yeast Pol II are homologous to RPC1 and RPC2 of Pol III, respectively. Asn620_Lys652 in RPC2 corresponds to Tyr679_Lys712 in RPB2. The deletion of Asn620_Lys652 (Tyr679_Lys712) would destroy a structural core of RPB2, leading to loss of RPB2 function. In addition, Arg768 (Arg852 in RPB2) interacts with the main-chain carbonyl group of Arg70 of the RPB12 subunit, and Asp926 (Asp1009 in RPB2) interacts with the side chain of Arg48 of the RPB10 subunit of Pol II (Figure 1D). Arg768His (Arg852His) and Asp926Glu (Asp1009Glu) substitutions are considered to disturb these subunit interactions, leading to dysfunction of the polymerase. Therefore, structural prediction suggests that the mutations in *POLR3B* (RPC2) could affect Pol III function. On the other hand, Ile897 and Arg1005 in RPC1 correspond to Val863 and Arg1036 in RPB1, respectively. Ile897 (Val863) has hydrophobic interactions with Leu170 and Pro176 of the RPB5 subunit and with Phe900 (Phe866) of the RPB1 subunit of Pol II (Figure 1E). Ile897Asn (Val863Asn) substitution is likely to disturb this interaction. Arg1005 (Arg1036) stabilizes interaction between RPB1 and RPB8 subunits (Figure 1F). The Arg1005Cys (Arg1036Cys) substitution appears to make this interaction unstable. Thus mutations in *POLR3A* are also predicted to affect Pol III function.

Clinical features of individuals with *POLR3A* or *POLR3B* mutations are presented in Table 1. MRI revealed high-intensity areas in the white matter in T2-weighted images, cerebellar atrophy, and a hypoplastic corpus callosum in all four individuals (Figure 3). Individuals 1 and 2 showed an extremely similar clinical course. They developed normally during their early infancy, i.e., walking unaided at 15 and 14 months, and uttering a few words at 12 and 13 months, respectively. After the age of 3, individual 1 presented with unstable walking and frequent stumbling and falling down, and individual 2 became poor at exercise. They both had severe myopia (corrected visual acuity of 0.7 and 0.5 at most, respectively). They graduated from elementary, junior high, and high schools with poor records, and the intelligence quotient (IQ) of individual 2 was 52 (WAIS-III). In individual 1, unstable walking was prominent at around 18 years, and he could not ride a bicycle because of ataxia; however, he could drive an automobile. Amenorrhea was noted in individual 2, and was successfully treated by hormone therapy. Individual 1 showed several signs of hypogonadism, including absence of underarm and mustache hair, thin pubic hair (Tanner II), and serum levels of testosterone, follicle stimulating hormone, and luteinizing hormone that were below normal for age 27. Neurological examination of both individuals revealed mild horizontal nystagmus, slowing of smooth-pursuit eye movement, and gaze limitation, especially in vertical gazing, hypotonia, mildly exaggerated deep-tendon reflex (patellar and Achilles tendon reflex) with negative Babinski reflex, and cerebellar signs and symptoms, including ataxic speech, wide-based ataxic gait, dysdiadochokinesis, and dysmetria. Clinical information for individual 3 has been reported previously.[6] Addi-
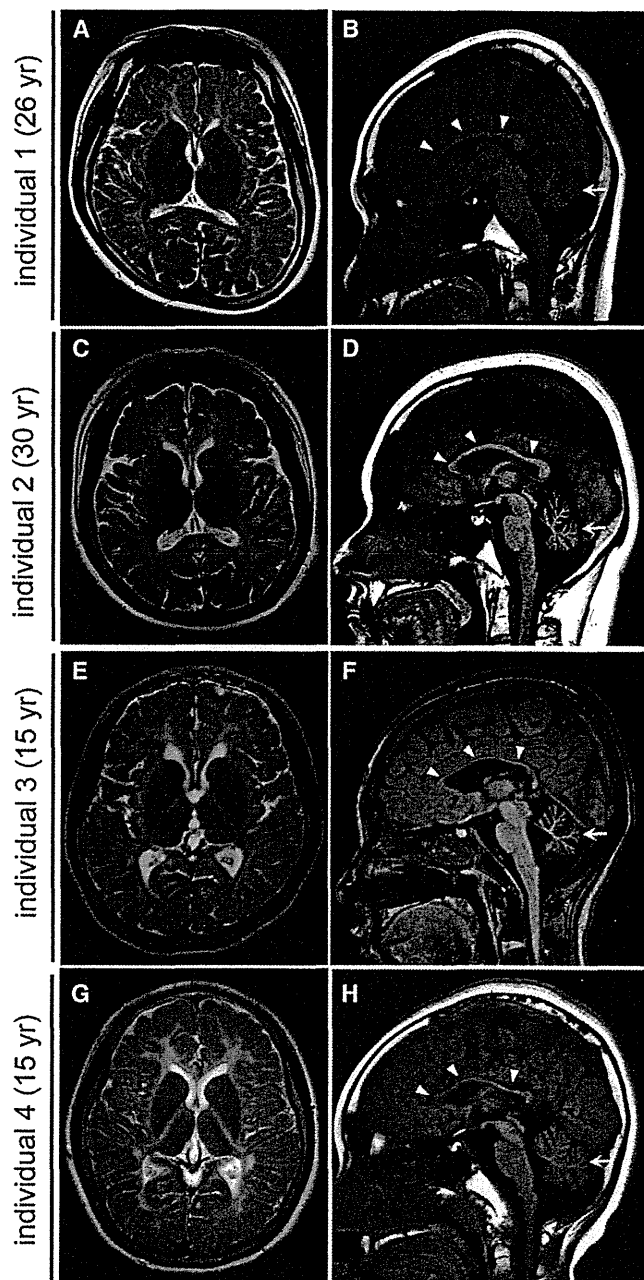

Figure 3. Brain MRI of Individuals with *POLR3B* and *POLR3A* Mutations
(A, C, E, and G) T2-weighted axial images through the basal ganglia. High-intensity areas in the white matter were observed in all individuals.
(B, D, F, and H) T1-weighted midline sagittal images. All the individuals showed hypoplastic corpus callosum (arrowheads) and atrophy of cerebellum (arrows).

tional findings are as follows: slowing of smooth-pursuit eye movement, gaze limitation in vertical gazing, normal auditory brain responses (ABR), cerebral symptoms with mild spasticity, and intellectual disability (an IQ of 43 according to the WISC-III test), and no myopia but hypermetropic astigmatism. She showed no deterioration besides a mild dysphagia and walks herself to a school for the disabled. Individual 4 developed normally during his

159

early infancy, had normal head control at 3 months, was speaking a few words at 12 months, and was walking unaided at 14 months. His parents noted mild tremors around 4 years. He had normal stature, weight, and head circumference. Although he had severe myopia, his eye movement was smooth with no limitation or nystagmus. He had sensory neuronal deafness on the left side. He showed normal muscle tone and had no spasticity or rigidity. His tendon reflexes were slightly elevated with a negative Babinski reflex. Cerebellar signs were noted; expressive ataxic explosive speech, intension tremor, poor finger to nose test, dysdiadochokinesis, dysmetria, and wide-based ataxic gait. His intelligence quotient was 57 (according to the WISC-III test). His peripheral nerve conduction velocity was within the normal range and his ABR showed normal responses on the right side. He suffered motor deterioration around age 14 and became wheelchair bound.

In this study, we successfully identified compound heterozygous mutations in POLR3A and POLR3B in individuals with HCAHC. Very recently, Bernard et al.[12] reported that POLR3A mutations cause three overlapping leukodystrophies, including 4H syndrome, suggesting that HCAHC is, at least in part, within a wide clinical spectrum caused by POLR3A mutations. The p.Arg1005Cys mutation was shared between individual 9 in their report and our individual 4. All 19 individuals with POLR3A mutations showed progressive upper motor neuron dysfunction and cognitive regression. In addition, individual 9 showed abnormal eye movement, hypodontia, and hypogonadism. None of these features were recognized in our individual 4; these differences further support phenotypic variability of POLR3A mutations.[12] Given the phenotypic similarities among 4H syndrome, HCAHC, and H-ABC, there is a possibility that H-ABC is also allelic and caused by recessive mutations in either POLR3A or POLR3B.

Pol III consists of 17 subunits and is involved in the transcription of small noncoding RNAs, such as 5S ribosomal RNA (rRNA), U6 small nuclear RNA (snRNA), 7SL RNA, RNase P, RNase MRP, short interspersed nuclear elements (SINEs), and all transfer RNAs (tRNAs). Pol III-transcribed genes are classified into three types based on promoter elements and transcription factors. 5S rRNA is a solo type I gene. Type II genes include tRNA, 7SL RNA, and SINEs. Type III genes include U6 snRNA, RNase P, and RNase MRP.[18-20] The Pol III system is important for cell growth in yeast, and its transcription is tightly regulated during the cell cycle.[20] In zebrafish, polr3b mutant larvae that have a deletion of 41 conserved amino acids (Δ239-279) from the Rpc2 protein showed a proliferation deficit in multiple tissues, including intestine, endocrine pancreas, liver, retina and terminal branchial arches.[21] In the mutants, the expression levels of tRNA were significantly reduced, whereas the level of 5S rRNA expression was not changed, suggesting that this polr3b mutation can differentially affect Pol III target promoters.[21] RPC2

contributes to the catalytic activity of the polymerase and forms the active center of the polymerase together with the largest subunit, RPC1.[22] Thus, it is reasonable to consider that mutations in POLR3A and POLR3B cause overlapping phenotypes. Indeed, three individuals with POLR3B mutations showed diffuse cerebral hypomyelination, atrophy of the cerebellum and corpus callosum, and abnormal eye movements that overlap with POLR3A abnormalities.[12] Furthermore, two out of three individuals showed hypogonadism, suggesting a common pathological mechanism between POLR3A and POLR3B mutations. In the zebrafish polr3b mutants there were no defects of the central nervous system other than a reduced size of the retina, probably reflecting species differences; however, the reduced level of tRNA in the polr3b mutants raises the possibility that defects of tRNA transcription by Pol III could be a common pathological mechanism underlying POLR3A and POLR3B mutations. Supporting this idea, mutations in two genes involved in aminoacylation activity of tRNA synthetase cause defects of myelination in central nervous system: DARS2 (MIM 610956) and AIMP (MIM 603605).[23,24] In addition, mutations in four genes encoding aminoacyl-tRNA synthetase cause Charcot-Marie-Tooth disease (MIM 613641, 613287, 601472, and 608323), resulting from demyelination of peripheral nerve axons: KARS (MIM 601421), GARS (MIM 600287), YARS (MIM 603623), and AARS (MIM 601065).[25-28] Thus, it is very likely that regulation of tRNA expression is essential for development and maintenance of myelination in both central and peripheral nervous systems.

An interesting clinical feature of POLR3B mutations is the absence of motor deterioration. All three individuals with POLR3B mutations could walk without support at ages 16, 27, and 30, whereas individual 3 with POLR3A mutations had motor deterioration around age 14. Bernard et al.[12] also reported progressive upper motor neuron dysfunction and cognitive regression in individuals with POLR3A mutations. Thus, there is a possibility that phenotypes caused by POLR3A mutations could be more severe and progressive than POLR3B mutant phenotypes. Identification of a greater number of cases with POLR3B mutations is required to confirm this hypothesis.

In conclusion, our data, together with that of a previous report,[12] demonstrate that mutations in Pol III subunits cause overlapping autosomal-recessive hypomyelinating disorders. Establishment of an animal model will facilitate our understanding of the pathophysiology of the multiple defects caused by Pol III mutations.

Supplemental Data

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

Acknowledgments

21. Yee, N.S., Gong, W., Huang, Y., Lorent, K., Dolan, A.C., Maraia, R.J., and Pack, M. (2007). Mutation of RNA Pol III subunit *rpc2/polr3b* Leads to Deficiency of Subunit Rpc11 and disrupts zebrafish digestive development. PLoS Biol. *5*, e312.

22. Werner, M., Thuriaux, P., and Soutourina, J. (2009). Structure-function analysis of RNA polymerases I and III. Curr. Opin. Struct. Biol. *19*, 740–745.

23. Scheper, G.C., van der Klok, T., van Andel, R.J., van Berkel, C.G., Sissler, M., Smet, J., Muravina, T.I., Serkov, S.V., Uziel, G., Bugiani, M., et al. (2007). Mitochondrial aspartyl-tRNA synthetase deficiency causes leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation. Nat. Genet. *39*, 534–539.

24. Feinstein, M., Markus, B., Noyman, I., Shalev, H., Flusser, H., Shelef, I., Liani-Leibson, K., Shorer, Z., Cohen, I., Khateeb, S., et al. (2010). Pelizaeus-Merzbacher-like disease caused by AIMP1/p43 homozygous mutation. Am. J. Hum. Genet. *87*, 820–828.

