Pfizer, Eli Lilly, and JCR Pharmaceuticals and has received speaker's fees from Chugai Pharmaceutical. All other authors state that they have no conflicts of interest.

Acknowledgments

We thank the patients and their family members who participated in this study. We also thank Minako Takaki and Reiko Onai for technical support, and Mitsuko Itoh for kind assistance in English usage. This study was supported by a Grant-in-Aid from the Ministry of Education, Science, Sports, and Culture of Japan.

Authors' roles: Study design: TI and SK. Patients' sample and clinical data collection: YO, ET, AY, TY, and RH. Study conduct and data collection: TI, JM, and HI. Exome sequencing data analysis: JM, KD, JY, HI, SM, and ST. Data interpretation: TI, JM, KD, SM, ST, and SK. Study support and intellectual input: SM, ST, and SK. Drafting manuscript: TI. Revising manuscript content: SM, ST, and SK. Approving final version of manuscript: all authors. TI, SM, and SK take responsibility for the integrity of the data analysis.

References

1. Kenny FM, Linarelli L. Dwarfism and cortical thickening of tubular bones. Transient hypocalcemia in a mother and son. Am J Dis Child. 1966;111(2):201–7.
2. Caffey J. Congenital stenosis of medullary spaces in tubular bones and calvaria in two proportionate dwarfs—mother and son; coupled with transitory hypocalcemic tetany. Am J Roentgenol Radium Ther Nucl Med. 1967;100(1):1–11.
3. Fanconi S, Fischer JA, Wieland P, et al. Kenny syndrome: evidence for idiopathic hypoparathyroidism in two patients and for abnormal parathyroid hormone in one. J Pediatr. 1986;109(3):469–75.
4. Lee WK, Vargas A, Barnes J, Root AW. The Kenny-Caffey syndrome: growth retardation and hypocalcemia in a young boy. Am J Med Genet. 1983;14(4):773–82.
5. Franceschini P, Testa A, Bogetti G, et al. Kenny-Caffey syndrome in two sibs born to consanguineous parents: evidence for an autosomal recessive variant. Am J Med Genet. 1992;42(1):112–6.
6. Sabry MA, Zaki M, Abul Hassan SJ, et al. Kenny-Caffey syndrome is part of the CATCH 22 haploinsufficiency cluster. J Med Genet. 1998;35(1):31–6.
7. Sabry MA, Farag TI, Shaltout AA, et al. Kenny-Caffey syndrome: an Arab variant? Clin Genet. 1999;55(1):44–9.
8. Parvari R, Hershkovitz E, Grossman N, et al. Mutation of TBCE causes hypoparathyroidism-retardation-dysmorphism and autosomal recessive Kenny-Caffey syndrome. Nat Genet. 2002;32(3):448–52.
9. Tadaki H, Tokuhiro E, Shiga K, Kikuchi N, Mukai N, Fujieda K. A case of a 2-year-old girl with Kenny-Caffey Syndrome Type 2. Clin Pediatr Endocrinol. 2005;14(2 Suppl):22.
10. Oda Y, Ono R, Hiwatari M, Iwasaki H, Namai Y, Iimori Y. A case report: three-year-old boy of Kenny-Caffey Syndrome Type 2. Clin Pediatr Endocrinol. 2000;9(2):140.
11. Yorifuji T, Muroi J, Uematsu A. Kenny-Caffey syndrome without the CATCH 22 deletion. J Med Genet. 1998;35(12):1054.
12. Izumi Y, Tanae A, Kuratsuji T, et al. A case of 12-year-old boy with Kenny syndrome associated with hypomagnesemia and humoral immunodeficiency [in Japanese]. Shoninaika. 1987;19(10):1503–2.
13. Unger S, Go´rna MW, Le Béchec A, et al. FAM111A mutations result in hypoparathyroidism and impaired skeletal development. Am J Hum Genet. 2013 May 14. [Epub ahead of print].
14. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754–60.
15. Li H, Handsaker B, Wysoker A, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25(16):2078–9.
16. Kong A, Frigge ML, Masson G, et al. Rate of de novo mutations and the importance of father's age to disease risk. Nature. 2012;488(7412):471–5.
17. Fine DA, Rozenblatt-Rosen O, Padi M, et al. Identification of FAM111A as an SV40 host range restriction and adenovirus helper factor. PLoS Pathog. 2012;8(10):e1002949.
18. Akamatsu S, Takata R, Haiman CA, et al. Common variants at 11q12, 10q26 and 3p11.2 are associated with prostate cancer susceptibility in Japanese. Nat Genet. 2012;44(4):426–9, S421.
19. Shiang R, Thompson LM, Zhu YZ, et al. Mutations in the transmembrane domain of FGFR3 cause the most common genetic form of dwarfism, achondroplasia. Cell. 1994;78(2):335–42.
20. Weinstein LS, Shenker A, Gejman PV, Merino MJ, Friedman E, Spiegel AM. Activating mutations of the stimulatory G protein in the McCune-Albright syndrome. N Engl J Med. 1991;325(24):1688–95.
21. Gensure RC, Mäkitie O, Barclay C, Chan C, Depalma SR, Bastepe M, Abuzahra H, Couper R, Mundlos S, Sillence D, Ala Kokko L, Seidman JG, Cole WG, Jüppner H. A novel COL1A1 mutation in infantile cortical hyperostosis (Caffey disease) expands the spectrum of collagen-related disorders. J Clin Invest. 2005;115(5):1250–7.
22. Courtens W, Wuyts W, Poot M, Szuhai K, Wauters J, Reyniers E, Eleveld M, Diaz G, Nöthen MM, Parvari R. Hypoparathyroidism-retardation-dysmorphism syndrome in a girl: A new variant not caused by a TBCE mutation--clinical report and review. Am J Med Genet A. 2006;140(6):611–7.
23. Shoback D, Clinical practice. Hypoparathyroidism. N Engl J Med. 2008;359(4):391–403.
24. Zajac JD, Danks JA. The development of the parathyroid gland: from fish to human. Curr Opin Nephrol Hypertens. 2008;17(4):353–6.
25. Boynton JR, Pheasant TR, Johnson BL, Levin DB, Streeten BW. Ocular findings in Kenny's syndrome. Arch Ophthalmol. 1979;97(5):896–00.
26. Parvari R, Diaz GA, Hershkovitz E. Parathyroid development and the role of tubulin chaperone E. Horm Res. 2007;67(1):12–21.
27. Wilson MG, Maronde RF, Mikity VG, Shinno NW. Dwarfism and congenital medullary stenosis (Kenny syndrome). Birth Defects Orig Artic Ser. 1974;10(12):128–32.

Journal of Human Genetics (2014) 59, 163–172

www.nature.com/jhg

ORIGINAL ARTICLE

Molecular epidemiology and clinical spectrum of hereditary spastic paraplegia in the Japanese population based on comprehensive mutational analyses

Hiroyuki Ishiura[1], Yuji Takahashi[1], Toshihiro Hayashi[1], Kayoko Saito[2], Hirokazu Furuya[3], Mitsunori Watanabe[4], Miho Murata[5], Mikiya Suzuki[6], Akira Sugiura[7], Setsu Sawai[8,9], Kazumoto Shibuya[10], Naohisa Ueda[11,12], Yaeko Ichikawa[1], Ichiro Kanazawa[13], Jun Goto[1] and Shoji Tsuji[1]

