and c.694C>T (p.Arg232Cys) mutants showed mislocalization, whereas c.466G>A (p.Asp156Asn) and c.899G>C (p.Cys300Ser) mutants showed normal localization (Figure S1). To investigate whether the *B3GALT6* missense mutations affect the enzyme function, the GalT-II activities of soluble FLAG-tagged proteins for WT and mutant B3GALT6 proteins were assayed. The GalT-II activities of p.Ser65Gly-, p.Pro67Leu-, p.Asp156Asn-, p.Arg232Cys-, and p.Cys300Ser-B3GALT6 were significantly decreased compared with WT-B3GALT6 (Figure 2C), indicating that these mutations resulted in a loss of enzyme function. On the other hand, there were no significant differences in the GalT-II activities between WT-B3GALT6 and p.Glu174Asp-B3GALT6, a common polymorphism (rs12085009) in the public database (Figure 2C).

All SEMD-JL1 individuals with the *B3GALT6* mutation had the characteristic skeletal abnormalities, including platyspondyly, short ilia, and elbow malalignment (Table 1 and Figure S2); however, some had a range of extraskeletal and connective tissue abnormalities that overlapped with those seen in Ehlers-Danlos syndrome, progeroid form (EDS-PF [MIM 130070]). EDS-PF is an autosomal-recessive connective tissue disorder characterized by sparse hair, wrinkled skin, and defective wound healing with atrophic scars.[10] In addition, skeletal abnormalities so far reported in EDS-PF are limited to generalized osteopenia and radial head dislocation, which are in contrast with the severe generalized dysplasias of the axial and appendicular skeleton observed in SEMD-JL1. Thus, both disorders at first glance appear as separate clinical entities, although they share the clinical features of short stature, joint laxity and dislocation, and facial dysmorphism. In two families with individuals with EDS-PF, recessive mutations of *B4GALT7* (MIM 604327) have been found.[11,12] *B4GALT7* (RefSeq NM_007255.2) encodes an enzyme, xylosylated protein β-1,4-galactosyltransferase, that catalyzes the second step of the GAG linker region biosynthesis (Figure 1). Therefore, we speculated that B3GALT6 and B4GALT7 deficiencies might show similar phenotypes. We then examined *B3GALT6* in four additional individuals (P9–P12) who had phenotypes compatible with EDS-PF (Table 1 and Figure S3) but in whom no *B4GALT7* mutations had been found. Sanger sequencing of the EDS-PF-like subjects revealed that all were compound heterozygotes for *B3GALT6* mutations (Table 2). There were two frameshift mutations and one missense mutation (c.925T>A [p.Ser309Thr]) common in two families (F8 and F9). We investigated the enzyme function of the missense mutation by using the same assay for SEMD-JL1 missense mutations. The GalT-II activities of p.Ser309Thr-B3GALT6 were significantly decreased (Figure S4).

Collectively, 11 different mutations in individuals from 10 families were identified in *B3GALT6* by a combination of exome and targeted sequencing (Table 2 and Figure S5). None of these mutations were detected in more than 200 ethnicity-matched controls or in public databases, including the 1000 Genomes database, indicating that they are unlikely to be polymorphisms. SEMD-JL1 and EDS-PF-like individuals had no common mutations (Table 2). The individuals with *B3GALT6* mutations were short at birth and their short stature worsened with age. Their common clinical features were a flat face with prominent forehead and kyphoscoliosis (Table 1). Kyphoscoliosis was noticed in infancy in most cases and even in utero in severe cases. Although skeletal changes were essentially the same, craniofacial and skin abnormalities, joint laxities, and muscular hypotonia were variable among the individuals with *B3GALT6* mutations. Common radiographic features were platyspondyly that becomes less conspicuous with age, short ilia, and elbow malalignment (Table 1). Prominent lesser trochanters and metaphyseal flaring were seen in most cases. No individuals showed generalized osteoporosis. The disease phenotype was very variable between families (mutations), but in two familial cases, phenotypes were similar between the pair of the sibs. As a corollary, our results indicate that EDS-PF is genetically heterogeneous, with a proportion of cases being caused by mutations in *B4GALT7* and another in *B3GALT6*.

Diseases caused by defects in enzymes involved in the biosynthesis of the GAG linker region are categorized as the GAG linkeropathy. The first member of GAG linkeropathy has been identified to arise from an EDS-PF/B4GALT7 deficiency. *B4GALT7* mutations have been identified in homozygous c.808C>T (p.Arg270Cys)[12] and compound heterozygous (c.557C>A [p.Ala186Asp] and c.617T>C [p.Leu206Pro])[11] states. Another member of GAG linkeropathy manifests itself as Larsen-like syndrome, B3GAT3 type (MIM 245600). A family with individuals harboring a homozygous *B3GAT3* (MIM 606374; RefSeq NM_012200.3) mutation (c.830G>A [p.Arg227Gln]) has been identified. The clinical features of five affected individuals of the family are characterized by dislocation and laxity of joints and congenital heart defects.[11] The former considerably overlaps with the phenotypes of SEMD-JL1 and EDS-PF, two other GAG linkeropathies; however, the association of heart defects has critically differentiated this disease from the others (Figure 1).

Given that the linker region biosynthesis is nonparallel and that the defects in the three enzymes simply affect the amounts of the linker region available to form GAGs (CS, HS, DS), phenotypic similarities of the three diseases are quite understandable. The quantitative difference of the phenotypes (severity of the diseases) most probably results from the difference in the degree of enzyme defects resulting from mutations. On the other hand, qualitative differences of the three diseases (e.g., scoliosis caused by the *B3GALT6* mutation, heart disease caused by the *B3GAT3* mutation, etc.) suggest other explanations. Tissue expression patterns of the three genes do not entirely explain the differences. We examined their mRNA expression in various human tissues, including cartilage, bone, and connective tissues by quantitative real-time PCR (Figure S6). We detected strong expression of *B3GALT6* in

Table 3. The Amount of GAGs in the Lymphoblastoid Cells from Individuals with Spondyloepimetaphyseal Dysplasia with Joint Laxity Type 1

	GAG (Disaccharides/mg Acetone Powder)[a] [pmol]			
Subject	**CS/DS**	**CS**	**DS**	**HS**
Control	62	48	29	128
SEMD-JL1				
P1	313	295	118	15
P2	345	175	60	21
P3	270	162	28	20

[a]Calculated based on the peak area in chromatograms of digests with a mixture of chondroitinases ABC and AC-II (CS/DS), chondroitinases AC-I and AC-II (CS), chondroitinase B (DS), and heparinases I and III (HS).

cartilage and bone but only weak expression in skin, ligament, and tendon. *B4GALT7* expression was stronger in cartilage than *B3GALT6* and also weak in skin and ligament. *B3GAT3* expression was not specific to heart. The qualitative difference may result from the difference in the effects of the three genes on GAG formation.

To examine how *B3GALT6* mutations affects the products of GAGs in vivo, we measured the amounts of CS and HS chains at the surface of lymphoblastoid cells from the subjects by flow cytometry by using CS-stub and HS-stub antibodies as previously described.[13–15] In brief, purified GAG fractions were treated individually with a mixture of chondroitinases ABC and AC-II, a mixture of chondroitinases AC-I (EC 4.2.2.5) (Seikagaku Corp.) and AC-II (EC 4.2.2.5) (Seikagaku Corp.), chondroitinase B (EC 4.2.2.19) (IBEX Technologies), or a mixture of heparinases-I and -III (IBEX Technologies) for analyzing the disaccharide composition of CS/DS, CS, DS, and HS, respectively. The digests were labeled with a fluorophore 2-aminobenzamide (2AB) and aliquots of the 2AB derivatives of CS/DS/HS disaccharides were analyzed by anion-exchange HPLC on a PA-03 column (YMC Co.). The HS-stub antibody (3G10) showed a markedly reduced binding to the epitopes on the subjects' cells (Figure S7). The relative numbers of the HS chains presented as the mean fluorescence intensity (MFI) of the cell population stained with the antibody for P1, P2, and P3 were 26%, 56%, and 35% of the control, respectively. On the other hand, the CS-stub antibody (2B6) showed a similar binding to the epitopes on the subjects' cells relative to those of the control (Figure S7). The MFI for P1, P2, and P3 were 114%, 104%, and 106% of the control, respectively. Furthermore, we measured disaccharide of GAG chains from lymphoblastoid cells by using anion-exchange HPLC after digestion with chondroitinase and heparinase. The amounts of the disaccharide from HS chains were significantly decreased, whereas CS and DS chains were ~5 times higher than those in the control (Table 3).

Previous biochemical studies on EDS-PF with *B4GALT7* mutations show a reduction in the synthesis of DS chains.[16,17] The c.830G>A (p.Arg227Gln) mutation in *B3GAT3* causes a drastic reduction in GlcAT-I activity in fibroblasts of the individual with SEMD-JL1 and numbers of CS and HS chains on the core proteins at the surface of the fibroblasts are decreased to about half of the controls.[11] Cultured lymphoblastoid cells from individuals with a c.419C>T (p.Pro140Leu) mutation in *B3GAT3* show that defective synthesis is more pronounced for CS than for HS.[11] Taken together with our results, these findings suggest that the effects of the deficiencies of the three enzymes on GAG synthesis are not identical. A possible explanation for the qualitative phenotypic differences may be that the biosynthesis of the GAG linker region is not a simple step-by-step addition but involves parallel processing and/or alternative pathways. Other glycosyltransferases may have similar biochemical functions to these three enzymes and thus complement their deficient activities to variable degrees in cell- and/or tissue-specific manners, leading to differences in the amount of GAGs in the tissues. It is known that B3GALT6 and B4GALT7 have several homologs.[18] It must be noted that all biochemical studies so far have been performed in vitro or in cultured cells, and therefore there is a severe limitation to our understanding of the pathogenesis at tissue and organ levels.

By exome sequencing, we identified loss-of-function mutations in *B3GALT6* in 12 individuals from 10 families. The mutations produced a spectrum of connective tissue disorders characterized by lax skin, muscle hypotonia, joint dislocation, and skeletal dysplasia and deformity, which include phenotypes previously known as SEMD-JL1 and EDS-PF (Figures S1 and S2). The pleiotropic phenotypes of *B3GALT6* mutations indicate that B3GALT6 plays critical roles in development and homeostasis of various tissues, including skin, bone, cartilage, tendons, and ligaments. Biochemical studies that used lymphoblastoid cells of the individuals with *B3GALT6* mutations showed a decrease of HS and a paradoxical increase of CS and DS of the cell surface. Further clinical, genetic, and biological studies are necessary to understand the pathological mechanism of the diseases caused by enzyme defects involved in the biosynthesis of the GAG linker region and roles of the region in GAG metabolism and function.

Supplemental Data

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

Acknowledgments

We thank the individuals with the disease and their family for their help to the study. We also thank the Japanese Skeletal Dysplasia Consortium. This study is supported by research grants from the Ministry of Health, Labor, and Welfare (23300101 to S.I. and N. Matsumoto; 23300201 to S.I.), by Grants-in-Aid for Young Scientists (23689052 to N. Miyake and 23790066 to S.M.) from the Japan Society for the Promotion of Science; by the Matching Program for Innovations in Future Drug Discovery and Medical Care

(K.S.); by The Ministry of Education, Culture, Sports, Science and Technology, Japan (MEXT); by a Grant-in-aid for Encouragement from the Akiyama Life Science Foundation (S.M.); by Swiss National Science Foundation Grants (31003A_141241 and 310030_132940); by The CoSMO-B project (Brazil and Switzerland); by the Leenaards Foundation (Switzerland); and by Research on intractable diseases, Health and Labour Sciences Research Grants, H23-Nanchi-Ippan-123 (S.I.).

Received: February 1, 2013
Revised: March 16, 2013
Accepted: April 5, 2013
Published: May 9, 2013

Web Resources

The URLs for data presented herein are as follows:

1000 Genomes, http://browser.1000genomes.org
ANNOVAR, http://www.openbioinformatics.org/annovar/
dbSNP, http://www.ncbi.nlm.nih.gov/projects/SNP/
GATK, http://www.broadinstitute.org/gatk/
MutationTaster, http://www.mutationtaster.org/
NHLBI Exome Sequencing Project (ESP) Exome Variant Server, http://evs.gs.washington.edu/EVS/
Novoalign, http://www.novocraft.com/main/page.php?s=novoalign
Online Mendelian Inheritance in Man (OMIM), http://www.omim.org/
Picard, http://picard.sourceforge.net/
PolyPhen, http://www.genetics.bwh.harvard.edu/pph2/
RefSeq, http://www.ncbi.nlm.nih.gov/RefSeq
SIFT, http://sift.bii.a-star.edu.sg/
UCSC Genome Browser, http://genome.ucsc.edu

