each DNA sample, we labeled the 5′ end of the forward primer with fluorescent isothiocyanate, and amplified it. PCR products were separated on the Automated DNA Sequencer II (ALF, Pharmacia Biotech, Umea, Sweden) and analyzed with ALF Win Manager as previously described (Ishikawa et al., 1996; Ishikawa et al. 1997). Eleven new markers with the highest polymorphism were then selected for genotyping and analysis for association with disease.

Haplotype analysis and linkage disequilibrium

Haplotype analysis was performed with 22 markers, including the 11 new markers, that showed enough polymorphism: 16q cen-D16S3043, D16S3031, CTATT01, TAGA02, GGAA05, D16S3086, TTCC01, GATA01, D16S421, TTTA001, CATG003, D16S3085, D16S3107, AAT01, D16S3025, TCTA01, CTTT01, D16S496, D16S3067, D16S3141, GT01, and D16S3095 16q ter-, in the order shown on our physical map (Fig. 1). We analyzed genotypes of all available subjects, including those who married into these families. Because the 22 markers spaced in a distance of approximately 5 Mb, we reconstructed haplotypes in every family that showed minimum recombination events. The association of disease phenotype and each allele was statistically analyzed by Fisher's exact probability test.

Results

Construction of a BAC-based physical map contig

A complete contig spanning markers D16S3043 and D16S3095 was constructed with 38 BAC clones (Fig. 1). The size of this contig was estimated to be no more than 4.8 Mb from the database ENSEMBL. According to the most recent NCBI database, 25 of the 38 BAC clones (65%) have been completely sequenced. On the other hand, the sequence of one BAC clone, TIGR 1109D8, is unknown.

The remaining 12 BAC clones have some gaps in their sequences as of March 2002, although approximately 70% of their entire inserts have been published in a draft sequence version. Therefore, the nucleotide sequences of approximately 90% of the entire contig is now available.

According to the NCBI and ENSEMBL databases, at least 100 annotated genes are present within the candidate region, and 65 of them are expressed in the human brain.

Identification of repeat sequences

To search anonymous repeat sequences by RepeatMasker, we first had to know how powerful this software was in detecting tandem repeats. We inserted known tandem repeats with different repeat units in the known 6-kb nucleotide sequences, and applied the software program. By this test, we found that the software sufficiently detects dinucleotide repeats with more than nine repeat units, trinucleotide repeats with more than six repeat units, tetranucleotide repeats with more four repeat units, and penta- or hexanucleotide repeats with more than three repeat units (Table 1).

We next applied this software to each BAC contig and found 146 dinucleotide, 83 trinucleotide, 213 tetranucleotide, and 91 penta- or hexanucleotide repeats in the entire sequenced region (Table 2). These repeats were confirmed to be unique by checking the flanking sequences. This would indicate that there are more than 500 tandem repeats within the entire contig.

Table 1. RepeatMasker program detectable repeat units and selected repeat units in this study

	Detectable repeat units	Selected repeat units
Dinucleotide	>9	>19
Trinucleotide	>6	>6 (ATT/TAA>14)
Tetranucleotide	>4	>9
Pentanucleotide	>3	>4 (except ATTTT/TAAAA)
Hexanucleotide	>3	>4 (except ATTTTT/TAAAAA)

Identification of 11 new polymorphic markers in 16q22.1

Among approximately 500 tandem repeats, we first selected repeats with higher numbers of repeat units, because repeats with larger repeat units tend to be more polymorphic. From this context, we chose 14 dinucleotide repeats that have more than 19 repeat units (Table 1). Similarly, we chose 37 trinucleotide repeats that have more than 6 repeat units, except for ATT/TAA repeats; 32 tetranucleotide repeats with more than 9 repeat units; and 9 penta- or hexanucleotide repeats with more than 4 repeat units. For ATT/TAA repeats, we chose 2 repeats that have more than 14 repeat units. We examined whether these 94 markers were polymorphic.

When 38 normal Japanese individuals were genotyped, 26% of the tested repeats were demonstrated to be polymorphic: 12 dinucleotide, 1 trinucleotide, 10 tetranucleotide, and 1 pentanucleotide.