Hereditary spastic paraplegia (HSP) is one of the most genetically heterogeneous neurodegenerative disorders characterized by progressive spasticity and pyramidal weakness of lower limbs. Because >30 causative genes have been identified, screening of multiple genes is required for establishing molecular diagnosis of individual patients with HSP. To elucidate molecular epidemiology of HSP in the Japanese population, we have conducted mutational analyses of 16 causative genes of HSP (*L1CAM*, *PLP1*, *ATL1*, *SPAST*, *CYP7B1*, *NIPA1*, *SPG7*, *KIAA0196*, *KIF5A*, *HSPD1*, *BSCL2*, *SPG11*, *SPG20*, *SPG21*, *REEP1* and *ZFYVE27*) using resequencing microarrays, array-based comparative genomic hybridization and Sanger sequencing. The mutational analysis of 129 Japanese patients revealed 49 mutations in 46 patients, 32 of which were novel. Molecular diagnosis was accomplished for 67.3% (33/49) of autosomal dominant HSP patients. Even among sporadic HSP patients, mutations were identified in 11.1% (7/63) of them. The present study elucidated the molecular epidemiology of HSP in the Japanese population and further broadened the mutational and clinical spectra of HSP.
Journal of Human Genetics (2014) **59**, 163–172; doi:10.1038/jhg.2013.139; published online 23 January 2014

Keywords: array-based comparative genomic hybridization; hereditary spastic paraplegia; resequencing microarray

INTRODUCTION

Hereditary spastic paraplegia (HSP) is a neurodegenerative disorder characterized by progressive lower limb spasticity and pyramidal weakness, which is one of the most genetically and clinically heterogeneous disorders.[1,2] HSP is clinically divided into two forms, pure and complicated forms, depending on whether the neurological symptoms are basically confined to spasticity and pyramidal weakness of the lower limbs or accompanied by additional neurological symptoms such as cognitive dysfunction, cerebellar signs, optic atrophy, retinitis pigmentosa, amyotrophy and peripheral neuropathy. HSP is characterized by enormous genetic heterogeneity; to date, more than 50 genetic loci (SPG1-57) and 37 causative genes have been identified: *L1CAM* (SPG1), *PLP1* (SPG2), *ATL1* (SPG3A), *SPAST* (SPG4), *CYP7B1* (SPG5A), *NIPA1* (SPG6), *SPG7* (SPG7), *KIAA0196* (SPG8), *KIF5A* (SPG10), *SPG11* (SPG11), *RTN2* (SPG12), *HSPD1* (SPG13), *SPG15/ZFYVE26* (SPG15), *BSCL2* (SPG17), *ERLIN2* (SPG18), *SPG20* (SPG20), *SPG21* (SPG21), *DDHD1* (SPG28), *KIF1A* (SPG30), *REEP1* (SPG31), *ZFYVE27* (SPG33), *FA2H* (SPG35), *PNPLA6* (SPG39), *SLC33A1* (SPG42), *GJC2* (SPG44), *GBA2* (SPG46), *AP4B1* (SPG47), *KIAA0415* (SPG48), *TECPR2* (SPG49), *AP4M1* (SPG50), *AP4E1* (SPG51), *AP4S1* (SPG52), *VPS37A* (SPG53), *DDHD2* (SPG54), *C12ORF65* (SPG55), *CYP2U1* (SPG56), and *TFG* (SPG57).

Because of the limited availability of information on genotype–phenotype correlations and locus heterogeneity, it is often difficult to prioritize genes for the mutational analysis of HSP. Therefore, it is essential to incorporate knowledge of the molecular epidemiology of HSP and relative frequencies of the types of mutations (substitution, insertion/deletion or rearrangement) in each gene into the algorithm of molecular diagnosis of HSP. We also need to be aware that different

[1]Department of Neurology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan; [2]Institute of Medical Genetics, Tokyo Women's Medical University, Tokyo, Japan; [3]Department of Neurology, Neuro-Muscular Center, National Omuta Hospital, Fukuoka, Japan; [4]Department of Neurology, Institute of Brain Science, Hirosaki University Graduate School of Medicine, Aomori, Japan; [5]Department of Neurology, National Center of Neurology and Psychiatry, Tokyo, Japan; [6]Department of Neurology, Higashisaitama Hospital, National Hospital Organization, Saitama, Japan; [7]Department of Neurology, Shizuoka Institute of Epilepsy and Neurological Disorders, Shizuoka, Japan; [8]Department of Molecular Diagnosis, Graduate School of Medicine, Chiba University, Chiba, Japan; [9]Division of Laboratory Medicine and Clinical Genetics, Chiba University Hospital, Chiba, Japan; [10]Department of Neurology, Graduate School of Medicine, Chiba University, Chiba, Japan; [11]Department of Neurology, Chigasaki Municipal Hospital, Kanagawa, Japan; [12]Department of Neurology, Yokohama City University School of Medicine, Kanagawa, Japan and [13]Graduate School, International University of Health and Welfare, Tokyo, Japan
Correspondence: Dr S Tsuji, Department of Neurology, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan.
E-mail: tsuji@m.u-tokyo.ac.jp
Received 13 September 2013; revised 16 November 2013; accepted 29 November 2013; published online 23 January 2014

methodologies are required to detect each type of mutations with high sensitivities. Although there have been studies on molecular epidemiology focusing on selected causative genes in large case series,[3–10] there have been only few studies based on comprehensive mutational analyses focusing on multiple genes as well as various types of mutation. Thus, the comprehensive molecular epidemiology of HSP is largely unestablished.

For these reasons, a comprehensive mutational analysis of multiple genes is necessary to efficiently provide molecular diagnosis for individual HSP patients and, furthermore, to clarify the molecular epidemiology of HSP. To accomplish high sensitivities for detection of various kinds of mutations, we have conducted comprehensive mutational analyses incorporating custom-made resequencing microarrays,[11] which enable comprehensive detection of single-nucleotide variations, custom-made comparative genomic hybridization (CGH) microarrays,[12] which enable efficient detection of large insertion/deletion variants, and Sanger sequencing, which enables detection of small insertions/deletions in addition to single base substitutions. We herein describe molecular epidemiology and the clinical spectrum of HSP based on a large-scale comprehensive mutational analysis of 129 Japanese patients with various forms of HSP.

SUBJECTS AND METHODS

Patients

One hundred and twenty-nine Japanese patients (75 male and 54 female) with a clinical diagnosis of HSP were enrolled in the study, including 45 patients who visited the University of Tokyo Hospital and 89 patients referred to our Department of Neurology, the University of Tokyo Tokyo, Japan for the molecular diagnosis of HSP from various regions in Japan between 1994 and July 2007. Genomic DNAs were prepared from peripheral blood leukocytes or an autopsied brain (1 patient) using a standard procedure. Written informed consent was obtained from all the participants or their family members. The study was approved by the institutional review board of the University of Tokyo.

Outline of mutational analysis system

The outline of the comprehensive mutational analysis is shown in Figure 1. All the samples were first subjected to resequencing microarray analysis for analyzing 13 causative genes of HSP. Among the patients in whom mutations were not detected by resequencing microarray analysis, direct nucleotide sequence analysis of *SPAST* and *REEP1* was carried out in patients with autosomal dominant HSP (AD-HSP) and sporadic pure-form HSP; in patients with a thin corpus callosum and cognitive impairment, direct nucleotide sequence analysis of *SPG11* was carried out. For those in which mutations were not detected by any of these methods, array-based CGH (aCGH) analysis was carried out.