References

1. Warman, M.L., Cormier-Daire, V., Hall, C., Krakow, D., Lachman, R., LeMerrer, M., Mortier, G., Mundlos, S., Nishimura, G., Rimoin, D.L., et al. (2011). Nosology and classification of genetic skeletal disorders: 2010 revision. Am. J. Med. Genet. A. *155A*, 943–968.
2. Beighton, P., Gericke, G., Kozlowski, K., and Grobler, L. (1984). The manifestations and natural history of spondylo-epimetaphyseal dysplasia with joint laxity. Clin. Genet. *26*, 308–317.
3. Nishimura, G., Satoh, M., Aihara, T., Aida, N., Yamamoto, T., and Ozono, K. (1998). A distinct subtype of "metatropic dysplasia variant" characterised by advanced carpal skeletal age and subluxation of the radial heads. Pediatr. Radiol. *28*, 120–125.
4. Boyden, E.D., Campos-Xavier, A.B., Kalamajski, S., Cameron, T.L., Suarez, P., Tanackovic, G., Andria, G., Ballhausen, D., Briggs, M.D., Hartley, C., et al. (2011). Recurrent dominant mutations affecting two adjacent residues in the motor domain of the monomeric kinesin KIF22 result in skeletal dysplasia and joint laxity. Am. J. Hum. Genet. *89*, 767–772.
5. Min, B.J., Kim, N., Chung, T., Kim, O.H., Nishimura, G., Chung, C.Y., Song, H.R., Kim, H.W., Lee, H.R., Kim, J., et al. (2011). Whole-exome sequencing identifies mutations of KIF22 in spondyloepimetaphyseal dysplasia with joint laxity, leptodactylic type. Am. J. Hum. Genet. *89*, 760–766.
6. Miyake, N., Elcioglu, N.H., Iida, A., Isguven, P., Dai, J., Murakami, N., Takamura, K., Cho, T.J., Kim, O.H., Hasegawa, T., et al. (2012). PAPSS2 mutations cause autosomal recessive brachyolmia. J. Med. Genet. *49*, 533–538.
7. Tsurusaki, Y., Okamoto, N., Ohashi, H., Kosho, T., Imai, Y., Hibi-Ko, Y., Kaname, T., Naritomi, K., Kawame, H., Wakui, K., et al. (2012). Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. Nat. Genet. *44*, 376–378.
8. Bai, X., Zhou, D., Brown, J.R., Crawford, B.E., Hennet, T., and Esko, J.D. (2001). Biosynthesis of the linkage region of glycosaminoglycans: cloning and activity of galactosyltransferase II, the sixth member of the beta 1,3-galactosyltransferase family (beta 3GalT6). J. Biol. Chem. *276*, 48189–48195.
9. Saunders, C.J., Minassian, B.E., Chow, E.W., Zhao, W., and Vincent, J.B. (2009). Novel exon 1 mutations in MECP2 implicate isoform MeCP2_e1 in classical Rett syndrome. Am. J. Med. Genet. A. *149A*, 1019–1023.
10. Kresse, H., Rosthøj, S., Quentin, E., Hollmann, J., Glössl, J., Okada, S., and Tønnesen, T. (1987). Glycosaminoglycan-free small proteoglycan core protein is secreted by fibroblasts from a patient with a syndrome resembling progeroid. Am. J. Hum. Genet. *41*, 436–453.
11. Baasanjav, S., Al-Gazali, L., Hashiguchi, T., Mizumoto, S., Fischer, B., Horn, D., Seelow, D., Ali, B.R., Aziz, S.A., Langer, R., et al. (2011). Faulty initiation of proteoglycan synthesis causes cardiac and joint defects. Am. J. Hum. Genet. *89*, 15–27.
12. Faiyaz-Ul-Haque, M., Zaidi, S.H., Al-Ali, M., Al-Mureikhi, M.S., Kennedy, S., Al-Thani, G., Tsui, L.C., and Teebi, A.S. (2004). A novel missense mutation in the galactosyltransferase-I (B4GALT7) gene in a family exhibiting facioskeletal anomalies and Ehlers-Danlos syndrome resembling the progeroid type. Am. J. Med. Genet. A. *128A*, 39–45.
13. Kinoshita, A., and Sugahara, K. (1999). Microanalysis of glycosaminoglycan-derived oligosaccharides labeled with a fluorophore 2-aminobenzamide by high-performance liquid chromatography: application to disaccharide composition analysis and exosequencing of oligosaccharides. Anal. Biochem. *269*, 367–378.
14. Miyake, N., Kosho, T., Mizumoto, S., Furuichi, T., Hatamochi, A., Nagashima, Y., Arai, E., Takahashi, K., Kawamura, R., Wakui, K., et al. (2010). Loss-of-function mutations of CHST14 in a new type of Ehlers-Danlos syndrome. Hum. Mutat. *31*, 966–974.
15. Mizumoto, S., and Sugahara, K. (2012). Glycosaminoglycan chain analysis and characterization (Glycosylation/Epimerization). In Methods in Molecular Biology. In Proteoglycans: Methods and Protocols, F. Rédini, ed. (New York, USA: Humana Press, Springer), pp. 99–115.
16. Okajima, T., Fukumoto, S., Furukawa, K., and Urano, T. (1999). Molecular basis for the progeroid variant of Ehlers-Danlos syndrome. Identification and characterization of two mutations in galactosyltransferase I gene. J. Biol. Chem. *274*, 28841–28844.
17. Quentin, E., Gladen, A., Rodén, L., and Kresse, H. (1990). A genetic defect in the biosynthesis of dermatan sulfate proteoglycan: galactosyltransferase I deficiency in fibroblasts from a patient with a progeroid syndrome. Proc. Natl. Acad. Sci. USA *87*, 1342–1346.
18. Togayachi, A., Sato, T., and Narimatsu, H. (2006). Comprehensive enzymatic characterization of glycosyltransferases with a beta3GT or beta4GT motif. Methods Enzymol. *416*, 91–102.

Journal of Human Genetics (2013) 58, 391–394

www.nature.com/jhg

SHORT COMMUNICATION

Exome sequencing identifies a novel *INPPL1* mutation in opsismodysplasia

Aritoshi Iida[1,9], Nobuhiko Okamoto[2,9], Noriko Miyake[3,9], Gen Nishimura[4], Satoshi Minami[5], Takuya Sugimoto[6], Mitsuko Nakashima[3], Yoshinori Tsurusaki[3], Hirotomo Saitsu[3], Masaaki Shiina[7], Kazuhiro Ogata[7], Shigehiko Watanabe[8], Hirofumi Ohashi[8], Naomichi Matsumoto[3] and Shiro Ikegawa[1]

Opsismodysplasia is an autosomal recessive skeletal disorder characterized by facial dysmorphism, micromelia, platyspondyly and retarded bone maturation. Recently, mutations in the gene encoding inositol polyphosphate phosphatase-like 1 (INPPL1) are found in several families with opsismodysplasia by a homozygosity mapping, followed by whole genome sequencing. We performed an exome sequencing in two unrelated Japanese families with opsismodysplasia and identified a novel *INPPL1* mutation, c.1960_1962delGAG, in one family. The mutation is predicted to result in an in-frame deletion (p.E654del) within the central catalytic 5-phosphate domain. Our results further support that *INPPL1* is the disease gene for opsismodysplasia and that opsismodysplasia has genetic heterogeneity.
Journal of Human Genetics (2013) **58**, 391–394; doi:10.1038/jhg.2013.25; published online 4 April 2013

Keywords: exome sequencing; INPPL1; opsismodysplasia

INTRODUCTION

Opsismodysplasia (OMIM 258480) is a rare skeletal dysplasia identifiable at birth. Its clinical features are rhizomelic micromelia and facial dysmorphism, including prominent brow, large fontanels, depressed nasal bridge and small anteverted nose with long philtrum, as well as short feet and hands with sausage-like fingers.[1] Its main radiological features include retarded bone maturation, marked shortness of the bones of hands and feet with concave metaphyses and thin, lamellar vertebral bodies. Some patients show severe phosphate wasting. Autosomal recessive inheritance is the most likely mode of inheritance; to date, at least three consanguineous families with opsismodysplasia are reported.[2–4]

Recently, Below *et al.*[5] performed a homozygosity mapping coupled with whole genome sequencing in a consanguineous family with opsismodysplasia, and identified *INPPL1* (inositol polyphosphate phosphatase-like 1) as a causative gene for opsismodysplasia. They first identified a homozygous missense mutation, p.Pro659Leu, in the consanguineous family, and then found *INPPL1* mutations in additional five unrelated families with opsismodysplasia. We performed a whole exome sequencing for two patients from two unrelated families and identified a homozygous in-frame deletion of *INPPL1* in one family.

SUBJECTS AND METHODS

Subjects and DNA samples

Two families with clinical diagnosis of opsismodysplaisa were included in the study. Family 1 consisted of parents and affected sibs (Figure 1a), and Family 2 consisted of parents and a patient. Genomic DNA was extracted by standard procedures from peripheral blood of the patients and their family members after informed consent. The study was approved by the ethical committee of RIKEN, Yokohama City University, and participating institutions.

Exome sequencing

Six individuals in the two families were analyzed by the whole exome sequence as described previously.[6] Briefly, 3 μg of genomic DNA was sheared by Covaris 2S system (Covaris, Woburn, MA, USA) and partitioned using SureSelect Human All Exon V4 (Agilent technology, Santa Clara, CA, USA) according to the manufacturer's instructions. The exon-enriched DNA libraries were sequenced using HiSeq2000 (Illumina, San Diego, CA, USA) with a 101-bp paired-end reads and a 7-bp index reads. Four samples (2.5 pM each, with different index) were run in one lane. HiSeq Control Software/Real-Time Analysis and CASAVA1.8.2 (Illumina) were used for image analysis and base calling. The mapping was performed to human genome hg19 using Novoalign (http://www.novocraft.com/main/page.php?s=novoalign). The aligned reads were processed by Picard to remove the polymerase chain reaction (PCR) duplicate (http://picard.sourceforge.net). The variants were called

[1]Laboratory for Bone and Joint Diseases, Center for Genomic Medicine, RIKEN, Tokyo, Japan; [2]Department of Medical Genetics, Osaka Medical Center and Research Institute for Maternal and Child Health, Osaka, Japan; [3]Department of Human Genetics, Yokohama City University Graduate School of Medicine, Yokohama, Japan; [4]Department of Pediatric Imaging, Tokyo Metropolitan Children's Medical Center, Fuchu, Japan; [5]Department of Gynecology and Obstetrics, Shingu Municipal Medical Center, Shingu, Japan; [6]Department of Pediatrics Shingu Municipal Medical Center, Shingu, Japan; [7]Department of Biochemistry, Yokohama City University Graduate School of Medicine, Yokohama, Japan and [8]Division of Medical Genetics, Saitama Children's Medical Center, Saitama, Japan
[9]These authors contributed equally to this work.
Correspondence: Professor S Ikegawa, Laboratory for Bone and Joint Diseases, Center for Genomic Medicine, RIKEN, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan.
E-mail: sikegawa@ims.u-tokyo.ac.jp
Received 5 January 2013; revised 6 March 2013; accepted 9 March 2013; published online 4 April 2013

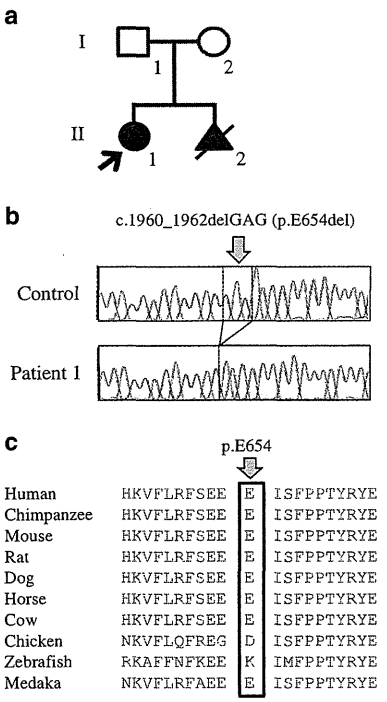

Figure 1 *INPPL1* mutation in a Japanese family with opsismodysplasia. (**a**) Pedigree, (**b**) an in-frame deletion c.1960_1962delGAG (p.E654del) within exon 17 and (**c**) conservation of p.E654 in INPPL1 among different species.

by Genome Analysis Toolkit 1.6-5 (GATK; http://www.broadinstitute.org/gsa/wiki/index.php/Main_Page) with the best practice variant detection with the GAKT v.3 (http://www.broadinstitute.org/gsa/wiki/index.php/Best_Practice_Variant_Detection_with_the_GATK_v3) and annotated by ANNOVAR (23 February 2012) (http://www.openbioinformatics.org/annovar/). Through this flow, common variants registered in dbSNP135 (minor allele frequeny $\geqslant$0.01) (http://genome.ucsc.edu/cgi-bin/hgTrackUi?hgsid=316787363&g=snp135Common&hgTracksConfigPage=configure) were removed.

Priority scheme

On the basis of the hypothesis that opsismodysplasia is inherited in an autosomal recessive manner, variants were filtered by following conditions using the script created by BITS (Tokyo, Japan). For the homozygous mutation model: (1) variant allele frequency (variant alleles/total alleles) in probands $\geqslant$0.8, (2) variant allele frequency in parents $\leqslant$0.8, (3) excluding synonymous changes and (4) excluding the variants observed in our in-house database (n=429). For the compound heterozygous mutation model: (1) mutation allele frequency in probands: 0.2–0.8, (2) variant allele frequency in parents $\leqslant$0.8, (3) excluding synonymous changes, (4) excluding the variants observed in our in-house database (n=429) and (5) selecting genes with compound heterozygous change. After combining variants selected by both models, genes commonly found in the two families were searched.

Sanger sequencing

We performed Sanger sequencing to confirm the deletion identified in the proband of Family 1 by the exome sequencing. We amplified exon 17 by PCR using primer sequences, 5′-AAGCACAAGGTCTTCCTTCGATTCA-3′ and 5′-CCATACCCTTGACCCAAATTCTTGAT-3′. We directly sequenced the PCR product using an Applied Biosystems 3730xl DNA analyzer (Life Technologies, Forster City, CA, USA). For the patient in Family 2, we screened 28 exons of *INPPL1* and exon–intron boundaries by direct sequencing of PCR products from genome DNA. The primer sequences are available on request.

Evaluation of polymorphism

We used the invader assay coupled with PCR[7] to exclude the possibility of polymorphism in 188 Japanese general populations. The deletion was evaluated by databases, PROVEAN v.1.1 (http://provean.jcvi.org/genome_submit.php), dbSNP (http://www.ncbi.nlm.nih.gov/projects/SNP/) and 1000 genomes (http://www.1000genomes.org/). We used Evola website to investigate the conservation of p.E654 of INPPL1 (http://www.h-invitational.jp/hinv/ahg-db/index.jsp).

RESULTS

Exome sequencing

By the whole exome sequencing, 3.8–5.1 Gb sequences uniquely mapped to all human RefSeq coding region were obtained. For all subjects, at least 95.9% of all coding regions were covered in five reads depth and more (Supplementary Table 1). No candidate genes that had mutations in the two families were identified.

Because *INPPL1* mutations have recently been identified in opsismodysplasia,[5] we checked *INPPL1* mutations in the exome sequence data. Five or more reads covered 100% of its coding regions (Supplementary Table 1). A homozygous deletion, c.1960_1962 (p.E654del), was found in the proband of Family 1 (Figure 1a). However, this deletion had been excluded as a candidate mutation because no *INPPL1* variant likely to be a mutation was detected in Family 2.

Confirmation and evaluation of c.1960_1962delGAG

We confirmed the deletion by direct sequence of PCR product from genomic DNA in the proband of Family 1 (Figure 1b). Next, we performed the invader assay coupled with PCR in the family. The parents were compound heterozygous for the deletion and the affected sibs were homozygous for it. The deletion was not found in 188 Japanese controls and in the public databases. The E654 is conserved between different species (Figure 1c). It is within the central catalytic 5-phosphate domain, but located at the position far from active site (25 amino acids) and within a loop region, which is thought to have structural flexibility in general. Inositol polyphosphate 5-phosphatase domain (ipp5c) of yeast synaptojanin in complex with inositol (1,4)-bisphosphate and calcium ion (PDB ID 1i9z) is the most analogous structure to the human INPPL1 catalytic domain among the currently available structures; however, its sequence identity with the human INPPL1 catalytic domain is low (26%). These make the structural assessment of the mutation equivocal. The PROVEAN database showed that p.E654del had a deleterious function against the gene product (score: −12.1).

Mutation screening of *INPPL1* in Family 2

We screened the *INPPL1* mutation in the patient of Family 2 by direct sequencing of the entire coding exons and their flanking regions. A total of nine SNPs were found, but no mutation was found in the patient.

