Table 1. Clinical Characteristics of Affected Subjects

Pedigree No.	Patient ID	Gender	Onset Age (yr)	Current Age (yr)	Ataxia	Motor Neuron Involvement			Genotype of GGCCTG Repeats
						Skeletal Muscle Atrophy	Skeletal Muscle Fasciculation	Tongue Atrophy/ Fasciculation	
1	III-5	M	50	70	+++	N.D.	N.D.	N.D.	g.263397_263402[6]+(1800)
	III-6	F	52	68	++	+	+	+	g.263397_263402[6]+(2300)
	IV-2	F	57	63	+	-	-	+	g.263397_263402[6]+(2300)
	IV-4	M	50	59	+	-	-	+	g.263397_263402[6]+(2300)
2	II-1	M	55	77	+++	++	+	+	g.263397_263402[6]+(2200)
	II-2	F	53	70	++	N.D.	N.D.	N.D.	g.263397_263402[6]+(2200)
3	II-3	M	58	77	++	++	+	+	g.263397_263402[3]+(2300)
	III-1	M	56	62	+	-	-	±	g.263397_263402[8]+(2200)
	III-2	M	51	61	++	+	+	+	g.263397_263402[6]+(1800)
4	I-1	M	57	died in 2001 at 83	++	N.D.	N.D.	N.D.	g.263397_263402[5]+(1800)
	II-1	F	48	61	++	+	±	++	g.263397_263402[6]+(2000)
5	I-1	M	57	86	++	+++	+	+	g.263397_263402[5]+(2000)
	II-1	F	47	58	++	+	+	+	g.263397_263402[8]+(1700)
	SCA#1	M	52	69	+++	+++	+++	+++	g.263397_263402[5]+(2200)
	SCA#2	F	43	53	+++	-	-	+	g.263397_263402[6]+(1800)
	SCA#3	M	55	60	++	-	-	++	g.263397_263402[8]+(1700)
	SCA#4	M	57	81	+++	+	+	+++	g.263397_263402[5]+(2200)
Mean			52.8						
SD			4.3						

N.D., not determined.

affected individuals, none of whom displayed severe lower limb spasticity or extensor plantar response. Electrophysiological studies were performed in an affected individual. Nerve conduction studies revealed normal findings in all of the cases that were examined; however, an electromyogram showed neurogenic changes only in cases with skeletal muscle atrophy, indicating that lower motor neuropathy existed in this particular disease. Progression of motor neuron involvement in this SCA was typically limited to the tongue and main proximal skeletal muscles in both upper and lower extremities, which is clearly different from typical ALS, which usually involves most skeletal muscles over the course of a few years, leading to fatal results within several years.

We conducted genome-wide linkage analysis for nine affected subjects and eight unaffected subjects in three informative families (pedigrees 1–3; Figure 1). For genotyping, we used an ABI Prism Linkage Mapping Set (Version 2; Applied Biosystems, Foster City, CA, USA) with 382 markers, 10 cM apart, for 22 autosomes. Fine-mapping markers (approximately 1 cM apart) were designed according to information from the uniSTS reference physical map in the NCBI database. A parametric linkage analysis was carried out in GENEHUNTER[8] with the assumption of an autosomal-dominant model. The disease allele frequency was set at 0.000001, and a phenocopy frequency of 0.000001 was assumed. Population allele frequencies were assigned equal portions of individual alleles. We performed multipoint analyses for autosomes and obtained LOD scores. We considered LOD scores above 3.0 to be significant.[8] Genome-wide linkage analysis revealed a single locus on chromosome 20p13 with a LOD score of 3.20. Fine mapping increased the LOD score to 4.60 (Figure 3). Haplotype analysis revealed two recombination events in pedigree 3, delimiting a1.8 Mb region (D20S906–D20S193) (Figure 1). We further tested whether the five cases shared the haplotype. As shown in Figure 1, pedigrees 4 and 5 were confirmed to have the same haplotype as pedigrees 1, 2, and 3, indicating that the 1.8 Mb region is very likely to be derived from a common ancestor.

The 1.8 Mb region harbors 44 genes (NCBI, build 37.1). We eliminated two pseudogenes and five genes (*LOC441938*, *LOC100289473*, *LOC100288797*, *LOC100289507*, and *LOC100289538*) from the candidates. Evidence view showed that the first, fourth, and fifth genes were not found in the contig in this region, whereas the second and third

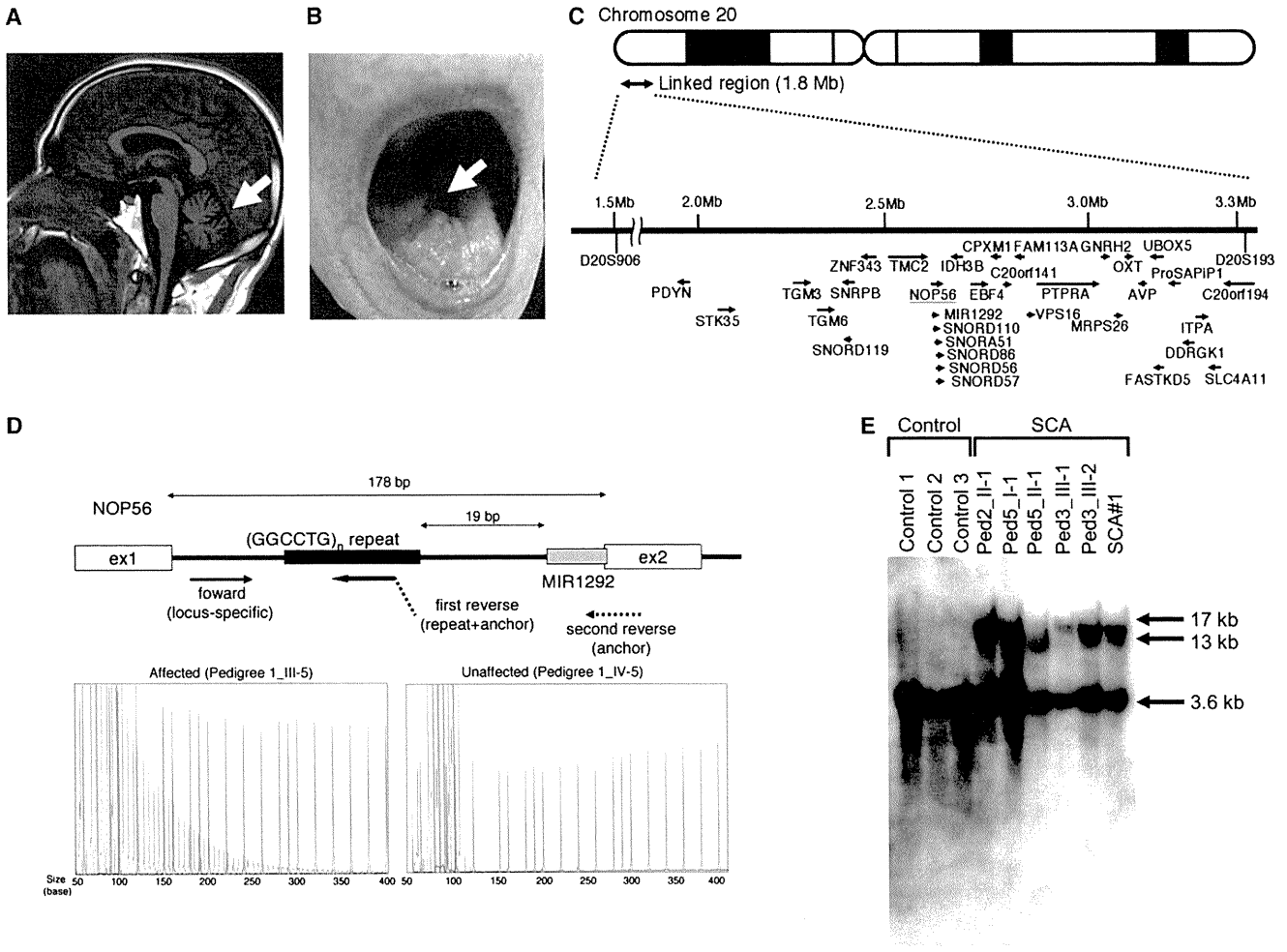


Figure 2. Motor Neuron Involvement and (GGCCTG)n Expansion in the First Intron of *NOP56*

(A) MRI of an affected subject (SCA#3) showed mild cerebellar atrophy (arrow) but no other cerebral or brainstem pathology.

(B) Tongue atrophy (arrow) was observed in SCA#1.

(C) Physical map of the 1.8-Mb linkage region from D20S906 (1,505,576 bp) to D20S193 (3,313,494 bp), with 33 candidate genes shown, as well as the direction of transcription (arrows).

(D) The upper portion of the panel shows the scheme of primer binding for repeat-primer PCR analysis. In the lower portion, sequence traces of the PCR reactions are shown. Red lines indicate the size markers. The vertical axis indicates arbitrary intensity levels. A typical saw-tooth pattern is observed in an affected pedigree.

(E) Southern blotting of LCLs from SCA cases and three controls. Genomic DNA (10 μg) was extracted from Epstein-Barr virus (EBV)-immortalized LCLs derived from six affected subjects (Ped2_II-1, Ped3_III-1, Ped3_III-2, Ped5_I-1, Ped5_II-1, and SCA#1) and digested with 2 U of *AvrII* overnight (New England Biolabs, Beverly, MA, USA). A probe covering exon 4 of *NOP56* (452 bp) was subjected to PCR amplification from human genomic DNA with the use of primers (Table S3) and labeled with ^{32}P-dCTP.

genes are not assigned to orthologous loci in the mouse genome. Sequence similarities among paralog genes defied direct sequencing of four genes: *SIRPD* [NM 178460.2], *SIRPB1* [NM 603889], *SIRPG* [NM 605466], and *SIRPA* [NM 602461]. Thus, we sequenced 33 of 37 genes (*PDYN* [MIM 131340], *STK35* [MIM 609370], *TGM3* [MIM 600238], *TGM6* [NM_198994.2], *SNRPB* [MIM 182282], *SNORD119* [NR_003684.1], *ZNF343* [NM_024325.4], *TMC2* [MIM 606707], *NOP56* [NM_006392.2], *MIR1292* [NR_031699.1], *SNORD110* [NR_003078.1], *SNORA51* [NR_002981.1], *SNORD86* [NR_004399.1], *SNORD56* [NR_002739.1], *SNORD57* [NR_002738.1], *IDH3B* [MIM 604526], *EBF4* [MIM 609935], *CPXM1* [NM_019609.4], *C20orf141* [NM_080739.2], *FAM113A* [NM_022760.3], *VPS16* [MIM 608550], *PTPRA* [MIM 176884], *GNRH2* [MIM 602352], *MRPS26* [MIM 611988], *OXT* [MIM 167050], *AVP* [MIM 192340], *UBOX5* [NM_014948.2], *FASTKD5* [NM_021826.4], *ProSAPiP1* [MIM 610484], *DDRGK1* [NM_023935.1], *ITPA* [MIM 147520], *SLC4A11* [MIM 610206], and *C20orf194* [NM_001009984.1]) (Figure 2C). All noncoding and coding exons, as well as the 100 bp up- and downstream of the splice junctions of these genes, were sequenced in two index cases (IV-4 in pedigree1 and III-1 in pedigree 3) and in three additional cases (II-1 in pedigree 2, II-1 in pedigree 4, and II-1 in pedigree 5) with the use of specific primers (Table S1 available online). Eight unregistered variants were found among the two index cases. Among these, there was a coding variant, c.795C>G

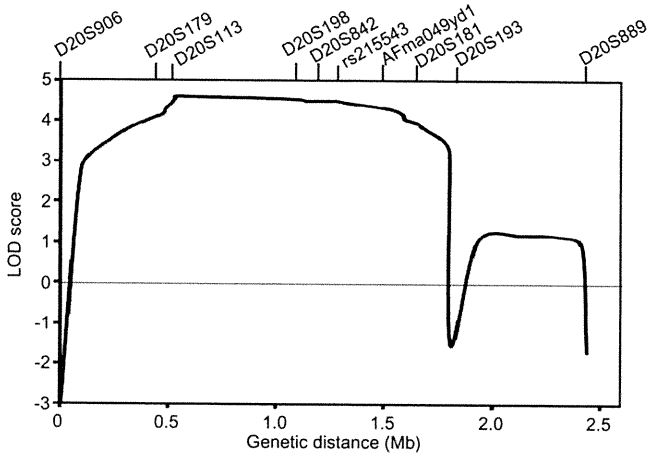


Figure 3. Multipoint Linkage Analysis with Ten Markers on Chromosome 20p13

(p.Phe265Leu), in *C20orf194*, whereas the other seven included one synonymous variant, c.1695T>A (p.Leu565-Leu), in *ZNF343* and six non-splice-site intronic variants (Table S2). We tested segregation by sequencing exon 11 of *C20orf194* in IV-2 and III-5 in pedigree 1. Neither IV-2 nor III-5 had this variant. We thus eliminated *C20orf194* as a candidate. Missense mutations in *PDYN* and *TGM6*, which have been recently reported as causes of SCA, mapped to 20p12.3-p13,[9,10] but none were detected in the five index cases studied here (Table S2).

Possible expansions of repetitive sequences in these 33 genes were investigated when intragenic repeats were indicated in the database (UCSC Genome Bioinformatics). Expansions of the hexanucleotide repeat GGCCTG (rs68063608) were found in intron 1 of *NOP56* (Figure 2D) in all five index cases through the use of a repeat-primed PCR method.[11–13] An outline of the repeat-primed PCR experiment is described in Figure 2D. In brief, the fluorescent-dye-conjugated forward primer corresponded to the region upstream of the repeat of interest. The first reverse primer consisted of four units of the repeat (GGCCTG) and a 5′ tail used as an anchor. The second reverse primer was an "anchor" primer. These primers are described in Table S3. Complete segregation of the expanded hexanucleotide was confirmed in all pedigrees, and the maximum repeat size in nine unaffected members was eight (data not shown).

