Glycobiology vol. 15 no. 10 pp. 912–923, 2005
doi:10.1093/glycob/cwi094
Advance Access publication on June 15, 2005

Characterization of the LARGE family of putative glycosyltransferases associated with dystroglycanopathies

Prabhjit K. Grewal[1,3], Jennifer M. McLaughlan[1], Christopher J. Moore[1], Claudia A. Browning, and Jane E. Hewitt[2]

Institute of Genetics, Queen's Medical Centre, University of Nottingham, Nottingham NG7 2UH, UK

Received on December 15, 2004; revised on June 6, 2005; accepted on June 10, 2005

The Large^{myd} mouse has a loss-of-function mutation in the putative glycosyltransferase gene *Large*. Mutations in the human homolog (*LARGE*) have been described in a form of congenital muscular dystrophy (MDC1D). Other genes (*POMT1*, *POMGnT1*, *fukutin*, and *FKRP*) that encode known or putative glycosylation enzymes are also causally associated with human congenital muscular dystrophies. All these diseases are associated with hypoglycosylation of the membrane protein α-dystroglycan (α-DG) and consequent loss of extracellular ligand binding. Hence, they are termed dystroglycanopathies. A paralogous gene for *LARGE* (*LARGE2* or *GYLTL1B*) may also have a role in DG glycosylation. Using database interrogation and reverse-transcriptase polymerase chain reaction (RT–PCR), we identified vertebrate orthologs of each of these *LARGE* genes in many vertebrates, including human, mouse, dog, chicken, zebrafish, and pufferfish. However, within invertebrate genomes, we were able to identify only single homologs. We suggest that vertebrate *LARGE* orthologs be referred to as *LARGE1*. RT–PCR, dot-blot, and northern analysis indicated that *LARGE2* has a more restricted tissue-expression profile than *LARGE1*. Using epitope-tagged proteins, we show that both LARGE1 and LARGE2 localize to the Golgi apparatus. The high similarity between the LARGE paralogs suggests that LARGE2 may also act on DG. Overexpression of *LARGE2* in mouse C2C12 myoblasts results in increased glycosylation of α-DG accompanied by an increase in laminin binding. Thus, there may be functional redundancy between LARGE1 and LARGE2. Consistent with this idea, we show that α-DG is still fully glycosylated in kidney (a tissue that expresses a high level of *LARGE2* mRNA) of Large^{myd} mutant mice.

Key words: dystroglycan/glycosyltransferase/Golgi/muscular dystrophy

Introduction

Mutations in at least six genes encoding known or putative glycosylation enzymes are associated with muscular dystrophies or myopathies (Hewitt and Grewal, 2003; Martin and Freeze, 2003). The majority of these disorders are severe congenital forms, often associated with brain and eye abnormalities. Muscle-eye-brain disease (MEB), Fukuyama-type congenital muscular dystrophy, and Walker–Warburg syndrome (WWS) are associated with mutations in the genes *POMGnT1*, *fukutin*, or *POMT1* (Kobayashi *et al*., 1998; Yoshida *et al*., 2001; Beltrán-Valero de Bernabé *et al*., 2002). Mutations in the gene encoding fukutin-related protein (*FKRP*) result in a wide spectrum of clinical phenotypes ranging from the severe congenital muscular dystrophies Type 1C (MDC1C), MEB, and WWS to the milder limb-girdle muscular dystrophy 2I (LGMD2I) (Mercuri *et al*., 2003; Beltrán-Valero de Bernabé *et al*., 2004). The *Large* gene, which encodes a putative bifunctional glycosyltransferase, is mutated in the myodystrophy (Large^{myd}) mouse (Grewal *et al*., 2001) and in congenital muscular dystrophy Type 1D (MDC1D) (Longman *et al*., 2003).

