


# Severe neurocognitive and growth disorders due to variation in *THOC2*, an essential component of nuclear mRNA export machinery

Raman Kumar<sup>1</sup>  | Alison Gardner<sup>1</sup> | Claire C. Homan<sup>1</sup> | Evelyn Douglas<sup>2</sup> | Heather Mefford<sup>3</sup> | Dagmar Wiczorek<sup>4,5</sup> | Hermann-Josef Lüdecke<sup>4,5</sup> | Zornitza Stark<sup>6,7</sup> | Simon Sadedin<sup>6,8</sup> | The Broad CMG<sup>9</sup> | Catherine Bearce Nowak<sup>10</sup> | Jessica Douglas<sup>10</sup> | Gretchen Parsons<sup>11</sup> | Paul Mark<sup>11</sup> | Lourdes Loidi<sup>12</sup> | Gail E. Herman<sup>13</sup> | Theresa Mihalic Mosher<sup>13</sup> | Meredith K. Gillespie<sup>14</sup> | Lauren Brady<sup>15</sup> | Mark Tarnopolsky<sup>15</sup> | Irene Madrigal<sup>16,17</sup> | Jesús Eiris<sup>18</sup> | Laura Domènech Salgado<sup>19</sup> | Raquel Rabionet<sup>19</sup> | Tim M. Strom<sup>20</sup> | Naoko Ishihara<sup>21</sup> | Hidehito Inagaki<sup>22</sup> | Hiroki Kurahashi<sup>22</sup> | Tracy Dudding-Byth<sup>23,24</sup>  | Elizabeth E. Palmer<sup>23,25</sup>  | Michael Field<sup>23</sup> | Jozef Gecz<sup>1,26</sup> 

<sup>1</sup>Adelaide Medical School and the Robinson Research Institute, The University of Adelaide, Adelaide, Australia

<sup>2</sup>Genetics and Molecular Pathology, SA Pathology, Adelaide, Australia

<sup>3</sup>Division of Genetic Medicine, Department of Pediatrics, University of Washington & Seattle Children's Hospital, Seattle, Washington

<sup>4</sup>Heinrich-Heine-University, Medical Faculty, Institute of Human Genetics, Düsseldorf, Germany

<sup>5</sup>Institut für Humangenetik, Universitätsklinikum Essen, Universität Duisburg-Essen, Essen, Germany

<sup>6</sup>Murdoch Children's Research Institute, Melbourne, Australia

<sup>7</sup>Department of Pediatrics, University of Melbourne, Melbourne, Australia

<sup>8</sup>Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, Massachusetts

<sup>9</sup>Broad's Center for Mendelian Genomics, Cambridge, Massachusetts

<sup>10</sup>The Feingold Center for Children at the Department of Genetics and Genomics, Boston Children's Hospital, Boston, Massachusetts

<sup>11</sup>Spectrum Health Medical Genetics, Grand Rapids, Michigan

<sup>12</sup>Fundación Pública Galega de Medicina Xenómica, Santiago de Compostela, Spain

<sup>13</sup>Nationwide Children's Hospital and The Ohio State University, Columbus, Ohio

<sup>14</sup>Children's Hospital of Eastern Ontario Research Institute, University of Ottawa, Ottawa, Canada

<sup>15</sup>Department of Pediatrics, McMaster University Medical Centre, Hamilton, Canada

<sup>16</sup>Biochemistry and Molecular Genetics Department, Hospital Clínic, IDIBAPS, Barcelona, Spain

<sup>17</sup>Centre for Biomedical Research on Rare Diseases (ISCIII), Barcelona, Spain

<sup>18</sup>Unidad de Neurología Pediátrica, Departamento de Pediatría, Hospital Clínico Universitario de Santiago de Compostela, Santiago de Compostela, Spain

<sup>19</sup>Centre for Genomic Regulation (CRG), Universitat Pompeu Fabra and CIBERESP, Barcelona Institute for Science and Technology, Barcelona, Spain

<sup>20</sup>Institut für Humangenetik, Universitätsklinikum Essen, Universität Duisburg-Essen, Essen, Germany

<sup>21</sup>Department of Pediatrics, Fujita Health University School of Medicine, Aichi, Japan

<sup>22</sup>Division of Molecular Genetics, Institute for Comprehensive Medical Science, Fujita Health University, Aichi, Japan

<sup>23</sup>Genetics of Learning Disability Service, Hunter Genetics, Waratah, NSW, Australia

<sup>24</sup>University of Newcastle, Australia Grow-Up-Well Priority Research Center, Callaghan, Australia

<sup>25</sup>School of Women's and Children's Health, University of New South Wales, Randwick, NSW, Australia

<sup>26</sup>Healthy Mothers, Babies and Children, South Australian Health and Medical Research Institute, Adelaide, SA, Australia

**Correspondence**

Jozef Gecz, School of Medicine, The Robinson Research Institute, The University of Adelaide, North Adelaide, SA 5005, Australia.  
Email: jozef.gecz@adelaide.edu.au

**Funding information**

Contract grant sponsors: National Health and Medical Research Council Program (APP1091593, APP1041920); Channel 7 Children's Research Foundation; Fondo de Investigación Sanitaria (APP19PI10/01710); Spanish Ministry of Economy and Competitiveness (SEV-2016-0571); MINECO Severo Ochoa (SVP-2013-0680066); Genome Canada; National Heart, Lung and Blood Institute (UM1 HG008900); Canadian Institutes of Health Research (CIHR); Ontario Genomics Institute; Ontario Research Fund; Génome Québec; Children's Hospital of Eastern Ontario Foundation; National Human Genome Research Institute; National Eye Institute; Undiagnosed Diseases Program-Victoria (UDP-Vic); Murdoch Children's Research Institute, Melbourne, Australia.

Communicated by Christine Van Broeckhoven

**Abstract**

Highly conserved TREX-mediated mRNA export is emerging as a key pathway in neuronal development and differentiation. TREX subunit variants cause neurodevelopmental disorders (NDDs) by interfering with mRNA export from the cell nucleus to the cytoplasm. Previously we implicated four missense variants in the X-linked *THOC2* gene in intellectual disability (ID). We now report an additional six affected individuals from five unrelated families with two *de novo* and three maternally inherited pathogenic or likely pathogenic variants in *THOC2* extending the genotypic and phenotypic spectrum. These comprise three rare missense *THOC2* variants that affect evolutionarily conserved amino acid residues and reduce protein stability and two with canonical splice-site *THOC2* variants that result in C-terminally truncated *THOC2* proteins. We present detailed clinical assessment and functional studies on a *de novo* variant in a female with an epileptic encephalopathy and discuss an additional four families with rare variants in *THOC2* with supportive evidence for pathogenicity. Severe neurocognitive features, including movement and seizure disorders, were observed in this cohort. Taken together our data show that even subtle alterations to the canonical molecular pathways such as mRNA export, otherwise essential for cellular life, can be compatible with life, but lead to NDDs in humans.

**KEYWORDS**

mRNA export, partial loss-of-function variants, protein stability, *THOC2*, XLID

**1 | INTRODUCTION**

Intellectual disability (ID), characterized by substantial limitations in both intellectual functioning and adaptive behavior, affects 1%–3% of the population starting before the age of 18 years and has significant impact on individuals, families, and communities (Vissers, Gilissen, & Veltman, 2016). Individuals with ID are more likely than members of the general population to experience poor physical and mental health, have a lower life expectancy, experience inequalities accessing health care, and frequently have limited or no specific therapies for their core symptoms (Bittles et al., 2002; Hosking et al., 2016). Both genetic and environmental factors contribute to the development of ID (Milani, Ronzoni, & Esposito, 2015). Over 120 of the identified >800 ID genes are located on the X-chromosome (Chiurazzi & Pirozzi, 2016; Schwartz, 2015), and diagnosis of X-linked causes of ID remain critically important for accurate genetic counseling of families (Ropers & Hamel, 2005). Dramatic improvements in high-throughput DNA sequencing technologies and analyses software has led to identification of new ID genes and additional variants in the known ID genes (Dickinson et al., 2016; Vissers, et al., 2016). A systematic review of clinical data suggests that ID affected individuals frequently have comorbid neurological, psychiatric, and behavioral disorders (Oeseburg, Dijkstra, Groothoff, Reijneveld, & Jansen, 2011; Vissers et al., 2016), and disease variants in different parts of a gene can lead to a broad range of complex neurocognitive disorders (Palmer et al., 2017; Zhu, Need, Petrovski, & Goldstein, 2014). This complexity contributes to heterogeneity in clinical symptoms and indistinct boundaries between syndromic and nonsyndromic forms of neurodevelopmental disorder (NDD).

In 2015, we reported genetic, molecular, and protein structural data on four missense variants in an X-linked essential gene *THOC2* (MIM# 300957; NM\_001081550.1; c.937C>T (p.Leu313Phe), c.1313T>C (p.Leu438Pro), c.2399T>C (p.Ile800Thr), and c.3034T>C

(p.Ser1012Pro), RNA not analyzed) (Kumar et al., 2015). The affected individuals had a syndromic NDD, characterized by borderline to severe ID, speech delay, short stature, and adult onset truncal obesity (Kumar et al., 2015). *THOC2* encodes for the *THOC2* protein—the largest subunit of the highly conserved TREX (Transcription-Export) mRNA export complex essential for exporting mRNA from the cell nucleus to the cytoplasm (Heath, Viphakone, & Wilson, 2016). The TREX complex is composed of a THO sub-complex (*THOC1*, *THOC2*, *THOC3*, *THOC5*, *THOC6*, and *THOC7*) and accessory proteins (*UAP56*, *UIF*, *Aly*, *CIP29*, *PDIP3*, *ZC11A*, *SRRT*, *Chtop*) (Heath et al., 2016). The TREX complex, besides its canonical role in mRNA export in the mammalian cells, has been shown to play critical roles in gene expression, 3' mRNA processing, stress responses, mitotic progression, and genome stability as well as developmental processes such as pluripotency maintenance and hematopoiesis (Yamazaki et al., 2010). We and others have recently demonstrated that subtle perturbations in mRNA export by gene variants or preferential cytoplasmic aggregation can lead to NDDs (Beaulieu et al., 2013; Coe et al., 2014; Kumar et al., 2015), neurodegeneration (Woerner et al., 2016), or cancer (Chinnam et al., 2014; Hautbergue, 2017; Liu et al., 2015; Viphakone et al., 2015). These alterations can have tissue-specific effects as TREX subunits are shown to have tissue-specific roles; for example, mouse *Thoc5* and *Thoc1* deficiency interferes with the maintenance of hematopoiesis (Guria et al., 2011; Mancini et al., 2010) and testis development (Wang et al., 2009). Taken together, altered TREX function can have diverse molecular and cellular consequences resulting in a range of diseases. Here, we present detailed information on the clinical presentations and functional investigations on an additional eight missense and two splice *THOC2* variants. These data reaffirm and extend our previous findings that *THOC2* variation plays a role in complex neurodevelopmental conditions with the core clinical presentation of ID.

## 2 | MATERIALS AND METHODS

### 2.1 | Molecular and cellular studies

RNA extraction, RT-qPCR (primers listed in Supp. Table S3), cycloheximide chase, and THOC2 immunofluorescence staining were performed as reported previously (Kumar et al., 2015). Molecular studies on the THOC2 exon35:c.4450-2A>G variant were performed using blood DNA and skin fibroblasts derived from the affected individual and his heterozygous carrier mother. Genomic DNA or cDNA (generated by reverse transcribing the fibroblast total RNAs using Superscript III reverse transcriptase; Life Technologies, Scoresby, Victoria, Australia) was amplified with KAPA HiFi PCR Kit (Kapa Biosystems, Mulgrave, Victoria, Australia) using hTHOC2-4326F/hTHOC2-4519-R (Supp. Table S3) at 95°C for 3 min, 35 cycles of 98°C–15 sec, 59°C–15 sec, 72°C–30 sec, incubation at 72°C for 10 min, gel purified (MinElute Gel Extraction kit; Qiagen, Chadstone Centre, Victoria, Australia) and Sanger sequenced using the same primers. For the THOC2 exon28:c.3503+4A>C, blood gDNA from unaffected father, carrier mother, and affected son was amplified with TaKaRa ExTaq (Clayton, Victoria, Australia) using THOC2-F/THOC2-R primers (Supp. Table S3) at 94°C for 2 min, 40 cycles of 94°C–30 sec, 60°C–30 sec, 72°C–30 sec, incubation at 72°C for 5 min. The cDNA was generated by reverse transcribing the white blood cell RNAs using Superscript III reverse transcriptase (Life Technologies) and amplified with TaKaRa ExTaq using THOC2-ex27F/THOC2-ex30R (Supp. Table S3) at 94°C for 2 min, 28 cycles of 94°C–30 sec, 60°C–30 sec, 72°C–30 sec, incubation at 72°C for 5 min. The amplified products were analyzed by Sanger sequencing.

### 2.2 | Generation of THOC2 variant expression constructs

Generation of the wild-type Myc-tagged human THOC2 expression plasmid was reported earlier (Kumar et al., 2015). Briefly, the THOC2 variants were introduced into the existing pCMV-Myc-THOC2 expression construct by overlap PCR method using the primers listed in Supp. Table S3. The variant plasmid sequences were confirmed by Sanger sequencing. Details relating generation of the THOC2 variant expression constructs are available on request.

### 2.3 | Transient expression and Western blotting

For transient expression experiments, HEK293T and HeLa cells were transfected with expression constructs (400 ng pCMV-Myc-THOC2 plasmid and 400 ng pEGFP-C1 plasmid/transfection for stability and cycloheximide assays and 4  $\mu$ g/transfection for immunofluorescence staining, IF) using Lipofectamine 3000 reagent according to manufacturer's protocol (Life Technologies). Twenty-four hours post-transfection, cells were either fixed with 4% formaldehyde for IF or collected and lysed in buffer containing 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% Triton-X 100, 1 mM EDTA, 50 mM NaF, 1 $\times$  Protease inhibitor/no EDTA and 0.1 mM Na<sub>3</sub>VO<sub>4</sub> for Western blot assay as reported previously (Kumar et al., 2015).

### 2.4 | In silico pathogenicity prediction

We used CADD v1.3 (includes PhyloP, GERP++ & PolyPhen2) (Kircher et al., 2014), Provean (Choi & Chan, 2015), and ACMG (Richards et al., 2015) on-line tools for *in silico* prediction of the pathogenicity of different variants (Table 1 and Supp. Table S1).

## 3 | RESULTS

### 3.1 | Identification of THOC2 variants

We previously implicated four missense THOC2 variants in 25 individuals with ID and a range of other clinical features (Table 1 and Supp. Table S1 and 2) (Kumar et al., 2015). We identified an additional five THOC2 variants (three missense; *de novo* c.2087C>T (p.Thr696Ile), *de novo* c.2138G>A (p.Gly713Asp), maternally inherited c.3559C>T (p.His1187Tyr), and two splicing-defective; maternally inherited chrX:122747561 exon35:c.4450-2A>G and chrX:122757634 exon28:c.3503+4A>C; GenBank: NM\_001081550.1) variants in a further six affected individuals, including one pair of monozygotic twins (Table 1 and Figures 1 and 2). Whole exome (WES) or whole genome sequencing (WGS) of probands and parents was used to identify the variants that were confirmed by Sanger sequencing of the PCR amplified variant-carrying region of genomic DNA of the parents and affected individuals. The previously unreported THOC2 variants affect amino acids that are highly conserved (Supp. Figure S1), are absent in gnomAD database and are predicted to be pathogenic based on a number of *in silico* analyses tools (Table 1). We included in our study a *de novo* missense p.Tyr517Cys variant in a female with moderate-severe ID, speech problems, epileptic encephalopathy, cortical visual impairment, and gait disturbances identified using WES as part of the Epi4K Consortium & Epilepsy Phenome/Genome Project (Epi et al., 2013) (Table 1). We have also collected further, rare and potentially pathogenic variants through international collaboration (Table 1, Supp. Table S1 and 2, and Figures 1 and 3; see Supplementary information for methods used for identifying the variants) and performed functional testing on several of these. The following three variants: c.229C>T (p.Arg77Cys), c.3034T>C (p.Ser1012Pro), and c.3781A>C (p.Asn1261His) showed no clear evidence of altered stability of variant THOC2 proteins in our cell-based assay. The reported variants have been submitted to ClinVar database (<https://www.ncbi.nlm.nih.gov/clinvar/accession> numbers SCV00680065-SCV00680074).

### 3.2 | Clinical presentations

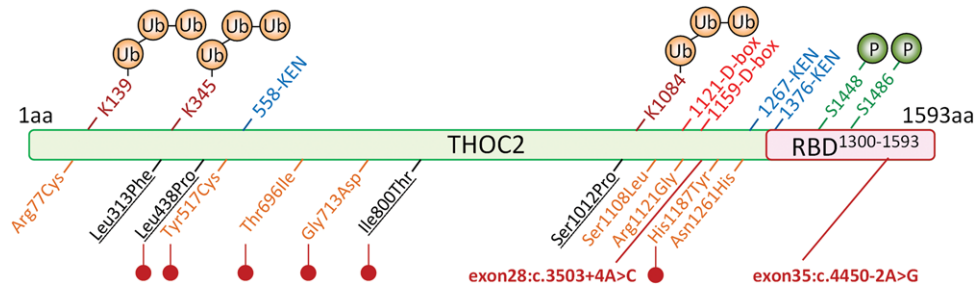
The clinical features of the five previously unreported affected individuals with (likely) pathogenic THOC2 variants, aged between 3 and 12 years, and the 10 year old female with *de novo* p.Tyr517Cys variant are summarized in Table 2 and photographs, when available, are shown in Figure 2. Detailed clinical information is available in the supplementary data. Each individual clinical center used local diagnostic criteria for determining degree of ID and diagnoses of comorbidities. ID was universal and at least moderate in severity: two out of seven were nonambulatory and three out of seven nonverbal. Behavioral problems

**TABLE 1** Detailed description of the THOC2 variants with supporting molecular evidence

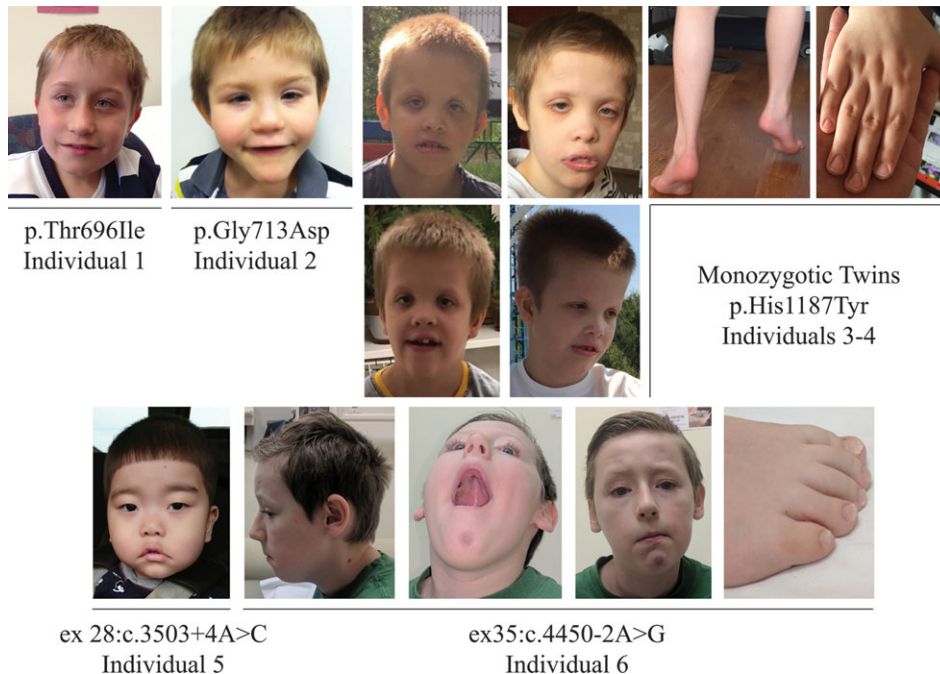
Individual	From	Method of identification	Position hg19	NM	Mode of inheritance	CADD	Provean score	Provean prediction	GERP++	PhyloP	gnomAD frequency	Polyphen2 <sup>a</sup>	Variant plasmid tested	Reduced Protein Stability	ACMG pathogenicity classification <sup>b</sup>
Missense	Kumar et al. (2015) AJHG	X-Chr Exome sequencing	122799566	c.1313T>C;p.Leu438Pro	Maternal inheritance	28.1	-6.08	Deleterious	5.7	1.902	Absent	D	YES	YES	LP
1	Australia	Trio WES	122767853	c.2087C>T;p.Thr696Ile	De novo	27.4	-5.47	Deleterious	5.03	0.963	Absent	D	YES	YES	DP
2	USA	WES	122766890	c.2138G>A;p.Gly713Asp	De novo	31	-4.69	Deleterious	5.73	2.412	Absent	D	YES	YES	DP
3-4	Canada/ Germany/ Russia Identical twins	WES	122757079	c.3559C>T;p.His1187Tyr	Maternal inheritance; Mother skewed (99.9:0.1%)	23.1	-5.07	Deleterious	6.07	2.571	Absent	P	YES	YES	LP
5	Japan	Trio WES	122757634	Exon28:c.3503+4A>C	Maternal inheritance; Mother skewed 98.2%	10.8	N/A	N/A	5.57	1.86	Absent	N/A	N/A	ND	LP
6	Canada	WES	122747561	Exon35:c.4450-2A>G	Maternal inheritance; Mother skewed 94.6%	23.7	N/A	N/A	5.25	1.735	Absent	N/A	Fibroblasts of the affected male proband and carrier mother	NO	LP
Missense 7	USA Epi4K Consortium & Epilepsy Phenome/Genome Project; Nature 501:217-221, 2013	WES	122778639	c.1550A>G;p.Tyr517Cys	De novo	26.6	-7.87	Deleterious	5.84	1.955	Absent	D	YES	YES	DP

<sup>a</sup>D, probably damaging; P, possibly damaging; N/A, not applicable; ND, not determined.<sup>b</sup>DP, de novo pathogenic; LP, likely pathogenic.





**FIGURE 1** Location of variant amino acids and structural features in THOC2 protein. Ubiquitinated (K139 and K1084 (Kim et al., 2011) and K345 (Lopitz-Otsoa et al., 2012; Wagner et al., 2011)), phosphorylated (S1448 and 1486 (Olsen et al., 2006)) amino acid residues, potential RNA binding domain (RBD) and destruction box (D-box) and KEN box sequences that interact with the Anaphase Promoting Complex/Cyclosome (APC/C) for protein ubiquitination and subsequent destruction by the proteasome (Morgan, 2013) are shown. Unreported (orange) and published (black: (Kumar et al., 2015)) missense variants effecting THOC2 protein stability are marked with red lollipops. The positions of two splice variants are shown in red



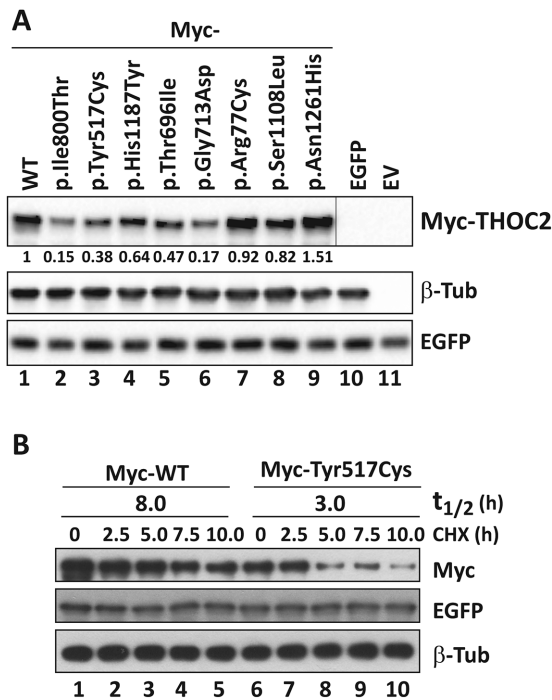
**FIGURE 2** Front and side facial views of the affected individuals with THOC2 variants

were reported in four individuals, with one meeting diagnostic criteria for autism spectrum disorder. Additional neurological features were common. Four of the seven had infantile hypotonia and two of the seven had tremor. The monozygotic twins (individuals 3 and 4) had a tendency to toe walking, which was considered behavioral as it was not associated with neurological signs of lower limb spasticity. Confirmed seizure disorder was only present in the affected female (individual 7) but suspected in individual 2. Neuroradiological studies were performed in five individuals: this was within normal limits for three individuals, with neuroanatomical differences reported in two. Individual 2 had complex neuroanatomical findings (see supplementary clinical description and Supp. Figure S2A) including changes in cortical gyral morphology, which in the inferior temporal lobes appeared finely nodular, as well as hypoplasia of the corpus callosum and reduced brainstem volume and individual 5 had mild dilatation of the lateral ventricles, mildly delayed myelination and an abnormal white matter lesion in the periventricular area close to the anterior horn (Supp. Figure S2B). Growth abnormalities were common

including low birth weight (three out of seven), microcephaly (two out of seven), and short stature (two out of seven). Facial features are shown in Figure 2. Appropriate consent for reporting variants, clinical data and photographs of the affected individuals was obtained from their parents or legal guardians. The research has been approved by the Women's and Children's Health Network Human Research Ethics Committee in Adelaide, Australia.

### 3.3 | THOC2 variant protein localization and stability

Without access to affected individuals' derived cells, we generated Myc-tagged THOC2 missense variant expression constructs to determine protein stability and localization. The THOC2 protein stability was determined in HEK293T cells and localization in both the HEK293T and HeLa cells. Total protein lysates of HEK293T cells ectopically-expressing the wild-type or variant Myc-THOC2 proteins were Western blotted for THOC2, EGFP, and  $\beta$ -Tubulin. We used HEK293T cells expressing Myc-p.Ile800Thr THOC2 as a control for



**FIGURE 3** Functional testing of THOC2 missense variants. **A:** THOC2 variant protein stability is reduced in HEK293T cells. pCMV-Myc-THOC2 wild-type or variant expression constructs and pEGFP-C1 plasmid (transfection control) were transfected into HEK293T cells. Total protein lysates of cells 24 hr post-transfection were analyzed by Western blotting with mouse anti-Myc (clone 9E10; Sigma), mouse anti-EGFP (clones 7.1 and 13.1; Roche) and rabbit anti- $\beta$ -tubulin (loading control; Abcam) antibodies. pCMV-Myc-THOC2 p.Ile800Thr construct expressing the p.Ile800Thr protein shown to have reduced stability was used as a control (Kumar et al., 2015). Western blot signals were quantified using ImageJ software. Averages of the Myc-THOC2 proteins normalized to the housekeeping  $\beta$ -tubulin signal from two independent runs are shown. **B:** Myc-p.Tyr517Cys THOC2 protein half-life is substantially reduced in HEK293T cells. pCMV-Myc-THOC2 or pCMV-Myc-THOC2-p.Tyr517Cys expression constructs and pEGFP-C1 plasmid (transfection control) were transfected into HEK293T cells. Next day the cells were cultured in the presence of 100  $\mu$ g/ml translation inhibitor cycloheximide and harvested at the time points shown. Total protein lysates were analyzed by Western blotting with mouse anti-Myc, mouse anti-EGFP, and rabbit anti- $\beta$ -tubulin (loading control) antibodies

protein stability assay as this variant is shown to cause reduced protein stability (Kumar et al., 2015). The results showed reduced stability of p.Tyr517Cys, p.Thr696Ile, p.Gly713Asp, and p.His1187Tyr THOC2 compared with the wild-type protein (Figure 3A). Presence of comparable levels of EGFP in the cells transfected with different expression constructs indicated that the reduced levels of THOC2 protein were not due to difference in transfection efficiency (Figure 3A). We also determined the turnover rate of Myc-p.Tyr517Cys THOC2 protein by cycloheximide chase. For this assay, the HEK293T cells transfected with pCMV-Myc-WT or pCMV-Myc-p.Tyr517Cys THOC2 and pEGFP-C1 transfection control plasmids were cultured in presence of translation inhibitor cycloheximide for different durations and Western blotted for THOC2, EGFP, and  $\beta$ -Tubulin. The results showed that p.Tyr517Cys THOC2 turnover rate was 3 hr compared with 8 hr

for the wild-type protein (Figure 3B). THOC2 variant proteins, similar to the wild-type, were mainly localized to the nucleus in both the HEK293T and HeLa cells (Supp. Figure S3).

### 3.4 | THOC2 splice variant: exon35:c.4450-2A>G, p.Arg1483fs52\*

Sanger sequencing of amplified target region from affected son and mother's blood genomic DNA showed that the affected boy inherited chrX:122747561 exon35:c.4450-2A>G variant from his unaffected heterozygous carrier mother (Figure 4C). A -2 A>G change in the intron-exon splicing site boundary (acceptor AG) is predicted to abolish splicing (Ohno, Takeda, & Masuda, 2018). To validate this possibility, we generated skin fibroblast cultures from the heterozygous carrier mother and the affected son. We PCR amplified their fibroblast cDNAs using primers with binding sequences located within exon 34 and 35. Amplification of a 194 bp DNA fragment from the mother indicated normal splicing but a 537 bp product from the affected son indicated retention of the intron located between these exons (Figure 4A and B). We confirmed this result by Sanger sequencing of the PCR products generated from genomic DNA that showed presence of A/G nucleotides in the carrier mother but only G (A>G) nucleotide in the affected son (Figure 4C). The cDNA sequence showed presence of normally-spliced mRNA in the mother but retention of intronic sequence upstream of the exon 35 in the affected son indicating defective splicing due to presence of -2 G variant at the intron-exon 35 junction sequence (Figure 4C). The presence of normally spliced mRNA in the unaffected mother is consistent with X-inactivation (94% skewing) of the variant allele in her fibroblasts. We predicted that a retention of intron between exon 34–35 in the affected fibroblasts would result in loss of 110 C-terminal amino acids of the 1,593 wild-type THOC2 protein (that is, 1,483 amino acids); however, overall the variant protein would be 58 amino acid smaller as it would now be a 1,535 amino acid protein comprised of 1,483 amino acids of wild-type THOC2 and 52 translated from intronic sequence in the defective mRNA (Figure 4A). Consistent with our prediction, the Western blot data showed presence of a slightly smaller THOC2 protein band in the affected son's fibroblasts than his unaffected mother. Many independent Western blot runs showed presence of two closely located THOC2 bands—similar to the fruit fly THO2 (Rehwinkel et al., 2004)—in the unaffected mother but a single highly intense band in the affected son's fibroblasts (Figure 4D). The observed difference in levels of THOC2 protein was post-translational as we found comparable amounts of THOC2 mRNA, as assayed by real time RT-qPCR, in the mother and son (Figure. 4E). Finally, we observed no difference in THOC2 localization in fibroblasts of the affected son and his unaffected mother (Figure 4F).