In addition to the SCA cases in five pedigrees, four unrelated cases (SCA#1–SCA#4) were found to have a (GGCCTG)n allele through screening of the cohort SCA patients (Table 1). Neurological examination was reevaluated in these four cases, revealing both ataxia and motor neuron dysfunction with tongue atrophy and fasciculation (Table 1). In total, nine unrelated cases were found in the 251 cohort patients with SCA (3.6%). For confirmation of the repeat expansions, Southern blot analysis was conducted in six affected subjects (Ped2_II-1, Ped3_III-1, Ped3_III-2, Ped5_I-1, Ped5_II-1, and SCA#1). The data showed >10 kb of repeat expansions in the lymphoblastoid cell lines (LCLs) obtained from the SCA patients (Figure 2E). Furthermore, the numbers of GGCCTG repeat expansion were estimated by Southern blotting in 11 other cases. The expansion analysis revealed approximately 1500 to 2500 repeats in 17 cases (Table 1). There was no negative association between age at onset and the number of GGCCTG repeats ($n = 17$, $r = 0.42$, $p = 0.09$; Figure S1) and no obvious anticipation in the current pedigrees.

To investigate the disease specificity and disease spectrum of the hexanucleotide repeat expansions, we tested the repeat expansions in an Alzheimer disease (MIM 104300) cohort and an ALS cohort followed by the Department of Neurology, Okayama University Hospital. We also recruited Japanese controls, who were confirmed to be free from brain lesions through MRI and magnetic resonance angiography, which was performed as described previously.[14] Screening of the 27 Alzheimer disease cases and 154 ALS cases failed to detect additional cases with repeat expansions. The GGCCTG repeat sizes ranged from 3 to 8 in 300 Japanese controls (5.9 ± 0.8 repeats), suggesting that the >10 kb repeat expansions were mutations.

Expression of *Nop56*, an essential component of the splicing machinery,[15] was examined by RT-PCR with the use of primers for wild-type mouse *Nop56* cDNA (Table S3). Expression of *Nop56* mRNA was detected in various tissues, including CNS tissue, and a very weak signal was detected in spinal cord tissue (Figure 4A). Immunohistochemistry using an anti-mouse Nop56 antibody (Santa Cruz Biotechnology, Santa Cruz, CA, USA) detected the Nop56 protein in Purkinje cells of the cerebellum as well as motor neurons of the hypoglossal nucleus and the spinal cord anterior horn (Figure 4B), suggesting that these cells may be responsible for tongue and muscle atrophy in the trunk and limbs, respectively. Immunoblotting also confirmed the presence of Nop56 in neural tissues (Figure 4C), where Nop56 is localized in both the nucleus and cytoplasm.

Alterations of *NOP56* RNA expression and protein levels in LCLs from patients were examined by real-time RT-PCR and immunoblotting. The primers for quantitative PCR of human *NOP56* cDNA are described in Table S3. Immunoblotting was performed with the use of an anti-human NOP56 antibody (Santa Cruz Biotechnology, Santa Cruz, CA, USA). We found no decrease in *NOP56* RNA expression or protein levels in LCLs from these patients (Figure 5A). To investigate abnormal splicing variants of *NOP56*, we performed RT-PCR using the primers covering the region from the 5′ UTR to exon 4 around the repeat expansion (Table S3); however, no splicing variant was observed in LCLs from the cases (Figure 5B). We also performed immunocytochemistry for NOP56 and coilin, a marker of the Cajal body, where NOP56 functions.[16] NOP56 and coilin distributions were not altered in LCLs of the SCA patients (Figure 5C), suggesting that qualitative or quantitative changes in the Cajal body did not occur. These results indicated that haploinsufficiency could not explain the observed phenotype.

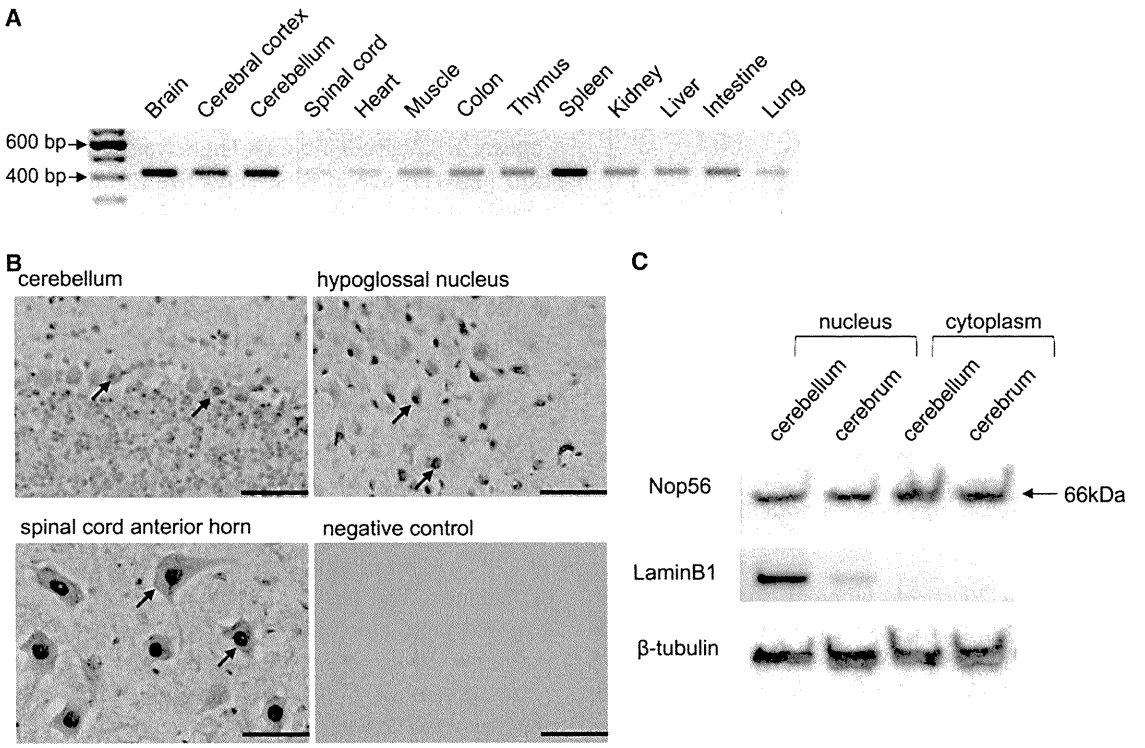


Figure 4. Nop56 in the Mouse Nervous System
(A) RT-PCR analysis of Nop56 (422 bp) in various mouse tissues. cDNA (25 ng) collected from various organs of C57BL/6 mice was purchased from GenoStaf (Tokyo, Japan).
(B) Immunohistochemical analysis of Nop56 in the cerebellum, hypoglossal nucleus, and spinal cord anterior horn in wild-type male Slc:ICR mice at 8 wks of age (Japan SLC, Shizuoka, Japan). The arrows indicate anti- Nop56 antibody staining. The negative control was the cerebellar sample without the Nop56 antibody treatment. Scale bar represents 100 μm.
(C) Immunoblotting of Nop56 (66 kDa) in the cerebellum and cerebrum. Protein sample (10 μg) was subjected to immunoblotting. LaminB1, a nuclear protein, and beta-tubulin were used as loading controls.

We performed fluorescence in situ hybridization to detect RNA foci containing the repeat transcripts in LCLs from patients, as previously described.[17,18] Lymphoblastoid cells from two SCA patients (Ped2_II-2 and Ped5_I-1) and two control subjects were analyzed. An average of 2.1 ± 0.5 RNA foci per cell were detected in 57.0% of LCLs (n = 100) from the SCA subjects through the use of a nuclear probe targeting the GGCCUG repeat, whereas no RNA foci were observed in control LCLs (n = 100) (Figure 6A). In contrast, a probe for the CGCCUG repeat, another repeat sequence in intron 1 of *NOP56*, detected no RNA foci in either SCA or control LCLs (n = 100 each) (Figure 6A), indicating that the GGCCUG repeat was specifically expanded in the SCA subjects. The specificity of the RNA foci was confirmed by sensitivity to RNase A treatment and resistance to DNase treatment (Figure 6A).

Several reports have suggested that RNA foci play a role in the etiology of SCA through sequestration of specific RNA-binding proteins.[5–7] In silico searches (ESEfinder 3.0) predicted an RNA-binding protein, SRSF2 (MIM 600813), as a strong candidate for binding of the GGCCUG repeat. Double staining with the probe for the GGCCUG repeat and an anti-SRSF2 antibody (Sigma-Aldrich, Tokyo, Japan) was performed. The results showed colocalization of RNA foci with SRSF2, whereas NOP56 and coilin were not colocalized with the RNA foci (Figure 6B), suggesting a specific interaction of endogenous SRSF2 with the RNA foci in vivo.

To further confirm the interaction, gel-shift assays were carried out for investigation of the binding activity of SRSF2 with $(GGCCUG)_n$. Synthetic RNA oligonucleotides (200 pmol), $(GGCCUG)_4$ or $(CUG)_6$, which is the latter part of the hexanucleotide, as well as the repeat RNA involved in myotonic dystrophy type 1 (DM1 [MIM 160900])[18] and SCA8 (MIM 608768),[5] were denatured and immediately mixed with different amounts (0, 0.2, or 0.6 μg) of recombinant full-length human SRSF2 (Abcam, Cambridge, UK). The mixtures were incubated, and the protein-bound probes were separated from the free forms by electrophoresis on 5%–20% native polyacrylamide gels. The separated RNA probes were detected with SYBR Gold staining (Invitrogen, Carlsbad, CA, USA). We found a strong association of $(GGCCUG)_4$ with SRSF2 in vitro in comparison to $(CUG)_6$ (Figure 6C). Collectively, we concluded that (GGCCUG)n interacts with SRSF2.

It is notable that *MIR1292* is located just 19 bp 3′ of the GGCCTG repeat (Figure 2D). MiRNAs such as *MIR1292* are small noncoding RNAs that regulate gene expression by inhibiting translation of specific target mRNAs.[19,20] MiRNAs are believed to play important roles in key molecular

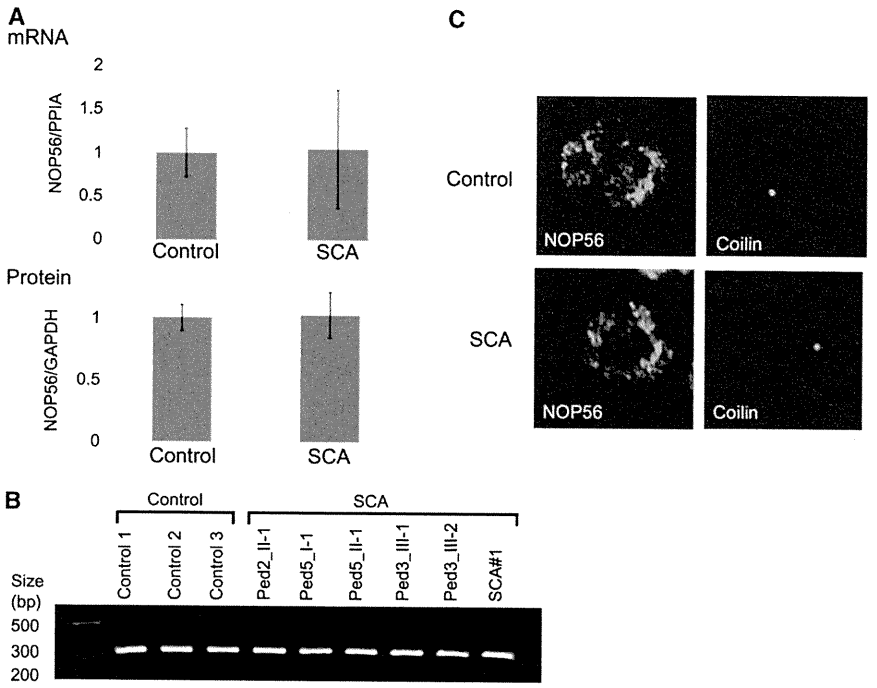


Figure 5. Analysis of NOP56 in LCLs from SCA Patients

(A) mRNA expression (upper panel) and protein levels (lower panel) in LCLs from cases (n = 6) and controls (n = 3) were measured by RT-PCR and immunoblotting, respectively. cDNA (10 ng) was transcribed from total RNA isolated from LCLs and used for RT-PCR. Immunoblotting was performed with the use of a protein sample (40 μg) extracted from LCLs. The data indicate the mean ± SD relative to the levels of *PPIA* and GAPDH, respectively. There was no significant difference between LCLs from controls and cases.

(B) Analysis for splicing variants of NOP56 cDNA. RT-PCR with 10 ng of cDNA and primers corresponding to the region from the 5′ UTR to exon 4 around the repeat expansion was performed. The PCR product has an expected size of 230 bp.

(C) Immunocytochemistry for NOP56 and coilin. Green signals represent NOP56 or coilin. Shown are representative samples from 100 observations of controls or cases.

pathways by fine-tuning gene expression.[19,20] Recent studies have revealed that miRNAs influence neuronal survival and are also associated with neurodegenerative diseases.[21,22] In silico searches (Target Scan Human 5.1) predicted glutamate receptors (*GRIN2B* [MIM 138252] and *GRIK3* [MIM 138243]) to be potential target genes. Real-time RT-PCR using TaqMan probes for miRNA (Invitrogen, Carlsbad, CA, USA) revealed that the levels of both mature and precursor *MIR1292* were significantly decreased in SCA LCLs (Figure 6D), indicating that the GGCCTG repeat expansion decreased the transcription of *MIR1292*. A decrease in *MIR1292* expression may upregulate glutamate receptors in particular cell types; e.g., *GRIK3* in stellate cells in the cerebellum,[23] leading to ataxia because of perturbation of signal transduction to the Purkinje cells. In addition, it has been suggested, on the basis of ALS mouse models,[24,25] that excitotoxicity mediated by a type of glutamate receptor, the NMDA receptor including *GRIN2B*, is involved in loss of spinal neurons. A very slowly progressing and mild form of the motor neuron disease, such as that described here, which is limited to mostly fasciculation of the tongue, limbs and trunk, may also be compatible with such a functional dysregulation rather than degeneration.