25. Latour, P., Thauvin-Robinet, C., Baudelet-Méry, C., Soichot, P., Cusin, V., Faivre, L., Locatelli, M.C., Mayençon, M., Sarcey, A., Broussolle, E., et al. (2010). A major determinant for binding and aminoacylation of tRNA(Ala) in cytoplasmic Alanyl-tRNA synthetase is mutated in dominant axonal Charcot-Marie-Tooth disease. Am. J. Hum. Genet. *86*, 77–82.

26. McLaughlin, H.M., Sakaguchi, R., Liu, C., Igarashi, T., Pehlivan, D., Chu, K., Iyer, R., Cruz, P., Cherukuri, P.F., Hansen, N.F., et al. (2010). Compound heterozygosity for loss-of-function lysyl-tRNA synthetase mutations in a patient with peripheral neuropathy. Am. J. Hum. Genet. *87*, 560–566.

27. Antonellis, A., Ellsworth, R.E., Sambuughin, N., Puls, I., Abel, A., Lee-Lin, S.Q., Jordanova, A., Kremensky, I., Christodoulou, K., Middleton, L.T., et al. (2003). Glycyl tRNA synthetase mutations in Charcot-Marie-Tooth disease type 2D and distal spinal muscular atrophy type V. Am. J. Hum. Genet. *72*, 1293–1299.

28. Jordanova, A., Irobi, J., Thomas, F.P., Van Dijck, P., Meerschaert, K., Dewil, M., Dierick, I., Jacobs, A., De Vriendt, E., Guergueltcheva, V., et al. (2006). Disrupted function and axonal distribution of mutant tyrosyl-tRNA synthetase in dominant intermediate Charcot-Marie-Tooth neuropathy. Nat. Genet. *38*, 197–202.

161

REPORT

Exome Sequencing Reveals a Homozygous SYT14 Mutation in Adult-Onset, Autosomal-Recessive Spinocerebellar Ataxia with Psychomotor Retardation

Hiroshi Doi,[1,2] Kunihiro Yoshida,[3] Takao Yasuda,[4] Mitsunori Fukuda,[4] Yoko Fukuda,[5] Hiroshi Morita,[6] Shu-ichi Ikeda,[6] Rumiko Kato,[7] Yoshinori Tsurusaki,[1] Noriko Miyake,[1] Hirotomo Saitsu,[1] Haruya Sakai,[1] Satoko Miyatake,[1] Masaaki Shiina,[8] Nobuyuki Nukina,[9] Shigeru Koyano,[2] Shoji Tsuji,[5] Yoshiyuki Kuroiwa,[2] and Naomichi Matsumoto[1,*]

Autosomal-recessive cerebellar ataxias (ARCAs) are clinically and genetically heterogeneous disorders associated with diverse neurological and nonneurological features that occur before the age of 20. Currently, mutations in more than 20 genes have been identified, but approximately half of the ARCA patients remain genetically unresolved. In this report, we describe a Japanese family in which two siblings have slow progression of a type of ARCA with psychomotor retardation. Using whole-exome sequencing combined with homozygosity mapping, we identified a homozygous missense mutation in SYT14, encoding synaptotagmin XIV (SYT14). Expression analysis of the mRNA of SYT14 by a TaqMan assay confirmed that SYT14 mRNA was highly expressed in human fetal and adult brain tissue as well as in the mouse brain (especially in the cerebellum). In an in vitro overexpression system, the mutant SYT14 showed intracellular localization different from that of the wild-type. An immunohistochemical analysis clearly showed that SYT14 is specifically localized to Purkinje cells of the cerebellum in humans and mice. Synaptotagmins are associated with exocytosis of secretory vesicles (including synaptic vesicles), indicating that the alteration of the membrane-trafficking machinery by the SYT14 mutation may represent a distinct pathomechanism associated with human neurodegenerative disorders.

Hereditary ataxias are genetically heterogeneous neurological disorders: autosomal-dominant, autosomal-recessive, X-linked, and mitochondrial types are known. Among ataxias, spinocerebellar ataxia (SCA) is relatively common and involves the cerebellum, brainstem, or spinocerebellar long tracts.[1] Autosomal-recessive cerebellar ataxias (ARCAs) are generally associated with diverse neurological and nonneurological attributes, resulting in complex phenotypes. ARCAs include congenital nonprogressive ataxias and progressive ataxias such as SCAs.[2] The clinical onset of ARCAs usually occurs before the age of 20, even if congenital types are excluded.[1,3,4] Currently, more than 20 defective genes have been identified in ARCAs.[2,5,6] These genes have variable recognized functions, including those involving mitochondrial energy generation, cellular metabolisms, DNA repair, chaperone-mediated protein folding, RNA processing, and ion channels.[1,3,6] Approximately half of the patients with ARCAs remain genetically unresolved.[4] Therefore, more investigations of ARCAs are required. In this study, we describe a Japanese family with two siblings showing psychomotor retardation and the slowly progressive type of SCA without involvement of pyramidal tracts or peripheral nerves. Exome sequencing combined with homozygosity mapping successfully identified a causative mutation.

Clinical information and blood materials were obtained from the family members after written informed consent was secured. Experimental protocols were approved by IRBs of the Yokohama City University and the Shinshu University. Among the children of first-cousin parents, two siblings (IV-3 and IV-4) were found to be affected, whereas the other two (IV-1 and IV-2) were healthy (Figure 1A). No similar patients were recognized within the family. IV-3 had mild psychomotor retardation from childhood. He found a job after graduating from an ordinary junior high school. At 35 years of age, he lost his job for social reasons. Although he had some gait disturbances from childhood, he could independently go shopping and walk a dog even after leaving his occupation. At the age of ~56, he developed obvious gait unsteadiness and began to stumble frequently. At 58, he started to choke on food. These symptoms gradually worsened, and he sought medical examination at 59 years of age. He displayed disturbances of smooth-pursuit eye movements, dysarthria, mild limb ataxia, and moderate truncal ataxia. His muscle tone was normal, and no involuntary
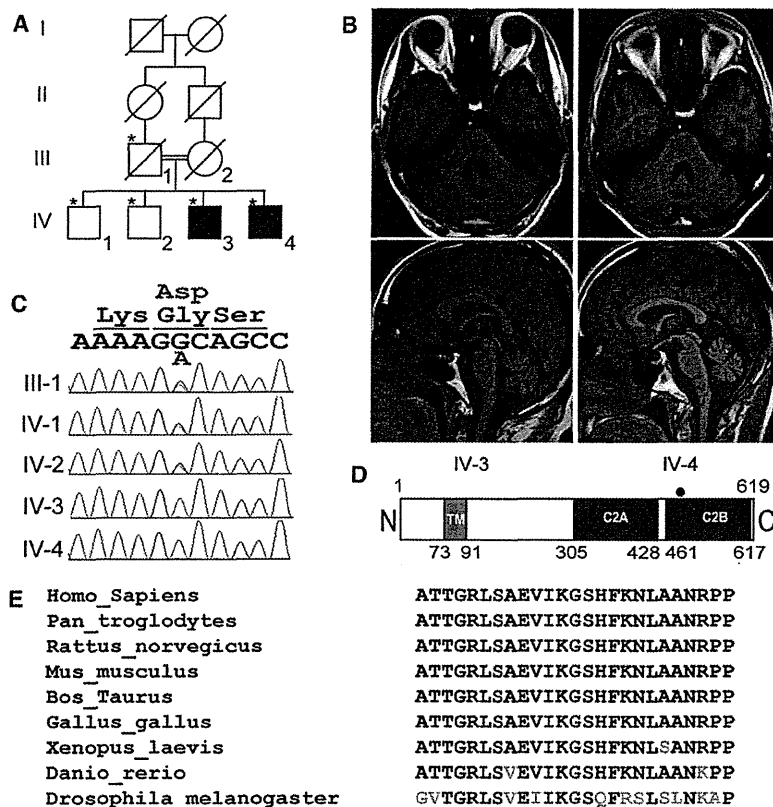
Figure 1. Familial Pedigree, Brain MRI of Patients, and the *SYT14* Mutation Identified

(A) Familial pedigree of the patients with autosomal-recessive spinocerebellar ataxia. *An asterisk indicates members whose genomic DNA was available for this study.

(B) Brain MRI of IV-3 at 59 years of age (left panels) and IV-4 at 56 years of age (right panels). Axial (upper panels) and sagittal (lower panels) sections of a T1-weighted image are shown.

(C) Electropherograms of unaffected (III-1, IV-1, and IV-2) and affected (IV-3 and IV-4) members, who show the mutation.

(D) Schematic presentation of SYT14. The red dot indicates the location of the mutation in the C2B domain.

(E) The missense mutation occurred at an evolutionarily conserved amino acid (in red).

movements were observed. Laboratory examination, including analysis of serum albumin, vitamin E, and α-fetoprotein, was normal. A nerve-conduction study (NCS) indicated no neuropathy. No retinitis pigmentosa was recognized by ophthalmologic evaluation. Brain magnetic resonance imaging (MRI) revealed mild atrophy of the cerebellar vermis and hemispheres but no apparent atrophy of the brain stem or the cerebral cortex. (Figure 1B, left panels).

Similar to IV-3, IV-4 also had psychomotor retardation from childhood, but this retardation was more severe than that of IV-3. After graduation from a school for disabled children at the age of 15, he entered a facility for disabled people. He showed gait disorder, but he was able to walk without a cane. At an age of ~50, his gait disturbance worsened, and he went for a medical check at a hospital when he was 56 years old. He displayed disturbance of smooth-pursuit eye movements, gaze-evoked horizontal nystagmus, dysarthria, mild limb ataxia, and moderate truncal ataxia. No involuntary movements were observed. His laboratory tests, including those for serum albumin, vitamin E, and α-fetoprotein, were normal. NCS was normal. A brain MRI was similar to that of IV-3 (Figure 1B, right panels). The clinical manifestations of these two patients are summarized in Table 1.

To search for the disease locus, we conducted genome-wide SNP genotyping of III-1, IV-1, IV-2, IV-3, and IV-4 by using the Genome-wide Human SNP Array 6.0 (SNP

6.0 array) (Affymetrix, Inc., Santa Clara, CA) according to the manufacturer's instructions. Then, SNP 6.0 array data were subjected to homozygosity mapping with HomozygosityMapper software.[7] For the linkage analysis, a subset of 7339 SNPs with high heterozygosity (mean heterozygosity 0.49) was extracted from the SNP 6.0 array data with the program Linkdatagen, for which the bin size was set to 0.5cM and the allele frequency of the Japanese population was used.[8] The multipoint LOD score was calculated with Allegro version 2[9] on the basis of the model of autosomal-recessive and X-linked-recessive inheritance, respectively. In both models, complete penetrance and a disease-allele frequency of 0.0001 were adopted. The homozygosity mapping revealed a total of three regions, which together were approximately 11.35 Mb in size, as candidate loci, where genes known to be mutated in ARCA did not exist (Table 2). In the model of autosomal-recessive inheritance, a total of ten regions with a LOD score greater than 1.5 in the multipoint linkage analysis were identified (Table S1). The three homozygous regions in accordance with the linked regions still survived as candidate regions. On the basis of X-linked recessive inheritance, a total of three regions with positive LOD scores (maximum LOD score = 0.9031) were highlighted; together, these three regions were approximately 101.03 Mb (Table S1).

