Table 2 Splice site strength of *COL1A1* exon 45 predicted by the NetGene2 and the Splice Site Prediction by Neural Network

Sequence variation	NetGene2		Splice Site Prediction by Neural Network	
	Confidence		Score	
	Acceptor	Donor	Acceptor	Donor
Wild type	0.97	0.93	0.98	0.95
rs1800215	–	–	–	–
rs1800217	–	–	–	–
c.3226G>A	–	0.89	–	–
c.3226G>A+rs1800215	0.94	0.89	–	–
c.3226G>A+rs1800217	–	0.89	–	–
c.3226G>T	0.94	0.87	–	–
c.3226G>T+rs1800215	0.94	0.86	–	–
c.3226G>T+rs1800217	0.94	0.87	–	–
c.3235G>A	0.94	0.87	–	–
c.3235G>A+rs1800215	0.94	0.86	–	–
c.3235G>A+rs1800217	0.94	0.87	–	–
c.3244G>T	0.94	0.86	–	–
c.3244G>T+rs1800215	0.94	0.86	–	–
c.3244G>T+rs1800217	0.94	0.86	–	–
c.3253G>A	0.94	0.88	0.52[a]	–
c.3253G>A+rs1800215	0.94	0.87	0.52[a]	–
c.3253G>A+rs1800217	–	0.89	–	–

– symbol represents being identical to the wild-type

[a] In addition to the native splice acceptor site of 0.98, a cryptic splice acceptor site 'AG' is generated at c.3253_3254

Table 3 Twelve SNPs identified by exome resequensing in five out of ten genes associated with hyperuricemia

Ch	Gene	Position	Nuc.	Amino acid	AF	dbSNP	II-2	II-3	VIII-2
1	*AGL*	100,336,361	C > T	Syn.	0.7	rs2230306	T/T	–/T	–/T
4	*ABCG2*[b]	89,034,551	G > A	Syn.	0.02	rs35622453	–/–	–/–	–/A
		89,052,323	C > A	Q141K	0.31	rs2231142	–/A	A/A	–/–
		89,061,114	G > A	V12M	0.19	rs2231137	–/A	–/–	–/A
4	*SLC2A9*[b]	9,909,923	C > T	P350L	0.33	rs2280205	–/T	–/T	–/–
		9,922,130	G > A	R294H	0.72	rs3733591	–/A	–/–	–/A
		9,998,440	G > A	Syn.	0.54	rs10939650	–/A	A/A	–/A
		10,022,981	G > A	G25R	0.43	rs2276961	–/A	–/A	–/–
		10,027,542	G > A	A17T	0.06	rs6820230	–/–	–/A	–/A
11	*SLC22A12*[b]	64,359,286	C > T	Syn.	0.21	rs3825016	–/T	T/T	–/T
		64,360,274	C > T	Syn.	0.81	rs11231825	–/T	–/–	–/T
12	*PFKM*		n.d.						
16	*UMOD*[b]		n.d.						
17	*G6PC*		n.d.						
X	*HPRT1*[a]		n.d.						
X	*PRPS1*[a]		n.d.						
X	*MAOA*	43,591,036	G > T	Syn.	0.3	rs6323	–/–	T/T	T/T

[a] The gene is associated with purine metabolism

[b] The gene is associated with renal excretion of urate

– symbol represents in the patients' genotypes mean being identical to the reference nucleotides

Syn synonymous nucleotide change, *AF* allelic frequency of the changed nucleotide, *n.d.* no SNPs are detected

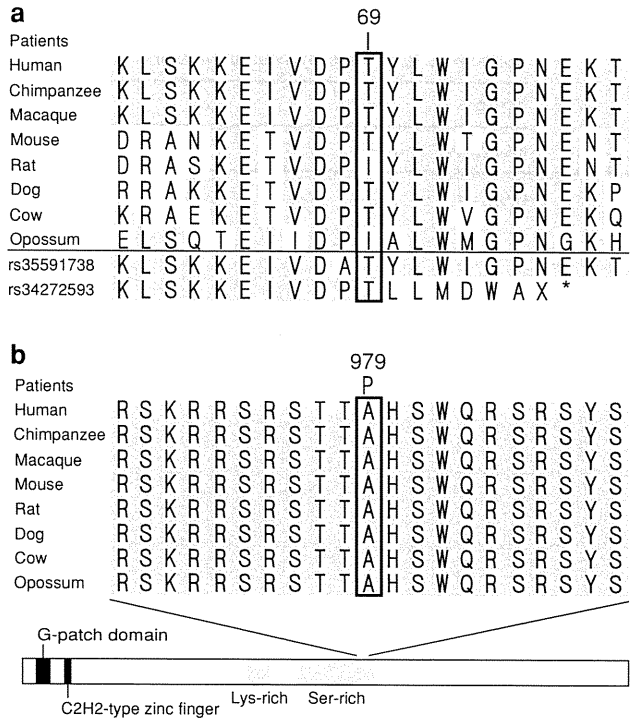


Fig. 3 Conservation of 21 amino acid segments encoded by *ZPBP2* and *GPATCH8* in mammals. Locations of non-synonymous variants identified in the Japanese family (F1) are *boxed*. Amino acids identical to human are *shaded*. **a** c.206C>T predicts p.T69I in *ZPBP2*. T69 is not conserved in rat and opossum. An SNP rs35591738 predicts p.P68A, and an SNP rs34272593 induces a frameshift. **b** c.2935G>C predicts p.A979P in *GPATCH8*. The mutation is located at the C-terminal end of the serine-rich region

Discussion

Phenotypic variabilities of OI mutations

We identified a heteroallelic c.3235G>A mutation in *COL1A1* exon 45 in a Japanese family with mild OI type I. The c.3235G>A mutation has been previously reported in six families with OI types I (Hartikka et al. 2004; Mottes et al. 1992; Roschger et al. 2008) and IV (Marini et al. 2007). *COL1A1* exon 45 encodes six of the 338 Gly-X–Y triplet repeats. Four additional mutations have been reported in exon 45 (Constantinou et al. 1989; Lund et al. 1997; Marini et al. 2007), and all substitute Ser or Cys for Gly (Fig. 2a) with mild to lethal phenotypes. Among the five mutations, two mutations introducing Cys result in a lethal type II, whereas three mutations introducing Ser give rise to milder types I, III and IV. This notion, however, cannot be applied to the other exons according to the human type I collagen mutation database (http://www.le.ac.uk/genetics/collagen/).

Being prompted by a report that more than 16–20% of exonic mutations disrupt an ESE (Gorlov et al. 2003), we asked if a mutation disrupting an ESE in exon 45 causes skipping of an inframe exon 45 and exhibits a severe dominant negative phenotype. Three Web-based programs predict that all the five mutations and the two SNPs affect 16 putative exonic splicing *cis*-elements (Table 1). We thus constructed and analyzed 18 minigenes carrying all possible combinations of the five mutations and two SNPs (Fig. 2b), but found that none affected pre-mRNA splicing (Fig. 2c). Our analysis suggests that the currently available algorithms of splicing *trans*-factors cannot efficiently predict splicing *cis*-elements. This is likely because the recognition motifs of splicing *trans*-factors are mostly determined by in vitro SELEX experiments. A recently developed technique, the high throughput sequencing coupled to crosslinking immunoprecipitation method (HITS-CLIP), enables us to extensively determine RNA segments recognized by a specific splicing *trans*-factor in vivo (Licatalosi et al. 2008; Yeo et al. 2009). Accumulation of knowledge with the HITS-CLIP technology will enable us to construct dependable algorithms to efficiently predict splicing *cis*-elements.