In the present study, we have conducted genetic analysis to find a genetic cause for the unique SCA with motor neuron disease. With extensive sequencing of the 1.8 Mb linked region, we found large hexanucleotide repeat expansions in *NOP56*, which were completely segregated with SCA in five pedigrees and were found in four unrelated cases with a similar phenotype. The expansion was not found in 300 controls or in other neurodegenerative diseases. We further proved that repeat expansions of *NOP56* induce RNA foci and sequester SRSF2. We thus concluded that hexanucleotide repeat expansions are considered to cause SCA by a toxic RNA gain-of-function mechanism, and we name this unique SCA as SCA36. Haplotype analysis indicates that hexanucleotide expansions are derived from a common ancestor. The prevalence of SCA36 was estimated at 3.6% in the SCA cohort in Chugoku district, suggesting that prevalence of SCA36 may be geographically limited to the western part of Japan and is rare even in Japanese SCAs.

Expansion of tandem nucleotide repeats in different regions of respective genes (most often the triplets CAG and CTG) has been shown to cause a number of inherited diseases over the past decades. An expansion in the coding region of a gene causes a gain of toxic function and/or reduces the normal function of the corresponding protein at the protein level. RNA-mediated noncoding repeat expansions have also been identified as causing eight other neuromuscular disorders: DM1, DM2 (MIM 602668), fragile X tremor/ataxia syndrome (FXTAS [MIM 300623]), Huntington disease-like 2 (HDL2 [MIM 606438]), SCA8, SCA10 (MIM 603516), SCA12 (MIM 604326), and SCA31 (MIM 117210).[26] The repeat numbers in affected alleles of SCA36 are among the largest seen in this group of diseases (i.e., there are thousands of repeats). Moreover, SCA36 is not merely a nontriplet repeat expansion disorder similar to SCA10, DM2, and SCA31, but is now proven to be a human disease caused by a large hexanucleotide repeat expansion. In addition, no or only weak anticipation has been reported for noncoding repeat expansion in SCA, whereas clear anticipation has been reported for most polyglutamine expansions in SCA.[2] As such, absence of anticipation in SCA36 is in accord with previous studies

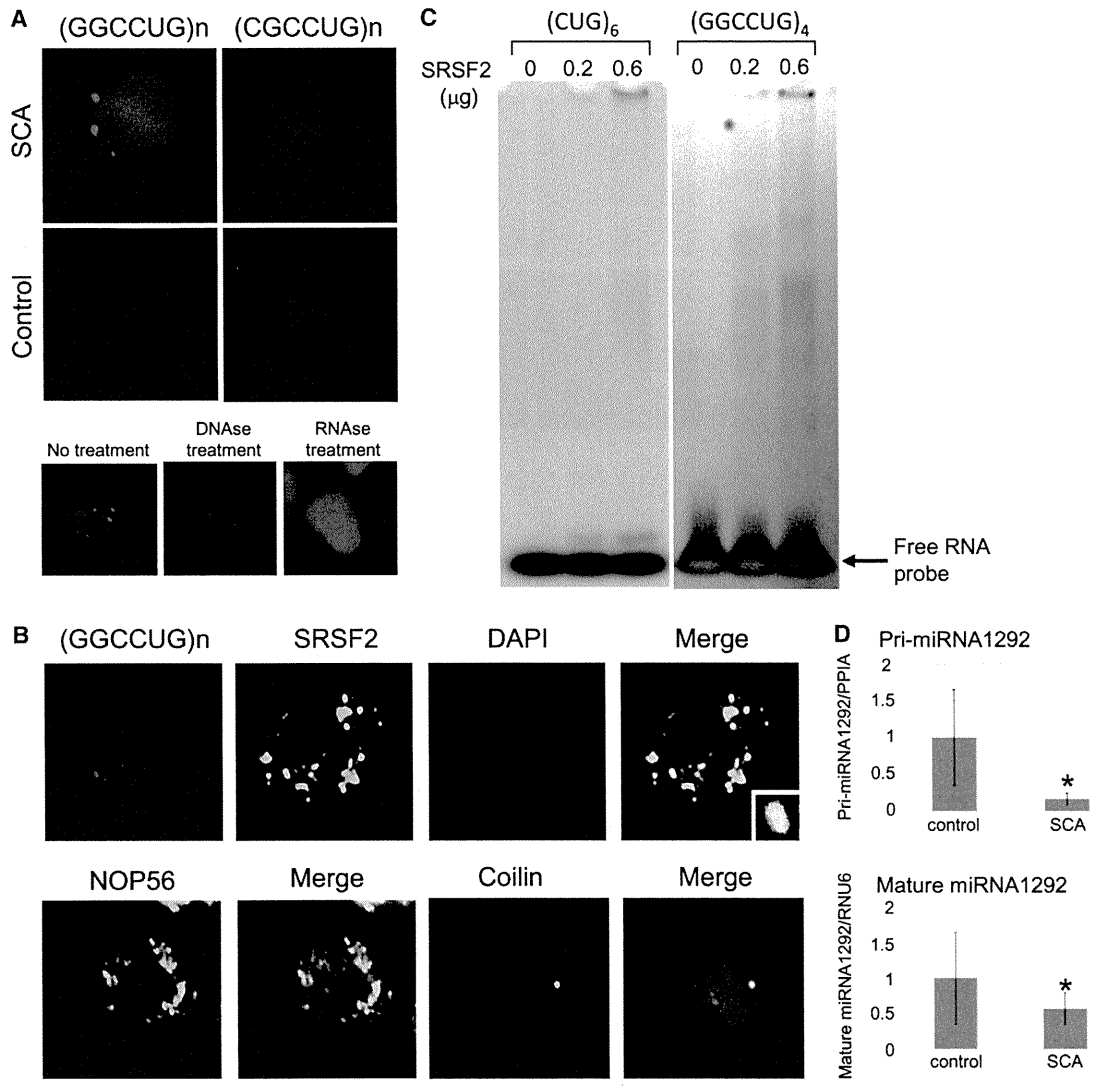


Figure 6. RNA Foci Formation and Decreased Transcription of *MIR1292*

(A) Cells were fixed on coverslips and then hybridized with solutions containing either a Cy3-labeled $C(CAGGCC)_2CAG$ or $G(CAGGCG)_2CAG$ oligonucleotide probe (1 ng/μl). For controls, the cells were treated with 1000 U/ml DNase or 100 μg/ml RNase for 1 hr at 37°C prior to hybridization, as indicated. After a wash step, coverslips were placed on the slides in the presence of ProLong Gold with DAPI mounting media (Molecular Probes, Tokyo, Japan) and photographed with a fluorescence microscope. The upper panels indicate LCLs from an SCA case and a control hybridized with $C(CAGGCC)_2CAG$ (left) or $G(CAGGCG)_2CAG$ (right). Red and blue signals represent RNA foci and the nucleus (DAPI staining), respectively. Similar RNA foci formation was confirmed in LCLs from another index case. The lower panels show RNA foci in SCA LCLs treated with DNase or RNase.

(B) Double staining was performed with the probe for $(GGCCUG)_n$ (red) and anti-SRSF2, NOP56, or coilin antibody (green).

(C) Gel-shift assays revealed specific binding of SRSF2 to $(GGCCUG)_4$ but little to $(CUG)_6$.

(D) RNA samples (10 ng) were extracted from LCLs of controls ($n = 3$) and cases ($n = 6$). MiRNAs were measured with the use of a TaqMan probe for precursor (Pri-) and mature *MIR1292*. The data indicate the mean ± SD, relative to the levels of *PP1A* or *RNU6*. *: $p < 0.05$.

on SCAs with noncoding repeat expansions. The common hallmark in these noncoding repeat expansion disorders is transcribed repeat nuclear accumulations with respective repeat RNA-binding proteins, which are considered to primarily trigger and develop the disease at the RNA level. However, multiple different mechanisms are likely to be involved in each disorder. There are at least two possible explanations for the motor neuron involvement of SCA36: gene- and tissue-specific splicing specificity of *SRSF2* and involvement of miRNA. In SCA36, there is the possibility that the adverse effect of the expansion mutation is mediated by downregulation of miRNA expression. The biochemical implication of miRNA involvement cannot be evaluated in this study, because availability of tissue samples from affected cases was limited to LCLs. Given definitive downregulation of miRNA 1292 in LCLs, we should await further study to substantiate its involvement in affected tissues. Elucidating which mechanism(s) plays a critical role in the pathogenesis will be required for determining whether cerebellar degeneration and motor neuron disease occur through a similar scenario.

In conclusion, expansion of the intronic GGCCTG hexanucleotide repeat in *NOP56* causes a unique form of SCA, SCA36, which shows not only ataxia but also motor neuron dysfunction. This characteristic disease phenotype can be explained by the combination of RNA gain of function and *MIR1292* suppression. Additional studies are required to investigate the roles of each mechanistic component in the pathogenesis of SCA36.

Supplemental Data

Supplemental Data include one figure and three tables and can be found with this article online at http://www.cell.com/AJHG/.

Acknowledgments

This work was supported mainly by grants to A.K. and partially by grants to T.M., Y.I., H.K., and K.A. We thank Norio Matsuura, Kokoro Iwasawa, and Kouji H. Harada (Kyoto University Graduate School of Medicine).

Received: February 23, 2011
Revised: May 8, 2011
Accepted: May 18, 2011
Published online: June 16, 2011

Web Resources

The URLs for data presented herein are as follows:

ESEfinder 3.0, http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home
NCBI, http://www.ncbi.nlm.nih.gov/
Target Scan Human 5.1, http://www.targetscan.org/
UCSC Genome Bioinformatics, http://genome.ucsc.edu

References

1. Harding, A.E. (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.
2. Matilla-Dueñas, A., Sánchez, I., Corral-Juan, M., Dávalos, A., Alvarez, R., and Latorre, P. (2010). Cellular and molecular pathways triggering neurodegeneration in the spinocerebellar ataxias. Cerebellum *9*, 148–166.
3. Schöls, L., Bauer, P., Schmidt, T., Schulte, T., and Riess, O. (2004). Autosomal dominant cerebellar ataxias: clinical features, genetics, and pathogenesis. Lancet Neurol. *3*, 291–304.
4. Ohta, Y., Hayashi, T., Nagai, M., Okamoto, M., Nagotani, S., Nagano, I., Ohmori, N., Takehisa, Y., Murakami, T., Shoji, M., et al. (2007). Two cases of spinocerebellar ataxia accompanied by involvement of the skeletal motor neuron system and bulbar palsy. Intern. Med. *46*, 751–755.
5. Daughters, R.S., Tuttle, D.L., Gao, W., Ikeda, Y., Moseley, M.L., Ebner, T.J., Swanson, M.S., and Ranum, L.P. (2009). RNA gain-of-function in spinocerebellar ataxia type 8. PLoS Genet. *5*, e1000600.
6. Sato, N., Amino, T., Kobayashi, K., Asakawa, S., Ishiguro, T., Tsunemi, T., Takahashi, M., Matsuura, T., Flanigan, K.M., Iwasaki, S., et al. (2009). Spinocerebellar ataxia type 31 is associated with "inserted" penta-nucleotide repeats containing (TGGAA)n. Am. J. Hum. Genet. *85*, 544–557.
7. White, M.C., Gao, R., Xu, W., Mandal, S.M., Lim, J.G., Hazra, T.K., Wakamiya, M., Edwards, S.F., Raskin, S., Teive, H.A., et al. (2010). Inactivation of hnRNP K by expanded intronic AUUCU repeat induces apoptosis via translocation of PKCdelta to mitochondria in spinocerebellar ataxia 10. PLoS Genet. *6*, e1000984.
8. Kruglyak, L., Daly, M.J., Reeve-Daly, M.P., and Lander, E.S. (1996). Parametric and nonparametric linkage analysis: a unified multipoint approach. Am. J. Hum. Genet. *58*, 1347–1363.
9. Bakalkin, G., Watanabe, H., Jezierska, J., Depoorter, C., Verschuuren-Bemelmans, C., Bazov, I., Artemenko, K.A., Yakovleva, T., Dooijes, D., Van de Warrenburg, B.P., et al. (2010). Prodynorphin mutations cause the neurodegenerative disorder spinocerebellar ataxia type 23. Am. J. Hum. Genet. *87*, 593–603.
10. Wang, J.L., Yang, X., Xia, K., Hu, Z.M., Weng, L., Jin, X., Jiang, H., Zhang, P., Shen, L., Guo, J.F., et al. (2010). TGM6 identified as a novel causative gene of spinocerebellar ataxias using exome sequencing. Brain *133*, 3510–3518.
11. Cagnoli, C., Michielotto, C., Matsuura, T., Ashizawa, T., Margolis, R.L., Holmes, S.E., Gellera, C., Migone, N., and Brusco, A. (2004). Detection of large pathogenic expansions in FRDA1, SCA10, and SCA12 genes using a simple fluorescent repeat-primed PCR assay. J. Mol. Diagn. *6*, 96–100.
12. Matsuura, T., and Ashizawa, T. (2002). Polymerase chain reaction amplification of expanded ATTCT repeat in spinocerebellar ataxia type 10. Ann. Neurol. *51*, 271–272.
13. Warner, J.P., Barron, L.H., Goudie, D., Kelly, K., Dow, D., Fitzpatrick, D.R., and Brock, D.J. (1996). A general method for the detection of large CAG repeat expansions by fluorescent PCR. J. Med. Genet. *33*, 1022–1026.
14. Hashikata, H., Liu, W., Inoue, K., Mineharu, Y., Yamada, S., Nanayakkara, S., Matsuura, N., Hitomi, T., Takagi, Y., Hashimoto, N., et al. (2010). Confirmation of an association of single-nucleotide polymorphism rs1333040 on 9p21 with familial and sporadic intracranial aneurysms in Japanese patients. Stroke *41*, 1138–1144.
15. Wahl, M.C., Will, C.L., and Lührmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. Cell *136*, 701–718.
16. Lechertier, T., Grob, A., Hernandez-Verdun, D., and Roussel, P. (2009). Fibrillarin and Nop56 interact before being co-assembled in box C/D snoRNPs. Exp. Cell Res. *315*, 928–942.
17. Liquori, C.L., Ricker, K., Moseley, M.L., Jacobsen, J.F., Kress, W., Naylor, S.L., Day, J.W., and Ranum, L.P. (2001). Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. Science *293*, 864–867.
18. Taneja, K.L., McCurrach, M., Schalling, M., Housman, D., and Singer, R.H. (1995). Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. J. Cell Biol. *128*, 995–1002.
19. Winter, J., Jung, S., Keller, S., Gregory, R.I., and Diederichs, S. (2009). Many roads to maturity: microRNA biogenesis pathways and their regulation. Nat. Cell Biol. *11*, 228–234.
20. Zhao, Y., and Srivastava, D. (2007). A developmental view of microRNA function. Trends Biochem. Sci. *32*, 189–197.
21. Eacker, S.M., Dawson, T.M., and Dawson, V.L. (2009). Understanding microRNAs in neurodegeneration. Nat. Rev. Neurosci. *10*, 837–841.