To find a gene mutation within the candidate loci, we performed whole-exome sequencing on IV-3 and IV-4. Three micrograms of genomic DNA was processed with the SureSelect Human All Exon Kit v.1 (approximately 180,000 exons covering 38 Mb of the CCDS database) (Agilent Technologies, Santa Clara, CA) according to the manufacturer's instructions. Captured DNAs were sequenced on an Illumina GAIIx (Illumina, San Diego, CA) with 76 bp pair-end reads. Of the possible eight lanes of the flow cell, two lanes for IV-3 and three lanes for IV-4 (Illumina) were used. Image analysis and base calling

Table 1. Clinical Features of the Patients		
Clinical Features	IV-3	IV-4
Age at present	61	58
Sex	male	male
Age of obvious ataxia	56	around 50
Mental retardation	mild	moderate
Ocular apraxia	no	no
Ophthalmoplegia	no	no
Nystagmus	no	+
Dysarthria	+	+
Truncal ataxia	++	++
Limb ataxia	+	+
Extrapyramidal signs	no	no
Involuntary movements	no	no
Sensory involvements	no	no
Tendon reflex	normal-increased	normal-decreased
Plantar responses	normal	normal
Peripheral neuropathy	no	no
Pes cavus	no	no
SARA[a]	12/40	15/40
Cerebellar atrophy on MRI	+	+
others	normal level of serum albumin, vitamin E, and α-fetoprotein	normal level of serum albumin, vitamin E, and α-fetoprotein

[a] SARA: Scale for the assessment and rating of ataxia.[32]

Table 2. Regions of Homozygosity			
Chromosome	Chromosomal Position (rsID)	Size (Mb)	LOD
1	207226930 (rs2761781)–213992561 (rs1857229)	6.77	2.0537
4	181929079 (rs918401)–185188999 (rs7690914)	3.26	2.0554
22	45676443 (rs3905396)–47003473 (rs2013591)	1.33	2.0545

Regions of homozygosity were identified by HomozygosityMapper, and the LOD scores were calculated by multipoint linkage analysis, for which SNPs were extracted from SNP 6.0 array data via Linkdatagen.

were performed by sequence control software (SCS) real-time analysis (Illumina) and CASAVA software v1.6 (Illumina). Reads were aligned to the human reference genome sequence (UCSC hg18, NCBI build 36.1) via the ELAND v2 program (Illumina). Coverage was calculated statistically with a script created by BITS (Tokyo, Japan). Approximately 71 million reads from IV-3 and 148 million reads from IV-4 (these numbers of reads passed quality-control [Path Filter]) were mapped to the human reference genome with Mapping and Assembly with Qualities (MAQ)[10] and NextGENe software v2.00 (SoftGenetics, State College, PA) under the default settings. MAQ aligned 59,491,138 and 126,159,746 reads to the whole genome for IV-3 and IV-4, respectively. A script created by BITS was used for extraction of SNPs and indels from the alignment data; dbSNP build 130 served as a reference for registered SNPs. A consensus quality score of 40 or more was adopted for the SNP analysis in MAQ. Coverage analysis revealed that 65.0% (IV-3) and 71.3% (IV-4) of the coding sequences (CDS) were completely covered (100%), and 77.7% (IV-3) and 80.3% (IV-4) of CDS were mostly covered by reads (90% or more) through the whole genome. 79.0% (IV-3)

and 79.7% (IV-4) of total CDS were covered by ten reads or more (50 reads or more in 66.4% and 77.1%, respectively).

To identify the pathogenic mutation, we adopted a prioritization scheme, which has been used in recent studies.[11–13] First, we excluded the variants registered in dbSNP130 from all the detected variants and then picked up homozygous mutations and variants in coding regions and the intronic regions within 50 bp from coding sequences. Of the homozygous mutations and variants, we focused on those within the candidate regions. As a result, only two missense mutations or variants, p.Gly484Asp (c.1451G>A) (NM_001146261.1) in exon 8 of *SYT14* (1q32.2, [MIM 610949]) and p.Gln4203Arg (c.12608A>G) (NM_206933.2) in exon 63 of *USH2A* (1q41, [MIM 608400]) remained as candidates for both cases (Table S2). Sanger sequencing with ABI 3500xL (Life Technologies, Carlsbad, CA) confirmed that the c.1451G>A of *SYT14* was homozygous in IV-3 and IV-4 and heterozygous in III-1 (father), IV-1, and IV-2, whereas the c.12608A>G of *USH2A* was homozygous in IV-2 as well as IV-3 and IV-4 (Figure 1C and data not shown). The *SYT14* missense mutation occurred at an evolutionarily conserved amino acid among different species and resides in the second C2 (C2B) domain (Figures 1D and 1E). In silico analysis incorporating different tools, including Polyphen, Polyphen2, SIFT, and Align GVGD, consistently indicated that the change was damaging (Table S3). The mutation was not detected in 576 Japanese control chromosomes, indicating that the mutation is very rare. On the basis of the X-linked recessive model, no pathological hemizygous mutation of protein-coding genes was detected in the possible candidate loci (Table S4).

We considered the *SYT14* mutation to be the causative agent and used the Sanger method to conduct mutation screening of all the coding regions of *SYT14* in 65 simplex SCA cases and 37 SCA familial cases, including three with autosomal-recessive inheritance. Only p.Gly183Glu (c.548G>A) was found in one family with autosomal-dominant SCA; however, the change was not consistent with the SCA phenotype in the family (Table S3). Thus, we could not detect any other pathological changes in *SYT14*. This was probably due to the small number of cases tested.

164

Synaptotagmin XIV (SYT14), which is encoded by *SYT14*, is a member of the synaptotagmins (SYTs), which are membrane-trafficking proteins, and SYT14 is conserved across many organisms.[14] Although the original report indicated that *SYT14* was not expressed in mouse brain,[14] multiple lines of evidence, including from the Allen brain Atlas, suggest that *SYT14* is expressed in the central nervous system (CNS) of the fly, mouse, and human brains.[15,16] To confirm *SYT14* expression in the CNS, we performed TaqMan quantitative real-time PCR analysis with cDNAs of adult human tissue (Human MTC Panel I, #636742) (Clontech Laboratories, Mountain View, CA), fetal human tissue (Human Fetal MTC Panel, #636747) (Clontech Laboratories), mouse tissue (Mouse MTC Panel I, #636745) (Clontech Laboratories), and various regions of the mouse brain (GSMBRSET) (NIPPON Genetics, Tokyo, Japan) as templates. Predesigned TaqMan probe sets for human *SYT14* (Hs00950169_m1), mouse *Syt14* (Mm00805319_m1), human *β-actin* (*ACTB*, 4326315E), and mouse *Actb* (43522341E) from Applied Biosystems were used. PCR reactions (total volume of 20 μl) contained 10 μl of the TaqMan Gene Expression Master Mix (Applied Biosystems), 1 μl of 20 × TaqMan reagents for *ACTB/Actb* and *SYT14/Syt14*, and 1 μl of cDNA (containing 1 ng cDNA in MTC panels and 25 ng cDNA in GSMBRSET) as the template. PCR was performed on a Rotor-Gene Q (QIAGEN, Valencia, CA) as follows: 2 min at 50°C and 10 min at 95°C, then 40 cycles of 95°C for 15 s and 60°C for 1 min. Expression levels were calculated with the Rotor-Gene Q Series Software (QIAGEN) by the $2^{-\Delta\Delta Ct}$ method. The cycling threshold (Ct) of the target gene was compared with the Ct of *ACTB* cDNA, and ΔCt was expressed as Ct of *SYT14* − Ct of *ACTB*. ΔΔCt was expressed as ΔCt of the control sample − ΔCt of each sample, and relative concentration was determined as $2^{-\Delta\Delta Ct}$. Expression in the kidney and the cerebral cortex was used as the control in Figures 2A–2D. *SYT14* was predominantly expressed in human adult and fetal brain tissues (Figures 2A and 2B). Even in mice, substantial expression in the brain was confirmed (but, not predominant) (Figure 2C). Among various brain regions in mice, *SYT14* was mostly expressed in the cerebellum (Figure 2D).

Intracellular distribution of SYT14 in cultured cells was investigated. The full-length *SYT14* PCR product amplified from human brain cDNA (MHS4426-99239810, Open Biosystems, Huntsville, AL) was used as a template and subcloned into pDONR221 (the entry vector of Gateway system, Invitrogen). We used site-directed mutagenesis to produce the *SYT14* mutant and variants by using a mutagenesis kit (Toyobo, Osaka, Japan). Variants include c.611C>T and c.810_812del, which are registered in dbSNP130, and c.548G>A, which was detected in an SCA patient with autosomal-dominant inheritance but did not segregate with the phenotype, indicating that it is nonpathogenic (Table S3). All constructs were verified by DNA sequencing. Each construct was recloned into the pEF-DEST51 mammalian expression vector (Invitrogen) and transfected to COS-1 cells with the FuGENE[R]6 transfection reagent (Roche Applied Science, Mannheim, Germany) according to the manufacturer's instructions. Localization of the mutant (p.Gly484Asp) was clearly different from that of the wild-type and other (normal) variants. Whereas the wild-type and other variants were localized to the perinuclear and submembranous regions, p.Gly484Asp was localized in the cytoplasm (significant amounts were in the perinuclear region) but formed a characteristic reticular pattern without showing any submembranous distribution (Figures 2E and S2B). Confocal microscopic analysis showed that the p.Gly484Asp mutant was colocalized with an endoplasmic reticulum (ER) marker, protein disulfide isomerase (PDI), throughout the cells, whereas the wild-type colocalized with PDI dominantly in perinuclear regions (Figure 2F). immunoblot analysis combined with subcellular fractionation of the transfected cells further confirmed that the mutant was distributed differently from the wild-type. The wild-type and the mutant (p.Gly484Asp) were distributed in the nucleus and Golgi apparatus fractions; however, only the mutant was detected in microsome fractions containing ER fragments together with an ER membrane marker, calnexin (Figure S1).[17] These data suggest that improper folding of the mutant protein results in abnormal retention in the ER.

To investigate the effect of the p.Gly484Asp mutation in the C2B domain on phospholipid binding activity, we amplified cDNA of C2B domains from the wild-type and the p.Gly484Asp mutant from SYT14-expressing vectors by using the following primers: sense, 5′-GGATCCGAAA GTACATCCTCATGTCA-3′; and antisense, 5′-TCATGAC TCTAGCAACGCAT-3′. We then recloned the cDNA into *Escherichia coli* (*E. coli*) expressing vector (pGEX-4T-3). The C2B domain of SYT14 fused to glutathione S-transferase (GST) was expressed in *E. coli* JM109 and purified by standard protocols. Both GST-SYT14-C2B (WT) and GST-SYT14-C2B (p.Gly484Asp) could be mostly purified of contamination by degradation products, but the amount of GST-SYT14-C2B (p.Gly484Asp) obtained was at least four times smaller than that of GST-SYT14-C2B (WT) (data not shown). Liposome (phosphatidylcholine and phosphatidylserine, 1:1, w/w) cosedimentation assay with purified GST-SYT14-C2B was performed as described previously.[18] The result showed that the SYT14-C2B (p.Gly484Asp) bound liposomes similarly to SYT14-C2B (WT) (Figure 2G), indicating that the p.Gly484Asp mutation had no effect on the Ca^{2+}-independent phospholipid-binding activity of the SYT14-C2B domain.

The Allen Mouse Brain Atlas indicates that *Syt14* is expressed in Purkinje cells of the cerebellum in mice; however, SYT14 localization has not been fully investigated.[15] A rabbit polyclonal anti-SYT14 antibody (Ab-SYT14) was generated for immunoblotting and immunocytochemistry (Operon Biotechnologies, Tokyo, Japan) (Figure S2). Immunohistochemical analysis of mouse and human brains was
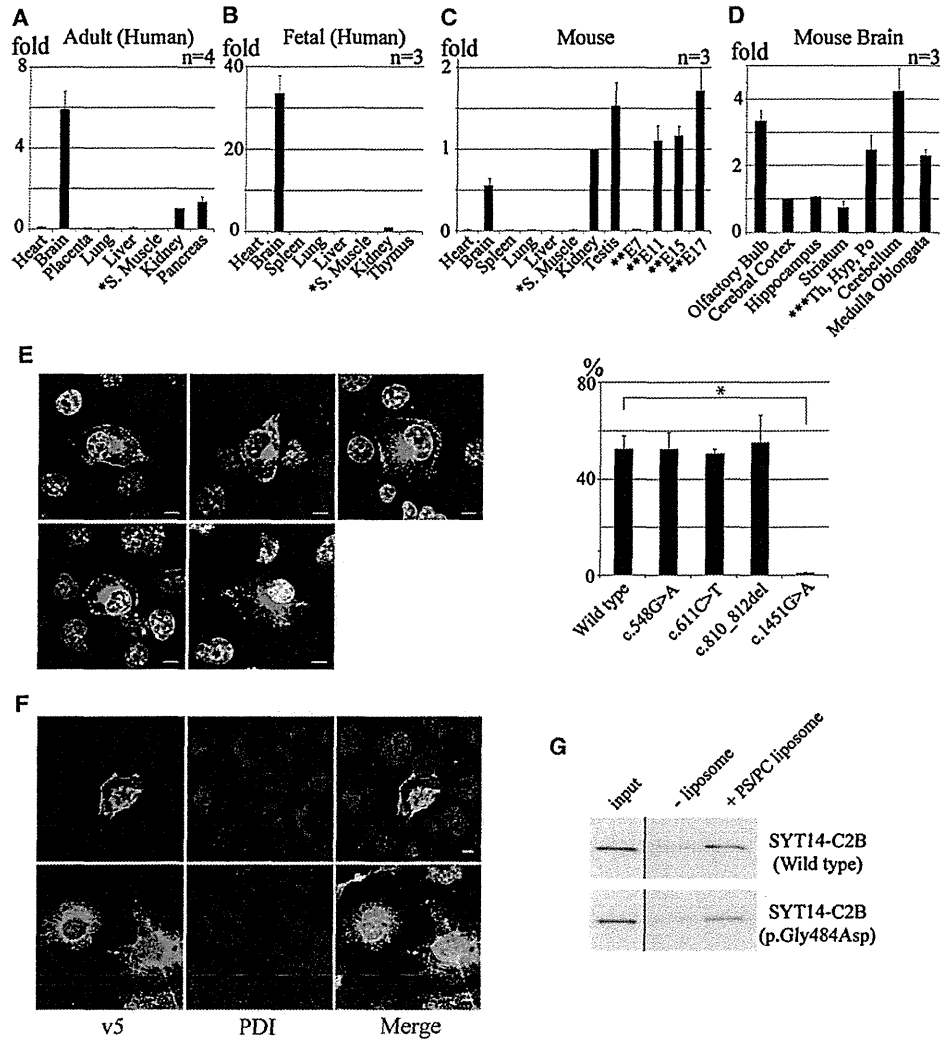
Figure 2. Expression Studies of *SYT14/Syt14* cDNA in Human and Mouse Tissues and Localization of SYT14 in Transfected COS-1 Cells

(A–D) The results of a TaqMan quantitative real-time PCR assay in which the first-strand cDNA of human adult tissues (A), human fetal tissues (B), mouse tissues (C), and various regions of mouse brain (D) were used as templates. The relative cDNA concentrations were determined from cDNA concentrations of the kidney (human adult tissues, human fetal tissues, and mouse tissues) or cerebral cortex (various regions of the mouse brain). Error bars represent the standard deviation. *S. Muscle indicates skeletal muscle. **E7, **E11, **E15, and **E17 indicate mouse embryos at 7, 11, 15, and 17 days of embryonic development, respectively. ***Th, Hyp and Po indicate thalamus, hypothalamus, and pons.