Clinical information of the patients with the *INPPL1* mutation

The proband of Family 1 (II-1 in Figure 1a) was a 9-year-old girl born to non-consanguineous healthy parents. Family history was unremarkable. She was referred to one of us because fetal echogram revealed short extremities. She was born at 40 weeks' of gestation. Her birth weight was 2119 g (<3 percentile), length 38.0 cm (<3 percentile) and head circumference 35.1 cm (<3 percentile). She had a wide fontanelle, widely patent sutures, frontal bossing, flat nasal

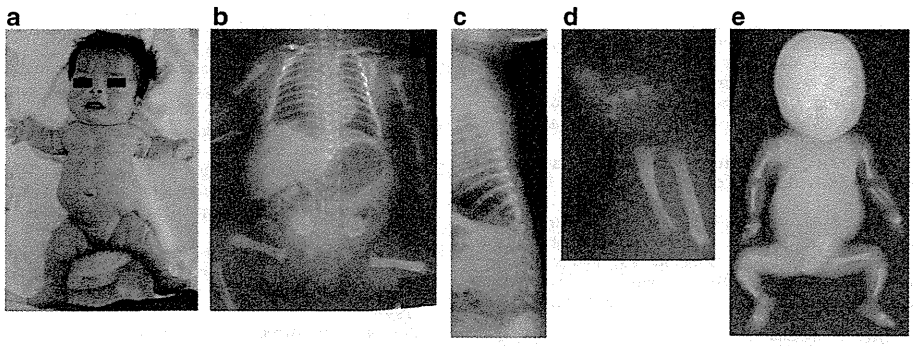

Figure 2 Phenotype of patients in Family 1. (**a**) Appearance of the proband in Family 1. Rhizomelic micromelia, frontal bossing, flat nasal bridge, low set ears, anteverted nostrils, micrognathia, narrow thorax and distended abdomen were noted. Radiographs of the proband (II-1) at birth (**b–d**) and the aborted fetus (II-2) (**e**). Characteristics of opsismodysplasia including retarded bone maturation, shortness of the bones of hands and feet, concave metaphyses and thin, lamellar vertebral bodies were noted.

bridge, low set ears, anteverted nostrils, micrognathia, narrow thorax and distended abdomen, and her extremities were remarkably short (Figure 2a). Her respiratory activity was weak and inspiratory wheezing was noted. Tracheal intubation became necessary 4 h after birth. Radiological investigations of her skeleton showed characteristics of opsisimodysplasia (Figures 2b–d). She was repeatedly admitted because of respiratory insufficiency due to infections. At 2 years of age, tracheotomy was performed to care for respiratory problems. She was noticed to show low serum phosphate levels at around 1 year and since then had been treated on phosphate supplements and/or alfacalcidol (1α-OH-D_3). At age 9 years, her height was 65 cm (<−6 s.d.) and weight 9 kg (−4 s.d). Her intellectual development was normal and was attending an elementary school.

In the second pregnancy, similar conditions were found by a fetal echogram. Artificial abortion was carried out. The post-mortem radiograph showed skeletal findings similar to the proband (Figure 2e).

DISCUSSION

Below *et al.*[5] examined *INPPL1* in a total of 12 unrelated families with opsismodysplasia and found its mutations in seven families. The list of mutation includes missense, nonsense and splicing mutations; all are predicted to be loss of function mutations. In one family, we also found a deletion mutation in *INPPL1* that is predicted to be a loss of function mutation, but in another family, we could not detect an *INPPL1* mutation. These results further support the results of the previous study that *INPPL1* is the disease gene for opsismodysplasia and that opsismodysplasia has genetic heterogeneity.[5] In retrospect, the patient of Family 2 showed significant platyspondyly, yet some of the radiographic features for opsismodysplasia that include hypoplasia of the base of the skull on lateral views and lateral spikes of the acetabular roof were absent. Further, the fragmented epiphyses and coning of the distal femora are not characteristically seen in opsismodysplasia. This case is also different from the other cases with an opsismodysplasia phenotype that do not have *INPPL1* mutations (Prof. Debora Krakow, personal communication). Further collection of *INPPL1* mutation-proven cases would help in defining the phenotype of opsismodysplasia. While we were preparing the manuscript, another study reporting the identification of *INPPL1* as the cause of opsismodysplasia was published.[8] It reports identification of the *INPPL1* mutation in all 10 families examined.

INPPL1 (also known as SHIP2) is a member of the inositol 5′-phosphatase family that hydrolyzes phosphatidylinositol 3,4,5-triphosphate (PtdIns(3,4,5)P_3) and generates phosphatidylinositol 3,4-bisphosphate (PtdIns(3,4)P_2).[9] *INPPL1* encodes a 142-kDa protein with a variety of protein interaction domains, including an N-terminal SH2 domain, a central catalytic 5-phosphatase domain, a C-terminal proline-rich domain, an NPXY site and a sterile a motif domain in the C-terminal region.[10] At least 12 proteins of binding partners for INPLL1, such as Shc, APS, filamin and EphA2, have been identified.[10] The genes for these binding partners are good candidates for the disease gene for the opsismodysplasia-like phenotype.

Biological roles of INPPL1 remain unclear. *INPPL1* expression is particularly high in heart, skeletal muscle and placenta.[11,12] Its proposed roles are cell adhesion and spreading, actin cytoskeletal remodeling and receptor internalization. INPPL1 negatively regulates insulin signaling through its catalytic PtdIns(3,4,5) P_3 5-phosphatase activity.[9] The *INPPL1*$^{-/-}$ mice show a shortened snout and grow more slowly than wild-type littermates.[13] After 6 weeks of age, they showed a substantial reduced body length and body weight; however, radiographic analysis showed no gross skeletal deficit. Further studies are necessary to clarify the role of *INPPL1* in skeletal development and homeostasis.

ACKNOWLEDGEMENTS

We thank the patients and their family for their help to the study. We also thank the Japanese Skeletal Dysplasia Consortium. This study is supported by research grants from the Ministry of Health, Labor and Welfare (23300101 to SI and NMat.; 23300201 to SI), by Grants-in-Aid for Young Scientists (23689052 to NMiy.) from the Japan Society for the Promotion of Science; by Research on intractable diseases, Health and Labour Sciences Research Grants, H23-Nanchi-Ippan-123 (SI) and by grants from the Japan Science and Technology Agency, the Strategic Research Program for Brain Sciences (11105137 to NMat.), a Grant-in-Aid for Scientific Research on Innovative Areas (Transcription Cycle) from the Ministry of Education, Culture, Sports, Science and Technology of Japan (12024421 to NMat.), a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (12020465 to NMat.) and the Takeda Science Foundation (to N Miy. and N Mat.). We thank Professors Debora Krakow and Michael Bamshad for their comments on the patients' phenotypes. We also thank Ms Tomoko Kusadokoro for technical assistance.

1 Cormier-Daire, V., Delezoide, A. L., Philip, N., Marcorelles, P., Casas, K., Hillion, Y. *et al.* Clinical, radiological, and chondro-osseous findings in opsismodysplasia: survey of a series of 12 unreported cases. *J. Med. Genet.* **40,** 195–200 (2003).
2 Beemer, F. A. & Kozlowski, K. S Additional case of opsismodysplasia supporting autosomal recessive inheritance. *Am. J. Med. Genet.* **49,** 344–347 (1994).
3 Santos, H. G. & Saraiva, J. M Opsismodysplasia: another case and literature review. *Clin. Dysmorphol.* **4,** 222–226 (1995).
4 Tyler, K., Sarioglu, N. & Kunze, J Five familial cases of opsismodysplasia substantiate the hypothesis of autosomal recessive inheritance. *Am. J. Med. Genet.* **83,** 47–52 (1999).
5 Below, J. E., Earl, D. L., Shively, K. M., McMillin, M. J., Smith, J. D., Turner, E. H. *et al.* Whole-genome analysis reveals that mutations in inositol polyphosphate phosphatase-like 1 cause opsismodysplasia. *Am. J. Hum. Genet.* **92,** 137–143 (2013).
6 Miyake, N., Elcioglu, N. H., Iida, A., Isguven, P., Dai, J., Murakami, N. *et al.* PAPSS2 mutations cause autosomal recessive brachyolmia. *J. Med. Genet.* **49,** 533–538 (2012).
7 Ohnishi, Y., Tanaka, T., Ozaki, K., Yamada, R., Suzuki, H. & Nakamura, Y A high-throughput SNP typing system for genome-wide association studies. *J. Hum. Genet.* **46,** 471–477 (2001).
8 Huber, C., Faqeih, E. A., Bartholdi, D., Bole-Feysot, C., Borochowitz, Z., Cavalcanti, D. P. *et al.* Exome sequencing identifies INPPL1 mutations as a cause of opsismodysplasia. *Am. J. Hum. Genet.* **92,** 144–149 (2013).
9 Dyson, J. M., Kong, A. M., Wiradjaja, F., Astle, M. V., Gurung, R. & Mitchell, C. A The SH2 domain containing inositol polyphosphate 5-phosphatase-2: SHIP2. *Int. J. Biochem. Cell. Biol.* **37,** 2260–2265 (2005).
10 Suwa, A., Kurama, T. & Shimokawa, T. SHIP2 and its involvement in various diseases. *Expert Opin. Ther. Targets* **14,** 727–737 (2010).
11 Hejna, J. A., Saito, H., Merkens, L. S., Tittle, T. V., Jakobs, P. M., Whitney, M. A. *et al.* Cloning and characterization of a human cDNA (INPPL1) sharing homology with inositol polyphosphate phosphatases. *Genomics* **29,** 285–287 (1995).
12 Pesesse, X., Deleu, S., De Smedt, F., Drayer, L. & Erneux, C Identification of a second SH2-domain-containing protein closely related to the phosphatidylinositol polyphosphate 5-phosphatase SHIP. *Biochem. Biophys. Res. Commun.* **239,** 697–700 (1997).
13 Sleeman, M. W., Wortley, K. E., Lai, K. M., Gowen, L. C., Kintner, J., Kline, W. O. *et al.* Absence of the lipid phosphatase SHIP2 confers resistance to dietary obesity. *Nat. Med.* **11,** 199–205 (2005).

Supplementary Information accompanies the paper on Journal of Human Genetics website (http://www.nature.com/jhg)

Whole genome sequencing in patients with retinitis pigmentosa reveals pathogenic DNA structural changes and *NEK2* as a new disease gene

Koji M. Nishiguchi[a,b], Richard G. Tearle[c], Yangfan P. Liu[d], Edwin C. Oh[d,e], Noriko Miyake[f], Paola Benaglio[a], Shyana Harper[g], Hanna Koskiniemi-Kuendig[a], Giulia Venturini[a], Dror Sharon[h], Robert K. Koenekoop[i], Makoto Nakamura[b], Mineo Kondo[b], Shinji Ueno[b], Tetsuhiro R. Yasuma[b], Jacques S. Beckmann[a,j,k], Shiro Ikegawa[l], Naomichi Matsumoto[f], Hiroko Terasaki[b], Eliot L. Berson[g], Nicholas Katsanis[d], and Carlo Rivolta[a,1]

[a]Department of Medical Genetics, University of Lausanne, 1005 Lausanne, Switzerland; [b]Department of Ophthalmology, Nagoya University School of Medicine, Nagoya 466-8550, Japan; [c]Complete Genomics, Inc., Mountain View, CA 94043; [d]Center for Human Disease Modeling and [e]Department of Neurology, Duke University, Durham, NC 27710; [f]Department of Human Genetics, Yokohama City University Graduate School of Medicine, Yokohama 236-0004, Japan; [g]Berman-Gund Laboratory for the Study of Retinal Degenerations, Harvard Medical School, Massachusetts Eye and Ear Infirmary, Boston, MA 02114; [h]Department of Ophthalmology, Hadassah-Hebrew University Medical Center, Jerusalem 91120, Israel; [i]McGill Ocular Genetics Laboratory, McGill University Health Centre, Montreal, QC, Canada H3H 1P3; [j]Service of Medical Genetics, Lausanne University Hospital, 1011 Lausanne, Switzerland; [k]Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland; and [l]Laboratory for Bone and Joint Diseases, Center for Genomic Medicine, RIKEN, Tokyo 108-8639, Japan

Edited by Jeremy Nathans, Johns Hopkins University, Baltimore, MD, and approved August 15, 2013 (received for review May 1, 2013)

We performed whole genome sequencing in 16 unrelated patients with autosomal recessive retinitis pigmentosa (ARRP), a disease characterized by progressive retinal degeneration and caused by mutations in over 50 genes, in search of pathogenic DNA variants. Eight patients were from North America, whereas eight were Japanese, a population for which ARRP seems to have different genetic drivers. Using a specific workflow, we assessed both the coding and noncoding regions of the human genome, including the evaluation of highly polymorphic SNPs, structural and copy number variations, as well as 69 control genomes sequenced by the same procedures. We detected homozygous or compound heterozygous mutations in 7 genes associated with ARRP (*USH2A*, *RDH12*, *CNGB1*, *EYS*, *PDE6B*, *DFNB31*, and *CERKL*) in eight patients, three Japanese and five Americans. Fourteen of the 16 mutant alleles identified were previously unknown. Among these, there was a 2.3-kb deletion in *USH2A* and an inverted duplication of ~446 kb in *EYS*, which would have likely escaped conventional screening techniques or exome sequencing. Moreover, in another Japanese patient, we identified a homozygous frameshift (p.L206fs), absent in more than 2,500 chromosomes from ethnically matched controls, in the ciliary gene *NEK2*, encoding a serine/threonine-protein kinase. Inactivation of this gene in zebrafish induced retinal photoreceptor defects that were rescued by human *NEK2* mRNA. In addition to identifying a previously undescribed ARRP gene, our study highlights the importance of rare structural DNA variations in Mendelian diseases and advocates the need for screening approaches that transcend the analysis of the coding sequences of the human genome.

medical genetics | ophthalmology | ciliopathy | retinal blindness

The identification of the genetic causes of rare Mendelian diseases is becoming increasingly important following some success with gene-based therapy, as recently reported for patients with a form of Leber congenital amaurosis (LCA), a severe autosomal recessive hereditary retinal dystrophy (1–3). The evidence that restoring a gene in the diseased retina could yield therapeutic effects has stimulated the pursuit of the genetic causes of other retinal dystrophies, including retinitis pigmentosa (RP).

RP is the name given to a group of hereditary retinal conditions in which degeneration of rod photoreceptors, responsible for vision under starlight or moonlight conditions, is more pronounced than that of cone photoreceptors, which mediate daylight vision. Individuals with RP typically experience night blindness at first, followed by progressive and unstoppable visual impairment in daytime conditions as well (4). Their visual fields become reduced gradually and sight is lost from the midperiphery to the periphery and then from the midperiphery to the center, resulting eventually in complete or near-complete blindness if left untreated. Most patients show intraretinal pigment in a bone spicule configuration around the fundus periphery, for which this condition was named. In addition, they typically show retinal arteriolar attenuation, elevated final dark adapted thresholds, and reduced and delayed electroretinograms (ERGs) (4). Vitamin A supplementation in combination with an omega-3 rich diet can slow the course of retinal degeneration and preserve visual acuity among adults with this condition (5, 6). Autosomal, recessively inherited RP (ARRP) is the most common form of hereditary retinal degeneration in humans. To date, over 50 genes have been associated with ARRP and allied disorders, among patients who are predominantly of European ancestry (RetNet; www.sph.uth.tmc.edu/retnet/home.htm). However, despite this high number of identified disease genes, ~40–50% of all diagnosed cases have no mutations in recognized loci (7). Furthermore, genetic defects in RP are also population specific. For example, a screening of 193 unrelated Japanese patients with isolate or autosomal recessive RP

Significance

Retinitis pigmentosa (RP) is a genetic disease that causes progressive blindness and that is caused by mutations in more than 50 genes. Conventional methods for identification of both RP mutations and novel RP genes involve the screening of DNA sequences spanning coding exons. In our work, we conversely test the use of whole genome sequencing, a technique that takes into account all variants from both the coding and noncoding regions of the human genome. In our approach, we identify a number of unique RP mutations, a previously undescribed disease gene, as well as pathogenic structural DNA rearrangements originating in introns.

Author contributions: K.M.N. and C.R. designed research; K.M.N., Y.P.L., E.C.O., N. Miyake, P.B., H.K.-K., and G.V. performed research; S.H., D.S., R.K.K., M.N., M.K., S.U., T.R.Y., S.I., N. Matsumoto, H.T., and E.L.B. contributed new reagents/analytic tools; K.M.N., R.G.T., Y.P.L., E.C.O., N. Miyake, P.B., H.K.-K., G.V., J.S.B., S.I., N. Matsumoto, N.K., and C.R. analyzed data; and K.M.N., E.C.O., J.S.B., E.L.B., N.K., and C.R. wrote the paper.