22. Hébert, S.S., and De Strooper, B. (2009). Alterations of the microRNA network cause neurodegenerative disease. Trends Neurosci. *32*, 199–206.
23. Tsuzuki, K., and Ozawa, S. (2005). Glutamate Receptors. Encyclopedia of life sciences. John Wiley and Sons, Ltd., http://onlinelibrary.com/doi/10.1038/npg.els.0005056.
24. Nutini, M., Frazzini, V., Marini, C., Spalloni, A., Sensi, S.L., and Longone, P. (2011). Zinc pre-treatment enhances NMDAR-mediated excitotoxicity in cultured cortical neurons from SOD1(G93A) mouse, a model of amyotrophic lateral sclerosis. Neuropharmacology *60*, 1200–1208.
25. Sanelli, T., Ge, W., Leystra-Lantz, C., and Strong, M.J. (2007). Calcium mediated excitotoxicity in neurofilament aggregate-bearing neurons in vitro is NMDA receptor dependant. J. Neurol. Sci. *256*, 39–51.
26. Todd, P.K., and Paulson, H.L. (2010). RNA-mediated neurodegeneration in repeat expansion disorders. Ann. Neurol. *67*, 291–300.

Comparisons of acoustic function in SCA31 and other forms of ataxias

Yoshio Ikeda[1], Makiko Nagai[1], Tomoko Kurata[1], Toru Yamashita[1], Yasuyuki Ohta[1], Shoko Nagotani[1], Kentaro Deguchi[1], Yasushi Takehisa[1], Yoshihiko Shiro[2], Tohru Matsuura[1], Koji Abe[1]

[1]Department of Neurology, Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University, Japan, [2]Department of Neurology, Kobe City Medical Center West Hospital, Japan

Objective: To investigate whether acoustic impairment can be one of the characteristic extracerebellar symptoms in sporadic and hereditary ataxias including spinocerebellar ataxia type 31 (SCA31).
Methods: We investigated genotypes of dominant ataxia families, and determined a frequency of each form in our cohort of 154 families. Acoustic function in the groups of various forms of ataxia with multiple system atrophy of cerebellar predominance (MSA-C), cortical cerebellar atrophy (CCA), and hereditary ataxias including SCA31 was evaluated by using audiogram and brainstem auditory evoked potentials (BAEPs).
Results: Genetic analysis of dominant ataxia families revealed that a frequency of SCA31 in our cohort was fewer than that reported from other areas of Japan, indicating that SCA31 is not widely distributed throughout Japan. Results of audiogram showed no significant difference of hearing levels among ataxic groups, and those of BAEPs did not support inner ear dysfunction in SCA31 in which hearing loss had initially been suggested as one of its characteristic symptoms.
Conclusion: This study suggests that acoustic impairment is neither specific to SCA31, MSA-C and CCA nor useful in making a differential diagnosis among them.

Keywords: Ataxia, Audiogram, Brainstem auditory evoked potential, Hearing level, SCA31

Introduction

Both sporadic and hereditary forms of ataxias are a heterogeneous group of neurodegenerative disorders characterized by gait and truncal instability, limb incoordination, and dysarthria.[1] Extracerebellar symptoms in ataxias sometimes help make a diagnosis of the specific form of the disease. Progressive deterioration of hearing acuity is one of the characteristic symptoms in rare hereditary ataxic disorders such as mitochondrial encephalomyopathy, Refsum disease, Usher syndrome, and so on. On the other hand, acoustic involvement in other forms of ataxia has not been fully investigated to date. We studied and demonstrated the results of audiogram and brainstem auditory evoked potentials (BAEPs) to clarify whether acoustic impairment could be a frequent neurological deficit in sporadic and hereditary forms of ataxias including recently identified spinocerebellar ataxia type 31 (SCA31) [MIM #117210], previously known as chromosome 16q22.1-linked autosomal dominant cerebellar ataxia (16q-ADCA), in which hearing loss had been suggested as one of its characteristic symptoms.[2–4] We also investigated genotypes of dominant ataxia families, and determined a frequency of each form in our cohort of 154 families.

Correspondence to: Yoshio Ikeda, Department of Neurology, Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University, 2-5-1 Shikatacho, Okayama 700-8558, Japan. Email: ikeda006@cc.okayama-u.ac.jp

Materials and Methods

We recruited both sporadic and dominant ataxia cases for this study. Individuals with multiple system atrophy of cerebellar predominance (MSA-C) fulfilled the diagnostic criteria of probable MSA according to the second consensus statement.[5] Sporadic cortical cerebellar atrophy (CCA) was defined as an adult onset (>20 years old), progressive pure cerebellar ataxia with little or no overt involvement of other central nervous system parts.[6] CCA cases did not fulfill the criteria for MSA, and their brain magnetic resonance imaging or computed tomography showed confined cerebellar atrophy. Ataxia cases with established symptomatic factors that can cause cerebellar ataxia (alcoholism, cerebrovascular, metabolic, neoplastic, autoimmune or inflammatory diseases, thiamine or vitamin E deficiency, and chronic intake of antiepileptic drugs) were excluded from this study.

DOI 10.1179/174313281oY.0000000011

Neurological Research ner2346.3d 30/9/10 22:50:57
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

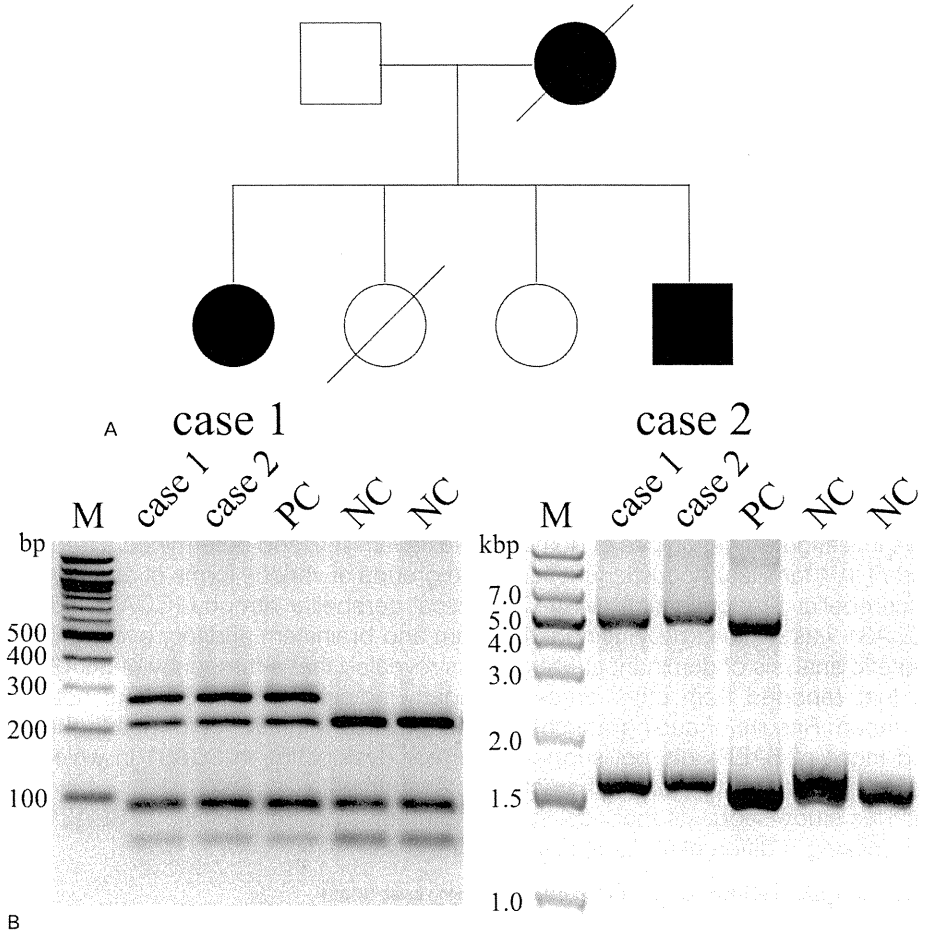


Figure 1 Pedigree and genetic analysis of an SCA31 family (F1). (A) A pedigree of an SCA31 family (F1) evaluated in this study. Symbols for individuals affected by ataxia are shaded. A diagonal line denotes individuals who are deceased. Squares and circles represent males and females, respectively. (B) Left panel; PCR-based restriction fragment length polymorphism analysis for −16 C-to-T change in puratrophin-1 gene. A region containing a −16 C-to-T change is PCR-amplified, digested by EcoNI, and analyzed on 3% agarose gel. Both cases from F1 show an undigested fragment due to a −16 C-to-T change. Right panel; long PCR analysis for the SCA31-associated pentanucleotide repeat insertion. A region containing an insertion is PCR-amplified, and analyzed on 1% agarose gel. Both cases from F1 show an upper band at a size around 5.0 kb that corresponds to a PCR product with an insertion. Lower bands appeared at a size around 1.5 kb correspond to normal alleles with polymorphic size changes. Lane M: DNA size marker; PC: SCA31-positive control; NC: normal control.

After obtaining informed consent, genomic DNAs of cases in dominantly inherited ataxia families were extracted from peripheral blood leukocytes, and screened for CAG triplet dynamic mutations in SCA types 1–3, 6–8, and dentatorubral and pallidoluysian atrophy (DRPLA) using a method described in our previous reports.[7,8] A diagnosis of SCA31/16q-ADCA is genetically investigated by confirming both an insertion of long penta-nucleotide repeat stretch containing (TGGAA)n by recently developed long polymerase chain reaction (PCR) analysis, and a C-to-T substitution at 5′ UTR of the puratrophin-1 gene by conventional PCR-based restriction fragment length polymorphism analysis as reported in elsewhere.[2,3]

Acoustic function in cases with ataxia was evaluated using audiogram and BAEPs. An audiogram was performed in 32 MSA-C cases, 25 CCA cases, four cases with SCA31 from three families, one SCA2, three SCA3/Machado–Joseph disease (MJD), two SCA6 and three DRPLA cases from independent families. Ages at audiogram examined are matched among patients in MSA-C, CCA, and SCA31 groups ($P<0.05$). BAEPs were recorded in two cases from an SCA31 family (Fig. 1A). Each subject was interviewed for occupational experience and history of otological diseases to rule out the presence of any pathological conditions that will affect the auditory function. Subjects with noise-induced hearing loss caused by occupational noise exposure or past medical history of otological disorders such as infection in auditory system or administration of drugs that can affect acoustic function were excluded from this study. The pure-tone air-conduction hearing thresholds were collected for each ear at frequencies of 125, 250, 500, 1000, 2000, 4000, and 8000 Hz using a duly calibrated diagnostic audiometer measured by trained audiometric technicians.

The pure-tone average (PTA) of air-conduction hearing thresholds was measured as the average of hearing thresholds at the four frequencies (500, 1000, 2000, and 4000 Hz).[9] BAEPs were recorded from both earlobes (A1 and A2) against Cz, following condensant-click stimulation at 65 dB above the individual click hearing threshold and at 10 Hz. One thousand responses were averaged twice to confirm reproducibility, and the latencies of the first five positive peaks and interpeak latencies were measured.

Statistical analysis of mean values of each group was evaluated by Kruskal–Wallis test. $P<0.05$ were considered to be statistically significant.

Results

We genetically analyzed in total 154 dominantly inherited ataxia families in this study. Of these families, genotypes of probands from 115 families are confirmed with respective mutations. Frequency of genotypes in dominantly inherited ataxia families is shown in Table 1. SCA6 was the most common form of dominant ataxia (35.1%), and DRPLA was the second (20.8%). Frequency of a pure form of cerebellar ataxia, which is also known as autosomal dominant cerebellar ataxia type III (ADCA III) proposed by Harding,[10,11] and in which SCA6, SCA8, and SCA31 are considered to be forms of this category, was 57.1% (88 families). Frequency of a non-pure cerebellar ataxia, which is also known as ADCA I exhibiting various extracerebellar neurological features,[10,11] and in which SCA1, SCA2, SCA3/MJD, and DRPLA are considered to be forms of this category, was 42.9% (66 families). We could not find any mutations in 39 dominant ataxia families (25.4%), of which 22 families were pure cerebellar form (ADCA III), and 17 were non-pure cerebellar form (ADCA I).

We confirmed five dominant ataxia families showing both an SCA31-associated pentanucleotide repeat insertion and a C-to-T change in puratrophin-1 gene (Fig. 1B). The lengths of the pentanucleotide repeat insertion were subsequently determined by HaeIII digestion of the long PCR products and agarose-gel electrophoresis according to a method described by Sato *et al.* (data not shown).[2] Two cases from an SCA31 family (family F1; Fig. 1A) exhibited the same repeat length with 3.5 kb, and other four SCA31 cases from independent families exhibited 3.4 kb in length, respectively.