(E) Immunocytochemistry of COS-1 cells transfected with expression vectors of v5/His-tagged wild-type (upper left), p.Gly183Glu (c.548G>A) (upper middle), p.Pro203Leu (c.611C>T) (upper right), p.Glu270del (c.810_812del) (lower left), or p.Gly484Asp (c.1451G>A) (lower middle) SYT14. The SYT14 was detected with the anti-v5 antibody (Alexa fluor 488 as the secondary antibody). Nuclei were stained (white) with 4′,6-diamidino-2-phenylindole (DAPI). The horizontal bars indicate 10 μm. The bar graph indicates the ratio of the cells in which overexpressed proteins were accumulated in submembranous regions. A total of 120 cells per each transfectant in triplicated experiments were counted. Submembranous localization of the mutant (p.Gly484Asp) was mostly unseen, in contrast to the wild-type (*p < 0.001).

(F) Immunocytochemical analysis of COS-1 cells transfected with expression vectors of v5/His-tagged wild-type (upper panels) or the p.Gly484Asp mutant (lower panels). The SYT14 was detected with the anti-v5 antibody (Alexa fluor 488 as the secondary antibody), and PDI (protein disulfide isomerase) was visualized with an anti-PDI antibody (Alexa fluor 546 as the secondary antibody). Nuclei were stained (white) with DAPI. The scale bar represents 10 μm. The anti-v5 and anti-PDI antibodies and the Alexa-488-conjugated secondary antibody were all used at a dilution of 1:1000.

(G) Phospholipid binding activity of the C2B domain of the wild-type SYT14 and the p.Gly484Asp mutant. Liposomes and GST-fusion proteins (2 μg) were incubated in 50 mM HEPES-KOH (pH 7.2) in the presence of 2 mM EGTA for 15 min at room temperature. After centrifugation at 12,000 × g for 10 min, the supernatants (non-binding fraction) and pellets (phospholipid-binding fraction) were separated as described previously.[18] The pelleted samples and input samples (100 ng) were subjected to 10% SDS-PAGE followed by immunoblotting with horseradish peroxidase-conjugated anti-GST antibody (Santa Cruz Biotechnology, Santa Cruz, CA).
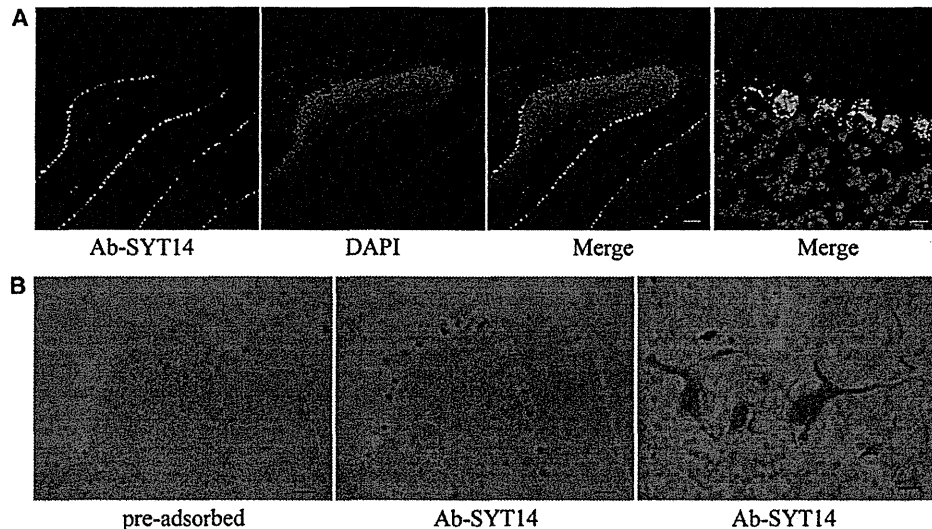
Figure 3. Selective Localization of SYT14 in Purkinje Cells of the Cerebellum in Mice and Humans
(A) Immunohistochemical analysis with the Ab-SYT14 antibody of the cerebellum from an adult mouse at 12 weeks of age. Nuclei were stained with DAPI (the scale bar represents 100 μm). A magnified image is shown in the first right panel (the scale bar represents 10 μm).The Ab-SYT14 antibody (0.9 mg/dl) was used at a dilution ration of 1:2000, and the Alexa-488-conjugated secondary antibody dilution was 1:1000.
(B) Immunohistochemical analysis with the Ab-SYT14 antibody of the cerebellum from the human control. Ab-SYT14 antibodies were preincubated with (left panel) or without (middle panel) peptide antigen before immunostaining. Nuclei were stained with hematoxylin (scale bars represent 100 μm). A magnified image is shown in the right panel (the scale bar represents 20 μm). The Ab-SYT14 antibody (0.9 mg/dl) was used at a dilution of 1:500.

performed with Ab-SYT14, as previously described.[19–21] Mouse brain sections were prepared at the RIKEN Brain Science Institute. Mouse experimental protocols were approved by the animal experiment committee of the RIKEN Brain Science Institute. The frozen brain of C57BL/6J mouse was mounted in Tissue-Tek and sliced to 10 μm sections with a freezing microtome. A human adult brain specimen was obtained through the postmortem examination of a brain from a control subject without neurodegenerative disorders. Informed consent was obtained from the family on the basis of the IRB-approved protocol of Yokohama City University School of Medicine. The human brain was fixed in 10% formalin and cut into 1-cm-thick slices. Sliced tissues were embedded in paraffin wax, and 5 μm sections were immunostained with primary antibodies and visualized with the Vectastain ABC kit (Vector Laboratories, Burlingame, CA). Selective localization of SYT14/Syt14 in Purkinje cells of the mouse cerebellum (Figure 3A) and human cerebellum (Figure 3B) were recognized, indicating that SYT14 plays an important role in the cerebellum. These data are in agreement with a scenario in which the SYT14 mutation causes cerebellar degeneration in this family.

In this study, only one p.Gly484Asp mutation of SYT14 was identified in association with SCA. Quintero-Rivera et al.[16] previously described a 12-year-old female with cerebral atrophy, absence seizures, developmental delay with a WISC III score of 58 for full IQ, and de novo t(1;3)(q32.1;q25.1) disrupting SYT14. Her brain MRI showed diffuse cerebral atrophy, including that of the cere-

bellar hemisphere and vermis. Although the inheritance modes are different (recessive impact on our family and dominant on the female patient), mild to moderate mental retardation and cerebellar atrophy are common among patients with SYT14 abnormalities. It will be important to assess the future phenotype of the female patient studied by Quintero-Rivera et al.[16]

Relatively common ARCAs in Japan include ataxia, early-onset; oculomotor apraxia, hypoalbuminemia/ataxia-oculomotor apraxia 1 (EAOH/AOA1 [MIM 208920]); ataxia-oculomotor apraxia 2 (AOA2 [MIM 606002]); spastic ataxia; Charlevoix-Saguenay type (SACS [MIM 270550]); ataxia with isolated vitamin E deficiency (AVED [MIM 277460]); and ataxia-telangiectasia (AT [MIM 208900]). (Friedrich ataxia 1 [FRDA (MIM 229300)] has never been described in the Japanese population.) In this family, patients never showed ocularmotor apraxia, spasticity, peripheral neuropathy, retinal abnormality, immunological abnormality, or other systemic involvements. As an adult-onset type of pure ARCA, SYNE1-related ARCA (also known as spinocerebellar ataxia, autosomal-recessive 8; SCAR8 [MIM 610743]) is found to be caused by mutations of the gene encoding synaptic nuclear envelope protein 1.[22] Furthermore, these patients were not associated with psychomotor retardation. Thus, SYT14-mutated ARCA, described here, should be categorized to a distinct type of ARCA.

SYTs is a large family of transmembrane proteins associated with exocytosis of secretory vesicles (including synaptic vesicles).[23] The mammalian SYT family is composed

of 17 members. SYTs are anchored to the secretory vesicles via a single transmembrane domain (TM) close to its N terminus and have tandem cytoplasmic domains, C2A and C2B.[24] Among SYTs, SYT1 (MIM 185605) is involved in neurotransmitter release and has been intensively studied. The crystal structure of the C2 domains consists of a compact eight-stranded β-barrel with two protruding loops (loops 1 and 3) that form the Ca^{2+}-binding pockets.[25] SYT1 binds three and two Ca^{2+} ions via loops 1 and 3 of C2A and C2B, respectively. Ca^{2+} binding triggers the rapid penetration of the C2 domains into membranes harboring negatively charged phospholipids. Ca^{2+} also promotes SYT1 binding to t-SNAREs (target-membrane-soluble N-ethylmaleimide-sensitive factor attachment protein receptors). SYT1 is a key sensor for evoked and synchronous neurotransmitter release in many classes of neurons.[23] SYT14 also has TM, C2A, and C2B domains, but it has no conserved Ca^{2+}-binding motif that includes the conserved aspartic acid residues in loops 1 and 3 of C2A and C2B.[26] Although the roles of SYTs as Ca^{2+} sensors have been studied extensively, little is known about Ca^{2+}-independent SYTs, which might inhibit the SNARE-catalyzed fusion in both the absence and presence of Ca^{2+}.[27] Recently, Zhang et al.[28] suggested that Ca^{2+}-independent SYT4 (MIM 600103) negatively regulates exocytosis, regardless of its inability to induce Ca^{2+}-dependent exocytosis.

SYT14 has phospholipid-binding activity that is Ca^{2+} independent.[14] The glycine residue mutated in the family is located around the C2B domain loop 1, which plays an important role in binding to phospholipids in SYT1.[25] We confirmed that, compared to the wild-type, the mutation did not alter the binding activity of SYT14 to phospholipids. In an overexpression system, wild-type SYT14 as well as normal variants were distributed in the cytoplasm close to the plasma membrane, showing in-line accumulation along with the membrane. In contrast, the p.Gly484Asp mutant showed a different (reticular) distribution pattern. In the ER, several cotranslational and posttranslational modifications that are required for the correct folding of transmembrane and secretory proteins take place.[29,30] Incompletely folded proteins are generally excluded from ER exit sites.[29] The fact that the p.Gly484Asp was not properly transferred from the ER suggests that the mutant protein might not fold correctly. The lower yield of the mutant protein as compared to the wild-type in the bacterial expression system we performed also supports the improper folding of the mutant. Abnormal distribution in the ER might result in the loss of function of SYT14 or in ER dysfunction.

In conclusion we have shown that SYT14 is localized specifically in Purkinje cells of mouse and human cerebellum. The results strongly support the involvement of SYT14 in the pathogenesis of SCA and are consistent with the atrophy of the cerebellum seen in both patients. A possible relationship between SYTs and neurodegeneration has been suggested previously,[31] and here we provide data that support the idea that disruption of an SYT protein is involved in human neurodegeneration and that exocytosis machinery can be involved in one of the pathomechanisms of neurodegeneration.

Supplemental Data

Supplemental Data include two figures and five tables and can be found with this article online at http://www.cell.com/AJHG/.