Mutational analysis using custom-made resequencing microarrays

We developed resequencing microarrays using GeneChip CustomSeq (Affymetrix, Santa Clara, CA, USA).[11] We utilized custom-designed microarrays of the 30-kb format that contain tiled sequences for *SPAST* (NM_014946.3) and *ATL1* (NM_015915) (TKYPD01), those for *SPG7* (NM_003119) (TKYALS01), those for *L1CAM* (NM_000425) and *PLP1* (NM_000533) (TKYAD01) and those for *NIPA1* (NM_144599), *KIF5A* (NM_004984) and *SPG20* (NM_015087) (TKYPD02), as previously described.[13–15] In this study, we additionally designed two 50-kb-format microarrays. One was TKYPD03, which contained tiled sequences for *SPAST*, *ATL1* and *REEP1*, and the other was TKYALS02, which contained tiled sequences for 10 causative genes (*L1CAM* (NM_000425), *PLP1* (NM_000533), *NIPA1* (NM_144599), *SPG7* (NM_003119), *KIAA0196* (NM_014846), *KIF5A* (NM_004984), *HSPD1* (NM_002156), *BSCL2* (NM_001122955), *SPG20* (NM_015087) and *SPG21* (NM_016630)), enabling mutational analysis of 13 causative genes of HSP. Experiments were performed following the manufacturer's instructions (Supplementary Information S1). Data were analyzed using GeneChip DNA Analysis Software version 2.0 (GDAS2.0) for 30-kb-format microarrays,[16] or updated GeneChip Sequence Analysis Software version 4.0 (GSEQ4.0) adjunctively with a custom-designed program (Supplementary Information S2). All the called mutations were verified by direct nucleotide sequence analysis. The frequencies of detected nonsynonymous variations in the populations were checked using dbSNP (http://www.ncbi.nlm.nih.gov/snp/), the 1000 genomes database (http://www.1000genomes.org/) and by screening of >150 control chromosomes by direct nucleotide sequence analysis or restriction fragment length polymorphism analysis.

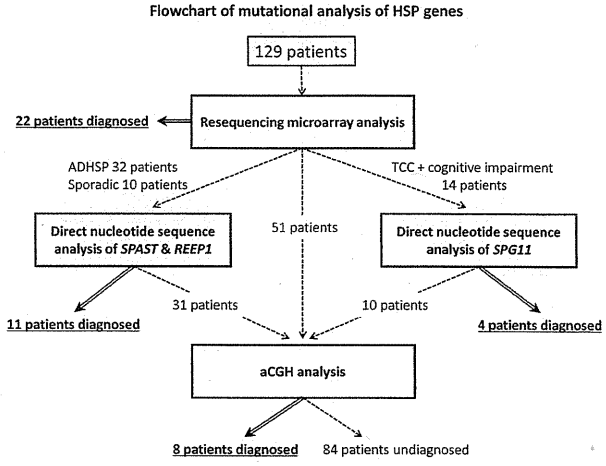

Figure 1 Flowchart of mutational analysis of HSP genes. We first performed resequencing microarray analysis, which could analyze 13 causative genes of HSP. Then, samples of mutation-negative AD-HSP and 10 sporadic pure HSP patients were subjected to direct nucleotide sequence analysis of *SPAST* and *REEP1*, because small insertion/deletion mutations are relatively frequent in these genes. Samples of mutation-negative HSP patients with a thin corpus callosum and cognitive impairment were subjected to direct nucleotide sequence analysis of *SPG11*. Finally, aCGH analysis was performed in the 92 mutation-negative patients.

Mutational analysis using aCGH

We custom-designed a CGH microarray (Agilent Technology, Santa Clara, CA, USA) in the 8 × 15 K format.[12] The genomic sequences of 16 causative genes of HSP and the flanking regions (*L1CAM*, *PLP1*, *ATL1*, *SPAST*, *CYP7B1*, *NIPA1*, *SPG7*, *KIAA0196*, *KIF5A*, *SPG11*, *HSPD1*, *BSCL2*, *SPG20*, *SPG21*, *REEP1* and *ZFYVE27*) were tiled on the array (Supplementary Information S1). The average distance between probes was 200 bp. When insertion/deletion mutations were detected by the aCGH analysis, breakpoints were determined by PCR analysis using primer pairs flanking the breakpoints and direct nucleotide sequence analysis.[17,18]

Direct nucleotide sequence analysis

Direct nucleotide sequence analysis was performed using ExoSAP-IT (USB, Cleveland, OH, USA), a BigDye Terminator v3.1 kit, and XTerminator using an ABI PRISM3100 sequencer (Applied Biosystems, Foster City, CA, USA). Primers and the amplification condition are described in Supplementary Information S1.

RESULTS

Demographic characteristics of patients

The demographic characteristics of the 129 HSP patients enrolled in this study are summarized in Table 1. The ages at onset of the patients classified on the basis of the clinical form of HSP (Supplementary Figure S1A) revealed a bimodal distribution in patients with pure-form HSP, whereas one large peak for juvenile onset and one small

Table 1 Characteristics of the HSP patients

No. of HSP patients		129
Male		75 (58.1%)
Female		54 (41.9%)
Male: female		1.4: 1
Ages at onset		0–70 y.o. (30 ± 20 y.o.)
Clinical phenotype		
Pure		82 (63.5%)
Complicated		47 (36.4%)
Family history		
Positive family history		
ADHSP		49
	Pure	44 (89.8% = 44/49)
	Complicated	5 (10.2% = 5/49)
ARHSP		11
	Pure	4 (36.4% = 4/11)
	Complicated	7 (63.6% = 7/11)
Familial (undetermined mode of inheritance)		6
	Pure	3 (50% = 3/6)
	Complicated	3 (50% = 3/6)
Sporadic		63
Sporadic with consanguinity		9
	Pure	5 (55.6% = 5/9)
	Complicated	4 (44.4% = 4/9)
Sporadic without consanguinity		54
	Pure	26 (48.1% = 26/53)
	Complicated	28 (51.9% = 28/53)

peak of adult to late onset were observed in patients with complicated-form HSP. Focusing on the mode of inheritance (Supplementary Figure S1B), the ages at onset of AD-HSP patients and sporadic HSP patients showed a similar bimodal distribution, whereas those of autosomal recessive HSP (AR-HSP) patients showed a skewed distribution.

We found 14 patients with complicated-form HSP with a thin corpus callosum and cognitive impairment. There were no patients with AD-HSP with motor neuropathy clinically diagnosed as Silver syndrome.[19]

Mutational analysis by resequencing microarray analysis

All the samples were first subjected to resequencing microarray analysis (Figure 1). The analysis detected 22 mutations, all of which were nucleotide substitutions (Table 2). Representative resequencing microarray data on a heterozygous mutation in *SPAST* (c.1493 + 2 T > C) are shown (Figures 2a–c). Using GDAS2.0 or GSEQ4.0, the overall call rate was about 90%. Except for the nucleotides for which base calling was difficult because of high GC contents, G/C stretches or locally repetitive polymorphic sequences (such as a GCG stretch in exon 1 of *NIPA1*), the overall rate of base calling using GDAS/GSEQ4.0 in combination with visual inspection was as high as 99.9%.

The custom-designed program detected one mutation (c.1741C > T, p.R581* in *SPAST*), which was not detected by GSEQ4.0, increasing the sensitivity of mutation detection (Supplementary Information S2). No additional base substitutions were detected by the subsequent direct nucleotide sequence analysis of *SPAST* and *REEP1* in AD-HSP patients.

Mutational analysis by Sanger sequencing

We applied Sanger sequencing of *SPAST* and *REEP1* in patients with AD-HSP and sporadic pure-form HSP, in which mutations were not detected by the resequencing microarray analysis, mainly to detect insertion/deletion mutations that are common in these diseases (Figure 1). We detected 10 small insertion/deletion mutations (1–41 bp) in *SPAST* and 1 insertion in *REEP1* (Table 2).

In patients with a thin corpus callosum and cognitive impairment among the cases in which mutations were not detected by the resequencing microarray analysis, we then applied Sanger sequencing of *SPG11* (Figure 1). We found homozygous or compound heterozygous mutations of *SPG11* in five patients with family histories consistent with the autosomal recessive mode of inheritance. In a sporadic HSP patient, we found compound heterozygous mutations of c.1735 + 2 delT and c.6999 + 5 delG, both of which were considered pathogenic because no other pathogenic alleles were detected, and *in silico* analysis of splicing scores of c.6999 + 5 delG showed scores that decreased from 10.1 to 7.1 (http://rulai.cshl.edu/new_alt_exon_db2/HTML/score.html) and from 1.0 to 0.62 (http://www.fruitfly.org/seq_tools/splice.html). Another patient had only one null allele (p.R2031*) in *SPG11* and no other pathogenic alleles were found.

