



**HAL**  
open science

## **Congenital Muscular Dystrophy type 1D (MDC1D) due to a large intragenic insertion/deletion involving intron 10 of the gene**

Pascale Guicheney, Nigel F Clarke, Svetlana Maugenre, Aurélie Vandebrouck, J Andoni Urtizbera, Tobias Willer, Rachel Peat, Françoise Gray, Céline Bouchet, Hiroshi Manya, et al.

► **To cite this version:**

Pascale Guicheney, Nigel F Clarke, Svetlana Maugenre, Aurélie Vandebrouck, J Andoni Urtizbera, et al.. Congenital Muscular Dystrophy type 1D (MDC1D) due to a large intragenic insertion/deletion involving intron 10 of the gene. *European Journal of Human Genetics*, 2011, 10.1038/ejhg.2010.212 . hal-00609410

**HAL Id: hal-00609410**

**<https://hal.science/hal-00609410>**

Submitted on 19 Jul 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Congenital Muscular Dystrophy type 1D (MDC1D) due to a large intragenic insertion/deletion involving intron 10 of the *LARGE* gene**

Nigel F Clarke, MB ChB, PhD<sup>1,2\*</sup>, Svetlana Maugenre, MS<sup>1,2\*</sup>, Aurélie Vandebrouck, PhD<sup>1,2</sup>, J. Andoni Urtizberea, MD<sup>3</sup>, Tobias Willer, PhD<sup>4</sup>, Rachel Peat, PhD<sup>1,2</sup>, Françoise Gray, MD, PhD<sup>5</sup>, Céline Bouchet, PhD<sup>6</sup>, Hiroshi Manyá, PhD<sup>7</sup>, Sandrine Vuillaumier-Barrot, PhD<sup>6</sup>, Tamao Endo, PhD<sup>7</sup>, Eliane Chouery, PhD<sup>8</sup>, Kevin P. Campbell, PhD<sup>4</sup>, André Mégarbané, MD, PhD<sup>8,9,10</sup>, Pascale Guicheney, PhD<sup>1,2</sup>

1. Inserm, U956, Faculté de médecine Pierre et Marie Curie, Paris, France
2. UPMC Univ Paris 06, UMR\_S956, IFR14, Paris, France
3. AP-HP, Hôpital Marin, Hendaye, France
4. Howard Hughes Medical Institute, Departments of Molecular Physiology and Biophysics, Neurology, and Internal Medicine, University of Iowa Roy J. and Lucille A. Carver College of Medicine, Iowa City, Iowa 52242-1101 USA
5. AP-HP, Hôpital Lariboisière, Service Central d'Anatomie et de Cytologie Pathologiques, Paris, France
6. AP-HP, Hôpital Bichat-Claude Bernard, Biochimie Métabolique, Paris, France
7. Tokyo Metropolitan Institute of Gerontology, Glycobiology Research Group, Foundation for Research on Aging and Promotion of Human Welfare, Itabashi-ku, Tokyo, Japan
8. Université Saint Joseph, Faculté de Médecine Unité de Génétique Médicale, Beirut, Lebanon
9. Institut Jérôme Lejeune, Paris, France
10. Inserm UMR 910, Génétique Médicale et Génomique Fonctionnelle, Faculté de Médecine de la Timone, Marseille, France.

\* These authors wish to be considered equal first authors

For correspondence:

Dr Pascale Guicheney, Unité UMR\_S956, Faculté de Médecine Pierre et Marie Curie, site

Pitié-Salpêtrière, 91 boulevard de l'Hôpital, 75634 Paris Cedex 13, France

Email: pascale.guicheney@upmc.fr

Tel: (33) 1 40 77 98 05 Fax: (33) 1 40 77 96 45

Running title: MDC1D due to a large intragenic insertion/deletion

Word counts

Title: 128 characters (including spaces)

Abstract: 223 words

Body of paper: 3158 words

Key words: congenital muscular dystrophy 1D, alpha-dystroglycan, DNA duplication,  
muscle-eye-brain disease