Conflict of interest statement: R.G.T. is an employee and shareholder of Complete Genomics, Inc.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: carlo.rivolta@unil.ch.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1308243110/-/DCSupplemental.

for 30 disease genes identified commonly within North American or European patients revealed candidate pathogenic mutations in only 14% of the cohort (8).

Recent advances in massively parallel sequencing have enabled the analysis of large amounts of sequences (genes) at reasonable costs, revolutionizing the traditional approach of exon-by-exon Sanger sequencing (9). The two major forms of sequencing strategies allowing large-scale analyses are whole genome sequencing (WGS) and whole exome sequencing (WES). The former reads the entire genome with no distinction between exons and non-exonic regions. It allows the detection of intergenic variants, copy number variations (CNVs), and other structural rearrangements, as well as unrecognized exonic sequences. The latter technique relies on targeted DNA capture and focuses on the analysis of the known exonic content of the genome, performed according to the genomic annotation available at a given point in time.

In this study, we performed WGS as a method for mutation discovery in a highly genetically heterogeneous Mendelian disease; to this end, we evaluated 16 unrelated RP patients from diverse ethnic backgrounds.

Results

Genome Sequencing. Genome sequencing in the 16 analyzed patients produced an average mapping yield of 200.8 ± 17.9 (mean ± SD) Gb and an average coverage of 66.1 ± 2.4 (mean ± SD) reads per base (*SI Appendix*, Table S1). This covered a genomic fraction of 0.968 ± 0.004, in which roughly 3.8 million putative variations were identified. Of these, ~7.7% were not reported in dbSNP build 131 and were classified as novel variants. Variations present within transcripts were classified further as synonymous and nonsynonymous, and analyzed separately for the North American and Japanese sets of patients. Scoring of large structural variations (SVs) could be achieved only for seven genomes, as the remaining DNA samples, possibly because of their older age, did not produce reliable mate pair information (*SI Appendix*, *Results S1*).

Assessment of pathogenic variants was performed by a series of filtering steps, summarized in Fig. 1.

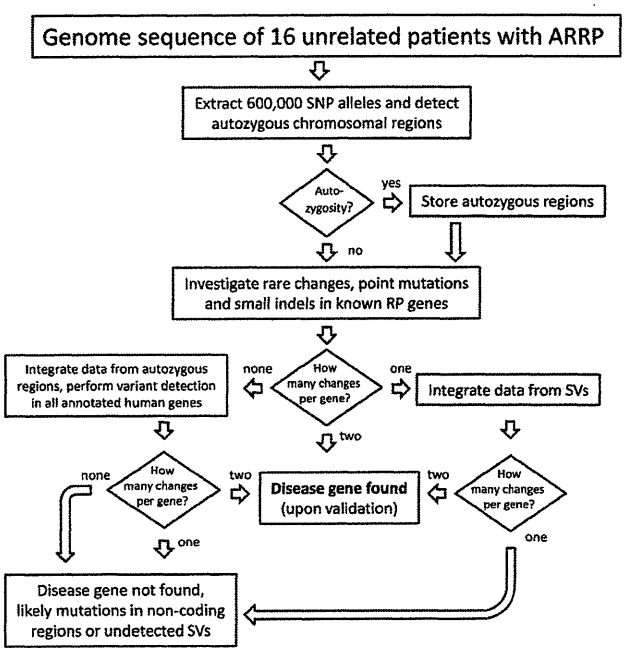

Fig. 1. Flowchart of the filtering process applied in this work.

Assessment of Autozygous Regions. Each genome was evaluated for known or undocumented parental consanguinity as well as for possible founder mutation events by extracting genotypes of known polymorphic SNPs and by searching for long intervals with high degrees of homozygosity (at least 500 consecutive SNP markers, or ~2.2 Mb on average), indicative of identity by descent (IBD). Significant genomic homozygosity was observed only in the five Japanese patients (individual IDs: R14, R15, R16, R18, and R19) who had documented parental consanguinity. The areas of IBD had essentially no overlap among these patients except for a 10-Mb interval on chromosome 1 shared by R15 and R19. Haplotype analysis indicated the shared intervals to be of different origins. No other patients carried genomic areas indicative of IBD; this was consistent with their family history reporting no parental consanguinity.

Sequence Analyses of Known RP Genes. We first focused our analyses on genes known to be associated with ARRP. We investigated both small variants (from 1 to 50 bp) from the mapping of short reads and, whenever possible, large SVs. Our results are summarized in *SI Appendix*, Table S2; detailed results are provided in *SI Appendix*, *Results S1*, Figs. S1 and S3, and Table S3.

In addition to point mutations and short indels (insertion/deletions), we detected pathogenic SVs in *USH2A* and *EYS* in patients 003–019 and R9, respectively, by combining information from sequence coverage and abnormal junctions/mate pair distance. In the genome of patient 003–019, we identified a ~2-kb deletion that removed exon 27 of *USH2A*, whereas patient R9 was found to carry a 446-kb head-to-head inverted duplication of the portion of chromosome 6 that included exons 23–29 of *EYS* (Fig. 2).

We found two pathogenic alleles, in either a homozygous or compound heterozygous state, in 8 of the 16 patients, 5 Americans and 3 Japanese, in seven different genes (*SI Appendix*, Table S2). Six patients carried mutations in one of the following genes: *USH2A*, *RDH12*, *CNGB1*, *EYS*, *PDE6B*, and *DFNB31*; 2 patients had mutations in *CERKL*. None of these mutations were found in the control cohorts of 95 healthy North American or 95 Japanese individuals. None of these mutations were reported previously, except p.R257X in *CERKL* and p.G76R in *RDH12* (10, 11). All mutations cosegregated with RP as recessive, pathogenic alleles in all family members of the index patients for whom DNA samples were available (Fig. 3).

Systematic Screening of All Genes. Based on the data from the analysis of known RP genes, we adopted a pipeline to perform a systematic analysis targeting all annotated genes in the genomes of patients with unsolved genetic etiology (*SI Appendix*, Fig. S2). With the aim of selecting a restricted number of candidate genes, more aggressive filtering was adopted with respect to the one used for the screening of known disease genes. The major differences in the analytical pipeline included removal of all entries in dbSNP. We safely applied this filtering because, given the low frequency of individual mutations in ARRP genes (including undetected ones), the risk of eliminating pathogenic DNA variants that could be fortuitously included in dbSNP build 131 is negligible. Further, to validate this approach, we applied it again retrospectively to the genomes for which mutations in RP genes were already detected. All of identified RP mutations were present in the final list of variants, supporting the sensitivity of the strategy. Detailed results are provided in *SI Appendix*, *Results S2* and are summarized in *SI Appendix*, Figs. S2 and S3 and Table S4.

In R19, in whom we did not find any clear-cut mutations in known ARRP genes, we found a homozygous frameshift variant (p.L206fs, c.617_624delTGTATGAGinsA) in the never in mitosis gene A (NIMA)-related kinase 2 (*NEK2*) gene. This variant was present within a highly homologous genomic stretch of 19.6 Mb of chromosome 1q32, predicted to be IBD (*SI Appendix*, Fig. S4).

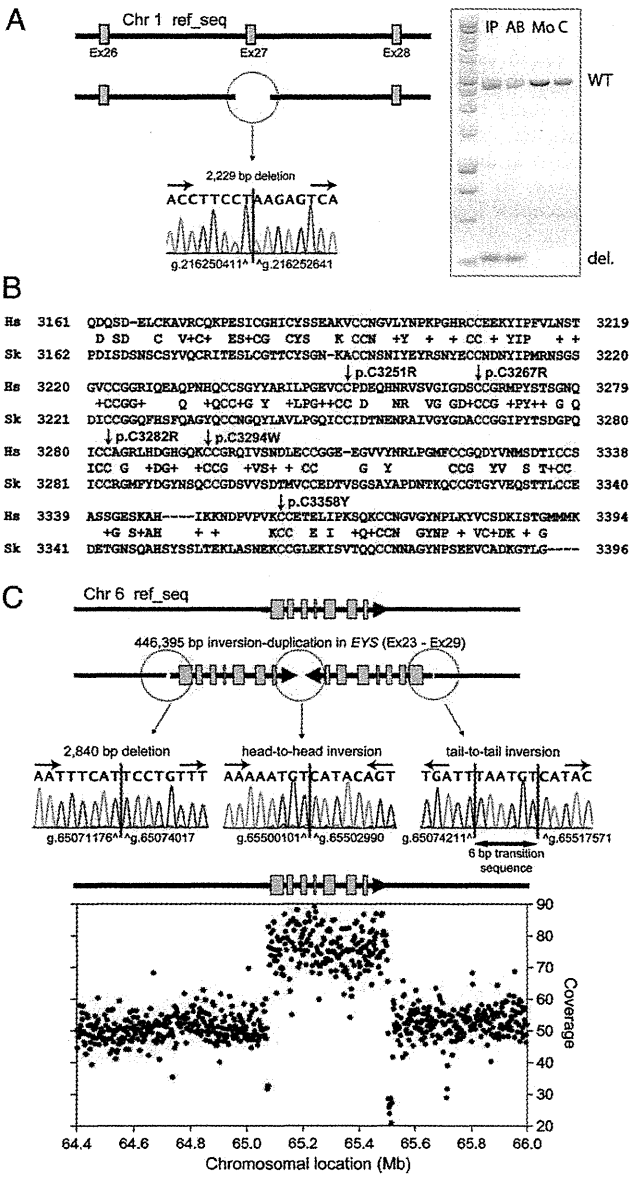

Fig. 2. Pathogenic structural variations identified. (*A*) Sequence of the heterozygous *USH2A* 2,229-bp deletion in patient 003–019 (*Left*) and electrophoresis of the PCR fragments showing a smaller fragment carrying the deletion in the index patient (IP) and her affected brother (AB) but not in her mother (Mo) or a control DNA (C). del, deleted; WT, wild type. (*B*) Alignment of the USH2A protein from *Homo sapiens* (Hs) and *Saccoglossus kowalevskii* (Sk, acorn worm) showing the conservation of 13 CC repeat motifs (red) and the location of the mutation p.C3294W, newly identified in patient 003–019 and her sister. Four previously reported disease-associated missense changes (p.C3251R, p.C3267R, p.C3282R, and p.C3358Y) also affect neighboring CC repeats. (*C*) Schematic representation and DNA sequence of the junctions characterizing the chromosomal rearrangement detected in patient R9 and involving the *EYS* gene. Integration of the information obtained by Sanger DNA sequencing and WGS coverage of the region allows identifying an inverted duplication encompassing exons 23–29.

Similar to most frameshifts producing a premature termination codon, p.L206fs is predicted to result in an mRNA allele that is subject to nonsense-mediated mRNA decay, and therefore in no protein product. Targeted DNA screening revealed that c.617_624delTGTATGAGinsA was absent from 1,273 Japanese and 95 North American control individuals. The entire coding sequence of the *NEK2* gene was then analyzed in a mixed cohort of 190 American patients with ARRP, in 64 Japanese patients with isolate RP, as well as in 13 patients found previously to show linkage between recessive retinal degeneration and the *NEK2* region. However, other than known polymorphisms (rs1056729, rs12031285, and rs45623136), we found only a few isolated heterozygous missense variants (p.R26Q, c.77G>A; p.V137I, c.409G>A; p.I265V, c.793A>G; p.N189S, c.566A>G; and p.K103E, c.307A>G; none were present in dbSNP) insufficient to account for ARRP. Notably, an additional Japanese male with ARRP was found to carry the same frameshift variant p.L206fs, but heterozygously, with no other variants in the *NEK2* coding sequence. This same patient (R51) was later found to carry the retinitis pigmentosa GTPase regulator (*RPGR*) mutation c.2405_2406delAG; p.E802fs (Human Gene Mutation Database entry: CD004115), described previously to be a sufficient cause of RP (12).

In light of a recent study reporting the involvement of noncoding RNA in the pathogenesis of retinal degeneration in mice (13), variants in noncoding RNA were also analyzed. After the removal of variants observed in 52 publicly available control genomes, only isolated heterozygous variants each with one entry per gene remained, insufficient to account for ARRP.

***nek2* Inactivation and Rescue in Zebrafish.** To validate the pathogenic role of *NEK2* deficiency in RP, we suppressed the sole ortholog of *NEK2* in zebrafish embryos and asked whether this manipulation might give rise to photoreceptor phenotypes. Upon injection of 6 ng of *nek2* splice-blocking morpholino, we observed gross ocular defects, including microphthalmia and enlarged eye sockets in 5-d postfertilization (dpf) morphant (MO) embryos (Fig. 4*A*). Whereas 63% of MO embryos displayed such phenotypes, only 21% of embryos expressing both MO and wild-type human *NEK2* mRNA did, suggesting that the ocular phenotypes are specific to the *nek2* suppression ($P < 0.001$) (Fig. 4*B*).

We next asked whether, in addition to overt structural abnormalities that may not directly inform the involvement of this gene to RP in humans, suppression of *nek2* might also give rise to photoreceptor defects consistent with those of patients with ARRP. We therefore embedded and paraffin sectioned control and MO embryos. In addition to the small eye phenotype, we detected alterations in the photoreceptor layer. Specifically, after

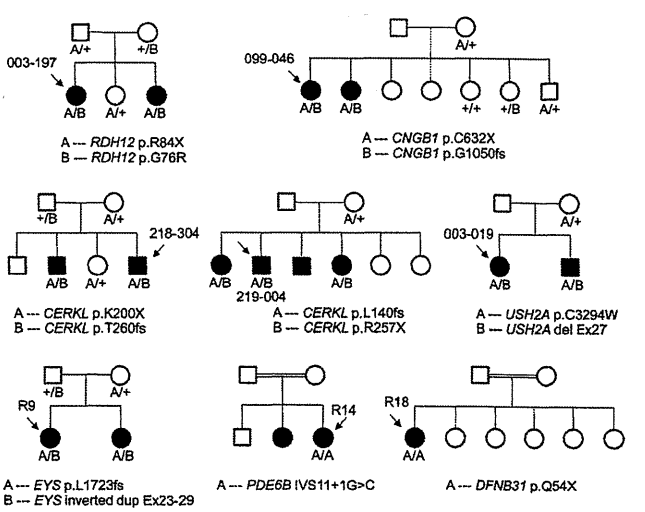

Fig. 3. Cosegregation analyses. All mutations analyzed cosegregated with the disease according to an autosomal recessive pattern of inheritance.