Acknowledgments

Web Resources

The URLs for data presented herein are as follows:

Align GVGD, http://agvgd.iarc.fr/
Allen Human brain Atlas, http://human.brain-map.org/
Allen Mouse Brain Atlas, http://mouse.brain-map.org/
HomozygosityMapper, http://www.homozygositymapper.org/
Online Mendelian Inheritance in Man (OMIM), http://www.omim.org/
PolyPhen, http://genetics.bwh.harvard.edu/pph/
PolyPhen2, http://genetics.bwh.harvard.edu/pph2/
SIFT, http://blocks.fhcrc.org/sift/SIFT.html

Rerefences

1. Fogel, B.L., and Perlman, S. (2007). Clinical features and molecular genetics of autosomal recessive cerebellar ataxias. Lancet Neurol. 6, 245–257.
2. Palau, F., and Espinós, C. (2006). Autosomal recessive cerebellar ataxias. Orphanet J. Rare Dis. 1, 47.
3. Embiruçu, E.K., Martyn, M.L., Schlesinger, D., and Kok, F. (2009). Autosomal recessive ataxias: 20 types, and counting. Arq. Neuropsiquiatr. 67, 1143–1156.
4. Anheim, M., Fleury, M., Monga, B., Laugel, V., Chaigne, D., Rodier, G., Ginglinger, E., Boulay, C., Courtois, S., Drouot, N., et al. (2010). Epidemiological, clinical, paraclinical and

molecular study of a cohort of 102 patients affected with autosomal recessive progressive cerebellar ataxia from Alsace, Eastern France: Implications for clinical management. Neurogenetics *11*, 1–12.

5. Manto, M., and Marmolino, D. (2009). Cerebellar ataxias. Curr. Opin. Neurol. *22*, 419–429.

6. Vermeer, S., Hoischen, A., Meijer, R.P., Gilissen, C., Neveling, K., Wieskamp, N., de Brouwer, A., Koenig, M., Anheim, M., Assoum, M., et al. (2010). Targeted next-generation sequencing of a 12.5 Mb homozygous region reveals *ANO10* mutations in patients with autosomal-recessive cerebellar ataxia. Am. J. Hum. Genet. *87*, 813–819.

7. Seelow, D., Schuelke, M., Hildebrandt, F., and Nurnberg, P. (2009). HomozygosityMapper—An interactive approach to homozygosity mapping. Nucleic Acids Res. *37*, W593–W599.

8. Bahlo, M., and Bromhead, C.J. (2009). Generating linkage mapping files from Affymetrix SNP chip data. Bioinformatics *25*, 1961–1962.

9. Gudbjartsson, D.F., Thorvaldsson, T., Kong, A., Gunnarsson, G., and Ingolfsdottir, A. (2005). Allegro version 2. Nat. Genet. *37*, 1015–1016.

10. Li, H., Ruan, J., and Durbin, R. (2008). Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res. *18*, 1851–1858.

11. Gilissen, C., Arts, H.H., Hoischen, A., Spruijt, L., Mans, D.A., Arts, P., van Lier, B., Steehouwer, M., van Reeuwijk, J., Kant, S.G., et al. (2010). Exome sequencing identifies *WDR35* variants involved in Sensenbrenner syndrome. Am. J. Hum. Genet. *87*, 418–423.

12. Tsurusaki, Y., Osaka, H., Hamanoue, H., Shimbo, H., Tsuji, M., Doi, H., Saitsu, H., Matsumoto, N., and Miyake, N. (2011). Rapid detection of a mutation causing X-linked leucoencephalopathy by exome sequencing. J. Med. Genet., in press. Published online March 17, 2011. 10.1136/jmg.2010.083535.

13. Becker, J., Semler, O., Gilissen, C., Li, Y., Bolz, H.J., Giunta, C., Bergmann, C., Rohrbach, M., Koerber, F., Zimmermann, K., et al. (2011). Exome sequencing identifies truncating mutations in human *SERPINF1* in autosomal-recessive osteogenesis imperfecta. Am. J. Hum. Genet. *88*, 362–371.

14. Fukuda, M. (2003). Molecular cloning, expression, and characterization of a novel class of synaptotagmin (Syt XIV) conserved from *Drosophila* to humans. J. Biochem. *133*, 641–649.

15. Adolfsen, B., Saraswati, S., Yoshihara, M., and Littleton, J.T. (2004). Synaptotagmins are trafficked to distinct subcellular domains including the postsynaptic compartment. J. Cell Biol. *166*, 249–260.

16. Quintero-Rivera, F., Chan, A., Donovan, D.J., Gusella, J.F., and Ligon, A.H. (2007). Disruption of a synaptotagmin (*SYT14*) associated with neurodevelopmental abnormalities. Am. J. Med. Genet. A. *143*, 558–563.

17. Michelsen, U., and von Hagen, J. (2009). Isolation of subcellular organelles and structures. Methods Enzymol. *463*, 305–328.

18. Fukuda, M., Kojima, T., and Mikoshiba, K. (1996). Phospholipid composition dependence of Ca^{2+}-dependent phospho-

19. Doi, H., Mitsui, K., Kurosawa, M., Machida, Y., Kuroiwa, Y., and Nukina, N. (2004). Identification of ubiquitin-interacting proteins in purified polyglutamine aggregates. FEBS Lett. *571*, 171–176.

20. Jana, N.R., Tanaka, M., Wang, G., and Nukina, N. (2000). Polyglutamine length-dependent interaction of Hsp40 and Hsp70 family chaperones with truncated N-terminal huntingtin: Their role in suppression of aggregation and cellular toxicity. Hum. Mol. Genet. *9*, 2009–2018.

21. Oyama, F., Miyazaki, H., Sakamoto, N., Becquet, C., Machida, Y., Kaneko, K., Uchikawa, C., Suzuki, T., Kurosawa, M., Ikeda, T., et al. (2006). Sodium channel beta4 subunit: down-regulation and possible involvement in neuritic degeneration in Huntington's disease transgenic mice. J. Neurochem. *98*, 518–529.

22. Gros-Louis, F., Dupré, N., Dion, P., Fox, M.A., Laurent, S., Verreault, S., Sanes, J.R., Bouchard, J.P., and Rouleau, G.A. (2007). Mutations in *SYNE1* lead to a newly discovered form of autosomal recessive cerebellar ataxia. Nat. Genet. *39*, 80–85.

23. McCue, H.V., Haynes, L.P., and Burgoyne, R.D. (2010). The diversity of calcium sensor proteins in the regulation of neuronal function. Cold Spring Harb. Perspect. Biol. *2*, a004085.

24. Bai, J., and Chapman, E.R. (2004). The C2 domains of synaptotagmin—partners in exocytosis. Trends Biochem. Sci. *29*, 143–151.

25. Chapman, E.R. (2008). How does synaptotagmin trigger neurotransmitter release? Annu. Rev. Biochem. *77*, 615–641.

26. Rickman, C., Craxton, M., Osborne, S., and Davletov, B. (2004). Comparative analysis of tandem C2 domains from the mammalian synaptotagmin family. Biochem. J. *378*, 681–686.

27. Bhalla, A., Chicka, M.C., and Chapman, E.R. (2008). Analysis of the synaptotagmin family during reconstituted membrane fusion. Uncovering a class of inhibitory isoforms. J. Biol. Chem. *283*, 21799–21807.

28. Zhang, G., Bai, H., Zhang, H., Dean, C., Wu, Q., Li, J., Guariglia, S., Meng, Q., and Cai, D. (2011). Neuropeptide exocytosis involving synaptotagmin-4 and oxytocin in hypothalamic programming of body weight and energy balance. Neuron *69*, 523–535.

29. Ellgaard, L., and Helenius, A. (2003). Quality control in the endoplasmic reticulum. Nat. Rev. Mol. Cell Biol. *4*, 181–191.

30. Colgan, S.M., Hashimi, A.A., and Austin, R.C. (2011). Endoplasmic reticulum stress and lipid dysregulation. Expert Rev. Mol. Med. *13*, e4.

31. Glavan, G., Schliebs, R., and Zivin, M. (2009). Synaptotagmins in neurodegeneration. Anat. Rec. (Hoboken) *292*, 1849–1862.

32. Schmitz-Hübsch, T., du Montcel, S.T., Baliko, L., Berciano, J., Boesch, S., Depondt, C., Giunti, P., Globas, C., Infante, J., Kang, J.S., et al. (2006). Scale for the assessment and rating of ataxia: Development of a new clinical scale. Neurology *66*, 1717–1720.

ORIGINAL INVESTIGATION

Rapid detection of gene mutations responsible for non-syndromic aortic aneurysm and dissection using two different methods: resequencing microarray technology and next-generation sequencing

Haruya Sakai · Shinichi Suzuki · Takeshi Mizuguchi · Kiyotaka Imoto · Yuki Yamashita ·
Hiroshi Doi · Masakazu Kikuchi · Yoshinori Tsurusaki · Hirotomo Saitsu ·
Noriko Miyake · Munetaka Masuda · Naomichi Matsumoto

Abstract Aortic aneurysm and/or dissection (AAD) is a life-threatening condition, and several syndromes are known to be related to AAD. In this study, two new technologies, resequencing array technology (ResAT) and next-generation sequencing (NGS), were used to analyze eight genes associated with syndromic AAD in 70 patients with non-syndromic AAD. Eighteen sequence variants were detected using both ResAT and NGS. In addition one of these sequence variants was detected by ResAT only and two additional variants by NGS only. Three of the 18 variants are likely to be pathogenic (in 4.3% of AAD patients and in 8.6% of a subset of patients with thoracic AAD), highlighting the importance of genetic analysis in non-syndromic AAD. ResAT and NGS similarly detected most, but not all, of the variants. Resequencing array technology was a rapid and efficient method for detecting most nucleotide substitutions, but was unable to detect short insertions/deletions, and it is impractical to update custom arrays frequently. Next-generation sequencing was able to detect almost all types of mutation, but requires improved informatics methods.

H. Sakai · T. Mizuguchi · Y. Yamashita · H. Doi · M. Kikuchi ·
Y. Tsurusaki · H. Saitsu · N. Miyake · N. Matsumoto (✉)
Department of Human Genetics, Yokohama City University
Graduate School of Medicine, 3-9 Fukuura, Kanazawa-ku,
Yokohama 236-0004, Japan
e-mail: naomat@yokohama-cu.ac.jp

S. Suzuki · M. Masuda
Department of Surgery, Yokohama City University Graduate
School of Medicine, Yokohama, Japan

K. Imoto
Cardiovascular Center, Yokohama City University Medical
Center, 4-57 Urafune, Minami-ku, Yokohama 232-0024, Japan

Introduction

Aortic aneurysm and/or dissection (AAD) is a life-threatening condition. As significant symptoms do not usually appear before the rupture of the AAD, which can be lethal, it is often difficult to prevent death from AAD. Timely cardiovascular surgery may prevent AAD rupture and save the patient's life. Approximately 20% of patients with thoracic aortic disease have a family history of the disease, which is typically inherited in an autosomal dominant manner with decreased penetrance and variable expressivity (Wang et al. 2010). Therefore, if a causative mutation is detected in a patient, it is worth checking for the mutation in their asymptomatic family members to prevent future aortic events by medical and/or surgical intervention. Several genes are known to be associated with syndromes presenting with hereditary AAD and vascular disruption: *FBN1* (Dietz et al. 1991; Lee et al. 1991a), *TGFBR2* (Mizuguchi et al. 2004), *TGFBR1* (Loeys et al. 2005), *MYH11* (Zhu et al. 2006), *ACTA2* (Guo et al. 2007), *COL3A1* (Superti-Furga et al. 1988), *PLOD1* (Hautala et al. 1993), and *SLC2A10* (Coucke et al. 2006) (Table 1). Most AAD patients who have been surgically treated are not affected by these syndromes. However, the contribution of these genes to non-syndromic AAD has not been thoroughly investigated. A comprehensive study of these genes by conventional Sanger sequencing is a huge and expensive undertaking. Even high-resolution melting methods and denaturing high performance liquid chromatography require the amplification of at least 210 exons from these eight genes (Table 1). Therefore, it has been unrealistic for most laboratories to analyze these genes in multiple samples.

Table 1 Overview of genes associated with AAD analyzed in this study

Gene	GenBank accession no.	Disorder	Type	Exon (CDE)	ORF (bp)	Amplicon
FBN1	NM_000138	MFS, SGS, TAAD	AD	66 (65)	8,616	39
TGFBR2	NM_001024847	MFS2, LDS, SGS, TAAD	AD	8 (8)	1,779	8
TGFBR1	NM_004612	MFS2, LDS, SGS, TAAD	AD	9 (9)	1,512	7
COL3A1	NM_000090	EDS type IV	AD	51 (51)	4,401	16
PLOD1	NM_000302	EDS type VI	AR	19 (19)	2,184	13
MYH11	NM_001040113	TAAD	AD	43 (41)	5,838	30
SLC2A10	NM_030777	ATS	AR	5 (5)	1,626	5
ACTA2	NM_001613	TAAD	AD	9 (8)	1,134	6

CDE coding exon, ORF open reading frame, MFS Marfan syndrome, MFS2 Marfan syndrome type II, LDS Loeys–Dietz syndrome, SGS Shprint-zen–Goldberg syndrome, TAAD thoracic aortic aneurysm and dissection, EDS Ehlers–Danlos syndrome, ATS arterial tortuosity syndrome, AD autosomal dominant, AR autosomal recessive

Resequencing array technology (ResAT) enables the investigation of multiple genes on one chip. This technology has been used for multiple-gene analysis in childhood hearing loss (Kothiyal et al. 2010), breast-ovarian cancer syndrome (Schroeder et al. 2010), dilated cardiomyopathy (Zimmerman et al. 2010), X-linked intellectual disability (Jensen et al. 2011), familial hypercholesterolemia (Chiou et al. 2011), and hypertrophic cardiomyopathy (Fokstuen et al. 2011). Different research groups have shown ResAT to be a highly efficient, relatively accurate, cost-effective, and rapid method. However, several drawbacks have been pointed out, including its insensitivity in detecting nucleotide insertions/deletions (indels) and nucleotide changes in GC-rich regions and repeat sequences.