Detection of large deletions and duplications by aCGH analysis

We applied aCGH analysis in patients in whom mutations were not detected by resequencing microarray analysis and direct nucleotide sequencing analysis (Figure 1). Representative results for a heterozygous large deletion in *KIAA0196* are shown in Figure 2d, in which the breakpoint sequence was clearly determined using an appropriate primer pair (Figure 2e).

In total, we identified 7 large deletions (4 in *SPAST*, 1 in *REEP1*, 1 in *KIAA0196*, and 1 in *SPG11*) and 1 duplication in *SPAST* in 92 patients examined (Table 2). All the breakpoints but one (in which the deletion was beyond the tiled sequences on the array) were unequivocally determined at the nucleotide level. The sizes of the deletions/duplication ranged from 4634 bp to > 170 kb.

Five (four *SPAST* deletions and one *SPG11* deletion) of the seven deletion breakpoints were inside *Alu* sequences. Among the breakpoint sequences of the remaining deletions, the deletion in *REEP1* and the deletion in *KIAA0196* showed microhomology of 3 bp. The breakpoint sequence of the *SPAST* duplication showed no homologous sequences. In total, five of the eight breakpoints (62.5%) were inside *Alu* sequences.

Molecular epidemiology of HSP in Japanese population

In summary, we found 49 mutations in 46 patients (Table 2). The relative frequencies of individual HSP types classified on the basis of the clinical presentations and the mode of inheritance are summarized in Figures 3a and b.

Focusing on all AD-HSP patients, SPG4 (55.1%) was the most frequent. SPG3A (2.0%), SPG8 (4.1%), and SPG31 (4.1%) were relatively rare in Japanese HSP patients. The frequency of SPG3A is lower than that in the Caucasian populations.[5] The frequencies of SPG8 (4.1%) and SPG31 (4.1%) in AD-HSP are comparable to those reported in Caucasian populations.[20,21]

In the AR-HSP group, we found five families with SPG11 and one family with SPG21. Among the 14 patients with a thin corpus

Table 2 Mutation and clinical summary of Japanese HSP patients

	Mutation	*Amino-acid change*	*Detecting method*	*Family history*	*Age at onset of index patient*	*Clinical phenotype*		*Reference*
SPG4								
Exon 1	c.139 A>T	p.K47*	R	AD	40s	Pure	Novel	
Exon 1	c.155 A>G	p.Y52C	R	Sporadic with consanguineous parents	49 y.o.	Pure	Novel	
Exon 1	c.283_323 del	p.A95Afs	D	AD	40 y.o.	Pure	Novel	
Exon 1	c.343_352 dup	p.V118Afs	D	AD	35 y.o.	Pure	Novel	
Exon 2	c. 422–425 delAGAA	p.Q141fs	D	AD	36 y.o.	Pure	Novel	
Exon 2	c. 422–425 delAGAA	p.Q141fs	D	AD	51 y.o.	Pure	Novel	
Exon 2	c. 422–425 delAGAA	p.Q141fs	D	Sporadic	35 y.o.	Pure	Novel	
Exon 3	c.532 C>T	p.Q178*	R	AD	33 y.o.	Pure	Novel	
Exon 5	c.734 C>G	p.S245*	R	AD	teens	Pure	Known	35
Exon 5	c.838 C>T	p.Q280*	R	AD	~6 y.o.	Pure	Novel	
Exon 6	c.871 delG	p.G291Vfs	D	AD	20 y.o.	Pure	Novel	
Intron 6	c.1005-2 A>G	IVS6-2A>G	R	AD	2 y.o.	Pure	Known	4
Exon 7	c. 1014 delT	p.A338Afs	D	AD	40s	Pure	Novel	
Exon 8	c.1105 A>C	p.T369P	R	AD	38 y.o.	Pure	Novel	
Exon 8	c.1141 C>T	p.F381L	R	Sporadic	<6 y.o.	Pure	Known	4
Exon 8	c.1141 C>T	p.F381L	R	AD	late 50s	Pure	Known	4
Intron 8	c.1173+1 G>A	IVS8+1G>A	R	AD	46 y.o.	Pure	Known	3
Exon 11	c.1378 C>T	p.R460C	R	AD	27 y.o.	Pure	Known	36
Exon 12	c.1426_1427 delGG	p.G476Rfs	D	AD	39 y.o.	Pure	Novel	
Intron 12	c.1493+2 T>C	IVS12+2T>C	R	Sporadic	40 y.o.	Pure	Known	35
Exon 13	c.1504 A>T	p.K502*	R	AD	30 y.o.	Pure	Novel	
Exon 13	c.1507 C>T	p.R503W	R	AD	~ 10 y.o.	Pure	Known	37
Exon 15	c.1646 insT	p.L549Lfs	D	AD	34 y.o.	Pure	Novel	
Exon 15	c.1646 insT	p.L549Lfs	D	AD	47 y.o.	Pure	Novel	
Exon 15	c.1646 T>C	p.L549P	R	AD	< 15 y.o.	Pure	Novel	
Exon 15	c.1688 G>A	p.R562Q	R	AD	~10 y.o.	Pure	Known	38
Exon 17	c.1741 C>T	p.R581*	R	AD	14 y.o.	Pure	Known	39
Promoter~intron 1	del Chr2:32136286–32145830 (9545 bp)	Del ex1	aCGH	AD	40 y.o.	Pure	Novel	
Intron 1~3′ region	(>170 kb)	Del ex2-17	aCGH	Affected sibling	24 y.o.	Pure	Novel	
Intron 16~3′ region	del Chr2:32290425–32231940 (58 482 bp)	Del ex17	aCGH	AD	58 y.o.	Pure	Novel	
Intron 16~3′ UTR	del Chr2:32229622–32234715 (5094 bp)	Del ex17	aCGH	AD	52 y.o.	Pure	Novel	
Exon 4~intron 7	dup Chr2:32177411–32199467 (22 057 bp)+insAGT	Tandem duplication (part of ex4-ex7)	aCGH	AD	< 6 y.o.	Pure	Novel	
SPG3A								
Exon 12	c.1243 C>T	p.R415W	R	AD	12 y.o.	Pure	Known	40
Exon 12	c.1483 C>T	p.R495W	R	Sporadic	~12 y.o.	Pure	Known	41
SPG8								
Exon 13	c.1749 A>C	p.R583S	R	AD	50 y.o.	Pure	Novel	
Intron 10~exon15	del Ch8:126 138 189–126 142 822 (4634 bp)	Del exon11-15	aCGH	AD	64 y.o.	Pure	Novel	
SPG17								
Exon 2	c. 107 G>A (c. 299 G>A)	p.C36Y (p.C100Y)	R	AD	~10 y.o.	Complicated	Novel	
SPG31								
Exon 2	c.87 insA	p.K30Kfs	D	AD	8 y.o.	Pure	Novel	
Intron 3~intron 5	del Ch2: 86 326 358–86 338 428 (12 064 bp)	Del exon 4-5	aCGH	AD	~12 y.o.	Pure	Novel	

Table 2 (Continued)

	Mutation	*Amino-acid change*	*Detecting method*	*Family history*	*Age at onset of index patient*	*Clinical phenotype*	*Reference*
SPG11							
Intron 18	c.3291+1 G>T	IVS18+1G>T (homozygous)	D	Consanguinity affected siblings	20 y.o.	Complicated Known	42
Intron 18	c.3291+1 G>T	IVS18+1G>T (homozygous)	D	Consanguinity affected sibling	25 y.o.	Complicated Known	42
Intron 8 and intron 38	c.1735+2 delT, c.6999+5 delG	IVS8+2 delT, IVS38+5 delG	D, D	Sporadic	22 y.o.	Complicated Novel	
Exon 20 and exon 28	c.3491G>A, c.4840T>A	p.W1164*, p.K1614*	D, D	Sporadic	18 y.o.	Complicated Novel	
Intron 7 ~ intron 8 and exon 25	del Chr15:42709367–42715955 (6589 bp), c.4426 insAT	Del exon8, p.C1476Yfs	aCGH, D	Affected sibling	2 y.o.	Complicated Novel	
SPG21 Exon 4	c.322 G>C	p.A108P (homozygous)	R	Familial	60 y.o.	Complicated Novel	

Abbreviations: aCGH, array-based comparative genomic hybridization analysis; AD, autosomal dominant; D, direct nucleotide sequence analysis; Del, deletion; dup, duplication; y.o., years old; R, resequencing microarray analysis; UTR, untranslated region.
All the patients presented the pure form. +1 of nucleotides is the first A of the start codon (ATG). The NCBI36/hg18 assembly is used as the reference genome.