All these human diseases, and the Large^{myd} mouse, are associated with hypoglycosylation of the muscle membrane protein α-dystroglycan (α-DG) and concomitant loss of extracellular ligand binding (Grewal *et al*., 2001; Michele *et al*., 2002). Many of the clinical findings can be recapitulated in the muscle and brain of mice by conditional depletion of DG (Côté *et al*., 1999; Moore *et al*., 2002), indicating that this protein is a key downstream target. Hence, these muscular dystrophies are now commonly termed dystroglycanopathies (Michele and Campbell, 2003). Altered glycosylation of α-DG might also play a role in cancer, as hypoglycosylation of the protein and loss of laminin binding have been demonstrated in invasive carcinoma cells (Singh *et al*., 2004).

α- and β-DG are essential components of the dystrophin–glycoprotein complex (Ervasti and Campbell, 1991; Ibraghimov-Beskrovnaya *et al*., 1992), which in muscle links dystrophin and the actin cytoskeleton inside the cell to the extracellular matrix. Both subunits are encoded by a single pro-peptide that is proteolytically cleaved; the subunits remain associated through non-covalent interactions (Michele and Campbell, 2003). α-DG is highly glycosylated in a tissue-specific manner with both *N*- and *O*-linked glycans, and the carbohydrate can amount to over half of its molecular mass. These O-linked structures are necessary for binding of DG to extracellular ligands such as laminin and agrin (Martin, 2003). β-DG is a membrane protein with less extensive glycosylation that is limited to *N*-linked glycans.

[1]These authors contributed equally to this work.
[2]To whom correspondence should be addressed; e-mail: jane.hewitt@nottingham.ac.uk
[3]Present address: Howard Hughes Medical Institute, 9500 Gilman Drive-0625, University of California San Diego, La Jolla, CA 92093

DG is widely expressed, with roles in muscle stabilization, cell migration, basement membrane assembly, and signaling (Winder, 2001). Homologs have been identified in a range of organisms including mammals (Ibraghimov-Beskrovnaya *et al*., 1992), chicken (Brancaccio *et al*., 1995), zebrafish (Parsons *et al*., 2002), *Xenopus* (Cohen *et al*., 1995), *Drosophila* (Greener and Roberts, 2000), and nematodes (Grisoni *et al*., 2002).

LARGE was originally named because it covers approximately 650 kb of human genomic DNA (Peyrard *et al*., 1999), despite a coding region only 2268 bp. The predicted protein structure contains an N-terminal cytoplasmic domain, a transmembrane region, a coiled-coil motif, and two putative catalytic domains (Grewal *et al*., 2001). The proximal catalytic domain has closest homology to Waaj, a bacterial family GT8 glycosyltransferase (Coutinho *et al*., 2003) involved in lipooligosaccharide synthesis (Heinrichs *et al*., 1998). The distal domain shows the most similarity to human UDP-GlcNAc:Galβ1,3-*N*-acetylglucosaminyltransferase (iGnT), a family GT49 glycosyltransferase (Coutinho *et al*., 2003) responsible for synthesis of the poly-*N*-acetyllactosamine-based erythrocyte i antigen (Sasaki *et al*., 1997). It has recently been shown that overexpression of *LARGE* in cell lines derived from patients with WWS and MEB results in hyperglycosylation of α-DG and restoration of laminin-binding activity to the protein (Barresi *et al*., 2004). As the loss of ligand binding activity by α-DG appears to contribute significantly to the pathology of these diseases, up-regulation of LARGE activity might represent a therapeutic approach for these muscular dystrophies (Barresi *et al*., 2004).

We previously reported a mouse gene closely related to *Large* (*Large-like*) identified using a BLAST search of mouse genome sequences (Grewal and Hewitt, 2002). This gene is also referred to as *Gyltl1b* (GenBank Accession number NM_152312) (Brockington *et al*., 2005) or *LARGE2* (Fujimura *et al*., 2005). In this article, we report the gene structures, evolutionary conservation, and the expression profiles of *LARGE* and *LARGE2*. In addition, epitope-tagged LARGE and LARGE2 proteins are both shown to localize to the Golgi apparatus, consistent with a function in glycosylation. Transient overexpression of *Large2* in the C2C12 mouse myoblast cell line results in an increase in α-DG glycosylation and laminin binding.