For publication: Tables: 0 Figures: 5

Supplementary data: Tables: 1 Figures: 2

**Abstract**

Mutation of the *LARGE* gene is the rarest of the six known genetic causes of  $\alpha$ -dystroglycanopathy. We report a further family with MDC1D due to a complex genomic rearrangement that was not apparent on standard sequencing of *LARGE*. Two sisters in a consanguineous family had moderate mental retardation and cerebellar malformations together with dystrophic changes and markedly reduced  $\alpha$ -dystroglycan glycosylation staining on muscle biopsy. There was homozygous linkage to the *LARGE* locus but sequencing of *LARGE* coding regions was normal. Analysis of *LARGE* cDNA showed an abnormal sequence inserted between exons 10 and 11 in most transcripts, predicted to introduce a premature stop codon. The abnormal sequence mapped to a spliced EST (DA935254) of unknown function, normally located 100 kb centromeric of *LARGE* on chromosome 22q12.3. Quantitative PCR analysis of the EST and adjacent regions showed twice the normal copy number in patient genomic DNA, consistent with a large intra-chromosomal duplication inserted into intron 10 of *LARGE* in a homozygous state. This insertion was associated with deletion of a central region of intron 10 but the exact breakpoints of the deletion/duplication were not found suggesting that an even more complex rearrangement may have occurred. The exact function of *LARGE*, a Golgi protein, remains uncertain. POMT and POMGnT enzyme activities were normal in patient lymphoblasts, suggesting that defects in *LARGE* do not affect initiation of *O*-mannosyl-glycans.

## 1. Introduction

Alpha-dystroglycan ( $\alpha$ DG) is a highly glycosylated cell surface protein that plays important roles in neuronal cell migration and cell-to-extracellular matrix interactions in muscle. Abnormalities in the glycosylation of  $\alpha$ DG is the hallmark histological abnormality and likely pathogenic mechanism for a group of congenital muscular dystrophies (CMD), collectively called the ' $\alpha$ -dystroglycanopathies'.<sup>1</sup> At present, six genes have been linked to this common clinicopathological presentation (*POMT1*, *POMT2*, *POMGNT1*, *FKTN*, *FKRP* and *LARGE*). All of these encode confirmed or putative enzymes that are thought to function in the *O*-mannosyl glycosylation pathway. Abnormal  $\alpha$ DG glycosylation correlates with an increasingly severe pattern of tissue effects that extends from an adult-onset limb-girdle muscular dystrophy with normal intelligence (LGMD type 2I) to severe congenital brain and eye malformations (Walker-Warburg syndrome, WWS).<sup>1</sup> The *LARGE* gene became a candidate for human CMD after the discovery of the myodystrophic (*myd*) mouse that has a spontaneous null mutation in *Large* due to a genomic deletion involving exons 5 to 7.<sup>2-4</sup> Another spontaneous murine mutant, named veils (*vls*) also results from an intragenic deletion leading to loss of exons 3 to 5 and premature protein termination.<sup>5</sup> The two strains share typical cerebral, ocular and muscle changes of an  $\alpha$ -dystroglycanopathy. MDC1D due to mutations in *LARGE* is the rarest of the  $\alpha$ -dystroglycanopathies identified to date, with only three confirmed families previously described.<sup>6-8</sup> Affected children have had typical neurological and muscle abnormalities associated with the  $\alpha$ -dystroglycanopathies, but with very different severities; one family had mild muscle-eye-brain disease (MEB) and the other two had typical WWS. Another WWS patient has been reported with a single heterozygous nonsense mutation in *LARGE*.<sup>9</sup> Confirmation that this patient has CMD due to *LARGE* awaits the discovery of a second mutation. In this paper, we describe the fourth confirmed family with MDC1D due to a homozygous *LARGE* mutation, in which two sisters from

consanguineous parents have moderate mental retardation, cerebellar and pontine hypoplasia and persistent white matter abnormalities on brain MRI, features consistent with mild MEB. We identified an unusual genetic mechanism for *LARGE* gene disruption, namely the splicing of an abnormal sequence into the *LARGE* mRNA transcript likely due to a duplication/deletion event involving intron 10. This report extends our knowledge of the clinical phenotype associated with *LARGE*, and is the second family where disease results from a large scale gene rearrangement.

The *POMT1*, *POMT2* and *POMGNT1* genes encode enzymes that together mediate the first two steps in the *O*-mannosyl glycosylation pathway. A recent study suggests that *LARGE* participates in post-phosphoryl glycosylation of a novel phosphorylated *O*-mannosyl glycan on the mucin-like domain of recombinant  $\alpha$ DG, which is required for laminin binding.<sup>10</sup> Evidence from cell culture and a *fukutin* transgenic mouse suggests that fukutin physically interacts with *POMGnT1* and can cause reduced *POMGnT1* activity when abnormal.<sup>11</sup> To investigate whether *LARGE* may also act by influencing the first steps of the *O*-mannosyl pathway, we assessed the activity levels of *POMT* and *POMGnT1* in cultured lymphoblasts from one of our patients.

## **2. Case reports**

The two affected girls are the only children of first-cousin parents from Lebanon. There is no other family history of neuromuscular disease. Both children were born full-term after uneventful pregnancies with normal fetal movements and normal deliveries. Birth weights, lengths and head circumferences were within normal limits.