Next-generation sequencing (NGS) is now regarded as the most powerful technology for detecting mutations (Ng et al. 2010; Tsurusaki et al. 2011). This platform is advantageous in finding almost all types of mutations including small indel mutations. The high throughput and multiplexing of NGS allows multiple genes to be sequenced in many samples in a single run (Farias-Hesson et al. 2010; Gabriel et al. 2009).

In this study, we analyzed the eight AAD-associated genes (*FBN1*, *TGFBR2*, *TGFBR1*, *COL3A1*, *PLOD1*, *MYH11*, *SLC2A10*, and *ACTA2*) in 70 patients with non-syndromic AAD by two methods: ResAT (all eight genes on one chip) and multiplex NGS. We describe here a comparison of the results.

Materials and methods

Patients

Seventy Japanese patients, who had surgery for AAD, were recruited from Yokohama City University Hospital and

Table 2 Clinical information of AAD patients

Clinical data	Number of patients (%)
Thoracic AAD[a]	35 (50.0)
Abdominal AAD[a]	30 (42.9)
Thoracic and abdominal AAD[a]	5 (7.1)
Age (years) (mean ± SD)	67.3 ± 10.2 (range 39–83)
Age (years) (median)	68.5
<50 years old	4 (5.7)
50–54 years old	5 (7.1)
55–59 years old	8 (11.4)
≥60 years old	53 (75.7)
Male	53 (75.7)
Female	17 (24.3)
Diabetes	9 (12.9)
Hyperlipidemia	32 (45.7)
Hypertension	54 (77.1)
Current smoker	15 (21.4)
Past smoker	30 (42.9)
Never smoked	23 (32.9)

[a] Including current and past operations

Yokohama City University Medical Center. The patients' clinical information is summarized in Table 2. Thoracic AAD involves the aorta above the diaphragm and abdominal AAD is located along the portion of the aorta passing through the abdomen. None of the patients in this study had any clinical test results supporting a diagnosis of syndromic AAD. Experimental protocols were approved by the Institutional Review Board of Yokohama City University School of Medicine. Informed consent for genetic analysis was obtained from the patients. DNA was extracted from peripheral blood leukocytes using a QuickGene-610L kit (Fujifilm, Tokyo, Japan).

Array design

Eight genes (*FBN1, TGFBR2, TGFBR1, COL3A1, PLOD1, MYH11, SLC2A10* and *ACTA2*) (Table 1) associated with AAD were selected for one custom chip (Affymetrix, Santa Clara, CA). All coding exons as well as 29 bp of sequence from each intron (21 bp on the 5'-side and 8 bp on the 3'-side of each exon) were analyzed. Repetitive sequences and intragenic low complexity regions larger than 25 bp were excluded from the chip. A total of 33,116 bp from the eight genes could be sequenced using this chip.

PCR amplification, purification, hybridization, scanning, and data analysis

The targeted regions were amplified as 124 fragments by PCR (ranging from 965 to 2,999 bp) using Blend Taq Plus (TOYOBO, Osaka, Japan) or KOD FX (TOYOBO) and genomic DNA as a template in a 20 μL volume. The PCR conditions were: denaturing at 94°C, 35 cycles of 94°C for 30 s, 62°C for 30 s, and 72°C for 3 min, and a final extension at 72°C for 7 min. The DNA concentration of the amplicons was determined using a Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen, Carlsbad, CA, USA) with a Spectra Fluor F129003 (Tecan, Männedorf, Switzerland). The PCR amplicons were pooled in equimolar quantities (110 fmol). The mixed samples were purified and the volume was reduced using a Microcon YM-100 filter (Millipore, Brussels, Belgium). Fragmentation of the products, labeling with biotin, hybridization, washing, and scanning procedures were carried out based on the CustomSeq resequencing array protocol version 2.1 (Affymetrix). An FS450 fluidics station (Affymetrix) was used for washing and staining and a GCS3000 7G scanner (Affymetrix) was used for scanning. To test the efficiency of mutation detection, PCR products containing 20 known heterozygous mutations (Table 3) from three genes (*FBN1, TGFBR2*, and *TGFBR1*), as well as another 104 PCR products amplified from normal control DNA, covering all the other exons, were analyzed using the chip. Affymetrix GCOS and GSEQ software were used to process the raw data and analyze the nucleotide sequences, respectively. The default settings of GSEQ were adopted.

Multiplex next-generation sequencing

The PCR amplicons from one patient were mixed and processed using a multiplexing sequencing primers and PhiX control kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions but with minor changes. In brief, amplicons were fragmented with Covaris S1 (Covaris, Woburn, MA, USA), and purified using Agencourt AMPure (Beckman Coulter, Brea, CA, USA) instead of gel extraction. DNA quality was checked with an Agilent 2100

Table 3 Known mutations used as positive controls for testing ResAT

Nucleotide substitution		Small deletion or insertion	
Gene	Mutation	Gene	Mutation
FBN1	c.400T > G	*FBN1*	c.937delT
	c.772C > T		c.1876delG
	c.1011C > A		c.4283–4284insG
	c.1285C > T		c.7039–7040delAT
	c.2413T > C		
	c.2942G > C		
	c.4099T > C		
	c.4495A > T		
	c.5539T > C		
	c.5788G + 5G > A		
	c.6236C > G		
	c.6773G > A		
TGFBR2	c.1142G > C		
	c.1411G > A		
	c.1624C > T		
TGFBR1	c.1135A > G		

All mutations are previously reported (Sakai et al. 2006; Togashi et al. 2007)

bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) and a bar code DNA tag (Illumina) was ligated on. The bar code DNA tags contain unique 6 bp sequences and allow the processing of up to 96 DNA fragments in a single run using an Illumina GAIIx (Illumina). Twelve processed DNA fragments, each with a different tag, were mixed and analyzed with single 76 bp reads in one lane of the flow cell. Six lanes were necessary for the analysis of 70 samples. Image analysis and base calling were performed by sequence control software real-time analysis (Illumina) and offline Basecaller software v1.8.0 (Illumina). The reads were aligned to the human reference genome sequence (UCSC hg19, GRCh37) using the ELAND v2 algorithm in CASAVA software v1.7.0 (Illumina).

Mapping strategy and variant annotation

An average of 2.4 million reads (ranging from 1.7 to 4.0 million reads) for each sample passed quality control (Path Filter) and were mapped to the human reference genome using mapping and assembly with qualities (MAQ) (Li et al. 2008), NextGENe software v2.00 (SoftGenetics, State College, PA, USA), and Burrows-Wheeler Aligner (BWA)/ sequence alignment/map tools (SAMtools) (Li and Durbin 2010; Li et al. 2009). Single nucleotide polymorphisms (SNPs) and indels were extracted from the alignment data using an original script created by BITS, Tokyo, Japan along with information on the registered SNPs (dbSNP131). A consensus quality score of 40 or more was used for the

SNP analysis in MAQ. SNPs in MAQ-passed reads were annotated using the SeattleSeq website (http://gvs.gs. washington.edu/SeattleSeqAnnotation/). A minimum base quality of 13, a minimum root mean square mapping quality for SNPs of 10, and a minimum read depth of 2 were used in BWA/SAMtools (Li and Durbin 2010; Li et al. 2009). NextGENe (SoftGenetics) was also used to analyze the reads, employing default settings apart from using the no-condensation mode. For base substitutions, we focused on variants detected in common by both MAQ and Next-GENe. Small indel variants were classified as positive if found by both BWA and NextGENe.

Validation of novel variants

Novel variants (not in dbSNP131, the 1,000 genomes dataset or our in-house database) identified by ResAT and NGS were validated by Sanger sequencing. Surplus PCR products were treated with ExoSAP IT (GE Healthcare, Piscataway, NJ) and sequenced using a standard protocol using BigDye terminators (Applied Biosystems, Foster City, CA, USA) on an ABI PRISM 3100 genetic analyzer (Applied Biosystems). Furthermore, novel variants were screened in 94 Japanese controls by high-resolution melt curve analysis (LightCycler 480; Roche Diagnostics, Basel, Switzerland) and subsequent Sanger sequencing. Novel variants were evaluated using web-based programs including PolyPhen (http://genetics. bwh.harvard.edu/pph/), PolyPhen2 (http://genetics.bwh. harvard.edu/pph2/), Mutation Taster (http://www.mutationt aster.org/), and ESEfinder (http://rulai.cshl.edu/cgi-bin/tools/ ESE3/esefinder.cgi?process=home).

Results

Array performance

Across all 70 samples, the mean nucleotide call rate was 95.7% (range 87.3–97.6%) using the default settings of GSEQ. We observed an improvement of the call rate as the number of samples increased. For example, the call rate by GCOS for the first two samples was 90.1 and 90.6% and was 93.3 and 93.9% when 10 samples were analyzed, and was 94.9 and 95.5% when 33 samples were analyzed. However, between 34 and 70 samples, the call rate did not greatly improve (only by 1%). We had constant difficulty in reading approximately 4% of the sequences per array (i.e., no sequence called), mostly in regions of high GC and CC content.

Detection of known mutations by ResAT

To validate the quality of mutation detection in our resequencing array, we analyzed amplicons containing 16 known nucleotide substitutions, three small deletions (1–2 bp), and one 1 bp insertion, plus all the other normal exons (Sakai et al. 2006; Togashi et al. 2007) (Table 3). Fourteen out of 16 nucleotide substitutions were detected (87.5%) by GSEQ in the automated mode. Two mutations (c.772C > T in *FBN1* and c.1142G > C in *TGFBR2*) were not detected. The former was insensitive, and the latter was indicated as a no-call. Visual inspection in the manual mode enabled easy detection of the *TGFBR2* mutation. The mutation detection rate was 93.8% (15/16) using both the automated and manual modes. None of the small indels were detected by our array in either the automated or manual modes.

Variant detection by ResAT

We detected 70 nucleotide substitutions in the automated mode in the 70 patients analyzed (0–3 variants per sample). Fifty-one variants were already registered in dbSNP131 and/or in our in-house database (Supplementary table). The remaining 19 novel variants were validated by Sanger sequencing (Table 4). One variant (c.976–16C > T in *PLOD1*) was homozygous and the others were heterozygous. No indel mutations were detected.

Variant detection by NGS

The target regions were completely covered by NGS reads (100%). The average read depth (coverage of sequence reads) was approximately 600 for each gene (Table 5). The NextGENe software detected a mean of 876 variants in the 70 patients with mutation scores of 10 or more (ranging from 581 to 1209 with SD = 131). MAQ and SeattelSeq detected a mean of 271 variants (ranging from 111 to 384 with SD = 52). Semi-automatic exclusion of variants that were out of the target regions (22 bp or more away from the 5′-end of exons and 9 bp or more away from the 3′-end of exons) or were known variants in dbSNP131 was performed using Excel 2008 for Mac (Microsoft, Redmond, WA, USA), narrowing the data down to 0–6 variants per sample. Twenty novel variants were detected by both MAQ and NextGENe, which were further validated by Sanger sequencing. No indel mutations were detected by MAQ, NextGENe, or BWA/SAMtools.