Results

LARGE is part of a highly conserved gene family

We previously reported a mouse gene closely related to *Large* identified by BLAST searches of mouse genome sequences (Grewal and Hewitt, 2002). We initially named this gene *Large-like*; it is also now known as *Gyltl1b* (GenBank Accession number NM_152312) or LARGE2 (Fujimura *et al*., 2005). We confirmed the complete coding sequences of human and mouse *LARGE2* by sequencing reverse-transcriptase polymerase chain reaction (RT–PCR) products amplified from kidney and testis RNA (GenBank Accession numbers AY742914 and AY742915 for mouse and NM_152312 for human). We then used database interrogation to identify homologs of both genes from other organisms. RT–PCR was used to confirm the full-length coding sequences of *LARGE* and *LARGE2* from zebrafish (AY662338 and AY662339) and chicken (AY662336 and AY662337). Within invertebrate genome sequences, we were able to identify single homologs in the nematode worms *Caenorhabditis elegans* (Peyrard *et al*., 1999) and *C. briggsae* (data not shown), and in the honeybee *Apis mellifera*. However, no homolog was found in the *Drosophila* genome. Until their biochemical functions are defined, we propose the generic name *LARGE* for this family with the paralogous genes referred to as *LARGE1* and *LARGE2*. The invertebrate homolog is therefore defined as *LARGE*. A ClustalX alignment of the predicted amino acid sequences of all the vertebrate homologs and the honeybee LARGE protein is shown in Figure 1. We were unable to establish the 5′ coding region of the bee gene with certainty, and therefore, this has been omitted from the alignment.

The predicted LARGE1 and LARGE2 proteins have a similar domain structure; they are type II membrane-anchored proteins with short (eight to ten amino acids) N-terminal cytoplasmic tails, short (16–23 amino acids) transmembrane domains, variable stem regions, and two catalytic domains. The ClustalX alignment shows the extensive amino acid similarity between paralogous vertebrate LARGE proteins over the catalytic domains (for example, mouse Large1 and Large2 share 68% amino acid identity). The conservation between species is also high; honeybee LARGE has ~53% amino acid identity with the other homologs over the catalytic domains. The MDC1D case reported by Longman *et al*. (2003) is a compound heterozygote for a truncating and a missense mutation (Glu509Lys) within the second catalytic domain. We note that Glu509 (indicated by a diamond in Figure 1) is conserved in all the LARGE homologs we identified, including the bee protein.

In Figure 1, the level of amino acid conservation between the LARGE proteins appears much lower over the N-terminal and transmembrane regions. Close examination of the alignment shows that the mammalian LARGE2 proteins, which lack predicted coiled-coil domains, show little similarity to LARGE1 in the N-terminal cytoplasmic and transmembrane regions. In contrast, the other vertebrate LARGE2 proteins have predicted coiled-coil domains and have N-terminal regions similar to the LARGE1 proteins.

LARGE1 and LARGE2 arose via a gene-duplication event

Figure 2 shows a rooted maximum likelihood tree based on the ClustalX alignment of the amino acid sequences of the predicted catalytic domains of the LARGE family proteins. The vertebrate LARGE1 and LARGE2 proteins each form a separate cluster within this tree, indicating they originated from a gene-duplication event in the vertebrate lineage.