*Patient 1:* The older sibling first presented to medical attention at age 2 ½ months with hypotonia and regurgitation. Acquisition of all motor skills was delayed. She sat at age 9 months, stood unsupported at age 2½ years and at age 8 years she could climb stairs with difficulty but could not run. Her full scale IQ was 61. On examination at age 8 years, she weighed 24 kg (50<sup>th</sup> centile), her height was 122.5 cm (35<sup>th</sup> centile) and the occipitofrontal circumference (OFC) was 53 cm (75<sup>th</sup> centile). Her speech was restricted to dysarthric, simple words and she followed simple commands. She had a myopathic face with an open-mouth posture, and mild tongue hypertrophy. There was generalised muscle hypertrophy, most prominent in the calves, thighs and shoulder girdle. This contrasted with moderate proximal lower limb weakness that manifested as a waddling gait and a positive Gowers' manoeuvre. Upper limb strength was within normal limits. There were no joint contractures, except for mild tightness of the Achilles' tendons, associated with toe-walking. The only abnormalities found on eye assessment were mild myopia and strabismus. There was no history of seizures. Cardiac and respiratory examinations were normal. Brain Magnetic Resonance Imaging (MRI) performed at ages 3.5 and 6 years showed marked cerebellar atrophy (particularly affecting the vermis) and cerebellar cysts (figure 1). There was mild generalized cerebral and pontine atrophy, dilatation of all ventricular spaces and pachygyria of the frontal lobes (figure 1). Diffuse symmetrical high signal abnormalities were seen within the periventricular and deep white matter of both cerebral hemispheres, with sparing of the subcortical U fibres. These hyperintensities were unchanged at age 6 years. A generalized skeletal X-ray survey, abdominal ultrasound, echocardiogram, electroencephalogram, and auditory evoked potentials with brainstem responses were all normal. Creatine kinase (CK) levels were 30 times higher than the upper limit of normal.

*Patient 2:* The younger sister followed a similar clinical course to her sister but was less delayed in her motor and cognitive skills. She stood unsupported at age 16 months and

walked independently at age 2 years. On examination at age 4½ years, her weight was 15 kg (25<sup>th</sup> centile), height 106 cm (75<sup>th</sup> centile) and OFC 51 cm (50<sup>th</sup> centile). There was generalised hypotonia, firm hypertrophic calves and a mild equinovarus foot position. Mild myopia was the only abnormality on eye examination. Myopathic features were noted on EMG while nerve conduction studies were normal. Brain MRI performed at 2 years showed diffuse high signal white matter abnormalities mainly in the frontal and parieto-occipital regions, dilated ventricular spaces and marked cerebellar hypoplasia. CK levels were 60 times the upper limit of normal.

### ***Muscle biopsy findings***

Muscle and skin biopsies were taken from Patient 1 from the right gastrocnemius muscle. Standard histochemical analysis revealed a dystrophic pattern characterised by fibre size variation, markedly increased connective tissue, internalised nuclei and scattered necrotic fibres. Immunohistochemistry (IHC) for  $\alpha$ DG showed markedly reduced staining using VIA4-1 (Millipore, Billerica, MA, USA), an antibody that recognizes a glycosylated epitope. Laminin  $\alpha$ 2 expression was mildly reduced by IHC using antibodies that recognize both the 80 kDa (clone 5H2; Millipore, Billerica, MA, USA) and the 300 kDa (NCL-merosin; Novocastra Labs, Newcastle, UK) fragments. Staining for  $\beta$ -dystroglycan ( $\beta$ DG), dystrophin, collagen VI, and sarcoglycans ( $\alpha$  and  $\gamma$ ) was normal.

### **3. Western blot analysis**

Western blotting was used to investigate the degree of abnormal  $\alpha$ DG glycosylation using a skin fibroblast cell line from Patient 1 as frozen muscle was not available. Fibroblasts were maintained at 37°C and 5% CO<sub>2</sub> in DMEM medium plus 20% fetal bovine serum, 0.5% penicillin-streptomycin (Gibco). Cultured cells were lysed and glycoproteins from cell



homogenates were enriched with WGA lectin as previously described.<sup>11</sup> We performed immunoblots (for  $\alpha$ DG and  $\beta$ DG) and laminin ligand overlay assays using polyvinylidene difluoride membranes and blots were developed by horseradish peroxidase (HRP) enhanced chemiluminescence (Pierce).<sup>12</sup> The monoclonal antibodies IIH6 and VIA4-1, that are specific for functionally glycosylated  $\alpha$ DG,<sup>13</sup> showed no detectable binding affinity to  $\alpha$ DG from patient fibroblast cells even though  $\beta$ DG (monoclonal antibody 7D11) was normally present, consistent with a glycosylation defect in  $\alpha$ DG (figure 2). This defect in  $\alpha$ DG glycosylation also resulted in loss of laminin-1 ligand binding (figure 2).