In addition to the phenotypic variability among similar mutations in the same exon, the same mutation often exhibits variable phenotypes, although the variability is usually less (Lund et al. 1996). This is also true for our families. In the Japanese family (F1), two sons (II-1 and II-3) experienced many fractures, whereas the father (I-1) and another son (II-2) had no history of fractures. In the Italian family (F2), the son suffered from many fractures, but his affected mother did not (Mottes et al. 1992). In the Canadian family (F3), the father was classified as OI type IV, whereas his daughter as OI type I (Roschger et al. 2008). Variable clinical phenotypes of the c.3235G>A mutation is likely due to differences in environmental factors or to SNPs in disease-modifying genes, but the molecular bases have not been elucidated in any type of OI.

Molecular basis of hyperuricemia

In F1, hyperuricemia cosegregated with OI type 1. Although no known genes causing hyperuricemia are on chr 17 where *COL1A1* is located, two candidate genes of *PRPSAP1* and *PRPSAP2* are on chr 17. Capillary sequencing of these genes, however, detected no mutation. We thus employed exome-capture resequencing analysis of two siblings in F1 to search for a responsible gene for hyperuricemia. We first looked into SNPs in the 10 candidate genes that are associated with hyperuricemia, purine metabolism and renal excretion, and found 12 SNPs in 10 genes (Table 3). Among them, three SNPs in *ABCG2* (rs2231142) and *SLC22A12* (rs3825016 and rs11231825) are previously reported markers for hyperuricemia and/or gout.

Dehghan et al. (2008) report that rs2231142 in *ABCG2* is associated with gout by a genome-wide association study (OR = 1.74 and 1.71 in white and black participants, respectively). Woodward et al. (2009) showed a significant association between rs2231142 and hyperuricemia (OR = 1.68) in a population-based study of 14,783 individuals. Kolz et al. (2009) demonstrated that rs2231142 elevated the serum urate concentration more strongly in men than in women by meta-analysis of 28,141 individuals. Stark et al. (2009) analyzed 683 patients with gout and indicated a significant association between rs2231142 and gout (OR = 1.37). Although rs2231142 is an attractive causative SNP, our normouricemic subjects were also heterozygous for rs2231142.

Graessler et al. (2006) analyzed 389 German individuals with primary hyperuricemia and found that rs3825016 and rs11231825 in *SLC22A12* were significantly associated with reduced fractional excretion of urate in the kidney. Tabara et al. (2010) analyzed 1,526 normal Japanese individuals retrospectively and longitudinally, and clarified that rs11231825 was associated with reduced urate excretion and with future development of hyperuricemia. Again, although the two SNPs are attractive causes of hyperuricemia, we observe variable dosages of these SNPs even in our normouricemic subjects.

We next looked into neighboring genes of *COL1A1* without considering the functions of the gene products, and identified that two missense variants in *ZPBP2* and *GPATCH8* cosegregated with the *COL1A1* mutation in F1. Neither variant was detected in 300 normal alleles or in dbSNP132. *ZPBP2* p.T69I, however, is unlikely to be pathogenic for three reasons: lack of conservation in mammals; two missense/frameshifting SNPs at or close to the variant site (Fig. 3a); and the benign predicted outcome by PolyPhen-2 and SIFT. On the other hand, p.A979P in *GPATCH8* substitutes an amino acid in the highly conserved serine-rich region (Fig. 3b), and the substitution is predicted to damage the structure and function of the protein by in silico analysis. *GPATCH8* encodes the G patch domain-containing protein 8 that harbors both an RNA-processing domain and a zinc finger domain. *GPATCH8* is expressed in a wide variety of human tissues including skeletal muscles, brain, heart, pancreas, liver and kidney (McKinney et al. 2004). Functions of the *GPATCH8* gene product, however, have not been studied to date. The p.A979P variant in *GPATCH8* is highly likely to be associated with hyperuricemia in F1, but it may also cause another yet unidentified phenotype that cosegregates with OI.

Acknowledgments We would like to thank the families for their participation in this study. We are grateful to Dr. Kunio Ihara at the Center for Gene Research of Nagoya University for the SOLiD sequencing analysis and Keiko Itano for technical assistance. This work was supported by Grants-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and the Ministry of Health, Labor and Welfare of Japan.

References

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR (2010) A method and server for predicting damaging missense mutations. Nat Methods 7:248–249

Alanay Y, Avaygan H, Camacho N, Utine GE, Boduroglu K, Aktas D, Alikasifoglu M, Tuncbilek E, Orhan D, Bakar FT, Zabel B, Superti-Furga A, Bruckner-Tuderman L, Curry CJ, Pyott S, Byers PH, Eyre DR, Baldridge D, Lee B, Merrill AE, Davis EC, Cohn DH, Akarsu N, Krakow D (2010) Mutations in the gene encoding the RER protein FKBP65 cause autosomal-recessive osteogenesis imperfecta. Am J Hum Genet 87:572–573

Allen GE, Rogers FB, Lansbury J (1955) Osteogenesis imperfecta tarda with hyperuricemia and gout: report of three cases. Am J Med Sci 230:30–32

Baldridge D, Schwarze U, Morello R, Lennington J, Bertin TK, Pace JM, Pepin MG, Weis M, Eyre DR, Walsh J, Lambert D, Green A, Robinson H, Michelson M, Houge G, Lindman C, Martin J, Ward J, Lemyre E, Mitchell JJ, Krakow D, Rimoin DL, Cohn DH, Byers PH, Lee B (2008) CRTAP and LEPRE1 mutations in recessive osteogenesis imperfecta. Hum Mutat 29:1435–1442

Bodian DL, Madhan B, Brodsky B, Klein TE (2008) Predicting the clinical lethality of osteogenesis imperfecta from collagen glycine mutations. Biochemistry 47:5424–5432

Brunak S, Engelbrecht J, Knudsen S (1991) Prediction of human mRNA donor and acceptor sites from the DNA sequence. J Mol Biol 220:49–65

Cabral WA, Chang W, Barnes AM, Weis M, Scott MA, Leikin S, Makareeva E, Kuznetsova NV, Rosenbaum KN, Tifft CJ, Bulas DI, Kozma C, Smith PA, Eyre DR, Marini JC (2007) Prolyl 3-hydroxylase 1 deficiency causes a recessive metabolic bone disorder resembling lethal/severe osteogenesis imperfecta. Nat Genet 39:359–365

Cartegni L, Chew SL, Krainer AR (2002) Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat Rev Genet 3:285–298

Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR (2003) ESEfinder: a web resource to identify exonic splicing enhancers. Nucleic Acids Res 31:3568–3571

Christiansen HE, Schwarze U, Pyott SM, AlSwaid A, Al Balwi M, Alrasheed S, Pepin MG, Weis MA, Eyre DR, Byers PH (2010) Homozygosity for a missense mutation in SERPINH1, which encodes the collagen chaperone protein HSP47, results in severe recessive osteogenesis imperfecta. Am J Hum Genet 86:389–398