Among these 24 repeats, we selected the most polymorphic 11 repeats as "new" markers to analyze haplotypes in patients with ADCA type III (Tables 3, 4, 5). The sequences of primers of these 11 markers are described in Table 3, and the size of PCR products and allele frequencies among 38 control individuals are described in Table 4.

Haplotype and association analyses

Haplotypes of eight 16q-linked ADCA type III families are shown in Table 5. The P4 family from our previous study (Ishikawa et al. 1997; Nagaoka et al. 2000; Takashima et al. 2001) had the haplotype 3(316 bp)-5(194 bp)-6(260 bp), which was different from the haplotype of other

Table 2. Tandem repeats in the candidate region

Experimented/ total number	Types of repeats				
Dinucleotide 14/146	(CA/TG)n 11/100	(TA/AT)n 1/37	(GA/TC)n 2/9		
Trinucleotide 39/83	(CAG/GTC)n 9/9	(CGG/GCC)n 4/7	(CAA/GTT)n 16/24	(TAA/ATT)n 6/38	
	(ATG/TAC)n 0/1	(ACC/TGG)n 3/3	(TCC/AGG)n 1/1		
Tetranucleotide 32/213	(TTCC/GGAA)n 10/16	(TTCA/AAGT)n 6/14	(TAAA/ATTT)n 2/89	(TTTG/AAAC)n 0/42	(TGGA/ACCT)n 2/13
	(GGGT/CCCA)n 0/1	(CAAT/GTTA)n 1/6	(TCCC/AGGG)n 1/4	(TCTA/AGAT)n 1/3	
	(TTTC/AAAG)n 2/11	(CATA/GTAT)n 0/2	(TTAA/AATT)n 0/1	(ATAC/TATG)n 0/3	
	(CAGG/GTCC)n 1/1	(GATA/CTAT)n 3/3	(TGAA/ACTT)n 2/3	(CACG/GTGC)n 1/1	
Penta-, hexanucleotide 9/91	(TAAAA/ATTTT)n 1/25	(TTTTG/AAAAC)n 1/42	(TTTTC/AAAAG)n 0/8	(GTCTG/CAGAC)n 1/1	(CTATT/GATAA)n 1/1
	(GGGGT/CCCCA)n 1/4	(GGATG/CCTAC)n 1/1	(CGGGG/GCCCC)n 0/2	(GGATG/CCTAC)n 1/1	(GGGAGA)n 0/3
	(CATATA)n 1/1	(ATGGTG)n 1/2			

Table 3. New markers developed in this study

Marker name	Located BAC clone	Repeat	Primer sequence (5′-3′)	Standard PCR product size (bp)
CTATT01	RP11-292B23	(CTATT)12	F:GGTGAGGATCTATCTAGATT R:GCCGCTGTGTGAATTAGAA	316
TAGA02	RP11-292B23	(TAGA)17	F:AGTGGGTAGAGGAGATGTTA R:GGGCAGAGAAAGAGTAACAA	194
GGAA05	RP11-93H5	(GA)6(GGGA)6(GGAA)13	F:GTTGCAGTGAGCGGAAATT R:CCTACCTCCAGTCCAGATT	244
TTCC01	RP11-361L15	(TTCC)23	F:CCATCTTATCTCTTAGTCCGT R:GTGAGCCAAGATCATGCCAT	168
GATA01	RP11-5A19	(GATA)16	F:GGGTCACTGAGATGTCCAGTA R:CGCTGTGAGCACTGCTCTCA	157
TTTA001	RP11-297D21	(TTTA)11	F:CCGAAATAGGCGTACCTAGGCTA R:AGTGAGCTGAGTTTGCACCACTG	218
CATG003	RP11-297D21	(CA)20	F:ACCACAGAGCTGAGGCCT R:AGCAAGTCTGCCTTGGAGAGG	192
ATT01	479C5	(ATT)15	F:CGAATCACAGACACACCCAC R:CAGAGATTGCAGTGAGCCAA	230
TCTA01	R-615I2	(TCTA)11	F:GGGAGGCAGTTTAATTCACTT R:GAGCGAGACTCCGTCTCAA	218
CTTT01	354N7	(CTTT)20	F:GCCCCAGATACGCCAGAG R:CCTCTCTACACTCCAGCCT	242
GT01	70O5	(GT)23	F:GGGCACTAAGGACACTTCTT R:GGCTCTGTCATCCATGTTATG	183