We then determined the gene sizes, locations, and exon organization for these genes, on the basis of available genomic sequence data (Table I, Figure 3). Owing to poor reliability of gene prediction programs for 5′ UTRs and low coverage of EST data for most species, only the coding exons are shown in Figure 3. However, we note that as both the human and mouse *LARGE1* genes have introns of at least 100 kb within their 5′ UTRs, this figure probably

significantly underestimates gene sizes. There is currently insufficient genomic sequence in the zebrafish assembly to determine the complete gene structure, and therefore, we have included the predicted puffer fish (*Takifugu rubripes*) genes instead. In general, there is a striking difference in gene size between the paralogs, for example the human *LARGE1* gene is over 100-fold bigger than *LARGE2*, and this difference is because of significantly greater intron sizes in all the *LARGE1* homologs.

Where possible, we also determined the chromosomal location of each gene (Table I). Comparison of the gene locations with regions of conserved gene synteny shows

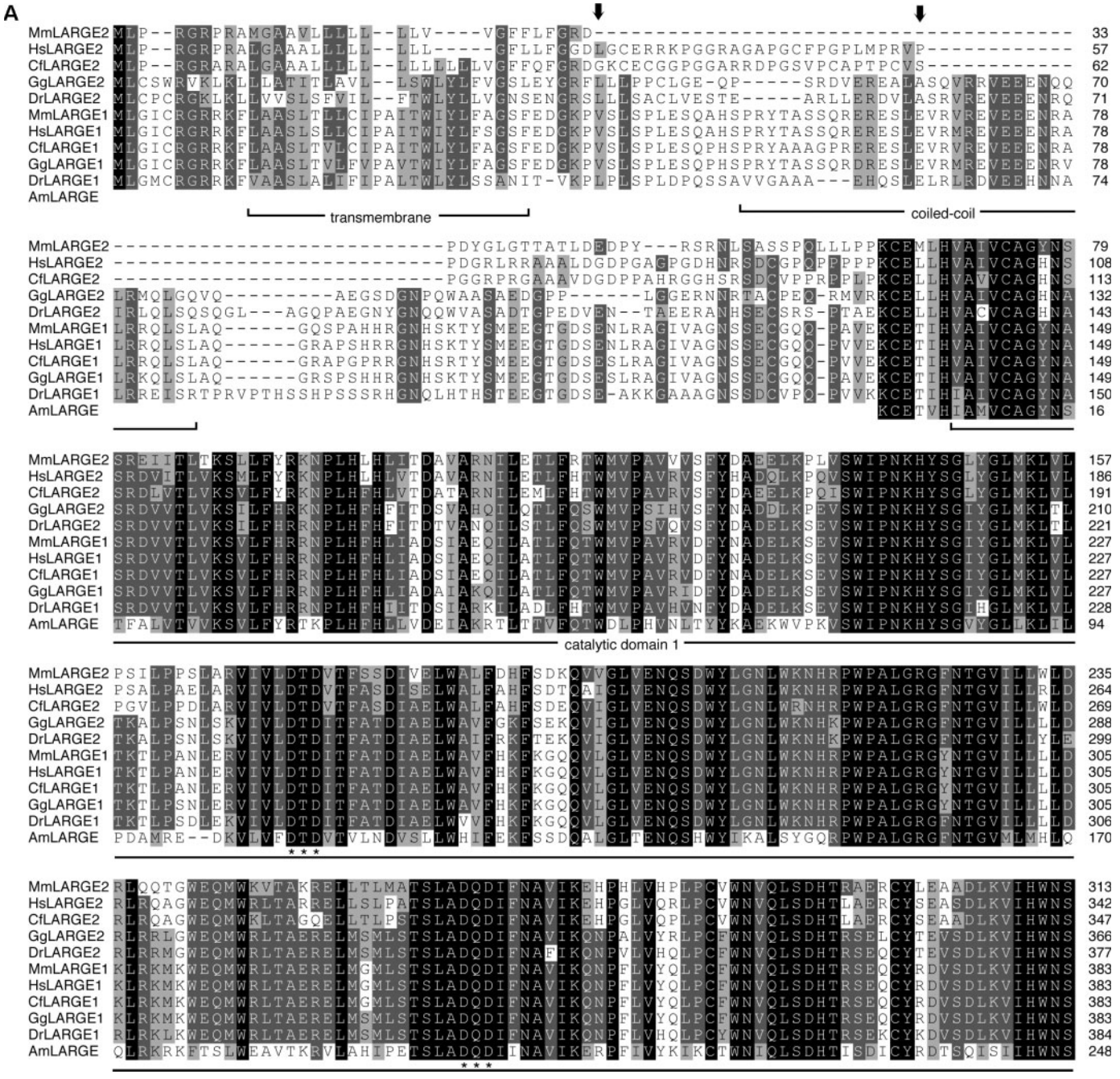


Fig. 1. ClustalX alignment of LARGE homologs. Predicted amino acid sequences of the orthologs of LARGE (designated LARGE1) and of LARGE2 were aligned using ClustalX. Identical residues in all proteins in the alignment are shown in white text on black. Conserved (white on dark gray) and similar residues (black on light gray) are calculated at a 60% threshold. The locations of predicted protein domains are shown under the alignment. The arrowheads indicate the region missing in the mouse protein due to recent intron formation in rodents. The diamond indicates the location of the Glu509Lys mutation in the MDC1D patient (Longman *et al.*, 2003). Asterisks denote DxD motifs, characteristic of many glycosyltransferase catalytic domains (Wiggins and Munro, 1998). Am, *Apis mellifera* (honey bee); Cf, *Canis familiaris* (dog); Dr, *Danio rerio* (zebrafish); Gg, *Gallus gallus* (chicken); Hs, *Homo sapiens* (human); Mm, *Mus musculus* (mouse).