### 3.5 | THOC2 splice variant: exon28:c.3503+4A>C, p.Gly1168fs7\*, and normal 1,593 aa protein

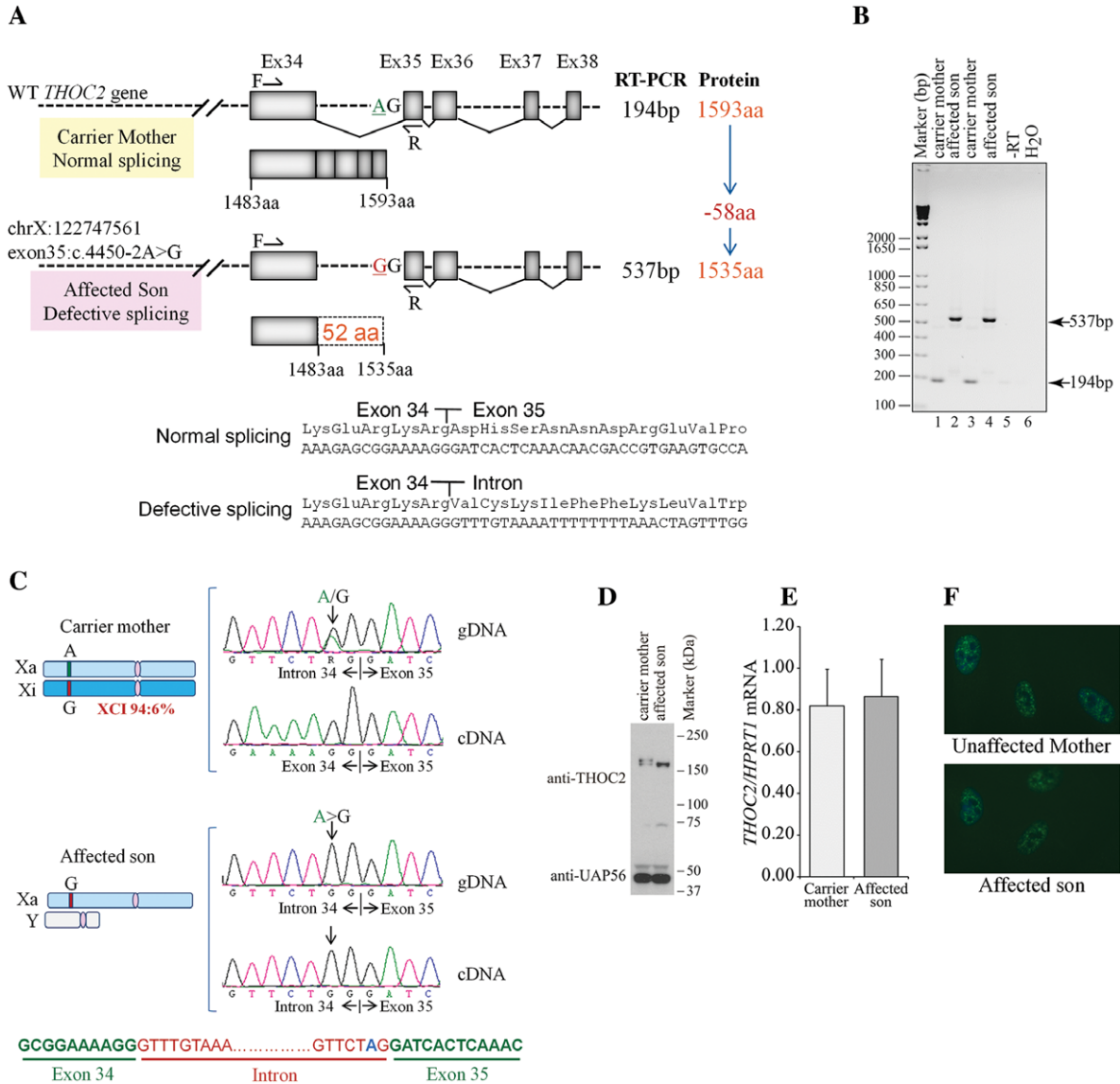
For the second splice variant chrX:122757634 exon28:c.3503+4A>C, molecular studies were performed on the white blood cells of the unaffected father, carrier mother, and the male proband. Sanger sequencing of target region amplified from the unaffected father and

**TABLE 2** Summary of clinical data of *THOC2* variants with supporting molecular evidence

Individual	(Likely) pathogenic						
	1	2	3	4	5	6	7
<b>Variant details</b>	c.2087C>T: p.Thr696Ile	c.2138G>A: p.Gly713Asp	c.3559C>T: p.His1187Tyr Twin 1	c.3559C>T: p.His1187Tyr Twin 2	Exon28: c.3503+4A>C	Exon35:c.4450- 2A>G	c.1550A>G: p.Tyr517Cys
<b>Gender</b>	Male	Male	Male	Male	Male	Male	Female
<b>Age (years)</b>	12	5	7	7	3	10	10
<b>Perinatal features</b>							
Gestation (weeks)	36	37	37	37	37	41	NA
Low birth weight (<2.5kg)	Yes	No	Yes	Yes	No	No	NA
Birth weight (g)	2,000	2,650	1,990	2,420	3,018	4,365	NA
<b>Neurologic features</b>							
Intellectual disability	Severe	Mod+	Mod+	Mod+	Severe	Severe	Profound
Speech delay	Yes, single words, signs	Yes, nonverbal	Yes	Yes	Yes, nonverbal	Yes, nasal dysarthria	Yes, non-verbal
Hypotonia	No	Yes	NA	NA	Yes	Yes, central hypotonia	Yes
Spasticity	No	No	No	No	No	Yes- appendicular spasticity	No
Hyperkinesia	No	No	Yes	Yes	Yes	No	No
Tremor	No	Yes, intermittent	No	No	No	Yes	No
Epilepsy	No	Suspected	No	No	No	No	Yes, epileptic encephalopathy
Gait disturbances	No	Yes, gait/balance problems	Yes, toe walking <sup>a</sup>	Yes, toe walking <sup>a</sup>	Non ambulatory	Yes, ataxia/broad based gait	Non ambulatory
Behavior problems	No	Yes	Yes	Yes	Yes, ASD	No	NR
Anxiety	No	No	No	No	No	No	NR
Depression	No	No	No	No	No	No	NR
Brain MRI/CT	MRI normal	Thin corpus callosum, low brainstem volume, variability in gyral pattern.	ND	ND	Ventricular dilatation, delayed myelination, periventricular white matter lesion	MRI within normal limits	MRI normal
<b>Growth parameters</b>							
Microcephaly ( $\leq 3\%$ )	Yes	Yes, <1%	No	No	No	Yes, 2%	No, 5%
Short stature ( $\leq 3\%$ )	Yes	Yes	No	No	No	No	No
Overweight (BMI $\geq 25$ )	No	No	No	No	No	No	No
Broad high forehead	Yes	Yes	Yes	Yes	No	No	NR
<b>Other features</b>		Mild joint laxity, subluxed hips, disordered sleep, feeding difficulties (g-tube dependency), laryngomalacia, micrognathia, abnormal palmar creases	Noonan facies, pes planus, hypospadias	Noonan facies, pes planus, hypospadias		Clinodactyly, nystagmus, abnormality soft palate	Cortical visual impairment

**Abbreviations:** %, centile; ASD, autism spectrum disorder; CT, computerized tomography scan; g-tube, gastrostomy tube; mod+, at least moderate severity; MRI, magnetic resonance imaging; ND, not done; NA, not available; NR, not reported; NICU, neonatal intensive care unit; VOUS, variant of uncertain significance.

<sup>a</sup>Toe walking in absence of neurological signs of lower limb spasticity, therefore considered a behavioral manifestation.

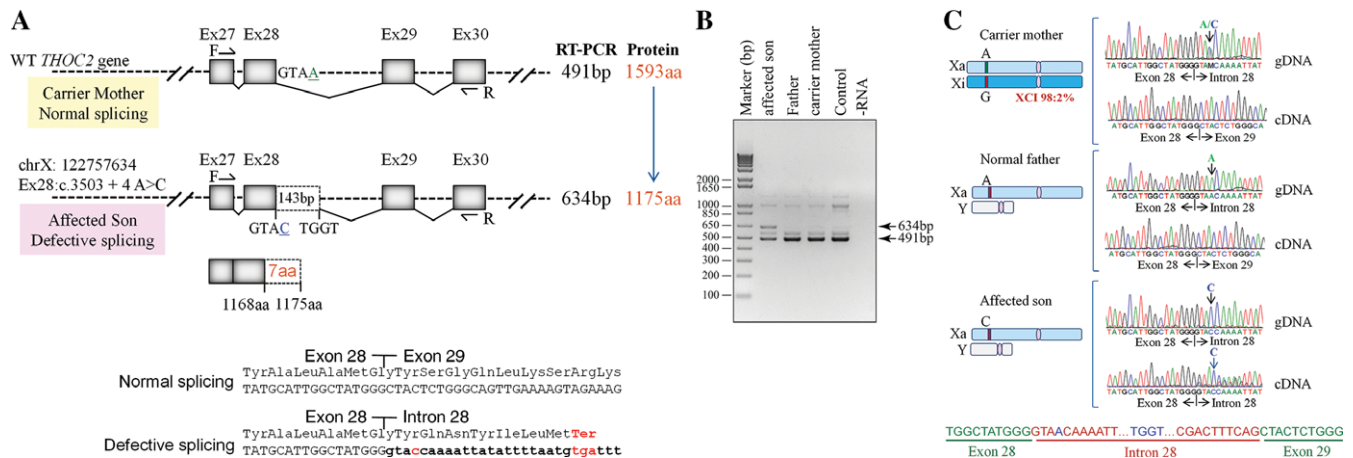


**FIGURE 4** Exon35:c.4450-2A>G variant abolishes splicing of intron between exons 34–35. **A:** Part of the *THOC2* gene showing location of the A/G nucleotide in the heterozygous carrier mother and A>G splice variant in the affected son. The C-terminal part of the 1,593 amino acid wild-type and 1,535 amino acid (that contains 1,483 normal and 52 amino acids coded by the unspliced intron) *THOC2* protein in the affected boy are also shown. **B:** Gel showing a 194 bp RT-PCR product from the normally spliced mRNA of the heterozygous carrier mother and a 537 bp product from defective splicing of mRNA causing retention of an intron between exon 34–35 in the affected son. RT-PCR products from total RNA isolated from passage 3 (lanes 1–2) and 5 (lanes 3–4) fibroblasts. Location of the forward and reverse primers within exons 34 and 35 is shown. **C:** Sanger sequencing chromatograms of PCR products amplified from genomic and cDNA of the affected son and his heterozygous carrier mother using primers located within exon 34 and 35. Genomic DNA around the Exon34-Intron-Exon35 region is shown. **D:** Western blot showing *THOC2* protein in the affected son and his carrier mother's skin fibroblasts. TREX subunit UAP56 was used as a loading control. **E:** RT-qPCR showing levels of the *THOC2* mRNA in the affected son and his carrier mother's skin fibroblasts. **F:** Immunofluorescence detection of *THOC2* in skin fibroblasts of the unaffected mother and affected son

mother, and affected son's genomic DNA showed that the affected son inherited the A>C change from his unaffected carrier mother who had A/C nucleotides at this position (Figure 5). The intronic nucleotide change A>C at +4 position of the 5' exon-intron donor splicing site sequence is predicted to cause aberrant splicing ([https://www.med.nagoya-u.ac.jp/neurogenetics/SD\\_Score/sd\\_score.html](https://www.med.nagoya-u.ac.jp/neurogenetics/SD_Score/sd_score.html)). To confirm this possibility, we amplified cDNA generated by reverse transcribing blood RNA of the father, mother, and the affected son using primers located within exon 27 and exon 30 (Figure 5A and Supp.

Table S3). Interestingly, whereas a 491 bp PCR product was observed in highly skewed carrier mother (98:2%) and normal father, 491 and 634 bp PCR products were detected in the affected son. A 491 bp amplified product indicated normal splicing in the mother and father, and 491 bp and 634 bp bands suggested partially defective splicing in the affected son. Amplification of a 634 bp instead of a 994 bp fragment that would have resulted from a complete retention of intron between exon 28–29 indicated aberrant splicing event in the affected son (Figure 5). Sanger sequencing of 491 and 634 bp PCR products from





**FIGURE 5** Exon28:c.3503+4A>C variant causes aberrant splicing of intron between exons 28–29. **A:** Part of the *THOC2* gene showing location of the A/C nucleotide in the heterozygous carrier mother and A>C splice variant in the affected son. The C-terminal part of the 1,593 amino acid wild-type and 1,175 amino acid (that contains 1,168 normal and 7 amino acids coded by the unspliced intron) *THOC2* protein in the affected son are also shown. **B:** Gel showing a 491 bp RT-PCR product from the normally-spliced heterozygous carrier mother and unaffected father, and 491 and 634 bp (retaining 143 bp of the 503 bp intron between exons 28–29) RT-PCR products derived from the normally and aberrantly spliced mRNAs, respectively, in the affected son. Location of the forward and reverse primers within exons 27 and 30 is shown. **C:** Sanger sequencing chromatograms of PCR products amplified using primers located within exons 27 and 30 from genomic and cDNA of unaffected father and mother and the affected son

the mother, father and son confirmed normal splicing in the mother and father and aberrant splicing in the affected son. The sequence showed retention of a 143 bp instead of complete 503 bp fragment due to activation of a cryptic splice site within the intron between exon 28–29 in the son (Figure 5). Retention of 143 bps from intron between exon 28–29 in the mRNA is predicted to result in a 1175 amino acid truncated *THOC2* protein containing 1,168 wild-type amino acids and 7 amino acids from the translation of the intronic sequence retained in the defective mRNA. This aberrant product would be present in addition to the wild-type 1,593 amino acid protein from the normally spliced mRNA in the affected son.

## 4 | DISCUSSION

Here, we present detailed clinical information, and molecular and functional studies, on five previously unreported *THOC2* variants in six affected males (two *de novo* variants and one maternally inherited variant in monozygotic twins) and on one affected female with a previously reported *de novo* p.Tyr517Cys variant. We present evidence that extends the genotypic spectrum beyond the four *THOC2* missense variants that we reported previously (Kumar et al., 2015) by including two intronic variants that affect splicing, and four missense variants that affect protein stability in a cell-based assay system. According to ACMG criteria, they were classified as pathogenic or likely pathogenic (Table 1) (Richards et al., 2015). These findings, along with the four missense variants reported earlier (Kumar et al., 2015), add to the existing evidence that alterations in essential mRNA export pathway cause NDDs (Amos et al., 2017; Beaulieu et al., 2013; Kumar et al., 2015).

We confirm that the core clinical feature of *THOC2*-related disorder in hemizygous males is ID, with several individuals having additional features including behavioral disorders, hypotonia, gait disturbance,

tremor, low birth weight, short stature, microcephaly, and variable neuroimaging findings. Although the range of neurodevelopmental features is similar, our original cohort contained males with ID in the mild or borderline range of intellectual functioning (Kumar et al., 2015), whereas all individuals in this cohort have ID which is at least in the moderate range. Individuals 2 and 6 had neurological signs that could be consistent with cerebellar dysfunction including tremor and a broad-based gait for individual 2 and nystagmus, tremor, and an ataxic broad-based gait for individual 6, in the absence of significant cerebellar abnormalities on MRI. This is interesting given the female patient with knockdown of *THOC2* function due to a *de novo* X;8 translocation that created a *PTK2-THOC2* fusion had congenital cerebellar hypoplasia and prominent cerebellar signs with mild ID (Di Gregorio et al., 2013). We used computerized face-matching technology to specifically evaluate the cohort to assess if a characteristic facial gestalt was evident across individuals with pathogenic or likely pathogenic variants across our original and this expanded clinical cohort (Supp. Figure S4) (Dudding-Byth et al., 2017). Although a clearly recognizable facial gestalt was not obvious, there are some similarities. The facial gestalt spectrum associated with *THOC2* pathogenic variants will continue to emerge as more individuals are reported.

As was the case in our original cohort, heterozygous mothers were clinically unaffected, and, where available, X-chromosome inactivation (XCI) was highly skewed (Table 1). In contrast individual 7, with a *de novo* missense variant (p.Tyr517Cys) is a female with a particularly severe neurocognitive presentation. This is consistent with other reported severely affected females with *de novo* variants in X linked genes (de Lange et al., 2016; Palmer et al., 2016; Snijders Blok et al., 2015; Zweier et al., 2014). Unfortunately, we did not have access to individual 7's genomic DNA to test XCI status.

A range of protein–protein interactions are required for mRNA export (Chi et al., 2013). Proteins with altered stability (Hirayama et al.,



2008), localization (Beaulieu et al., 2013) (e.g., THOC6 p.Gly46Arg implicated in syndromic ID), or interaction (Chi et al., 2013) can impact mRNA export and consequently disrupt normal cell function. We did not observe mislocalization of the THOC2 variant proteins in cultured cells and did not test alterations in their interaction with the other known or unknown TREX proteins. However, reduced levels of a number of new (p.Tyr517Cys, p.His1187Tyr, p.Thr696Ile, p.Gly713Asp) and published (p.Leu438Pro, p.Ile800Thr; Kumar et al., 2015) missense THOC2 variant proteins are due to impaired protein stability or reduced levels of normal mRNA due to aberrant splicing (exon28:c.3503+4A>C). We also noted increased stability of p.Asn1261His THOC2 protein. We and others have shown that THOC2 controls TREX function by maintaining the stability of THOC1, 3, 5, and 7 subunits (Chi et al., 2013; Kumar et al., 2015). Reduced levels of THOC2 missense variant proteins are most likely due to enhanced proteasome-mediated degradation as THOC2 is ubiquitinated (Lopitz-Otsoa et al., 2012). THOC2 depletion has been reported to have different consequences in diverse organisms. For example, shRNA-mediated Thoc2 knockdown leads to significant increase in length of neurites in cultured rat primary hippocampal neurons (Di Gregorio et al., 2013) although effects on neurons with persistently reduced THOC2 variant proteins in the affected individuals may be different and *Caenorhabditis elegans thoc2* knockouts, that are completely immobile, slow-growing, sterile, have functional defects in specific sensory neurons and die prematurely (Di Gregorio et al., 2013). *Danio rerio Thoc2* is essential for embryonic development (Amsterdam et al., 2004) and in *Drosophila melanogaster* S2 cells *Thoc2* knockdown inhibits mRNA export and cell proliferation (Rehwinkel et al., 2004). THOC2 depletion also results in chromosome alignment, mitotic progression, and genomic stability in human HeLa cells (Yamazaki et al., 2010). Finally, *Thoc2* and *Thoc5* knockdown experiments have shown their role in regulation of embryonic stem cell (ESC) self-renewal (Wang et al., 2013).

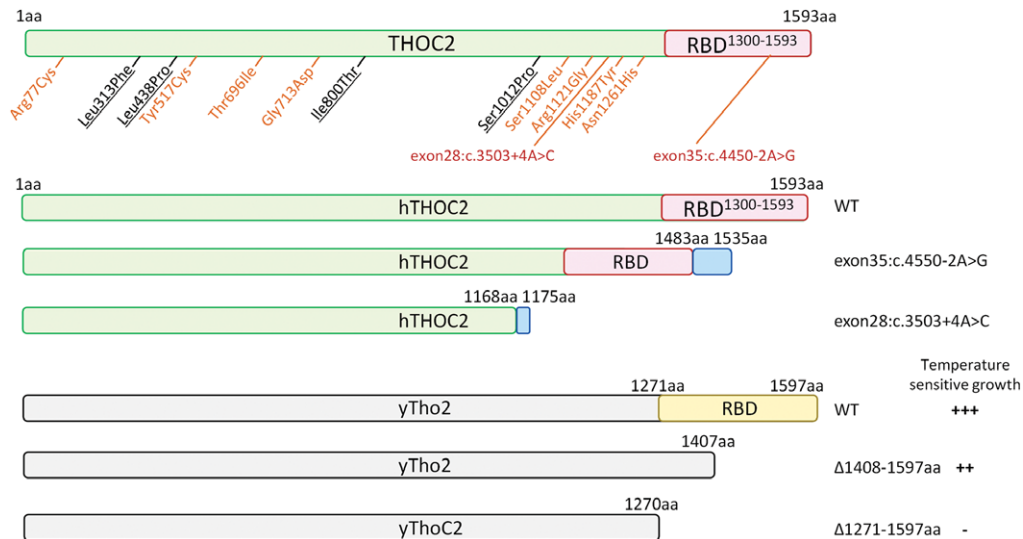
Both the affected individuals carrying the splice-variants presented with severe neurocognitive features. The exon35:c.4450-2A>G and exon28:c.3503+4A>C THOC2 splice variants present interesting biological scenarios; the former resulting in a 1,535 amino acid truncated protein that is present at higher level and the latter with both normal (albeit potentially much reduced) and a 1,175 amino acid truncated THOC2 protein. We postulate that the clinical outcomes in the exon35:c.4450-2A>G individual are caused by partial loss of function due to loss of 110 amino acid C-terminal region and accumulation of the truncated THOC2 protein. However, pathogenicity in exon28:c.3503+4A>C affected individual is most likely caused by reduced levels of normal and potential dominant-negative effects of the C-terminally truncated THOC2 protein. That reduced THOC2 protein levels are associated with ID and other clinical symptoms is emerging as a frequent theme; for example, due to reduced THOC2 protein stability caused by missense variants (see above) or aberrant splicing. Indeed reduced THOC2 levels are shown to destabilize the TREX complex in humans (Chi et al., 2013; Kumar et al., 2015) and removal of any THO subunit causes destabilization of other TREX components in yeast (Pena et al., 2012).

Systematic functional analysis of the Tho2 C-terminal RNA binding region in yeast provides interesting explanation as to how the

truncated THOC2 protein can perturb normal mRNA export function in human cells (Pena et al., 2012) (Figure 6). The data showed that whereas  $\Delta$ Tho2 yeast strain does not grow at 37°C (restrictive temperature), Tho2 $\Delta_{1,408-1,597}$  and Tho2 $\Delta_{1,271-1,597}$  growth is considerably reduced, suggesting that C-terminal 1,271–1,597 amino acids are required for cell survival at restrictive temperature (Pena et al., 2012). If the exon28:c.3503+A>C variant caused complete splicing defect retaining intron between exon 28–29 in all mRNAs, the cells would translate only 1,175 amino acids (with 1,168 normal) THOC2 protein; essentially lacking the C-terminal region encompassing the RNA binding domain (RBD) that when deleted in yeast Tho2 $\Delta_{1,271-1,597}$  strain restricts its growth at 37°C. However, the affected boy carrying a single allele of the exon28:c.3503+A>C THOC2 variant, although with severe clinical symptoms, is alive. This could be explained by presence of reduced levels of THOC2 protein produced from translation of about 2/3rd normally spliced mRNA in the affected white blood cells. Taken together, clinical outcomes in the affected boy may be due to perturbed mRNA export caused by reduced levels of THOC2 protein and perhaps also C-terminally truncated THOC2 protein translated from about 1/3rd aberrantly spliced mRNAs that retain a part of intron sequence between exon 28–29.

We also identified a set of previously unreported THOC2 missense variants that, according to ACMG criteria, are variants of uncertain clinical significance (VOUS) (Supp. Table S1): namely, p.Arg77Cys, p.Ser1108Leu, p.Arg1121Gly, and p.Asn1261His. Nevertheless, these variants have supportive evidence pointing toward potential pathogenicity as they are rare (absent from ExAC/gnomAD databases of reference individuals) (Lek et al., 2016), affect highly evolutionarily conserved amino acid residues, are predicted to be pathogenic by *in silico* analyses and are within the clinical presentations spectrum of those seen in individuals with confirmed THOC2-related ID. However, they lack supportive evidence from our existing functional assays. These variants may still have a detrimental effect on THOC2 function due to altered protein structure impacting protein–RNA and/or protein–protein interactions with known or unknown TREX subunits [e.g., Boehringer et al., 2017]. The challenge of proving causality for previously unreported missense variants in NDD genes is well recognized and speaks to the need for ongoing intertwined clinical and research efforts to clarify causality of VOUS (Wright et al., 2018). We therefore report detailed variant and clinical data (see Supp. materials, Supp. Table S1 and S2, and Supp. Figure S4) with the intention of alerting researchers and clinicians to these variants, as future studies, for example identification of their recurrence in affected individuals with overlapping clinical phenotypes or pathophysiological investigations, may help clarify their clinical significance.

THOC2 is ubiquitously expressed in all human tissues (Thul et al., 2017) and more specifically in the developing and mature human brain (Johnson et al., 2009; Kumar et al., 2015; Uhlen et al., 2015) and mouse brain, with higher abundance in frontal cortex and cerebellum (Di Gregorio et al., 2013; Kumar et al., 2015). THOC2 is an essential mRNA export factor as its siRNA-mediated depletion results in almost complete retention of mRNAs in the cell nucleus (Chi et al., 2013), potentially toxic to the cell. These data are consistent with the findings that THOC2 is a highly constrained gene (Samocha et al., 2014)



**FIGURE 6** Summary of truncated human THOC2 proteins translated from aberrantly spliced mRNAs and functional outcomes of yeast C-terminal Tho2 deletion strains (Pena et al., 2012). Blue boxes depict the 52 and 7 amino acids coded by unspliced intron sequences of exon35:c.4550-2T>C and exon28:c.3503+4A>C variants, respectively. WT, +++ = normal,  $\Delta$ 1,408–1,597aa, ++ and  $\Delta$ 1,271–1,597aa = reduced growth at restrictive temperature

and *THOC* (e.g., *THOC1*, 3, 5, 6, and 7) genes are essential for cell survival (Blomen et al., 2015). Taken together, as *THOC2* knockout cells will not survive due to complete mRNA nuclear retention, we predict that the identified *THOC2* variants represent partial loss-of-function that disrupt normal mRNA export in neuronal and possibly other cell types, potentially causing variable clinical presentations.

TREX complex couples transcription and mRNA biogenesis with nuclear mRNA export, and has emerged as an essential pathway in embryogenesis, organogenesis and differentiation (Heath et al., 2016). For example, *Thoc2* and *Thoc5* selectively bind and regulate export of mRNAs (e.g., *Nanog*, *Sox2*, *Esrrb*, and *Klf4* mRNAs) involved in maintenance of pluripotency of mouse ESCs (Wang et al., 2013) and *Thoc5* in maintenance of hematopoiesis and HSP70 mRNA export (Katahira, Inoue, Hurt, & Yoneda, 2009; Mancini et al., 2010). Mouse modeling shows that both *Thoc1* and *Thoc5* knockouts are embryonic lethal (Mancini et al., 2010; Wang, Chang, Li, Zhang, & Goodrich, 2006). However, *Thoc1* and *Thoc5* expression in a range of developing and adult tissues may indicate that the two genes have a more essential role in early embryonic development compared to less stringent requirement during later stages of embryonic or adult development (Mancini et al., 2010; Wang et al., 2006); a functional pattern most likely followed by the *THOC2* gene. Essentiality of *THOC2* gene indicates that *THOC2* knockout will also be lethal. However, reduced levels or perturbed functionality can lead to a range of NDD phenotypes as observed for a cohort of *THOC2* variants identified by us. It is now well established that development of brain depends on tightly regulated and complex sequence of events involving neuronal and glial cell proliferation, migration, and maturation (Chiurazzi & Pirozzi, 2016). Therefore, it is not surprising that our *THOC2* variant data and published work (Dickinson et al., 2016) provides strong evidence that even subtle alterations to the canonical molecular pathways such as mRNA export, otherwise essential for cellular life, can be tolerated but at a cost of a NDD.

In summary, we present detailed clinical data on seven individuals with *THOC2*-associated ID caused by both missense and splice variants that meet ACMG criteria for (likely) pathogenicity. They have a core phenotype of ID, and common findings of behavioural disorders, infantile hypotonia, gait disturbance and growth impairment, similar to the affected males with *THOC2*-associated ID we previously reported (Kumar et al., 2015). Other than the affected female with a *de novo* missense variant, heterozygote carrier females are typically unaffected. We also present data on five individuals with four previously unreported rare missense variants that show clinical overlap with our core group, but where convincing evidence for causality is still required. The significance of these variants may be clarified as additional individuals with *THOC2* variants are reported. We have also “adopted” *THOC2* on the Human Disease Gene (HDG) Website Series (<https://humandiseasesgenes.nl/thoc2>) in an effort to continue to explore the phenotypic-genotypic spectrum for *THOC2*-related ID.

#### ACKNOWLEDGMENTS

We thank the individuals and families for their contribution to this study.

#### DISCLOSURE STATEMENT

The authors declare no conflict of interest.

#### ORCID

Raman Kumar <http://orcid.org/0000-0001-7976-8386>

Tracy Dudding-Byth <http://orcid.org/0000-0002-9551-1107>

Elizabeth E. Palmer <http://orcid.org/0000-0003-1844-215X>

Jozef Gecz <http://orcid.org/0000-0002-7884-6861>

## REFERENCES

- Allen, A. S., Berkovic, S. F., Cossette, P., Delanty, N., Dlugos, D., Eichler, E. E., ... Winawer, M. R. (2013). De novo mutations in epileptic encephalopathies. *Nature*, *501*(7466), 217–221.
- Amos, J. S., Huang, L., Thevenon, J., Kariminedjad, A., Beaulieu, C. L., Masurel-Paulet, A., ... Boycott, K.M. (2017). Autosomal recessive mutations in THOC6 cause intellectual disability: Syndrome delineation requiring forward and reverse phenotyping. *Clinical Genetics*, *91*(1), 92–99.
- Amsterdam, A., Nissen, R. M., Sun, Z., Swindell, E. C., Farrington, S., & Hopkins, N. (2004). Identification of 315 genes essential for early zebrafish development. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(35), 12792–12797.
- Beaulieu, C. L., Huang, L., Innes, A. M., Akimenko, M. A., Puffenberger, E. G., Schwartz, C., ... Boycott, K. M. (2013). Intellectual disability associated with a homozygous missense mutation in THOC6. *Orphanet Journal of Rare Diseases*, *8*(1), 62.
- Bittles, A. H., Petterson, B. A., Sullivan, S. G., Hussain, R., Glasson, E. J., & Montgomery, P. D. (2002). The influence of intellectual disability on life expectancy. *Journal of Gerontology A Biological Sciences and Medical Sciences*, *57*(7), M470–2.
- Blomen, V. A., Majek, P., Jae, L. T., Bigenzahn, J. W., Nieuwenhuis, J., Staring, J., ... Brummelkamp, T. R. (2015). Gene essentiality and synthetic lethality in haploid human cells. *Science*, *350*(6264), 1092–1096.
- Boehringer, A., Garcia-Mansfield, K., Singh, G., Bakkar, N., Pirrotte, P., & Bowser, R. (2017). ALS associated mutations in Matrin 3 alter protein-protein interactions and impede mRNA nuclear export. *Science Reports*, *7*(1), 14529.
- Chi, B., Wang, Q., Wu, G., Tan, M., Wang, L., Shi, M., ... Cheng, H. (2013). Aly and THO are required for assembly of the human TREX complex and association of TREX components with the spliced mRNA. *Nucleic Acids Research*, *41*(2), 1294–1306.
- Chinnam, M., Wang, Y., Zhang, X., Gold, D. L., Khoury, T., Nikitin, A. Y., ... Goodrich, D. W. (2014). The Thoc1 ribonucleoprotein and prostate cancer progression. *Journal of the National Cancer Institute*, *106*(11), dju306.
- Chiurazzi, P., & Pirozzi, F. (2016). Advances in understanding - genetic basis of intellectual disability. *F1000Res*, *5*, 599.
- Choi, Y., & Chan, A. P. (2015). PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics*, *31*(16), 2745–2747.
- Coe, B. P., Witherspoon, K., Rosenfeld, J. A., van Bon, B. W., Vulto-van Silfhout, A. T., Bosco, P., ... Eichler, E. E. (2014). Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nature Genetics*, *46*(10), 1063–1071.
- de Lange, I. M., Helbig, K. L., Weckhuysen, S., Moller, R. S., Velinov, M., Dolzhanskaya, N., ... Koeleman, B. P. C. (2016). De novo mutations of KIAA2022 in females cause intellectual disability and intractable epilepsy. *Journal of Medical Genetics*, *53*(12), 850–858.
- Di Gregorio, E., Bianchi, F. T., Schiavi, A., Chiotto, A. M., Rolando, M., Verdun di Cantogno, L., ... Brusco, A. (2013). A de novo X;8 translocation creates a PTK2-THOC2 gene fusion with THOC2 expression knockdown in a patient with psychomotor retardation and congenital cerebellar hypoplasia. *Journal of Medical Genetics*, *50*(8), 543–551.
- Dickinson, M. E., Flenniken, A. M., Ji, X., Teboul, L., Wong, M. D., White, J. K., ... Murray, S. A. (2016). High-throughput discovery of novel developmental phenotypes. *Nature*, *537*(7621), 508–514.
- Dudding-Byth, T., Baxter, A., Holliday, E. G., Hackett, A., O'Donnell, S., White, S. M., ... Lovell, B. C. (2017). Computer face-matching technology using two-dimensional photographs accurately matches the facial gestalt of unrelated individuals with the same syndromic form of intellectual disability. *BMC Biotechnology*, *17*(1), 90.
- Guria, A., Tran, D. D., Ramachandran, S., Koch, A., El Bounkari, O., Dutta, P., ... Tamura, T. (2011). Identification of mRNAs that are spliced but not exported to the cytoplasm in the absence of THOC5 in mouse embryo fibroblasts. *RNA*, *17*(6), 1048–1056.
- Hautbergue, G. M. (2017). RNA Nuclear Export: From neurological disorders to cancer. *Advances in Experimental Medicine and Biology*, *1007*, 89–109.
- Heath, C. G., Viphakone, N., & Wilson, S. A. (2016). The role of TREX in gene expression and disease. *Biochemical Journal*, *473*(19), 2911–2935.
- Hirayama, S., Yamazaki, Y., Kitamura, A., Oda, Y., Morito, D., Okawa, K., ... Nagata, K. (2008). MKKS is a centrosome-shuttling protein degraded by disease-causing mutations via CHIP-mediated ubiquitination. *Molecular Biology of the Cell*, *19*(3), 899–911.
- Hosking, F. J., Carey, I. M., Shah, S. M., Harris, T., DeWilde, S., Beighton, C., & Cook, D. G. (2016). Mortality among adults with intellectual disability in England: Comparisons with the general population. *American Journal of Public Health*, *106*(8), 1483–1490.
- Johnson, M. B., Kawasawa, Y. I., Mason, C. E., Krsnik, Z., Coppola, G., Bogdanovic, D., ... Sestan, N. (2009). Functional and evolutionary insights into human brain development through global transcriptome analysis. *Neuron*, *62*(4), 494–509.
- Katahira, J., Inoue, H., Hurt, E., & Yoneda, Y. (2009). Adaptor Aly and co-adaptor Thoc5 function in the Tap-p15-mediated nuclear export of HSP70 mRNA. *EMBO Journal*, *28*(5), 556–567.
- Kim, W., Bennett, E. J., Huttlin, E. L., Guo, A., Li, J., Possemato, A., ... Gygi, S. P. (2011). Systematic and quantitative assessment of the ubiquitin-modified proteome. *Molecular Cell*, *44*(2), 325–340.
- Kircher, M., Witten, D. M., Jain, P., O'Roak, B. J., Cooper, G. M., & Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nature Genetics*, *46*(3), 310–315.
- Kumar, R., Corbett, M. A., van Bon, B. W., Woenig, J. A., Weir, L., Douglas, E., ... Gecz, J. (2015). THOC2 mutations implicate mRNA-export pathway in X-linked intellectual disability. *American Journal of Human Genetics*, *97*(2), 302–310.
- Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., ... MacArthur, D. G. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, *536*(7616), 285–291.
- Liu, C., Yue, B., Yuan, C., Zhao, S., Fang, C., Yu, Y., & Yan, D. (2015). Elevated expression of Thoc1 is associated with aggressive phenotype and poor prognosis in colorectal cancer. *Biochemical and Biophysical Research Communication*, *468*(1–2), 53–58.
- Lopitz-Otsoa, F., Rodriguez-Suarez, E., Aillet, F., Casado-Vela, J., Lang, V., Matthiesen, R., ... Rodriguez, M. S. (2012). Integrative analysis of the ubiquitin proteome isolated using Tandem Ubiquitin Binding Entities (TUBEs). *Journal of Proteomics*, *75*(10), 2998–3014.
- Mancini, A., Niemann-Seyde, S. C., Pankow, R., El Bounkari, O., Klebbafarber, S., Koch, A., ... Tamura, T. (2010). THOC5/FMIP, an mRNA export TREX complex protein, is essential for hematopoietic primitive cell survival in vivo. *BMC Biology*, *8*, 1.
- Milani, D., Ronzoni, L., & Esposito, S. (2015). Genetic advances in intellectual disability. *Journal of Pediatric Genetics*, *4*(3), 125–127.
- Morgan, D. O. (2013). The D box meets its match. *Molecular Cell*, *50*(5), 609–610.
- Oeseburg, B., Dijkstra, G. J., Groothoff, J. W., Reijneveld, S. A., & Jansen, D. E. (2011). Prevalence of chronic health conditions in children with intellectual disability: A systematic literature review. *Intellectual and Developmental Disabilities*, *49*(2), 59–85.
- Ohno, K., Takeda, J. I., & Masuda, A. (2018). Rules and tools to predict the splicing effects of exonic and intronic mutations. *Wiley Interdisciplinary Reviews RNA*, *9*(1), e1451.

- Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., & Mann, M. (2006). Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell*, 127(3), 635–648.
- Palmer, E. E., Kumar, R., Gordon, C. T., Shaw, M., Hubert, L., Carroll, R., ... Gecz, J. (2017). A recurrent de novo nonsense variant in ZSWIM6 results in severe intellectual disability without frontonasal or limb malformations. *American Journal of Human Genetics*, 101(6), 995–1005.
- Palmer, E. E., Stuhlmann, T., Weinert, S., Haan, E., Van Esch, H., Holvoet, M., ... Kalscheuer, V. M. (2016). De novo and inherited mutations in the X-linked gene CLCN4 are associated with syndromic intellectual disability and behavior and seizure disorders in males and females. *Molecular Psychiatry*, 23(2), 222–230.
- Pena, A., Gewartowski, K., Mroczek, S., Cuellar, J., Szykowska, A., Prokop, A., ... Dziembowski, A. (2012). Architecture and nucleic acids recognition mechanism of the THO complex, an mRNP assembly factor. *EMBO Journal*, 31(6), 1605–1616.
- Rehwinkel, J., Herold, A., Gari, K., Kocher, T., Rode, M., Ciccarelli, F. L., ... Izaurralde, E. (2004). Genome-wide analysis of mRNAs regulated by the THO complex in *Drosophila melanogaster*. *Nature Structural and Molecular Biology*, 11(6), 558–566.
- Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., ... Rehm, H. L. (2015). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine*, 17(5), 405–424.
- Ropers, H. H., & Hamel, B. C. (2005). X-linked mental retardation. *Nature Reviews Genetics*, 6(1), 46–57.
- Samocha, K. E., Robinson, E. B., Sanders, S. J., Stevens, C., Sabo, A., McGrath, L. M., ... Daly, M. J. (2014). A framework for the interpretation of de novo mutation in human disease. *Nature Genetics*, 46(9), 944–950.
- Schwartz, C. E. (2015). *X-linked intellectual disability genetics*. Wiley Online Library; Chichester, UK: John Wiley & Sons, Ltd.
- Snijders Blok, L., Madsen, E., Juusola, J., Gilissen, C., Baralle, D., Reijnders, M. R., ... Kleefstra, T. (2015). Mutations in DDX3X are a common cause of unexplained intellectual disability with gender-specific effects on Wnt signaling. *American Journal of Human Genetics*, 97(2), 343–352.
- Thul, P. J., Akesson, L., Wiking, M., Mahdessian, D., Geladaki, A., Ait Blal, H., ... Lundberg, E. (2017). A subcellular map of the human proteome. *Science*, 356(6340).
- Uhlen, M., Fagerberg, L., Hallstrom, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., ... Ponten, F. (2015). Proteomics. Tissue-based map of the human proteome. *Science*, 347(6220), 1260419.
- Viphakone, N., Cumberbatch, M. G., Livingstone, M. J., Heath, P. R., Dickman, M. J., Catto, J. W., & Wilson, S. A. (2015). Luszp4 defines a new mRNA export pathway in cancer cells. *Nucleic Acids Research*, 43(4), 2353–2366.
- Vissers, L. E., Gilissen, C., & Veltman, J. A. (2016). Genetic studies in intellectual disability and related disorders. *Nature Reviews Genetics*, 17(1), 9–18.
- Wagner, S. A., Beli, P., Weinert, B. T., Nielsen, M. L., Cox, J., Mann, M., & Choudhary, C. (2011). A proteome-wide, quantitative survey of in vivo ubiquitylation sites reveals widespread regulatory roles. *Molecular and Cellular Proteomics*, 10(10), M111.013284.
- Wang, L., Miao, Y. L., Zheng, X., Lackford, B., Zhou, B., Han, L., ... Hu, G. (2013). The THO complex regulates pluripotency gene mRNA export and controls embryonic stem cell self-renewal and somatic cell reprogramming. *Cell Stem Cell*, 13(6), 676–690.
- Wang, X., Chang, Y., Li, Y., Zhang, X., & Goodrich, D. W. (2006). Thoc1/Hpr1/p84 is essential for early embryonic development in the mouse. *Molecular and Cellular Biology*, 26(11), 4362–4367.
- Wang, X., Chinnam, M., Wang, J., Wang, Y., Zhang, X., Marcon, E., ... Goodrich, D. W. (2009). Thoc1 deficiency compromises gene expression necessary for normal testis development in the mouse. *Molecular and Cellular Biology*, 29(10), 2794–2803.
- Woerner, A. C., Frottin, F., Hornburg, D., Feng, L. R., Meissner, F., Patra, M., ... Hipp, M. S. (2016). Cytoplasmic protein aggregates interfere with nucleocytoplasmic transport of protein and RNA. *Science*, 351(6269), 173–176.
- Wright, C. F., McRae, J. F., Clayton, S., Gallone, G., Aitken, S., FitzGerald, T. W., ... Firth, H. V. (2018). Making new genetic diagnoses with old data: Iterative reanalysis and reporting from genome-wide data in 1,133 families with developmental disorders. *Genetics in Medicine*, <https://doi.org/10.1038/gim.2017.246>
- Yamazaki, T., Fujiwara, N., Yukinaga, H., Ebisuya, M., Shiki, T., Kurihara, T., ... Weis, K. (2010). The closely related RNA helicases, UAP56 and URH49, preferentially form distinct mRNA export machineries and coordinately regulate mitotic progression. *Molecular Biology of Cell*, 21(16), 2953–2965.
- Zhu, X., Need, A. C., Petrovski, S., & Goldstein, D. B. (2014). One gene, many neuropsychiatric disorders: Lessons from Mendelian diseases. *Nature Neuroscience*, 17(6), 773–781.
- Zweier, C., Rittinger, O., Bader, I., Berland, S., Cole, T., Degenhardt, F., ... Wiczorek, D. (2014). Females with de novo aberrations in PHF6: Clinical overlap of Borjeson-Forssman-Lehmann with Coffin-Siris syndrome. *American Journal of Medical Genetics C: Seminars in Medical Genetics*, 166C(3), 290–301.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Kumar R, Gardner A, Homan CC, et al. Severe neurocognitive and growth disorders due to variation in *THOC2*, an essential component of nuclear mRNA export machinery. *Human Mutation*. 2018;39:1126–1138. <https://doi.org/10.1002/humu.23557>



CASE REPORT

Open Access



# Exonic duplication of the *OTC* gene by a complex rearrangement that likely occurred via a replication-based mechanism: a case report

Katsuyuki Yokoi<sup>1,2</sup>, Yoko Nakajima<sup>1</sup>, Hidehito Inagaki<sup>2</sup>, Makiko Tsutsumi<sup>2</sup>, Tetsuya Ito<sup>1</sup> and Hiroki Kurahashi<sup>2\*</sup> 

## Abstract

**Background:** Ornithine transcarbamylase deficiency (OTCD) is an X-linked recessive disorder involving a defect in the urea cycle caused by *OTC* gene mutations. Although a total of 417 disease-causing mutations in *OTC* have been reported, structural abnormalities in this gene are rare. We here describe a female OTCD case caused by an exonic duplication of the *OTC* gene (exons 1–6).

**Case presentation:** A 23-year-old woman with late-onset OTCD diagnosed by biochemical testing was subjected to subsequent genetic testing. Sanger sequencing revealed no pathogenic mutation throughout the coding exons of the *OTC* gene, but multiplex ligation-dependent probe amplification (MLPA) revealed duplication of exons 1–6. Further genetic analyses revealed an inversion of duplicated exon 1 and a tandem duplication of exons 2–6. Each of the junctions of the inversion harbored a microhomology and non-templated microinsertion, respectively, suggesting a replication-based mechanism. The duplication was also of de novo origin but segregation analysis indicated that it took place in the paternal chromosome.

**Conclusion:** We report the first OTCD case harboring an exonic duplication in the *OTC* gene. The functional defects caused by this anomaly were determined via structural analysis of its complex rearrangements.

**Keywords:** Ornithine transcarbamylase deficiency, Exonic duplication, Complex rearrangement, Fork stalling and template switching (FoSTeS), Non-homologous end joining (NHEJ)

## Background

Ornithine transcarbamylase (*OTC*) is a mitochondrial urea cycle enzyme that catalyzes the reaction between carbamyl phosphate and ornithine to form citrulline and phosphate [1]. Ornithine transcarbamylase deficiency (OTCD) is one of the most common urea cycle disorders [2] with an estimated prevalence of 1 in 14,000–77,000 [1]. The human *OTC* gene, located on the short arm of the X chromosome (Xp11.4), is 73 kb with 10 exons and 1062 bp of coding sequence [3–5]. Because OTCD is inherited in an X-linked manner, deficient hemizygous males usually develop this disorder. However, a remarkable feature of OTCD is that a

substantial subset of heterozygous females also develop this condition. The symptoms of carrier females vary in terms of onset and severity. Since the *OTC* gene is subject to X-inactivation, it is believed that this phenotypic variability depends on a skewed degree of this in the livers of carrier females [5].

In 85–90% of patients with a biochemical phenotype of OTCD, a mutation can be identified through sequencing or deletion/duplication testing [6]. A total of 417 disease-causing mutations in the *OTC* gene have been reported to date [1]. Exonic deletions have also been described but no prior case of OTCD caused by exonic duplication has previously been reported [7]. In our current case report, we describe a female patient with OTCD caused by a partial duplication of *OTC* exons 1–6.

\* Correspondence: [kura@fujita-hu.ac.jp](mailto:kura@fujita-hu.ac.jp)

<sup>2</sup>Division of Molecular Genetics, Institute for Comprehensive Medical Science, Fujita Health University, 1-98 Dengakugakubo, Kutsukake-cho, Toyoake, Aichi 470-1192, Japan

Full list of author information is available at the end of the article





## Case presentation

### Patient

The current study patient was a 23-year-old woman with normal psychomotor development and healthy nonconsanguineous parents. She had frequent episodes of nausea, vomiting, stomachache and temporary elevated transaminase from about 4 years of age. Ammonia and plasma amino acid levels were measured when she was 5 years old. Her serum ammonia was 220  $\mu\text{g}/\text{dl}$  (normal range 12 ~ 60  $\mu\text{g}/\text{dl}$ ) and she showed high levels of glutamine (1212 nmol/ml; normal value, 420–700), lower normal limits of citrulline (18.4 nmol/ml; normal value, 17–43), and lower plasma levels of arginine (32.2 nmol/ml; normal value, 54–130). A urine metabolic screen indicated a gross elevation in orotate (orotate/creatinine ratio 234.3  $\mu\text{mol}/\text{g}$  creatinine; normal value, 4.7 ~ 15.9  $\mu\text{mol}/\text{g}$  creatinine). These findings were consistent with OTC deficiency. She was therefore biochemically diagnosed with OTCD and her blood ammonia level has been well controlled since by a protein-restricted diet and by oral sodium phenylbutyrate and arginine. Recently, we performed genetic analysis to identify the genetic alterations of the *OTC* gene in this patient. However, Sanger sequencing revealed no pathogenic mutation.

### Genetic analysis

#### Mutational analyses

Sanger sequencing was performed to screen for genetic variations at the nucleotide level throughout all coding exons of the *OTC* gene (Additional file 1). We used UCSC genome browser (<http://genome-asia.ucsc.edu/>) as human genome assembly. To screen for exonic deletions or duplications, multiplex ligation-dependent probe amplification (MLPA) was performed using the SALSA P079-A3 *OTC* MLPA kit (MRC Holland, Amsterdam, The Netherlands), in accordance with the manufacturer's recommendations. MLPA products were separated by capillary electrophoresis on an ABI3730 genetic analyzer and then processed using GeneMapper software. The peak heights of the samples were compared with control probes and the ratios of these peaks were calculated for all exons. If the dosage quotient was 1.0, the results were considered normal. Thresholds for deletions and duplications were set at 0.5 and 1.5, respectively.

#### Quantitative real time PCR

To demarcate the duplicated region, quantitative real-time PCR was conducted on blood DNA from the patient and a male control subject using the Applied Biosystems 7300 real time PCR system (Thermo Fisher Scientific). Several primer pairs were designed for *OTC* (upstream of exon 1 and intron 6) and *RPP30* that was used as an autosomal single copy gene reference to

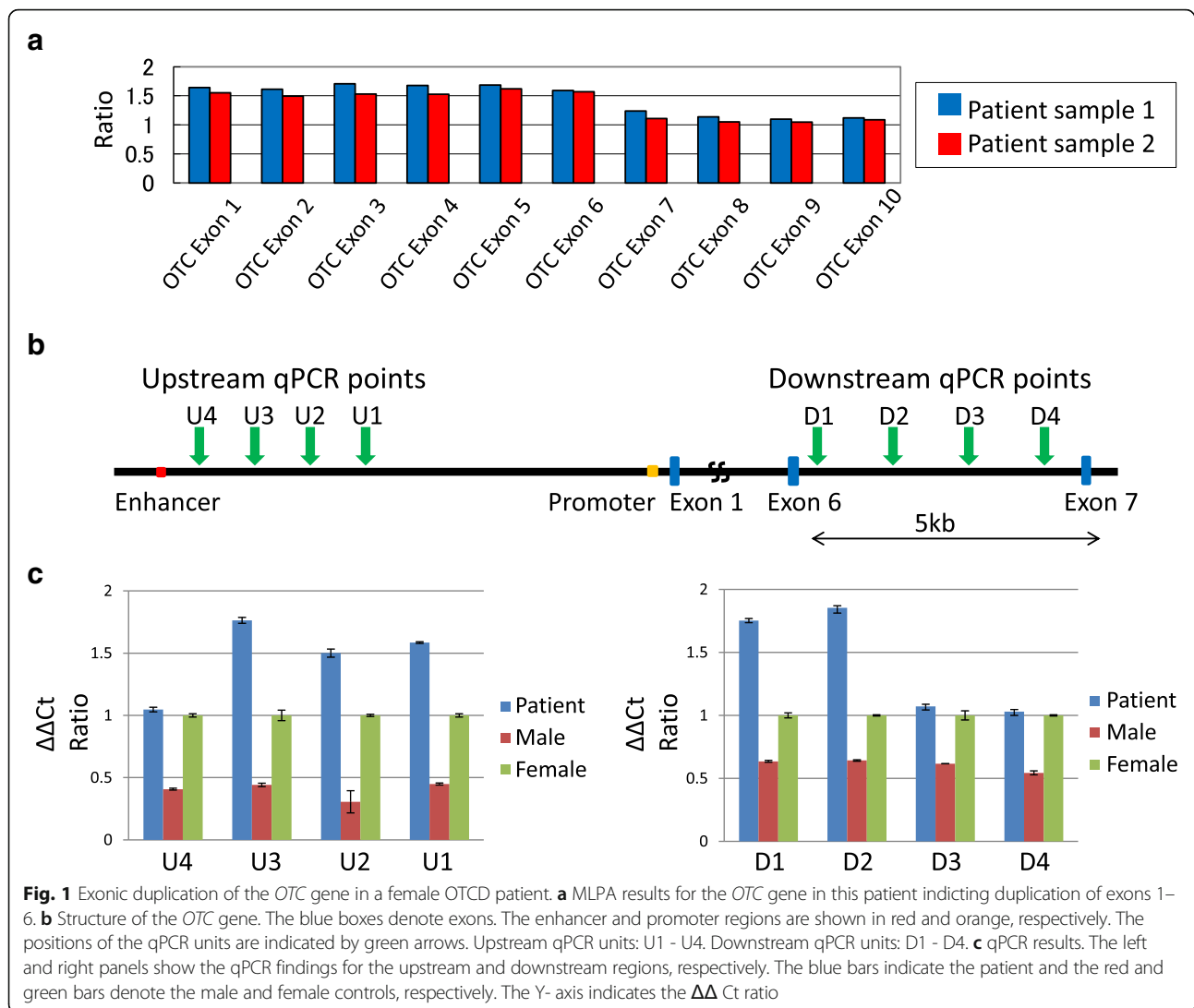
generate amplicons suitable for real-time PCR (Fig. 1b, Additional file 1). The PCR reaction was performed in a 15  $\mu\text{L}$  reaction system, containing 2  $\mu\text{L}$  of template DNA (5 ng/ $\mu\text{L}$ ), 0.6  $\mu\text{L}$  of each primer set (10  $\mu\text{mol}/\text{L}$ ), 0.3  $\mu\text{L}$  ROX Reference Dye, 4  $\mu\text{L}$  distilled water, and 7.5  $\mu\text{L}$  of 2xTB Green Premix Ex TaqII (Tli RNaseH Plus, TaKaRa). Two parallel PCR reactions were prepared for each sample. The amplification cycling conditions were as follows: 95 °C for 30 s, followed by 40 cycles at 95 °C 5 s and 60 °C for 1 min. Data evaluation was carried out using the 7300 system SDS software and Microsoft Excel. The threshold cycle number (Ct) was determined for all PCR reactions and the same threshold and baseline were set for all samples. The starting copy number of the samples was determined using the  $\Delta\Delta\text{Ct}$ -Method.  $\Delta\Delta\text{Ct}$  method was a modification of the method described in Livak et al. for quantifying mRNA [8].  $\Delta\text{Ct}$  represents the mean Ct value of each sample and was calculated for *OTC* and *RPP30*. The starting copy number of the unknown samples was determined relative to the known copy number of the control sample using the following formula:

$$\Delta\Delta\text{Ct} = [\Delta\text{Ct } OTC(\text{patient}) - \Delta\text{Ct } RPP30(\text{patient})] - [\Delta\text{Ct } OTC(\text{female}) - \Delta\text{Ct } RPP30(\text{female})].$$

The relative gene copy number was calculated by the expression  $2^{-\Delta(\Delta\text{Ct})}$ . The starting copy number of male control was also determined as a reference value.

#### Inverse PCR

Inverse PCR were performed using restriction enzyme *TaqI* (TaKaRa, Shiga, Japan) to isolate the unknown sequences adjacent to the duplicated region of the *OTC* gene in the study patient. ApE – A plasmid Editor software was used to identify the recognition sites for the restriction enzyme. The restriction enzyme was chosen based on the following criteria: (1) no cutting of the expected breakpoint area; and (2) endonuclease activity would be unaffected by CpG methylation of the target sequence. A 100 ng aliquot of genomic DNA from both our patient and a control female was digested with the selected restriction enzyme in a total volume of 30  $\mu\text{L}$  at 65 °C for 90 min. The reaction was inactivated using the QiaQuick PCR Purification Kit. A 20  $\mu\text{L}$  sample of digested DNA was then mixed with 23  $\mu\text{L}$  of DW, 5  $\mu\text{L}$  of 10 × T4 ligase buffer (TaKaRa, Shiga, Japan) and 2  $\mu\text{L}$  of T4 DNA ligase to make a final volume of 50  $\mu\text{L}$ . Ligation reactions were incubated at 16 °C for 16 h. For subsequent PCR, 1  $\mu\text{L}$  of digested and re-ligated DNA template was used in a total reaction volume of 25  $\mu\text{L}$  with Tks Gflex DNA Polymerase (TaKaRa, Shiga, Japan). Primers were designed to avoid repetitive sequences (Additional file 1). The PCR conditions were as follows: 30 cycles of 10 s at 98 °C, 15 s at 60 °C, and 1 min at 68 °C. Amplified products were analyzed by gel



electrophoresis and were purified following nested PCR (Additional file 1). The purified PCR products were sequenced via the standard Sanger method.

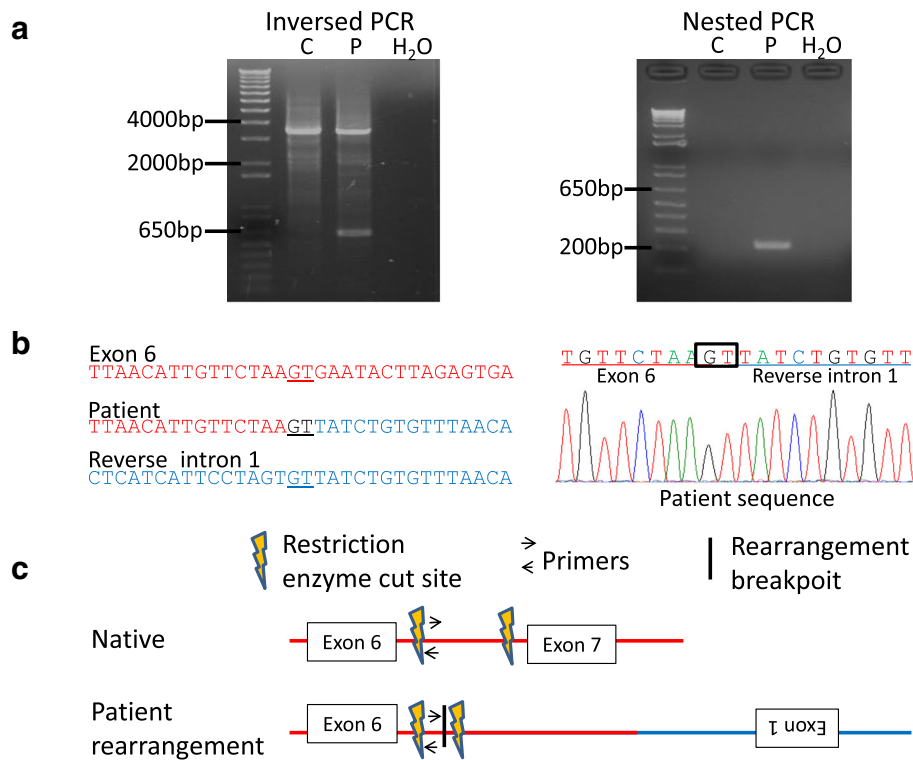
#### Breakpoint analysis on the other side

PCR was performed using Tks Gflex (TaKaRa, Shiga, Japan) to confirm the other side of the breakpoint sequence. Primer R which was previously designed for real-time PCR analysis of *OTC* upstream of exon 1 (i.e. *OTC* intron 1) was used as primer F in this reaction (Additional file 1). The PCR conditions and Sanger methodology were similar to those described above.

MLPA revealed the duplication of exons 1–6 of the *OTC* gene in our current study patient (Fig. 1a). We determined the range of the duplication using quantitative real-time PCR (Fig. 1b). We designed four qPCR experiments (U1–U4) between the promoter and enhancer regions to identify the upstream breakpoint. Likewise, we

designed four qPCR assays (D1–D4) within intron 6 to identify the downstream breakpoint. In contrast to the male or female controls that showed  $\Delta\Delta$ Ct ratios of 0.5 or 1.0, respectively, the patient's samples showed a  $\Delta\Delta$ Ct ratio > 1.5 in some of these qPCR assays, suggesting that these regions were duplicated in this patient (Fig. 1c). The results indicated that the putative upstream breakpoints were located between PCR U3 and U4, and that the downstream breakpoints were between PCR D2 and D3.

We next performed inverse PCR to analyze the genomic structure of the duplicated region. *TaqI*-digested DNA was used as a template to produce a 3.5 kb PCR product when amplified with inversely oriented intron 6 primers (Fig. 2a, c). However, an additional small PCR product was detected by agarose gel electrophoresis in the patient sample (Fig. 2a). The amplified products were sequenced after nested PCR (Fig. 2a). As expected,



**Fig. 2** Identification of the duplication junction via inverse PCR. **a** Isolation of the junction fragment. Two distinct inverse PCR products were observed following agarose gel electrophoresis. The larger product was derived from a normal allele and the small product from a rearranged allele (left). The amplified products were purified following nested PCR (right). P, patient; C, control; H, H<sub>2</sub>O. **b** Sanger sequencing of the PCR products including the junction. The unknown sequence next to the junction was identified as intron 1 of the *OTC* gene in the reverse orientation. The normal exon 6 and intron 1 sequences are aligned in red and blue typeface, respectively. Underlined nucleotides indicate microhomology at the breakpoint junction. **c** Predicted structure of the junction. Horizontal arrows indicate the recognition sites of the primers used for inverse PCR and the vertical arrows denote the *TaqI* restriction sites

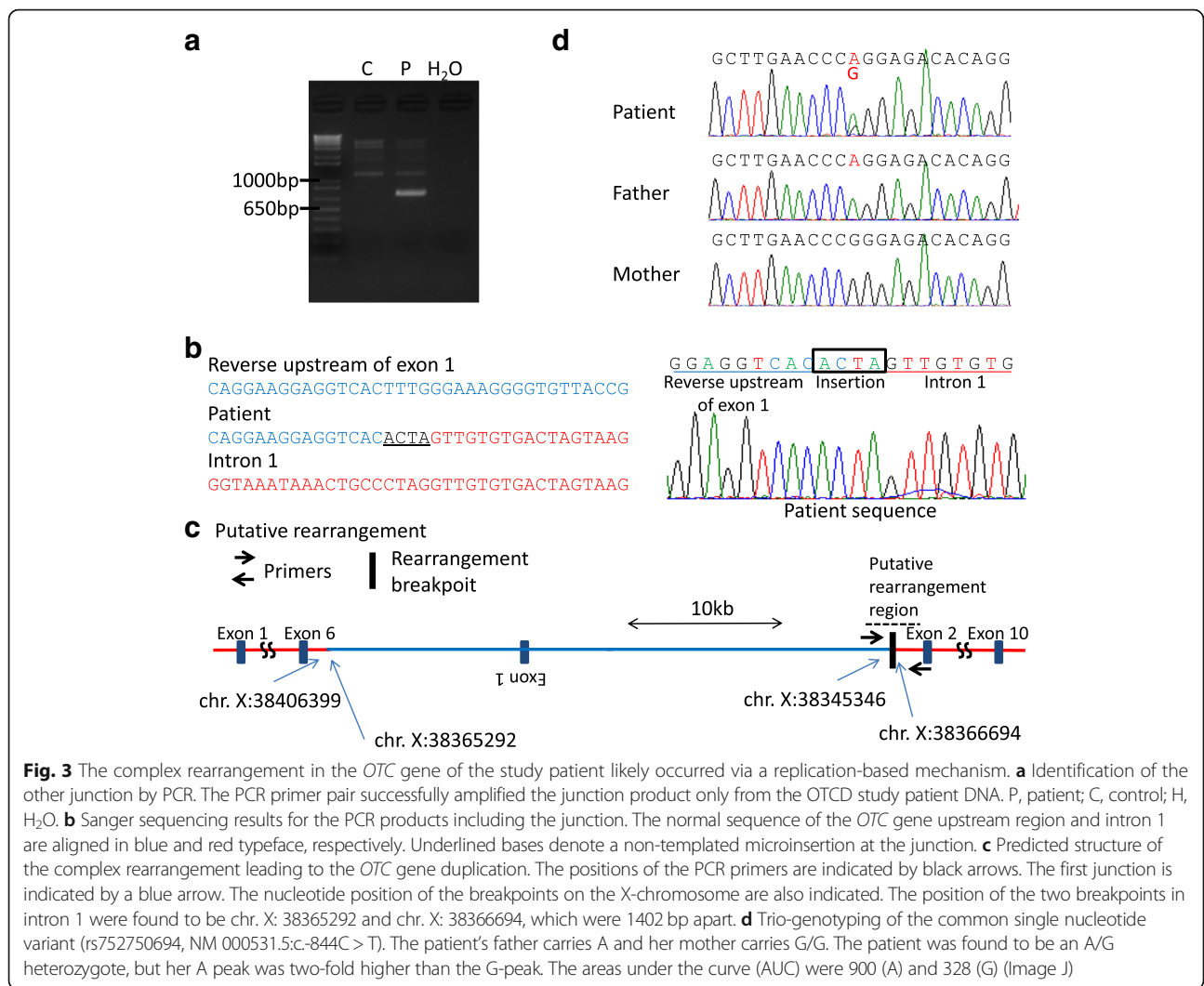
the breakpoint was located within intron 6 (Fig. 2b, c). Unexpectedly however, this breakpoint was found to be connected with intron 1 of the *OTC* gene in the reverse orientation. The breakpoint junction contained 2 nucleotides of microhomology at the fusion junction (Fig. 2b).

The other side breakpoint was analyzed using standard PCR with primers for the upstream breakpoint region and the breakpoint region in intron 1. The primer pair amplified only products from the patient's DNA (Fig. 3a). By Sanger sequencing, the upstream region of the *OTC* gene was found to make an inverted connection with 1 (Fig. 3b, c). This breakpoint junction contained an additional 4 nucleotides (ACTA) of unknown origin (Fig. 3b). The positions of the two breakpoints in intron 1 were found to be chrX: 38365292 and chrX: 38366694, which were 1402 bp apart (Fig. 3c). We performed the same PCR amplification of both junctions in the patient's parents but detected no products, suggesting that this complex rearrangement arose de novo. The patient's duplicated region included a common single nucleotide variant (rs752750694, NM\_000531.5:c.-844C > T). The patient's father carries an A whereas the mother carries

a G/G at this site (Fig. 3d). The patient was found to be an A/G heterozygote, but the peak of the A nucleotide was two-fold greater than the G-peak, suggesting that the patient carries two copies of A. These data suggest that the de novo duplication was of paternal origin.

## Discussion and conclusions

We here report the first documented case of OTCD caused by an exonic duplication of the *OTC* gene. Although the MLPA results for this case indicated a simple duplication of exons 1–6, further analysis indicated that it resulted from complex rearrangements. Two possible mechanisms have been proposed for such rearrangements: one is chromothripsis that is caused by chromosome shattering followed by reunion, and the other is chromoanasythesis that is a replication-based mechanism also known as fork stalling and template switching (FoSTeS)/microhomology-mediated break-induced replication (MMBIR). According to the replication-based model, the active replication fork can stall and switch templates using complementary template microhomology to anneal and prime DNA replication. This



**Fig. 3** The complex rearrangement in the *OTC* gene of the study patient likely occurred via a replication-based mechanism. **a** Identification of the other junction by PCR. The PCR primer pair successfully amplified the junction product only from the OTCD study patient DNA. P, patient; C, control; H, H<sub>2</sub>O. **b** Sanger sequencing results for the PCR products including the junction. The normal sequence of the *OTC* gene upstream region and intron 1 are aligned in blue and red typeface, respectively. Underlined bases denote a non-templated microinsertion at the junction. **c** Predicted structure of the complex rearrangement leading to the *OTC* gene duplication. The positions of the PCR primers are indicated by black arrows. The first junction is indicated by a blue arrow. The nucleotide position of the breakpoints on the X-chromosome are also indicated. The position of the two breakpoints in intron 1 were found to be chr. X: 38365292 and chr. X: 38366694, which were 1402 bp apart. **d** Trio-genotyping of the common single nucleotide variant (rs752750694, NM 000531.5:c.-844C > T). The patient's father carries A and her mother carries G/G. The patient was found to be an A/G heterozygote, but her A peak was two-fold higher than the G-peak. The areas under the curve (AUC) were 900 (A) and 328 (G) (Image J)

mechanism enables the joining or template-driven juxtaposition of different sequences from discrete genomic positions and can result in complex rearrangements [9].

In our current OTCD case, one junction presented 2 nucleotides of microhomology (GT), and the other junction manifested 4 nucleotides as a microinsertion (ACTA). Copy number variation with complex rearrangements and the presence of microhomology is indicative of a replication-based mechanism but the evidence for a non-templated microinsertion is noteworthy. Microinsertions are often observed in non-proofing DNA repair processes such as non-homologous end joining (NHEJ), which is activated by double-strand breaks [10]. However, a considerable body of evidence now suggests that microinsertions can be identified at junctions mediated by DNA replication-based mechanisms [11, 12]. A recent study has also suggested that an NHEJ-like pathway mediated by Polθ, which is an alternative NHEJ mechanism, may be induced by replication stress [13]. Taken together, an alternative NHEJ pathway might be activated during aberrant

replication to restore DNA integrity, thus leading to chromoanagenesis.

The evidence to date also suggests that de novo mutations occur more frequently in paternal alleles [14]. This bias is attributed to the higher number of DNA replication events in spermatogenesis than in oogenesis. Likewise, chromosomal structural variations are more frequently derived from the father [15]. The complex genomic rearrangements in our present patient were found to be of de novo origin but genotyping of a single nucleotide variant in the *OTC* gene demonstrated that the rearrangement allele originated from her father. Given the higher chance of the DNA replication errors during spermatogenesis, it might also reflect the replication-based mechanism.

MLPA can be used in the molecular diagnosis of several genetic diseases whose pathogenesis is related to the presence of deletions or duplications of specific genes [16]. Although deletions are clearly pathogenic, this is less certain in the case of duplications. In case of the

*OTC* gene for example, duplications of the entire gene are innocuous and present as a normal variant in the general population [7]. In cases of partial duplication as seen in our current patient, gene function may not be necessarily be affected when the additional sequence is inserted into another genomic locus. Even in cases of a tandem duplication, it is feasible that one copy of the *OTC* gene may maintain an intact structure. In the current OTCD case, the inversion of exon 1 occurred together with its duplication. We predicted in this instance that this complex rearrangement would generate a tandem duplication of exons 2–6 and the production of truncated *OTC* proteins with defective function due to a frameshift or null protein expression due to nonsense-mediated mRNA decay. The functional defects caused by this mutant allele were therefore the cause of the OTCD in this woman.

In conclusion, we report the first case of OTCD caused by a complex rearrangement resulting in exonic duplication of the *OTC* gene. Our present report also emphasizes the necessity of fully investigating whether pathogenicity has resulted from a genomic duplication.

## Additional file

**Additional file 1:** PCR primers and genomic coordinates. (a) Primers for Sanger sequences of *OTC* exons. (b) Primers for qRT-PCR. (c) Other PCR. (XLSX 13 kb)

## Abbreviations

FoSTeS: Fork stalling and template switching; MLPA: Multiplex ligation-dependent probe amplification; MMBIR: Microhomology-mediated break-induced replication; NHEJ: Non-homologous end joining; *OTC*: Ornithine transcarbamylase; OTCD: Ornithine transcarbamylase deficiency

## Acknowledgements

We thank the patient and her family for their participation in this study. We also thank past and present members of our laboratory.