BAC, Bacterial artificial chromosome; PCR, polymerase chain reaction

families with the new markers CTATT01, TAGA02, and GGAA05, located in the centromeric region of the candidate interval. The new family T2 also had a different 2(244 bp) allele from other families for the marker GGAA05. However, all families had the same haplotype from the marker D16S3086 to the most telomeric marker GT01 (Table 5): 16 cen-D16S3086: allele 2-TTCC01: allele 4-GATA01: allele 2-D16S421: allele 3-TTTA001: allele 4-CATG003: allele 4-D16S3085: allele 2-D16S3107: allele 7-AAT01: allele 2-D16S3025: allele 4-TCTA01: allele 2-CTTT01: allele 8-D16S496: allele 5-D16S3067: allele 5-D16S3141: allele 3-GT01: allele 6–16q ter. This indicates that a new critical interval of this disease is a region between GGAA05 and D16S3095, no more than 3.8 Mb in size (Fig.1 and Table 5). Among 16 markers spaced in the new critical interval 10 showed significant association with the disease: TTCC01 ($P = 0.0462$), GATA01 ($P = 0.002$), D16S3107 ($P < 0.0001$), AAT01 ($P = 0.0258$), TCTA01 ($P = 0.0227$), CTTT01 ($P = 0.0009$), D16S496 ($P = 0.032$), D16S3067 ($P = 0.008$), D16S3141 ($P = 0.011$), and GT01 ($P < 0.0001$) (Table 5). In particular, three of these markers, D16S3107, CTTT01, and GT01, were in significant linkage disequilibrium (Tables 4, 5). In contrast, none of the 38 control individuals had the particular haplotype 2-4-2-3-4-4-2-7-2-4-2-8-5-5-3-6, indicating that this haplotype is significantly associated with the disease and may be useful for molecular diagnosis of 16q-ADCCA.

Screening of dynamic mutations in 16q-linked ADCA type III patients

In the entire sequenced region, we found nine CAG/CTG repeats that have more than seven repeat units by RepeatMasker (Fig. 1). Four of these were seen in the coding region of three genes, the E2F4 gene (Ginsberg et al. 1994), CTGB43a mRNA, and the NFAT5 gene (Lopez-Rodriguez et al. 1999; Hebinck et al. 2000). CAG repeats in *CTGB43a* and *NFAT5* encode polyglutamine tracts and a CAG repeat in *E2F4* is translated into a polyserine stretch. The other two CAG/CTG repeats have already been annotated as polymorphic CAG repeats in the Cooperative Human Linkage Center (CHLC) database (www.chlc.org). The remaining three CAG/CTG repeats, newly identified in this study, were not located in any known genes.

Amplifying these nine CAG/CTG repeats revealed that none of them were expanded in our patients with 16q-linked ADCA type III. Direct sequencing of the three genes, the E2F4 gene, CTGB43a mRNA, and the NFAT5 gene, disclosed that there are also no mutations in the coding regions of these genes.

Similarly, we tested 30 trinucleotide repeats, 32 tetranucleotide repeats, and 9 penta- or hexanucleotide repeats for polymorphism in control individuals (Table 2). However, we have so far not detected any abnormal expansions in 16q-linked ADCA type III patients.