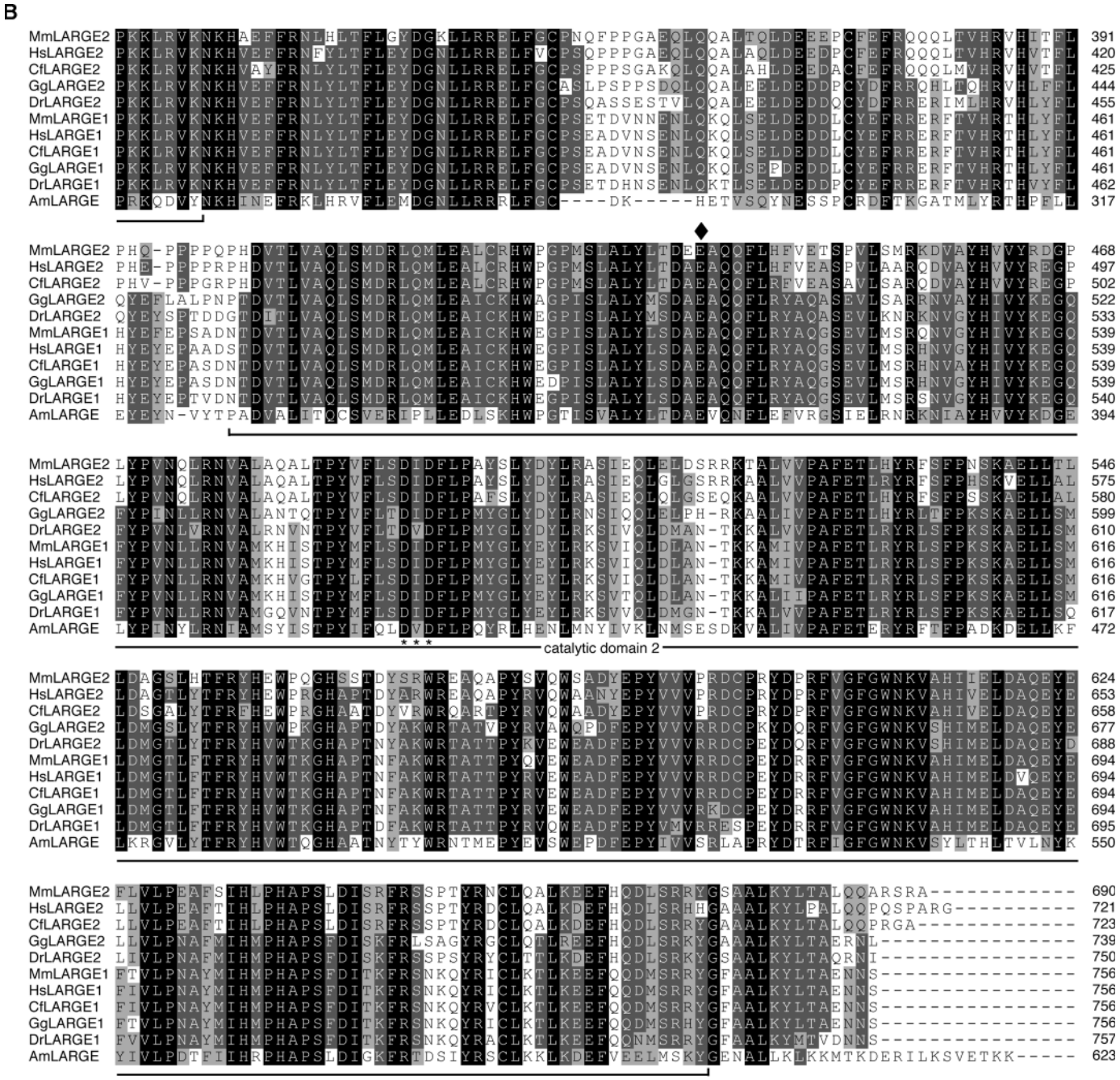


Fig. 1. continued

that all the mammalian and chicken *LARGE1* genes map to homologous genomic locations, as do the *LARGE2* genes (Table I). This confirms that vertebrate *LARGE1* and *LARGE2* genes are true paralogs that arose via a gene-duplication event.

As expected for a gene duplication, the regions encoding the two predicted catalytic domains show a very similar organization in all the vertebrate genes; they all span 12 exons and have intron boundaries occurring at equivalent locations (Figure 3). However, from examination of both the alignments and the genomic organization, it is apparent that the mammalian *LARGE1* orthologs lack the exon encoding the coiled-coil domain. All the other vertebrate genes investigated in this study encode this domain, as do the frog *Xenopus laevis* and *X. tropicalis LARGE2* genes (data not shown). Furthermore, the chromosomal position of chicken *LARGE2* is homologous to the human and mouse *LARGE2* gene locations, indicating they are true orthologs. Together with the lack of conservation of both amino acid sequence and intron locations at the 5′ end of the genes, this suggests that the mammalian *LARGE2* genes have acquired an alternative initial coding exon and that