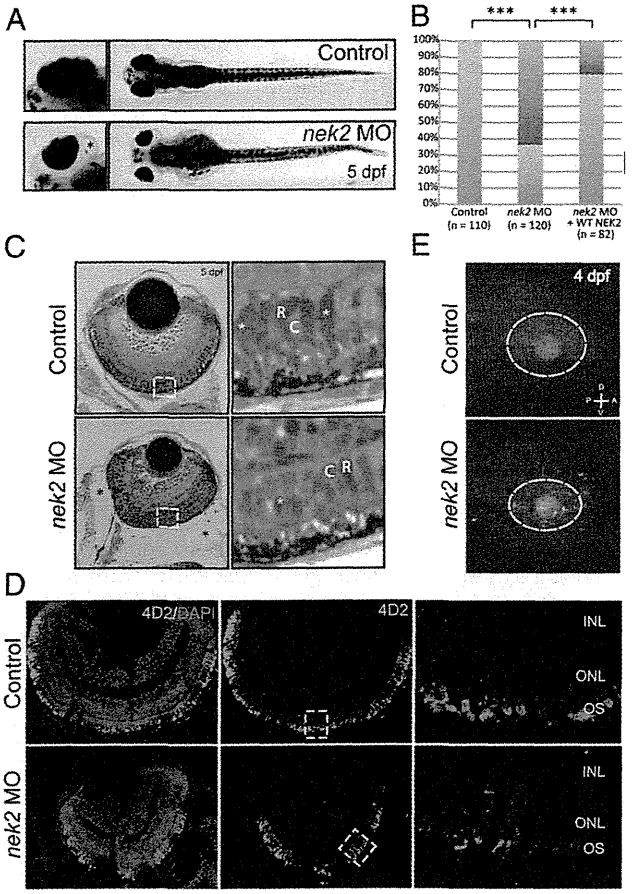

Fig. 4. In vivo functional evaluation of *nek2* loss in zebrafish. (*A*) Bright-field representation of 5-dpf control and *nek2* morphant zebrafish embryos. Magnified *Insets* highlight ocular phenotypes including microphthalmia and enlarged eye sockets (marked by the black asterisk). (*B*) Ocular phenotypes including microphthalmia and enlarged eye sockets vs. normal phenotypes (red bars and blue bars, respectively) are quantified in control and *nek2* morphant embryos, as well as in morphant animals rescued with human WT *NEK2* mRNA. Asterisks indicate statistically significant differences between groups ($P < 0.001$). (*C*) Histology of control and *nek2* morphant embryos also show enlarged eye sockets (marked by black asterisks) and microphthalmia. Magnified *Insets* show a decrease in the number of photoreceptors with apparent changes in domains of condensed chromatin (white asterisks). C, cones; R, rods. (*D*) Immunohistochemical analyses of retinal cryosections from control and *nek2* MO embryos, stained with DAPI (blue) and the 4D2 antibody against rhodopsin (green). Suppression of *nek2* results in the depletion of rods and in the mislocalization of rod opsin from the outer segment (OS) of photoreceptors. INL, inner nuclear layer; ONL, outer nuclear layer. (*E*) TUNEL immunofluorescent images of 4-dpf embryos, showing an increase in the number of apoptotic cells in *nek2* morphant embryos. The dotted ovals indicate the position of the eye. A, anterior; D, dorsal; P, posterior; V, ventral.

serial sectioning of 10–20 embryos injected with sham, MO, or MO + human *NEK2* mRNA, we observed a persistent decrease in the number of photoreceptors with large central domains of condensed chromatin. This phenotype was seen in all *nek2* MO embryos evaluated, but was absent from embryos injected with either sham or MO + human *NEK2* mRNA, suggesting a loss of rod photoreceptors specific to the suppression of *nek2* (Fig. 4*C* and *SI Appendix*, Fig. S5). To verify this observation, we used a rhodopsin (4D2) antibody to stain retinal cryosections from embryos injected with sham, MO, or MO + human *NEK2* mRNA (Fig. 4*D* and *SI Appendix*, Fig. S5). Immunohistochemical analyses of cross-sections from each condition demonstrated that the suppression of *nek2* resulted in the depletion of ~24% of 4D2-positive rod photoreceptors. In addition, mislocalization of rod opsin throughout the photoreceptor cells was evident in the central retina of *nek2* MO specimens, consistent with the hypothesis that nek2 is required for the appropriate trafficking of rhodopsin to the outer segments (Fig. 4*D*).

Further, to ask whether apoptosis, a major mechanism of photoreceptor loss in most known forms of RP (14), might account for some of the observed loss of photoreceptors, we performed TUNEL analysis. Masked scoring of embryos (~50 embryos per injection mixture) revealed a sevenfold increase in the number of TUNEL-positive cells in the eye and head region of *nek2* morphant embryos. By sharp contrast, we did not observe more than 1–10 TUNEL-positive cells in embryos injected with MO + *NEK2* mRNA (Fig. 4*E*).

Finally, we were intrigued by the discovery of a heterozygous frameshift variant p.L206fs in *NEK2* and the bona fide *RPGR* mutation p.E802fs in a patient with RP. We therefore asked whether the *RPGR* variant may interact genetically with the *NEK2* locus. To test this possibility, we coinjected subeffective doses of the *nek2* MO and *rpgr* MO and compared embryos with single or double MO ($n > 100$ at subeffective doses). Approximately 28% of embryos carrying subeffective doses of both *nek2* and *rpgr* MO revealed ocular ($P < 0.001$) and rod photoreceptor phenotypes (serial sectioning of 10 embryos per genotype) that exceeded the number of affected embryos induced by either *nek2* (3%) or *rpgr* (10%) MO alone, suggesting that the RPGR allele interacts *in trans* with the NEK2 locus to exacerbate photoreceptor defects (*SI Appendix*, Fig. S6).

Discussion

Massively parallel sequencing has proven to have a high potential to detect mutations in patients with rare Mendelian diseases (15). To date, most reports focus on monogenic conditions with no genetic heterogeneity, for which mutations can be recognized from benign variants since they invariantly affect the same gene in different patients.

In this study, we explored the efficacy of WGS in identifying mutations in unrelated patients from diverse ethnic backgrounds and presenting with a disease that is clinically the same but that has different genetic drivers. Whereas the small number of genomes analyzed in this study precludes an accurate analysis of quantitative measures, such as sensitivity of the WGS to detect mutations in known RP genes, we observed a few features that allowed us to make some valid comparisons between the different techniques currently available for genetic diagnosis. First, the majority of the pathogenic mutations identified were never reported before. This implies that tools that rely on systematic search for known pathogenic variants, both via mutation-centered resequencing and chip-based hybridization, may not be adequate for ARRP. Second, thanks to full-genome data, we detected complex structural variants whose junctions were located deep in noncoding regions. Because of their nature, these disease-causing variants would have been invisible to standard screening methods, or even to WES. Coverage-based analysis of CNV in exome sequencing has been attempted, with variable results. Limitations of this approach include the uneven efficiency of target DNA capture (and hence sequence coverage, on which assessment of number of copies is based) over different probes and, above all, the low probability of detecting junctions defining the SVs, which are more likely to be found in the nonexonic sequences composing ~98% of our genome. Unambiguous detection of abnormal junctions and mate pair information are crucial parameters in defining a SV; for instance, they allow distinguishing a tandem duplication from an inverted one. Third, because we had access to the full wealth of genomic information, we could integrate many sources of information

at once (e.g., SNP genotypes, phasing, etc.) that allowed us to accurately filter DNA variants that were related to the disease.

Genetic defects in *EYS* were proposed recently to be one of the major causes of ARRP in the Japanese population (16). We found that one of the pathogenic *EYS* alleles was a large SV (446 kb) with a complex genomic rearrangement. This finding supports the notion that SVs represent frequent pathogenic mutations in this gene (17). A homozygous nonsense mutation in exon 6 of *DFNB31* was identified in R18, a patient with nonsyndromic ARRP. The *DFNB31* gene encodes whirlin, a PDZ scaffold protein with expression in both hair cell stereocilia and retinal photoreceptor cells. Whirlin binds to the protein encoded by *USH2A* (18), a gene associated with both Usher syndrome type II (ARRP accompanied by hearing loss) and nonsyndromic ARRP (19). Whereas mutations in *DFNB31* have been reported as rare causes of Usher syndrome type II (20, 21), no DNA changes in its sequence have yet been associated with nonsyndromic ARRP. However, at the age of 66, the past medical history of this patient was significant for only hyperlipidemia and she did not report any hearing loss. We could not perform an auditory examination because she was no longer reachable.

In patients from consanguineous families, regions of IBD allowed restricting the search for pathogenic mutations to only a fraction of the genome. However, these same regions were susceptible to carrying other rare but nonpathogenic homozygous changes as well. Indeed, a higher number of candidate genes/mutations remained among Japanese patients with parental consanguinity compared with those without it (*SI Appendix*, Table S4). These results suggest that even if the analysis should be restricted to areas of IBD, genomes with high homozygosity do not necessarily offer an extra advantage in mutation detection, when comprehensive genomic sequencing in single individuals is performed.

In three patients we identified clear-cut pathogenic but heterozygous mutations in known ARRP genes that could not be associated directly with the disease. This was particularly evident for patient R14, who carried a heterozygous frameshift in *DFNB31* but was also homozygous for a mutation inactivating *PDE6B* (22). These findings are not surprising, given the elevated number of recessive ARRP mutations that are predicted to be present in the general population. Based both on theoretical assessments and on experimental data from control cohorts, we estimated that 1 in 3–7 individuals could be potential heterozygous carriers of an ARRP mutation (23, 24) or, as in the present case, 3 in 16.

The reasons why no candidate mutations of similar quality (i.e., two mutations, at least one of them being clearly deleterious in nature) to those revealed in known RP genes was uncovered in most of the unresolved genomes are unknown. Explanations for this observation may include the presence of variants or SVs that were undetected because of problems inherent in the mapping or sequencing procedure, or of less obvious pathogenic changes that alter splicing or transcription. These would include variants located in introns or in promoter regions, synonymous changes, or changes lying within important yet unannotated exons, genes, or genetic elements that have not been explored in the current study. Diseases caused by oligogenic modes of inheritance, or perhaps attributable to missense mutations for which efficient prioritization is difficult, is another possible explanation. De novo mutations in unknown dominant RP genes could also be evoked.

The search for mutations in unknown disease-causing genes revealed a number of genes with two nonsynonymous changes, which were mostly previously undescribed missenses. Application of more stringent filtering criteria by imposing the presence of at least one deleterious mutation followed by targeted annotation highlighted a single candidate, *NEK2*, in a Japanese patient who carried a homozygous frameshift in this gene. The serine/threonine-protein kinase NEK2 is known to play an important role in regulation of cell cycle progression through localization to the centrosomes and interaction with microtubules (25). The identified frameshift would result either in the creation of premature stop codon yielding a null allele or (less likely) a truncated protein lacking kinase activity and loss of microtubule binding. Importantly, defects in members of the Nek kinase family have been linked to impaired ciliogenesis and polycystic kidney disease (26). Recently, a role for *Nek2* in the left–right patterning of vital organs (a phenotype associated with ciliary function) was established in *Xenopus laevis* (27). In the same work, in situ hybridization revealed the expression of *nek2* transcripts in the eye (27). Furthermore, because NEK2 interacts with and can phosphorylate rootletin, a component of photoreceptor cilia (28, 29), *NEK2* was considered to be an important candidate for ARRP.

Our zebrafish studies showed that lack of Nek2 induces microphthalmia as a gross morphological phenotype. More importantly, in *nek2* morphants, we observed mistrafficking of rhodopsin, a hallmark of photoreceptor disease (30), and a reduced number of rod photoreceptors, likely via apoptotic processes. These phenotypes were rescued by injection of wild-type human *NEK2* mRNA, validating the specificity of the induced defects. Microphthalmia is a phenotype that is difficult to interpret in the present context but that is not uncommon to zebrafish models of RP (31, 32). Meanwhile, photoreceptor death, mistrafficking of rhodospsin, and reduction of the outer retinal layers are classical features of RP in both patients and animal models (7, 14, 33). Indeed, no microphthalmia was noted in patient R19.

Intriguingly, the *NEK2* frameshift identified in R19 was also present in R51, another patient with RP who had a deleterious mutation in *RPGR*. As the *RPGR* mutation in itself could explain the disease, an obvious question was whether the *NEK2* mutation might in fact represent a common benign allele. We therefore searched for this variant in 1,273 control Japanese individuals and found that none carried it (allele frequency $<3.9 \times 10^{-4}$). The p.L206fs mutation in *NEK2* is therefore exceedingly rare, such that its presence in a homozygous state in a patient is a strong argument in favor of its being an uncommon cause for ARRP. Although it is possible to attribute the presence of both *NEK2* and *RPGR* mutations in R51 to chance, a more parsimonious explanation is that mutations in these two genes, both expressed in the connecting cilium, act synergistically to define a severe RP phenotype, due to the established principles of mutational load and oligogenic interactions of pathogenic alleles (34). In turn, this would increase the likelihood for the patient of being examined at earlier ages and analyzed genetically. Multiple genetic modifier genes have been reported for cilia-encoding genes and especially for RPGR (35). These modifiers may account in part for the wide phenotypic spectrum associated with genetic defects in this gene, ranging from localized macular atrophy to retinitis pigmentosa of variable severity. To investigate the possibility of the cooperative effect between deficiencies in these two ciliary genes, we performed in vivo genetic interaction studies and showed that loss of Rpgr function can exacerbate Nek2 ocular phenotypes, including defects comprising the trapping of rhodopsin in the inner segment. Taken together, our genetic and functional data indicate that *NEK2* is a disease gene and that the retinal phenotype that results from its deficiency may represent a newly recognized ciliopathy.

To date, WGS has not been as widely explored as WES in the context of mutation detection. This can be attributed mainly to cost-related issues, because WGS is at least twice as expensive as WES procedures ensuring the same average coverage. We believe that the additional features displayed by WGS are worth the difference in price; however, this is a rather subjective matter that also depends on the disease that is being investigated. In the present case, WGS was essential to identify two pathogenic structural variations originating in introns. This is a significant finding, considering that only seven genomes could undergo SV analysis. Therefore, as a general rule, WGS is probably the strategy of

choice when detection of structural variants or mutations in noncoding regions represents an important element of investigation. In the long term, considering that costs associated with massively parallel sequencing technology is expected to fall further and that analysis pipelines continue to evolve, it is probable that WGS would be just as workable economically and physically as WES. Limitations of WGS include the requirement of high-quality DNA to explore the full leverage of the mate-pair mapping and the lack of reliable pipelines to detect SVs ranging in size from 50 to a few hundred base pairs. Unexpectedly, the difficulty accompanied by handling the large amount of data produced by WGS was not a significant obstacle, given the power of desktop computers presently available on the market. Whereas samples with suitable quality could be obtained through careful preparation of fresh DNA samples, under detection of SVs may be a more problematic issue to solve. This occurs because current mapping is based on two steps: mapping of the short reads aimed at detecting variations between 1 and 50 bp and mate-pair mapping for detection of SVs larger than a few hundred bases; to our knowledge, a solution that could fill the gap between these two mapping approaches remains to be found.

In conclusion, in this study we identified clear-cut causative mutations among the overwhelming number of DNA variants present in the human genome, in single patients from genetically diverse populations. This happened without ambiguities in a highly heterogeneous disease, ARRP, and in more than 50% of the individuals analyzed. Furthermore, two cases presented mutations involving noncoding parts of the genome. Considering that the majority of patients referred for molecular genetics diagnosis are isolated individuals, our results are relevant not only to basic research, but also to future clinical genetic testing.

Methods

Our research protocol involving humans and animals was approved by the institutional review boards of our respective universities and organizations. Written informed consent for providing medical information and blood samples was obtained from each patient. Experimental procedures are described in detail in *SI Appendix, Methods*.