#### 4. Genetic analysis

We performed linkage analysis using microsatellite markers to the five  $\alpha$ -dystroglycanopathy loci known at the time (*POMGNT1*, *FKRP*, *POMT1*, *POMT2* and *FKTN*) and the *DAG1* gene that encodes DG but we identified no homozygous linkage. A genome-wide linkage study using 400 microsatellite markers showed potential linkage to the newly identified *LARGE* locus and analysis of further markers supported homozygous linkage (figure 3). We sequenced the coding regions of *LARGE* bi-directionally by standard techniques. Oligonucleotide primers at least 80 bp outside the exons were chosen for PCR amplification and sequencing on ABI3730 capillary electrophoresis system (Applied Biosystems), but no abnormalities were found. Subsequently mRNA was extracted from cultured patient EBV-transformed lymphoblasts from Patient 1 (Trizol method), cDNA was generated using oligodT Superscript II First-Strand Synthesis (Invitrogen, Carlsbad, CA, USA) and the *LARGE* gene transcript was sequenced using 14 overlapping primer pairs. Two abnormally large PCR products, together with a faint product of the expected size were obtained for primers that spanned exons 9 and 12 (see figure 4A). Sequencing revealed that most mRNA transcripts contained an abnormal 197 base pair (bp) sequence inserted between exons 10 and

11. Low levels of a 171 bp insertion (identical to the 197 bp sequence except lacking 26 bases from the 3' end) and a normally spliced transcript were also obtained. Identical splicing abnormalities were also seen in mRNA from cultured patient fibroblasts, except a greater proportion of transcripts showed normal splicing (figure 4A). The abnormal sequences mapped to part of a spliced EST (GenBank Accession No: DA935254) of unknown function, originally identified in a human splenic cDNA library,<sup>14</sup> normally located 100 kb centromeric to the *LARGE* gene on chromosome 22q12.3 (figure 5). The presence of the two most abundant abnormal inserted sequences are predicted to immediately introduce a premature stop codon in the *LARGE* mRNA transcript, resulting in truncation of the *LARGE* protein mid-way through translation (figure 4B). Analysis of the inserted EST sequence identified potential splice acceptor and donor sites for both the larger and smaller insertions that likely mediate the abnormal splicing (E-figure 1). Archived frozen muscle was no longer available to verify the abnormalities in *LARGE* splicing in muscle.

To investigate whether the abnormal *LARGE* splicing arose due to a chromosomal duplication, we performed quantitative PCR (qPCR) analysis of the EST sequence from patient genomic DNA (gDNA) using the LightCycler® 480 (Roche Diagnostics, Basel, Switzerland) and the SYBR Green I detection system. We normalised gDNA concentrations between individual samples using PCR primers for a distant gene (*PTPLA* on chromosome 10). E-Table 1 lists the key primer pairs used. The EST (DA935254) was present in double the expected copy number in both patients compared with two healthy unrelated individuals (figure 4C), consistent with a homozygous duplication of the EST sequence in both patients (ie they have four copies instead of two). To investigate the size of the putative duplication, we repeated qPCR analysis for genomic regions at variable distances up- and downstream of the EST sequence (E-Table 1). These studies indicate that between 40825 and 42937 bp of

genomic DNA surrounding the EST sequence is present at twice normal levels in gDNA in both patients (figure 4C). Both parents had 1.5 times the normal copy number, consistent with each having one copy of the chromosome 22 duplication and a normal allele. The 3' breakpoint is likely located in a 1313 bp region between regions 3b and 3c (Chr 22:33548218- 33549530; GRCh37/hg19 assembly) and the 5' breakpoint in a 799 bp region between regions 5b and 5c (Chr 22:33590356-33591154).

Together, these results indicate that the patients are homozygous for a duplication of gDNA around EST (DA935254) and the most likely site of insertion for the duplicated DNA is intron 10 of *LARGE*, leading to the abnormal splicing of the EST into the *LARGE* mRNA transcript between exons 10 and 11.

To identify the site of the insertion in intron 10 of *LARGE*, we designed twelve pairs of overlapping primers (E-figure 2) to cover the whole length of the intron (44117 bp). With two primer pairs (5aF-5aR, 6aF-6aR) we amplified the expected bands in controls and parents but not in the two patients indicating that the insertion should be in this region (E-figure 2). We repeated the same strategy with several other primer pairs chosen inside this region and in the bounding regions to finally determine a region of 3012 to 4175 bp that could never be amplified in the patients, which likely corresponds to a deleted region associated with the insertion. By qPCR analysis using primer pairs L5F-L6R and L6F and L7R (Supplemental E-Table 1), both parents had reduced amplification compared to controls, consistent with them being heterozygous for a deletion in this region. We were not able to amplify any product with primers chosen on the borders of the insertion and on the non deleted sides of the intron, and thus did not identify the exact breakpoints for the insertion/deletion rearrangement.