Comparison of ResAT and NGS variants

Eighteen novel variants were detected by ResAT and NGS. One was detected by ResAT only and two by NGS only. The two variants undetected by ResAT were c.1388G > A (p.Arg463Gln) in *PLOD1* and c.136A > C (p.Ser46Arg) in *TGFBR2*. The former was indicated as a no-call, but was detected later in the manual mode. The latter was within a

173

Table 4 Novel variants detected by ResAT and/or NGS

Mutation		Amino acid change	Methods of detection	Read depth in NGS	PolyPhen	PolyPhen2	Mutation taster	Patients	Controls (total number)
Gene	Mutation								
TGFBR2	c.136A > C[c]	p.Ser46Arg	NGS	472	Benign	Benign (0.099)	Polymorphism	1	0 (94)
	c.403G > T	p.Asp135Tyr	ResAT and NGS	1,257	Possibly damaging	Possibly damaging (0.682)	Polymorphism	1	0 (94)
	c.692C > T[d]	p.Thr231Met	ResAT and NGS	989	Benign	Possibly damaging (0.670)	Polymorphism	1	0 (93)
TGFBR1	c.1032T > C	p.Asn344Asn	ResAT and NGS	939	Unknown	–	Polymorphism	1	3 (94)
COL3A1	c.1815 + 5G > A		ResAT	255	Unknown	–	Disease-causing	1	0 (94)
	c.84T > C[d]	p.Val28Val	ResAT and NGS	644	Unknown	–	Polymorphism	1	0 (94)
	c.119C > T[e]	p.Ala40Val	ResAT and NGS	1,402	Unknown	Unknown	Polymorphism	1	0 (94)
	c.3133G > A	p.Ala1045Thr	ResAT and NGS	630	Benign	Probably damaging (0.979)	Polymorphism	1	0 (94)
	c.3776C > T	p.Ala1259Val	ResAT and NGS	872	Unknown	Unknown	Disease-causing	1	0 (94)
PLOD1	c.976–16C > T[a]		ResAT and NGS	631	unknown	–	polymorphism	1	2[a] (94)
	c.1098–8C > G		ResAT and NGS	633	Unknown	–	Disease-causing	1	0 (94)
	c.1388G > A	p.Arg463Gln	NGS	624, 768	Unknown	Probably damaging (0.961)	Disease-causing	2	4 (94)
	c.1495C > T	p.Arg499Trp	ResAT and NGS	509, 532, 568, 679	Probably damaging	Probably damaging (0.992)	Disease-causing	4	2 (94)
MYH11	c.4600–13G > A		ResAT and NGS	1,336	unknown	–	Polymorphism	1	2 (94)
	c.4625G > A[b]	p.Arg1542Gln	ResAT and NGS	1,254	Possibly damaging	Probably damaging (0.994)	Disease-causing	1	0 (94)
	c.4963C > T[b]	p.Arg1655Cys	ResAT and NGS	2,711	Probably damaging	Probably damaging (1.000)	Disease-causing	1	0 (94)
SLC2A10	c.315C > T	p.Arg105Arg	ResAT and NGS	543	Unknown	–	Polymorphism	1	0 (94)
	c.330C > T[e]	p.Phe110Phe	ResAT and NGS	500	Unknown	–	Polymorphism	1	0 (94)
	c.1220T > G[b]	p.Leu407Arg	ResAT and NGS	382	Benign	possibly damaging (0.925)	Disease-causing	1	0 (94)
ACTA2	c.130–18T > C[c]		ResAT and NGS	607, 647	Unknown	–	Polymorphism	2	2 (94)
	c.482T > C	p.Val161Ala	ResAT and NGS	752	Probably damaging	Beningn (0.013)	Disease-causing	1	0 (94)

The underlined mutation is highly likely to be pathogenic

[a] Homozygous substitution

[b] Mutations detected in patient 16 patient

[c] Mutations detected in patient 24

[d] Mutations detected in patient 28

[e] Mutations detected in patient 89

Table 5 Gene-based read depth in NGS

Gene	Mean depth[a]
FBN1	655
TGFBR2	613
TGFBR1	568
COL3A1	596
PLOD1	607
MYH11	643
SLC2A10	571
ACTA2	543

[a] Based on NextGENe calculation

repetitive sequence. One variant (c.1815 + 5G > A in *COL3A1*) was undetected by NGS due to our set criteria (the variant was detected by MAQ, but not by NextGENe or BWA/SAMtools).

Pathological significance of the variants

We realized that none of the known pathogenic mutations were identified. The pathological impact of the variants was considered if none of the healthy controls showed the same change, if the variants altered evolutionarily conserved amino acids in functional repeats/domains, or if they were predicted to cause abnormal splicing resulting in protein truncation or degradation. Moreover, homozygous and compound heterozygous changes that were found in *PLOD1* and *SLC2A10* may confer autosomal recessive effects. At least three heterozygous variants were considered as putative pathogenic gene alterations (Table 6):

1. c.1815 + 5G > A in *COL3A1* (patient 29). A similar mutation, c.1815 + 5G > T, associated with the skipping of exon 25, was reported in a patient with Ehlers–Danlos syndrome type IV (EDS IV) (Lee et al. 1991b). ESEfinder suggested that the binding position of the splice donor matrix was changed similarly by c.1815 + 5G > A and c.1815 + 5G > T. Thus, C.1815 + 5G > A is highly likely to be pathogenic.
2. c.4963C > T (p.Arg1655Cys) in *MYH11* (patient 16). In addition to this mutation, the patient had two novel

heterozygous variants: c.4625G > A (p.Arg1542Gln) in *MYH11* and c.1220T > G (p.Leu407Arg) in *SLC2A10*. Mutations in *SLC2A10* cause autosomal recessive arterial tortuosity syndrome (MIM #208050) (Coucke et al. 2006), although it is unknown whether the heterozygous variant we identified would be related to this, assuming a second-hit model of recessive disease. Both p.Arg 1542Gln and p.Arg1655Cys in *MYH11* were similarly predicted to be pathogenic by three programs (PolyPhen, PolyPhen2, and Mutation Taster). These residues are located in the coiled-coil region, and both are evolutionarily conserved amino acids (Fig. 1). Paircoil2 (http://groups.csail.mit.edu/cb/paircoil2/) was used to predict the effect of variants on the parallel coiled coil fold using pairwise residue probabilities (McDonnell et al. 2006). Paircoil2 indicated that p.Arg1655Cys altered the p score from 0.00096 (wild type) to 0.00579 (mutation), while p.Arg1542Gln did not alter the p score, 0.00016 (mutation) and 0.00018 (wild type) (Fig. 1). Thus, p.Arg1655Cys was more likely than p.Arg1542Gln to be pathogenic.
3. c.482T > C (p.Val161Ala) in *ACTA2* (patient 27). The patient was found retrospectively to suffer from familial thoracic AAD. The patient has an affected brother, but his DNA was unavailable. Valine at amino acid 161 is evolutionarily conserved and located within the actin domain. However, as we could not analyze the DNA of the affected brother, it may be more appropriate to call this variant 'of unknown significance'.

Discussion

Exon-by-exon Sanger sequencing is the gold standard for genetic analysis, but multiple-gene analysis in many patients is a huge task in terms of time and cost. In this study, we applied two emerging technologies providing rapid and efficient analysis of eight genes in 70 AAD patients. We also compared the results of the two technologies.

The overall mean call rate of our custom array by GSEQ software was 95.7%, which is comparable with previous

Table 6 Pathogenic variants found in the patients

Patient ID	Sex	Mutation	Clinical diagnosis	Age[a]	Age[b]	Family history
Patient 16	M	*MYH11* c.4963C > T p.Arp1655Cys	Thoracic and abdominal AAD	80	80	None
Patient 27	F	*ACTA2* c.482T > C p.Val161Ala	Thoracic AAD	57	46	Affected brother
Patient 29	F	*COL3A1* c.1815 + 5G > A	Thoracic AAD	80	67	None

M male, *F* female

[a] At blood collection

[b] At the first surgery

a

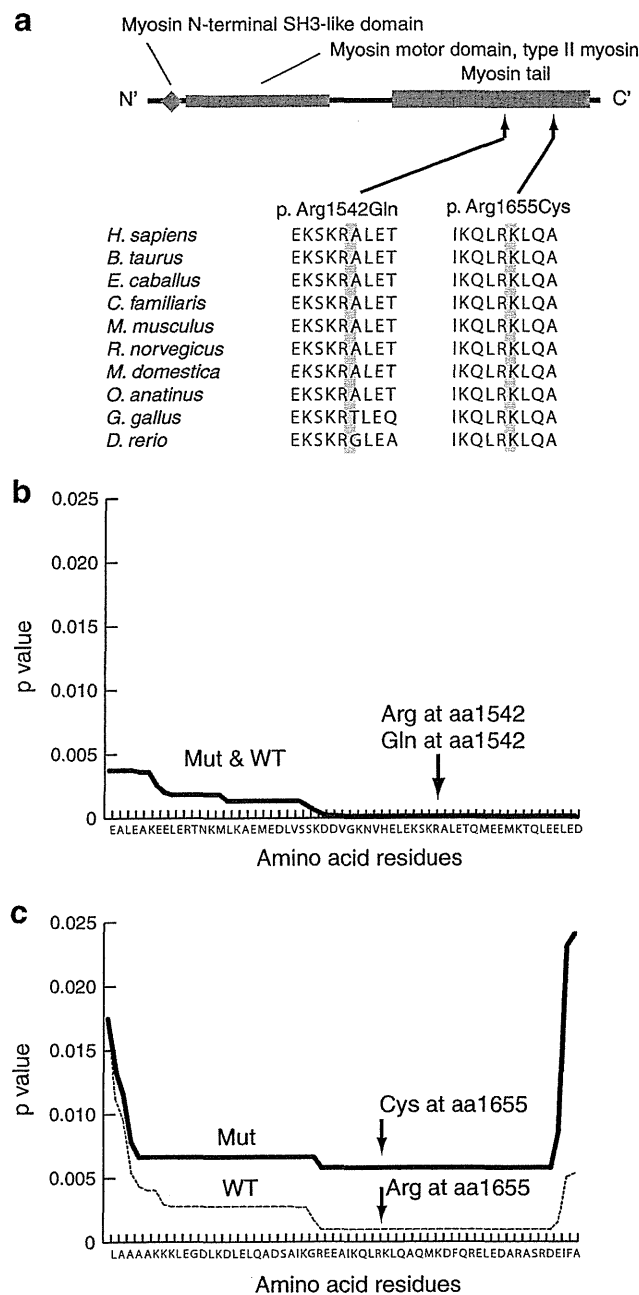|	p. Arg1542Gln	p. Arg1655Cys
H. sapiens	EKSKRALET	IKQLRKLQA
B. taurus	EKSKRALET	IKQLRKLQA
E. caballus	EKSKRALET	IKQLRKLQA
C. familiaris	EKSKRALET	IKQLRKLQA
M. musculus	EKSKRALET	IKQLRKLQA
R. norvegicus	EKSKRALET	IKQLRKLQA
M. domestica	EKSKRALET	IKQLRKLQA
O. anatinus	EKSKRALET	IKQLRKLQA
G. gallus	EKSKRTLEQ	IKQLRKLQA
D. rerio	EKSKRGLEA	IKQLRKLQA

b

c

Fig. 1 Double mutations in *MYH11*. **a** Schematic representation of the MYH11 protein. Three functional domains are indicated: the myosin N-terminal SH3-like domain, the myosin motor domain for type II myosin, and the myosin tail. Both the mutations are located in the myosin tail. **b, c** Paircoil2 analysis showing a significantly decreased probability of coiled-coil formation for p.Arg1655Cys relative to the wild-type sequence, but no change for p.Arg1542Gln

studies (Bruce et al. 2010; Chiou et al. 2011; Jensen et al. 2011; Schroeder et al. 2010). The call rates became higher as the number of patients increased. Approximately 33 samples were necessary to attain the maximum call-rate in GSEQ. A similar observation was described previously (Fokstuen et al. 2011). No-call regions are one of the

problems of ResAT. Other groups have previously suggested that most of the no-call regions are GC- and CC-rich (Bruce et al. 2010; Chiou et al. 2011; Fokstuen et al. 2011). In our custom array, approximately 4% of the target sequences were difficult to obtain (no-calls) in most of the samples.

The mean detection rate of known variants using our custom array and GSEQ with the default settings (automated analysis) was 87.5%. This rate increased to 93.8% after manual inspection. For our ResAT data, the detection rate of nucleotide substitutions in the automated mode was higher, and that in the manual mode was slightly lower, compared with detection rates in previous studies (82.1 vs. 81%, respectively, in automated mode, and 97.4 vs. 100%, respectively, in manual mode) (Bruce et al. 2010; Chiou et al. 2011). Our ResAT analysis was unable to detect any small indel mutations; this is similar to other studies (Hartmann et al. 2009; Kothiyal et al. 2010). In the human gene mutation database (HGMD; http://www.hgmd.cf.ac.uk/ac/index. php), insertions/deletions account for a substantial proportion of the total registered mutations in our genes of interest: *FBN1* 23.6%, *TGFBR2* 6.4%, *TGFBR1* 10%, *COL3A1* 12.8%, *PLOD1* 46.2%, *MYH11* 20%, *SLC2A10* 21.1%, and *ACTA* 20%. Thus, the incapability of ResAT to detect indel mutations is one of its most significant drawbacks.