Constantinou CD, Nielsen KB, Prockop DJ (1989) A lethal variant of osteogenesis imperfecta has a single base mutation that substitutes cysteine for glycine 904 of the alpha 1(I) chain of type I procollagen. The asymptomatic mother has an unidentified mutation producing an overmodified and unstable type I procollagen. J Clin Invest 83:574–584

Dalgleish R (1997) The human type I collagen mutation database. Nucleic Acids Res 25:181–187

Dehghan A, Kottgen A, Yang Q, Hwang SJ, Kao WL, Rivadeneira F, Boerwinkle E, Levy D, Hofman A, Astor BC, Benjamin EJ, van Duijn CM, Witteman JC, Coresh J, Fox CS (2008) Association of three genetic loci with uric acid concentration and risk of gout: a genome-wide association study. Lancet 372:1953–1961

Doring A, Gieger C, Mehta D, Gohlke H, Prokisch H, Coassin S, Fischer G, Henke K, Klopp N, Kronenberg F, Paulweber B, Pfeufer A, Rosskopf D, Volzke H, Illig T, Meitinger T, Wichmann HE, Meisinger C (2008) SLC2A9 influences uric acid concentrations with pronounced sex-specific effects. Nat Genet 40:430–436

Fairbrother WG, Yeh RF, Sharp PA, Burge CB (2002) Predictive identification of exonic splicing enhancers in human genes. Science 297:1007–1013

Gibbs RA, Caskey CT (1987) Identification and localization of mutations at the Lesch–Nyhan locus by ribonuclease A cleavage. Science 236:303–305

Glorieux FH, Rauch F, Plotkin H, Ward L, Travers R, Roughley P, Lalic L, Glorieux DF, Fassier F, Bishop NJ (2000) Type V osteogenesis imperfecta: a new form of brittle bone disease. J Bone Miner Res 15:1650–1658

Glorieux FH, Ward LM, Rauch F, Lalic L, Roughley PJ, Travers R (2002) Osteogenesis imperfecta type VI: a form of brittle bone disease with a mineralization defect. J Bone Miner Res 17:30–38

Goren A, Ram O, Amit M, Keren H, Lev-Maor G, Vig I, Pupko T, Ast G (2006) Comparative analysis identifies exonic splicing regulatory sequences—the complex definition of enhancers and silencers. Mol Cell 22:769–781

Gorlov IP, Gorlova OY, Frazier ML, Amos CI (2003) Missense mutations in hMLH1 and hMSH2 are associated with exonic splicing enhancers. Am J Hum Genet 73:1157–1161

Graessler J, Graessler A, Unger S, Kopprasch S, Tausche AK, Kuhlisch E, Schroeder HE (2006) Association of the human urate transporter 1 with reduced renal uric acid excretion and hyperuricemia in a German Caucasian population. Arthritis Rheum 54:292–300

Hart TC, Gorry MC, Hart PS, Woodard AS, Shihabi Z, Sandhu J, Shirts B, Xu L, Zhu H, Barmada MM, Bleyer AJ (2002) Mutations of the UMOD gene are responsible for medullary cystic kidney disease 2 and familial juvenile hyperuricaemic nephropathy. J Med Genet 39:882–892

Hartikka H, Kuurila K, Korkko J, Kaitila I, Grenman R, Pynnonen S, Hyland JC, Ala-Kokko L (2004) Lack of correlation between the type of COL1A1 or COL1A2 mutation and hearing loss in osteogenesis imperfecta patients. Hum Mutat 24:147–154

Hebsgaard SM, Korning PG, Tolstrup N, Engelbrecht J, Rouze P, Brunak S (1996) Splice site prediction in *Arabidopsis thaliana* pre-mRNA by combining local and global sequence information. Nucleic Acids Res 24:3439–3452

Ishizuka T, Ahmad I, Kita K, Sonoda T, Ishijima S, Sawa K, Suzuki N, Tatibana M (1996) The human phosphoribosylpyrophosphate synthetase-associated protein 39 gene (PRPSAP1) is located in the chromosome region 17q24–q25. Genomics 33:332–334

Katashima R, Iwahana H, Fujimura M, Yamaoka T, Itakura M (1998) Assignment of the human phosphoribosylpyrophosphate synthetase-associated protein 41 gene (PRPSAP2) to 17p11.2–p12. Genomics 54:180–181

Kolz M, Johnson T, Sanna S, Teumer A, Vitart V, Perola M, Mangino M, Albrecht E, Wallace C, Farrall M, Johansson A, Nyholt DR, Aulchenko Y, Beckmann JS, Bergmann S, Bochud M, Brown M, Campbell H, Connell J, Dominiczak A, Homuth G, Lamina C, McCarthy MI, Meitinger T, Mooser V, Munroe P, Nauck M, Peden J, Prokisch H, Salo P, Salomaa V, Samani NJ, Schlessinger D, Uda M, Volker U, Waeber G, Waterworth D, Wang-Sattler R, Wright AF, Adamski J, Whitfield JB, Gyllensten U, Wilson JF, Rudan I, Pramstaller P, Watkins H, Doering A, Wichmann HE, Spector TD, Peltonen L, Volzke H, Nagaraja R, Vollenweider P, Caulfield M, Illig T, Gieger C (2009) Meta-analysis of 28,141 individuals identifies common variants within five new loci that influence uric acid concentrations. PLoS Genet 5:e1000504

Kumar P, Henikoff S, Ng PC (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 4:1073–1081

Lalonde E, Albrecht S, Ha KC, Jacob K, Bolduc N, Polychronakos C, Dechelotte P, Majewski J, Jabado N (2010) Unexpected allelic heterogeneity and spectrum of mutations in Fowler syndrome revealed by next-generation exome sequencing. Hum Mutat 31:918–923

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10:R25

Lapunzina P, Aglan M, Temtamy S, Caparros-Martin JA, Valencia M, Leton R, Martinez-Glez V, Elhossini R, Amr K, Vilaboa N, Ruiz-Perez VL (2010) Identification of a frameshift mutation in Osterix in a patient with recessive osteogenesis imperfecta. Am J Hum Genet 87:110–114

Licatalosi DD, Mele A, Fak JJ, Ule J, Kayikci M, Chi SW, Clark TA, Schweitzer AC, Blume JE, Wang X, Darnell JC, Darnell RB (2008) HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature 456:464–469

Lund AM, Schwartz M, Skovby F (1996) Variable clinical expression in a family with OI type IV due to deletion of three base pairs in COL1A1. Clin Genet 50:304–309

Lund AM, Skovby F, Schwartz M (1997) Serine for glycine substitutions in the C-terminal third of the alpha 1(I) chain of collagen I in five patients with nonlethal osteogenesis imperfecta. Hum Mutat 9:378–382

Marini JC, Forlino A, Cabral WA, Barnes AM, San Antonio JD, Milgrom S, Hyland JC, Korkko J, Prockop DJ, De Paepe A, Coucke P, Symoens S, Glorieux FH, Roughley PJ, Lund AM, Kuurila-Svahn K, Hartikka H, Cohn DH, Krakow D, Mottes M, Schwarze U, Chen D, Yang K, Kuslich C, Troendle J, Dalgleish R, Byers PH (2007) Consortium for osteogenesis imperfecta mutations in the helical domain of type I collagen: regions rich in lethal mutations align with collagen binding sites for integrins and proteoglycans. Hum Mutat 28:209–221