As repetitive sequences are implicated in the pathogenesis of recurrent chromosomal deletion or duplication syndromes (eg. CMT1A)<sup>15</sup> we analysed intron 10 and the duplicated region around EST DA935254 for homologous sequences. Both regions shared multiple SINE, LINE and LTR repeats that may have played roles in the genesis of the mutation, but no extended regions of homology.

### **5. POMT and POMGnT1 activity in patient lymphoblasts**

The POMT and POMGnT1 enzymes mediate the first and second steps respectively of the *O*-mannosyl glycosylation of  $\alpha$ DG. To investigate whether *LARGE* mutations may disrupt the glycosylation of  $\alpha$ DG by influencing the activity of either POMTs or POMGnT1, we assessed the enzymatic activity of these enzymes in cultured patient lymphoblasts using previously validated assays.<sup>16</sup> The POMT activity was based on the amount of [<sup>3</sup>H]mannose transferred from [<sup>3</sup>H]mannosylphosphoryldolichol to a glutathione-S-transferase fusion  $\alpha$ DG, and the POMGnT1 activity was based on the amount of [<sup>3</sup>H]GlcNAc transferred from UDP-[<sup>3</sup>H]GlcNAc to a mannosylpeptide. In brief, we incubated microsomal membrane fractions prepared from patient and control EBV-immortalised B lymphocytes with an acceptor substrate and radio-labelled sugar donor substrate. We then assessed the incorporation of radioactivity into each acceptor substrate. We normalised control and patient samples using the activity level of GnT1, a Golgi enzyme involved in the *N*-glycosylation pathway, which functions independently of *O*-glycosylation. Results showed no difference between the patient and eight control samples for POMT/GnT1 activity (patient 0.283 vs control 0.287  $\pm$  0.067 standard deviation) or for POMGnT1/GnT1 activity (patient 0.068 vs control 0.071  $\pm$  0.019).

### **6. Discussion**

In this paper we report the fourth confirmed family with MDC1D due to mutations in *LARGE*. The clinical abnormalities are typical for the  $\alpha$ -dystroglycanopathy group of disorders. Both sisters have typical features of an alpha-dystroglycanopathy, namely raised CK levels, dystrophic changes on muscle histology associated with reduced glycosylated alpha-dystroglycan staining, intellectual disability and brain malformations. In terms of severity of neurological involvement, the two sisters are best classified as having mild MEB disease, even though there were only minor ocular abnormalities. The findings were similar to the first family identified with mutations in *LARGE* with mild/moderate developmental abnormalities of the cerebellum, pons, brainstem and cerebrum although our patients are the first reported with cerebellar cysts.<sup>6</sup>

Even though standard sequencing of the coding regions of *LARGE* was normal in the proband of this family, we persisted with *LARGE* analysis because there was homozygous linkage to the gene locus in the context of consanguinity. Analysis of cDNA, generated from mRNA derived from patient lymphoblasts, showed an abnormally large product for a PCR reaction that spanned exons 9 to 12. Sequencing of this product showed that an abnormal sequence of either 197 or 171 bases was spliced between exons 10 and 11 in almost all transcripts. The abnormal sequence is normally situated on chromosome 22 centromeric to the 3' end of *LARGE*. In this position, it would not usually be incorporated into the *LARGE* transcript, even though the sequence contains the necessary splice donor and acceptor sites. To explain its incorporation into the *LARGE* mRNA, we hypothesised that the EST sequence was translocated into intron 10 of *LARGE* in our family, possibly through duplication of this genomic region. We assessed the copy number of the EST sequence by qPCR relative to other genes in patient genomic DNA and showed twice the expected levels, consistent with a genomic duplication in a homozygous state. Both parents had a 1.5-fold increase in copy

number (an increase from two to three copies of the sequence), consistent with both of them being asymptomatic heterozygous carriers of the duplication. To map the approximate size of the duplication, we repeated qPCR analysis up- and downstream of the EST sequence and showed the duplication spans between 40 and 43 kb of genomic DNA. Even though we identified the likely site of insertion in the central part of intron 10 and delimited a large intronic region that cannot be amplified in the patients, mapping the precise breakpoints of the proposed duplication/insertion was not feasible, possibly due to the presence of a more complex rearrangement than anticipated.