ACKNOWLEDGMENTS. We thank Anna M. Siemiatkowska and Frans P. M. Cremers for sharing material from a person with RP, Adriana Ransijn for technical help, as well as Andrea Superti-Furga, and Luisa Bonafé for fruitful suggestions. Data storage was ensured by the Vital-IT Center for high-performance computing of the Swiss Institute of Bioinformatics. This work was supported by the Swiss National Science Foundation (Grant 310030_138346) and the Gebert Rüf Foundation, Switzerland (Rare Diseases-New Technologies Grant) (both to C.R.); a Center Grant from the Foundation Fighting Blindness (to E.L.B.); National Institutes of Health Grants DK072301 and MH-084018 (to N.K.); Ministry of Health, Labor and Welfare (MHLW) of Japan [Grant 23300101 (to S.I. and N. Matsumoto) and Grant 23300201 (to S.I.)]; MHLW, the Japan Science and Technology Agency, and the Strategic Research Program for Brain Sciences (N. Matsumoto); and a Grant-in-Aid for Scientific Research on Innovative Areas (transcription cycle) from the Ministry of Education, Culture, Sports, Science and Technology of Japan and the Takeda Science Foundation (to N. Matsumoto).

1. Maguire AM, et al. (2008) Safety and efficacy of gene transfer for Leber's congenital amaurosis. *N Engl J Med* 358(21):2240–2248.
2. Bainbridge JW, et al. (2008) Effect of gene therapy on visual function in Leber's congenital amaurosis. *N Engl J Med* 358(21):2231–2239.
3. Cideciyan AV, et al. (2008) Human gene therapy for RPE65 isomerase deficiency activates the retinoid cycle of vision but with slow rod kinetics. *Proc Natl Acad Sci USA* 105(39):15112–15117.
4. Berson EL (1993) Retinitis pigmentosa. The Friedenwald Lecture. *Invest Ophthalmol Vis Sci* 34(5):1659–1676.
5. Berson EL, Rosner B, Sandberg MA, Weigel-DiFranco C, Willett WC (2012) ω-3 intake and visual acuity in patients with retinitis pigmentosa receiving vitamin A. *Arch Ophthalmol* 130(6):707–711.
6. Berson EL, et al. (1993) A randomized trial of vitamin A and vitamin E supplementation for retinitis pigmentosa. *Arch Ophthalmol* 111(6):761–772.
7. Hartong DT, Berson EL, Dryja TP (2006) Retinitis pigmentosa. *Lancet* 368(9549): 1795–1809.
8. Jin ZB, et al. (2008) Identifying pathogenic genetic background of simplex or multiplex retinitis pigmentosa patients: A large scale mutation screening study. *J Med Genet* 45(7):465–472.
9. Tucker T, Marra M, Friedman JM (2009) Massively parallel sequencing: The next big thing in genetic medicine. *Am J Hum Genet* 85(2):142–154.
10. Tuson M, Marfany G, Gonzàlez-Duarte R (2004) Mutation of CERKL, a novel human ceramide kinase gene, causes autosomal recessive retinitis pigmentosa (RP26). *Am J Hum Genet* 74(1):128–138.
11. Aldahmesh MA, et al. (2009) Molecular characterization of retinitis pigmentosa in Saudi Arabia. *Mol Vis* 15:2464–2469.
12. Vervoort R, et al. (2000) Mutational hot spot within a new RPGR exon in X-linked retinitis pigmentosa. *Nat Genet* 25(4):462–466.
13. Sanuki R, et al. (2011) miR-124a is required for hippocampal axogenesis and retinal cone survival through Lhx2 suppression. *Nat Neurosci* 14(9):1125–1134.
14. Cottet S, Schorderet DF (2009) Mechanisms of apoptosis in retinitis pigmentosa. *Curr Mol Med* 9(3):375–383.
15. Rabbani B, Mahdieh N, Hosomichi K, Nakaoka H, Inoue I (2012) Next-generation sequencing: Impact of exome sequencing in characterizing Mendelian disorders. *J Hum Genet* 57(10):621–632.
16. Hosono K, et al. (2012) Two novel mutations in the EYS gene are possible major causes of autosomal recessive retinitis pigmentosa in the Japanese population. *PLoS ONE* 7(2):e31036.
17. Pieras JI, et al. (2011) Copy-number variations in EYS: A significant event in the appearance of arRP. *Invest Ophthalmol Vis Sci* 52(8):5625–5631.
18. van Wijk E, et al. (2006) The DFNB31 gene product whirlin connects to the Usher protein network in the cochlea and retina by direct association with USH2A and VLGR1. *Hum Mol Genet* 15(5):751–765.
19. Rivolta C, Sweklo EA, Berson EL, Dryja TP (2000) Missense mutation in the USH2A gene: Association with recessive retinitis pigmentosa without hearing loss. *Am J Hum Genet* 66(6):1975–1978.
20. Ebermann I, et al. (2007) A novel gene for Usher syndrome type 2: Mutations in the long isoform of whirlin are associated with retinitis pigmentosa and sensorineural hearing loss. *Hum Genet* 121(2):203–211.
21. Yang J, et al. (2010) Ablation of whirlin long isoform disrupts the USH2 protein complex and causes vision and hearing loss. *PLoS Genet* 6(5):e1000955.
22. McLaughlin ME, Sandberg MA, Berson EL, Dryja TP (1993) Recessive mutations in the gene encoding the beta-subunit of rod phosphodiesterase in patients with retinitis pigmentosa. *Nat Genet* 4(2):130–134.
23. Rivolta C, Sharon D, DeAngelis MM, Dryja TP (2002) Retinitis pigmentosa and allied diseases: Numerous diseases, genes, and inheritance patterns. *Hum Mol Genet* 11(10): 1219–1227.
24. Nishiguchi KM, Rivolta C (2012) Genes associated with retinitis pigmentosa and allied diseases are frequently mutated in the general population. *PLoS ONE* 7(7):e41902.
25. Fry AM, Meraldi P, Nigg EA (1998) A centrosomal function for the human Nek2 protein kinase, a member of the NIMA family of cell cycle regulators. *EMBO J* 17(2): 470–481.
26. Quarmby LM, Mahjoub MR (2005) Caught Nek-ing: Cilia and centrioles. *J Cell Sci* 118(Pt 22):5161–5169.
27. Fakhro KA, et al. (2011) Rare copy number variations in congenital heart disease patients identify unique genes in left-right patterning. *Proc Natl Acad Sci USA* 108(7): 2915–2920.
28. Bahe S, Stierhof YD, Wilkinson CJ, Leiss F, Nigg EA (2005) Rootletin forms centriole-associated filaments and functions in centrosome cohesion. *J Cell Biol* 171(1):27–33.
29. Yang J, et al. (2002) Rootletin, a novel coiled-coil protein, is a structural component of the ciliary rootlet. *J Cell Biol* 159(3):431–440.
30. Hollingsworth TJ, Gross AK (2012) Defective trafficking of rhodopsin and its role in retinal degenerations. *Int Rev Cell Mol Biol* 293:1–44.
31. Luo N, Lu J, Sun Y (2012) Evidence of a role of inositol polyphosphate 5-phosphatase INPP5E in cilia formation in zebrafish. *Vision Res* 75:98–107.
32. Patil SB, Hurd TW, Ghosh AK, Murga-Zamalloa CA, Khanna H (2011) Functional analysis of retinitis pigmentosa 2 (RP2) protein reveals variable pathogenic potential of disease-associated missense variants. *PLoS ONE* 6(6):e21379.
33. Chang GQ, Hao Y, Wong F (1993) Apoptosis: Final common pathway of photoreceptor death in rd, rds, and rhodopsin mutant mice. *Neuron* 11(4):595–605.
34. Davis EE, Katsanis N (2012) The ciliopathies: A transitional model into systems biology of human genetic disease. *Curr Opin Genet Dev* 22(3):290–303.
35. Fahim AT, et al. (2011) Allelic heterogeneity and genetic modifier loci contribute to clinical variation in males with X-linked retinitis pigmentosa due to RPGR mutations. *PLoS ONE* 6(8):e23021.

Epilepsia, 54(7):1262–1269, 2013
doi: 10.1111/epi.12203

FULL-LENGTH ORIGINAL RESEARCH

Targeted capture and sequencing for detection of mutations causing early onset epileptic encephalopathy

***Hirofumi Kodera, †Mitsuhiro Kato, ‡[1]Alex S. Nord, ‡Tom Walsh, ‡Ming Lee, §Gaku Yamanaka, ¶Jun Tohyama, *†Kazuyuki Nakamura, #Eiji Nakagawa, **Tae Ikeda, ††Bruria Ben-Zeev, ‡‡Dorit Lev, ‡‡Tally Lerman-Sagie, §§Rachel Straussberg, ¶¶Saori Tanabe, ##Kazutoshi Ueda, ***Masano Amamoto, †††Sayaka Ohta, ‡‡‡Yutaka Nonoda, *Kiyomi Nishiyama, *Yoshinori Tsurusaki, *Mitsuko Nakashima, *Noriko Miyake, †Kiyoshi Hayasaka, ‡Mary-Claire King, *Naomichi Matsumoto, and *Hirotomo Saitsu**

***Department of Human Genetics, Yokohama City University Graduate School of Medicine, Yokohama, Japan; †Department of Pediatrics, Yamagata University Faculty of Medicine, Yamagata, Japan; ‡Department of Genome Sciences and Department of Medicine, University of Washington, Seattle, Washington, U.S.A.; §Department of Pediatrics, Tokyo Medical University, Tokyo, Japan; ¶Department of Pediatrics, Nishi-Niigata Chuo National Hospital, Niigata, Japan; #Department of Child Neurology, National Center of Neurology and Psychiatry, Tokyo, Japan; **Division of Pediatric Neurology, Osaka Medical Center and Research Institute for Maternal and Child Health, Osaka, Japan; ††The Edmond and Lily Safra Children's Hospital, Sheba Medical Center, Ramat Gan, Israel; ‡‡Metabolic Neurogenetic Clinic, Wolfson Medical Center, Holon, Israel; §§Department of Neurogenetics, Schneider's Children Medical Center, Petah Tiqwa, Israel; ¶¶Department of Pediatrics, Nihonkai General Hospital, Sakata, Japan; ##Department of Pediatrics, Kitano Hospital, Osaka, Japan; ***Pediatric Emergency Center, Kitakyusyu City Yahata Hospital, Kitakyushu, Japan; †††Department of Pediatrics, Graduate School of Medicine, University of Tokyo, Tokyo, Japan; and ‡‡‡Department of Pediatrics, School of Medicine, Kitasato University, Sagamihara, Japan**

Summary

Purpose: Early onset epileptic encephalopathies (EOEEs) are heterogeneous epileptic disorders caused by various abnormalities in causative genes including point mutations and copy number variations (CNVs). In this study, we performed targeted capture and sequencing of a subset of genes to detect point mutations and CNVs simultaneously.

Methods: We designed complementary RNA oligonucleotide probes against the coding exons of 35 known and potential candidate genes. We tested 68 unrelated patients, including 15 patients with previously detected mutations as positive controls. In addition to mutation detection by the Genome Analysis Toolkit, CNVs were detected by the relative depth of coverage ratio. All detected events were confirmed by Sanger sequencing or genomic microarray analysis.

Key Findings: We detected all positive control mutations. In addition, in 53 patients with EOEEs, we detected 12 pathogenic mutations, including 9 point mutations (2 nonsense, 3 splice-site, and 4 missense mutations), 2 frameshift mutations, and one 3.7-Mb microdeletion. Ten of the 12 mutations occurred de novo; the other two had been previously reported as pathogenic. The entire process of targeted capture, sequencing, and analysis required 1 week for the testing of up to 24 patients.

Significance: Targeted capture and sequencing enables the identification of mutations of all classes causing EOEEs, highlighting its usefulness for rapid and comprehensive genetic testing.

KEY WORDS: Target capture, Sequencing, Mutation, Copy number variation, Genetic testing.

Early onset epileptic encephalopathies (EOEEs), occurring before 1 year of age, are characterized by impairment of cognitive, sensory, and motor development by recurrent clinical seizures or prominent interictal epileptiform discharges (Berg et al., 2010). Ohtahara syndrome (OS), West syndrome (WS), early myoclonic encephalopathy (EME), migrating partial seizures in infancy (MPSI), and Dravet syndrome (DS) are the best known epileptic encephalopathies recognized by the International League Against Epilepsy (ILAE; Berg et al., 2010). However, many infants with similar features do not strictly fit the parameters of these syndromes.

To date, 11 genes have been shown to be associated with EOEEs (Mastrangelo & Leuzzi, 2012). The identification of

Accepted March 21, 2013; Early View publication May 10, 2013.
Address correspondence to Hirotomo Saitsu, Department of Human Genetics, Yokohama City University Graduate School of Medicine, 3-9 Fukuura, Kanazawa-ku, Yokohama 236-0004, Japan. E-mail: hsaitsu@yokohama-cu.ac.jp
[1]Present address: Genomics Division, Lawrence Berkeley National Laboratory, Berkeley, California, U.S.A.

causative mutations associated with EOEEs and their related phenotypes is useful for genetic counseling, and possibly for management of the patients; however, it is time-consuming and arduous to screen all known disease-causing genes one by one using Sanger sequencing or high-resolution melting curve analysis (Wittwer, 2009). In addition, copy number variations (CNVs) involving causative genes can also cause EOEEs (Saitsu et al., 2008; Mei et al., 2010; Saitsu et al., 2011, 2012b). Array comparative genomic hybridization (CGH) and multiplex ligation-dependent probe amplification (MLPA) are well established for the detection of CNVs; however, it is often difficult for array CGH to detect small CNVs such as a single-exon deletion and for MLPA to screen multiple genes at a time (Schouten et al., 2002; Dibbens et al., 2011; Mefford et al., 2011; Stuppia et al., 2012). Therefore, an integrated method that detects both point mutations and CNVs for multiple genes would be useful for comprehensive genetic testing in EOEEs.

Recent progress in massively parallel DNA sequencing in combination with target capturing has facilitated rapid mutation detection (Ng et al., 2009). It has been reported that CNVs involving disease-causing genes in patients with breast or ovarian cancer can be detected by target capture sequencing using the relative depth of coverage ratio (Walsh et al., 2010, 2011; Nord et al., 2011). Targeted capture and sequencing of patients with epileptic disorders has successfully identified potential disease-causing mutations in 16 of 33 patients (Lemke et al., 2012), revealing its efficacy for detecting mutations. However, the detection of both point mutations and CNVs has not been reported in patients with epilepsy.

In this study, we performed targeted capture and sequencing of a subset of 35 genes to detect mutations and CNVs simultaneously in 68 patients with EOEEs. By analyzing the relative depth of coverage ratio, we were able to detect microdeletions, in which the numbers of deleted exons varied from a single exon to all exons of two genes. In combination with rapid sequencing using a benchtop next-generation sequencer, our method provides a fast, comprehensive, and cost-effective method for genetic testing of patients with EOEE.

Methods

Patients

We examined 68 patients (36 male and 32 female) with EOEEs (20 patients with OS, 20 with WS, 3 with EME, 4 with MPSI, 2 with DS, and 19 with unclassified epileptic encephalopathy). Diagnoses were based on clinical features and characteristic patterns on electroencephalography. In 15 of 68 patients (10 male and five female), disease-causing mutations or CNVs had been previously identified in our laboratory, so these mutations were used as positive controls (Table 1) (Saitsu et al., 2008, 2010a,b, 2011, 2012b,c; Nonoda et al., 2013). Genomic DNA was isolated from blood leukocytes according to standard methods. Experimental protocols were approved by the Yokohama City University School of Medicine Institutional Review Board for Ethical Issues. Written informed consent for genetic testing was obtained from the guardians of all tested individuals prior to analysis.