Discussion

In the present study, we constructed a 4.8-Mb physical map consisting of 38 BAC clones covering the candidate region between D16S3043 and D16S3095 (Takashima et al. 2001). When we started constructing this contig, very few BAC or yeast artificial chromosome clones had been mapped to this region. In addition, the precise order of microsatellite DNA markers, used in examining the haplotype in our previous set of six families with ADCA type III, was not known (Dib et al. 1996; Takashima et al. 2001). By constructing the

Table 4. Size of PCR products and allele frequencies among control individuals on each newly developed DNA marker

Allele	CTATT01 (frequency)	TAGA02 (frequency)	GGAA05 (frequency)	TTCC01 (frequency)	GATA01 (frequency)	TTTA001 (frequency)	CATG003 (frequency)	AAT01 (frequency)	TCTA01 (frequency)	CTTT01 (frequency)	GT01 (freuqnecy)
1	306 bp(0.014)	178 bp (0.094)	**240 bp (0.014)**	155 bp (0.013)	154 bp (0.191)	204 bp (0.05)	188 bp (0.064)	224 bp (0.012)	215 bp (0.013)	226 bp (0.025)	179 bp (0.026)
2	**311 bp (0.186)**	182 bp (0.014)	**244 bp (0.094)**	159 bp (0.026)	**158 bp (0.441)**	208 bp (0.025)	190 bp (0.013)	**230 bp (0.475)**	**219bp (0.46)**	228 bp (0.036)	181 bp (0.171)
3	**316 bp (0.229)**	**186 bp (0.122)**	248 bp (0.149)	163 bp (0.08)	162 bp (0.294)	212 bp (0.275)	192 bp (0.038)	233 bp (0.363)	223 bp (0.408)	230 bp (0.013)	183 bp (0.184)
4	321 bp (0.386)	190 bp (0.432)	252 bp (0.338)	**167 bp (0.605)**	166 bp (0.059)	**216 bp (0.65)**	**194 bp (0.769)**	236 bp (0.113)	227 bp (0.08)	232 bp (0.025)	185 bp (0.237)
5	326 bp (0.143)	**194 bp (0.297)**	256 bp (0.243)	171 bp (0.25)	170 bp (0.015)		196 bp (0.064)	239 bp (0.012)	231 bp (0.039)	234 bp (0.05)	187 bp (0.211)
6	331 bp (0.014)	198 bp (0.041)	**260 bp (0.135)**	175 bp (0.026)			198 bp (0.052)		251 bp (0.025)	236 bp (0.063)	**189 bp (0.158)**
7	336 bp (0.014)		264 bp (0.027)							238 bp (0.087)	
8	341 bp (0.014)									**242 bp (0.288)**	
9										246 bp (0.175)	
10										250 bp (0.087)	
11										254 bp (0.113)	
12										260 bp (0.025)	
13										264 bp (0.013)	

Alleles shared among 16q-linked ADCA type III are highlighted in bold
PCR, Polymerase chain reaction; ADCA III, autosomal dominant cerebellar ataxia type III

physical map of the candidate interval, we now know that the entire length of the interval is no more than 4.8 Mb, and we further refined the critical interval of 16q-linked ADCA type III by the exact order of DNA markers shared among our affected individuals.

By applying the RepeatMasker program, we discovered 11 new highly polymorphic DNA markers that could be used to analyze haplotypes in our eight families. In the present study, we added two new 16q-linked ADCA type III families for haplotype analysis. This allowed us to further narrow a critical interval that shows a common, "founder" haplotype flanked by GGAA05 and D16S3095. The size of this new critical region was estimated to be 3.8 Mb. The present finding not only refines the candidate interval of this disease, but also corroborates our previous finding of a founder effect in the Japanese families with 16q-linked ADCA type III. In addition, analysis of this common haplotype would be useful for the molecular diagnosis of 16q-linked ADCA type III, which was previously diagnosed only with linkage analysis. Because large informative families are fairly uncommon for such a late-onset disease, finding a diagnostic haplotype would be important.