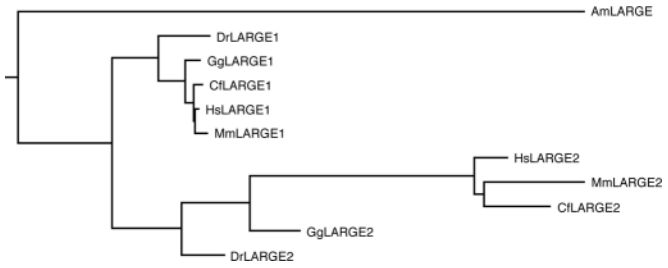


Fig. 2. Rooted maximum likelihood tree of LARGE proteins. The tree was generated from the ClustalX amino acid alignment using ProML and displayed using Dendromaker. Gapped residues and all those N-terminal of the highly conserved KCE motif (residues 64–66 of human LARGE2) were excluded. Am, *Apis melifera* (honeybee); Cf, *Canis familiaris* (dog); Dr, Danio rerio (zebrafish); Gg, *Gallus gallus* (chicken); Hs, *Homo sapiens* (human); Mm, *Mus musculus* (mouse).

the ancestral gene resembled *LARGE1*. Furthermore, this mammalian exon shows evidence for recent formation of an intron. In mouse (but not human or dog), this first coding exon is interrupted by a 75-bp intron (Figure 3), removing a 25-amino acid part of the putative stem region (indicated by arrowheads in Figure 1). Although an open-reading frame has been maintained across this intron, it is spliced in all the RT–PCR products we obtained. The intron is also present and correctly spliced in the rat *Large2* gene (data not shown), indicating the formation of this intron occurred within the rodent lineage.