Target capture sequencing and variant detection

A custom-made SureSelect oligonucleotide probe library (Agilent Technologies, Santa Clara, CA, U.S.A.) was designed to capture the coding exons of 35 genes; 5 of them were potential candidates for EOEEs based on unpublished data (for a list of the 30 of 35 genes, see Table 2). We designed 120-bp capture probes with 3× centered probe-tiling, and avoiding 20-bp overlap to repeat region using the Agilent e-Array Web-based design tool. To cover regions

Table 1. Known mutations and copy number variants used as positive controls

	Case	Sex	Chr	Genes	Reported mutations or copy number variants (positive controls)	Type	Deletion size (kb)	Refs
SNVs	27	F	9	*STXBP1*	c.1328T>G (p.Met443Arg)	Missense		Saitsu et al. (2008)
	69	M	X	*CASK*	c.1A>G	Missense		Saitsu et al. (2012b)
	241	M	X	*CDKL5*	c.145G>A (p.Glu49Lys)	Missense		–
Indels	95	M	9	*STXBP1*	c.388_389del (p.Leu130AspfsX11)	Deletion		Saitsu et al. (2010a)
	313	M	X	*CASK*	c.227_228del (p.Glu76ValfsX6)	Deletion		–
	26	F	9	*SPTAN1*	c.6619_6621del (p.Glu2207del)	Deletion		Saitsu et al. (2010b)
	220	M	9	*STXBP1*	c.1381_1390del (p.Lys461GlyfsX82)	Deletion		–
	16	M	9	*SPTAN1*	c.6923_6928dup (p.Arg2308_Met2309dup)	Duplication		Saitsu et al. (2010b)
	309	M	9	*SPTAN1*	c.6908_6916dup (p.Asp2303_Leu2305dup)	Duplication		Nonoda et al. (2013)
CNVs	12	F	9	*STXBP1*, *SPTAN1*	Del(9)(q33.33–q34.11)	Microdeletion	2150	Saitsu et al. (2008)
	22	M	9	*STXBP1*	*STXBP1* Ex4 deletion	Microdeletion	4.6	Saitsu et al. (2011)
	83	M	X	*CASK*	*CASK* Ex2 deletion	Microdeletion	111	Saitsu et al. (2012b)
	102	F	X	*MECP2*	Del(X)(q28)	Microdeletion		–
	204	M	9	*STXBP1*, *SPTAN1*	Del(9)(q33.33–q34.11)	Microdeletion	2850	Saitsu et al. (2011)
	214	F	X	*CDKL5*	Del(X)(q22.13)	Microdeletion	137	Saitsu et al. (2011)

SNVs, single nucleotide variants; Indels, insertion/deletions; CNVs, copy number variations.

Table 2. Sequence performance for 30 target genes

Gene	Cytoband	No. of coding exons	Mean read depth	%bases above 5× depth (%)	%bases above 10× depth (%)
ARHGEF9	Xq11.1–q11.2	10	206	100	100
ARX	Xp21.3	5	44	59.4–94.4	38.7–90.6
CASK	Xp11.4	27	201	95.9–100	95.9–100
CDKL5	Xp22.13	20	238	100	100
COL4A1	13q34	52	287	98.3–100	98.3–100
COL4A2	13q34	47	190	100	99.1–100
FOXG1	14q12	1	231	86.5–100	81.1–96.4
GABRG2	5q34	11	300	92.3	92.3
GRIN2A	16p13.2	13	310	100	100
KCNQ2	20q13.33	17	135	100	97.7–100
MAGI2	7q21.11	22	255	96–98.3	94.5–97.5
MAPK10	4q21.3	12	304	100	100
MECP2	Xq28	3	217	96.2	96.2
MEF2C	5q14.3	10	270	100	100
NTNG1	1p13.3	9	298	100	100
PCDH19	Xq22.1	6	212	100	100
PLCB1	20p12.3	32	293	100	100
PNKP	19q13.33	17	208	100	98.5–100
PNPO	17q21.32	7	210	100	100
SCN1A	2q24.3	26	345	100	100
SCN2A	2q24.3	26	323	100	100
SLC25A22	11p15.5	9	121	100	100
SLC2A1	1p34.2	10	209	100	98.8–100
SNPH	20p13	4	179	100	100
SPTAN1	9q34.11	56	277	100	100
SRGAP2	1q32.1	20	320	96.6	96.6
ST3GAL5	2p11.2	8	302	93.6–100	93.6–99.9
STXBP1	9q34.11	20	306	100	100
SYN1	Xp11.23	13	131	93.4–100	81–100
SYP	Xp11.23	6	146	100	99.1–100

where we could not design probes with the above settings, some probes from the SureSelect Human All Exon 50-Mb kit (Agilent Technologies) were added to the probe libraries. A total of 2,738 probes, covering 156 kb, were prepared. Exon capture, enrichment, and indexing were performed according to the manufacturer's instructions. Twenty-four captured libraries were mixed and sequenced on an Illumina MiSeq (Illumina, San Diego, CA, U.S.A.) with 150-bp paired-end reads. Image analysis and base calling were performed using the Illumina Real Time Analysis Pipeline version 1.13 and CASAVA software v.1.8 (Illumina) with default parameters. Sequence reads were aligned to the reference human genome (GRCh37: Genome Reference Consortium human build 37) with Novoalign (Novocraft Technologies, Selangor, Malaysia). After conversion of the SAM file to a BAM file with SAMtools (Li et al., 2009), duplicate reads were marked using Picard (http://picard.sourceforge.net/) and excluded from downstream analysis. Local realignment around insertion/deletions (indels) and base quality score recalibration were performed using the Genome Analysis Toolkit (DePristo et al., 2011). Single-nucleotide variants (SNVs) and indels were identified using the Genome Analysis Toolkit UnifiedGenotyper and filtered according to the Broad Institute's best-practice guidelines v.3 except for HaplotypeScore filtering. We excluded variants found in 147 exomes from healthy individuals previously sequenced in our laboratory. Variants were annotated using ANNOVAR (Wang et al., 2010). Candidate disease-causing mutations were confirmed by Sanger sequencing on a 3500xL Genetic Analyzer (Applied Biosystems, Foster City, CA, U.S.A.). The Human Gene Mutation Database professional 2012.3 (BIOBASE GmbH, Wolfenbuettel, Germany) was used to check whether the variants had been previously reported.

Copy number analysis using target capture sequence data

Copy number changes were analyzed based on the relative depth of coverage ratios (Nord et al., 2011). Raw coverage on the target regions was calculated by SAMtools using BAM files, in which duplicate reads were excluded. Raw coverage was normalized and corrected for GC content and bait capture bias. Next, the ratios were calculated by comparing the sample-corrected coverage to the median-corrected coverage for the other 23 samples. A sliding window (20 bp) was used to identify CNVs for which the majority of bases had a ratio ≤ 0.6 (loss) or ≥ 1.4 (gain). We visually inspected the ratio

data and judged whether the call was true or likely to be a false positive. A flow chart of our variant detection and copy number analysis scheme is illustrated in Fig. S1.

Genomic microarray analysis and cloning of deletion breakpoints

The microdeletion involving *SCN1A* and *SCN2A* was confirmed using a CytoScan HD Array (Affymetrix, Santa Clara, CA, U.S.A.) according to the manufacturer's protocol. Copy number alterations were analyzed using the Chromosome Analysis Suite (ChAS; Affymetrix) with NA32 (hg19) annotations. The junction fragment spanning the deletion was amplified by long polymerase chain reaction (PCR) using several primer sets based on putative breakpoints according to the microarray data. Long PCR was performed in a 20-μl volume, containing 30 ng genomic DNA, 1× buffer for KOD FX, 0.4 mM each dNTP, 0.3 μM each primer, and 0.3 U KOD FX polymerase (Toyobo, Osaka, Japan). The deletion junction fragments were obtained using the following primers: #409-F (5′-TCCACAGTTTACAAACATCTTTTCATGG-3′) and #409-R (5′-AGAAATTGGCTTGGTCAGTACCAGCA-3′) (1.6-kb amplicon). PCR products were electrophoresed on agarose gels stained with ethidium bromide, purified with ExoSAP (USB Technologies, Cleveland, OH, U.S.A.), and sequenced with BIGDYE TERMINATOR CHEMISTRY v.3 according to the manufacturer's protocol (Applied Biosystems).

Results

Target capture sequencing yielded an average of 26 Mb per sample (range 17–41 Mb per sample) on the target regions, resulting in an average read depth of 255 (range across all samples: 173–437). The coverage of the protein-coding sequences of the 30 target genes is shown in Table 2. Overall, 98.6% of targeted coding sequence bases were covered by 10 reads or more; however, some genes such as *ARX* and *FOXG1* were less well covered because of embedded repeat sequences (Fig. S2). To validate the performance of target capture sequencing for detecting mutations and CNVs, we analyzed 15 samples in which disease-causing mutations or microdeletions had been identified previously in our laboratory (Saitsu et al., 2008, 2010a,b, 2011, 2012b; Nonoda et al., 2013). All nine control point mutations and six control microdeletions were detected (Table 1; Fig. 1). These data indicate that our target capture sequencing method was able to detect both point mutations and microdeletions, including deletion of a single exon.

Examination of 53 previously unresolved EOEE patients by targeted capture and sequencing revealed mutations in 12 patients (Table 3). Every patient harbored a different

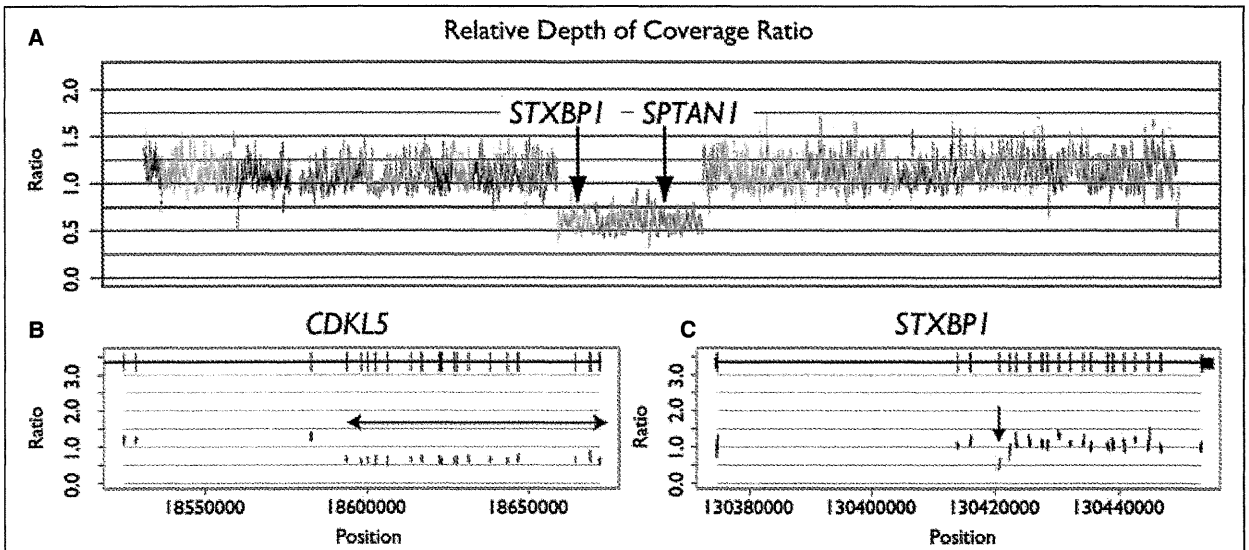

Figure 1.

Detection of three known microdeletions by target capture sequencing. (**A**) Relative depth of coverage ratio for patient 12. Coverage ratios for each target gene are indicated by different colors. A microdeletion including *STXBP1* and *SPTAN1* is clearly observed. (**B**, **C**) Relative depth of coverage ratio for patient 214 in the *CDKL5* region and patient 22 in the *STXBP1* region, respectively. Black vertical lines indicate exons and horizontal lines indicate introns (top). Red vertical lines show bait regions that were judged to be "deleted." A number of exons of *CDKL5* were deleted in patient 214 (bidirectional arrow in **B**), and a single exon of *STXBP1* was deleted in patient 22 (arrow in **C**).

Table 3. Mutations in 53 patients with EOEEs detected by targeted capture and sequencing

	Case	Sex	Diagnosis	Chr	Gene	Mutation	Type	Deletion size (kb)	Inheritance	References
SNVs	329	M	OS/EME	9	*STXBP1*	c.247-2A>G	Splice site		De novo	–
	402	M	OS	9	*STXBP1*	c.902+1G>A	Splice site		De novo	Milh et al. (2011)
	423	F	OS	9	*STXBP1*	c.246+1G>A	Splice site		De novo	–
	403	F	MAE or DS	2	*SCN1A*	c.580G>A (p.Asp194Asn)	Missense		Not found in the mother	Mancardi et al. (2006)
	415	F	EOEE	2	*SCN1A*	c.3714A>C (p.Glu1238Asp)	Missense		Not determined	Harkin et al. (2007)
	416	M	EOEE	X	*CDKL5*	c.533G>A (p.Arg178Gln)	Missense		De novo	Liang et al. (2011)
	418	F	WS, severe hypotonia	2	*SCN2A*	c.632G>A (p.Gly211Asp) in NM_001040143 (variant 3)	Missense		De novo	–
	244	F	Epilepsy + PCH	X	*CASK*	c.55G>T (p.Gly19X)	Nonsense		De novo	–
	404	F	EOEEs	X	*MECP2*	c.844C>T (p.Arg282X)	Nonsense		De novo	–
Indels	336	F	OS	9	*STXBP1*	c.1056del (p.Asp353ThrfsX3)	Deletion		De novo	–
	397	F	DS	2	*SCN1A*	c.342_344delinsAGGAGTT (p.Phe114LeufsX6)	Deletion–insertion		De novo	–
CNV	409	F	MPSI	2	*SCN2A*, *SCN1A*	Microdeletion	Microdeletion	3,726	De novo	–

OS, Ohtahara syndrome; EME, early myoclonic encephalopathy; MAE, myoclonic astatic epilepsy; DS, Dravet syndrome; WS, West syndrome; PCH, pontocerebellar hypoplasia; MPSI, malignant migrating partial seizures in infancy; SNVs, single nucleotide variants; CNVs, copy number variations; EOEEs, early onset epileptic encephalopathies.

mutation. Of these 12 mutations, 9 were single-nucleotide variants (2 nonsense, 3 splice-site, and 4 missense mutations) and two were small indels leading to frameshifts. The other mutation was a microdeletion. All these 11 point mutations were confirmed by Sanger sequencing. Four of the mutations (*STXBP1* c.902+1G>A, *SCN1A* c.580G>A, *SCN1A* c.3714A>C, and *CDKL5* c.533G>A) have been reported in individuals with EOEEs, so are recurrent (Mancardi et al., 2006; Harkin et al., 2007; Azmanov et al., 2010; Liang et al., 2011; Milh et al., 2011). Nine of the 11 mutations occurred de novo. The other two could not be tested because the paternal sample for one patient (*SCN1A* c.580G>A) and parental samples for another patient (*SCN1A* c.3714A>C) were unavailable.

CNV analysis of the 53 patients revealed a microdeletion involving *SCN1A* and *SCN2A* at 2q24.3 in patient 409 (Fig. 2A). To investigate this mutation further, we performed genomic microarray analysis and identified an approximately 3.7-Mb microdeletion (Fig. 2B). The deletion contained 13 RefSeq genes including *SCN2A* and *SCN1A*. Breakpoint-specific PCR analysis of the patient and her parents confirmed that the rearrangement occurred de novo (Fig. 2C). The sequence of the junction fragment confirmed a 3,726,029-bp deletion (chr2: 164,420,771–168,146,801) (Fig. 2D).