After refining the critical interval, we next searched a causative mutation in our patients. Because 16q-linked ADCA type III showed a mild phenomenon of anticipation, in that difference in the age-of-onset between parents and offsprings was 4.9 years (Nagaoka et al. 2002), we first searched for an expansion of trinucleotide CAG/CTG repeats detected by the software RepeatMasker, although none of the nine CAG/CTG repeats in the entire candidate region were expanded. When we tested the threshold of detecting CAG/CTG repeat sequence by the RepeatMasker program, we found that the program faithfully detected CAG/CTG repeats with more than six repeat units. Therefore, it seems probable that the CAG/CTG repeat, which has at least six repeat units in the normal population, is not expanded in our patients with 16q-linked ADCA type III. The smallest CAG/CTG repeats that are known to cause human neurodegenerative disease are the SCA6 gene and the myotonin protein kinase (*MtPK*) gene (Strachan and Read 1996; Zhuchenko et al. 1997; Ishikawa et al. 1997). These repeats are as small as 4 to 6 repeats in the normal population, and they expand to more than 20 repeats in SCA6 and more than 50 repeats in myotonic dystrophy type 1 (Strachan and Read 1996; Zhuchenko et al. 1997; Ishikawa et al. 1997). Considering that the threshold of detecting CAG/CTG repeats in our system is high, the present finding might indicate that the cause of the 16q-linked ADCA type III is not the expansion of the CAG/CTG repeat.

We also searched for any expansion of tri-, tetra-, penta-, or hexanucleotide repeats, as the ATTCT repeat expansion in SCA10 or the complex CCTG repeat expansion in myotonic dystrophy type 2 (Liquori et al. 2001). However, we did not observe any expansion of repeat sequences among a group of repeats that showed polymorphism in the general population. These observations further indicate that 16q-ADCA type III may be caused by a mutation other than a repeat expansion, such as a missense

Table 5. Linkage disequilibrium and haplotype analysis for 22 markers in the candidate region of 16q22.1-linked ADCA type III 8 families

DNA marker	Locus BAC	Haplotype of disease chromosome for 16q-linked ADCA type III pedigrees								Allele pattern	Allele frequency in general population (%)	Linkage disequilibrium (P value)
		P2	P4	P6	P12	P14	P15	T1	T2			
D16S3043	RP11-120P24	2	2	3	4	6	3	4	1	6	—	—
D16S3031	RP11-120P24	9	9	9	9	9	9	9	9	9	68.1	0.113
CTATT01	RP11-292B23	2	3	2	2	2	2	2	2	8	18.6/22.9	—
TAGA02	RP11-292B23	3	5	3	3	3	3	3	3	6	12.2/29.7	—
GGAA05	RP11-93H5	1	6	1	1	1	1	1	2	7	1.4/9.4/13.5	—
D16S3086	RP11-361L15	2	2	2	2	2	2	2	2	3	65.7	0.093
TTCC01	RP11-361L15	4	4	4	4	4	4	4	4	6	60.5	0.0462*
GATA01	RP11-5A19	2	2	2	2	2	2	2	2	5	44.1	0.002*
D16S421	CTC277H1	3	3	3	3	3	3	3	3	4	75.7	0.205
TTTA001	RP11-297D21	4	4	4	4	4	4	4	4	4	65	0.0566
CATG003	RP11-297D21	4	4	4	4	4	4	4	4	6	76.9	0.0519
D16S3085	RP1-167P11	2	2	2	2	2	2	2	2	4	81.8	0.32
D16S3107	RP1-167P11	7	7	7	7	7	7	7	7	7	13.9	<.0001*
AAT01	CTC479C5	2	2	2	2	2	2	2	2	6	47.5	0.0258*
D16S3025	R-615I2	4	4	4	4	4	4	4	4	6	75	0.215
TCTA01	R-615I2	2	2	2	2	2	2	2	2	5	46	0.0227*
CTTT01	RP11-354N7	8	8	8	8	8	8	8	8	13	28.2	0.0009*
D16S496	RP11-521L9	5	5	5	5	5	5	5	5	5	54.2	0.032*
D16S3067	RP11-123C5	5	5	5	5	5	5	5	5	7	41.9	0.008*
D16S3141	RP11-123C5	3	3	3	3	3	3	3	3	7	44.4	0.011*
GT01	RP11-70O5	6	6	6	6	6	6	6	6	6	15.8	<.0001*
D16S3095	RP11-419C5	2	3	1	1	1	1	1	1	5	—	—

* P values show significant linkage disequilibrium by Fisher's exact probability test ($P < 0.05$)

mutation. If this is proven to be true, the mild anticipation observed in our families might be caused by another mechanism, as in SPG4 (Hazan et al. 1999). Obviously, it is still possible that a certain repeat that has a small repeat unit not detected by the RepeatMasker is specifically expanded in our patients. We also do not preclude the possibility of pathogenic expansion of a certain repeat hidden in the remaining small regions, whose nucleotide sequences have not been published. Screening not only dynamic expansion, but also other types of mutation is needed to discover the cause of 16q-linked ADCA type III.