McKinney JL, Murdoch DJ, Wang J, Robinson J, Biltcliffe C, Khan HM, Walker PM, Savage J, Skerjanc I, Hegele RA (2004) Venn analysis as part of a bioinformatic approach to prioritize expressed sequence tags from cardiac libraries. Clin Biochem 37:953–960

Morello R, Bertin TK, Chen Y, Hicks J, Tonachini L, Monticone M, Castagnola P, Rauch F, Glorieux FH, Vranka J, Bachinger HP, Pace JM, Schwarze U, Byers PH, Weis M, Fernandes RJ, Eyre DR, Yao Z, Boyce BF, Lee B (2006) CRTAP is required for prolyl 3-hydroxylation and mutations cause recessive osteogenesis imperfecta. Cell 127:291–304

Mottes M, Sangalli A, Valli M, Gomez Lira M, Tenni R, Buttitta P, Pignatti PF, Cetta G (1992) Mild dominant osteogenesis imperfecta with intrafamilial variability: the cause is a serine for glycine alpha 1(I) 901 substitution in a type-I collagen gene. Hum Genet 89:480–484

Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, Lee C, Shaffer T, Wong M, Bhattacharjee A, Eichler EE, Bamshad M, Nickerson DA, Shendure J (2009) Targeted capture and massively parallel sequencing of 12 human exomes. Nature 461:272–276

Ng SB, Bigham AW, Buckingham KJ, Hannibal MC, McMillin MJ, Gildersleeve HI, Beck AE, Tabor HK, Cooper GM, Mefford HC, Lee C, Turner EH, Smith JD, Rieder MJ, Yoshiura K, Matsumoto N, Ohta T, Niikawa N, Nickerson DA, Bamshad MJ, Shendure J (2010a) Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. Nat Genet 42:790–793

Ng SB, Buckingham KJ, Lee C, Bigham AW, Tabor HK, Dent KM, Huff CD, Shannon PT, Jabs EW, Nickerson DA, Shendure J,

Bamshad MJ (2010b) Exome sequencing identifies the cause of a Mendelian disorder. Nat Genet 42:30–35

Ohno K, Milone M, Shen X-M, Engel AG (2003) A frameshifting mutation in *CHRNE* unmasks skipping of the preceding exon. Hum Mol Genet 12:3055–3066

Rauch F, Glorieux FH (2004) Osteogenesis imperfecta. Lancet 363:1377–1385

Reese MG, Eeckman FH, Kulp D, Haussler D (1997) Improved splice site detection in Genie. J Comput Biol 4:311–323

Roessler BJ, Nosal JM, Smith PR, Heidler SA, Palella TD, Switzer RL, Becker MA (1993) Human X-linked phosphoribosylpyrophosphate synthetase superactivity is associated with distinct point mutations in the PRPS1 gene. J Biol Chem 268:26476–26481

Roschger P, Fratzl-Zelman N, Misof BM, Glorieux FH, Klaushofer K, Rauch F (2008) Evidence that abnormal high bone mineralization in growing children with osteogenesis imperfecta is not associated with specific collagen mutations. Calcif Tissue Int 82:263–270

Sillence DO, Senn A, Danks DM (1979) Genetic heterogeneity in osteogenesis imperfecta. J Med Genet 16:101–116

Stark K, Reinhard W, Grassl M, Erdmann J, Schunkert H, Illig T, Hengstenberg C (2009) Common polymorphisms influencing serum uric acid levels contribute to susceptibility to gout, but not to coronary artery disease. PLoS One 4:e7729

Tabara Y, Kohara K, Kawamoto R, Hiura Y, Nishimura K, Morisaki T, Kokubo Y, Okamura T, Tomoike H, Iwai N, Miki T (2010) Association of four genetic loci with uric acid levels and reduced renal function: the J-SHIPP Suita study. Am J Nephrol 32:279–286

van Dijk FS, Nesbitt IM, Zwikstra EH, Nikkels PG, Piersma SR, Fratantoni SA, Jimenez CR, Huizer M, Morsman AC, Cobben JM, van Roij MH, Elting MW, Verbeke JI, Wijnaendts LC, Shaw NJ, Hogler W, McKeown C, Sistermans EA, Dalton A, Meijers-Heijboer H, Pals G (2009) PPIB mutations cause severe osteogenesis imperfecta. Am J Hum Genet 85:521–527

Vitart V, Rudan I, Hayward C, Gray NK, Floyd J, Palmer CN, Knott SA, Kolcic I, Polasek O, Graessler J, Wilson JF, Marinaki A, Riches PL, Shu X, Janicijevic B, Smolej-Narancic N, Gorgoni B, Morgan J, Campbell S, Biloglav Z, Barac-Lauc L, Pericic M, Klaric IM, Zgaga L, Skaric-Juric T, Wild SH, Richardson WA, Hohenstein P, Kimber CH, Tenesa A, Donnelly LA, Fairbanks LD, Aringer M, McKeigue PM, Ralston SH, Morris AD, Rudan P, Hastie ND, Campbell H, Wright AF (2008) SLC2A9 is a newly identified urate transporter influencing serum urate concentration, urate excretion and gout. Nat Genet 40:437–442

Wang Z, Rolish ME, Yeo G, Tung V, Mawson M, Burge CB (2004) Systematic identification and analysis of exonic splicing silencers. Cell 119:831–845

Ward LM, Rauch F, Travers R, Chabot G, Azouz EM, Lalic L, Roughley PJ, Glorieux FH (2002) Osteogenesis imperfecta type VII: an autosomal recessive form of brittle bone disease. Bone 31:12–18

Woodward OM, Kottgen A, Coresh J, Boerwinkle E, Guggino WB, Kottgen M (2009) Identification of a urate transporter, ABCG2, with a common functional polymorphism causing gout. Proc Natl Acad Sci USA 106:10338–10342

Yeo GW, Coufal NG, Liang TY, Peng GE, Fu XD, Gage FH (2009) An RNA code for the FOX2 splicing regulator revealed by mapping RNA–protein interactions in stem cells. Nat Struct Mol Biol 16:130–137

Yi X, Liang Y, Huerta-Sanchez E, Jin X, Cuo ZX, Pool JE, Xu X, Jiang H, Vinckenbosch N, Korneliussen TS, Zheng H, Liu T, He W, Li K, Luo R, Nie X, Wu H, Zhao M, Cao H, Zou J, Shan Y, Li S, Yang Q, Asan Ni P, Tian G, Xu J, Liu X, Jiang T, Wu R, Zhou G, Tang M, Qin J, Wang T, Feng S, Li G, Huasang Luosang J, Wang W, Chen F, Wang Y, Zheng X, Li Z, Bianba Z, Yang G, Wang X, Tang S, Gao G, Chen Y, Luo Z, Gusang L, Cao Z, Zhang Q, Ouyang W, Ren X, Liang H, Zheng H, Huang Y, Li J, Bolund L, Kristiansen K, Li Y, Zhang Y, Zhang X, Li R, Li S, Yang H, Nielsen R, Wang J, Wang J (2010) Sequencing of 50 human exomes reveals adaptation to high altitude. Science 329:75–78

Zhang XH, Chasin LA (2004) Computational definition of sequence motifs governing constitutive exon splicing. Genes Dev 18:1241–1250

Zhang XH, Kangsamaksin T, Chao MS, Banerjee JK, Chasin LA (2005) Exon inclusion is dependent on predictable exonic splicing enhancers. Mol Cell Biol 25:7323–7332