Table 1 Frequency of genotypes in dominantly inherited ataxia families

Genotype	No. of families (%)
SCA1	1 (0.6)
SCA2	4 (2.6)
SCA3/MJD	12 (7.8)
SCA6	54 (35.1)
SCA7	0 (0)
SCA8	7 (4.5)
SCA31	5 (3.2)
DRPLA	32 (20.8)
Unknown	39 (25.4)
Total	154 (100)

Note: SCA=spinocerebellar ataxia; MJD=Machado–Joseph disease; DRPLA=dentatorubral and pallidoluysian atrophy.

Clinical data and hearing levels evaluated by audiogram in respective ataxia group are shown in Table 2. Average duration of illness at audiogram examined in MSA-C group (3.6 years) is apparently shorter than that in CCA or SCA31 group (8.0 and 9.7 years, respectively), indicating that MSA-C patients tend to be diagnosed earlier after onset probably because of faster disease progression than CCA or SCA31. In the SCA group, average ages at onset and audiogram examined are almost 10 years earlier, and average duration of illness at audiogram examined is longer than those of other ataxia groups.

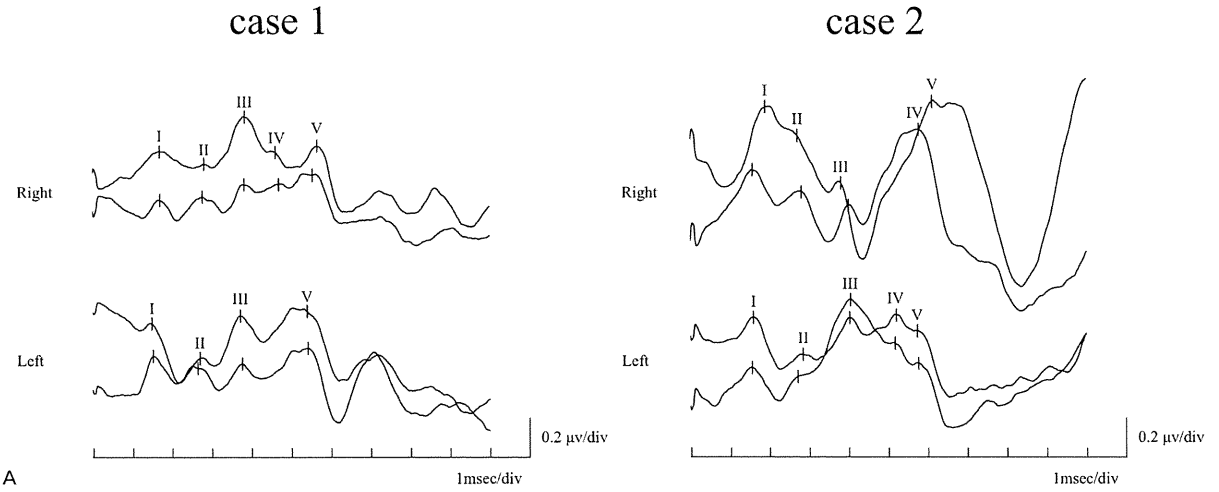
Ages at onset with truncal instability in cases 1 and 2 in an SCA31 family (family F1; Fig. 1A) were 50 and 48 years old, respectively. Neurological examination of both cases revealed quite similar positive findings for truncal and gait ataxia, limb incoordination, ataxic dysarthria, and impaired smooth pursuit, but negative for gaze-evoked nystagmus, ophthalmoplegia, pyramidal tract signs, sensory disturbance, parkinsonism, cognitive impairment, and urinary disturbance. These neurological findings are consistent with the previously reported clinical characteristics that SCA31 is relatively pure cerebellar ataxia.[3,4,12–21] All SCA31 cases who were clinically evaluated did not complain acoustic impairment that can affect verbal communication in daily life.

The average hearing levels (PTA) are not significantly different between groups in MSA-C (29.8/30.3 dB at right/left side, respectively) and CCA (25.6/29.2 dB). The average hearing levels in four SCA31 cases, examined at 64.3 years old, are 26.5/25.5 dB, with no significant difference from either MSA-C or CCA (Table 2). BAEPs were evaluated in two patients from an SCA31 family (F1) to examine their inner ear function. The I–V waves in these two cases were clearly confirmed, and latencies to each wave-peak and the I–III, III–V, and I–V interpeak latencies were not obviously delayed (Fig. 2).

Discussion

This is the first report showing that acoustic impairment cannot be a specific neurological deficit in MSA-C or CCA. This study also clarified that acoustic impairment is not frequently observed in SCA31 cases with recently identified insertion of long penta-nucleotide repeat stretch containing

Neurological Research ner2346.3d 30/9/10 22:51:01
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)



A

B

case ID	Stim. Side	Latency (msec) I	II	III	IV	V	Interpeak latency (msec) I-III	III-V	I-V
F1-case 1	Right	1.69	2.81	3.84	4.61	5.66	2.15	1.82	3.97
	Left	1.49	2.70	3.72	N/A	5.40	2.23	1.68	3.91
F1-case 2	Right	1.74	2.76	3.91	N/A	6.00	2.17	2.05	4.22
	Left	1.58	2.79	4.04	5.20	5.76	2.47	1.72	4.18

Figure 2 Results of brainstem auditory evoked potentials of an SCA31 family (F1). (A) BAEPs of cases 1 and 2 from an SCA31 family (F1). Upper BAEPs of each case indicate responses (Cz-A2) elicited by right ear stimulation, and lower BAEPs indicate responses (Cz-A1) elicited by left ear stimulation. The BAEPs clearly demonstrate I wave formation with normal latency to their peaks. (B) Latencies of the five peaks and interpeak latencies in cases 1 and 2 from family F1.

(TGGAA)n. The average hearing level in Japanese controls examined at ages 65–69 years old (n=386) was reported to be 33.4 dB in a previous report.[22] Average ages at audiogram examined and average hearing levels in MSA-C, CCA and SCA31 groups are comparable to those in the controls, indicating that acoustic function in these ataxic patients is not impaired by their disease process.

Table 2 Summary of clinical data and hearing levels evaluated by audiogram in each ataxia group

Group	No. of cases	Age at AG (years, mean)	Age at onset (years, mean)	Duration (years, mean)	PTA (dB, mean) Right	Left
MSA-C	32	62.6	58.8	3.6	29.8	30.3
CCA	25	64.2	55.4	8.0	25.6	29.2
SCA31	4	64.3	51.7	9.7	26.5	25.5
Other SCAs	9	53.1	42.0	11.1	27.1	25.0

Individual data in SCA31 and other dominantly inherited ataxias

SCA type	Case ID	Age at AG (years)	Age at onset (years)	Duration (years)	PTA (dB) Right	Left
SCA31	F1-case 1	65	50	15	28	28
	F1-case 2	55	48	7	11	10
	F2-case 1	73	N/A	N/A	36	36
	F3-case 1	64	57	7	31	28
SCA2	F1-case 1	48	45	3	11	6
SCA3	F1-case 1	71	52	19	39	33
	F2-case 1	53	41	12	18	20
	F3-case 1	38	36	2	30	15
SCA6	F1-case 1	72	63	9	31	44
	F2-case 1	48	34	14	20	24
DRPLA	F1-case 1	57	40	17	56	36
	F2-case 1	47	36	11	19	16
	F3-case 1	44	31	13	20	31

Note: Age at AG=age at audiogram examined; duration=duration of illness at audiogram examined; PTA=pure-tone average; SCA=spinocerebellar ataxia; other SCAs=SCAs other than SCA31; F1–F3 indicate family numbers in respective forms of hereditary ataxias; DRPLA=dentatorubral and pallidoluysian atrophy.

Neurological Research ner2346.3d 30/9/10 22:51:02
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

SCA31 is a dominant form of hereditary ataxia frequently reported from Japan with a significant founder effect.[2,3,12,14,17–19,23] A C-to-T transition at position −16 nucleotide upstream of translation initiation codon of puratrophin-1 gene was initially reported to be a potential genetic defect responsible for SCA31. Although >99% Japanese SCA31 cases have this nucleotide change, two SCA31 families in which this single-nucleotide change was not co-segregated with ataxia were subsequently found.[3,18,23] These findings indicate that −16 C-to-T change in puratrophin-1 gene is a rare polymorphism in strong linkage disequilibrium with SCA31 but not a disease-causing mutation. Recently, an inserted sequence that consists of complex penta-nucleotide repeats containing (TGGAA)n was found to be the most likely candidate for the SCA31 mutation. Both a −16 C-to-T change in puratrophin-1 and an insertion of long penta-nucleotide repeat stretch containing (TGGAA)n were confirmed in our five SCA31 families.

Genetic analysis of dominantly inherited ataxia families revealed that the frequency of SCA31 in our cohort was 3.2% (five families) as the fifth most common form. We mainly investigated families living in the south-eastern part of Chugoku District, which is located on western part of the main island of Japan. The frequency of SCA31 in our study is considered to be fewer than that reported from other areas of Japan, because in these former studies SCA31 was the first to the fourth most frequent form of dominant ataxia at a frequency of 8.2–42.7%.[12–21] So far, SCA31 was found only in families of Japanese descent and a single Korean family, but not in a European population,[24–27] and in Japan it is reported to be prevalent at a frequency similar to another pure cerebellar form, SCA6. In our study, the frequency of SCA6 (35.1%) is almost 11-fold higher than that of SCA31, indicating that SCA31 is not widely distributed throughout Japan.

The previous study showed that puratrophin-1 is expressed in a variety of organs including brain. It is described that puratrophin-1 is also expressed in epithelial hair cells in the cochlea, and this is potentially associated with relatively high frequency of hearing impairment in SCA31 cases.[3,4] Ishikawa *et al.* reported that 42.9% of SCA31 families had hearing impairment.[3] Owada *et al.* reported that two SCA31 patients from a single family showed hearing impairment in audiograms and poor I wave formation in their BAEPs.[4] However, our SCA31 patients do not show severe hearing impairment with comparable PTA values to other groups, and the results of BAEPs in two SCA31 patients do not support their inner ear impairment even in a patient with longer disease duration (15 years). It is also obvious that normal findings of interpeak latencies in BAEPs are not suggestive of brainstem involvement, indicating that SCA31 is a predominantly cerebellar ataxia. In conclusion, the present study confirmed that acoustic impairment is neither specific to SCA31, MSA-C and CCA nor useful in making a differential diagnosis among them.

Acknowledgements

We would like to thank the patients and doctors who participated in this work. This study was supported in part by Grants-in-Aid for Scientific Research (C) 21591084 (to YI), Scientific Research on Innovative Areas 2020078 (to TM), and Scientific Research (B) 21390267 (to KA) from the Ministry of Education, Culture, Sports, Science and Technology, Japan, as well as Grants-in-Aid from the Research Committees (Y. Itoyama, T. Imai, I. Nakano, G. Sobue, and M. Nishizawa), the Ministry of Health, Labour and Welfare, Japan (to YI and KA).

References

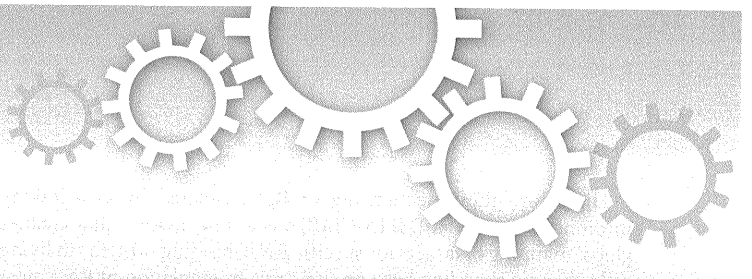
1 Schols L, Bauer P, Schmidt T, Schulte T, Riess O. Autosomal dominant cerebellar ataxias: clinical features, genetics, and pathogenesis. Lancet Neurol 2004;3:291–304.
2 Sato N, Amino T, Kobayashi K, Asakawa S, Ishiguro T, Ysunemi T, *et al.* Spinocerebellar ataxia type 31 is associated with 'inserted' penta-nucleotide repeats containing (TGGAA)n. Am J Hum Genet 2009;85:544–57.
3 Ishikawa K, Toru S, Tsunemi T, Li M, Kobayashi K, Yokota T, *et al.* An autosomal dominant cerebellar ataxia linked to chromosome 16q22.1 is associated with a single-nucleotide substitution in the 5′ untranslated region of the gene encoding a protein with spectrin repeat and Rho guanine-nucleotide exchange-factor domains. Am J Hum Genet 2005;77:280–96.
4 Owada K, Ishikawa K, Toru S, Ishida G, Gomyoda M, Tao O, *et al.* A clinical, genetic, and neuropathologic study in a family with 16q-linked ADCA type III. Neurology 2005;65:629–32.
5 Gilman S, Wenning GK, Low PA, Brooks DJ, Mathias CJ, Trojanowski JQ, *et al.* Second consensus statement on the diagnosis of multiple system atrophy. Neurology 2008;71:670–6.
6 Abele M, Minnerop M, Urbach H, Specht K, Klockgether T. Sporadic adult onset ataxia of unknown etiology : a clinical, electrophysiological and imaging study. J Neurol 2007;254:1384–9.
7 Ikeda Y, Dalton JC, Moseley ML, Gardner KL, Bird TD, Ashizawa T, *et al.* Spinocerebellar ataxia type 8: molecular genetic comparisons and haplotype analysis of 37 families with ataxia. Am J Hum Genet 2004;75:3–16.
8 Ikeda Y, Shizuka M, Watanabe M, Okamoto K, Shoji M. Molecular and clinical analyses of spinocerebellar ataxia type 8 in Japan. Neurology 2000;54:950–5.
9 Chia EM, Wang JJ, Rochtchina E, Cumming RR, Newall P, Mitchell P. Hearing impairment and health-related quality of life: the Blue Mountains Hearing Study. Ear Hear 2007;28:187–95.
10 Harding AE. 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 1982;105:1–28.
11 Harding AE. Clinical features and classification of inherited ataxias. Adv Neurol 1993;61:1–14.
12 Hirano R, Takashima H, Okubo R, Okamoto Y, Maki Y, Ishida S, *et al.* Clinical and genetic characterization of 16q-linked autosomal dominant spinocerebellar ataxia in South Kyushu, Japan. J Hum Genet 2009;54:377–81.
13 Onodera Y, Aoki M, Mizuno H, Warita H, Shiga Y, Itoyama Y. Clinical features of chromosome 16q22.1 linked autosomal dominant cerebellar ataxia in Japanese. Neurology 2006;67:1300–2.
14 Shimizu Y, Yoshida K, Okano T, Ohara S, Hashimoto T, Fukushima Y, *et al.* Regional features of autosomal-dominant