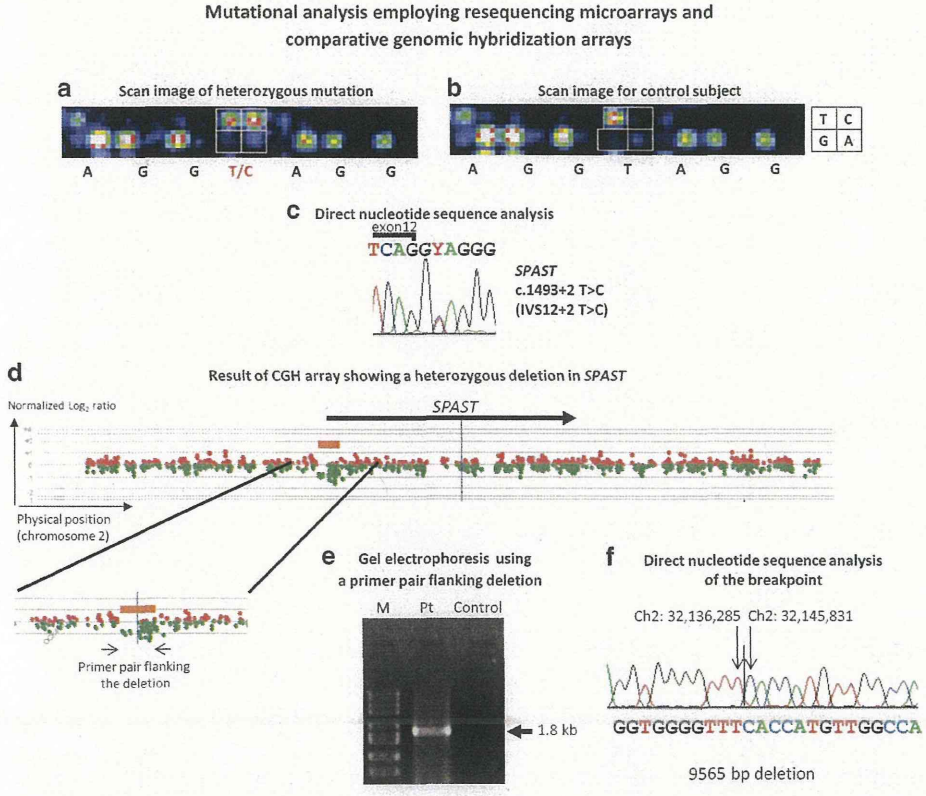

Figure 2 Mutational analysis using resequencing microarrays and comparative genomic hybridization microarrays. (**a**) This figure shows a scan image obtained by resequencing microarray analysis (TKYPD03) of a sporadic HSP patient. Each tile in a square indicates one of the four nucleotides. Depending on the nucleotide of each allele, each quadrant provides a fluorescent signal. As shown in a square that corresponds to the position of c.1493+2 of *SPAST*, the upper left tile and the upper right tile, which correspond to T and C, respectively, provided similarly intense hybridization signals. The signal pattern indicates the existence of the T allele (wild type) and the C allele (variant) in that position. (**b**) Scan image of the same positions of the resequencing microarray as those in panel (**a**) obtained from a mutation-negative patient, where only the upper left tile corresponding to 'T' gives an intense fluorescent signal. (**c**) Heterozygous c.1493+2 T>C mutation confirmed by direct nucleotide sequence analysis, which is expected to disrupt the consensus splice donor site. (**d**) Example of comparative genomic hybridization analysis. The vertical axis indicates the $\log_2$ ratio of hybridization signal intensities obtained from a patient with SPG4 and a male control subject. The horizontal axis indicates the physical position of oligonucleotide probes. If copy number variations do not exist, the $\log_2$ ratios of the hybridization signal intensities are expected to be near 0. In the region indicated by an orange bar, the $\log_2$ ratio of hybridization signal intensities is approximately −1, which indicates a heterozygous deletion (halved gene dosage) in *SPAST*. (**e**) PCR analysis using primers flanking the deletions revealed that the truncated band corresponding to 1.8 kb was detected only in the patient. No PCR product was detected in a control, because the distance between the primer pair was too long to amplify (~11 kb). (**f**) Direct nucleotide sequence analysis determines breakpoints with a deletion size of 9565 bp.

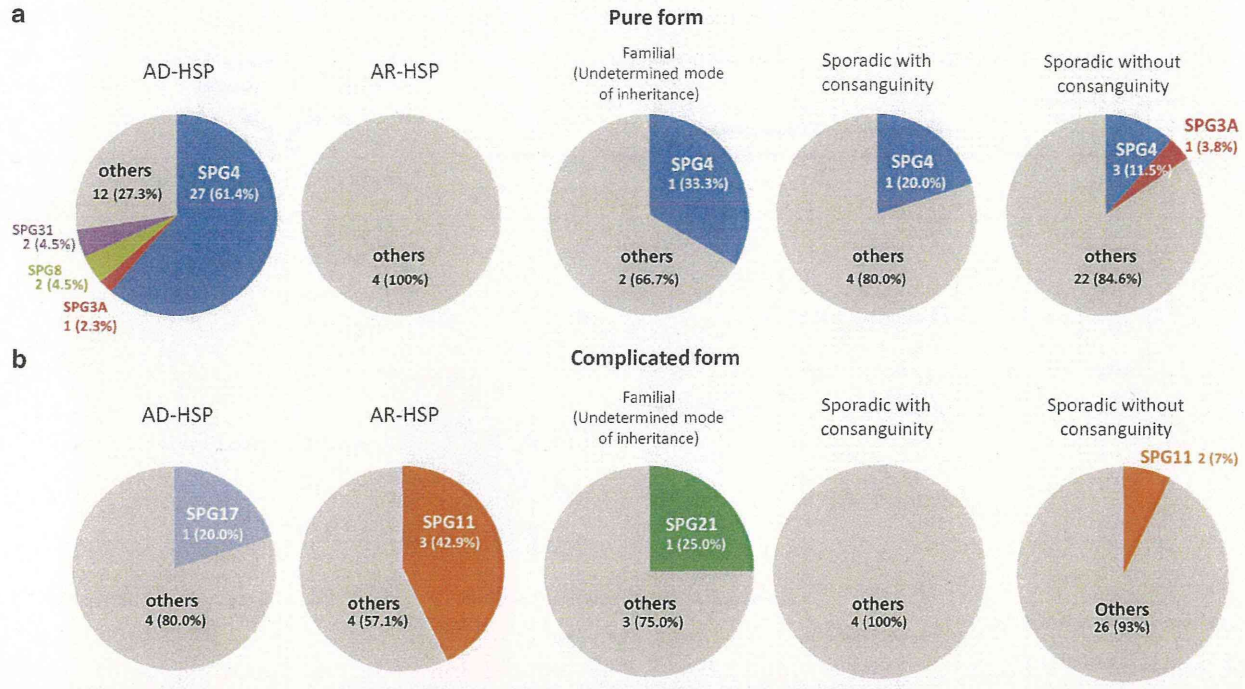

Figure 3 Relative frequencies of individual HSP types in groups classified on the basis of the clinical presentations and mode of inheritance. The figure shows the relative frequencies of individual HSP types in our cohort. (**a**) Pure form and (**b**) complicated form. The family history of each subgroup is indicated above the figures. Mutations were detected in a total of 67.3% of all the AD-HSP patients or 72.7% of the patients with pure-form AD-HSP. Focusing on sporadic HSP patients, six patients (four SPG4, one SPG3A and one SPG11) were identified, which accounted for 9.8% (6/61). Of note, *SPAST* mutations were present in 6.6% of all sporadic HSP patients, and particularly in 12.9% (4/31) of sporadic pure-form HSP patients, suggesting the usefulness of mutational analysis of *SPAST* in sporadic cases, particularly in patients with the pure form. Others, patients with unidentified mutation.

callosum and cognitive impairment, 35.7% (5/14) carried *SPG11* mutations.