Finally, the fruitfulness of this study may be important not only for 16q-linked ADCA type III, but also for other diseases linked to this chromosomal region. The finding of 11 new polymorphic markers as well as the present BAC contig in chromosome 16q22.1 might be valuable for diseases linked to 16q22.1, such as candidate tumor suppressor genes in breast cancer (Callen et al. 2002).

In conclusion, the finding of a complete BAC contig in 16q22.1, a number of new polymorphic markers, and a common haplotype in 16q-ADCA type III is very useful information for identifying the mutation for 16q-linked ADCA type III. Identification of the disease gene will greatly facilitate diagnosis of SCA patients and will also have an impact on our understanding of complex molecular mechanisms underlying ADCAs.

References

Callen DF, Crawford J, Derwas C, Cleton-Jansen AM, Cornelisse CJ, Baker E (2002) Defining regions of loss of heterozygosity of 16q in breast cancer cell lines. Cancer Genet Cytogenet 133:76–82

David G, Abbas N, Stevanin G, Durr A, Yvert G, Cancel G, Weber C, Imbert G, Saudou F, Antoniou E, Drabkin H, Gemmill R, Giunti P, Benomar A, Wood N, Ruberg M, Agid Y, Mandel JL, Brice A (1997) Cloning of the *SCA7* gene reveals a highly unstable CAG repeat expansion. Nat Genet 17:65–70

Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J (1996) A comprehensive genetic map of the human genome based on 5264 microsatellites. Nature 380:152–154

Flanigan K, Gardner K, Alderson K, Galster B, Otterud B, Leppert MF, Kaplan C, Ptacek LJ (1996) Autosomal dominant spinocerebellar ataxia with sensory axonal neuropathy (SCA4): clinical description and genetic localization to chromosome 16q22.1. Am J Hum Genet 59:392–399

Gardner K, Alderson K, Galster B, Kaplan C, Leppert M, Ptacek L (1994) Autosomal dominant spinocerebellar ataxia: clinical description of a distinct hereditary ataxia and genetic localization to chromosome 16 (SCA4) in a Utah kindred. Neurology 44:A361

Ginsberg D, Vario G, Chittenden T, Xiao ZX, Xu G, Wydner KL, DeCaprio JA, Lawrence JB, Livingston DM (1994) E2F-4, a new member of the E2F transcription factor family, interacts with p107. Genes Dev 8:2665–2679

Harding AE (1982) The clinical features and classification of the late onset autosomal dominant cerebellar ataxias. A study of 11 families, including descendants of the "the Drew family of Walworth." Brain 105:1–28

Hazan J, Fonkechten N, Mavel D, Paternotte C, Samson D, Artiguenave F, Davoine CS, Cruaud C, Durr A, Wincker P, Brottier P, Cattolico L, Barbe V, Burgunder JM, Prud'homme JF, Brice A, Fontaine B, Heilig B, Weissenbach J (1999) Spastin, a new AAA protein, is altered in the most frequent form of autosomal dominant spastic paraplegia. Nat Genet 23:296–303

Hebinck A, Dalski A, Engel H, Mattei M-G, Hawken R, Schwinger E, Zuhlke C (2000) Assignment of transcription factor NFAT5 to human chromosome 16q22.1, murine chromosome 8D and porcine chromosome 6p1.4 and comparison of the polyglutamine domains. Cytogenet Cell Genet 90:68–70