Neurological Research ner2346.3d 30/9/10 22:51:03
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

cerebellar ataxia in Nagano: clinical and molecular genetic analysis of 86 families. J Hum Genet 2004;49:610–6.
15 Yoshida K, Shimizu Y, Morita H, Okano T, Sakai H, Ohata T, *et al.* Severity and progression rate of cerebellar ataxia in 16q-linked autosomal dominant cerebellar ataxia (16q-ADCA) in the endemic Nagano Area of Japan. Cerebellum 2009;8:46–51.
16 Ouyang Y, Sakoe K, Shimazaki H, Namekawa M, Ogawa T, Ando Y, *et al.* 16q-linked autosomal dominant cerebellar ataxia: a clinical and genetic study. J Neurol Sci 2006;247:180–6.
17 Nozaki H, Ikeuchi T, Kawakami A, Kimura A, Koide R, Tsuchiya M, *et al.* Clinical and genetic characterizations of 16q-linked autosomal dominant spinocerebellar ataxia (AD-SCA) and frequency analysis of AD-SCA in the Japanese population. Mov Disord 2007;22:857–62.
18 Ohata T, Yoshida K, Sakai H, Hamanoue H, Mizuguchi T, Shimizu Y, *et al.* A −16C>T substitution in the 5′ UTR of the puratrophin-1 gene is prevalent in autosomal dominant cerebellar ataxia in Nagano. J Hum Genet 2006;51:461–6.
19 Hirano R, Takashima H, Okubo R, Tajima K, Okamoto Y, Ishida S, *et al.* Fine mapping of 16q-linked autosomal dominant cerebellar ataxia type III in Japanese families. Neurogenetics 2004;5:215–21.
20 Hayashi M, Adachi Y, Mori M, Nakano T, Nakashima K. Clinical and genetic epidemiological study of 16q22.1-linked autosomal dominant cerebellar ataxia in western Japan. Acta Neurol Scand 2007;116:123–7.
21 Basri R, Yabe I, Soma H, Sasaki H. Spectrum and prevalence of autosomal dominant spinocerebellar ataxia in Hokkaido, the northern island of Japan: a study of 113 Japanese families. J Hum Genet 2007;52:848–55.
22 Yagi M, Kawabata I, Sato T, Toriyama M, Yamashita K, Makishima K, *et al.* [Hearing acuity in the elderly in Japan]. Nippon Jibiinkoka Gakkai Kaiho 1996;99:869–74.
23 Amino T, Ishikawa K, Toru S, Ishiguro T, Sato N, Tsunemi T, *et al.* Redefining the disease locus of 16q22.1-linked autosomal dominant cerebellar ataxia. J Hum Genet 2007;52:643–9.
24 Zhou Y, Song X, Yi J, Jiang H, Wang J, Liao S, *et al.* [Study on the single-nucleotide substitution (c. −16C to T) of the PURATROPHIN-1 gene in Chinese patients with spinocerebellar ataxia]. Zhonghua Yi Xue Yi Chuan Xue Za Zhi 2008;25:646–8.
25 Cagnoli C, Brussino A, Di Gregorio E, Brusco A, Stevanin G, Durr A, *et al.* The (−16C>T) substitution in the PLEKHG4 gene is not present among European ADCA patients. Mov Disord 2007;22:752–3.
26 Wieczorek S, Arning L, Alheite I, Epplen JT. Mutations of the puratrophin-1 (PLEKHG4) gene on chromosome 16q22.1 are not a common genetic cause of cerebellar ataxia in a European population. J Hum Genet 2006;51:363–7.
27 Lee PH, Park HY, Jeong SY, Hong JH, Kim HJ. 16q-linked autosomal dominant cerebellar ataxia in a Korean family. Eur J Neurol 2007;14:e16–7.

Neurological Research ner2346.3d 30/9/10 22:51:03
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

SCIENTIFIC REPORTS



OPEN

CUGBP1 and MBNL1 preferentially bind to 3′ UTRs and facilitate mRNA decay

SUBJECT AREAS:
CELLULAR NEUROSCIENCE
MOTOR SYSTEM
GENE REGULATION
TRANSCRIPTOME

Received
25 August 2011

Accepted
8 December 2011

Published
4 January 2012

Correspondence and requests for materials should be addressed to K.O. (ohnok@med.nagoya-u.ac.jp)

* These authors contributed equally to this work.

Akio Masuda[1], Henriette Skovgaard Andersen[2]*, Thomas Koed Doktor[2]*, Takaaki Okamoto[1], Mikako Ito[1], Brage Storstein Andresen[2] & Kinji Ohno[1]

[1]Division of Neurogenetics, Center for Neurological Diseases and Cancer, Nagoya University Graduate School of Medicine, Nagoya, Japan, [2]Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense M, Denmark.

CUGBP1 and MBNL1 are developmentally regulated RNA-binding proteins that are causally associated with myotonic dystrophy type 1. We globally determined the *in vivo* RNA-binding sites of CUGBP1 and MBNL1. Interestingly, CUGBP1 and MBNL1 are both preferentially bound to 3′ UTRs. Analysis of CUGBP1- and MBNL1-bound 3′ UTRs demonstrated that both factors mediate accelerated mRNA decay and temporal profiles of expression arrays supported this. Role of CUGBP1 on accelerated mRNA decay has been previously reported, but the similar function of MBNL1 has not been reported to date. It is well established that CUGBP1 and MBNL1 regulate alternative splicing. Screening by exon array and validation by RT-PCR revealed position dependence of CUGBP1- and MBNL1-binding sites on the resulting alternative splicing pattern. This study suggests that regulation of CUGBP1 and MBNL1 is essential for accurate control of destabilization of a broad spectrum of mRNAs as well as of alternative splicing events.

In recent years, new technologies, such as microarray analysis and high throughput sequencing, have dramatically changed our knowledge on gene expression and revealed that extensive regulation takes place during posttranscriptional RNA processing[1]. Numerous RNA processing elements and regulatory RNA-binding proteins play together in a finely tuned interplay to ensure that different mRNAs are made from the primary transcript from a gene and are present in the right cell at the right time and in the correct amounts. Such complex regulation is of course vulnerable and a rapidly increasing number of human diseases are now known to be caused by misregulated RNA processing[2]. An intriguing example where this kind of disease mechanism is in operation is myotonic dystrophy type 1 (DM1), where aberrant regulation of two RNA-binding proteins, CUG-binding protein 1 (CUGBP1) and muscleblind-like 1 (MBNL1) co-operationally cause some of the disease symptoms. DM1 is the most common form of myotonic dystrophy (DM), and is caused by an expansion of CTG-repeats in the 3′ untranslated region (UTR) of the DM protein kinase gene (*DMPK*) on chromosome 19[3–5]. DM1 is a multisystemic disorder and the clinical features include myotonia, muscle degeneration, heart failure, ocular cataracts, impaired glucose tolerance, and mental retardation[6,7]. A dominant negative effect of the *DMPK* mutant allele through RNA gain-of-function has been proposed as the molecular disease mechanism. Many studies support a mechanism where toxic *DMPK* RNA with expanded CUG repeats binds to and sequesters proteins that are important for RNA metabolism including transcription, RNA transport, alternative splicing, translation, and yet unknown processes[6]. The expanded CUG repeats in the *DMPK* mRNA bind to and sequestrate MBNL1 in discrete nuclear foci, which results in depletion of functional MBNL1[8,9]. By a yet unknown mechanism the expanded CUG repeats also activate protein kinase C (PKC), which phosphorylates and stabilizes CUGBP1[10]. Thus, the expanded CUG repeats contribute to DM1 pathogenesis by causing loss of MBNL1 and gain of CUGBP1 activity[11].

Both CUGBP1 and MBNL1 regulate postnatal transitions in alternative splicing patterns during striated muscle development[9,12,13]. Representative targets of CUGBP1 splicing regulation, which are misregulated in DM1 striated muscles, include genes for cardiac troponin T (*TNNT2*)[14,15], insulin receptor (*INSR*)[16], and chloride channel 1 (*CLCN1*)[15,17]. MBNL1 contains four CCCH-type zinc fingers that recognize a YGCY motif that is indeed observed in the CUG repeat (CUGCUG)[18–21]. Mice deficient in *Mbnl1* show aberrant splicing of *Clcn1*, *Tnnt2*, and *Tnnt3*, but not *Insr*[22]. Very recently, MBNL1 was shown to regulate *BIN1* alternative splicing, and dysregulation of *BIN1* splicing in DM1 muscles was suggested to be part of the disease pathology resulting in muscle weakness[23]. Besides an important role in splicing regulation, CUGBP1 mediates mRNA decay of short-lived transcripts by interaction with GU-rich elements in the 3′ UTR[24–27]. In addition, CUGBP1 increases the translation of *CDKN1A*[28] and *Mef2a*[29]. In contrast to the multiple functionalities in posttranscriptional gene regulation of CUGBP1, MBNL1 has so far been exclusively recognized as a splicing regulatory *trans*-factor.

High-throughput sequencing of RNA isolated by crosslinking immunoprecipitation (HITS-CLIP)[30] is a new method that enables global mapping of targets for specific RNA-binding proteins in living cells, thereby shedding light on their role in regulation of RNA processing of known and unknown targets.

In the present study we performed HITS-CLIP analysis for CUGBP1 and MBNL1 on the mouse myoblast cell line C2C12 to extensively characterize their RNA-binding sites and functional roles in RNA processing. We identified position-dependence of CUGBP1/MBNL1-binding sites in regulating exon inclusion or skipping. Interestingly, we discovered that both CUGBP1 and MBNL1 preferentially bind to the 3′ UTR and destabilize target mRNAs. This points to a new important role of MBNL1 and suggests that binding to the 3′ UTRs and destabilization of mRNA are likely to be a fundamental function shared by CUGBP1 and MBNL1.

Results

Genome-wide CUGBP1/MBNL1-RNA interaction maps. In order to determine global CUGBP1/MBNL1-binding sites *in vivo*, we performed HITS-CLIP experiments using the mouse myoblast cell line, C2C12.

In C2C12 cells, CUGBP1 is constantly expressed throughout myoblast differentiation, whereas expression of MBNL1 is low in undifferentiated cells and gradually increases during differentiation (Supplementary Fig. S1), as previously described[8]. We thus performed HITS-CLIP analysis of CUGBP1 and MBNL1 using undifferentiated and differentiated C2C12 cells, respectively. We also performed CLIP of MBNL1 using undifferentiated cells in three independent experiments, but this yielded an insufficient amount of RNA-protein complexes and failed to yield cDNA libraries suitable for high-throughput sequencing. In the HITS-CLIP analysis of CUGBP1, our first experiment yielded 34,733,815 CLIP tags of 32 nt, of which 29,545,067 (85.06%) were mapped to the mm9 genome allowing at most 2 mismatches and placing reads mapping to multiple locations to a single random site. A second CLIP experiment yielded 10,079,185 CLIP tags of 36 nt, of which 8,516,256 (84.49%) were mapped. In the first MBNL1 CLIP experiment, we obtained 13,218,685 CLIP tags, of which 11,044,152 (83.55%) were mapped, while the second CLIP experiment yielded 13,474,600 CLIP tags with 11,455,886 (85.02%) tags mapped to the mm9 genome. For the analysis of binding motif and binding region annotation, we selected only reads that were aligned uniquely in the genome and removed all potential PCR duplicates by collapsing reads with an identical 5′ start into a single read. This resulted in 177,013 and 130,828 CLIP tags from the two CUGBP1 CLIP experiments, while the two MBNL1 experiments yielded 59,156 and 583,841 CLIP tags respectively.

In an effort to confirm the specificity of our CLIP experiments, we performed CLIP analysis of polypyrimidine tract-binding protein (PTB), a multifunctional RNA-binding protein, using undifferentiated mouse C2C12 cells. We identified 12,841,778 CLIP tags of which 11,184,829 (87.10%) were mapped to the mouse mm9 genome. Removal of non-uniquely aligned reads and PCR duplicates yielded 307,995 unambiguous CLIP tags.

Consensus motifs. To determine RNA-binding motifs associated with CUGBP1/MBNL1 *in vivo*, we used the motif-finding algorithm, Multiple EM for Motif Elicitation (MEME)[31]. We used SeqMonk to identify likely binding regions, and identified 1,841 CUGBP1-binding regions and 302 MBNL1-binding regions. Comparison of SeqMonk's maximum depth scores between samples indicates that binding regions in each replicated experiment are highly overlapping, while PTB binding regions did not overlap with those of the other four CLIP experiments (Supplementary Fig. S2). The lower number of identified MBNL1 regions supported by two independent experiments (Supplementary Fig. S2b) was likely due to the large difference in the number of CLIP tags in the two MBNL1 experiments. The regions demonstrate enrichment of GU-rich motifs for CUGBP1 and YGCY-containing motifs for MBNL1 (Fig. 1).

Our *in vivo* binding motifs are in accordance with previously suggested binding motifs for CUGBP1[25,32] and MBNL1[21,33–35]. We identified 1,824 PTB binding regions in the mouse genome and detected a CU-rich motif, which is essentially identical to the motif for PTB recently identified by HITS-CLIP analysis of a human cell line[36].