Our NGS analysis missed one of 21 variants (c.1815 + 5G > A in *COL3A1*). Our protocol focused on variants identified by two different informatics methods, to increase the true-positive rate. For example, MAQ (single-end reads) can detect nucleotide substitutions well, but is not good at detecting small indels (Li et al. 2008). BWA is more sensitive at detecting small indels because it can align gapped sequence (Krawitz et al. 2010). NextGENe is based on the Burrows-Wheeler transform algorithm, which is good at detecting small indels. NGS needs more efficient informatics methods to extract all the nucleotide changes correctly with lower error rates.

In this study, concomitant variants in two genes were detected in four patients (Table 4): c.4625G > A and c.4963C > T in *MYH11*, and c.1220T > G in *SLC2A10* (patient 16); c.136A > C in *TGFBR2* and c.130–18T > C in *ACTA2* (patient 24); c.84T > C in *COL3A1* and c.692C > T in *TGFBR2* (patient 28); c.119C > T in *COL3A1* and c.330C > T in *SLC2A10* (patient 89) (Table 4). It may be quite difficult to detect variants in two or more genes by conventional methods. ResAT and NGS permitted us to find multiple variants in multiple genes easily and rapidly. Double or triple mutations in unusual clinical cases will also be found using such technologies.

Three different putative pathological mutations in a heterozygous state in three of 70 patients were found in this study (4.3%). Interestingly, all the three patients suffered from thoracic AAD. Considering only those patients with

thoracic AAD ($n = 35$), the rate increased to 8.6%. Thus, non-syndromic AAD (especially thoracic AAD) can be explained to some extent by aberrations of genes related to Mendelian disorders, although our sample size was small. Interestingly, among these three patients, only patient 29 showed hyperlipidemia and the other two (patients 16 and 27) did not, which supports the genetic origin of thoracic AAD.

In this study, we compared ResAT and NGS. Considering the drawbacks of ResAT, including its inability to detect small indels and its no-call regions, we believe that NGS is the better technology for comprehensive analysis of multiple genes, especially with improved informatics methods, as it can detect all types of mutations with no bias. Another advantage of NGS is its flexibility. Resequencing array technology requires a custom-made sequencing array. It is not easy or practical to update arrays frequently. However, NGS is currently quite expensive for most laboratories. Next-generation sequencing combined with the pooled genomic DNA method with indexing may improve its cost-effectiveness (Calvo et al. 2010; Druley et al. 2009).

In conclusion, we found that 4.3% of non-syndromic AAD patients (8.5% of thoracic AAD patients) have abnormalities in genes that cause Mendelian disorders. ResAT and NGS enabled multiple genes to be analyzed efficiently. In addition to the 70 AAD patients, a patient with familial Marfan syndrome and a patient with Loeys–Dietz syndrome were initially included before their diagnosis was known. We detected c.6793T > G (p.Cys2265Gly) in *FBN1* in the Marfan syndrome patient [by ResAT (NGS was not done)] and c.797A > G (p.Asp266Gly) in *TGFBR1* in the Loeys–Dietz patient (by ResAT and NGS). We excluded these two patients from this study because they are syndromic AAD patients, but the efficient detection of their mutations highlights the validity of our approach. Finally, high throughput technologies have the potential to routinely identify novel variants of known or unknown significance in clinical settings. Therefore, more sophisticated methods to evaluate gene variants as well as databases containing normal (rare) variants are needed.

References

Bruce CK, Smith M, Rahman F, Liu ZF, McMullan DJ, Ball S, Hartley J, Kroos MA, Heptinstall L, Reuser AJ, Rolfs A, Hendriksz C,

Kelly DA, Barrett TG, MacDonald F, Maher ER, Gissen P (2010) Design and validation of a metabolic disorder resequencing microarray (BRUM1). Hum Mutat 31:858–865

Calvo SE, Tucker EJ, Compton AG, Kirby DM, Crawford G, Burtt NP, Rivas M, Guiducci C, Bruno DL, Goldberger OA, Redman MC, Wiltshire E, Wilson CJ, Altshuler D, Gabriel SB, Daly MJ, Thorburn DR, Mootha VK (2010) High-throughput, pooled sequencing identifies mutations in NUBPL and FOXRED1 in human complex I deficiency. Nat Genet 42:851–858

Chiou KR, Charng MJ, Chang HM (2011) Array-based resequencing for mutations causing familial hypercholesterolemia. Atherosclerosis 216(2):383–389

Coucke PJ, Willaert A, Wessels MW, Callewaert B, Zoppi N, De Backer J, Fox JE, Mancini GM, Kambouris M, Gardella R, Facchetti F, Willems PJ, Forsyth R, Dietz HC, Barlati S, Colombi M, Loeys B, De Paepe A (2006) Mutations in the facilitative glucose transporter GLUT10 alter angiogenesis and cause arterial tortuosity syndrome. Nat Genet 38:452–457

Dietz HC, Cutting GR, Pyeritz RE, Maslen CL, Sakai LY, Corson GM, Puffenberger EG, Hamosh A, Nanthakumar EJ, Curristin SM et al (1991) Marfan syndrome caused by a recurrent de novo missense mutation in the fibrillin gene. Nature 352:337–339

Druley TE, Vallania FL, Wegner DJ, Varley KE, Knowles OL, Bonds JA, Robison SW, Doniger SW, Hamvas A, Cole FS, Fay JC, Mitra RD (2009) Quantification of rare allelic variants from pooled genomic DNA. Nat Methods 6:263–265

Farias-Hesson E, Erikson J, Atkins A, Shen P, Davis RW, Scharfe C, Pourmand N (2010) Semi-automated library preparation for high-throughput DNA sequencing platforms. J Biomed Biotechnol 2010:617469

Fokstuen S, Munoz A, Melacini P, Iliceto S, Perrot A, Ozcelik C, Jeanrenaud X, Rieubland C, Farr M, Faber L, Sigwart U, Mach F, Lerch R, Antonarakis SE, Blouin JL (2011) Rapid detection of genetic variants in hypertrophic cardiomyopathy by custom DNA resequencing array in clinical practice. J Med Genet 48(8):572–576

Gabriel C, Danzer M, Hackl C, Kopal G, Hufnagl P, Hofer K, Polin H, Stabentheiner S, Proll J (2009) Rapid high-throughput human leukocyte antigen typing by massively parallel pyrosequencing for high-resolution allele identification. Hum Immunol 70:960–964

Guo DC, Pannu H, Tran-Fadulu V, Papke CL, Yu RK, Avidan N, Bourgeois S, Estrera AL, Safi HJ, Sparks E, Amor D, Ades L, McConnell V, Willoughby CE, Abuelo D, Willing M, Lewis RA, Kim DH, Scherer S, Tung PP, Ahn C, Buja LM, Raman CS, Shete SS, Milewicz DM (2007) Mutations in smooth muscle alpha-actin (ACTA2) lead to thoracic aortic aneurysms and dissections. Nat Genet 39:1488–1493

Hartmann A, Thieme M, Nanduri LK, Stempfl T, Moehle C, Kivisild T, Oefner PJ (2009) Validation of microarray-based resequencing of 93 worldwide mitochondrial genomes. Hum Mutat 30:115–122

Hautala T, Heikkinen J, Kivirikko KI, Myllyla R (1993) A large duplication in the gene for lysyl hydroxylase accounts for the type VI variant of Ehlers–Danlos syndrome in two siblings. Genomics 15:399–404

Jensen LR, Chen W, Moser B, Lipkowitz B, Schroeder C, Musante L, Tzschach A, Kalscheuer VM, Meloni I, Raynaud M, van Esch H, Chelly J, de Brouwer AP, Hackett A, van der Haar S, Henn W, Gecz J, Riess O, Bonin M, Reinhardt R, Ropers HH, Kuss AW (2011) Hybridisation-based resequencing of 17 X-linked intellectual disability genes in 135 patients reveals novel mutations in ATRX, SLC6A8 and PQBP1. Eur J Hum Genet 19(6):717–720

Kothiyal P, Cox S, Ebert J, Husami A, Kenna MA, Greinwald JH, Aronow BJ, Rehm HL (2010) High-throughput detection of mutations responsible for childhood hearing loss using resequencing microarrays. BMC Biotechnol 10:10

Krawitz P, Rodelsperger C, Jager M, Jostins L, Bauer S, Robinson PN (2010) Microindel detection in short-read sequence data. Bioinformatics 26:722–729

Lee B, Godfrey M, Vitale E, Hori H, Mattei MG, Sarfarazi M, Tsipouras P, Ramirez F, Hollister DW (1991a) Linkage of Marfan syndrome and a phenotypically related disorder to two different fibrillin genes. Nature 352:330–334

Lee B, Vitale E, Superti-Furga A, Steinmann B, Ramirez F (1991b) G to T transversion at position +5 of a splice donor site causes skipping of the preceding exon in the type III procollagen transcripts of a patient with Ehlers-Danlos syndrome type IV. J Biol Chem 266:5256–5259

Li H, Durbin R (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26:589–595

Li H, Ruan J, Durbin R (2008) Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res 18:1851–1858

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079

Loeys BL, Chen J, Neptune ER, Judge DP, Podowski M, Holm T, Meyers J, Leitch CC, Katsanis N, Sharifi N, Xu FL, Myers LA, Spevak PJ, Cameron DE, De Backer J, Hellemans J, Chen Y, Davis EC, Webb CL, Kress W, Coucke P, Rifkin DB, De Paepe AM, Dietz HC (2005) A syndrome of altered cardiovascular, craniofacial, neurocognitive and skeletal development caused by mutations in TGFBR1 or TGFBR2. Nat Genet 37:275–281

McDonnell AV, Jiang T, Keating AE, Berger B (2006) Paircoil2: improved prediction of coiled coils from sequence. Bioinformatics 22:356–358

Mizuguchi T, Collod-Beroud G, Akiyama T, Abifadel M, Harada N, Morisaki T, Allard D, Varret M, Claustres M, Morisaki H, Ihara M, Kinoshita A, Yoshiura K, Junien C, Kajii T, Jondeau G, Ohta T, Kishino T, Furukawa Y, Nakamura Y, Niikawa N, Boileau C, Matsumoto N (2004) Heterozygous TGFBR2 mutations in Marfan syndrome. Nat Genet 36:855–860

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 (2010) Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. Nat Genet 42:790–793

Sakai H, Visser R, Ikegawa S, Ito E, Numabe H, Watanabe Y, Mikami H, Kondoh T, Kitoh H, Sugiyama R, Okamoto N, Ogata T, Fodde R, Mizuno S, Takamura K, Egashira M, Sasaki N, Watanabe S, Nishimaki S, Takada F, Nagai T, Okada Y, Aoka Y, Yasuda K, Iwasa M, Kogaki S, Harada N, Mizuguchi T, Matsumoto N (2006) Comprehensive genetic analysis of relevant four genes in 49 patients with Marfan syndrome or Marfan-related phenotypes. Am J Med Genet A 140:1719–1725

Schroeder C, Stutzmann F, Weber BH, Riess O, Bonin M (2010) High-throughput resequencing in the diagnosis of BRCA1/2 mutations using oligonucleotide resequencing microarrays. Breast Cancer Res Treat 122:287–297

Superti-Furga A, Gugler E, Gitzelmann R, Steinmann B (1988) Ehlers-Danlos syndrome type IV: a multi-exon deletion in one of the two COL3A1 alleles affecting structure, stability, and processing of type III procollagen. J Biol Chem 263:6226–6232

Togashi Y, Sakoda H, Nishimura A, Matsumoto N, Hiraoka H, Matsuzawa Y (2007) A Japanese family of typical Loeys-Dietz syndrome with a TGFBR2 mutation. Intern Med 46:1995–2000

Tsurusaki Y, Osaka H, Hamanoue H, Shimbo H, Tsuji M, Doi H, Saitsu H, Matsumoto N, Miyake N (2011) Rapid detection of a mutation causing X-linked leucoencephalopathy by exome sequencing. J Med Genet 48(9):606–609

Wang L, Guo DC, Cao J, Gong L, Kamm KE, Regalado E, Li L, Shete S, He WQ, Zhu MS, Offermanns S, Gilchrist D, Elefteriades J, Stull JT, Milewicz DM (2010) Mutations in myosin light chain kinase cause familial aortic dissections. Am J Hum Genet 87:701–707

Zhu L, Vranckx R, Khau Van Kien P, Lalande A, Boisset N, Mathieu F, Wegman M, Glancy L, Gasc JM, Brunotte F, Bruneval P, Wolf JE, Michel JB, Jeunemaitre X (2006) Mutations in myosin heavy chain 11 cause a syndrome associating thoracic aortic aneurysm/aortic dissection and patent ductus arteriosus. Nat Genet 38:343–349

Zimmerman RS, Cox S, Lakdawala NK, Cirino A, Mancini-DiNardo D, Clark E, Leon A, Duffy E, White E, Baxter S, Alaamery M, Farwell L, Weiss S, Seidman CE, Seidman JG, Ho CY, Rehm HL, Funke BH (2010) A novel custom resequencing array for dilated cardiomyopathy. Genet Med 12:268–278