Herman-Bert A, Stevanin G, Netter JC, Rascol O, Brassat D, Calvas P, Camuzat A, Yuan Q, Schalling M, Dürr A, Brice A (2000) Mapping of spinocerebellar ataxia 13 to chromosome 19q13.3–q13.4 in a family with autosomal dominant cerebellar ataxia and mental retardation. Am J Hum Genet 67:229–235

Holmes SE, O'Hearn EE, McInnis MG, Gorelick-Feldman DA, Kleiderlein JJ, Callahan C, Kwak NG, Ingersoll-Ashworth RG, Sherr M, Sumner AJ, Sharp AH, Ananth U, Seltzer WK, Boss MA, Vieria-Saecker AM, Epplen JT, Riess O, Ross CA, Margolis RL (1999) Expansion of a novel CAG trinucleotide repeat in the 5′ region of PPP2R2B is associated with SCA12. Nat Genet 23:391–392

Imbert G, Saudou F, Yvert G, Devys D, Trottier Y, Garnier JM, Weber C, Mandel JL, Cancel G, Abbas N, Durr A, Didierjean O, Stevanin G, Agid Y, Brice A (1996) Cloning of the gene for spinocerebellar ataxia 2 reveals a locus with high sensitivity to expanded CAG/glutamine repeats. Nat Genet 14:285–291

Ishikawa K, Mizusawa H, Saito M, Tanaka H, Nakajima N, Kondo N, Kanazawa I, Shoji S, Tsuji S (1996) Autosomal dominant pure cerebellar ataxia. A clinical and genetic analysis of eight Japanese families. Brain 119:1173–1182

Ishikawa K, Tanaka H, Saito M, Ohkoshi N, Fujita T, Yoshizawa K, Ikeuchi T, Watanabe M, Hayashi A, Takiyama Y, Nishizawa M, Nakano I, Matsubayashi K, Miwa M, Shoji S, Kanazawa I, Tsuji S, Mizusawa H (1997) Japanese families with autosomal dominant pure cerebellar ataxia map to chromosome 19p13.1–p13.2 and are strongly associated with mild CAG expansions in the spinocerebellar ataxia type 6 gene in chromosome 19p13.1. Am J Hum Genet 61:336–346

Kawaguchi Y, Okamoto T, Taniwaki M, Aizawa M, Inoue M, Katayama S, Kawakami H, Nakamura S, Nishimura M, Akiguchi I (1994) CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1. Nat Genet 8:221–228

Koide R, Kobayashi S, Shimohata T, Ikeuchi T, Maruyama M, Saito M, Yamada M, Takahashi H, Tsuji S (1999) A neurological disease caused by an expanded CAG trinucleotide repeat in the TATA-binding protein gene: a new polyglutamine disease? Hum Mol Genet 8:2047–2053

Koob MD, Moseley ML, Schut LJ, Benzow KA, Bird TD, Day JW, Ranum LPW (1999) An untranslated CTG expansion causes a novel form of spinocerebellar ataxia (SCA8). Nat Genet 21:379–384

Liquori CL, Ricker K, Moseley ML, Jacobsen JF, Kress W, Naylor SL, Day JW, Ranum LP (2001) Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. Science 293:864–867

Lopez-Rodriguez C, Aramburu J, Rakeman AS, Rao A (1999) NFAT5, a constitutively nuclear NFAT protein that does not cooperate with Fos and Jun. Proc Natl Acad Sci USA 96:7214–7219

Matsuura T, Yamagata T, Burgess DL, Rasmussen A, Grewal RP, Watase K, Khajavi M, McCall AE, Davis CF, Zu L, Achari M, Pulst SM, Alonso E, Noebels JL, Nelson DL, Zoghbi HY, Ashizawa T (2000) Large expansion of the ATTCT pentanucleotide repeat in spinocerebellar ataxia type 10. Nat Genet 26:191–194

Miyoshi Y, Yamada T, Tanimura M, Taniwaki T, Arakawa K, Ohyagi Y, Furuya H, Yamamoto K, Sakai K, Sasazuki T, Kira J (2001) A novel autosomal dominant spinocerebellar ataxia (SCA16) linked to chromosome 8q22.1–24.1. Neurology 57:96–100