## Funding

No funding was received.

## Availability of data and materials

All data generated or analysed during this study are included in this published article [and its Additional files].

## Authors' contributions

KY did most of the experiments, retrieved the data, drafted and revised the manuscript. YN and TI discovered the patients and provided many data. HI and MT supported and supervised experiments. HK have contributed equally to the manuscript. All authors contributed to and reviewed the manuscript. All authors read and approved the final manuscript.

## Ethics approval and consent to participate

All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2005(5). The study protocol was approved by the Ethical Review Board for Human Genome Studies at Fujita Health University.

## Consent for publication

Written informed consent to publish medical information and images was obtained from all patients reported in this publication.

## Competing interests

The authors declare that they have no competing interests.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Author details

<sup>1</sup>Department of Pediatrics, Fujita Health University School of Medicine, Toyoake, Japan. <sup>2</sup>Division of Molecular Genetics, Institute for Comprehensive Medical Science, Fujita Health University, 1-98 Dengakugakubo, Kutsukake-cho, Toyoake, Aichi 470-1192, Japan.

Received: 13 September 2018 Accepted: 3 December 2018

Published online: 12 December 2018

## References

- Caldovic L, Abdikarim I, Narain S, Tuchman M, Morizono H. Genotype-phenotype correlations in ornithine transcarbamylase deficiency: a mutation update. *J Genet Genomics*. 2015;42(5):181–94.
- Seminara J, Tuchman M, Krivitzy L, Krischer J, Lee HS, Lemons C, et al. Establishing a consortium for the study of rare diseases: the urea cycle disorders consortium. *Mol Genet Metab*. 2010;100(Suppl 1):S97–105.
- Lindgren V, de Martinville B, Horwich AL, Rosenberg LE, Francke U. Human ornithine transcarbamylase locus mapped to band Xp21.1 near the Duchenne muscular dystrophy locus. *Science*. 1984;226(4675):698–700.
- Horwich AL, Fenton WA, Williams KR, Kalousek F, Kraus JP, Doolittle RF, et al. Structure and expression of a complementary DNA for the nuclear coded precursor of human mitochondrial ornithine transcarbamylase. *Science*. 1984;224(4653):1068–74.
- Yorifuji T, Muroi J, Uematsu A, Tanaka K, Kiwaki K, Endo F, et al. X-inactivation pattern in the liver of a manifesting female with ornithine transcarbamylase (*OTC*) deficiency. *Clin Genet*. 1998;54(4):349–53.
- Jang YJ, LaBella AL, Feeney TP, Braverman N, Tuchman M, Morizono H, et al. Disease-causing mutations in the promoter and enhancer of the ornithine transcarbamylase gene. *Hum Mutat*. 2018;39(4):527–36.
- Shchelochkov OA, Li FY, Geraghty MT, Gallagher RC, Van Hove JL, Lichter-Konecki U, et al. High-frequency detection of deletions and variable rearrangements at the ornithine transcarbamylase (*OTC*) locus by oligonucleotide array CGH. *Mol Genet Metab*. 2009;96(3):97–105.
- Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods*. 2001;25(4):402–8.
- Zhang F, Carvalho CM, Lupski JR. Complex human chromosomal and genomic rearrangements. *Trends Genet*. 2009a;25(7):298–307.
- Lieber MR. The mechanism of human nonhomologous DNA end joining. *J Biol Chem*. 2008;283(1):1–5.
- Hastings PJ, Lupski JR, Rosenberg SM, Ira G. Mechanisms of change in gene copy number. *Nat Rev Genet*. 2009;10(8):551–64.
- Zhang F, Khajavi M, Connolly AM, Towne CF, Batish SD, Lupski JR. The DNA replication FoSTeS/MMBIR mechanism can generate genomic, genic and exonic complex rearrangements in humans. *Nat Genet*. 2009b;41(7):849–53.
- Masset H, Hestand MS, Van Esch H, Kleinfinger P, Plaisancié J, Afenjar A, et al. A distinct class of Chromoanagenesis events characterized by focal copy number gains. *Hum Mutat*. 2016;37(7):661–8.
- Jónsson H, Sulem P, Kehr B, Kristmundsdóttir S, Zink F, Hjartarson E, et al. Parental influence on human germline de novo mutations in 1,548 trios from Iceland. *Nature*. 2017;549(7673):519–22.
- Kurahashi H, Bolor H, Kato T, Kogo H, Tsutsumi M, Inagaki H, et al. Recent advance in our understanding of the molecular nature of chromosomal abnormalities. *J Hum Genet*. 2009;54(5):253–60.
- Stuppia L, Antonucci I, Palka G, Gatta V. Use of the MLPA assay in the molecular diagnosis of gene copy number alterations in human genetic diseases. *Int J Mol Sci*. 2012;13(3):3245–76.



# SCIENTIFIC REPORTS



OPEN

## Genotype determination of the *OPN1LW/OPN1MW* genes: novel disease-causing mechanisms in Japanese patients with blue cone monochromacy

Satoshi Katagiri<sup>1</sup>, Maki Iwasa<sup>2</sup>, Takaaki Hayashi<sup>1,3</sup>, Katsuhiko Hosono<sup>4</sup>, Takahiro Yamashita<sup>5</sup>, Kazuki Kuniyoshi<sup>6</sup>, Shinji Ueno<sup>7</sup>, Mineo Kondo<sup>8</sup>, Hisao Ueyama<sup>9</sup>, Hisakazu Ogita<sup>9</sup>, Yoshinori Shichida<sup>5</sup>, Hidehito Inagaki<sup>10</sup>, Hiroki Kurahashi<sup>10</sup>, Hiroyuki Kondo<sup>11</sup>, Masahito Ohji<sup>2</sup>, Yoshihiro Hotta<sup>4</sup> & Tadashi Nakano<sup>1</sup>

Blue cone monochromacy (BCM) is characterized by loss of function of both *OPN1LW* (the first) and *OPN1MW* (the downstream) genes on the X chromosome. The purpose of this study was to investigate the first and downstream genes in the *OPN1LW/OPN1MW* array in four unrelated Japanese males with BCM. In Case 1, only one gene was present. Abnormalities were found in the promoter, which had a mixed unique profile of first and downstream gene promoters and a  $-71A > C$  substitution. As the promoter was active in the reporter assay, the cause of BCM remains unclear. In Case 2, the same novel mutation, M273K, was present in exon 5 of both genes in a two-gene array. The mutant pigments showed no absorbance at any of the wavelengths tested, suggesting that the mutation causes pigment dysfunction. Case 3 had a large deletion including the locus control region and entire first gene. Case 4 also had a large deletion involving exons 2–6 of the first gene. As an intact LCR was present upstream and one apparently normal downstream gene was present, BCM in Case 4 was not ascribed solely to the deletion. The deletions in Cases 3 and 4 were considered to have been caused by non-homologous recombination.

The human retina contains three types of cone photoreceptors: long-wavelength sensitive cones (L cones), medium-wavelength sensitive cones (M cones), and short-wavelength sensitive cones (S cones). These cone photoreceptors express respective visual pigments, L, M, and S opsins. Among these, the genes encoding L opsin (*OPN1LW*, OMIM; \*300822) and M opsin (*OPN1MW*, OMIM; \*300821) are present in tandem on the human X chromosome<sup>1,2</sup>, forming an L/M pigment gene array. In individuals with normal color vision, the first gene in the array is an L gene, and the downstream (the second and later) gene(s) is/are M gene(s). Abnormalities in the array are reportedly associated with protan and deutan color vision deficiencies<sup>3</sup>, blue cone monochromacy (BCM)<sup>4</sup>, and Bornholm eye disease<sup>5</sup>.

<sup>1</sup>Department of Ophthalmology, The Jikei University School of Medicine, Tokyo, Japan. <sup>2</sup>Department of Ophthalmology, Shiga University of Medical Science, Shiga, Japan. <sup>3</sup>Department of Ophthalmology, Katsushika Medical Center, The Jikei University School of Medicine, Tokyo, Japan. <sup>4</sup>Department of Ophthalmology, Hamamatsu University School of Medicine, Shizuoka, Japan. <sup>5</sup>Department of Biophysics, Graduate School of Science, Kyoto University, Kyoto, Japan. <sup>6</sup>Department of Ophthalmology, Kindai University Faculty of Medicine, Osaka, Japan. <sup>7</sup>Department of Ophthalmology, Nagoya University Graduate School of Medicine, Aichi, Japan. <sup>8</sup>Department of Ophthalmology, Mie University Graduate School of Medicine, Mie, Japan. <sup>9</sup>Department of Biochemistry and Molecular Biology, Shiga University of Medical Science, Shiga, Japan. <sup>10</sup>Division of Molecular Genetics, Institute for Comprehensive Medical Science, Fujita Health University, Aichi, Japan. <sup>11</sup>Department of Ophthalmology, University of Occupational and Environmental Health, Fukuoka, Japan. Satoshi Katagiri and Maki Iwasa contributed equally to this work. Correspondence and requests for materials should be addressed to T.H. (email: [taka@jikei.ac.jp](mailto:taka@jikei.ac.jp)) or H.U. (email: [datt@belle.shiga-med.ac.jp](mailto:datt@belle.shiga-med.ac.jp))

BCM (OMIM; #303700) is a rare congenital color vision deficiency with an X-linked inheritance pattern<sup>4,6</sup>. Cases of BCM typically present with severely impaired color discrimination, reduced visual acuity, nystagmus, photophobia, and diminished L/M cone function despite retention of rod and blue cone function<sup>6,7</sup>. The dysfunction in both L and M cones in BCM is reportedly caused by one of the three genotypes. One genotype involves deletion of the locus control region (LCR)<sup>4,8–12</sup>, which is located upstream of the L/M gene array (−3,681 to −3,021 from the cap site of the first gene) and believed to be involved in the mutually exclusive expression of L and M genes<sup>13,14</sup>. Therefore, neither gene is expressed in the absence of the LCR. Another genotype involves a deleterious mutation in a single-gene array (either the L or M gene present alone in the array). The derivation of this genotype has two obvious steps: first, non-homologous recombination between the L and M genes to form a single-gene array followed by an inactivating mutation in the single gene (reverse order is also possible). The most common mutation is C203R<sup>4,8,15–17</sup>, but other mutations, such as P307L<sup>8</sup>, R247X<sup>8</sup>, and deletion of exon 2<sup>16</sup>, have also been documented. The LIAVA haplotype in exon 3, which affects splicing<sup>18</sup>, was also reported in a single L gene<sup>17</sup>. The third genotype involves inactivating mutations in both the L and M genes. Although the C203R mutation has been documented in this genotype<sup>8,16,19</sup>, the LIAVA haplotype (or a very similar haplotype) in exon 3 of both genes seems to be frequent<sup>17,20</sup>.

Although little is known about the prevalence of BCM in the Japanese population, to date, only two BCM families have been described in the literature, demonstrating the mechanism of deletion of the LCR in both families<sup>11,21</sup>.

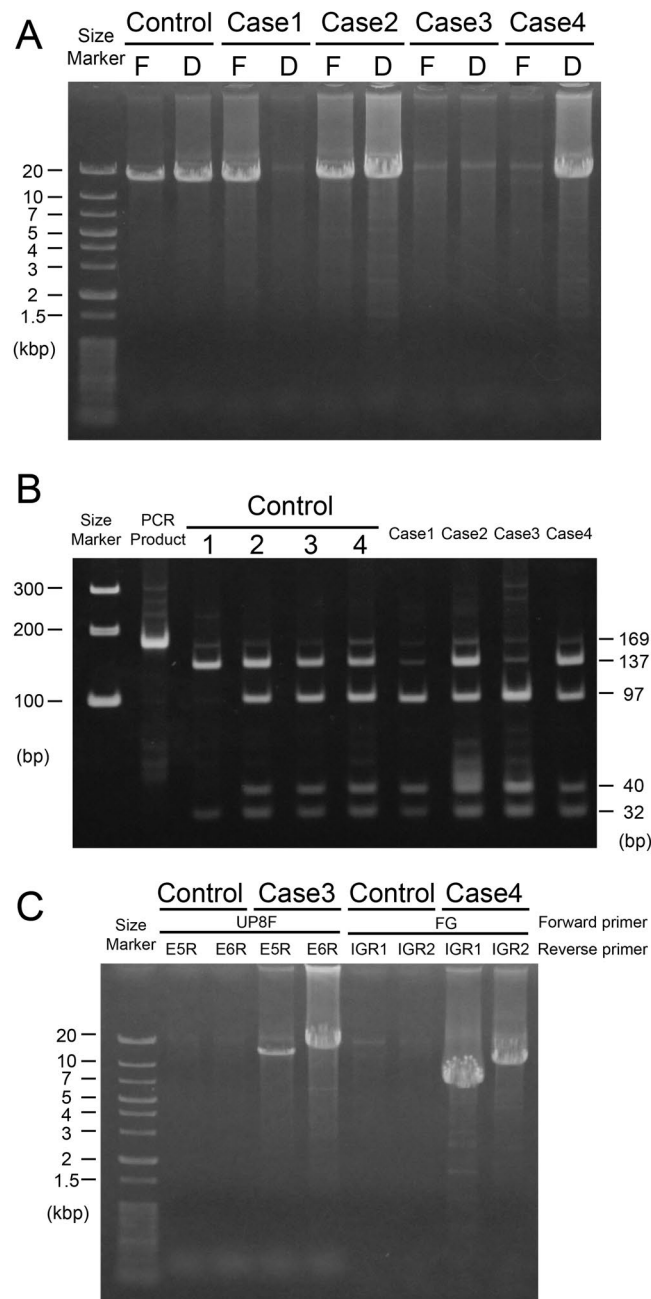
In the current study, the L/M pigment gene arrays in four unrelated Japanese males with BCM were analyzed. The purpose of this study was to investigate their genotypes in the L/M pigment gene array, which could be categorized into one of the three above-mentioned genotypes, but others were unreported mechanisms and differed from each other.

## Results

**Case 1 (JU#1299).** Long-range polymerase chain reaction (PCR) was successful for the first gene but not for downstream gene(s) (Fig. 1A). Promoter analysis of gene number, by contrast, showed that the first gene promoter was absent (only the downstream gene promoter was detected) (Fig. 1B). From the results of repeated long-range PCR analysis of downstream genes, we concluded that this subject had a single gene (no downstream genes) in the array. The LCR was present upstream of the single gene, and its sequence had no aberrations. The curious result that this subject had only the downstream gene promoter (Fig. 1B) was later found to be due to the unique sequence of the promoter. The first and downstream gene promoters differ by 14 nucleotides, but in the promoter region of the single gene, 8 upstream sites were associated with the first gene, whereas the other 6 sites were a random mixture of these nucleotides (Fig. 2). Moreover, the promoter had a −71 A > C substitution, which has been reported to be associated with deutan color vision deficiency<sup>22</sup> (Fig. 2). The G at the −30 position in the <sup>−34</sup>GCCGGT<sup>−29</sup> sequence (the number is from the cap site of the first gene) in the promoter analysis indicated that the first gene promoter was absent (Fig. 1B). The −30 (A or G) site discriminates the promoters of the first and downstream genes by *Cfr*10I (recognition sequence = RCCGGY). Our conclusion was that Case 1 had a single gene array because the sequencing of the promoter showed only one curious pattern as shown in Fig. 2. The luciferase activity of the promoter of the single gene was more than twice that of the usual first gene promoter in the reporter assay (Supplementary Fig. S1). No abnormalities were found in exons 1–6 and their adjacent introns in nucleotide sequencing. Exons 2–5 were M type, and the haplotype of exon 3 was MVAIA rather than LIAVA (Table 1). The curious promoter found in Case 1 has not been reported previously.

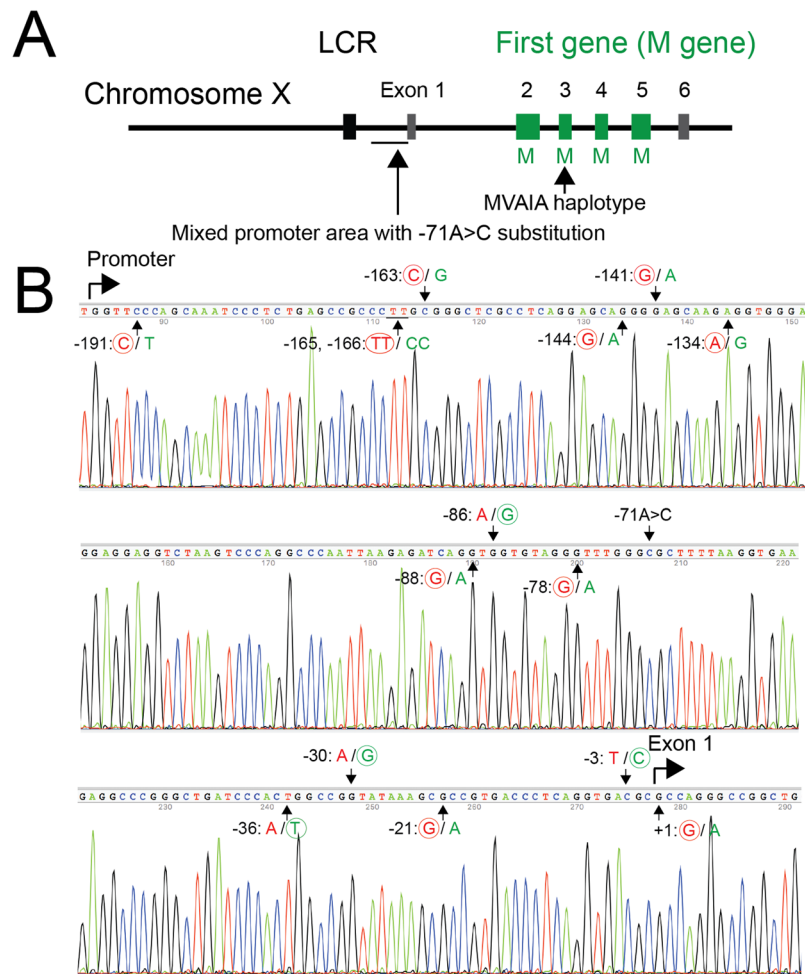
**Case 2 (JU#1311, KINKI-125-70).** Products of both the first and downstream genes were obtained using long-range PCR (Fig. 1A). Promoter analysis showed that the subject had a 2-gene array (Fig. 1B). Both genes had the same missense mutation (c.818 T > A, M273K) in exon 5 (Fig. 3A,B). The chromosome positions (GRCh38.p7) of the mutation are 154,156,367 (L gene) and 154,193,481 (M gene). The M273K mutation has not been reported previously, and not found in the Single Nucleotide Polymorphism Database (<https://www.ncbi.nlm.nih.gov/projects/SNP/>), Genome Aggregation Database (<http://gnomad.broadinstitute.org/>), Exome Aggregation Consortium (<http://exac.broadinstitute.org/>) and Human Genetic Variation Database (<http://www.hgvd.genome.med.kyoto-u.ac.jp/>). The analysis of the recombinant proteins of M273K mutants revealed that the opsin with the M273K mutation was significantly detectable in the Western blot and cultured cells (Supplementary Fig. S2) but showed no absorbance at any of the wavelengths tested after reconstitution with 11-*cis*-retinal (Fig. 3C). These results indicated that the M273K mutation in both genes results in dysfunctional opsin protein, probably because of a lack of the ability to bind to 11-*cis*-retinal. We therefore ascribed the BCM phenotype in this subject to the mutation. In the first gene, exons 2, 3, and 4 were L type, exon 5 was M type, and the haplotype of exon 3 was LV AIS. In the downstream gene, exons 2–5 were M type, and the haplotype of exon 3 was MVAIA (Table 1).

**Case 3 (JU#1318, MIE-050-0071).** Long-range PCR was unsuccessful for both the first and downstream genes (Fig. 1A). Promoter analysis for determining gene number showed that the first gene promoter was absent (only the downstream gene promoter was detected) (Fig. 1B). Amplification of the LCR was also unsuccessful, indicating a deletion including both the LCR and first gene. To determine the exact deletion breakpoints, 11 sets of PCR primers were designed to cover the sequence −53,930 to −9,320 (number is from the cap site of the first gene) (NT\_025965.12: 707,760 to 752,370). PCR products were obtained when using the UP8F and UP8R pair (Supplementary Table S1) but not the UP9F and UP9R pair (Supplementary Table S1), suggesting that the upstream breakpoint of the deletion was between −32,015 and −28,150 (NT\_025965.12: 729,665 to 733,530). Long-range PCR using the primer sets UP8F/E5R and UP8F/E6R was successful in this subject but not in a color-normal control subject (Fig. 1C). According to the human genome database (NT\_025965.12), the distance between UP8F- and E5R-corresponding regions and UP8F- and E6R-corresponding regions were 81,131 bp and



**Figure 1.** Long-range PCR and promoter analysis. (A) First and downstream genes in the L/M gene array were amplified separately by long-range PCR. The control was a color-normal subject having both the first and downstream genes. F, first gene; D, downstream gene(s). Thin bands of approximately 20 kb are not amplified products but the templates (genomic DNA, usually approximately 100 ng per reaction). (B) Promoter analysis of gene number. Promoters were amplified by PCR using primers common to the first and downstream genes. PCR products (169 bp) digested with *Cfr*10I were loaded onto a polyacrylamide gel. Controls 1–4 have gene numbers 1–4, respectively<sup>28</sup>. (C) Long-range PCR beyond the deletion. Combinations of the primers UP8F/E5R and UP8F/E6R were used for long-range PCR in the control and Case 3. Combinations of the primers FG/IGR1 and FG/IGR2 were used for long-range PCR in the control and Case 4.

83,925 bp, respectively (the E5R- and E6R-corresponding regions are those of the downstream gene), which were too far for long-range PCR. In Case 3, however, due to the large deletion including the first gene, the distances between the regions had been reduced to about 12 kbp and 15 kbp, respectively, and therefore, long-range PCR products were obtained in this case (Fig. 1C). The 15-kbp product contained not only exons 1–6 (exons 2, 4, and 5 were L type, exon 3 was M type, with the haplotype of MVAIA) (Table 1) but also the downstream gene promoter. By sequencing the 15-kbp product using the UP12F primer (Supplementary Table S1), the upstream breakpoint of the deletion was determined to be somewhere in the sequence <sup>-31,241</sup>GAACTCCTGACCTCAGG<sup>-31,225</sup> (the number is from the cap site of the first gene) (NT\_025965.12: 730,439 to 730,455), and the downstream breakpoint



**Figure 2.** Genotype of Case 1. **(A)** Overview of the genotype of Case 1. Case 1 had an intact LCR and a single M gene array in which no aberrations were found. The promoter regions had a unique profile including a  $-71 A > C$  substitution. **(B)** The promoter of the single M-gene array. Black arrows indicate the 14 nucleotides differing between first and downstream genes and the  $-71 A > C$  substitution. At each position, the usual nucleotide of the first gene promoter is shown on the left side in red and that of the downstream gene promoter is shown on the right side in green. The nucleotides in Case 1 are circled.

was determined to be somewhere in the sequence  $^{-407}GAACTCCTGACCTCAGG^{-391}$  (the number is from the cap site of the downstream gene) (NT\_025965.12: 799,682 to 799,698) (Fig. 4). The reason why long-range PCR was unsuccessful for the downstream gene (Fig. 1A) is that the deletion includes the region corresponding to the forward primer for long-range PCR, DG ( $-748$  to  $-728$  from the cap site of the downstream gene). The estimated size of the deletion was 69,243 bp. The long-range PCR products in Fig. 1C were calculated to be exactly 11,888 bp and 14,682 bp. As a LCR is reportedly necessary for the expression of L/M genes<sup>13,14</sup>, the BCM phenotype in this subject was ascribed to the deletion.

**Case 4 (JU#1368, Nagoya-140).** Long-range PCR was successful for downstream gene(s) but not for the first gene (Fig. 1A). Promoter analysis of gene number showed a 1:1 ratio for the first and downstream genes (Fig. 1B). As the first gene promoter was shown to be present, the FG primer–corresponding region (upstream of the promoter) should also be present. PCR analysis using combinations of the FG primer and various intragenic reverse primers revealed that the upstream breakpoint of the deletion was within intron 1 (between the primer I1R1– and primer I1R2–corresponding regions) and that the deletion expanded beyond exon 6. The failure of long-range PCR for the first gene was ascribed to the absence of exon 6 (primer E6R corresponds to exon 6). To determine the downstream breakpoint of the deletion, 15 reverse primers specific to the intergenic region (between the first and downstream genes) were designed for long-range PCR. PCR products were obtained for two primer pairs (FG/IGR1, and FG/IGR2) in this subject but not in the control (Fig. 1C). According to the human genome database (NT\_025965.12), the distances between the FG- and IGR1-corresponding regions and between the FG- and IGR2-corresponding regions were 30,137 bp and 34,286 bp, respectively, which were too long for long-range PCR. In Case 4, however, due to the large deletion, the distances had been reduced to approximately 7 kbp and 11 kbp, respectively, and therefore, long-range PCR products were obtained (Fig. 1C).

		Exon 2				Exon 3								Exon 4					Exon 5									Haplotype in Exon 3*		
Reference	Nucleotide position	194	300	331	347	453	457	465	511	513	521	532	538	689	697	698	699	706	820	823	825	828	830	835	849	853	888	892	926	
	L gene	C	A	A	C	G	C	G	G/A	G/T	C/T	A/G	T	T	G	C	T	A	A	T	T	G	A	G	C	A	T	G	A	
	M gene	T	G	G	A	A	A	C					G	C	A	G	C	G	G	C	G	A	T	T	A	G	C	C	T	
	Amino acid Position	65	100	111	116	151	153	155	171-1	171-3	174	178	180	230	233-1	233-2	233-3	236	274	275-1	275-3	276	277	279	283	285	296	298	309	
	L gene	T	L	I	S	R	L	V					S	I	A			M	I	F		A	Y	V	P	T	G	A	Y	
	M gene	I	L	V	Y	R	M	V	V/I		A/V	I/V	A	T	S			V	V	L		A	F	F	P	A	G	P	F	
Case 1 First (single) gene	Nucleotide	T	G	G	A	A	A	C	G	G	C	A	G	C	A	G	C	G	G	C	G	A	T	T	A	G	C	C	T	MVAIA
	Amino acid	I	L	V	Y	R	M	V	V		A	I	A	T	S			V	V	L		A	F	F	P	A	G	P	F	
Case 2 First gene	Nucleotide	C	A	A	C	G	C	G	G	G	C	A	T	T	G	C	T	A	G	C	G	A	T	T	A	G	C	C	T	LVAIS
	Amino acid	T	L	I	S	R	L	V	V		A	I	S	I	A			M	V	L		A	F	F	P	A	G	P	F	
Case 2 Downstream gene	Nucleotide	T	G	G	A	A	A	C	G	G	C	A	G	C	A	G	C	G	G	C	G	A	T	T	A	G	C	C	T	MVAIA
	Amino acid	I	L	V	Y	R	M	V	V		A	I	A	T	S			V	V	L		A	F	F	P	A	G	P	F	
Case 3 Downstream gene (First gene was deleted)	Nucleotide	C	A	A	C	A	A	C	G	G	C	A	G	T	G	C	T	A	A	T	T	G	A	G	C	A	T	G	A	MVAIA
	Amino acid	T	L	I	S	R	M	V	V		A	I	A	I	A			M	I	F		A	Y	V	P	T	G	A	Y	
Case 4 Downstream gene (First gene was deleted)	Nucleotide	T	G	G	A	G	C	G	G	G	T	G	G	C	A	G	C	G	G	C	G	A	T	T	A	G	C	C	T	LVVVA
	Amino acid	I	L	V	Y	R	L	V	V	V	V	V	A	T	S			V	V	L		A	F	F	P	A	G	P	F	

**Table 1.** Nucleotides in each gene of the four cases. The positions of nucleotides different between wild-type L and M genes, and polymorphic nucleotide positions 511, 513, 522 and 532 as well, are shown. \*Haplotype in Exon 3 was determined by amino acid residues at 153, 171, 174, 178, and 180.

The IGR1 primer corresponds to the region +14,867 to +14,887 (the number is from the stop codon in exon 6) (NT\_025965.12: 791,231 to 791,251), and the IGR2 primer corresponds to the region +15,232 to +15,252 (NT\_025965.12: 767,324 to 767,329). Using the 11-kbp PCR product and primer I1F (Supplementary Table S1), the upstream breakpoint of the deletion was determined to be somewhere in <sup>+5,492</sup>TGAGCC<sup>+5,497</sup> (the number is from 5' splice site of intron 1 of the first gene) (NT\_025965.12: 767,324 to 767,329), and the downstream breakpoint was determined to be somewhere in <sup>+14,714</sup>TGAGCC<sup>+14,719</sup> (the number is from the stop codon in exon 6) (NT\_025965.12: 790,713 to 790,718) (Fig. 5). The estimated size of the deletion was 23,389 bp. The long-range PCR products shown in Fig. 1C were calculated to be exactly 6,748 bp and 10,897 bp. PCR confirmed the presence of the LCR, and its sequence had no aberrations. The downstream gene had no abnormalities in the promoter, exons 1–6, or their adjacent introns. Exons 2–5 were M type, and the haplotype of exon 3 was LVVVA (Table 1). As the downstream gene was apparently normal, the cause of BCM in this subject could not be confidently determined.

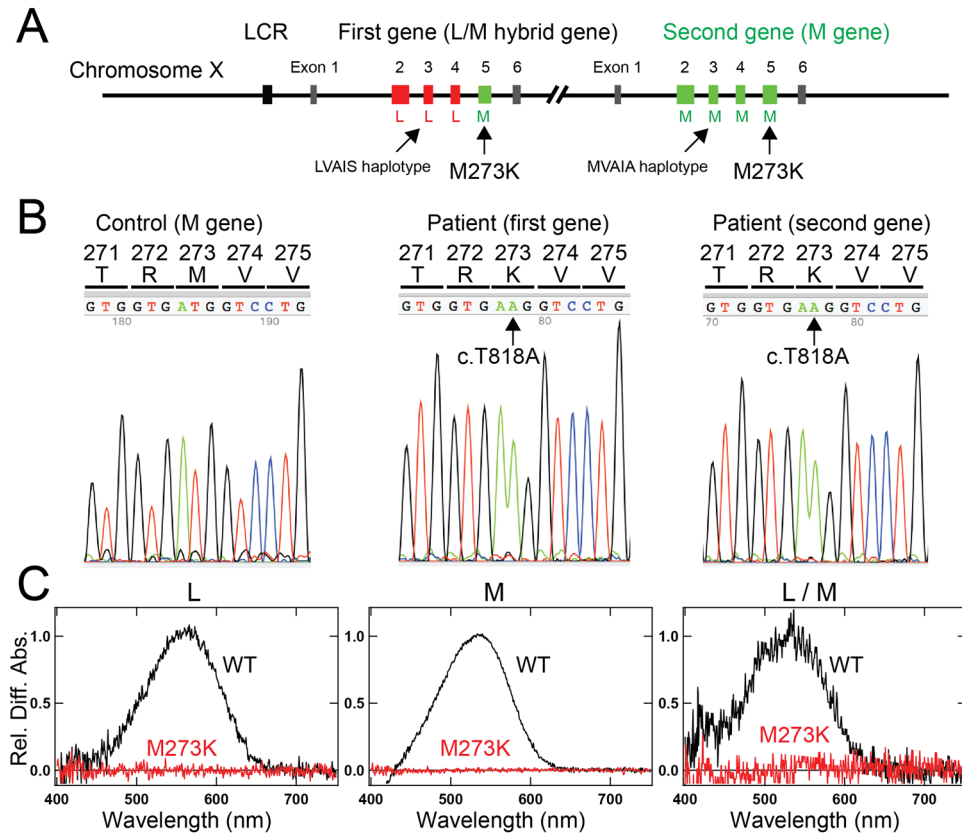
## Discussion

In this study, we reported the results of gene analyses in four cases of BCM. Their genotypes were unreported and different from each other.

Case 1 had a single-gene array having a curious promoter sequences with a  $-71 A > C$  substitution. The  $-71 A > C$  substitution was reported to be associated with deutan color vision deficiency due to decreased promoter activity<sup>22</sup>. We hypothesized that the  $-71 A > C$  substitution causes dysfunction of the single gene. However, rather than being low, the activity of the promoter in Case 1 was more than twice that of the control in the reporter assay. It is reported that not only the LCR but also normal promoters in the L/M gene array were essential for expression of both L and M genes<sup>13,14</sup>. The LCR was not contained in the constructs for our reporter assay system. Although the dysfunction of the single opsin gene in Case 1 remains unclear, it might be possible that the curious promoter sequences (Fig. 2) might interfere with the LCR binding to the promoter.