We also analyzed the CUGBP1 and MBNL1 motifs enriched in regions containing reads with multiple potential mapping locations (Supplementary Fig. S3), and compared them with the motifs with unique mapping (Fig. 1). Following removal of potential PCR duplicates, we observed 699,382 tags that were non-uniquely aligned in the 1st CUGBP1 CLIP experiment, 219,128 tags in the 2nd CUGBP1 CLIP experiment, 105,432 and 216,882 tags in the two MBNL1 CLIP experiments respectively and finally 851,324 tags in the PTB CLIP experiment. We observed that enriched motifs in these regions (Supplementary Fig. S3) are very similar to the CUGBP1 and MBNL1 motifs enriched in the binding regions containing uniquely aligned reads (Fig. 1), suggesting that these regions share the same properties as the uniquely aligned regions and that they may contain functional binding sites.

HITS-CLIP analysis of splicing targets. We next studied the effects of CUGBP1/MBNL1 binding on alternative splicing. CUGBP1 tags are clustered in intronic regions flanking alternative rather than

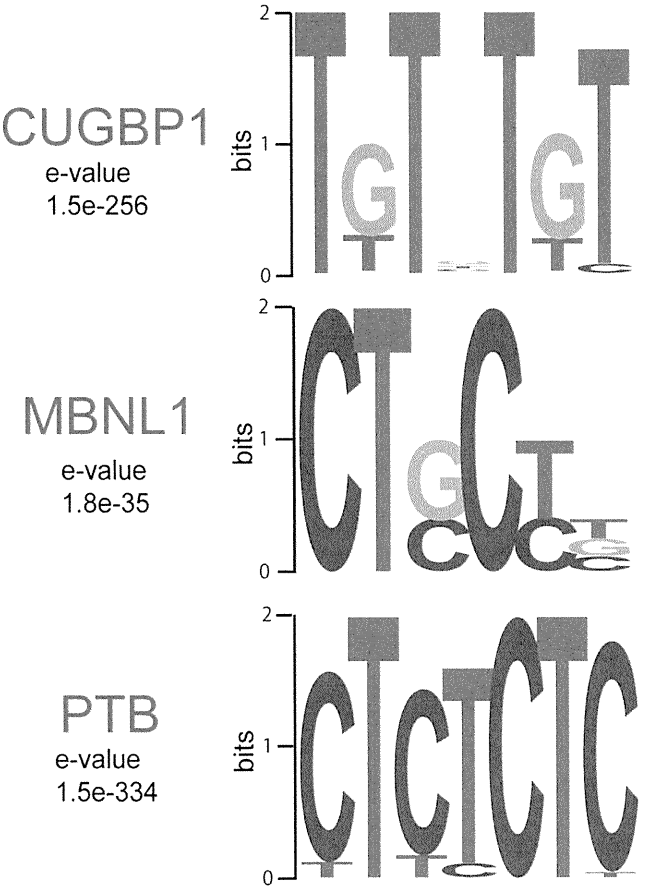


Figure 1 | Binding motif analysis. WebLogos of consensus binding motifs of CUGBP1, MBNL1, and PTB generated by the MEME motif analysis tool. The likelihood of finding the indicated motif by chance is indicated as an E-value.

constitutive exons (Fig. 2a). MBNL1 tags are similarly clustered in intronic regions flanking alternative exons, and are also enriched in alternative and constitutive exons. In order to investigate if and how CUGBP1/MBNL1 binding around splice sites regulate alternative splicing, we knocked down these factors by siRNA in undifferentiated C2C12 cells (Supplementary Fig. S4a). We analyzed alterations of splicing globally using the Affymetrix Mouse Exon 1.0 ST Array (GEO accession number, GSE29990) and identified 8 CUGBP1-responsive and 24 MBNL1-responsive exons (Supplementary Table 1, Figs. S5 and S6abc). We also analyzed 29 CUGBP1-tagged and 51 MBNL1-tagged exons/introns known to be alternatively spliced according to the ENSEMBL version *e!*61, and identified 16 CUGBP1-responsive and 21 MBNL1-responsive exons by RT-PCR (Supplementary Figs. S5 and S6abc). We made the compiled dataset C, which is comprised of the 24 CUGBP1-regulated exons (15 skipped and 9 included), as well as the compiled dataset M consisting of the 45 MBNL1-regulated exons (25 skipped and 20 included). The datasets include 1 and 9 previously identified target exons of CUGBP1 and MBNL1, respectively (Supplementary Fig. S5). In addition, 9 exons are shared between datasets C and M. Mbnl1 siRNA sufficiently suppressed MBNL1 expression up to day 3 after differentiation (Supplementary Fig. S4b), and we observed that as many as 44 of the 45 MBNL1-regulated exons in dataset M respond similarly to MBNL1 knockdown in both differentiated and undifferentiated cells (Supplementary Figs. S4 and S5).

We also made dataset M2 that includes 26 additional MBNL1-dependent cassette exons (15 skipped and 11 included) that were previously identified in skeletal muscle of MBNL1 knockout mice

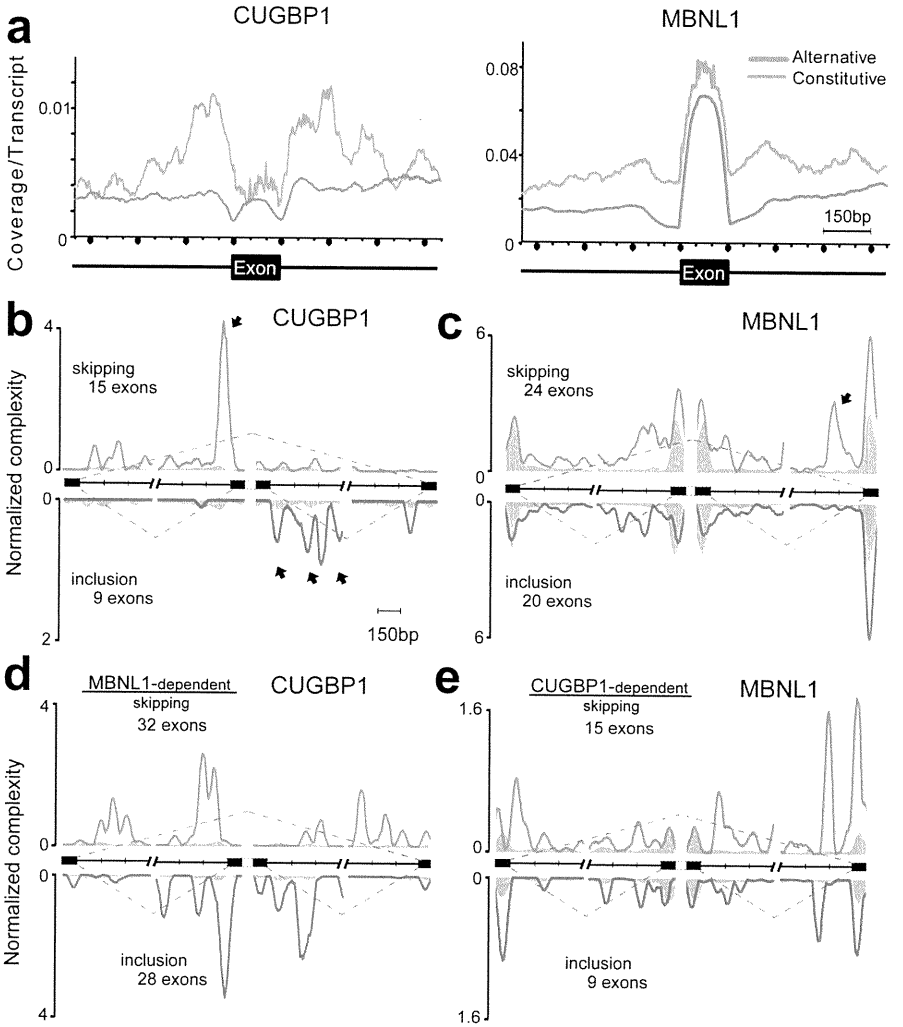


Figure 2 | Mapping of CLIP-tags on exon-intron structures. (a) Distributions of CLIP-tags on constitutively or alternatively spliced exons and the flanking intronic regions. The abscissa indicates an intron-exon-intron structure. The sizes of all the exons are normalized to 150 nucleotides. Numbers of exonic CLIP-tags are also normalized accordingly. Intronic CLIP-tags within 500 nucleotides upstream or downstream of exons are indicated. The number of CLIP-tags is normalized for the number of transcripts belonging to either category of constitutive and alternative exons. (b) Normalized complexity map of CUGBP1 at CUGBP1-dependent splice sites. Twenty-four CUGBP1-regulated splicing events in dataset C in undifferentiated C2C12 cells are compiled. (c) Normalized complexity map of MBNL1 at MBNL1-dependent splice sites. Forty-four MBNL1-regulated splicing events in differentiated C2C12 cells in dataset M are compiled. (d) Normalized complexity map of CUGBP1 at MBNL1-dependent splice sites. Sixty MBNL1-regulated splicing events in undifferentiated C2C12 cells in datasets M and M2 are compiled. (e) Normalized complexity map of MBNL1 at CUGBP1-dependent splice sites. Twenty-four CUGBP1-regulated splicing events in undifferentiated C2C12 cells in dataset C are compiled. Shaded areas represent an average of 100 sets of normalized complexity of 50 (b, c, and d) and 15 (e) randomly selected constitutive exons. Arrows indicate representative peaks that Graphs represents results of the 2nd CLIP experiments for both CUGBP1 and MBNL1. Resutls of the 1st CLIP experiments are shown in Supplmentary Fig. S7.

(Supplementary Table S2)[33], and found that 18 exons are similarly regulated by *Mbnl1* knockdown in undifferentiated C2C12 cells (Supplementary Fig. S6d and Table S2).

We combined datasets C and M into a single composite pre-mRNA and made integrated RNA maps from our HITS-CLIP reads mapped to the corresponding genomic regions as previously described for Nova[30] and PTB[36]. This showed that CUGBP1 binding to upstream intronic regions facilitates exon skipping, whereas CUGBP1 binding to downstream intronic regions promotes exon inclusion (closed arrows in Fig. 2b and Supplementary Fig. S7a). Results of the 2nd experiments are shown in Fig. 2 and those of the 1st experiments are in Supplementary Fig. S7. In contrast, although the binding sites of MBNL1 are more diffusely distributed and less abundant in regions flanking splice sites (Fig. 2c), MBNL1 binding close to the 3′ end of the downstream intron induces exon skipping (closed arrow in Fig. 2c and Supplementary Fig. S7b). The presence of a similar peak in dataset M2 (closed arrow in Supplementary Fig. S7c) further supports this observation.

We next analyzed the interaction between CUGBP1 and MBNL1 in splicing regulation. We made an RNA map of CUGBP1-binding sites in MBNL1-regulated exons from datasets M and M2 (Fig. 2d and Supplementary Fig. S7e), as well as an RNA map of MBNL1-binding sites in CUGBP1-regulated exons from dataset C (Fig. 2e and Supplementary Fig. S7f). Both RNA maps demonstrate the presence of CUGBP1 clusters in MBNL1-responsive exons and vice versa, which suggests that CUGBP1 and MBNL1 are likely to regulate alternative splicing of some of the same exons.

MBNL1 and CUGBP1 both preferentially bind to the 3′ UTR. MBNL1 has so far solely been categorized as an exon/intron-binding splicing regulatory protein[6], but to our surprise we found that the majority (55%) of MBNL1-binding regions are located in 3′ UTRs (Fig. 3a). The same pattern with preferential binding (53%) in 3′ UTRs is observed for CUGBP1, while only 2% of PTB binding regions are located in 3′ UTRs (Fig. 3a). Similarly, when HITS-CLIP tags are mapped to the size-normalized positions of all the genes in the mouse genome, CUGBP1 and MBNL1 CLIP tags, but not PTB CLIP-tags, are enriched close to the 3′ ends of genes (Fig. 3b). Additionally, 610 3′ UTRs, which constitutes 28.7% of the CUGBP1-tagged 3′ UTRs and 17.4% of the MBNL1-tagged 3′ UTRs, are shared

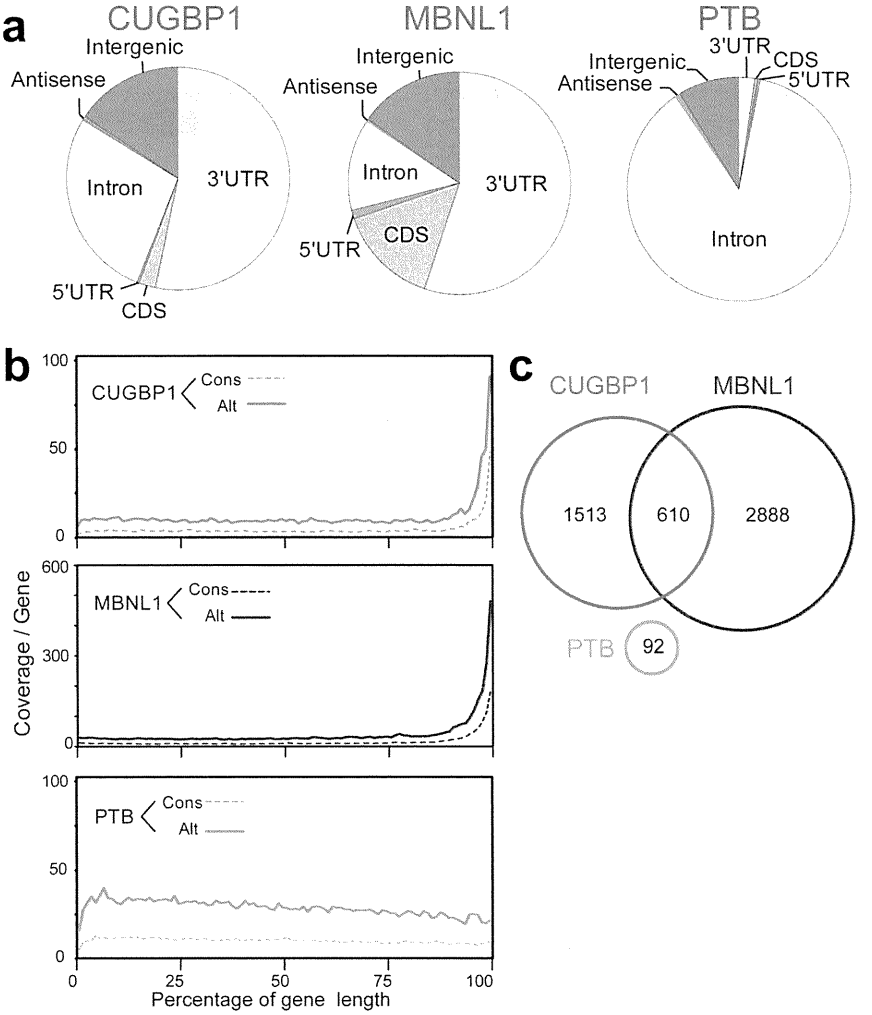


Figure 3 | Enrichment of CUGBP1 and MBNL1 CLIP-tags in the 3′ UTR. (a) Distributions of CUGBP1, MBNL1, and PTB binding regions. Binding regions are mapped to CDS (coding sequence), 5′ and 3′ UTRs, introns, intergenic regions (incl. tRNA and rRNA genes), or antisense within genes according to the UCSC knownGene annotation of the NCBI Build 37.1 mouse genome (mm9). Pie-charts show ratios of binding regions mapped to the indicated regions. (b) Distributions of CUGBP1, MBNL1, and PTB CLIP-tags mapped to the relative positions of all the mouse genes. The relative positions of the genes are shown in percentages of the gene length in abscissa. The broken lines represent 15,638 genes with constitutive transcriptional start and end sites (Cons), and the solid lines represent 7,477 genes with alternative transcriptional start or end site (Alt). (c) Venn diagram of the numbers of genes with CUGBP1-, MBNL1-, and PTB-binding regions within the 3′ UTR. Binding regions were identified using the SeqMonk software.