Expansion of Intronic GGCCTG Hexanucleotide Repeat in *NOP56* Causes SCA36, a Type of Spinocerebellar Ataxia Accompanied by Motor Neuron Involvement

Hatasu Kobayashi,[1,4] Koji Abe,[2,4] Tohru Matsuura,[2,4] Yoshio Ikeda,[2] Toshiaki Hitomi,[1] Yuji Akechi,[2] Toshiyuki Habu,[3] Wanyang Liu,[1] Hiroko Okuda,[1] and Akio Koizumi[1,*]

Autosomal-dominant spinocerebellar ataxias (SCAs) are a heterogeneous group of neurodegenerative disorders. In this study, we performed genetic analysis of a unique form of SCA (SCA36) that is accompanied by motor neuron involvement. Genome-wide linkage analysis and subsequent fine mapping for three unrelated Japanese families in a cohort of SCA cases, in whom molecular diagnosis had never been performed, mapped the disease locus to the region of a 1.8 Mb stretch (LOD score of 4.60) on 20p13 (D20S906–D20S193) harboring 37 genes with definitive open reading frames. We sequenced 33 of these and observed a large expansion of an intronic GGCCTG hexanucleotide repeat in *NOP56* and an unregistered missense variant (Phe265Leu) in *C20orf194*, but we found no mutations in *PDYN* and *TGM6*. The expansion showed complete segregation with the SCA phenotype in family studies, whereas Phe265Leu in *C20orf194* did not. Screening of the expansions in the SCA cohort cases revealed four additional occurrences, but none were revealed in the cohort of 27 Alzheimer disease cases, 154 amyotrophic lateral sclerosis cases, or 300 controls. In total, nine unrelated cases were found in 251 cohort SCA patients (3.6%). A founder haplotype was confirmed in these cases. RNA foci formation was detected in lymphoblastoid cells from affected subjects by fluorescence in situ hybridization. Double staining and gel-shift assay showed that (GGCCUG)n binds the RNA-binding protein SRSF2 but that $(CUG)_6$ does not. In addition, transcription of *MIR1292*, a neighboring miRNA, was significantly decreased in lymphoblastoid cells of SCA patients. Our finding suggests that SCA36 is caused by hexanucleotide repeat expansions through RNA gain of function.

Autosomal-dominant spinocerebellar ataxias (SCAs) are a heterogeneous group of neurodegenerative disorders characterized by loss of balance, progressive gait, and limb ataxia.[1–3] We recently encountered two unrelated patients with intriguing clinical symptoms from a community in the Chugoku region in western mainland Japan.[4] These patients both showed complicated clinical features, with ataxia as the first symptom, followed by characteristic late-onset involvement of the motor neuron system that caused symptoms similar to those of amyotrophic lateral sclerosis (ALS [MIM 105400]).[4] Some SCAs (SCA1 [MIM 164400], SCA2 [MIM 183090], SCA3 [MIM 607047], and SCA6 [MIM 183086]) are known to slightly affect motor neurons; however, their involvement is minimal and the patients usually do not develop skeletal muscle and tongue atrophies.[4] Of particular interest is that RNA foci have been recently demonstrated in hereditary disorders caused by microsatellite repeat expansions or insertions in the noncoding regions of their gene.[5–7] The unique clinical features in these families have seldom been described in previous reports; therefore, we undertook a genetic analysis.

A similar form of SCA was observed in five Japanese cases from a cohort of 251 patients with SCA, in whom molecular diagnosis had not been performed, who were followed by the Department of Neurology, Okayama University Hospital. These five cases originated from a city of 450,000 people in the Chugoku region. Thus, we suspected the presence of a founder mutation common to these five cases, prompting us to recruit these five families (pedigrees 1–5) (Figure 1, Table 1). This study was approved by the Ethics Committee of Kyoto University and the Okayama University institutional review board. Written informed consent was obtained from all subjects. An index of cases per family was investigated in some depth: IV-4 in pedigree 1, II-1 in pedigree 2, III-1 in pedigree 3, II-1 in pedigree 4, and II-1 in pedigree 5. The mean age at onset of cerebellar ataxia was 52.8 ± 4.3 years, and the disease was transmitted by an autosomal-dominant mode of inheritance. All affected individuals started their ataxic symptoms, such as gait and truncal instability, ataxic dysarthria, and uncoordinated limbs, in their late forties to fifties. MRI revealed relatively confined and mild cerebellar atrophy (Figure 2A). Unlike individuals with previously known SCAs, all affected individuals with longer disease duration showed obvious signs of motor neuron involvement (Table 1). Characteristically, all affected individuals exhibited tongue atrophy with fasciculation, although its degree of severity varied (Figure 2B). Despite severe tongue atrophy in some cases, their swallowing function was relatively preserved, and they were allowed oral intake even at a later point after onset. In addition to tongue atrophy, skeletal muscle atrophy and fasciculation in the limbs and trunk appeared in advanced cases.[4] Tendon reflexes were generally mildly to severely hyperreactive in most

[1]Department of Health and Environmental Sciences, Graduate School of Medicine, Kyoto University, Kyoto, Japan; [2]Department of Neurology, Graduate School of Medicine, Dentistry and Pharmaceutical Science, Okayama University, Okayama, Japan; [3]Radiation Biology Center, Kyoto University, Kyoto, Japan
[4]These authors contributed equally to this work
*Correspondence: koizumi.akio.5v@kyoto-u.ac.jp
DOI 10.1016/j.ajhg.2011.05.015.