Analysis of *LARGE* mRNA transcripts from patient lymphoblasts and fibroblasts suggests that most *LARGE* mRNA transcripts contain an abnormal sequence. Both versions of the abnormal insertion are predicted to immediately introduce a premature stop codon mid-way through protein translation, just before the second catalytic domain, and result in a non-functional protein (figure 4B). Small amounts of normally spliced transcript were also seen and the levels appeared slightly higher in fibroblasts than lymphoblasts, suggesting this may vary in different tissues. Unfortunately it was not possible to assess levels of normally spliced *LARGE* mRNA or protein function either in muscle or brain, the two main tissues that show pathology in the  $\alpha$ -dystroglycanopathies. We observed no functionally glycosylated  $\alpha$ DG by Western blot analysis in patient fibroblasts but it is likely that low levels of functional  $\alpha$ DG are present in patient tissues at levels below the threshold of detection by Western blot for several reasons. Residual glycosylated  $\alpha$ DG staining was present in the muscle biopsy and both sisters have less severe cognitive deficits and brain malformations than a patient with homozygous null mutations in *LARGE*, who had a WWS phenotype.<sup>7</sup> We also detected low levels of a normally spliced *LARGE* gene mRNA transcript in both lymphoblasts and fibroblasts, which should allow normal *LARGE* protein to be produced at low levels.

We also investigated whether *LARGE* may exert some of its effects on  $\alpha$ DG glycosylation through interactions with either the POMT or POMGnT1 enzymes. Both POMT and POMGnT1 activities were normal in patient lymphoblasts, strongly suggesting that the two initial steps of  $\alpha$ DG glycosylation occur independently of *LARGE*.

At present it appears that *LARGE* is a rare cause of CMD since the coding regions of the gene have been screened in large cohorts of patients with  $\alpha$ -dystroglycanopathy of unknown cause.<sup>6,8</sup> Although experience is very limited at present, it is notable that two of the four confirmed families with CMD due to *LARGE* mutation reported have had large gene rearrangements, either a deletion involving several exons<sup>7</sup> or a likely intragenic insertion/deletion (our family). In another family a heterozygous nonsense mutation (p.Trp516X) was identified on standard gene sequencing, leaving open the possibility that a large deletion/duplication or rearrangement on the other allele was overlooked.<sup>9</sup> In addition, consanguineous families are described that link to the *LARGE* locus but no mutation has been found by exon sequencing.<sup>7,17</sup> *LARGE* was so named for the enormous genomic region it occupies on chromosome 22q12.3 (647 kb), due to the presence of very large introns enriched in low copy repeats (LCRs). The structure of the *LARGE* gene may predispose to mutations that involve duplication or deletion of large regions of DNA due to non-allelic homologous recombination between LCRs, the proposed mechanism for many genomic disorders.<sup>18</sup> Sequencing of coding regions from genomic DNA will not identify many large gene rearrangements. Therefore, sequencing *LARGE* from cDNA and the application of CGH, qPCR, and other methods for detecting rearrangement mutations should be considered routinely.

### **Acknowledgements**

We thank Eric Villard for helpful discussion. NC was supported by a joint grant from the Australian National Health and Medical Research Council (NHMRC; ID 402861) and the Institut National de la Santé et de la Recherche Médicale, AV and RP by Marie Curie grants from the European commission (OIF N°039762 and IIF N°221685). This work was also supported by research grants from the Ministry of Health, Labour and Welfare of Japan (HM, TE; for Nervous and Mental Disorders (20A-13) and Research on Psychiatric and Neurological Diseases and Mental Health), from The Takeda Science Foundation (HM, TE) and from the Paul D. Wellstone Muscular Dystrophy Cooperative Research Center grant (1U54NS053672; KPC). KPC is an Investigator of the Howard Hughes Medical Institute.



**Conflict of interest:** The authors report no conflicts of interest

## Reference List

1. Martin PT: Mechanisms of disease: congenital muscular dystrophies-glycosylation takes center stage. *Nat Clin Pract Neurol* 2006; **2**: 222-230.
2. Lane PW, Beamer TC, Myers DD: Myodystrophy, a new myopathy on chromosome 8 of the mouse. *J Hered* 1976; **67**: 135-138.
3. Mathews KD, Rapisarda D, Bailey HL, Murray JC, Schelper RL, Smith R: Phenotypic and pathologic evaluation of the myd mouse. A candidate model for facioscapulohumeral dystrophy. *J Neuropathol Exp Neurol* 1995; **54**: 601-606.
4. Grewal PK, Holzfeind PJ, Bittner RE, Hewitt JE: Mutant glycosyltransferase and altered glycosylation of alpha-dystroglycan in the myodystrophy mouse. *Nat Genet* 2001; **28**: 151-154.
5. Lee Y, Kameya S, Cox GA *et al*: Ocular abnormalities in Large(myd) and Large(vls) mice, spontaneous models for muscle, eye, and brain diseases. *Mol Cell Neurosci* 2005; **30**: 160-172.
6. Longman C, Brockington M, Torelli S *et al*: Mutations in the human LARGE gene cause MDC1D, a novel form of congenital muscular dystrophy with severe mental retardation and abnormal glycosylation of alpha-dystroglycan. *Hum Mol Genet* 2003; **12**: 2853-2861.
7. van Reeuwijk J, Grewal PK, Salih MA *et al*: Intragenic deletion in the LARGE gene causes Walker-Warburg syndrome. *Hum Genet* 2007; **121**: 685-690.
8. Mercuri E, Messina S, Bruno C *et al*: Congenital muscular dystrophies with defective glycosylation of dystroglycan: a population study. *Neurology* 2009; **72**: 1802-1809.