Molecular and clinical spectra of individual HSP types

SPG3A. We found two patients with SPG3A carrying previously reported mutations (Table 2). Although both patients with SPG3A showed basically pure-form HSP with juvenile onset, one patient showed hypesthesia and hypalgesia in the distal lower limbs accompanied by decreased vibratory sensation in all extremities.

SPG4. Of the 32 patients with SPG4, 24 (75%) had nonsense, frameshift or large deletion/duplication mutations leading to truncated proteins, which were distributed throughout the genes (Supplementary Figure S2). On the other hand, seven out of the eight missense mutations were located in the AAA domain (ATPase associated with various cellular activities). We found a novel mutation (p.Y52C) outside the AAA domain. Note that large deletions/duplications in *SPAST* were detected by aCGH analysis, and small deletion mutations were detected by Sanger sequencing analysis in 22.7% (5/22)[22] and 45.5% (10/22) of AD-HSP patients, respectively, in whom no mutations were detected by the resequencing microarray analysis. The ages at onset of patients with SPG4 showed two peaks, in the teens and in 40 s (Supplementary Figure S3A). The types of the mutation in *SPAST* and age at onset did not correlate (Supplementary Figure S3B).

SPG8. We found a large deletion in *KIAA0196*, which has not been described to date. The breakpoints of the large deletion in *KIAA0196* are located in intron 10 and exon 15 (Figures 4a–c). RT-PCR and direct nucleotide sequence analyses revealed that exons 10–15 were deleted in cDNA, predicting a premature termination codon (Figures 4d and e). There are only three missense mutations reported to date, and in a previous paper, it was proposed that haploinsufficiency is the disease-causing mechanism of SPG8 on the basis of experiments using zebrafish.[20] The large deletion in *KIAA0196* detected in the present study further supported a disease mechanism of haploinsufficiency and indicate a necessity of screening for rearrangements of *KIAA0196* in AD-HSP. SPG8 has been reported to be an 'aggressive' subtype of HSP and the disease onset is in the 20s or 30s.[20] In contrast, two patients with SPG8 found in the study had adult-onset or late-onset HSP.

SPG11. The five patients with SPG11 showed complicated-form HSP with cognitive impairment and a thin corpus callosum. Notably, rearrangement in *SPG11* was found in a patient, and aCGH analysis was helpful for accurate diagnosis of the patient. The age at onset ranged from 2 to 25 years. Although SPG11 is allelic to juvenile amyotrophic lateral sclerosis (ALS5),[23] none of the patients showed the ALS phenotype.

SPG17. A novel *BSCL2* (NM_032667) p.C36Y substitution (which can also be called p.C100Y in NM_001122955 because there are two known start codons) was found in one AD-HSP patient. He suffered from early-onset spastic paraparesis with mild mental retardation and did not show amyotrophy. Clinical and genetic data of other family members were not available. C36 is conserved among species and is located in the first transmembrane domain,[24] raising a possibility that p.C36Y can change the function of seipin, the protein product of *BSCL2*. Because only p.N88S and p.S90L of seipin have been