Case 2 had a two-gene array, and both genes had a novel missense mutation (M273K) that would cause dysfunction of both gene products (Fig. 3C). The C203R mutation in exon 4 reportedly causes loss of function of the L/M gene products; this mutation causes deutan color vision deficiency when present in the M gene<sup>23,24</sup> and BCM when present in both the L and M genes<sup>8,16,19</sup>. The occurrence of C203R mutation in both the L and M genes was explained by gene conversion<sup>19</sup> (i.e., transfer of the C203R mutation present in the downstream M gene to the first L gene). The deleterious LIAVA haplotype in exon 3 of both the L and M genes was also explained by gene conversion<sup>20</sup>. Also, the occurrence of the M273K mutation in both the first and second genes (Fig. 3A) might be explained by the same mechanism of gene conversion seen in the C203R mutation<sup>19</sup>. Otherwise, because the first gene in this subject had L-type exons 2–4 and M-type exon 5 (with the mutation), we developed the



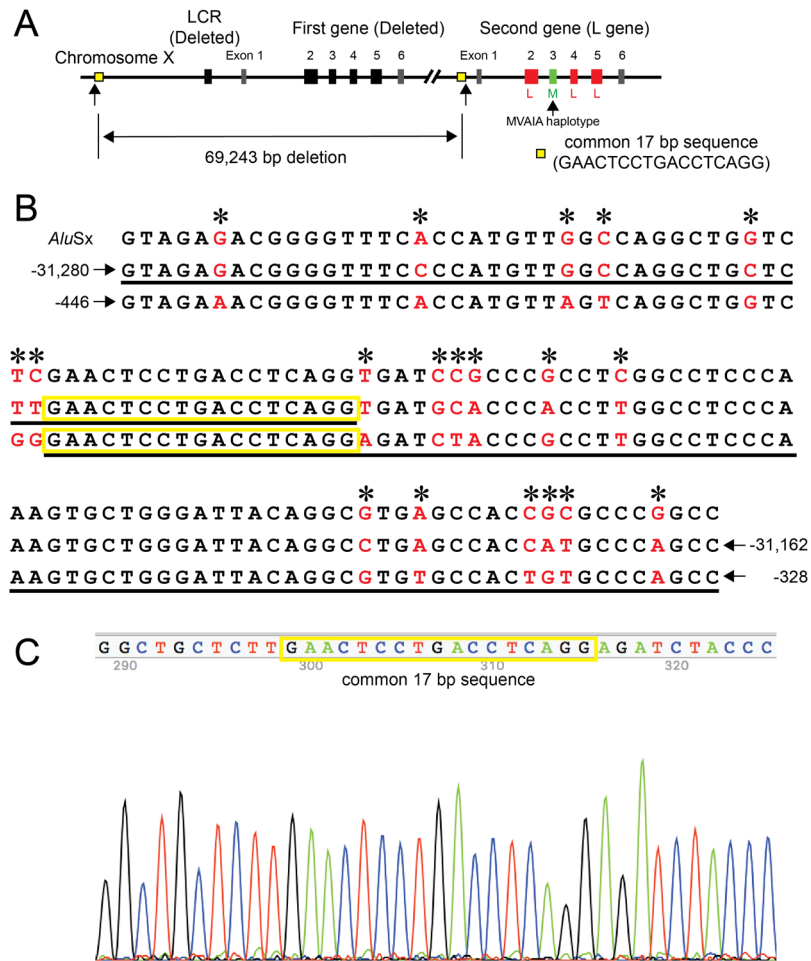


**Figure 3.** Genotype of Case 2. (A) Overview of the genotype of Case 2. Case 2 had an intact LCR and two genes. The first gene had exons 2–4 of L type, exon 3 with LVAIA haplotype, and exon 5 of M type. The second gene had exons 2–5 of M type and exon 3 with MVAIA haplotype. Both genes had the same missense mutation (c.818 T > A, M273K) in exon 5. (B) Partial sequence data around the missense mutation (c.818 T > A, M273K) in exon 5 in the control and two genes of Case 2. (C) Opsin reconstitution experiments. L, L opsin in which exons 2–5–derived amino acid sequences are all L type; M, M opsin in which exons 2–5–derived amino acid sequences are all M type, as in the product of the second gene of Case 2; L/M, M opsin in which exons 2–4–derived amino acid sequences are L type but exon 5–derived amino acid sequence is M type, as in the product of the first gene of Case 2. WT, wild-type opsin; M273K, mutant opsin with the M273K mutation. “Rel. Diff. Abs.” indicates relative difference absorption.

following hypothesis as the other alternative possibility. The M273K mutation occurred in the downstream M gene (the array was L-M\*; \*denotes the mutation), duplication of the M gene occurred (L-M\*-M\*), followed by non-homologous recombination (L/M\* hybrid-M\*) as shown in Fig. 3A. Duplication of the second gene was supported by the fact that (i) many color-normal individuals have multiple downstream M genes with the same nucleotide sequence<sup>25</sup> and (ii) the result that three (or more) downstream M genes had the same 11-bp deletion in a protanopia subject<sup>26</sup>. Case 2 had a genotype with two unique profiles; the M273K mutation was novel and present in both L and M genes.

Cases 3 and 4 showed large deletions of 62,934 bp including the LCR and 23,389 bp not including the LCR, respectively. The genotype of Case 3 was consistent with the known genotype of BCM. In Case 4, the first gene was obviously non-functional due to the absence of exons 2–6, but the downstream gene seemed to be functional, as no deleterious mutations were found in the promoter and exons, including their adjacent introns. The clinical phenotype of BCM in Case 4 indicated that the downstream gene was non-functional. The gene array revealed that the promoter of the second gene was directly connected to intron 1 of the first gene with absence of exons 2–6 (23,389 deletion) (Fig. 5). Although we cannot explain reasonable mechanisms underlying BCM in Case 4, the residual sequences (exon 1 and partial intron 1) of the first gene might impact on the promoter activity of the second gene, resulting in suppression of the second gene expression. The breakpoints of the deletion were within *Alu* elements in both cases, as in previously reported BCM cases<sup>11</sup>. Many reports have described deletions in the L/M gene array<sup>4,8–12,17</sup>. However, few studies have determined the exact breakpoints of the deletion at the nucleotide level<sup>4,11,16</sup>; the breakpoints have resulted from simple breakage and fusion<sup>4,16</sup> outside the repetitive sequence and simple breakage and fusion (and insertion) in *Alu* elements<sup>11</sup>. The breakpoints we determined differed from these; non-equal crossing-over occurred between the two *Alu* elements in the region of the same sequence (Figs 4 and 5).

The various haplotypes of five amino acid residues at positions 153 (L/M), 171 (V/I), 174 (A/V), 178 (I/V), and 180 (S/A) in exon 3 have been reported in subjects with normal color vision and subjects with color vision deficiencies<sup>18,20</sup>. The haplotypes have been roughly classified into four groups in terms of the magnitude of the



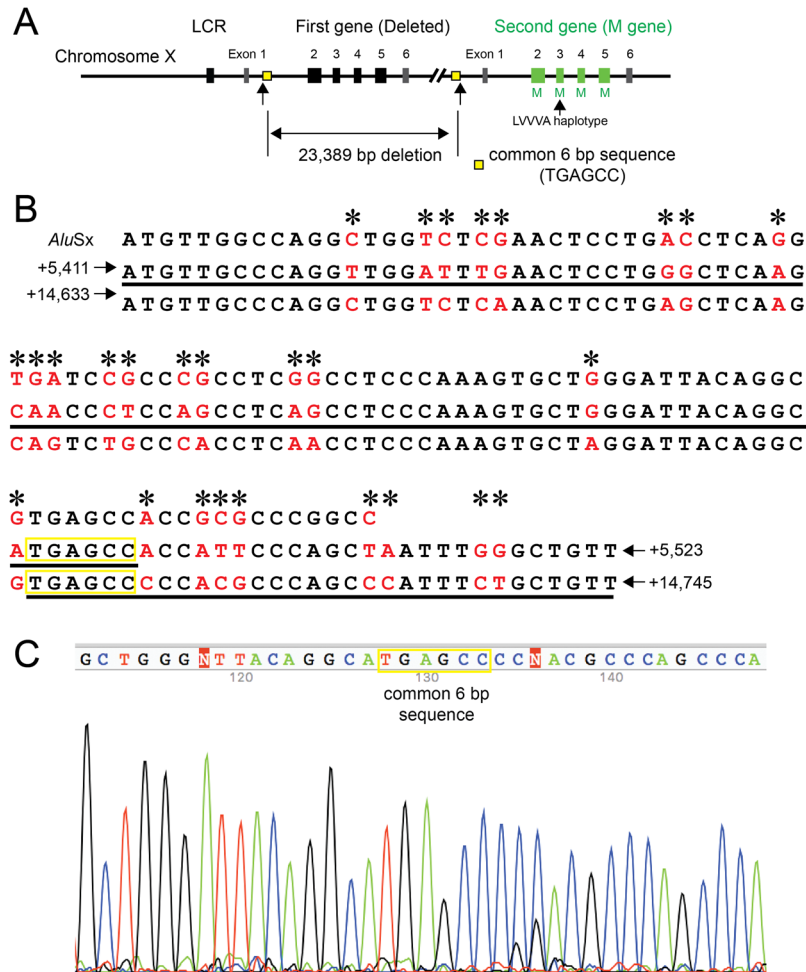
**Figure 4.** Genotype of Case 3. (A) Overview of the genotype of Case 3. Case 3 had a large deletion of 69,243 bp including the LCR and first gene. The remaining second gene was an L gene in which exons 2, 4, and 5 were L type. Exon 3 was M type with MVAIA haplotype. (B) Upper row shows the consensus sequence for the left monomer of the *Alu* element (complementary sequence of *AluSx*). Middle row shows a part of the upstream region of the first gene (number is from the cap site of the first gene). Lower row shows a part of the upstream region of the downstream gene (number is from the cap site of the downstream gene). The nucleotides differing among the regions are shown in red and by asterisks. The actual sequence obtained in Case 3 is underlined. (C) Partial sequence data around the breakpoint of the deletion. The breakpoint is somewhere in the common 17-bp sequence.

splicing defect<sup>20</sup>; highly deleterious haplotypes include LIAVA, MIAVA, and LVAVA; intermediately deleterious haplotypes include LIAIA, LIAVS, and MVAVA, minor deleterious haplotypes include LVAIA, LVAIS, MVAIA, and MVVVA, and the MVAIS haplotype exhibits no splicing defect. According to our data, the MVAIA and LVAIS haplotypes in Cases 1–3 would be expected to produce essentially correct splicing; however, the LVVVA haplotype in Case 4 could not be classified, as this haplotype was not described in the above-mentioned study<sup>20</sup>. We<sup>18</sup> and other researchers<sup>27</sup> examined the LVVVA haplotype using a mini-gene system and observed that the opsin mRNA retaining exon 3 was in clearly greater abundance than that lacking exon 3. Moreover, we reported one color-normal subject in which the exon 3 haplotype was LVVVA<sup>18</sup>. Based on these results, the BCM phenotype in Case 4 could not be ascribed to the LVVVA haplotype in the downstream gene.

In conclusion, we reported four novel and different genotypes in four unrelated Japanese patients with BCM. In two patients (Case 2 and Case 3), the genotypes were consistent with that of BCM (the same deleterious mutation in both opsin genes and deletion of the LCR), but in the other two cases (Case 1 and Case 4), the cause of BCM could not be clearly determined, although the patients exhibited very unique genotypes.

## Methods

The protocol for this study was approved by the Institutional Review Boards of The Jikei University of Medicine, Shiga University of Medical Science, Hamamatsu University School of Medicine, Kyoto University, Kindai University Faculty of Medicine, Nagoya University Graduate School of Medicine, Mie University Graduate School of Medicine, and Fujita Health University. The protocol adhered to the tenets of the Declaration of Helsinki, and informed consent was obtained from all participants.



**Figure 5.** Genotype of Case 4. **(A)** Overview of the genotype of Case 4. Case 4 had a large deletion of 23,389 bp including exons 2–6 of the first gene. The intact LCR and second gene were present. The second gene had exons 2–5 of M type and exon 3 with LVVVA haplotype. **(B)** Upper row shows the consensus sequence for the left monomer of the *Alu* element (complementary sequence of *AluSx*). Middle row shows a part of intron 1 of the first gene (the number is from the 5' splice site of intron 1). Lower row shows a part of the intergenic region (the number is from the stop codon in exon 6 of the first gene). The nucleotides differing among the regions are shown in red and by asterisks. The actual sequence obtained in Case 4 is underlined. **(C)** Partial sequence data around the breakpoint of the deletion. The breakpoint is somewhere in the common 6-bp sequence.

**Participants.** We recruited four unrelated Japanese male patients with BCM, whose diagnosis of BCM was made according to the findings reported<sup>6,21</sup>. In brief, all participants exhibited clinical findings of BCM, such as decreased visual acuity, severely impaired color discrimination in color vision tests, diminished L and M cone functions but retained S cone function in electroretinography, and an X-linked inheritance pattern in the family history. The detailed clinical findings are summarized in Supplementary Table S2.

**Molecular genetic analysis.** Genomic DNA was extracted from leucocytes in venous blood samples using a Gentra Puregene blood kit (Qiagen, Hilden, Germany). First and downstream genes of the L/M visual pigment gene array were separately amplified by long-range PCR using a QIAGEN LongRange PCR kit (Qiagen). Primers FG and E6R were used for the first gene, and primers DG and E6R were used for downstream gene(s) (Supplementary Table S1). The position of primers used in this study are schematically shown in Supplementary Fig. S3. Primer E6R was common to both genes, but primers FG and DG were designed specifically for the upstream region of each gene. The cycling parameters were: 93 °C for 3 min; 10 cycles of 93 °C for 30 s, 62 °C for 30 s, and 68 °C for 15 min; then 18 cycles of 93 °C for 30 s, 62 °C for 30 s, and 68 °C for 15 min, with a 20-s increment per cycle. The resulting PCR products were used as templates for sequencing the ‘promoter + exon 1’ and exons 2–6, including their adjacent introns, using a BigDye Terminator v3.1 Cycle Sequencing kit (Thermo Fisher Scientific, Waltham, MA, USA) and ABI 3130xl sequencer (Thermo Fisher Scientific). The primer pairs used for sequencing are listed in Supplementary Table S1.

The LCR, which is located about 3.5 kb upstream of the first gene, was amplified by PCR using primers LCRF and LCRR, and its nucleotide sequence was then determined. When deletion was suspected, multiple sets of primers were designed for PCR to determine the exact deletion breakpoints.

Array gene number was determined by promoter analysis, as previously described<sup>22</sup>. Briefly, the promoters were amplified by PCR using Takara *Taq* DNA polymerase (Takara Bio Inc., Kusatu, Japan) and the primer pair common to both genes (Supplementary Table S1). The 169-bp PCR product was digested with *Cfr*101 (Takara) and analyzed by electrophoresis on a polyacrylamide gel. The first gene promoter was expected to yield two DNA fragments (137 bp and 32 bp), whereas the downstream gene promoter was expected to yield three DNA fragments (97 bp, 40 bp, and 32 bp). Gene number was estimated from the fluorescence ratio ( $[(137 \text{ bp} + 97 \text{ bp})/137 \text{ bp}]$ ). Gene Ladder Wide 2 was used as the size marker (Nippon Gene Co., Ltd., Toyama, Japan). Genomic DNAs from four subjects in which the L/M gene number was confirmed to be 1–4 by pulsed-field gel electrophoresis and Southern blotting<sup>28</sup> were used as control templates.

**Promoter assay.** To evaluate the activity of the promoter region, a luciferase reporter assay was performed as previously described<sup>29</sup>. Briefly, the promoter region of interest (–190 to +41 from the cap site of the gene) was amplified by PCR using Phusion High-Fidelity DNA polymerase (New England BioLabs, Ipswich, MA, USA) and restriction site–tagged primers (*Nhe*I site upstream and *Hind*III site downstream). The PCR product was cloned between the *Nhe*I and *Hind*III sites of a luciferase reporter plasmid, pGL4.17 (Promega Corp., Fitchburg, WI, USA). The resulting plasmid was transfected into WERI Rb1 cells using X-tremeGENE 9 DNA transfection reagent (Sigma-Aldrich, St. Louis, MO, USA). Two days after transfection, the cells were collected and lysed using PicaGene Cell Culture Lysis Reagent Luc (Wako Chemicals, Osaka, Japan). Luciferase activity was measured using a luminometer (Lumicounter NU-2500, Microtech Co., Ltd., Funahashi, Japan) and PicaGene Luminescence kit (Wako). Transfection efficiency was monitored in cells co-transfected with a  $\beta$ -galactosidase–encoding plasmid (Promega).

**Analysis of M273K mutant pigments.** The cDNAs of human L and M pigments and respective hybrid pigment were tagged with the epitope sequence of the anti-bovine rhodopsin monoclonal antibody Rho1D4 (ETSQVAPA) at the C terminus and were inserted into the mammalian expression vector pMT4<sup>30</sup>. cDNAs harboring the mutation M273K were constructed using an In-Fusion cloning kit (Takara). For the spectral analysis, the plasmid DNA was transfected into HEK293 cells using the calcium-phosphate method<sup>31</sup>. After 2 days of incubation, the cells were collected by centrifugation and supplemented with 11-*cis*-retinal in buffer A (50 mM Hepes [pH 6.5] and 140 mM NaCl) to reconstitute the pigments. The reconstituted pigments were extracted using 0.75% CHAPS and 1 mg/mL phosphatidylcholine in buffer A. Absorption spectra of the extracted pigments were recorded at 0 °C using a Shimadzu UV-2450 spectrophotometer. The pigments were irradiated with orange light through an O58 cutoff filter (Toshiba, Tokyo, Japan) for 1 min. Difference spectra were calculated from spectra recorded before and after irradiation. For the western blot analysis, extracts from pigment-transfected or mock-transfected HEK293 cells were subjected to SDS-PAGE, transferred onto a polyvinylidene difluoride membrane, and probed with Rho1D4. Immunoreactive proteins were detected by ECL Western Blotting Detection Reagents (GE Healthcare, United Kingdom) and visualized by a luminescent image analyzer (LAS 4000mini, GE Healthcare). For the fluorescence microscopy analysis, pigment-transfected or mock-transfected HEK293 cells were seeded onto poly-L-lysine coated coverslips. After 24 h incubation, cells were fixed in cooled methanol for 5 min. After fixation, cells were washed three times in PBS and were incubated overnight with primary antibody, Rho1D4, in 10% normal goat serum at room temperature. Cells were washed three times in PBS and were incubated for 1 h with secondary antibody, Alexa Fluor 488 anti-mouse IgG, in 10% normal goat serum at room temperature. Cells were washed a final time and were mounted onto slides with home-made aqueous mounting media consisting of glycerol and polyvinyl alcohol.

## References

- Nathans, J., Thomas, D. & Hogness, D. S. Molecular genetics of human color vision: the genes encoding blue, green, and red pigments. *Science* **232**, 193–202 (1986).
- Vollrath, D., Nathans, J. & Davis, R. W. Tandem array of human visual pigment genes at Xq28. *Science* **240**, 1669–1672 (1988).
- Nathans, J., Piantanida, T. P., Eddy, R. L., Shows, T. B. & Hogness, D. S. Molecular genetics of inherited variation in human color vision. *Science* **232**, 203–210 (1986).
- Nathans, J. *et al.* Molecular genetics of human blue cone monochromacy. *Science* **245**, 831–838 (1989).
- McClements, M. *et al.* Variations in opsin coding sequences cause X-linked cone dysfunction syndrome with myopia and dichromacy. *Invest Ophthalmol Vis Sci* **54**, 1361–1369 (2013).
- Berson, E. L., Sandberg, M. A., Rosner, B. & Sullivan, P. L. Color plates to help identify patients with blue cone monochromatism. *Am J Ophthalmol* **95**, 741–747 (1983).
- Kellner, U. *et al.* Blue cone monochromatism: clinical findings in patients with mutations in the red/green opsin gene cluster. *Graefes Arch Clin Exp Ophthalmol* **242**, 729–735 (2004).
- Nathans, J. *et al.* Genetic heterogeneity among blue-cone monochromats. *Am J Hum Genet* **53**, 987–1000 (1993).
- Ayyagari, R. *et al.* Bilateral macular atrophy in blue cone monochromacy (BCM) with loss of the locus control region (LCR) and part of the red pigment gene. *Mol Vis* **5**, 13 (1999).
- Ayyagari, R. *et al.* Spectrum of color gene deletions and phenotype in patients with blue cone monochromacy. *Hum Genet* **107**, 75–82 (2000).
- Wang, C. *et al.* Novel *OPN1LW/OPN1MW* deletion mutations in 2 Japanese families with blue cone monochromacy. *Hum Genome Var* **3**, 16011 (2016).
- Yatsenko, S. A. *et al.* High-resolution microarray analysis unravels complex Xq28 aberrations in patients and carriers affected by X-linked blue cone monochromacy. *Clin Genet* **89**, 82–87 (2016).
- Wang, Y. *et al.* A locus control region adjacent to the human red and green visual pigment genes. *Neuron* **9**, 429–440 (1992).
- Smallwood, P. M., Wang, Y. & Nathans, J. Role of a locus control region in the mutually exclusive expression of human red and green cone pigment genes. *Proc Natl Acad Sci USA* **99**, 1008–1011 (2002).

15. Michaelides, M. *et al.* Blue cone monochromatism: a phenotype and genotype assessment with evidence of progressive loss of cone function in older individuals. *Eye (Lond)* **19**, 2–10 (2005).
16. Gardner, J. C. *et al.* Blue cone monochromacy: causative mutations and associated phenotypes. *Mol Vis* **15**, 876–884 (2009).
17. Gardner, J. C. *et al.* Three different cone opsin gene array mutational mechanisms with genotype-phenotype correlation and functional investigation of cone opsin variants. *Hum Mutat* **35**, 1354–1362 (2014).
18. Ueyama, H. *et al.* Unique haplotype in exon 3 of cone opsin mRNA affects splicing of its precursor, leading to congenital color vision defect. *Biochem Biophys Res Commun* **424**, 152–157 (2012).
19. Reyniers, E. *et al.* Gene conversion between red and defective green opsin gene in blue cone monochromacy. *Genomics* **29**, 323–328 (1995).
20. Buena-Atienza, E. *et al.* *De novo* intrachromosomal gene conversion from *OPN1MW* to *OPN1LW* in the male germline results in Blue Cone Monochromacy. *Sci Rep* **6**, 28253 (2016).
21. Terasaki, H. & Miyake, Y. Association of acquired color vision defects in blue cone monochromatism. *Jpn J Ophthalmol* **39**, 55–59 (1995).
22. Ueyama, H. *et al.* An A-71C substitution in a green gene at the second position in the red/green visual-pigment gene array is associated with deutan color-vision deficiency. *Proc Natl Acad Sci USA* **100**, 3357–3362 (2003).
23. Winderickx, J. *et al.* Defective colour vision associated with a missense mutation in the human green visual pigment gene. *Nat Genet* **1**, 251–256 (1992).
24. Jagla, W. M., Jägle, H., Hayashi, T., Sharpe, L. T. & Deeb, S. S. The molecular basis of dichromatic color vision in males with multiple red and green visual pigment genes. *Hum Mol Genet* **11**, 23–32 (2002).
25. Hayashi, S., Ueyama, H., Tanabe, S., Yamada, S. & Kani, K. Number and variations of the red and green visual pigment genes in Japanese men with normal color vision. *Jpn J Ophthalmol* **45**, 60–67 (2001).
26. Ueyama, H. *et al.* Analysis of L-cone/M-cone visual pigment gene arrays in Japanese males with protan color-vision deficiency. *Vision Res* **44**, 2241–2252 (2004).
27. Greenwald, S. H., Kuchenbecker, J. A., Rowlan, J. S., Neitz, J. & Neitz, M. Role of a dual splicing and amino acid code in myopia, cone dysfunction and cone dystrophy associated with L/M opsin interchange mutations. *Transl Vis Sci Technol* **6**, 2 (2017).
28. Ueyama, H., Tanabe, S., Muraki-Oda, S., Yamada, S. & Ohkubo, I. Protan color vision deficiency with a unique order of green-red as the first two genes of a visual pigment array. *J Hum Genet* **51**, 686–694 (2006).
29. Ueyama, H. *et al.* Analysis of introns and promoters of L/M visual pigment genes in relation to deutan color-vision deficiency with an array of normal gene orders. *J Hum Genet* **54**, 525–530 (2009).
30. Oprian, D. D., Molday, R. S., Kaufman, R. J. & Khorana, H. G. Expression of a synthetic bovine rhodopsin gene in monkey kidney cells. *Proc Natl Acad Sci USA* **84**, 8874–8878 (1987).
31. Yamashita, T., Nakamura, S., Tsutsui, K., Morizumi, T. & Shichida, Y. Chloride-dependent spectral tuning mechanism of L-group cone visual pigments. *Biochemistry* **52**, 1192–1197 (2013).

## Acknowledgements

We thank the patients for their participation in this study. We also thank Prof. R. S. Molday for the generous gift of a Rho1D4-producing hybridoma. This work was supported by grants from the Initiative on Rare and Undiagnosed Diseases for Adults (16ek0109151h0002 to Y.H.) and Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research (17K11447 to Y.H., 16K11284 to K.H., 25462711 to H.U., 17K11441 to H. Kondo, 25462738 to T.H., and 17K11434 to T.H.).

## Author Contributions

M.I. and H.U. performed the molecular genetic analyses in all four cases. K.H. and Y.H. performed part of the molecular genetic analysis in case 3. T.Y. and Y.S. performed analysis of M273K mutant pigments. S.K., M.I., T.H., and H.U. interpreted the data and wrote the manuscript. H.I., H. Kurahashi, H. Kondo, H.O., M.O., and T.N. assisted with data interpretation. S.K., T.H., K.K., S.U., and M.K. performed ophthalmic examinations at each institution. T.H. and H.U. designed and supervised the study. All authors have read and approved the final manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-29891-9>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018



## Twin pregnancy with chromosomal abnormalities mimicking a gestational trophoblastic disorder and coexistent foetus on ultrasound

Akiko Ohwaki, Haruki Nishizawa, Noriko Aida, Takema Kato, Asuka Kambayashi, Jun Miyazaki, Mayuko Ito, Makoto Urano, Yuka Kiriya, Makoto Kuroda, Masahiro Nakayama, Shin-Ichi Sonta, Kaoru Suzumori, Takao Sekiya, Hiroki Kurahashi & Takuma Fujii

To cite this article: Akiko Ohwaki, Haruki Nishizawa, Noriko Aida, Takema Kato, Asuka Kambayashi, Jun Miyazaki, Mayuko Ito, Makoto Urano, Yuka Kiriya, Makoto Kuroda, Masahiro Nakayama, Shin-Ichi Sonta, Kaoru Suzumori, Takao Sekiya, Hiroki Kurahashi & Takuma Fujii (2018): Twin pregnancy with chromosomal abnormalities mimicking a gestational trophoblastic disorder and coexistent foetus on ultrasound, *Journal of Obstetrics and Gynaecology*, DOI: [10.1080/01443615.2017.1401598](https://doi.org/10.1080/01443615.2017.1401598)

To link to this article: <https://doi.org/10.1080/01443615.2017.1401598>



[View supplementary material](#)



Published online: 09 Mar 2018.



[Submit your article to this journal](#)



Article views: 6



[View related articles](#)



[View Crossmark data](#)

CASE REPORT



## Twin pregnancy with chromosomal abnormalities mimicking a gestational trophoblastic disorder and coexistent foetus on ultrasound

Akiko Ohwaki<sup>a,b</sup>, Haruki Nishizawa<sup>a</sup>, Noriko Aida<sup>a</sup>, Takema Kato<sup>b</sup>, Asuka Kambayashi<sup>a</sup>, Jun Miyazaki<sup>a,b</sup>, Mayuko Ito<sup>a,b</sup>, Makoto Urano<sup>c</sup>, Yuka Kiriyama<sup>c</sup>, Makoto Kuroda<sup>c</sup>, Masahiro Nakayama<sup>d</sup>, Shin-Ichi Sonta<sup>e</sup>, Kaoru Suzumori<sup>e</sup>, Takao Sekiya<sup>a</sup>, Hiroki Kurahashi<sup>b</sup> and Takuma Fujii<sup>a</sup>

<sup>a</sup>Department of Obstetrics and Gynecology, Fujita Health University School of Medicine, Toyoake, Japan; <sup>b</sup>Division of Molecular Genetics, Institute for Comprehensive Medical Science, Fujita Health University, Toyoake, Japan; <sup>c</sup>Department of Diagnostic Pathology, Fujita Health University School of Medicine, Toyoake, Japan; <sup>d</sup>Department of Pathology and Laboratory Medicine, Osaka Medical Center and Research Institute for Maternal and Child Health, Izumi, Japan; <sup>e</sup>Fetal Life Science Center, Ltd, Nagoya, Japan

### Case report

Gestational trophoblastic disorder with a coexistent foetus occurs in 1 in 20,000–100,000 pregnancies (Wee and Jauniaux 2005) and mostly involves a partial hydatidiform mole with a live foetus and rarely a twin pregnancy with a complete hydatidiform mole and co-twin foetus (Gupta et al. 2015). Most cases of partial hydatidiform mole have triploidy with multiple structural anomalies and result in first trimester miscarriage (Toufaily et al. 2016). However, their management is complicated because the coexistent foetus is occasionally a normal healthy diploid foetus. Furthermore, this condition is often accompanied by severe complications such as hyperemesis, preeclampsia or thromboembolic disease (Matsui et al. 2000; Sebire et al. 2002). Thus, the diagnosis and management of gestational trophoblastic diseases with coexistent foetus are clinically important.

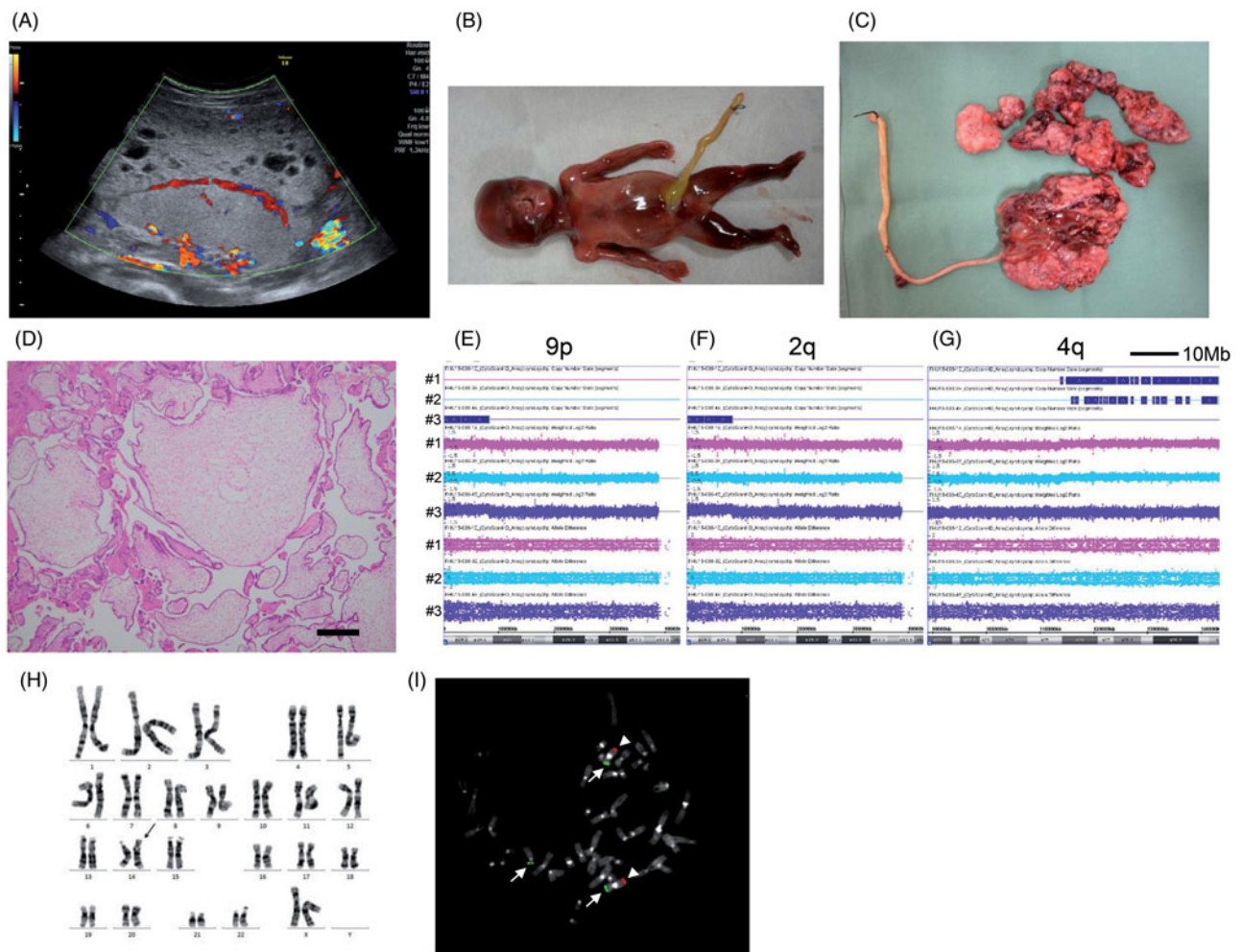
A gravid 33-year-old woman (gravid 4, para 3) was referred to our hospital with vaginal bleeding from 9 weeks of gestation. She was noted on prenatal ultrasound to have a normal foetus with an abnormally thickened space in the placental region. At 11 gestational weeks, a snowstorm pattern was observed on ultrasound examination, but it was slightly different from the typical pattern for hydatidiform mole. Multivesicular areas were prominent, but the other areas appeared relatively normal (Figure 1(A)). At 13 gestational weeks, the snowstorm pattern persisted with a foetal growth retardation of a biparietal diameter of 22.3 mm (−1.9 SD). The serum  $\beta$ -human chorionic gonadotropin ( $\beta$ -hCG) level was alarmingly elevated at 369,065 mIU/ml at 14 gestational weeks, whereas alpha-fetoprotein (AFP) showed a normal level of 109.5 ng/ml.  $\beta$ -hCG was persistently high at 207,336 mIU/ml at 16 gestational weeks, whereas AFP was 159.8 ng/ml.

The couple decided to terminate the pregnancy after considering the risks because the possibility of hydatidiform mole and coexistent foetus could not be excluded. After the curettage, the woman was in good condition and the  $\beta$ -hCG

level decreased to 4 mIU/ml. The delivered foetus had a median cleft lip and palate (Figure 1(B)). The placenta appeared to have patchy villous hydropic changes (Figure 1(C)). Histological examination revealed focal villous oedema. Trophoblast hyperplasia was not observed (Figure 1(D)). After receiving approval from the Ethical Review Board and obtaining written informed consent from the couple, we obtained samples from the foetal skin and from the oedematous and normal-seeming areas of the placenta.