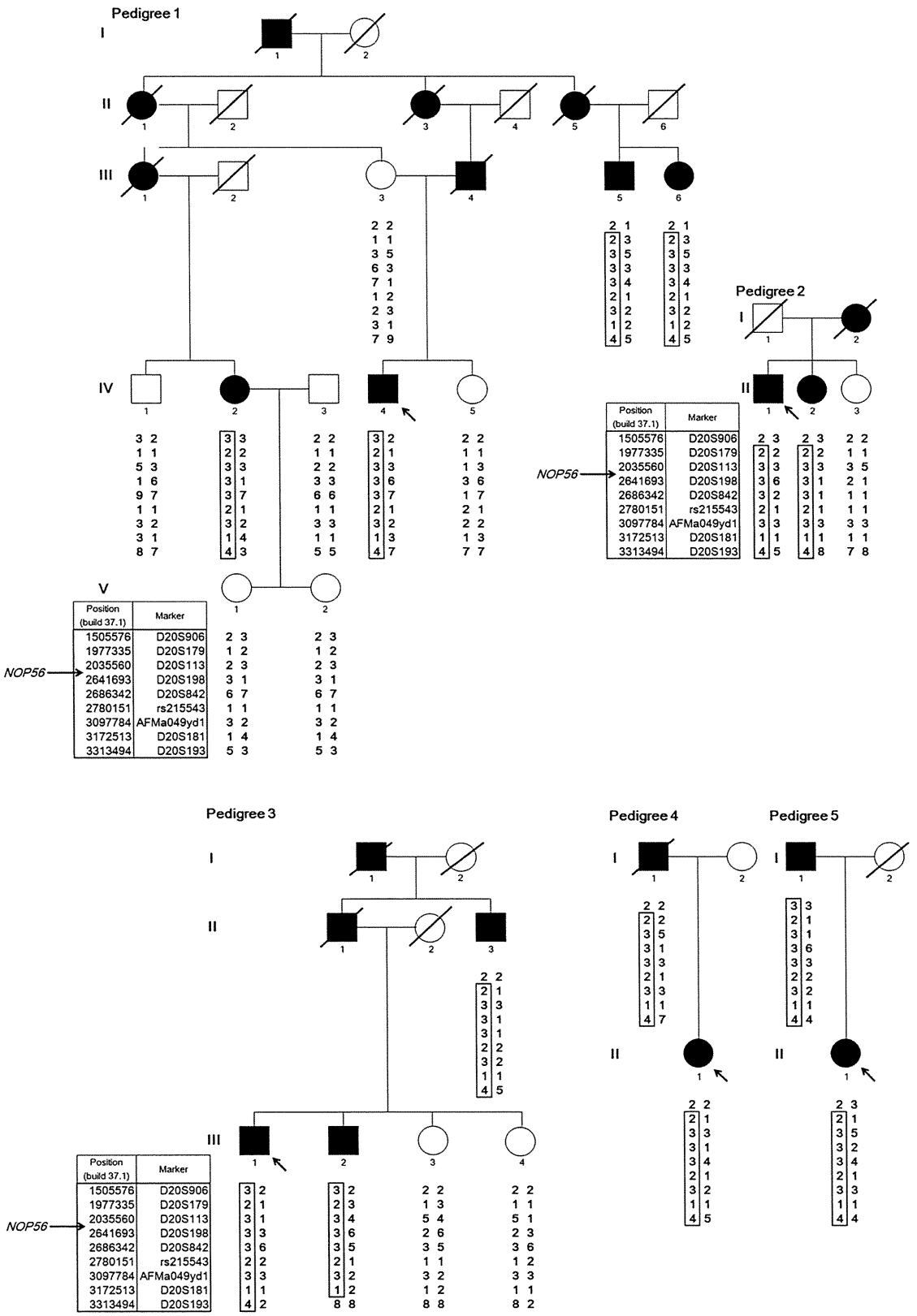


Figure 1. Pedigree Charts of the Five SCA Families
Haplotypes are shown for nine markers from D20S906 (1,505,576 bp) to D20S193 (3,313,494 bp), spanning 1.8 Mb on chromosome 20p13. *NOP56* is located at 2,633,254–2,639,039 bp (NCBI build 37.1). Filled and unfilled symbols indicate affected and unaffected individuals, respectively. Squares and circles represent males and females, respectively. A slash indicates a deceased individual. The putative founder haplotypes among patients are shown in boxes constructed by GENHUNTER.[8] Arrows indicate the index case. The pedigrees were slightly modified for privacy protection.

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

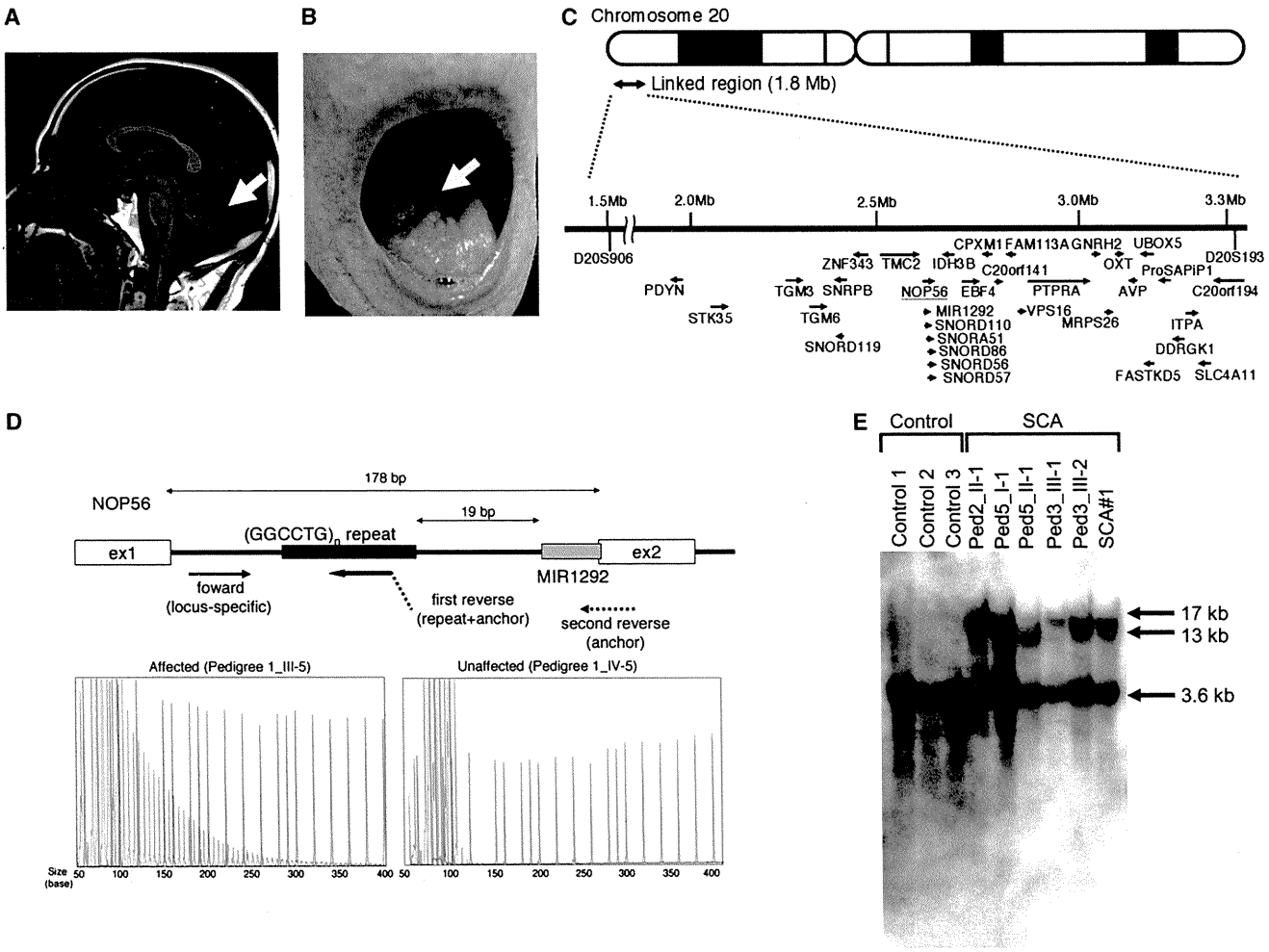


Figure 2. Motor Neuron Involvement and (GGCCTG)n Expansion in the First Intron of *NOP56*

(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

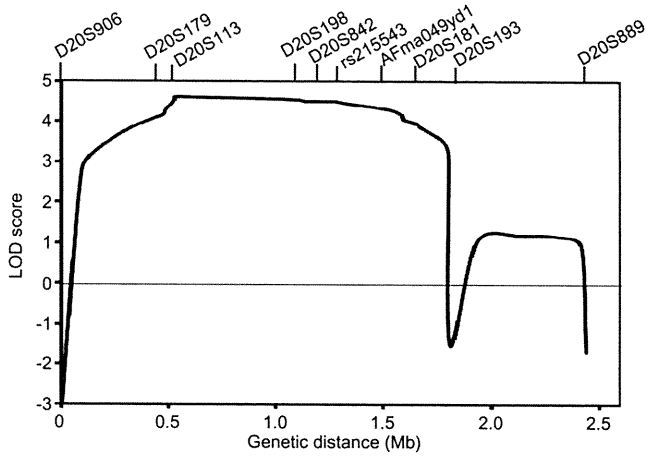


Figure 3. Multipoint Linkage Analysis with Ten Markers on Chromosome 20p13

(p.Phe265Leu), in *C20orf194*, whereas the other seven included one synonymous variant, c.1695T>A (p.Leu565Leu), 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.