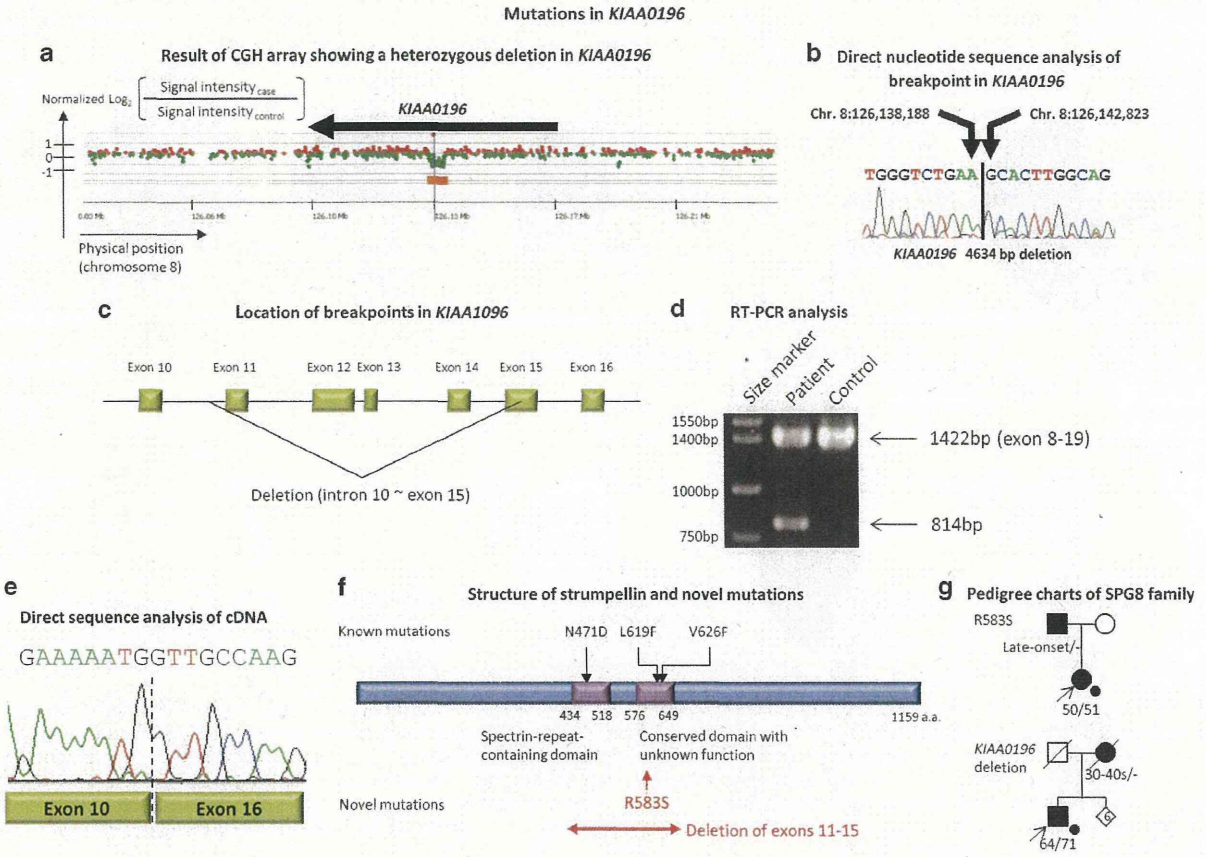

Figure 4 Mutations in *KIAA0196*. (**a**) Result of CGH array showing a heterozygous deletion in KIAA0196. An orange bar shows heterozygous deletion. (**b**) Direct nucleotide sequence analysis of the breakpoint in KIAA0196, which shows 4634 bp deletion. (**c**) Schematic presentation of the exon–intron structure of *KIAA0196*. The deletion detected by the array CGH analysis is shown. (**d**) RT-PCR analysis of species of RNAs extracted from the patient with the *KIAA0196* deletion and a control. In the control, only a single band with the expected size corresponding to 1422 bp was observed, while a truncated band with the size corresponding to 814 bp in addition to PCR products corresponding to 1422 bp was observed in the patient. (**e**) Direct nucleotide sequence analysis of the truncated PCR products revealed that exons 11–15 were absent in the *KIAA0196* mRNA as a result of a deletion in *KIAA0196*. (**f**) Schematic representation of strumpellin, the protein product of *KIAA0196*, and the mutations identified in patients with SPG8. The position of the large deletion (deletion of exons 11–14 and a part of exon 15) and the novel mutation found in the present study are shown (red). Previously reported mutations in *KIAA0196* (p.N471D, p.L619F and p.V626F) are located in the spectrin-repeat-containing domain (amino acids 434–518) or the conserved domain with unknown function (amino acids 576–649). The novel mutation (p.R583S) found in the present study is also located in the conserved domain with unknown function. (**g**) Pedigree charts of the Japanese SPG8 families. Age at onset and age at examination are indicated.

described in Silver syndrome/SPG17, we still need to be cautious about the pathogenicity of p.C36Y substitution.

SPG21. We found a novel homozygous amino-acid substitution (p.A108P) in *SPG21* encoding maspardin in a family with late-onset complicated-form HSP (Figure 5, Supplementary Tables S1 and S2). The two patients managed to walk with a cart or a cane in their 70s and 60s. In addition to cognitive decline, callosal disconnection syndrome, such as ideomotor apraxia predominantly of the left hand, agraphia of the left hand and constructional impairment predominantly on the right side, was observed, which was mild but progressed over 5 years in the index patient. There were no extrapyramidal signs, cerebellar signs or bulbar symptoms, as reported in the original family with an *SPG21* mutation.[25] Magnaetic resonance imaging of the index patient showed progressive thinning of the corpus callosum and predominantly frontotemporal atrophy (Figure 5e–i). ^{123}I-*N*-isopropyl-*p*-iodoamphetamine single-photon emission computed tomography revealed decreased blood flow in the frontal and temporal cortices (Figure 5j).

This family is the first family with SPG21 identified outside the Amish population.[25] Intriguingly, compared with the original Mast syndrome family with an *SPG21* mutation, the ages at onset of HSP symptoms in the patients in the new SPG21 family were strikingly late. Although characteristics such as cognitive decline and a thin corpus callosum were shared in common, characteristic clinical signs in Mast syndrome such as bulbar, extrapyramidal and cerebellar signs were not found in the new family (Supplementary Table S2), thus presenting dissimilar phenotypes. Because the mutation detected in the new family is a missense mutation (p.A108P) next to the active site of the alpha/beta-hydrolase domain (S109), dysfunction of alpha/beta-hydrolase activity of maspardin seems to be related to pathogenicity.

SPG31. The two novel mutations in *REEP1* were a frameshift (insertion of A) and a large deletion (Table 2), suggesting haploinsufficiency as the disease-causing mechanism. A large deletion detected in the study demanded a screening of rearrangement of *REEP1* in the diagnosis of SPG31. These two patients with SPG31 had

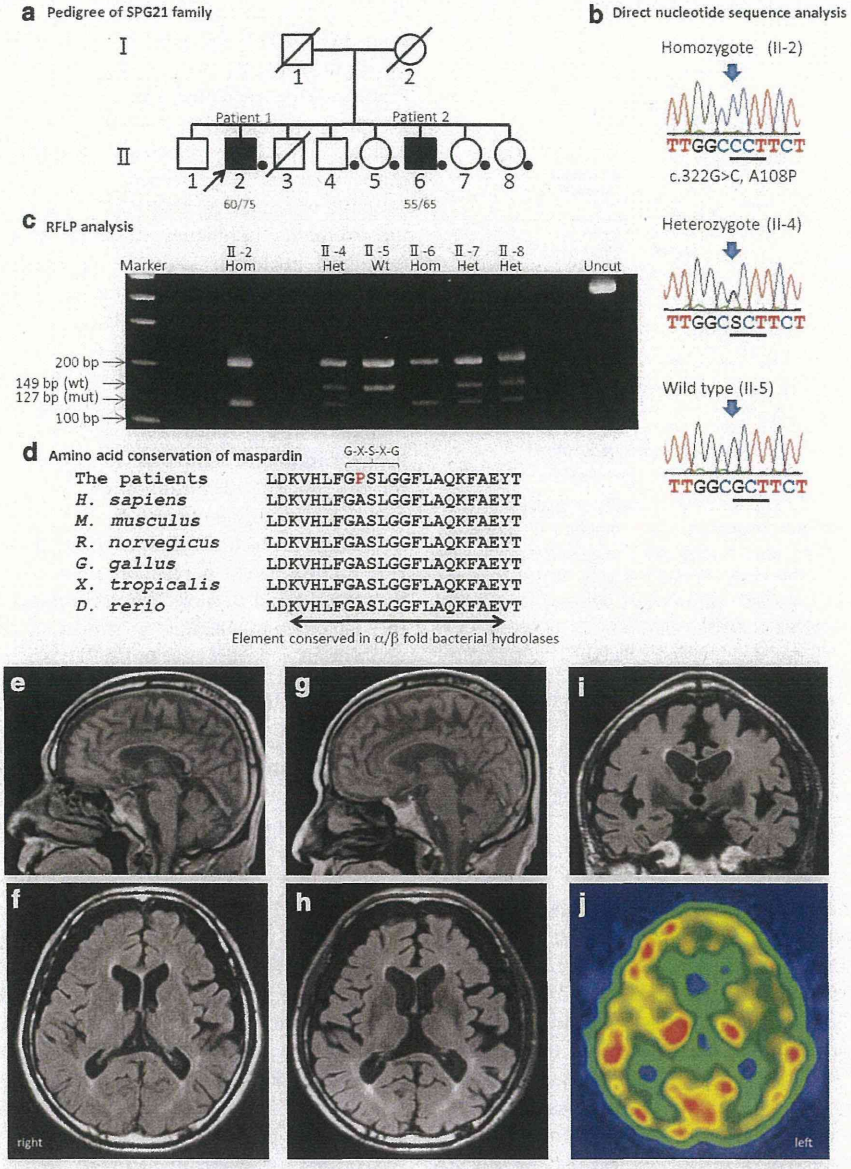

Figure 5 A family with SPG21 and molecular genetic analysis. (**a**) Pedigree tree of the family. Squares indicate males and circles indicate females. Black squares are affected members and the index patient (II-2) is indicated by an arrow. Symbols with a diagonal line indicate deceased members. Members with dots allowed us neurological and genetic examinations. (**b**) Electropherograms of the family members carrying homozygous c.322G>C mutation (II-2), heterozygous c.322G>C mutation (II-4) and wild-type allele (II-5). (**c**) PCR-restriction fragment length polymorphism (RFLP) analysis of family members. The uncut PCR fragment length is 344 bp. With *Hae*III digestion, the wild-type allele shows fragment sizes of 149 and 195 bp, whereas the mutant allele shows fragment sizes of 127, 22 and 195 bp. (**d**) Comparison of amino-acid sequence of ACP33/maspardin among species. A108 is located in the α/β-fold bacterial hydrolase domain, which is highly evolutionally conserved. The G-X-S-X-G motif at the nucleophile elbow is also shown. (**e** and **f**) A sagittal T1-weighted image (**e**) and a transverse fluid-attenuated inversion recovery (FLAIR) image (**f**) of patient 1 at the age of 70 years show a thin corpus callosum and mildly atrophic cerebrum. Atrophy in the brainstem and cerebellum is not observed. (**g**–**i**) A sagittal T1-weighted image (**g**), a transverse FLAIR image (**h**) and a coronal FLAIR image (**i**) of patient 1 at the age of 75 years shows progressive thinning of the corpus callosum mainly in the trunk and progressive atrophy of the cerebrum, which is marked in the frontal and temporal lobes. Slight white matter changes are observed around the lateral ventricles. Atrophy in the brainstem and cerebellum is not observed. (**j**) ^{123}I-*N*-isopropyl-*p*-iodoamphetamine single-photon emission computed tomography (SPECT) at the age of 75 years shows decreased blood flow in the frontal and temporal cortices. Wt, wild type; mut, mutant; Het, heterozygote; Homo, homozygote.

pure-form HSP and their disease started in their early teens, compatible with previous reports.[9,21]

Sporadic HSP. As much as 11.1% (7/63) of the patients with sporadic HSP were revealed to have mutations in the genes for monogenic diseases. Among sporadic pure-form HSP patients, 12.9% (4/31) had *SPAST* mutations, and 6.3% (2/32) of sporadic complicated-form HSP patients had *SPG11* mutations.