Initial cytogenetic analysis by Giemsa staining indicated a normal karyotype (data not shown). Cytogenetic microarray of the foetus revealed three copies of an 8-Mb region at the terminus of 9p, but monosomy 2q and trisomy 4q in the placenta (Figure 1(E–G)). Although hydatidiform moles generally result from dispermic triploidy or diandric diploidy with the paternal genome only, there was no evidence of triploidy or uniparental disomy. The foetus was found to carry  $\text{arr}[\text{hg}19]9\text{p}24.3\text{p}24.1(326,927\text{--}8,441,863)\times 3$ , which appeared to be mosaic with normal cells because the copy number (CN) state was 2.80. On the other hand, the placental tissue was found to carry  $\text{arr}[\text{hg}19]2\text{q}37.3(237,337,625\text{--}242,408,074)\times 1,4\text{q}25\text{q}35.2(113,816,349\text{--}190,957,473)\times 3$ . These appeared to be in mosaicism because the CN state was 1.35 and 2.67, respectively. Approximately 65–67% of cells showed monosomy 2q and trisomy 4q, and it is likely that the same cells had monosomy 2q and trisomy 4q simultaneously. The placental tissue also showed 9p trisomy at CN state 2.33, suggesting that 33% of cells carried the 9p trisomy identified in the foetus. On the other hand, we did not detect monosomy 2q and trisomy 4q in foetal tissue at all.

Microsatellite analysis of the DXS0767 locus revealed that there was only a small level of maternal tissue contamination in placental tissue (2–3%, data not shown) and none in foetal tissue. The pattern of whole-genome SNP genotyping also excluded the chimeric pattern but indicated a single zygote origin, suggesting that all of the foetus and placenta were derived from a monozygotic twin or somatic



**Figure 1.** Clinical phenotypes and cytogenetic analysis of the foetus and placenta. Cytogenetic microarray was performed using CytoScan HD Array (Affymetrix). #1: placenta that appeared relatively normal; #2: placenta that included villous hydrops lesions; and #3: foetus. (A) Ultrasound examination at 11 weeks of gestation. Two separate areas – a vesicular area (upper area) and relatively normal area (lower area) – were observed, which are atypical for gestational trophoblastic disease. (B) Foetus. A median cleft lip and palate were observed. (C) Macroscopic analysis of the placenta. Patchy villous hydropic changes were observed. (D) Histological specimen for chorionic villi. Focal villous oedema was observed. Scale bars, 100  $\mu$ m. (E) 9p and 9q. (F), 2q. (G), 4q. Scale bars, 10 Mb. (H), Giemsa staining. Additional material was observed at the terminal region of 14p. (I) FISH. Subtelomeric probes (Vysis ToTelVision, Abbott Molecular) revealed the presence of  $\text{der}(14)\text{t}(9;14)(\text{p}24;\text{p}11.2)$  (arrow). White arrows: 9p; and white arrow heads: 9q.

mosaicism of a single zygote. As the CN state showed that the cell population with 9p and that with monosomy 2q and trisomy 4q were mutually exclusive, we concluded that they were likely from monozygotic twins ([Supplementary Figure](#)).

Reexamination of Giemsa staining of the foetal fibroblasts showed additional material at the terminal of 14p. Subtelomeric FISH was performed to further characterise the CN abnormalities. Trisomy 9p was found to originate from  $\text{der}(14)\text{t}(9;14)(\text{p}24;\text{p}11.2)$  in all of the 20 metaphases examined ([Figure 1\(H,I\)](#)). As the CN states of monosomy 2q and trisomy 4q are reciprocal, the monosomy 2q and trisomy 4q found in the placenta were likely to have originated from unbalanced  $\text{t}(2;4)(\text{q}37.3;\text{q}25)$  translocation. However, subtelomeric FISH did not detect the  $\text{t}(2;4)$  translocation in any of the foetal cells. We did not study the karyotype of the couple because they did not want to undergo the required examinations.

We recommend careful performance of the differential diagnosis of abnormal placenta with snowstorm pattern, particularly in cases with a coexistent foetus. A molecular

cytogenetic study including zygosity test is necessary for differential diagnosis because it is possible that a chromosomal disorder might underlie placental abnormalities. The severities of the clinical symptoms in the foetus with such disorders vary widely. These disorders often result in lethality from multiple congenital anomalies, whereas cases with milder cytogenetic abnormalities can occasionally survive and live to a good age. Furthermore, confined placental mosaicism might affect the foetus to a lesser degree ([Johnson and Wapner 1997](#); [Lestou and Kalousek 1998](#)). Thus, the results of the cytogenetic test might seriously affect the choice of treatment for the ultrasound findings.

### Acknowledgements

We gratefully acknowledge the patients and their families for participating in this study.

### Disclosure statement

The authors report no conflicts of interest.

## Funding

This study was supported by the Ogyaa Donation Foundation from the Japan Association of Obstetricians & Gynecologists and by grants-in-aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology and from the Ministry of Health, Labour and Welfare of Japan.

## References

- Gupta K, Venkatesan B, Kumaresan M, Chandra T. 2015. Early detection by ultrasound of partial hydatidiform mole with a coexistent live fetus. *WMJ: Official Publication of the State Medical Society of Wisconsin* 114:208–211.
- Johnson A, Wapner RJ. 1997. Mosaicism: implications for postnatal outcome. *Current Opinion in Obstetrics & Gynecology* 9:126–135.
- Lestou VS, Kalousek DK. 1998. Confined placental mosaicism and intra-uterine fetal growth. *archives of disease in childhood. Fetal and Neonatal Edition* 79:F223–F226.
- Matsui H, Sekiya S, Hando T, Wake N, Tomoda Y. 2000. Hydatidiform mole coexistent with a twin live fetus: a national collaborative study in Japan. *Human Reproduction (Oxford, England)* 15:608–611.
- Sebire NJ, Foskett M, Paradinas FJ, Fisher RA, Francis RJ, Short D, et al. 2002. Outcome of twin pregnancies with complete hydatidiform mole and healthy co-twin. *Lancet (London, England)* 359:2165–2166.
- Toufaily MH, Roberts DJ, Westgate MN, Holmes LB. 2016. Triploidy: variation of phenotype. *American Journal of Clinical Pathology* 145:86–95.
- Wee L, Jauniaux E. 2005. Prenatal diagnosis and management of twin pregnancies complicated by a co-existing molar pregnancy. *Prenatal Diagnosis* 25:772–776.

# Prenatal diagnosis of premature chromatid separation/ mosaic variegated aneuploidy (PCS/MVA) syndrome

Tomoko Yamaguchi<sup>1</sup>, Masatoshi Yamaguchi<sup>1,2</sup>, Keiko Akeno<sup>1</sup>, Midori Fujisaki<sup>1</sup>,  
Kaeko Sumiyoshi<sup>1</sup>, Masanao Ohashi<sup>1,3</sup>, Hiroshi Sameshima<sup>1</sup>, Mamoru Ozaki<sup>4</sup>, Maki Kato<sup>5</sup>,  
Takema Kato<sup>5</sup>, Eriko Hosoba<sup>5</sup> and Hiroki Kurahashi<sup>5</sup>

<sup>1</sup>Department of Obstetrics and Gynecology, Faculty of Medicine, University of Miyazaki, <sup>2</sup>Division of Clinical Genetics, University of Miyazaki Hospital, <sup>3</sup>Department of Obstetrics and Gynecology, Miyazaki Medical Association Hospital, Miyazaki, <sup>4</sup>Division of Genomic Medicine, Department of Advanced Medicine, Medical Research Institute, Kanazawa Medical University, Uchinada and <sup>5</sup>Division of Molecular Genetics, Institute for Comprehensive Medical Science, Fujita Health University, Toyoake, Japan

## Abstract

Premature chromatid separation/mosaic variegated aneuploidy (PCS/MVA) syndrome is a rare genetic disorder. In this case report, we describe the prenatal diagnosis of PCS/MVA syndrome in a 24-year-old, gravida 1, para 1, woman who was referred to us in her second trimester due to fetal growth restriction and extreme microcephaly (−5.0 standard deviations). Amniocentesis and chromosomal analysis confirmed PCS in 80% of cultured fetal cells. PCS findings were positive in 9% of paternal cells and 11% of maternal cells, indicative that both were PCS carriers. Genetic analysis confirmed that the fetus carried a combined heterozygote of maternal G > A point mutation of the promoter area of the *BUB1B* gene and a paternal *Alu* sequence insertion between intron 8 and exon 9 of the *BUB1B* gene. As PCS/MVA syndrome is associated with the development of various malignancies in early life, prenatal diagnosis is important for effective planning of post-natal care.

**Key words:** *BUB1B* gene, case report, microcephaly, premature chromatid separation/mosaic variegated aneuploidy syndrome, prenatal diagnosis.

## Introduction

Premature chromatid separation [PCS, Mendelian Inheritance in Man (MIN) 176430], also known as mosaic variegated aneuploidy syndrome (MVA, MIN 257300), was first reported by Rudd *et al.* in 1983.<sup>1</sup> In this initial report, Rudd *et al.* described three cases in which cells cultured in colcemid showed increased cell mitosis, with separated centromeres and splayed chromatids. The clinical phenotype of premature chromatid separation/mosaic variegated aneuploidy (PCS/MVA) syndrome includes microcephaly and mental retardation or Dandy Walker syndrome.

Moreover, the PCS/MVA syndrome is associated with a high incidence of cancer formation in early life. Therefore, although the incidence of this syndrome is 1:1000000, prenatal diagnosis is important to inform effective post-natal care.

The *BUB1B* gene, known as a checkpoint gene, is one of the candidate genes involved in the PCA/MVA syndrome.<sup>2</sup> Therefore, the PCA/MVA syndrome is distinguishable on genetic analysis. Our aim in this case report is to describe the prenatal genetic diagnosis of PCS/MVA syndrome in a patient referred to our center for prenatal follow-up after abnormal findings on prenatal ultrasound examination.

Received: November 12 2017.

Accepted: February 10 2018.

Correspondence: Dr Masatoshi Yamaguchi, Department of Obstetrics and Gynecology, Division of Clinical Genetics, University of Miyazaki Hospital, 5200 Kihara, Kiyotake-Cho, Miyazaki 889-1692, Japan. Email: myam@med.miyazaki-u.ac.jp

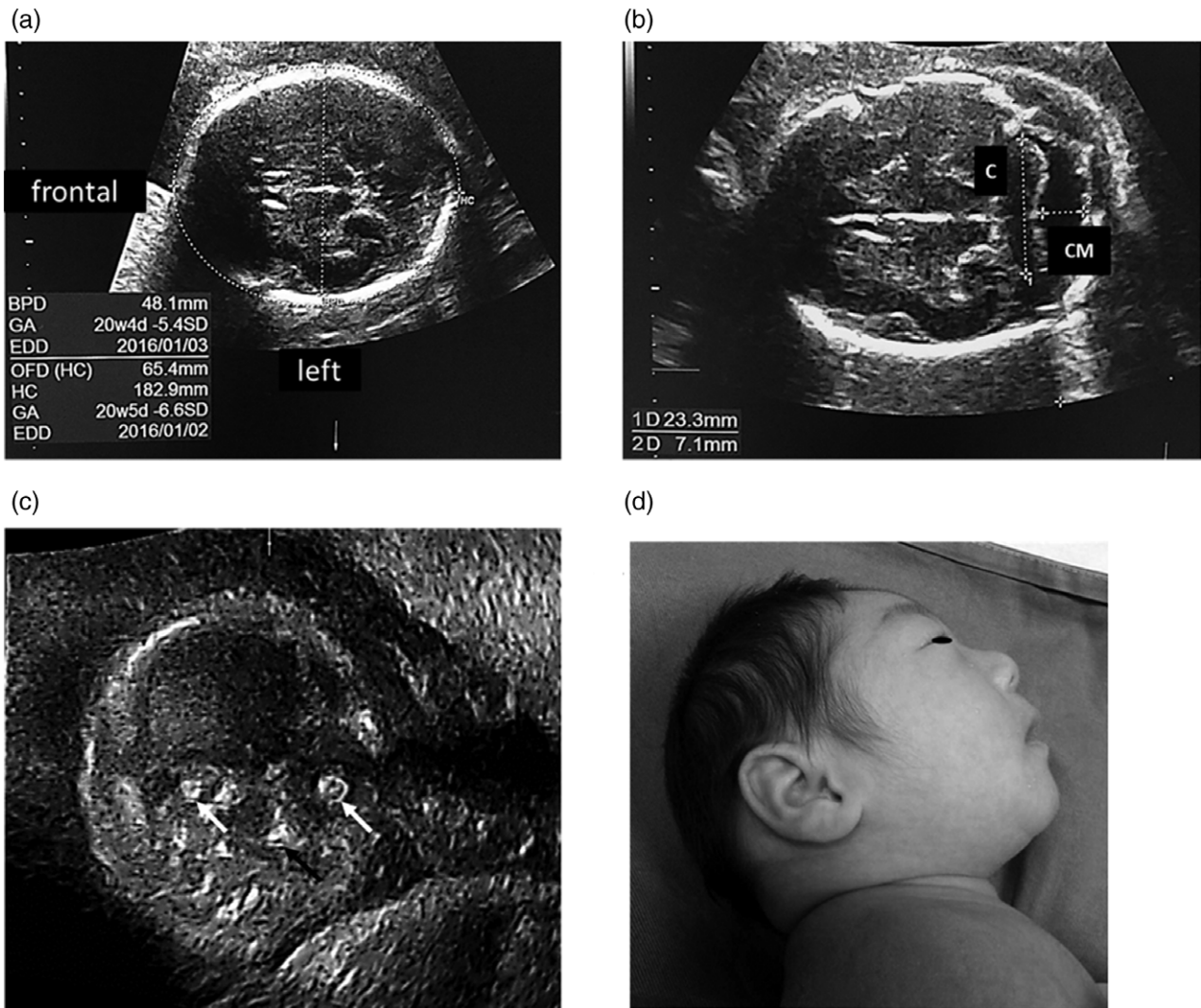


## Case Report

A 24-year-old gravida 1, para 1, woman was referred to our hospital in her second trimester because of abnormal fetal findings on prenatal ultrasound examination. She reported having regular menstrual cycles prior to pregnancy. Her husband was not a consanguineous partner, and there was no maternal or paternal family history of congenital malformation.

Fetal ultrasonography at 23 weeks of gestation showed fetal growth restriction (FGR) with a severe

microcephaly; the estimated fetal body weight measured 415 g ( $-2.4$  SD), and the head circumference measured in the  $-6.6$  SD for gestational age. The high echogenic eyeballs appeared to be congenital fetal cataracts (Fig. 1). However, the remainder of the fetal structures seemed to be normal, including the brain, heart, kidneys and other internal organ or external genitalia. Specifically, renal tumor could not be detected. These images and the low-avidity toxoplasma Ig G indicated the possibility of toxoplasmosis. At 24 weeks of gestation, we performed the amniocentesis for the diagnosis of toxoplasmosis

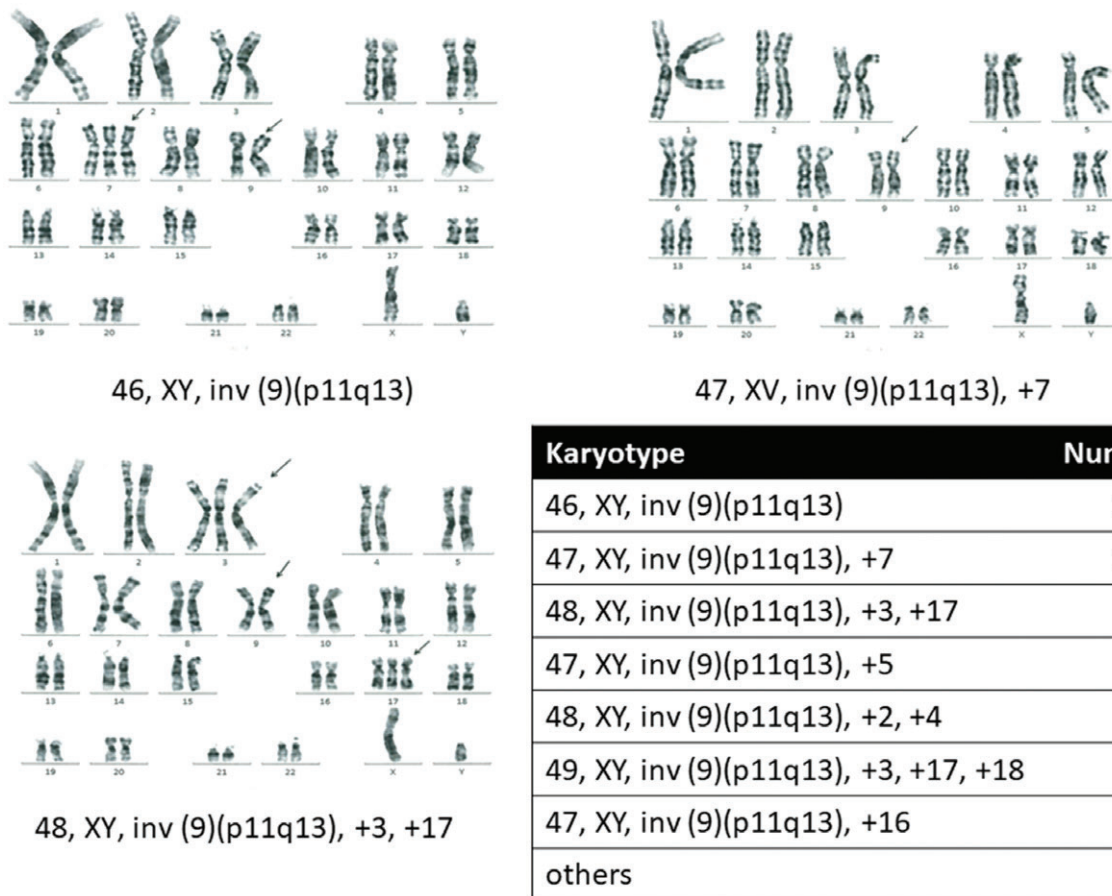


**Figure 1** Summary of fetal ultrasound findings showing: (a) microcephaly, with the head circumference being  $-6.6$  SD of the age-referenced normal range; (b) normal transverse cerebellar diameter (C) and no enlargement of the fetal cistern magna (CM);<sup>3</sup> (c) high echogenic fetal lens (white arrow) and nose bone (black arrow), high echogenic fetal lens indicated bilateral cataracts; and (d) neonatal head appearance, note that there was microcephaly with low-set ears, and micrognathia.

using PCR methods, and we wanted to rule out chromosomal aberrations. However, toxoplasmosis PCR was negative, and several karyotypes of PCS and PCS-related MVA were present on G-band chromosomal analysis (Fig. 2). Subsequent to these findings, we tested the karyotype of both parents using 200 cells of cultured peripheral white blood cells for each parent. PCS-positive cells were identified in both the father (9% of cells) and mother (11% of cells), indicative that both parents were likely to be carriers of the PCS/MVA syndrome (Table 1).

In order to confirm the diagnosis, DNA from a cultured amniotic cell was isolated for genetic diagnosis using a previously described method.<sup>2</sup> Genetic analysis confirmed that the fetus carried a combined heterozygote of maternal G > A point mutation of the promoter area of the *BUB1B* gene and a paternal *Alu* sequence insertion between intron 8 and exon 9 of the *BUB1B* gene.

At 38 weeks 3 days of gestation, the mother experienced abrupt vaginal bleeding and underwent an emergency cesarean section due to a suspected placental abruption. The male fetus was delivered. The neonate had no signs of respiratory distress, with Apgar scores of 8 at 1 min and 9 at 5 min. The birthweight was 1934 g (-3.2 SD), with a head circumference of 26.7 cm (-4.5 SD). Physical anomalies were noted, including bilateral congenital cataracts, severe microcephaly, prominent nasal bridge, upslanting palpebral fissure, low-set ears, micrognathia and ambiguous genitalia. The infant had a clinical follow-up especially for neurodevelopment and renal tumor. At birth, the neurological examination was unremarkable, but at 3 months of age, the infant developed a generalized seizure disorder and started anticonvulsants treatment. The sonographic examination at 6 months of age confirmed the presence of bilateral



**Figure 2** Fetal karyotype of the amniocentesis, showing several types, including: 46, XY, inv (9)(p11q13), 47, XY, inv (9)(p11q13),+7 and 48, XY, inv (9)(p11q13), +3 and +13. These findings were strongly suggestive of PCS/MVA syndrome.

**Table 1** Karyotype and *BUB1B* genotype for the fetus and the parents

	Karyotype	%PCS/MVA phenomenon	<i>BUB1B</i> genotype
Case	46, XY, inv (9)(p11q13) etc	78	Combined heterozygote
Mother	46, XX, 1qh+, inv (9)(p11q13)	11	A/G at ss802470619
Father	46, XY	9	<i>Alu</i> insertion intron8

White blood cells, collected from peripheral blood for both parents, were incubated with phytohemagglutinin (PHA) for karyotyping and genotyping.

renal tumor; laparotomy was performed to resect the bilateral kidneys, and peritoneal dialysis was induced. The renal tumor was diagnosed as Wilms tumor. At 26 months of age, he required medical treatments for renal failure and intractable epilepsy with severe neurodevelopment disorder.

## Discussion

Microcephaly is an important fetal abnormality related to atypical perinatal neurological development. Viral infections are the most common cause of microcephaly,<sup>4,5</sup> with chromosomal abnormality<sup>6</sup> and fetal alcohol syndrome<sup>7</sup> rarely being associated with microcephaly. Moreover, in most cases of microcephaly, the fetal head circumference is usually above  $-2.0$  SD, with extreme microcephaly being a rare occurrence.

Although PCS/MVA syndrome was first reported by Rudd *et al.*<sup>1</sup> in 1983, the syndrome was first reported in Japan by Kajii *et al.* in 1998.<sup>8</sup> Although the incidence rate of the PCA/MVA syndrome is very low (1:1000000), this syndrome should be considered in cases of severe fetal microcephaly and growth restriction. Amniocentesis and karyotyping is an important component of the prenatal diagnosis. In our case, 78% of fetal amniotic cells were positive for PCS/MVA.

PCS/MVA is an autosomal recessive disorder, and therefore, both parents may be carriers, as in our case, with  $\sim 10\%$  of the peripheral white blood cells showing PCS-positive traits in both parents. As an autosomal recessive disorder, there is a 25% chance of recurrence on subsequent pregnancies. Recently, microarray, rather than traditional cytogenetics, has been used for chromosomal analysis. However, the karyotype of each cell is variable in case of PCS/MVA syndrome. Therefore, microarray technology is not suitable for diagnosis of PCS/MVA syndrome.

*BUB1B* is a candidate gene of the PCS/MVA syndrome.<sup>9–11</sup> BubR1, a biproduct of the *BUB1B* gene, is an important component of the cell spindle

assembly checkpoint.<sup>12</sup> Impairment in the function of BubR1 typically results in chromosomal instability in cells and the development of malignancy.<sup>12</sup> Mutation of the *BUB1B* gene is a cause of malignancy in both children and adults.<sup>13,14</sup> Of concern is the early onset and severity of cancer development among Asian, compared to Caucasian, neonates with a *BUB1B* gene mutation.<sup>13</sup> Therefore, prenatal diagnosis of the PCA/MVA syndrome in Asian populations is important to inform effective post-natal care. Yet, we only identified two previous reports on the prenatal diagnosis of PCS/MVA syndrome.<sup>15,16</sup> In our case, our genetic testing confirmed a combined heterozygote mutation of the *BUB1B* gene, which has previously been described.<sup>2</sup> As PCS/MVA syndrome has a poor prognosis, genetic testing to determine the PCS status of the parents is important to provide parents with necessary information for the planning of future possible pregnancies, including prenatal genetic testing for diagnosis.

In conclusion, we described the prenatal diagnosis of PCS/MVA syndrome, which was performed due to the identification of severe fetal microcephaly on prenatal ultrasound examination at 23 weeks of gestation. Based on our experience, we propose that the PCS/MVA syndrome be considered one of the possible differential diagnoses in cases of severe fetal microcephaly.

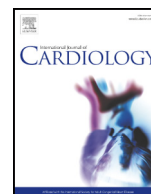
## Disclosure

None declared.

## References

- Rudd NL, Teshim IE, Martin RH, Siskin JE, Weksberg R. A dominantly inherited cytogenetic anomaly: A possible cell division mutant. *Hum Genet* 1983; **65**: 117–121.
- Kato M, Kato T, Hosoba E *et al.* PCS/MVA syndrome due to *Alu* insertion in the *BUB1B* gene. *Hum Genome Var* 2017; **4**: 17021.
- Snijders RJM, Nicolaidis KH. Fetal biometry at 14–40 weeks' gestation. *Ultrasound Obstet Gynecol* 1994; **4**: 34–48.

4. Degani S. Sonographic findings in fetal viral infections: A systematic review. *Obstet Gynecol Surv* 2006; **61**: 329–336.
5. Dahlgren L, Wilson RD. Prenatally diagnosed microcephaly: A review of etiologies. *Fetal Diagn Ther* 2001; **16**: 323–326.
6. Wagner N, Guengoer E, Mau-Holzmann UA *et al*. Prenatal diagnosis of a fetus with terminal deletion of chromosome 1 (q43) in first-trimester screening: Is there a characteristic antenatal 1q deletion phenotype? A case report and review of the literature. *Fetal Diagn Ther* 2011; **29**: 253–256.
7. Niccols A. Fetal alcohol syndrome and the developing socio-emotional brain. *Brain Cogn* 2007; **65**: 135–142.
8. Kajii T, Kawai T, Takumi T *et al*. Mosaic variegated aneuploidy with multiple congenital abnormalities: Homozygosity for total premature chromatid separation trait. *Am J Med Genet* 1998; **78**: 245–249.
9. Ochiai H, Miyamoto T, Kanai A *et al*. TALEN-mediated single-base-pair editing identification of an intergenic mutation upstream of BUB1B as causative of PCS (MVA) syndrome. *Proc Natl Acad Sci U S A* 2014; **111**: 1461–1466.
10. Matsuura S, Matsumoto Y, Morishima K *et al*. Monoallelic BUB1B mutations and defective mitotic-spindle checkpoint in seven families with premature chromatid separation (PCS) syndrome. *Am J Med Genet A* 2006; **140**: 358–367.
11. Kajii T, Ikeuchi T, Yang ZQ *et al*. Cancer-prone syndrome of mosaic variegated aneuploidy and total premature chromatid separation: Report of five infants. *Am J Med Genet* 2001; **104**: 57–64.
12. Yang F, Huang Y, Dai W. Sumoylated BubR1 plays an important role in chromosome segregation and mitotic timing. *Cell Cycle* 2012; **11**: 797–806.
13. Hanks S, Coleman K, Summersgill B *et al*. Comparative genomic hybridization and BUB1B mutation analyses in childhood cancers associated with mosaic variegated aneuploidy syndrome. *Cancer Lett* 2006; **239**: 234–238.
14. Sari N, Akyuz C, Aktas D *et al*. Wilms tumor, AML and medulloblastoma in a child with cancer prone syndrome of total premature chromatid separation and Fanconi anemia. *Pediatr Blood Cancer* 2009; **53**: 208–210.
15. Plaja A, Mediano C, Cano L *et al*. Prenatal diagnosis of a rare chromosomal instability syndrome: Variegated aneuploidy related to premature centromere division (PCD). *Am J Med Genet A* 2003; **117A**: 85–86.
16. Chen CP, Lee CC, Chen WL, Wang W, Tzen CY. Prenatal diagnosis of premature centromere division-related mosaic variegated aneuploidy. *Prenat Diagn* 2004; **24**: 19–25.



## Frequent intragenic microdeletions of elastin in familial supravalvular aortic stenosis



Satoshi Hayano <sup>a,1</sup>, Yusuke Okuno <sup>b,1</sup>, Makiko Tsutsumi <sup>c,1</sup>, Hidehito Inagaki <sup>c,1</sup>, Yoshie Fukasawa <sup>a,1</sup>, Hiroki Kurahashi <sup>c,1</sup>, Seiji Kojima <sup>a,1</sup>, Yoshiyuki Takahashi <sup>a,1</sup>, Taichi Kato <sup>a,\*,1</sup>

<sup>a</sup> Department of Pediatrics, Nagoya University Graduate School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya, Japan

<sup>b</sup> Center for Advanced Medicine and Clinical Research, Nagoya University Hospital, 65 Tsurumai-cho, Showa-ku, Nagoya, Japan

<sup>c</sup> Division of Molecular Genetics, Institute for Comprehensive Medical Science, Fujita Health University, 1-98 Dengakugakubo, Kutsukake-cho, Toyoake, Japan

### ARTICLE INFO

#### Article history:

Received 10 April 2018

Received in revised form 1 September 2018

Accepted 7 September 2018

Available online 13 September 2018

#### Keywords:

Supravalvular aortic stenosis

Elastin

Congenital heart defects

Whole exome sequencing

### ABSTRACT

**Background:** Supravalvular aortic stenosis (SVAS) is a congenital heart disease affecting approximately 1:25,000 live births. SVAS may occur sporadically, be inherited in an autosomal dominant manner, or be associated with Williams-Beuren syndrome, a complex developmental disorder caused by a microdeletion of chromosome 7q11.23. *ELN* on 7q11.23, which encodes elastin, is the only known gene to be recurrently mutated in less than half of SVAS patients.

**Methods:** Whole-exome sequencing (WES) was performed for seven familial SVAS families to identify other causative gene mutations of SVAS.

**Results:** Three truncating mutations and three intragenic deletions affecting *ELN* were identified, yielding a diagnostic efficiency of 6/7 (85%). The deletions, which explained 3/7 of the present cohort, spanned 1–29 exons, which might be missed in the course of mutational analysis targeting point mutations. The presence of such deletions was validated by both WES-based copy number estimation and multiplex ligation-dependent probe amplification analyses, and their pathogenicity was reinforced by co-segregation with clinical presentations.

**Conclusions:** The majority of familial SVAS patients appear to carry *ELN* mutations, which strongly indicates that elastin is the most important causative gene for SVAS. The frequency of intragenic deletions highlights the need for quantitative tests to analyze *ELN* for efficient genetic diagnosis of SVAS.

© 2018 Elsevier B.V. All rights reserved.

## 1. Introduction

Supravalvular aortic stenosis (SVAS; MIM #185500) is a congenital heart disease affecting approximately 1:25,000 live births [1]. Congenital narrowing of the lumen of the ascending aorta or peripheral pulmonary arteries provokes increased resistance to blood flow and causes elevated ventricular pressure and hypertrophy resulting in heart failure. Peripheral pulmonary stenosis (PPS) is known to occasionally coexist with SVAS [2]. Approximately 30% to 50% of patients with SVAS have Williams-Beuren Syndrome (WBS; MIM #194050) [2–4], which is a complex genetic disorder caused by 7q11.23 microdeletion and

characterized by growth failure, a characteristic facial appearance (so-called “Elfin face”), mental retardation, and SVAS [5].

On the other hand, Eisenberg et al. first reported non-syndromic “familial SVAS” with autosomal dominant inheritance in 1964 [3], accounting for 20% of SVAS cases (approximately 1:125,000 live births) [6]. These patients showed normal intelligence and lacked the dysmorphic features of WBS. Genetic analysis including linkage analysis identified *ELN*, which encodes elastin, as a causative gene of non-syndromic familial SVAS [1,7–17]. In harmony with the genetic findings, luminal obstruction of the aorta was shown in a transgenic mouse model carrying homozygous or heterozygous elastin gene deletion [18,19].

Metcalfe et al. sequenced *ELN* exons of patients with non-syndromic SVAS, which showed truncating mutations in 35 cases, but no causative variants were found in the remaining 64 patients (of which 8 were familial cases) [20]. Micale et al. also investigated *ELN* gene mutations in 14 familial and 10 sporadic cases of SVAS, resulting in 7 novel mutations, including 5 frameshift and 2 donor splice site mutations, but found no *ELN* gene abnormality in the remaining 17 cases [21]. Therefore, less than half of the cases could be explained by *ELN* mutations, whereas it still remains unclear whether *ELN* could explain the remaining cases

\* Corresponding author.

E-mail addresses: [jvauma@gmail.com](mailto:jvauma@gmail.com) (S. Hayano), [yusukeokuno@gmail.com](mailto:yusukeokuno@gmail.com) (Y. Okuno), [makiko@fujita-hu.ac.jp](mailto:makiko@fujita-hu.ac.jp) (M. Tsutsumi), [hinagaki@fujita-hu.ac.jp](mailto:hinagaki@fujita-hu.ac.jp) (H. Inagaki), [love.rodin.love@gmail.com](mailto:love.rodin.love@gmail.com) (Y. Fukasawa), [kura@fujita-hu.ac.jp](mailto:kura@fujita-hu.ac.jp) (H. Kurahashi), [kojimas@med.nagoya-u.ac.jp](mailto:kojimas@med.nagoya-u.ac.jp) (S. Kojima), [ytakaha@med.nagoya-u.ac.jp](mailto:ytakaha@med.nagoya-u.ac.jp) (Y. Takahashi), [ktachi@med.nagoya-u.ac.jp](mailto:ktachi@med.nagoya-u.ac.jp) (T. Kato).

<sup>1</sup> This author takes responsibility for all aspects of the reliability and freedom from bias of the data presented and their discussed interpretation.



**Abbreviations**

SVAS	supravalvular aortic stenosis
WES	whole-exome sequencing
PPS	peripheral pulmonary stenosis
WBS	Williams-Beuren syndrome
FISH	fluorescence in situ hybridization
MLPA	Multiplex ligation-dependent probe amplification

with SVAS, or there are unidentified causative genes. In this study, whole-exome sequencing (WES) was performed with careful assessment of *ELN* mutations, including copy number analysis, to elucidate the genetic background of SVAS.