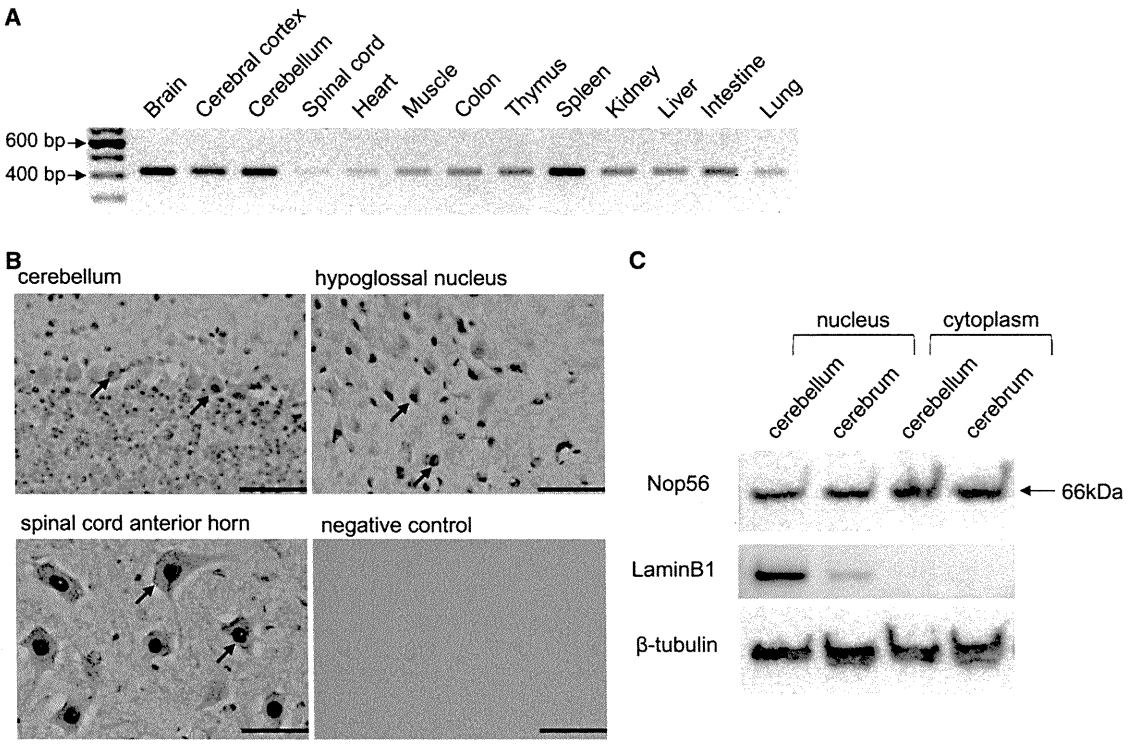


Figure 4. Nop56 in the Mouse Nervous System
(A) RT-PCR analysis of Nop56 (422 bp) in various mouse tissues. cDNA (25 ng) collected from various organs of C57BL/6 mice was purchased from GenoStaf (Tokyo, Japan).
(B) Immunohistochemical analysis of Nop56 in the cerebellum, hypoglossal nucleus, and spinal 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

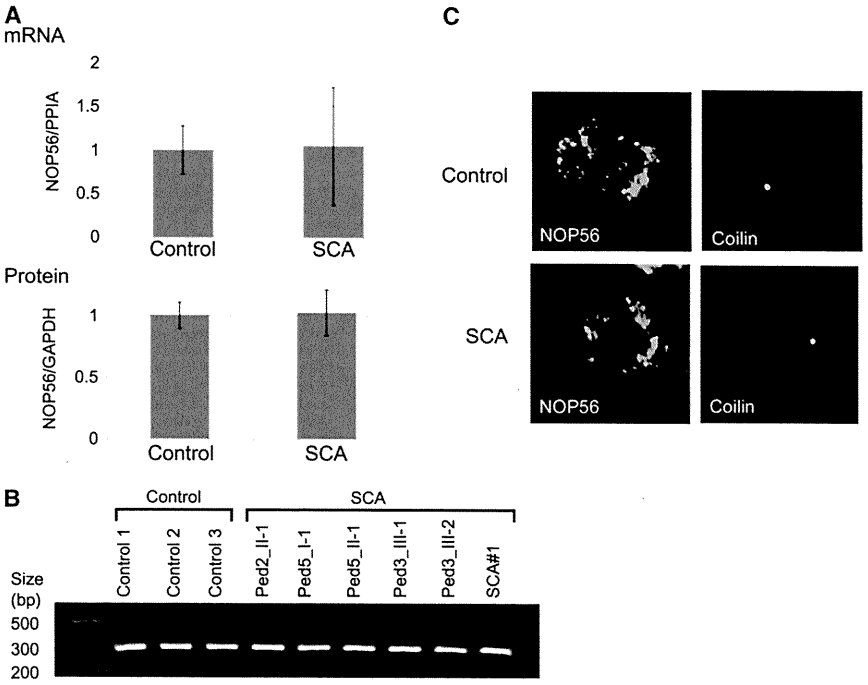


Figure 5. Analysis of NOP56 in LCLs from SCA Patients

(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 *PP1A* 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