Discussion

Several bench-top high-throughput sequencing platforms are now available (Glenn, 2011; Loman et al., 2012; Quail et al., 2012). We selected Illumina MiSeq because it provides reasonable sequence throughput (1.6 Gb per run), a low error rate, a short run time (27 h), and sufficiently long reads (150 bp). We captured genomic DNA fragments of target genes by 3× tiling complementary RNA oligonucleotide probes (Nord et al., 2011) and sequenced 24 samples per MiSeq run, achieving sufficient coverage (a mean read depth of 255) over the target regions. This high coverage enabled us to detect point mutations and CNVs simultaneously, and long reads enabled us to detect small indels (Krawitz et al., 2010). Mapping by Novoalign, we were able to detect indels ranging in size from a 10-bp deletion to a 9-bp duplication.

By evaluating depth of coverage ratios (Nord et al., 2011), we detected six control microdeletions and one novel microdeletion, ranging in size from 4.6 kb to 3.7 Mb. To date, CNVs causing EOEEs have been analyzed by array CGH and MLPA (Mulley & Mefford, 2011). Array CGH can detect genome-wide CNVs, but its standard resolution is relatively low (>10 kb). On the other hand, MLPA can detect CNVs in specific genes, including single exon deletions; however, it is difficult to screen many genes at a time because MLPA is limited to 50 target exons per reaction (Stuppia et al., 2012). In addition, copy number analysis using MLPA can be affected by single nucleotide variants and indels in regions corresponding to the MLPA probes (Stuppia et al., 2012). In contrast, targeted capture and sequencing can analyze all targeted genes to detect mutations and CNVs simultaneously. CNVs as small as a single exon can be identified. Because all the procedures—from

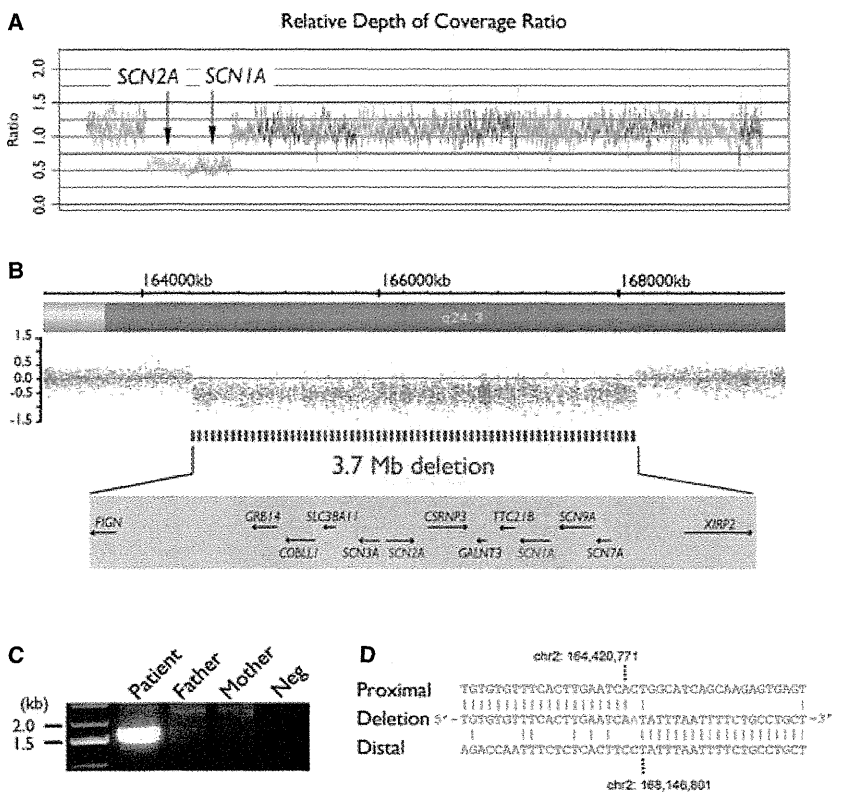

Figure 2.
A 3.7-Mb microdeletion including *SCN2A* and *SCN1A* in patient 409. (**A**) Relative depth of coverage ratio for patient 409 indicates a microdeletion encompassing *SCN2A* and *SCN1A*. Different colors distinguish the target genes. (**B**) The array profile clearly shows a 3.7-Mb microdeletion at 2q24.3 in this patient. Thirteen RefSeq genes, including *SCN2A* and *SCN1A*, lie within the microdeletion (bottom). (**C**) Breakpoint-specific PCR analysis of the patient's family. Primers flanking the deletion were able to amplify a 1,607-bp product from the patient only, indicating that the translocation occurred de novo. (**D**) Deletion junction sequence. The top, middle, and bottom strands show the proximal, deleted, and distal sequences, respectively. A single inserted nucleotide (colored in red) was identified at the breakpoint.

the capture of target genes to the detection of mutations and CNVs—can be done within a week, our workflow provides a fast, sensitive, and comprehensive genetic testing method for patients with epilepsy.

Whole-exome sequencing will reveal novel mutations in unexpected genes in patients with EOEEs. For example, *KCNQ2* mutations, which cause benign familial neonatal seizures (Biervert et al., 1998; Charlier et al., 1998), were identified in patients with OS by whole exome sequencing (Saitsu et al., 2012a). Similarly, screening known and potential candidate genes in patients with EOEEs will reveal novel mutations in unexpected genes, in addition to mutations in well-known genes.

In our target capture analysis, some exons of genes such as *ARX* and *FOXG1* were insufficiently sequenced because repeat sequences hampered the design of capture probes. Repeat sequences also interfere with appropriate mapping of sequence reads, resulting in low coverage. For these exons, Sanger sequencing should be added for complete analysis.

In conclusion, a rapid and efficient system of target capture sequencing can be applied to the comprehensive genetic analysis of EOEEs. Point mutations, small indels, and CNVs are all detected by this method, confirming the potential of this approach for efficient genetic testing.

Acknowledgments

We thank the patients and their families for their participation in this study. We also thank Aya Narita and Nobuko Watanabe for their technical assistance. This work was Supported by the Ministry of Health, Labour and Welfare of Japan (24133701, 11103577, 11103580, 11103340, 10103235), a Grant-in-Aid for Scientific Research (C) from the Japan Society for the Promotion of Science (24591500), a Grant-in-Aid for Young Scientists from the Japan Society for the Promotion of Science (10013428, 11001011, 12020465), the Takeda Science Foundation, the Japan Science and Technology Agency, the Strategic Research

Program for Brain Sciences (11105137), and a Grant-in-Aid for Scientific Research on Innovative Areas (Transcription Cycle) from the Ministry of Education, Culture, Sports, Science and Technology of Japan (12024421).

Disclosure

None of the authors has any conflicts of interest to disclose. We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

References

Azmanov DN, Zhelyazkova S, Dimova PS, Radionova M, Bojinova V, Florez L, Smith SJ, Tournev I, Jablensky A, Mulley J, Scheffer I, Kalaydjieva L, Sander JW. (2010) Mosaicism of a missense SCN1A mutation and Dravet syndrome in a Roma/Gypsy family. *Epileptic Disord* 12:117–124.

Berg AT, Berkovic SF, Brodie MJ, Buchhalter J, Cross JH, van Emde Boas W, Engel J, French J, Glauser TA, Mathern GW, Moshe SL, Nordli D, Plouin P, Scheffer IE. (2010) Revised terminology and concepts for organization of seizures and epilepsies: report of the ILAE Commission on Classification and Terminology, 2005–2009. *Epilepsia* 51:676–685.

Biervert C, Schroeder BC, Kubisch C, Berkovic SF, Propping P, Jentsch TJ, Steinlein OK. (1998) A potassium channel mutation in neonatal human epilepsy. *Science* 279:403–406.

Charlier C, Singh NA, Ryan SG, Lewis TB, Reus BE, Leach RJ, Leppert M. (1998) A pore mutation in a novel KQT-like potassium channel gene in an idiopathic epilepsy family. *Nat Genet* 18:53–55.

DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 43:491–498.

Dibbens LM, Kneen R, Bayly MA, Heron SE, Arsov T, Damiano JA, Desai T, Gibbs J, McKenzie F, Mulley JC, Ronan A, Scheffer IE. (2011) Recurrence risk of epilepsy and mental retardation in females due to parental mosaicism of PCDH19 mutations. *Neurology* 76: 1514–1519.

Glenn TC. (2011) Field guide to next-generation DNA sequencers. *Mol Ecol Resour* 11:759–769.

Harkin LA, McMahon JM, Iona X, Dibbens L, Pelekanos JT, Zuberi SM, Sadleir LG, Andermann E, Gill D, Farrell K, Connolly M, Stanley T, Harbord M, Andermann F, Wang J, Batish SD, Jones JG, Seltzer WK, Gardner A, Sutherland G, Berkovic SF, Mulley JC, Scheffer IE. (2007) The spectrum of SCN1A-related infantile epileptic encephalopathies. *Brain* 130:843–852.

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

Lemke JR, Riesch E, Scheurenbrand T, Schubach M, Wilhelm C, Steiner I, Hansen J, Courage C, Gallati S, Burki S, Strozzi S, Simonetti BG, Grunt S, Steinlin M, Alber M, Wolff M, Klopstock T, Prott EC, Lorenz R, Spaich C, Rona S, Lakshminarasimhan M, Kroll J, Dorn T, Kramer G, Synofzik M, Becker F, Weber YG, Lerche H, Bohm D, Biskup S. (2012) Targeted next generation sequencing as a diagnostic tool in epileptic disorders. *Epilepsia* 53:1387–1398.

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.

Liang JS, Shimojima K, Takayama R, Natsume J, Shichiji M, Hirasawa K, Imai K, Okanishi T, Mizuno S, Okumura A, Sugawara M, Ito T, Ikeda H, Takahashi Y, Oguni H, Imai K, Osawa M, Yamamoto T. (2011) *CDKL5* alterations lead to early epileptic encephalopathy in both genders. *Epilepsia* 52:1835–1842.

Loman NJ, Misra RV, Dallman TJ, Constantinidou C, Gharbia SE, Wain J, Pallen MJ. (2012) Performance comparison of benchtop high-throughput sequencing platforms. *Nat Biotechnol* 30:434–439.

Mancardi MM, Striano P, Gennaro E, Madia F, Paravidino R, Scapolan S, Dalla Bernardina B, Bertini E, Bianchi A, Capovilla G, Darra F, Elia M, Freri E, Gobbi G, Granata T, Guerrini R, Pantaleoni C, Parmeggiani A, Romeo A, Santucci M, Vecchi M, Veggiotti P, Vigevano F, Pistorio A, Gaggero R, Zara F. (2006) Familial occurrence of febrile seizures and epilepsy in severe myoclonic epilepsy of infancy (SMEI) patients with *SCN1A* mutations. *Epilepsia* 47:1629–1635.

Mastrangelo M, Leuzzi V. (2012) Genes of early-onset epileptic encephalopathies: from genotype to phenotype. *Pediatr Neurol* 46: 24–31.

Mefford HC, Yendle SC, Hsu C, Cook J, Geraghty E, McMahon JM, Eeg-Olofsson O, Sadleir LG, Gill D, Ben-Zeev B, Lerman-Sagie T, Mackay M, Freeman JL, Andermann E, Pelakanos JT, Andrews I, Wallace G, Eichler EE, Berkovic SF, Scheffer IE. (2011) Rare copy number variants are an important cause of epileptic encephalopathies. *Ann Neurol* 70:974–985.

Mei D, Marini C, Novara F, Bernardina BD, Granata T, Fontana E, Parrini E, Ferrari AR, Murgia A, Zuffardi O, Guerrini R. (2010) Xp22.3 genomic deletions involving the CDKL5 gene in girls with early onset epileptic encephalopathy. *Epilepsia* 51:647–654.

Milh M, Villeneuve N, Chouchane M, Kaminska A, Laroche C, Barthez MA, Gitiaux C, Bartoli C, Borges-Correia A, Cacciagli P, Mignon-Ravix C, Cuberos H, Chabrol B, Villard L. (2011) Epileptic and nonepileptic features in patients with early onset epileptic encephalopathy and *STXBP1* mutations. *Epilepsia.* 52:1828–1834.

Mulley JC, Mefford HC. (2011) Epilepsy and the new cytogenetics. *Epilepsia* 52:423–432.

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

Nonoda Y, Saito Y, Nagai S, Sasaki M, Iwasaki T, Matsumoto N, Ishii M, Saitsu H. (2013) Progressive diffuse brain atrophy in West syndrome with marked hypomyelination due to *SPTAN1* gene mutation. *Brain Dev* 35:280–283.

Nord AS, Lee M, King MC, Walsh T. (2011) Accurate and exact CNV identification from targeted high-throughput sequence data. *BMC Genomics* 12:184.

Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, Bertoni A, Swerdlow HP, Gu Y. (2012) A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* 13:341.

Saitsu H, Kato M, Mizuguchi T, Hamada K, Osaka H, Tohyama J, Uruno K, Kumada S, Nishiyama K, Nishimura A, Okada I, Yoshimura Y, Hirai S, Kumada T, Hayasaka K, Fukuda A, Ogata K, Matsumoto N. (2008) De novo mutations in the gene encoding STXBP1 (MUNC18-1) cause early infantile epileptic encephalopathy. *Nat Genet* 40:782–788.

Saitsu H, Kato M, Okada I, Orii KE, Higuchi T, Hoshino H, Kubota M, Arai H, Tagawa T, Kimura S, Sudo A, Miyama S, Takami Y, Watanabe T, Nishimura A, Nishiyama K, Miyake N, Wada T, Osaka H, Kondo N, Hayasaka K, Matsumoto N. (2010a) *STXBP1* mutations in early infantile epileptic encephalopathy with suppression-burst pattern. *Epilepsia* 51:2397–2405.

Saitsu H, Tohyama J, Kumada T, Egawa K, Hamada K, Okada I, Mizuguchi T, Osaka H, Miyata R, Furukawa T, Haginoya K, Hoshino H, Goto T, Hachiya Y, Yamagata T, Saitoh S, Nagai T, Nishiyama K, Nishimura A, Miyake N, Komada M, Hayashi K, Hirai S, Ogata K, Kato M, Fukuda A, Matsumoto N. (2010b) Dominant-negative mutations in alpha-II spectrin cause West syndrome with severe cerebral hypomyelination, spastic quadriplegia, and developmental delay. *Am J Hum Genet* 86:881–891.

Saitsu H, Kato M, Shimono M, Senju A, Tanabe S, Kimura T, Nishiyama K, Yoneda Y, Kondo Y, Tsurusaki Y, Doi H, Miyake N, Hayasaka K, Matsumoto N. (2011) Association of genomic deletions in the STXBP1 gene with Ohtahara syndrome. *Clin Genet* 81:399–402.

Saitsu H, Kato M, Koide A, Goto T, Fujita T, Nishiyama K, Tsurusaki Y, Doi H, Miyake N, Hayasaka K, Matsumoto N. (2012a) Whole exome sequencing identifies *KCNQ2* mutations in Ohtahara syndrome. *Ann Neurol* 72:298–300.

Saitsu H, Kato M, Osaka H, Moriyama N, Horita H, Nishiyama K, Yoneda Y, Kondo Y, Tsurusaki Y, Doi H, Miyake N, Hayasaka K, Matsumoto