9. Godfrey C, Clement E, Mein R *et al*: Refining genotype phenotype correlations in muscular dystrophies with defective glycosylation of dystroglycan. *Brain* 2007; **130**: 2725-2735.
10. Yoshida-Moriguchi T, Yu L, Stalnaker SH *et al*: O-mannosyl phosphorylation of alpha-dystroglycan is required for laminin binding. *Science* 2010; **327**: 88-92.
11. Xiong H, Kobayashi K, Tachikawa M *et al*: Molecular interaction between fukutin and POMGnT1 in the glycosylation pathway of alpha-dystroglycan. *Biochem Biophys Res Commun* 2006; **350**: 935-941.
12. Michele DE, Barresi R, Kanagawa M *et al*: Post-translational disruption of dystroglycan-ligand interactions in congenital muscular dystrophies. *Nature* 2002; **418**: 417-422.
13. Ervasti JM, Campbell KP: Membrane organization of the dystrophin-glycoprotein complex. *Cell* 1991; **66**: 1121-1131.
14. Kimura K, Wakamatsu A, Suzuki Y *et al*: Diversification of transcriptional modulation: large-scale identification and characterization of putative alternative promoters of human genes. *Genome Res* 2006; **16**: 55-65.
15. Pentao L, Wise CA, Chinault AC, Patel PI, Lupski JR: Charcot-Marie-Tooth type 1A duplication appears to arise from recombination at repeat sequences flanking the 1.5 Mb monomer unit. *Nat Genet* 1992; **2**: 292-300.
16. Manya H, Bouchet C, Yanagisawa A *et al*: Protein O-mannosyltransferase activities in lymphoblasts from patients with alpha-dystroglycanopathies. *Neuromuscul Disord* 2008; **18**: 45-51.

17. Manzini MC, Gleason D, Chang BS *et al*: Ethnically diverse causes of Walker-Warburg syndrome (WWS): FCMD mutations are a more common cause of WWS outside of the Middle East. *Hum Mutat* 2008; **29**: E231-E241.
18. Inoue K, Lupski JR: Molecular mechanisms for genomic disorders. *Annu Rev Genomics Hum Genet* 2002; **3**: 199-242.

Title for figure 1

Brain MRI images from Patient 1 performed at age 3.5 years

Legend for figure 1

A: Coronal T1 weighted image showing cerebellar and vermal hypoplasia (white arrows), cerebellar cysts (black arrow), generalised white matter atrophy and increased CSF spaces. B: Transverse T2-weighted image showing abnormal high white matter signal and dilated lateral ventricles. C: Sagittal T1-weighted image showing cerebellar (white arrow) and pontine hypoplasia.

Title for figure 2

Western blot analysis of  $\alpha$ -dystroglycan

Legend for figure 2

Immunoblot analysis of WGA-enriched homogenates from fibroblasts from Patient 1 and a healthy individual (control). Absent staining of antibodies IIIH6 and VIA4 that bind to glycosyl epitopes on  $\alpha$ -dystroglycan ( $\alpha$ -DG) and a laminin-1 (Lam) overlay assay show abnormalities in  $\alpha$ DG glycosylation in patient fibroblasts that is associated with loss of laminin-1 protein binding. Beta-dystroglycan ( $\beta$ -DG) is used as a loading control.

Title for figure 3

Genotyping for the *LARGE* locus

Legend for figure 3

Family tree showing analysis of microsatellite markers in close proximity to the *LARGE* locus (22q12.3). Both affected children are homozygous for four contiguous markers, consistent with homozygosity by descent for this chromosomal region.

Title for figure 4

Analysis of the *LARGE* gene insertion in genomic and cDNA

Legend for figure 4

A; Agarose gel showing the PCR products from patient (P) and control (C) cDNA generated from cultured lymphoblasts (Lymph) and fibroblasts (Fib) using primers that span exons 9 to 12 of *LARGE* (see E-Table 1). In patient cDNA, two abnormally high bands predominate that correspond to insertions of abnormal DNA sequences between exons 10 and 11. A faint band of the normal size is also seen (arrow). B; Diagram showing the amino acid sequence that is coded by normal and patient cDNA around the exon10/11 boundary. The abnormal insertion found in patient cDNA results in the introduction of a premature STOP codon, predicted to cause protein termination mid-way through the translation of the *LARGE* protein. C; Results from qPCR analysis of genomic DNA (gDNA) for EST DA935254 and the surrounding chromosomal region. gDNA concentrations between samples were adjusted using qPCR results from a distant gene (PTPLA) and results were scaled so that control samples had a mean copy number of 2 at each point (data not shown). In patient gDNA, qPCR analysis of the EST sequence and for five regions at variable distances upstream and downstream shows twice the expected copy number, consistent with homozygous duplication of this region. Genomic copy number returns to normal for markers 5b and 3b in patient gDNA, and therefore these markers define the maximum size of the duplication. Analysis of gDNA from both parents shows that both had three copies of the duplicated region, consistent with each