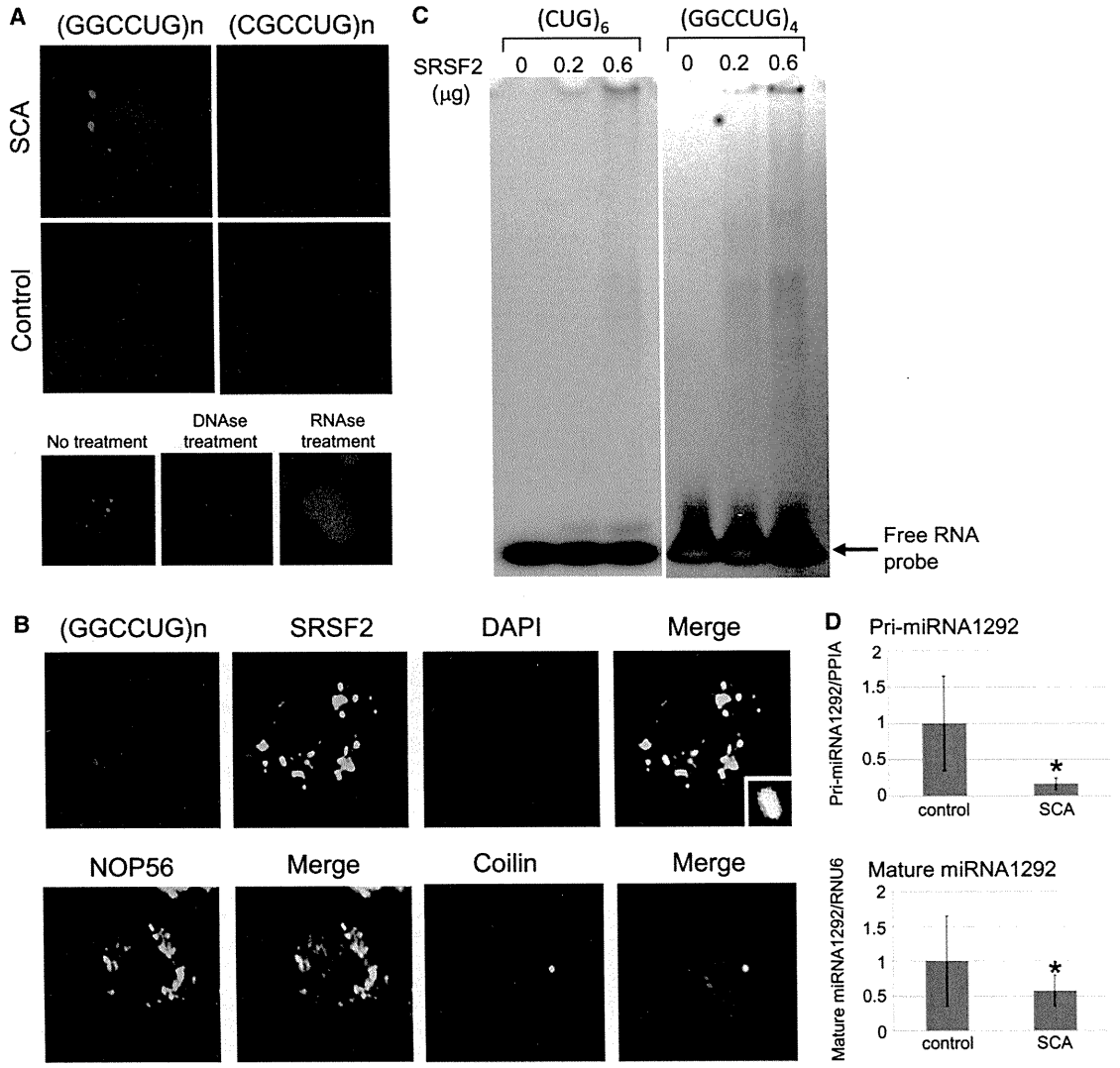


Figure 6. RNA Foci Formation and Decreased Transcription of *MIR1292*

(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)

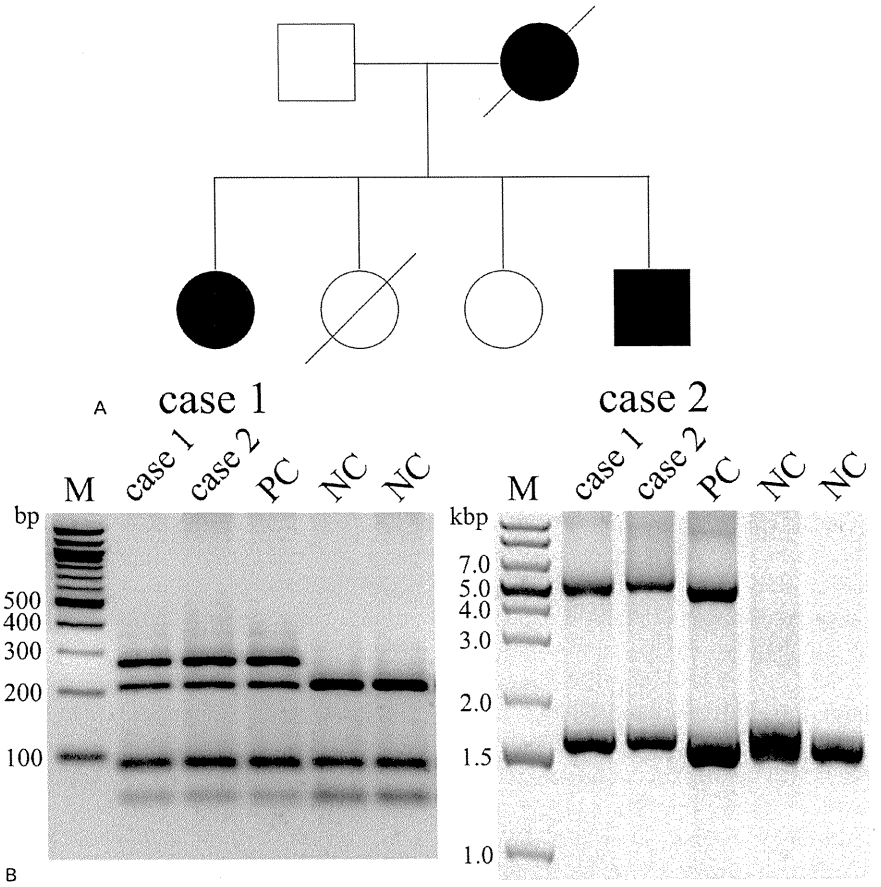


Figure 1 Pedigree and genetic analysis of an SCA31 family (F1). (A) A pedigree of an SCA31 family (F1) evaluated in this study. Symbols for individuals affected by ataxia are shaded. A diagonal line denotes individuals who are deceased. Squares and circles represent males and females, respectively. (B) 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.

Neurological Research ner2346.3d 30/9/10 22:50:58
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

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.

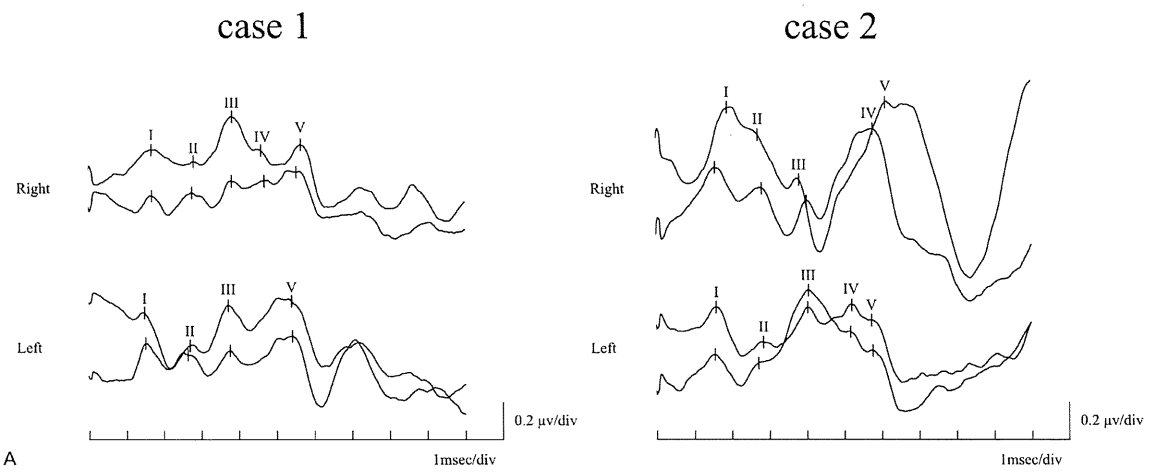
Ages at onset with truncal instability in cases 1 and 2 in an SCA31 family (family F1; Fig. 1A) were 50 and 48 years old, respectively. Neurological examination of both cases revealed quite similar positive findings for truncal and gait ataxia, limb incoordination, ataxic dysarthria, and impaired smooth 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)					Interpeak latency (msec)		
		I	II	III	IV	V	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)