DISCUSSION

We herein described a comprehensive mutational analysis of as many as 16 causative genes of HSP and applied it to the mutational analysis

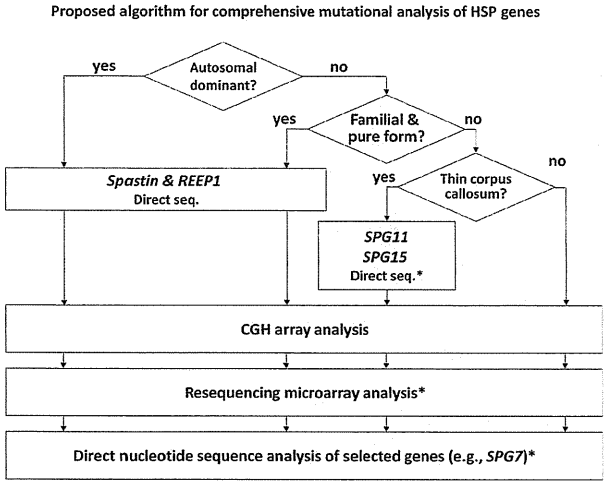

Figure 6 Proposed algorithm for comprehensive mutational analysis of HSP genes. Considering the types and frequencies of mutations in individual SPG genes, we propose an efficaceous strategy for a large-scale mutational analysis of HSP at the time of the study. In patients with ADHSP patients and in familial pure HSP patients with an unknown mode of inheritance, direct nucleotide sequence analysis of *spastin* and *REEP1* followed by CGH analysis is recommended, considering the relatively high frequency of small insertions/deletions in *spastin* and *REEP1* and large deletions/duplications in *spastin*. In patients with thin corpus callosum and/or cognitive dysfunction, *SPG11* and *SPG15* should be analyzed first. Next step is CGH analysis followed by resequencing microarray analysis, because throughput of CGH analysis is higher than that of resequencing microarray analysis. *In these days, these stages can be replaced by whole genome or exome sequencing. Direct seq., direct nucleotide sequence analysis.

of 129 Japanese HSP patients. An epidemiological study[26] based on the Registry of the Ministry of Health, Labour and Welfare, Japan in 2002 reported about 500 HSP patients. Although there remains a possibility that some patients may have not been registered for various reasons, the collection of 129 patients should represent a substantial proportion of Japanese HSP patients. In the 129 HSP patients, we identified 49 mutations, 32 of which were novel. Resequencing microarray and aCGH analyses were proved to be efficacious methods to detect nucleotide substitutions and large duplications/deletions, respectively. Indeed, the fact that we did not find additional base substitution mutations of *SPAST* and *REEP1* in AD-HSP patients by direct sequence analysis, for whom mutations were not detected by resequencing microarrays, indicates a false-negative rate of resequencing microarray analysis was low, if any, by tuning up by our algorithm (a computer program). However, note also that both resequencing microarray and aCGH analyses did not detect small insertion/deletion mutations, and direct nucleotide sequence analysis was needed to detect them. Our results revealed that the combination of these technologies, including resequencing microarray, aCGH, and direct nucleotide sequence analyses, are essential to detect various kinds of mutations, including base substitutions, and insertions/deletions of various sizes with high sensitivities.

Given the results of this study, we propose an algorithm for a comprehensive mutational analysis for HSP. To analyze genes that have relatively frequent small insertion/deletion mutations (for example, *SPAST*, *REEP1*, *SPG11* and *SPG15*), direct nucleotide sequence analysis is the first priority. To analyze genes in which most of the mutations are nucleotide substitutions (for example, *ATL1*, *NIPA1*, *KIF5A*, *KIAA0196*, *HSPD1* and *BSCL2*), resequencing microarray analysis is highly suitable. Considering the throughput, direct nucleotide sequence analysis becomes more laborious as the number of exons to be sequenced increases. In contrast, it is not the case for resequencing microarray and CGH array analyses. That is, the time required for analysis remains constant with increasing number of genes or exons to be sequenced until a limit determined by the structure of arrays. We propose a strategy of utilizing high-throughput microarray techniques and minimizing the use of time-consuming direct nucleotide sequence analysis considering the molecular epidemiology and the mutation types in individual genes (Figure 6). Although there remains a possibility that uncommon mutations (for example, insertions/deletions of intermediate length) or uncommon presentation (for example, *SPAST* mutation in a family having apparently autosomal recessive mode of inheritance or *SPG11* mutations in a pseudoautosomal dominant family) are missed and it might introduce some bias, the algorithm should be highly useful for the efficient identification of the majority, if not all, of the mutations responsible for HSP.

Utilizing the technologies, we elucidated molecular epidemiology of HSP in the Japanese population. Interestingly, the study revealed that the overall trend of molecular epidemiology of AD-HSP/AR-HSP in the Japanese population is similar to those in the Caucasian populations reported previously.[3,5,6,20,21,27] In contrast, considerable differences in the epidemiology of spinocerebellar ataxias[26] or amyotrophic lateral sclerosis (especially in those who have hexanucleotide repeat expansion mutation in *C9ORF72*)[28,29] have been demonstrated, which presumably reflect founder effects.[29,30] Thus, the similarity in the molecular epidemiology of HSP irrespective of ethnicity suggests that contribution of founder effects is limited in HSP.

We did not find causative mutations in 16 AD-HSP, 8 AR-HSP and 5 familial HSP patients. Although we cannot completely exclude the possibility of false-negative results in our analyses, we assume that these undiagnosed patients would have mutations in causative genes that have recently been identified after the study (*RTN2* or *GBA2*, for example) or mutations in as yet unidentified causative genes.

The extent to which mutations of causative genes account for apparently sporadic HSP is an important but unsolved issue. We found that 7 out of the 62 sporadic HSP patients (11.1%) had mutations of genes for HSP. In particular, we found that SPG4 and SPG11 are relatively frequent in sporadic pure-form HSP and complicated-form HSP patients, respectively. The findings indicate that careful genetic counseling of such patients and families will be required.

With recent progresses in massively parallel sequencing technologies, exome and targeted sequencing are now becoming a robust method for high-throughput resequencing analysis at a relatively reasonable cost.[31–34] Detection of large insertion/deletion mutations based on the short reads generated by next-generation sequencers, however, is still a challenging task. It is of note that a substantial proportion (7/49, 14.3%) of mutations found in the study were insertions/deletions detected by aCGH analysis. Thus, combining multiple technologies, as we did in the study, is indispensable to detect as many mutations as possible even in the next-generation sequencer era. In addition, information on the relative frequencies of HSP types and on the distribution of various types of mutations in each HSP gene as shown in the study is helpful for making strategies for mutational analyses.

In summary, we elucidated the molecular epidemiology of HSP in the Japanese population combining multiple technologies of