between these two proteins (Fig. 3c). All these data document that both CUGBP1 and MBNL1 preferentially bind to 3′ UTRs, indicating that this is a key function of both proteins in RNA processing. This suggests that the functional repertoire of MBNL1 should be expanded and that MBNL1, from being primarily regarded as regulator of alternative splicing, should also be considered as an important regulator of 3′ UTR-mediated processes, such as mRNA stability/degradation.

MBNL1 destabilize mRNAs. To analyze the function of CUGBP1/MBNL1 binding to 3′ UTRs, we made luciferase reporter constructs harboring CUGBP1/MBNL1-binding sites in the 3′ UTR. Since no CLIP tags were observed in the 3′ UTR of *Gapdh* (Supplementary Fig. S8), we made a *luciferase-Gapdh* 3′ UTR expression vector, and then inserted 12 repeats of GT and 7 repeats of CTG immediately after the stop codon of *luciferase* to introduce a CUGBP1-binding site (GU rich motif) and an MBNL1-binding site (YGCY motif), respectively (Fig. 4a). We also inserted 12 AC repeats as a control. Due to the high expression level of CUGBP1 in C2C12 cells we used HEK293 cells for transient transfection of these reporter constructs along with CUGBP1/MBNL1 expression vectors. For the constructs with *Gapdh* 3′ UTR alone or with AC repeats inserted, overexpression of CUGBP1 or MBNL1 had no effect on luciferase activity (Fig. 4b). For the GT repeat construct, overexpression of CUGBP1 decreased the luciferase activity, but MBNL1 had no effect. For the CTG repeat construct overexpression of MBNL1 dramatically decreased the luciferase activity, and also overexpression of CUGBP1 significantly reduced luciferase activity (Fig. 4b). In order to shed light on the mechanism underlying the observed decrease in luciferase activity we investigated the decay of *luciferase* mRNA. The SV40 promoter of the luciferase reporter constructs was replaced with a tet-repressible promoter, and HEK293 Tet-off cells were transiently transfected with these constructs. Doxycycline was added to the medium to stop transcription of the tet-responsive promoter, and the temporal profiles of *luciferase* and *GAPDH* mRNA levels were measured. Overexpression of MBNL1 together with the CTG repeat reporter construct resulted in highly increased decay of *luciferase* mRNA and CUGBP1 overexpression together with the GT repeat reporter construct also increased mRNA decay. Overexpression of either protein together with the *Gapdh* 3′ UTR control construct did not alter mRNA decay (Fig. 4c). These data demonstrate that binding of CUGBP1 and MBNL1 to the 3′ UTR promotes mRNA decay. To examine whether CUGBP1 and MBNL1 regulate decay of endogenous mRNAs, we next analyzed mRNA stability in actinomycin D treated C2C12 cells by expression arrays following siRNA knock down of CUGBP1 or MBNL1 (GEO accession number, GSE27583). To identify genes with reliable half-life estimates, we restricted our analysis to 195 transcripts using three conditions: (i) half-life between 2.5–5 hrs; (ii) correlation coefficient of fitting to an exponential decay greater than 0.9; and (iii) RMA-normalized signal values more than 100 at all time points. The median half-life of all the transcripts matching these criteria in the control is 3.56 hrs, whereas those from CUGBP1- and MBNL1-knocked down cells are significantly prolonged to 3.91 hrs and 3.73 hrs, respectively (Fig. 5a). We chose four additional representative mRNAs with a cluster of either CUGBP1- or MBNL1-tags in the 3′UTR, and confirmed by real time PCR that knockdown of either CUGBP1 or MBNL1 results in approximately two-fold increase in mRNA half-life of these target mRNAs (Fig. 5b). The half-lives of 100 out of 195 transcripts are prolonged both by knockdown of CUGBP1 and MBNL1, suggesting overlapping activity in the regulation of mRNA decay by CUGBP1 and MBNL1. We next analyzed the relationship between change in mRNA half-life and coverage of HITS-CLIP tags in the 3′ UTRs. We found that genes displaying prolongation of half-lives in response to CUGBP1 knockdown harbors more CUGBP1-tags in their 3′ UTRs, compared to those displaying shortening of half-lives (Fig. 5c). Similarly, genes that display prolongation of their half-lives in response to MBNL1 knockdown have more MBNL1-tags in their 3′ UTRs (Fig. 5c).

Gene Ontology analysis of CUGBP1/MBNL1-bound 3′ UTRs revealed that the terms 'cytoskeletal protein binding, 'transcription factor binding' and 'RNA binding' are significantly overrepresented for CUGBP1- and MBNL1-bound genes (Table 1).

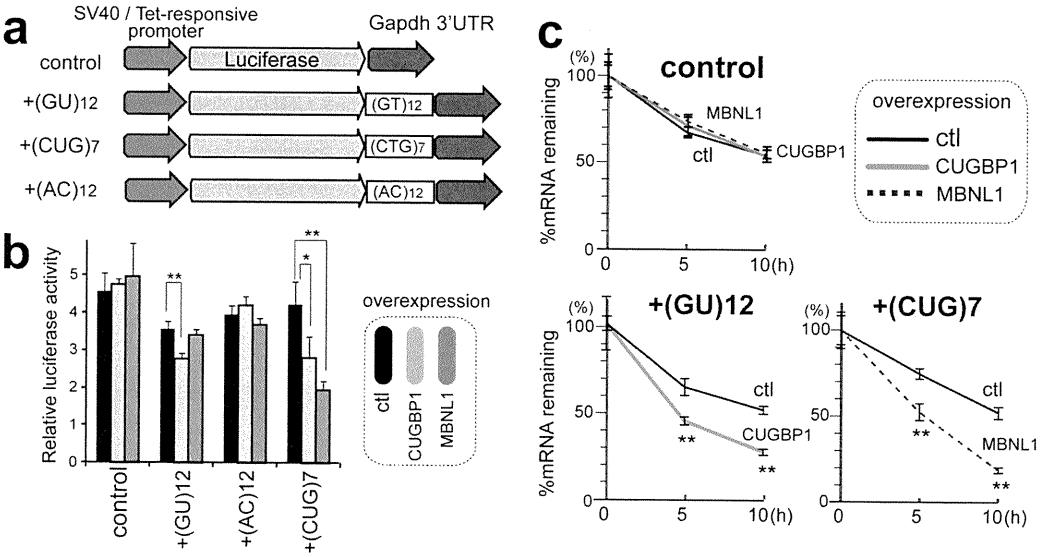


Figure 4 | Decay of *luciferase* mRNA by overexpression of CUGBP1/MBNL1. (a) Schemes of luciferase reporter plasmids harboring Gapdh 3′ UTR. Each construct was made carrying either SV40 or tet-responsive promoter. (b) Luciferase activity after overexpression of CUGBP1/MBNL1. HEK293 cells were transfected with the indicated SV40-driven luciferase reporter constructs. Luciferase activity is normalized for the transfection efficiency using co-transfection of pRL/SV40. (c) Decay of luciferase mRNA after overexpression of CUGBP1/MBNL1. HEK293 Tet-off cells were transfected with the indicated tet-responsive promoter-driven luciferase reporter constructs. Doxycycline was added to the medium to stop transcription at time 0. Temporal profiles of luciferase mRNA decay were quantified by real time RT-PCR and are normalized for Gapdh mRNA levels. All experiments were triplicated, and the mean and s.d. are indicated (* $p < 0.05$; ** $p < 0.01$).

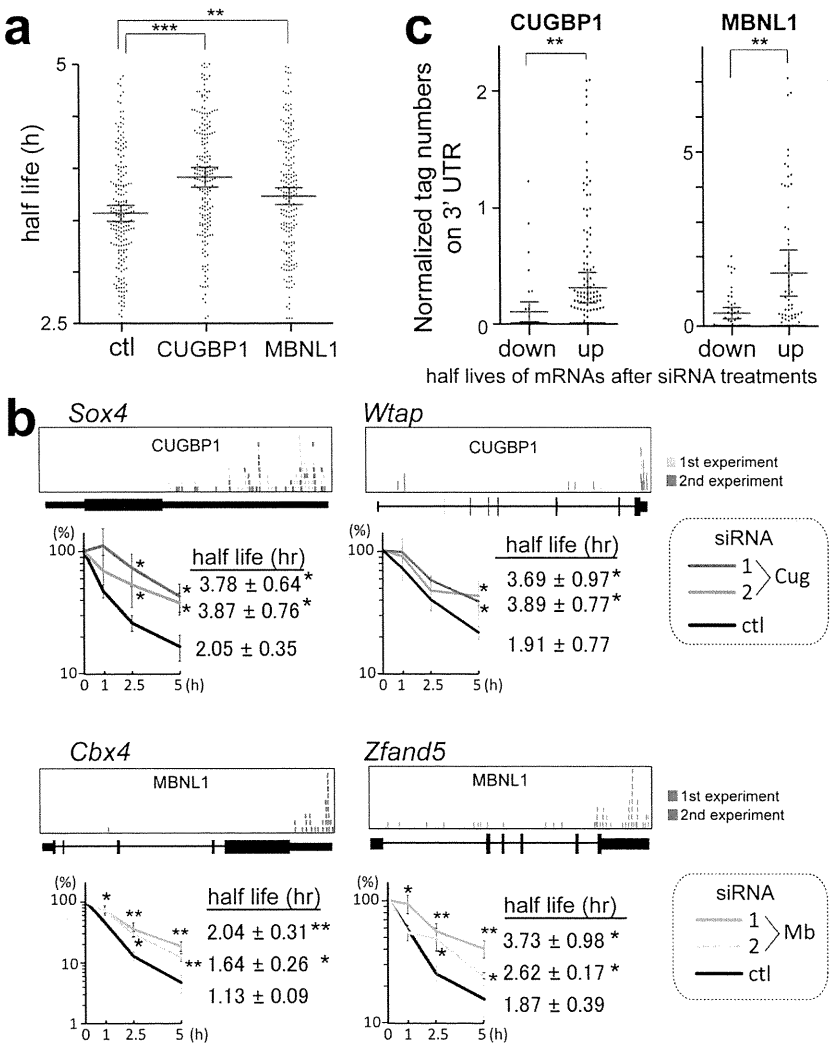


Figure 5 | Global analysis of mRNA decay by expression array of C2C12 cells treated with CUGBP1/MBNL1 siRNA. (a) Half-lives of mRNAs in C2C12 cells with the indicated siRNAs. Red lines represent means and 95% confidence intervals. ** $p < 0.01$ and *** $p < 0.001$. (b) Real-time RT-PCR analysis of the stability of four representative endogenous mRNAs, which were detected by expression arrays. CLIP-tag distributions are shown above each gene structure. C2C12 cells were treated with either control (ctl), CUGBP1 (Cug), or MBNL1 (Mb) siRNA. Actinomycin D was added to the medium to stop transcription at time 0. Temporal profiles of decay of the indicated genes were analyzed by real-time RT-PCR and are normalized for Gapdh mRNA levels. All experiments were triplicated, and the mean and s.d. are indicated (* $p < 0.05$ and ** $p < 0.01$). (c) Tag counts in the 3′ UTR of each gene are plotted in two categories of prolonging (up) and shortening (down) of half-lives after MBNL1 and CUGBP1 siRNAs. Red lines represent means and 95% confidence intervals. ** $p < 0.01$. Tag counts were normalized by the gene expression level at 0 h of cells treated with control siRNA.

Table 1 | The five most frequent Gene Ontology terms of mRNAs that are bound by CUGBP1 and MBNL1 to the 3′ UTR

CLIP data	GO ID	Term	P Value
CUGBP1	GO:0008092	cytoskeletal protein binding	1.58E-06
	GO:0003723	RNA binding	1.40E-04
	GO:0008134	transcription factor binding	9.65E-04
	GO:0051082	unfolded protein binding	0.003184
	GO:0019904	protein domain specific binding	0.006603
MBNL1	GO:0008092	cytoskeletal protein binding	7.31E-20
	GO:0008134	transcription factor binding	2.20E-08
	GO:0003723	RNA binding	0.001893
	GO:0019899	enzyme binding	0.002046
	GO:0032553	ribonucleotide binding	0.004210

We utilized the mRNAs that have more than 8-fold coverage of CLIP tags in their 3′UTR for the analysis by DAVID[53,54].