Nagafuchi S, Yanagisawa H, Ohsaki E, Shirayama T, Tadokoro K, Inoue T, Yamada M (1994) Structure and expression of the gene responsible for the triplet repeat disorder, dentatorubral and pallidoluysian atrophy (DRPLA). Nat Genet 8:177–182

Nagaoka U, Takashima M, Ishikawa K, Yoshizawa K, Yoshizawa T, Ishikawa M, Yamawaki T, Shoji S, Mizusawa H (2000) A gene on SCA4 locus causes dominantly inherited pure cerebellar ataxia. Neurology 54:1971–1975

Nakamura K, Jeong SY, Uchihara T, Anno M, Nagashima K, Nagashima T, Ikeda S, Tsuji S, Kanazawa I (2001) SCA17, a novel autosomal dominant cerebellar ataxia caused by an expanded polyglutamine in TATA-binding protein. Hum Mol Genet 10:1441–1448

Orr HT, Chung MY, Banfi S, Kwiatkowski TJ Jr, Servadio A, Beaudet AL, McCall AE, Duvick LA, Ranum LP, Zoghbi HY (1993) Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. Nat Genet 4:221–226

Ranum LP, Schut LJ, Lundgren JK, Orr HT, Livingston DM (1994) Spinocerebellar ataxia type 5 in a family descended from the grandparents of president Lincoln maps to chromosome 11. Nat Genet 8:280–284

Sanpei K, Takano H, Igarashi S, Sato T, Oyake M, Sasaki H, Wakisaka A, Tashiro K, Ishida Y, Ikeuchi T, Koide R, Saito M, Sato A, Tanaka T, Hanyu S, Takiyama Y, Nishizawa M, Shimizu N, Nomura Y, Segawa M, Iwabuchi K, Eguchi I, Tanaka H, Takahashi H, Tsuji S (1996) Identification of the spinocerebellar ataxia type 2 gene using a direct identification of repeat expansion and cloning technique, DIRECT. Nat Genet 14:277–284

Strachan T, Read AP (1996) Human molecular genetics. BIOS Scientific, Oxford UK, pp 266–267

Takashima M, Ishikawa K, Nagaoka U, Shoji S, Mizusawa H (2001) A linkage disequilibrium at the candidate gene locus for 16q-linked autosomal dominant cerebellar ataxia type III in Japan. J Hum Genet 46:167–171

Verbeek DS, Schelhaas JH, Ippel EF, Beemer FA, Pearson PL, Sinke RJ (2002) Identification of a novel SCA locus (SCA19) in a Dutch autosomal dominant cerebellar ataxia family on chromosome region 1p21–q21. Hum Genet 111:388–393

Vuillaume I, Devos D, Schraen-Maschke S, Dina C, Lemainque A, Vasseur F, Bocquillon G, Devos P, Kocinski C, Marzys C, Destee A, Sablonniere B (2002) A new locus for spinocerebellar ataxia (SCA21) maps to chromosome 7p21.3–p15.1. Ann Neurol 52:666–670

Worth PF, Giunti P, Gardner-Thorpe C, Dixon PH, Davis MB, Wood NW (1999) Autosomal dominant cerebellar ataxia type III: linkage in a large British family to a 7.6-cM region on chromosome 15q14–21.3. Am J Hum Genet 65:420–426

Yamashita I, Sasaki H, Yabe I, Fukazawa T, Nogoshi S, Komeichi K, Takada A, Shiraishi K, Takiyama Y, Nishizawa M, Kaneko J, Tanaka H, Tsuji S, Tashiro K (2000) A novel locus for dominant cerebellar ataxia (SCA14) maps to a 10.2-cM interval flanked by D19S206 and D19S605 on chromosome 19q13.4-qter. Ann Neurol 48:156–163

Zhuchenko O, Bailey J, Bonnen P, Ashizawa T, Stockton DW, Amos C, Dobyns WB, Subramony SH, Zoghbi HY, Lee CC (1997) Autosomal dominant cerebellar ataxia (SCA6) associated with small polyglutamine expansions in the alpha 1A-voltage-dependent calcium channel. Nat Genet 15:62–69