**2. Methods**

**2.1. Sample collection**

This study included seven families of Japanese ancestry with autosomal dominant inheritance of SVAS. There was no developmental delay or dysmorphic features suggestive of WBS or positive fluorescence in situ hybridization (FISH) on 7q11.23 in any family members. The vascular malformation (SVAS and PPS) was diagnosed if the sinotubular junction of the aorta was smaller than the diameter of the aortic annulus and significant pressure gradients were measurable by echocardiogram and/or angiographically [6]. Written, informed consent was obtained from patients or their parents, and whole blood or saliva was collected. Saliva samples were collected using an Oragene DNA self-collection kit (DNA Genotek, Ottawa, Canada). Genomic DNA was extracted from

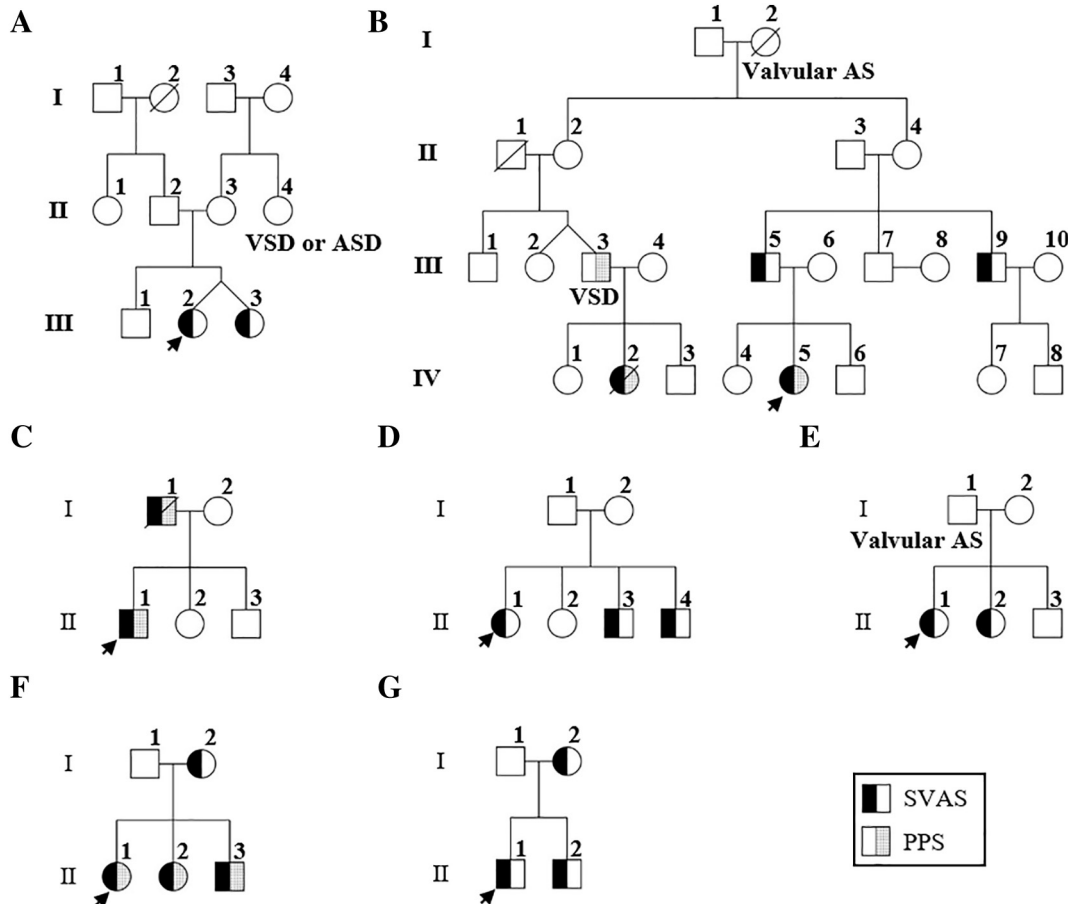
whole blood or saliva using the QIAamp DNA Blood Mini kit (Qiagen, Hilden, Germany), according to the manufacturer's instructions. The study was approved by the Ethics Committee of the Nagoya University Graduate School of Medicine (approval number 2015-0032).

**2.2. Whole-exome sequencing analysis**

Exome capture was performed on each proband using SureSelect Human All Exon V5 (Agilent Technologies, Santa Clara, CA), according to the manufacturer's instructions. Generated libraries were sequenced on a HiSeq 2500 platform (Illumina, San Diego, CA). Sequence data were analyzed using an in-house pipeline [22]. Briefly, reads were aligned to UCSC build hg19 reference genome using the Burrows-Wheeler Aligner [23]. Picard tools (<http://broadinstitute.github.io/picard>) were utilized to remove PCR duplicates. Variants were called using VarScan2, where a variant allele frequency of >0.20 was used as a cutoff [24]. ANNOVAR was used together with in-house scripts to annotate genetic variants [25]. The average depth of coverage across the whole exome for each sample achieved was 111.14 (range 90.68 to 127.66), and the number of mutations found per sample ranged from 25,534 to 25,920.

**2.3. Mutational analysis**

Mutational analysis to define each variant's pathogenicity was essentially based on the latest release of the American College of Medical Genetics (ACMG) guideline [26]. Briefly, variants outside of coding regions and common variants with >1% minor allele frequency in the National Heart, Lung, and Blood Institute ESP (Exome Sequencing Project) 6500 [27], 1000 Genomes Project [28], ExAC (Exome Aggregation Consortium) [29], HGVD (Human Genetic Variation Database) [30], or the in-house database were excluded. Variants expected to cause the disorders (eg, missense variants with reported pathogenicity and nonsense, frameshift insertion/deletion, and splice-site variants on genes known to cause a disease by inactivation) were validated by Sanger sequencing using PrimeSTAR GXL DNA polymerase (Takara, Shiga, Japan) and the Big Dye Terminator 3.1 Cycle Sequencing Kit (Thermo Fisher Scientific Inc., Waltham, MA) with ABI PRISM 3130xL (Applied Biosystems, Foster City, CA). Primer sequences are listed in Table S1.



**Fig. 1.** Pedigree chart of families with familial supravalvular aortic stenosis. Arrows indicate probands for whom whole-exome sequencing was performed. SVAS, supravalvular aortic stenosis; PPS, peripheral pulmonary stenosis; VSD, ventricular septal defect; ASD, atrial septal defect; AS, aortic stenosis.

## 2.4. Copy number analysis

Copy number analysis was performed by comparing the number of reads conveying each exon normalized by the mean depth of the entire sample with that of unrelated normal DNA samples, as we have previously shown [22]. Exons of normalized coverage  $>3$  standard deviations (SDs) or less than  $-3$  SDs from the mean coverage of reference samples were considered to be candidates for copy number variants. Multiplex ligation-dependent probe amplification (MLPA) according to the manufacturer's protocol with the SALSA MLPA P029-WBS probemix (MRC Holland, Amsterdam, Netherlands), which includes 10 exons of the *ELN* gene (Exon 1, 3, 4, 6, 9, 16, 20, 26, 27 and 33), was performed to validate candidate exonic deletions detected in the *ELN* gene by WES analysis. MLPA analysis software Coffalyser (MRC Holland) was used to identify CNVs.

## 3. Results

WES-based detection of point mutations and copy number alterations was performed in seven families of Japanese ancestry with SVAS showing an autosomal dominant mode of inheritance (Fig. 1). Three heterozygous pathogenic mutations in *ELN* (c.370delT, p.Ser124Leufs\*13 in family A, c.572-1G > A splice site mutation affecting the acceptor of exon 12 in family B, and c.218\_219insTG, p.Gly74Valfs\*49 in family C) were identified. All of these mutations were novel, and they were validated on all available family members by Sanger sequencing (Fig. S1). Mutations were present

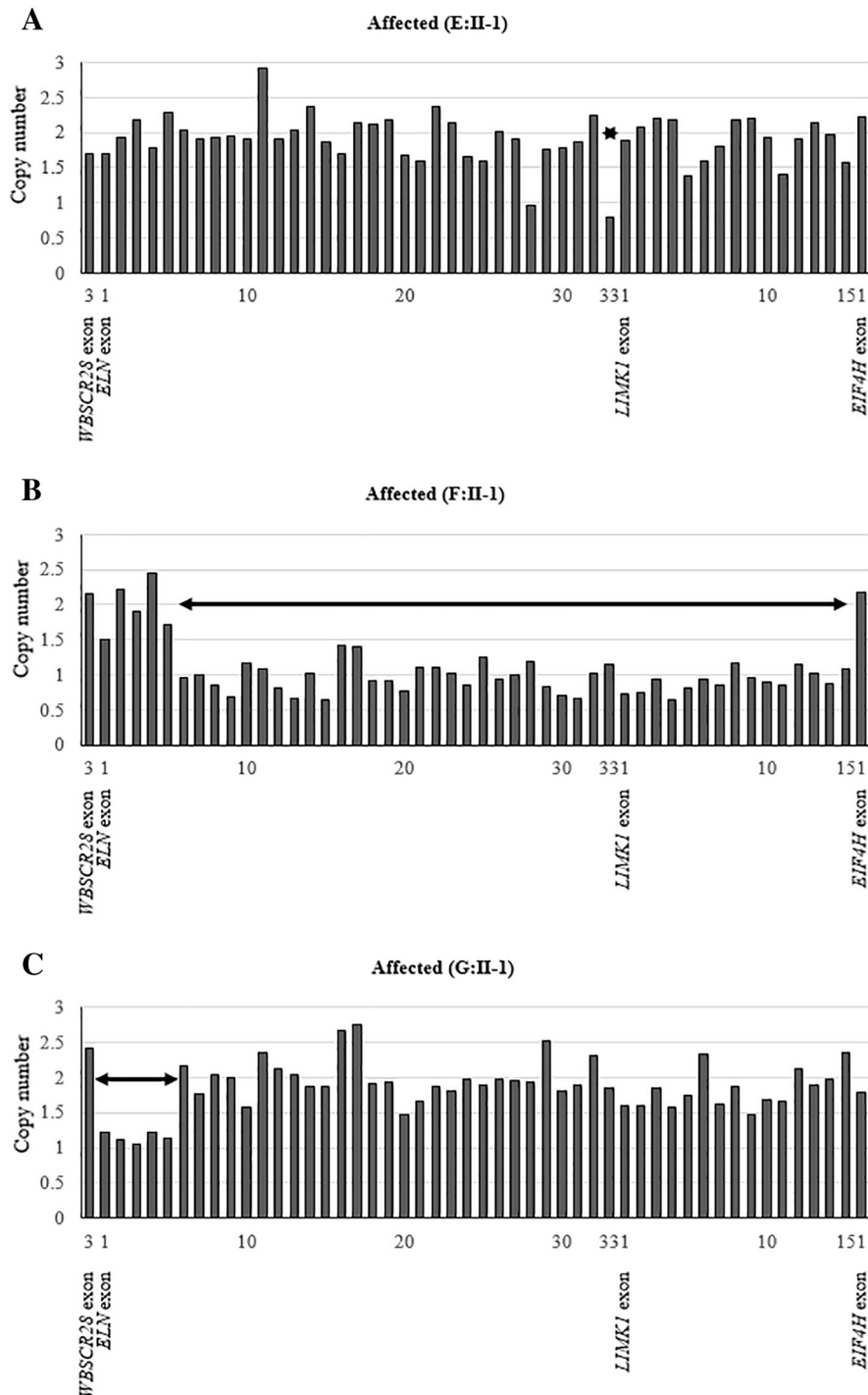


Fig. 2. Copy-number analysis The estimated copy number of each exon based on the number of reads in each exon in whole-exome sequencing. Each bar represents an exon, and the vertical axis represents the estimated copy number. Arrows indicate estimated deleted regions.

**Table 1**  
Phenotype and segregation of *ELN* mutations.

Family	Subject	Phenotype	<i>ELN</i> mutation
A	II-2	–	–
	II-3	–	c.370delT, p.Ser124Leufs*13
	III-1	–	–
	III-2	SVAS	c.370delT, p.Ser124Leufs*13
B	III-3	SVAS	c.370delT, p.Ser124Leufs*13
	III-3	PPS, VSD	c.572-1G > A splice site
	III-5	SVAS	c.572-1G > A splice site
	III-6	–	–
	IV-4	–	–
	IV-5	SVAS, PPS	c.572-1G > A splice site
C	IV-6	–	–
	I-2	–	–
	II-1	SVAS, PPS	c.218_219insTG, p.Gly74Valfs*49
D	II-2	–	–
	II-3	–	–
	II-3	–	c.218_219insTG, p.Gly74Valfs*49
E	II-1	SVAS	–
	I-1	Valvular AS	Microdeletion (exon 33)
F	I-2	–	–
	II-1	SVAS	Microdeletion (exon 33)
	II-2	SVAS	Microdeletion (exon 33)
	II-3	–	–
G	I-1	–	–
	I-2	SVAS	Microdeletion (exon 1–5)
	II-1	SVAS	Microdeletion (exon 1–5)
	II-2	SVAS	Microdeletion (exon 1–5)

All mutations were heterozygous. SVAS, supravalvular aortic stenosis; PPS, peripheral pulmonary stenosis; VSD, ventricular septal defect; AS, aortic stenosis.

in all patients and several family members without SVAS, indicating incomplete penetrance (Table 1).

Copy number aberrations in *ELN* were identified in three other families, all of which were deletions (exon 33 in family E, exons 5–33 in family F, and exons 1–5 in family G, Fig. 2). Such deletions were validated by MLPA (Fig. S2). All microdeletions showed complete cosegregation with clinical symptoms (Table 1).

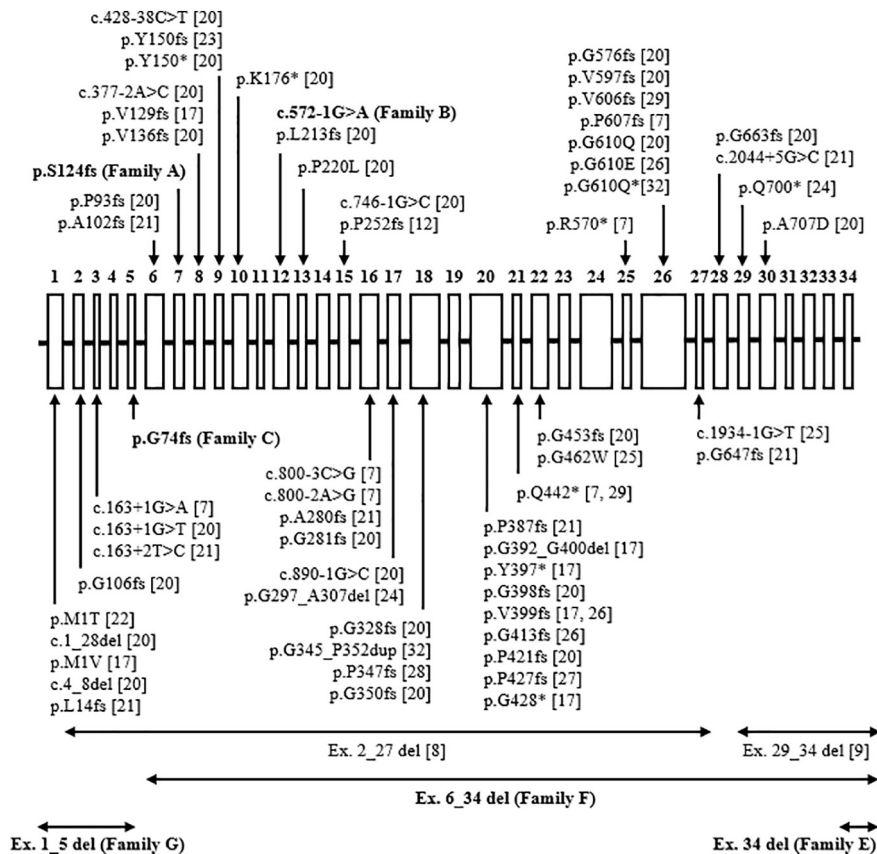
In the remaining family D, no diagnostic mutations associated with SVAS could be identified in the proband (D:II-1) either by WES or MLPA (Table S2, Fig. S3).

**4. Discussion**

Pathogenic mutations or deletions in *ELN* gene were identified in six of seven families with autosomal dominant inheritance of SVAS, including three novel point mutations and three intragenic deletions. These findings suggest that intragenic deletions in *ELN* gene could explain the genetic cause in half of so-far unexplained cases with familial SVAS in Japan. Updated by these findings, a comprehensive list of reported pathogenic SVAS mutations is provided (Fig. 3) [7–9,12,17,20,21,31–38].

*ELN* encodes elastin, which is expressed in various tissues and organs, including smooth muscle cells of the great arteries, and contributes to tissue elasticity [39,40]. The molecular mechanism of the pathogenesis of SVAS is not fully elucidated. However, considering accumulating knowledge from patients and transgenic mice [18,19,21,41], it seems likely that mutations of *ELN* impair vascular elasticity, and increased shear stress in the vascular wall could result in SVAS [39,40].

The microdeletions of *ELN* gene shown in the present study were not identified with existing FISH probes for WBS. There are a few case reports showing that microdeletions of *ELN* gene are the cause of SVAS [8,9]. The present finding raises the necessity to investigate the



**Fig. 3.** *ELN* cDNA showing the exons and mutations detected. This figure summarizes previously reported and newly identified mutations for familial and sporadic SVAS. The numbers above open boxes indicate the exon numbers. The present findings are shown in bold letters.

exon-spanning deletions affecting *ELN* using MLPA, array-CGH, or other methods to establish sufficient coverage for its mutations.

The present analysis showed co-segregation of symptoms and *ELN* mutations in the majority of analyzed individuals. The three mutations identified in this study (c.370delT in family A, c.572-1G > A splice site in family B, and c.218\_219insTG in family C) were highly pathogenic truncating mutations that result in premature stop codons (PTCs). A number of PTC mutations have actually been shown to be substrates of *ELN* mRNA insufficiency through nonsense-mediated decay in previous studies [17,21,35].

A highly variable phenotype within families with SVAS has been reported for large studies with many families with point mutations [20,21], and even for a family with apparently damaging 30 kb deletion involving multiple exons [7], ranging from asymptomatic mutation carrier to severe stenosis with multiple arteries. There were also two asymptomatic persons carrying *ELN* mutations (A:II-3 and C:II-3) in the present study. Factors affecting the variability of cardiovascular phenotypes in patients with *ELN* mutation are not yet fully understood, and there has as yet been no clear genotype-phenotype correlation reported for SVAS. Our comprehensive list of reported pathogenic SVAS mutations and deletions showed a universal distribution of variants over the entire *ELN* gene with no significant hotspot.

The primary mechanism for the pathogenesis of SVAS is proposed to be haploinsufficiency of *ELN*, as hemizyosity of *ELN* is established as the mechanism of SVAS in WBS [5]. Incomplete penetrance and a broad range in severity of cardiovascular phenotype are also seen in patients with WBS in whom one copy of *ELN* gene is totally lost [42,43]. Among possible causes affecting the severity of symptoms, the effect of the mutations in the remaining allele of *ELN* is very limited, as only two rare missense changes were identified through exon sequencing of 49 patients with WBS [44]. Currently, there is no definite explanation for the phenotypic variability associated with *ELN* mutations.

The present study showed that microdeletions of *ELN* gene could account for additional cases, around a half of previously unexplained cases, of familial SVAS, which would strengthen the causative role of *ELN* mutations in this disease entity. Therefore, quantitative genetic tests such as MLPA or array-CGH of *ELN* gene should be performed to genetically diagnose patients with familial SVAS to obtain satisfactory sensitivity. Further investigations of a larger cohort and so-far unexplained cases will be needed to elucidate the remaining molecular pathogenetic mechanisms of SVAS.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ijcard.2018.09.032>.

#### Authors' contributions

All authors developed the concept and designed the research; S.H., Y.O., and M.T. performed the experiments; S.H., Y.O., M.T., H.I., H.K., and T.K. analyzed the data; all authors interpreted the results of the experiments; S.H., Y.O., and T.K. prepared the figures; S.H., Y.O., M.T., H.I., Y.F., H.K., S.K., Y.T., and T.K. drafted the manuscript; S.H., Y.O., and T.K. edited the manuscript; all authors approved the final version of the manuscript.

#### Source of funding

This work was supported by funding from the Morinaga Foundation for Health and Nutrition to Dr. Taichi Kato. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

#### Competing interests

The authors report no relationships that could be construed as a conflict of interest.

#### Acknowledgements

The authors would like to thank all of the clinicians and families who made this study possible with the provision of samples. The authors would also like to thank Dr. Shinsuke Kataoka (Department of Pediatrics, Nagoya University Graduate School of Medicine, Nagoya, Japan) for valuable assistance. The authors would also like to express their great appreciation to Drs. Sayaka Mii, Daichi Fukumi, Tameo Hatano (Department of Pediatrics, Japanese Red Cross Nagoya Daiichi Hospital, Nagoya, Japan), Takahiro Okumura (Department of Cardiology, Nagoya University Graduate School of Medicine), Kentaro Omoya, Takashi Kuwahara (Department of Pediatric Cardiology, Gifu Prefectural General Medical Center, Gifu, Japan), Naoki Ohashi, Hiroshi Nishikawa, Masaki Matsushima (Department of Pediatric Cardiology, Chukyo Hospital, Nagoya, Japan), Takaya Ota, Kenji Kuraishi, Nobuo Tsuchi (Department of Pediatric Cardiology and Neonatology, Ogaki Municipal Hospital, Ogaki, Japan), and Noriko Nagai (Department of Pediatrics, Okazaki City Hospital, Okazaki, Japan) for their assistance with the collection of our samples and data. The authors would like to thank the Division for Medical Research Engineering, Nagoya University Graduate School of Medicine for technical support in next-generation sequencing. The authors also acknowledge the assistance of the Human Genome Center, Institute of Medical Science, University of Tokyo (<http://sc.hgc.jp/shirokane.html>), for providing supercomputing resources.

#### References



- [1] A.K. Ewart, C.A. Morris, G.J. Ensing, J. Loker, C. Moore, M. Leppert, et al., A human vascular disorder, supravalvular aortic stenosis, maps to chromosome 7, *Proc. Natl. Acad. Sci. U. S. A.* 90 (1993) 3226–3230.
- [2] A.J. Beuren, C. Schulze, P. Eberle, D. Harmjan, J. Apitz, The syndrome of supravalvular aortic stenosis, peripheral pulmonary stenosis, mental retardation and similar facial appearance, *Am. J. Cardiol.* 13 (1964) 471–483.
- [3] R. Eisenberg, D. Young, B. Jacobson, A. Boito, Familial supravalvular aortic stenosis, *Am. J. Dis. Child.* 108 (1964) 341–347.
- [4] J.C. Williams, B.G. Barratt-Boyes, J.B. Lowe, Supravalvular aortic stenosis, *Circulation* 24 (1961) 1311–1318.
- [5] B.R. Pober, Williams–Beuren syndrome, *N. Engl. J. Med.* 362 (2010) 239–252.
- [6] J.M.Z.J. Frieland-Little, R.J. Gajarski, Aortic stenosis, in: H.D.S.R. Allen, D.J. Penny, T.F. Feltes, F. Cetta (Eds.), *Moss & Adams Heart Disease in Infants, Children, and Adolescents: Including the Fetus and Young Adult*, 9th ed Wolters Kluwer, Philadelphia 2016, pp. 1085–1105.
- [7] D.Y. Li, A.E. Toland, B.B. Boak, D.L. Atkinson, G.J. Ensing, C.A. Morris, et al., Elastin point mutations cause an obstructive vascular disease, supravalvular aortic stenosis, *Hum. Mol. Genet.* 6 (1997) 1021–1028.
- [8] T.M. Olson, V.V. Michels, Z. Urban, K. Csiszar, A.M. Christiano, D.J. Driscoll, et al., A 30 kb deletion within the elastin gene results in familial supravalvular aortic stenosis, *Hum. Mol. Genet.* 4 (1995) 1677–1679.
- [9] A.K. Ewart, W. Jin, D. Atkinson, C.A. Morris, M.T. Keating, Supravalvular aortic stenosis associated with a deletion disrupting the elastin gene, *J. Clin. Invest.* 93 (1994) 1071–1077.
- [10] M.E. Curran, D.L. Atkinson, A.K. Ewart, C.A. Morris, M.F. Leppert, M.T. Keating, The elastin gene is disrupted by a translocation associated with supravalvular aortic stenosis, *Cell* 73 (1993) 159–168.
- [11] A.M. Jelsig, Z. Urban, V. Huchtagowder, H. Nissen, L.B. Ousager, Novel *ELN* mutation in a family with supravalvular aortic stenosis and intracranial aneurysm, *Eur. J. Med. Genet.* 60 (2017) 110–113.
- [12] A. Jakob, S. Unger, R. Arnold, J. Grohmann, C. Kraus, C. Schlensak, et al., A family with a new elastin gene mutation: broad clinical spectrum, including sudden cardiac death, *Cardiol. Young* 21 (2011) 62–65.
- [13] J. Katumba-Lunyanya, Two generations of identical twins with *ELN* deletion, *BMJ Case Rep.* 2009 (2009) (bcr06.2008.0036).
- [14] M. Martin, S. Secades, A.M. Plasencia, M.L. Rodriguez, C. Corros, A. Garcia-Campos, et al., Supravalvular aortic stenosis as a non-syndromic familial disease. Relevance of familial screening, *Int. J. Cardiol.* 172 (2014) 511–512.
- [15] G.M. Blue, E.P. Kirk, E. Giannoulatos, S.L. Dunwoodie, J.W. Ho, D.C. Hilton, et al., Targeted next-generation sequencing identifies pathogenic variants in familial congenital heart disease, *J. Am. Coll. Cardiol.* 64 (2014) 2498–2506.
- [16] T.M. Olson, V.V. Michels, N.M. Lindor, G.M. Pastores, J.L. Weber, D.J. Schaid, et al., Autosomal dominant supravalvular aortic stenosis: localization to chromosome 7, *Hum. Mol. Genet.* 2 (1993) 869–873.
- [17] Z. Urbán, V.V. Michels, S.N. Thibodeau, E.C. Davis, J.-P. Bonnefont, A. Munnich, et al., Isolated supravalvular aortic stenosis: functional haploinsufficiency of the elastin gene as a result of nonsense-mediated decay, *Hum. Genet.* 106 (2000) 577–588.
- [18] D.Y. Li, G. Faurly, D.G. Taylor, E.C. Davis, W.A. Boyle, R.P. Mecham, et al., Novel arterial pathology in mice and humans hemizygous for elastin, *J. Clin. Invest.* 102 (1998) 1783–1787.
- [19] D.Y. Li, B. Brooke, E.C. Davis, R.P. Mecham, L.K. Sorensen, B.B. Boak, et al., Elastin is an essential determinant of arterial morphogenesis, *Nature* 393 (1998) 276–280.

- [20] K. Metcalfe, A.K. Rucka, L. Smoot, G. Hofstadler, G. Tuzler, P. McKeown, et al., Elastin: mutational spectrum in supravalvular aortic stenosis, *Eur. J. Hum. Genet.* 8 (2000) 955–963.
- [21] L. Micale, M.G. Turturo, C. Fusco, B. Augello, L.A. Jurado, C. Izzi, et al., Identification and characterization of seven novel mutations of elastin gene in a cohort of patients affected by supravalvular aortic stenosis, *Eur. J. Hum. Genet.* 18 (2010) 317–323.
- [22] H. Muramatsu, Y. Okuno, K. Yoshida, Y. Shiraishi, S. Doisaki, A. Narita, et al., Clinical utility of next-generation sequencing for inherited bone marrow failure syndromes, *Genet. Med.* 19 (2017) 796–802.
- [23] H. Li, R. Durbin, Fast and accurate short read alignment with burrows-wheeler transform, *Bioinformatics* 25 (2009) 1754–1760.
- [24] D.C. Koboldt, Q. Zhang, D.E. Larson, D. Shen, M.D. McLellan, L. Lin, et al., VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing, *Genome Res.* 22 (2012) 568–576.
- [25] K. Wang, M. Li, H. Hakonarson, ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data, *Nucleic Acids Res.* 38 (2010) e164.
- [26] S. Richards, N. Aziz, S. Bale, D. Bick, S. Das, J. Gastier-Foster, et al., Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology, *Genet. Med.* 17 (2015) 405–424.
- [27] Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP). Seattle, WA, <http://evs.gs.washington.edu/EVS/>, Accessed date: 22 August 2017.
- [28] C. Genomes Project, A. Auton, L.D. Brooks, R.M. Durbin, E.P. Garrison, H.M. Kang, et al., A global reference for human genetic variation, *Nature* 526 (2015) 68–74.
- [29] M. Lek, K. Karczewski, E. Minikel, K. Samocha, E. Banks, T. Fennell, et al., Analysis of protein-coding genetic variation in 60,706 humans, *bioRxiv*, 2015.
- [30] K. Higasa, N. Miyake, J. Yoshimura, K. Okamura, T. Niihori, H. Saitsu, et al., Human genetic variation database, a reference database of genetic variations in the Japanese population, *J. Hum. Genet.* 61 (2016) 547–553.
- [31] J.J. Louw, G. Verleden, M. Gewillig, K. Devriendt, Haploinsufficiency of elastin gene may lead to familial cardiopathy and pulmonary emphysema, *Am. J. Med. Genet. A* 158a (2012) 2053–2054.
- [32] X. Ge, Y. Ren, O. Bartulos, M.Y. Lee, Z. Yue, K.Y. Kim, et al., Modeling supravalvular aortic stenosis syndrome with human induced pluripotent stem cells, *Circulation* 126 (2012) 1695–1704.
- [33] S. Park, E.J. Seo, H.W. Yoo, Y. Kim, Novel mutations in the human elastin gene (ELN) causing isolated supravalvular aortic stenosis, *Int. J. Mol. Med.* 18 (2006) 329–332.
- [34] L. Rodriguez-Revena, C. Badenas, A. Carrio, M. Mila, Elastin mutation screening in a group of patients affected by vascular abnormalities, *Pediatr. Cardiol.* 26 (2005) 827–831.
- [35] Z. Urban, S. Riazi, T.L. Seidl, J. Katahira, L.B. Smoot, D. Chitayat, et al., Connection between elastin haploinsufficiency and increased cell proliferation in patients with supravalvular aortic stenosis and Williams-Beuren syndrome, *Am. J. Hum. Genet.* 71 (2002) 30–44.
- [36] J. Dedic, A.S. Weiss, J. Katahira, B. Yu, R.J. Trent, Z. Urban, A novel elastin gene mutation (1281delC) in a family with supravalvular aortic stenosis: a mutation cluster within exon 20, *Hum. Mutat.* 17 (2001) 81.
- [37] T. Boeckel, A. Dierks, A. Vergopoulos, S. Bähring, H. Knoblauch, B. Müller-Myhsok, et al., A new mutation in the elastin gene causing supravalvular aortic stenosis, *Am. J. Cardiol.* 83 (1999) 1141–1143 (a9–10).
- [38] M. Tassabehji, K. Metcalfe, D. Donnai, J. Hurst, W. Reardon, M. Burch, et al., Elastin: genomic structure and point mutations in patients with supravalvular aortic stenosis, *Hum. Mol. Genet.* 6 (1997) 1029–1036.
- [39] C. Stamm, I. Friehs, S.Y. Ho, A.M. Moran, R.A. Jonas, P.J. del Nido, Congenital supravalvular aortic stenosis: a simple lesion? *Eur. J. Cardiothorac. Surg.* 19 (2001) 195–202.
- [40] G. Merla, N. Brunetti-Pierri, P. Piccolo, L. Micale, M.N. Loviglio, Supravalvular aortic stenosis: elastin arteriopathy, *Circ. Cardiovasc. Genet.* 5 (2012) 692–696.
- [41] Z. Urban, J. Zhang, E.C. Davis, G.K. Maeda, A. Kumar, H. Stalker, et al., Supravalvular aortic stenosis: genetic and molecular dissection of a complex mutation in the elastin gene, *Hum. Genet.* 109 (2001) 512–520.
- [42] L. Li, L. Huang, Y. Luo, X. Huang, S. Lin, Q. Fang, Differing microdeletion sizes and breakpoints in chromosome 7q11.23 in Williams-Beuren syndrome detected by chromosomal microarray analysis, *Mol. Syndromol.* 6 (2016) 268–275.
- [43] A. Wessel, R. Pankau, D. Kececioglu, W. Ruschewski, J.H. Bursch, Three decades of follow-up of aortic and pulmonary vascular lesions in the Williams-Beuren syndrome, *Am. J. Med. Genet.* 52 (1994) 297–301.
- [44] M. Delio, K. Pope, T. Wang, J. Samanich, C.R. Haldeman-Englert, P. Kaplan, et al., Spectrum of elastin sequence variants and cardiovascular phenotypes in 49 patients with Williams-Beuren syndrome, *Am. J. Med. Genet. A* 161A (2013) 527–533.



**RESEARCH ARTICLE**

# Chromosomally integrated human herpesvirus 6 in the Japanese population

Hiroki Miura MD<sup>1</sup> | Yoshiki Kawamura MD, PhD<sup>1</sup> | Fumihiko Hattori MD, PhD<sup>1</sup> |  
 Kei Kozawa MD<sup>1</sup> | Masaru Ihira PhD<sup>2</sup>  | Tamae Ohye PhD<sup>3</sup> |  
 Hiroki Kurahashi MD, PhD<sup>4</sup> | Tetsushi Yoshikawa MD, PhD<sup>1</sup> 

<sup>1</sup>Department of Pediatrics, Fujita Health University School of Medicine, Toyoake, Aichi, Japan

<sup>2</sup>Faculty of Clinical Engineering, Fujita Health University School of Health Sciences, Toyoake, Aichi, Japan

<sup>3</sup>Department of Clinical Laboratory Medicine, Graduate School of Health Sciences, Fujita Health University, Toyoake, Aichi, Japan

<sup>4</sup>Division of Molecular Genetics, Institute for Comprehensive Medical Science, Fujita Health University, Toyoake, Aichi, Japan

**Correspondence**

Tetsushi Yoshikawa, MD, PhD, Department of Pediatrics, Fujita Health University School of Medicine, 1-98 Dengakugakubo, Kutsukake-cho, Toyoake, Aichi, 470-1192 Japan.  
 Email: tetsushi@fujita-hu.ac.jp

**Funding information**

MEXT-Supported Program

The objectives of the work are to elucidate the incidence and virological findings of chromosomally integrated human herpesvirus 6 (ciHHV-6) in Japanese population and to analyze an association between ciHHV-6 and the clinical manifestation of exanthema subitum (ES). Real-time polymerase chain reaction was performed to determine HHV-6 DNA loads in 2347 cord blood samples from healthy neonates (cohort A), febrile children less than 5 years old (cohort B), and hematopoietic cell transplant recipients (cohort C). CiHHV-6 was confirmed by detection of high copy numbers of viral DNA in somatic cells. The integration site was determined by fluorescent in situ hybridization analysis. In the ciHHV-6 subjects of cohorts A and B, HHV-6 antibody titers were measured, the history of ES was obtained, and the incidence of ES was compared with non-ciHHV-6 children without primary HHV-6B infection in the cohort B. CiHHV-6 was detected in 14 (0.60%) of the 2347 samples: A (6/1006, 0.60%), B (6/790, 0.76%), and C (2/551, 0.36%). The integration sites were on chromosome 22q in seven cases, Yp in two cases, and 17q and Xp in one case. No past history of ES was observed in 11 of the 12 subjects. Nine children with ciHHV-6 underwent serological analysis and were found to be positive for HHV-6 IgG antibodies. Incidence of ES was statistically higher in the control subjects than the ciHHV-6 subjects ( $P = 0.0039$ ). In Japan, the frequency of ciHHV-6 was 0.60%. A high incidence of ciHHV-6A, specifically in chromosome 22, is a characteristic finding among the Japanese. CiHHV-6 may interfere with the clinical symptoms of primary HHV-6B infection.