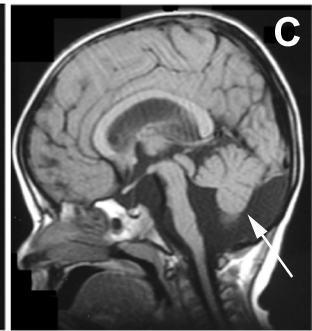
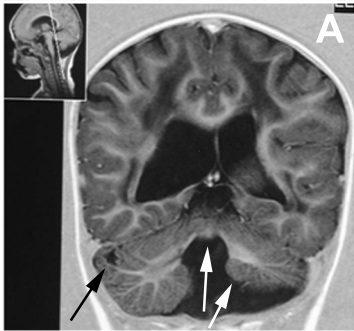
having the duplication in a heterozygous state. All results are shown relative to the reverse gDNA strand.

Title for figure 5

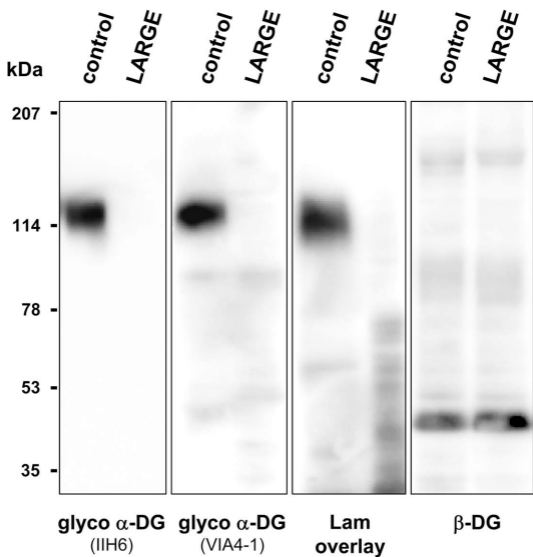
Diagram of the insertion/deletion mutation involving intron 10 of the *LARGE* gene

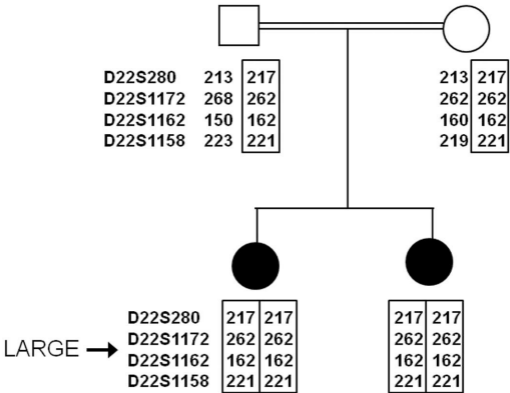
Legend for figure 5

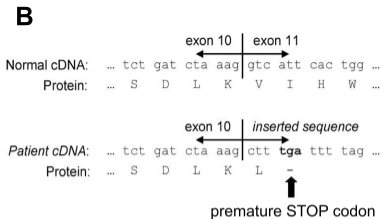
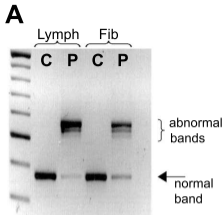
The EST DA935254, normally located 100 kb centromeric of *LARGE*, is duplicated together with ~40 kb of flanking DNA and likely inserted into intron 10 of *LARGE*. Numerous repetitive elements which may have contributed to the large intra-chromosomal rearrangement are present in both intron 10 and flanking the EST, including SINEs (green), LINEs (red) and LTRs (blue). bp = base pairs.



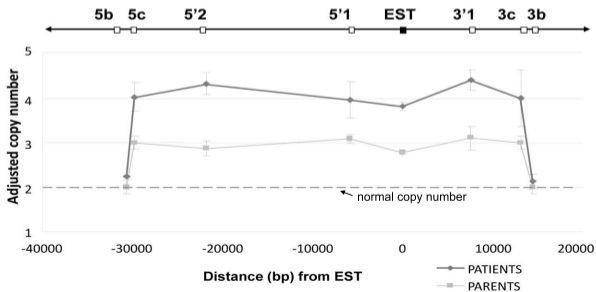








**C** Genomic DNA Copy Number at 22q12.3



chr22 (q12.3) 22p13.2 22p12 p11.2 11.21 12.1 12.3 13.1 13.2 13.31