**KEYWORDS**

antibody, chromosomally integrated human herpesvirus 6 (ciHHV-6), exanthema subitum (ES), Fluorescent in situ hybridization (FISH), integration site

## 1 | INTRODUCTION

Human herpesvirus 6 (HHV-6) is divided into two distinct species, HHV-6A and HHV-6B, that have an overall nucleotide sequence

identity of 90%.<sup>1-3</sup> Primary HHV-6B infection generally occurs in infancy and causes exanthema subitum (ES), which is a common febrile exanthematous disease.<sup>4,5</sup> In contrast to those of HHV-6B, the clinical features of primary HHV-6A infection remain unclear. Although horizontal transmission of HHV-6 is considered to be the main route of viral infection,<sup>6</sup> HHV-6 is genetically transmitted from a parent to child as an inherited chromosomally integrated HHV-6 (ciHHV-6).<sup>7-10</sup>

**Abbreviations:** ci, chromosomally integrated; ES, exanthema subitum; FISH, fluorescence in situ hybridization; HHV-6, human herpesvirus 6; NT, neutralization test.

Studies on viral reactivation from ciHHV-6 in immunocompromised patient<sup>11</sup> and pregnant women<sup>12,13</sup> have suggested that viral reactivation from the ciHHV-6 genome is similar to an individual latently infected with HHV-6B after primary viral infection. Moreover, *in vitro* experiments demonstrated the induction of viral reactivation in ciHHV-6-infected cells treated with a chemical compound.<sup>14</sup>

Individuals with ciHHV-6 demonstrate extremely high copy numbers of viral DNA, which may be more than 5.5 log<sub>10</sub> copies/mL in whole blood. Meanwhile, transient elevation of HHV-6B DNA loads in serum generally occurs in patients with primary HHV-6B infection, during viral reactivation.<sup>15</sup> High viral loads can cause misdiagnosis of active HHV-6 infection and unnecessary administration of antiviral drugs to patients.<sup>16</sup> Although the administration of unnecessary antiviral drugs is a major problem from a clinical point of view, the precise pathophysiological implications of ciHHV-6 are also important. Previous studies suggest that ciHHV-6 is associated with several clinical manifestations, such as encephalitis,<sup>17,18</sup> cognitive dysfunction and fatigue,<sup>19,20</sup> and angina pectoris.<sup>21</sup> The incidence of ciHHV-6 in the general population is approximately 0.8%–1.5% in the United States and United Kingdom<sup>22,23</sup> compared with 0.2% in Japan.<sup>7</sup> Such low incidence of ciHHV-6 has made elucidation of the clinical implications of this unique phenomenon difficult.

Our institute has been monitoring HHV-6 infection in two distinct patient populations—febrile children less than 5 years old visiting our outpatient clinic and emergency room for analysis of clinical features of primary HHV-6 infection, and hematopoietic stem cell transplant recipients. In addition, we have been exploring the isolation of HHV-6 DNA from cord blood mononuclear cells. Given the difference in incidences of ciHHV-6 in Japan versus the United States and Europe,<sup>7,8,15,22,23</sup> it is important to examine ciHHV-6-associated clinical problems in the Japanese population compared with the United States and Europe. Furthermore, it is very interesting to know what are the clinical symptoms of primary HHV-6 infection in ciHHV-6 individuals.

In the current study, we sought to elucidate the incidence of ciHHV-6, identify the species of HHV-6, and determine the chromosomal integration site in samples from cord blood, febrile children less than 5 years old, and hematopoietic stem cell transplant recipients. Furthermore, to determine whether ciHHV-6 impacts primary HHV-6B infection, we assessed the association between ciHHV-6 and the clinical manifestation of primary HHV-6B infection, ES.

## 2 | MATERIALS AND METHODS

### 2.1 | Subjects and samples

In this study, cord bloods harvested from healthy neonates (cohort A) born in Fujita Health University Hospital between April 1, 2013 and December 31, 2017 were used as samples for examining ciHHV-6. Serum samples were also harvested from febrile children less than 5 years old (cohort B) who visited our outpatient clinic or emergency room between June 1, 2014 and December 31, 2017. In addition, peripheral blood was collected from hematopoietic stem cell transplant recipients in Nagoya University Graduate School of Medicine and

Japanese Red Cross Nagoya First Hospital before transplant (cohort C) between September 1, 2004 and July 31, 2014, who received monitoring of HHV-6 infection at our institute. DNA was extracted from 200 µL of peripheral blood or serum samples using a QIAamp DNA Blood Mini kit (QIAGEN, Hilden, Germany). DNA samples were stored in -30°C until polymerase chain reaction (PCR) analysis.

### 2.2 | Determination of ciHHV-6

In the initial screening step, HHV-6 DNA loads in whole blood (cohorts A and C) or serum (cohort B) were measured by real-time PCR.<sup>24</sup> Additional whole blood samples were collected from ciHHV-6 subjects in cohort B for performing fluorescent *in situ* hybridization (FISH) assay and the HHV-6 DNA load in the whole blood was measured. When more than 1 × 10<sup>5</sup> copies/mL of viral DNA in whole blood or serum was detected, hair follicles and/or a buccal mucosa swab were collected from the subjects to confirm the presence of high copy numbers of viral DNA in the somatic cells. Finally, ciHHV-6 was confirmed by detection of high copy numbers of viral DNA in either whole blood or serum and somatic cells. To confirm inherited ciHHV-6, real-time PCR to measure HHV-6 DNA load was performed on hair follicle and/or buccal mucosa swab samples collected from the parents of the subjects in cohorts A and B. HHV-6 species was determined in the HHV-6 DNA positive samples based on restriction fragment length polymorphism analysis as previously described.<sup>25</sup>

### 2.3 | FISH analysis

To identify the site of chromosomal integration, 5 mL of heparinized peripheral blood was collected from the ciHHV-6 subjects for performing FISH analysis. The details of FISH analysis have been described elsewhere.<sup>26</sup> Briefly, phytohemagglutinin-stimulated lymphocytes or Epstein-Barr virus-transformed lymphoblasts were arrested by treatment with colcemid. Metaphase preparations were obtained by hypotonic treatment using 0.075 M KCl followed by methanol/acetate fixation. PCR products and a plasmid containing HHV-6 sequences were used as probes for FISH, which were labeled by nick-translation with biotin-16-dUTP or digoxigenin-11-dUTP. After hybridization, the probes were detected using either Alexa Fluor® 488-conjugated streptavidin or rhodamine-conjugated anti-digoxigenin, and visualized by counterstaining with 4',6-diamino-2-phenylindole. The following chromosome-specific probes were used as reference standards: human chromosome 22, RP11-186O8 (22q11.21) or TelVysion 22q SpectrumOrange (Abbott Molecular, Chicago, IL), CEP17 SpectrumGreen (Abbott Molecular), TelVysion Xq/Yq SpectrumOrange, and TelVysion Xp/Yp SpectrumGreen (Abbott Molecular).

### 2.4 | Past history of ES and serological assay for HHV-6

The past history of ES, which was defined as 3–4 days history of high fever and appearance of skin rash after the high fever subsided, was obtained from the parents of ciHHV-6 children in

cohorts A and B between January 1 and 31, 2018 by a telephonic interview. As a control group for comparison of the incidence of apparent infection of primary HHV-6B infection, the past history of ES was collected from 85 non-ciHHV-6 children without primary HHV-6B infection in cohort B by medical record, for whom information was available. HHV-6 IgG antibody titers were measured using an indirect immunofluorescence assay. The IgG antibody titer measured by immunofluorescence assay using HHV-6B-infected cells was defined as the reciprocal of the serum dilution that showed specific fluorescence.<sup>27</sup> Neutralization tests (NTs) against HHV-6A and HHV-6B were also carried out. Details of the NT have been previously described.<sup>28</sup> Cell-free virus for the NT was derived from the culture fluid of the representative strain (HHV-6A, U1102 strain and HHV-6B, HST strain) of virus-infected cord blood mononuclear cells. Serial two-fold serum dilutions were prepared on disposable plastic trays containing 96 wells and mixed with an equal volume (100  $\mu$ L) of the virus containing  $10^{2.5}$  tissue culture infective doses<sub>50</sub>. After 1 hour of incubation at 37°C,  $2 \times 10^5$  cord blood cells were added to each well, and incubation was continued for 7 days. The antibody titer was determined as the reciprocal of the dilution of serum that completely prohibited large cell formation.

## 2.5 | Statistical analysis

The statistical comparison of the incidence of ES and age at the time of the telephonic interview for investigating the past history of ES was evaluated using a Fisher exact test and a Wilcoxon signed-rank test,

respectively. Statistical significance was defined as two-sided  $P < 0.05$ . JMP version 12.2 (SAS Institute, Cary, NC) was used for analyses.

## 2.6 | Ethics approval

This study was approved by the Ethical Review Board Human Genome Studies at Fujita Health University (Accession number 90, approved on March 24, 2010). If the subjects were suspected of ciHHV-6 by HHV-6 DNA loads of blood samples, written informed consents were obtained from all of the participants or guardians of the subjects to carry out FISH analysis. Written informed consents were also collected from the parents for obtaining cord blood in cohort A. The consent of patients to participate in this study was obtained through an opt-out method in cohorts B and C.

## 3 | RESULTS

### 3.1 | Frequency of ciHHV-6 in Japan

A total of 2347 samples including whole blood (1557) or serum (790) samples were collected from the three cohorts. The numbers of subjects in cohorts A, B, and C were 1006, 790, and 551, respectively. Of the 2347 subjects, 14 subjects (0.60%) were identified as ciHHV-6 positive. The frequency of ciHHV-6 in cohorts A, B, and C were 0.60% (6/1006), 0.76% (6/790), and 0.36% (2/551), respectively. There was no statistical difference in the frequency of ciHHV-6 among the three cohorts ( $P = 0.598$ ). The clinical characteristics of ciHHV-6 cases are summarized in Table 1. Furthermore, both parents

**TABLE 1** Clinical characteristics of ciHHV-6 in the study population

Cohort	Case	Sex	Origin <sup>a</sup>	Species of integrated HHV-6	Site of integration	HHV-6 DNA copy number		
						Whole blood (copies/mL)	Buccal mucosa swab (copies/5 $\mu$ L DNA)	Hair follicle (copies/5 $\mu$ L DNA)
A	1	F	Maternal	HHV-6A	22q	$5.9 \times 10^6$	$4.6 \times 10^4$	$2.7 \times 10^3$
A	2	M	nd <sup>b</sup>	HHV-6A	nd	$4.6 \times 10^7$	$1.2 \times 10^4$	nd
A	3	M	Paternal	HHV-6B	22q	$1.6 \times 10^6$	$2.2 \times 10^5$	$1.5 \times 10^5$
A	4	M	Paternal	HHV-6B	nd	$9.8 \times 10^6$	$1.0 \times 10^5$	$2.0 \times 10^3$
A	5	F	Paternal	HHV-6A	22q	$1.9 \times 10^7$	$3.4 \times 10^5$	$1.3 \times 10^4$
A	6	M	Paternal	HHV-6B	Yp	$9.9 \times 10^7$	$3.1 \times 10^4$	$2.4 \times 10^4$
B	7	F	Paternal	HHV-6B	22q	$2.9 \times 10^8$	$1.3 \times 10^5$	$1.1 \times 10^4$
B	8	M	Paternal	HHV-6A	17q	$7.3 \times 10^7$	$3.5 \times 10^5$	$2.9 \times 10^3$
B	9	M	Paternal	HHV-6B	Yp	$7.0 \times 10^7$	$1.5 \times 10^4$	$1.5 \times 10^5$
B	10	F	Paternal	HHV-6A	22q	$2.4 \times 10^8$	$8.0 \times 10^4$	$1.1 \times 10^2$
B	11	F	Maternal	HHV-6A	nd	$1.0 \times 10^{6c}$	$1.5 \times 10^5$	$1.4 \times 10^4$
B	12	F	Paternal	HHV-6A	22q	$1.8 \times 10^8$	$8.3 \times 10^4$	$8.2 \times 10^3$
C	13	M	nd	HHV-6A	22q	$5.3 \times 10^6$	$1.8 \times 10^5$	nd
C	14	F	nd	HHV-6B	Xp	$1.9 \times 10^6$	$5.7 \times 10^4$	nd

ciHHV-6, chromosomally integrated human herpesvirus 6; F, Female; M, Male; N/A, not applicable; nd, not determined.

<sup>a</sup>Origin of chromosomally integrated HHV-6 is paternal or maternal.

<sup>b</sup>Mother was not ciHHV-6. Father's sample was not available.

<sup>c</sup>Serum.

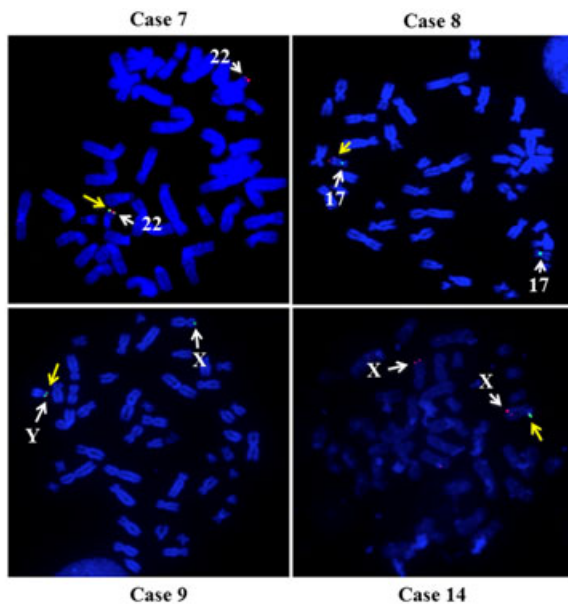
of 11 of 12 ciHHV-6 subjects in cohorts A and B were analyzed using real-time PCR. The parents' samples were not available in the two cases (cases 13 and 14) in cohort C, because these subjects were transplant recipients in other institute. At least one of the parents of each ciHHV-6 subject had high HHV-6 DNA loads, indicating inherited ciHHV-6. Only the mother was analyzed by real-time PCR in case 2, and HHV-6 DNA was not detected in hair follicle and buccal mucosa swab samples.

### 3.2 | Species and integration sites of ciHHV-6

As shown in Table 1, restriction fragment polymorphisms analyses indicated that 8 (57%) of the 14 ciHHV-6 cases were HHV-6A, and 6 (43%) of the 14 integrated cases were HHV-6B. Ten of the fourteen ciHHV-6 cases were analyzed by FISH analysis. The integration sites were on chromosomes 22q in seven cases, Yp in two cases, 17q in one case, and Xp in one case (Figure 1).

### 3.3 | Past history of ES and HHV-6 antibody titers

In cases 1 to 12, the past history of ES was collected from the parents by a telephonic interview. Case 1 had ES 18 months after birth. No past history of ES was reported in 11 of the 12 subjects (Table 2). Serological analysis was conducted on nine ciHHV-6 children (cases 1 and 5 to 12). Unfortunately, we were unable to collect serum samples for serological analyses of the other subjects. As shown in Table 3, all



**FIGURE 1** Characterization of ciHHV-6 by FISH. FISH analyses of metaphase chromosomes derived from the ciHHV-6 study subjects. A signal from the HHV-6 probe is detectable at the end of chromosome 22q (case 7), chromosome 17q (case 8), chromosome Yp (case 9), or chromosome Xp (case 14) (yellow arrowheads). The reference signals for chromosome 22q, 17q centromere, Xp/Yp, and Xq are indicated by white arrowheads. ciHHV-6, chromosomally integrated human herpesvirus 6; FISH, fluorescence in situ hybridization

**TABLE 2** Past history of exanthema subitum (ES)

Cohort	Case	Past history of ES	Age at the time of questionnaire for past history of ES
A	1	Yes	3 y 2 mo
A	2	No	2 y 10 mo
A	3	No	2 y 1 mo
A	4	No	2 y 0 mo
A	5	No	1 y 7 mo
A	6	No	0 y 6 mo
B	7	No	7 y 9 mo
B	8	No	6 y 11 mo
B	9	No	5 y 6 mo
B	10	No	3 y 4 mo
B	11	No	2 y 1 mo
B	12	No	1 y 5 mo
		Incidence of ES, no (%)	Median (IQR)
ciHHV-6 cases in cohort A and B		1/12 (8.3)	2 y 6 mo (1 y 8 mo-5 y 0 mo)
Non-ciHHV-6 children in cohort B		46/85 (54.1)	2 y 0 mo (1 y 3 mo-3 y 0 mo)
P value		<b>0.0039</b>	0.0696

m, month; N/A, not applicable; y, year.

Bold P value indicates statistical significance.

nine children were positive for the HHV-6 IgG antibody, as determined by immunofluorescence assay. In addition, both HHV-6A and B neutralizing antibodies were detected in cases 1, 6, 8, 10, and 12. Neutralizing antibodies against either HHV-6A or HHV-6B were detected in cases 5, 7, and 9. Despite detecting a low titer of HHV-6 IgG in case 11, we did not detect any neutralizing antibodies.

As shown in Table 2, 46 (54.1%) of the 85 control subjects, who were non-ciHHV-6 children without primary HHV-6B infection in cohort B, had a past history of ES. Incidence of ES was statistically higher in the control subjects than the ciHHV-6 subjects ( $P = 0.0039$ ). The ages for checking the past history of ES in the ciHHV-6 subjects and the control were 2 years 6 months old and 2 years 0 months old, respectively ( $P = 0.0696$ ).

## 4 | DISCUSSION

Given that the incidence of ciHHV-6 is extremely low, it may be difficult to determine the precise correlation between ciHHV-6 and its clinical manifestations. As we expected, there was no statistical difference in the incidence of ciHHV-6 among the three cohorts. Therefore, the incidence of ciHHV-6 was measured by using all of the subjects in the three cohorts, and it was 0.6%. A previous study in Japan demonstrated that the incidence of ciHHV-6 was 0.2%.<sup>7</sup> Thus, the incidence of ciHHV-6 in Japan appears to be low in comparison

**TABLE 3** Antibody titer against HHV-6 by IFA and NT

Cohort	Case	Species of integrated HHV-6	Age at the time of serological assay	IFA	NT	
					HHV-6A	HHV-6B
A	1	HHV-6A	3 y 6 mo	8	32	8
A	2	HHV-6A	N/A	nd	nd	nd
A	3	HHV-6B	N/A	nd	nd	nd
A	4	HHV-6B	N/A	nd	nd	nd
A	5	HHV-6A	3 mo	16	16	<8
A	6	HHV-6B	1 mo	32	64	32
B	7	HHV-6B	5 y 7 mo	16	<8	8
B	8	HHV-6A	5 y	16	16	16
B	9	HHV-6B	4 y 4 mo	64	8	<8
B	10	HHV-6A	2 y 8 mo	16	32	8
B	11	HHV-6A	10 mo	16	<8	<8
B	12	HHV-6A	1 y 3 mo	16	64	32

HHV-6, human herpesvirus 6; IFA, indirect immunofluorescence assay; m, month; N/A, not applicable; nd, not determined; NT, neutralizing test; y, year.

with previous studies in Western countries.<sup>22,23</sup> Previous studies revealed a similar frequency of ciHHV-6 between healthy controls and children with acute lymphoblastic leukemia,<sup>29</sup> patients with classical Hodgkin lymphoma,<sup>30</sup> and patients with breast cancer.<sup>31</sup> However, ciHHV-6 was associated with angina pectoris based on a large-genome cohort study.<sup>21</sup> Future studies examining adult patients, particularly samples collected from patients with angina pectoris, are needed to determine the role of ciHHV-6 in the pathogenesis of angina pectoris in the Japanese population.

As in the United States and Europe, HHV-6B infection is more prevalent than HHV-6A infection in Japan.<sup>32</sup> However, the frequency of ciHHV-6A (57%) was higher than ciHHV-6B (43%) in this study. In addition, we found seven ciHHV-6 cases in which the viral genome was integrated in chromosome 22q (five were ciHHV-6A and two were HHV-6B). To date, we have identified eight ciHHV-6A cases integrated in chromosome 22q and six ciHHV-6B cases integrated in chromosome 22q.<sup>11,26,33</sup> In contrast to our findings, most ciHHV-6 cases in the United States and Europe were integrated in chromosome 17.<sup>14,34-36</sup> These findings suggest distinct two possibilities: one is the founder effect and the other is that chromosome 22q is more susceptible to integration of the HHV-6 genome than other chromosomes. Recently, we demonstrated the founder effect for the integration of HHV-6A genome in chromosome 22q in Japanese population based on the microsatellite analysis of chromosome 22.<sup>37</sup> However, as some of the families had different microsatellites in that study, another possibility should be examined in the future study.

Most Japanese children acquire primary HHV-6B infection between 6 and 12 months after birth.<sup>27</sup> Timing of the telephonic interviews for past history of ES in ciHHV-6 subjects was considered to be older than a peak that of primary HHV-6B infection. Therefore, in this current study, it is considered that most of the ciHHV-6 subjects already received primary HHV-6B infection at the time of the telephonic interviews. In

our country, the majority of ES patients are generally diagnosed by physicians, and their parents remember past history of the disease. Therefore, telephonic interviews on the past history of the disease are considered reliable. Only 1 (8.3%) of the 12 cases reported ES (apparent infection), and no history of ES was detected in the seven patients older than 1 year in cohort A and B (case 5 and 7 to 12) despite their expression of HHV-6 IgG antibodies. Additionally, the past history of ES in the ciHHV-6 subjects was statistically lower than the subjects with past history of ES in the non-ciHHV-6 children ( $P = 0.0039$ ).

Although NT was still incomplete for removal of cross reaction between HHV-6A and B antibodies, a test was carried out to evaluate species-specific immune response in this study. In addition, all of the HHV-6 IgG antibody positive serum samples contained neutralizing antibodies against HHV-6A and/or B except case 11; indeed, it is possible that these antibodies might be induced by either integrated virus reactivation or infection of HHV-6B. Strenger et al<sup>38</sup> demonstrated that ciHHV-6 individuals have cell-mediated immune responses against HHV-6, which suggested that viral reactivation from an integrated viral genome may trigger the induction of host immune responses. These data suggest that host immune responses against HHV-6 probably induced by reactivation of the integrated virus may be associated with the low incidence of apparent primary HHV-6B infection (ES) in ciHHV-6 individuals. Although induced immune response against HHV-6 may cause autoimmune diseases, no correlation was demonstrated between ciHHV-6 and the diseases.<sup>19</sup> According to previous studies,<sup>11,33</sup> viral reactivation could occur in immunosuppressed ciHHV-6 individuals; indeed, an incidence of viral reactivation from the integrated viral genome is considered to be very low in immunocompetent ciHHV-6 individuals. Therefore, it is likely that synthesis of the target protein induced by viral reactivation is generally shut off.

The incidence of ciHHV-6 in the Japanese population is 0.6%. It was impossible to demonstrate the correlation between ciHHV-6 and



clinical manifestations due to the small number of the subjects for evaluation of such low incidence of phenomenon. A high incidence of ciHHV-6A, specifically in chromosome 22, is a characteristic finding in Japanese individuals with ciHHV-6. As the frequency of past history of ES was significantly lower in the ciHHV-6 subjects, ciHHV-6 possibly induced humoral immune responses may interfere with the clinical symptoms of primary HHV-6B infection.

## ACKNOWLEDGMENT

The authors thank Mrs Akiko Yoshikawa, Mrs Chieko Mori, and Mrs Yoko Osakabe for their technical support.

## CONFLICTS OF INTEREST

The authors have no conflicts of interest relevant to this article to disclose.

## AUTHORS' CONTRIBUTION

Sampling and collection of patients' information were carried out by Drs. Hiroki Miura, Yoshiki Kawamura, Fumihiko Hattori, Kei Kozawa and Tetsushi Yoshikawa. Drs. Hiroki Miura, Yoshiki Kawamura, and Masaru Ihira performed real-time PCR and serological analysis. FISH analysis was carried out by Drs. Tamae Ohye and Hiroki Kurahashi. Dr. Hiroki Miura drafted the initial manuscript, and Prof. Tetsushi Yoshikawa reviewed and revised the manuscript. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

## FUNDING

This study was supported by the MEXT-Supported Program for the Strategic Research Foundation at Private Universities (HK) Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

## FINANCIAL DISCLOSURE

The authors have no financial relationships relevant to this article to disclose.

## ORCID

Masaru Ihira  <http://orcid.org/0000-0002-1612-7060>

Tetsushi Yoshikawa  <http://orcid.org/0000-0002-2847-7682>

## REFERENCES

1. Ablashi D, Agut H, Berneman Z, et al. Human herpesvirus-6 strain groups: a nomenclature. *Arch Virol.* 1993;129(1-4):363-366.
2. Dominguez G, Dambaugh TR, Stamey FR, Dewhurst S, Inoue N, Pellett PE. Human herpesvirus 6B genome sequence: coding content and comparison with human herpesvirus 6A. *J Virol.* 1999;73(10):8040-8052.
3. Isegawa Y, Mukai T, Nakano K, et al. Comparison of the complete DNA sequences of human herpesvirus 6 variants A and B. *J Virol.* 1999;73(10):8053-8063.
4. Yamanishi K, Shiraki K, Kondo T, et al. Identification of human herpesvirus-6 as a causal agent for exanthem subitum. *Lancet.* 1988;1(8594):1065-1067.
5. Asano Y, Suga S, Yoshikawa T, Urisu A, Yazaki T. Human herpesvirus type 6 infection (exanthem subitum) without fever. *J Pediatr.* 1989;115(2):264-265.
6. Mukai T, Yamamoto T, Kondo T, et al. Molecular epidemiological studies of human herpesvirus 6 in families. *J Med Virol.* 1994; 42(3):224-227.
7. Tanaka-Taya K, Sashihara J, Kurahashi H, et al. Human herpesvirus 6 (HHV-6) is transmitted from parent to child in an integrated form and characterization of cases with chromosomally integrated HHV-6 DNA. *J Med Virol.* 2004;73(3):465-473.
8. Hall CB, Caserta MT, Schnabel K, et al. Chromosomal integration of human herpesvirus 6 is the major mode of congenital human herpesvirus 6 infection. *Pediatrics.* 2008;122(3):513-520.
9. Morissette G, Flamand L. Herpesviruses and chromosomal integration. *J Virol.* 2010;84(23):12100-12109.
10. Clark DA. Clinical and laboratory features of human herpesvirus 6 chromosomal integration. *Clin Microbiol Infect.* 2016;22(4):333-339.
11. Endo A, Watanabe K, Ohye T, et al. Molecular and virological evidence of viral activation from chromosomally integrated human herpesvirus 6A in a patient with X-linked severe combined immunodeficiency. *Clin Infect Dis.* 2014;59(4):545-548.
12. Hall CB, Caserta MT, Schnabel KC, et al. Transplacental congenital human herpesvirus 6 infection caused by maternal chromosomally integrated virus. *J Infect Dis.* 2010;201(4):505-507.
13. Gravel A, Hall CB, Flamand L. Sequence analysis of transplacentally acquired human herpesvirus 6 DNA is consistent with transmission of a chromosomally integrated reactivated virus. *J Infect Dis.* 2013; 207(10):1585-1589.
14. Arbuckle JH, Medveczky MM, Luka J, et al. The latent human herpesvirus-6A genome specifically integrates in telomeres of human chromosomes in vivo and in vitro. *Proc Natl Acad Sci USA.* 2010; 107(12):5563-5568.
15. Pellett PE, Ablashi DV, Ambros PF, et al. Chromosomally integrated human herpesvirus 6: questions and answers. *Rev Med Virol.* 2012; 22(3):144-155.
16. Lee SO, Brown RA, Razonable RR. Chromosomally integrated human herpesvirus-6 in transplant recipients. *Transpl Infect Dis.* 2012; 14(4):346-354.
17. Troy SB, Blackburn BG, Yeom K, Finley caulfield AK, Bhangoo MS, Montoya JG. Severe encephalomyelitis in an immunocompetent adult with chromosomally integrated human herpesvirus 6 and clinical response to treatment with foscarnet plus ganciclovir. *Clin Infect Dis.* 2008;47(12):e93-e96.
18. Wittekindt B, Berger A, Porto L, et al. Human herpes virus-6 DNA in cerebrospinal fluid of children undergoing therapy for acute leukaemia. *Br J Haematol.* 2009;145(4):542-545.
19. Pantry SN, Medveczky MM, Arbuckle JH, et al. Persistent human herpesvirus-6 infection in patients with an inherited form of the virus. *J Med Virol.* 2013;85(11):1940-1946.
20. Montoya JG, Neely MN, Gupta S, et al. Antiviral therapy of two patients with chromosomally-integrated human herpesvirus-6A presenting with cognitive dysfunction. *J Clin Virol.* 2012;55(1):40-45.
21. Gravel A, Dubuc I, Morissette G, Sedlak RH, Jerome KR, Flamand L. Inherited chromosomally integrated human herpesvirus 6 as a predisposing risk factor for the development of angina pectoris. *Proc Natl Acad Sci USA.* 2015;112(26):8058-8063.

22. Ward KN, Thiruchelvam AD, Couto-Parada X. Unexpected occasional persistence of high levels of HHV-6 DNA in sera: detection of variants A and B. *J Med Virol.* 2005;76(4):563-570.
23. Nam leong H, Tuke PW, Tedder RS, et al. The prevalence of chromosomally integrated human herpesvirus 6 genomes in the blood of UK blood donors. *J Med Virol.* 2007;79(1):45-51.
24. Tanaka N, Kimura H, Hoshino Y, et al. Monitoring four herpesviruses in unrelated cord blood transplantation. *Bone Marrow Transplant.* 2000;26(11):1193-1197.
25. Ihira M, Ohta A, Sugata K, Suga S, Asano Y, Yoshikawa T. Loop-mediated isothermal amplification for discriminating between human herpesvirus 6 A and B. *J Virol Methods.* 2008;154(1-2):223-225.
26. Ohye T, Kawamura Y, Inagaki H, et al. A simple cytogenetic method to detect chromosomally integrated human herpesvirus-6. *J Virol Methods.* 2016;228:74-78.
27. Yoshikawa T, Suga S, Asano Y, Yazaki T, Kodama H, Ozaki T. Distribution of antibodies to a causative agent of exanthem subitum (human herpesvirus-6) in healthy individuals. *Pediatrics.* 1989;84(4):675-677.
28. Suga S, Yoshikawa T, Asano Y, Yazaki T, Ozaki T. Neutralizing antibody assay for human herpesvirus-6. *J Med Virol.* 1990;30(1):14-19.
29. Gravel A, Sinnott D, Flamand L. Frequency of chromosomally-integrated human herpesvirus 6 in children with acute lymphoblastic leukemia. *PLoS One.* 2013;8(12):e84322.
30. Bell AJ, Gallagher A, Mottram T, et al. Germ-line transmitted, chromosomally integrated HHV-6 and classical Hodgkin lymphoma. *PLoS One.* 2014;9(11):e112642.
31. Gravel A, Dubuc I, Brooks-Wilson A, et al. Inherited chromosomally integrated human herpesvirus 6 and breast cancer. *Cancer Epidemiol Biomarkers Prev.* 2017;26(3):425-427.
32. Ablashi D, Agut H, Alvarez-Lafuente R, et al. Classification of HHV-6A and HHV-6B as distinct viruses. *Arch Virol.* 2014;159(5):863-870.
33. Miura H, Kawamura Y, Kudo K, et al. Virological analysis of inherited chromosomally integrated human herpesvirus-6 in three hematopoietic stem cell transplant patients. *Transpl Infect Dis.* 2015;17(5):728-731.
34. Torelli G, Barozzi P, Marasca R, et al. Targeted integration of human herpesvirus 6 in the p arm of chromosome 17 of human peripheral blood mononuclear cells in vivo. *J Med Virol.* 1995;46(3):178-188.
35. Clark DA, Nacheva EP, Leong HN, et al. Transmission of integrated human herpesvirus 6 through stem cell transplantation: implications for laboratory diagnosis. *J Infect Dis.* 2006;193(7):912-916.
36. Strenger V, Caselli E, Lautenschlager I, et al. Detection of HHV-6-specific mRNA and antigens in PBMCs of individuals with chromosomally integrated HHV-6 (ciHHV-6). *Clin Microbiol Infect.* 2014;20(10):1027-1032.
37. Kawamura Y, Ohye T, Miura H, et al. Analysis of the origin of inherited chromosomally integrated human herpesvirus 6 in the Japanese population. *J Gen Virol.* 2017;98:1823-1830.
38. Strenger V, Kayser S, Witte KE, et al. Individuals with inherited chromosomally integrated human herpes virus 6 (ciHHV-6) have functionally active HHV-6 specific T-cell immunity. *Clin Microbiol Infect.* 2016;22(2):209.e5-209.e8.

**How to cite this article:** Miura H, Kawamura Y, Hattori F, et al. Chromosomally integrated human herpesvirus 6 in the Japanese population. *J Med Virol.* 2018;90:1636-1642.  
<https://doi.org/10.1002/jmv.25244>