Large2 *is alternatively spliced in mouse*

From mouse EST information, we identified a potential alternative splice form of *Large2* that lacks one of the exons encoding the first catalytic domain (indicated by an asterisk in Figure 3). This alternative splicing was confirmed by RT–PCR in mouse testis and kidney RNA with products of 473 and 368 bp (Figure 4) and sequencing of the PCR products confirmed the exon was excluded. However, only one human EST was identified that lacked this exon and we were unable to amplify this alternative transcript from human RNA (Figure 4A). Therefore, the functional significance of this alternative splicing in the mouse is currently unclear. Exclusion of this exon does not alter the reading frame, but removes a highly conserved 35-amino acid region (residues 193–227 of mouse Large2). We found no evidence, either by RT–PCR or from examination of EST databases, for alternative splicing of this region of *LARGE1* in either human or mouse.

LARGE1 and LARGE2 have different tissue-expression profiles

We compared the tissue profiles of *Large1* and *Large2* mRNAs in mouse by RT–PCR (Figure 4B). In agreement with previous work (Peyrard *et al.*, 1999), we found that *Large1* is expressed in a range of adult tissues. *Large1* is also expressed during mouse embryogenesis; the earliest time-point studied was 7 dpc and we were also able to amplify the transcript in RNA from 12.5 dpc embryos (data not shown). *Large2* shows a more restricted pattern of expression in adult tissues than *Large1*, with high levels in the testis and kidney, but very low expression in the heart and brain. We were unable to detect significant expression in skeletal muscle (quadriceps femoris RNA was used in Figure 4). Like *Large1*, *Large2* is expressed during mouse embryogenesis, with RT–PCR products detected from 7 dpc.

Dot-blot analysis indicated that human *LARGE1* is also widely expressed, with highest levels in brain, heart, skeletal muscle, stomach, bladder, and pancreas (Figure 4C). Highest levels of *LARGE2* expression were observed in placenta, adult and fetal kidney, pancreas, and prostate and salivary glands. Only low levels of expression were detected in brain or skeletal muscle RNA (Figure 4C). This low expression of *LARGE2* in human skeletal muscle was confirmed by northern blotting of total RNA (Figure 4D). For *LARGE1*, a 4.5-kb transcript was seen in adult and fetal skeletal muscle, brain, and kidney. For *LARGE2*, a 2.5-kb transcript was identified in both fetal and adult kidney, with a faint

Table I. Sizes and chromosomal locations of *LARGE1* and *LARGE2* genes

Gene	Organism	Gene size (kb)	Genomic location	GenBank/Ensembl gene predictions
LARGE1	Human	668	22q12.3	NM_004737/ NM_133642
	Mouse	562	8C2	NM_010687
	Dog	387[a]	10 (contig_18722.1.205387)	GENSCAN00000048546
	Chicken	198[a]	1 (contig_7.158/7.160)	AY662337
	Fugu	16[a]	Scaffold_1084	SINFRUT00000159991
LARGE2	Human	6	11p11.2	NM_152312
	Mouse	6	2E2	AY742914/AY742915
	Dog	5	18 (contig_53049.1.136022)	GENSCAN0000056344
	Chicken	5	5 (contig_45.137/45.138)	AY662336
	Fugu	≥5[b]	Scaffold_4111	SINFRUT00000133672

[a]The estimated sizes for these genes encompass only the coding region.
[b]The genomic assembly for fugu *LARGE2* lacks the first coding exon.

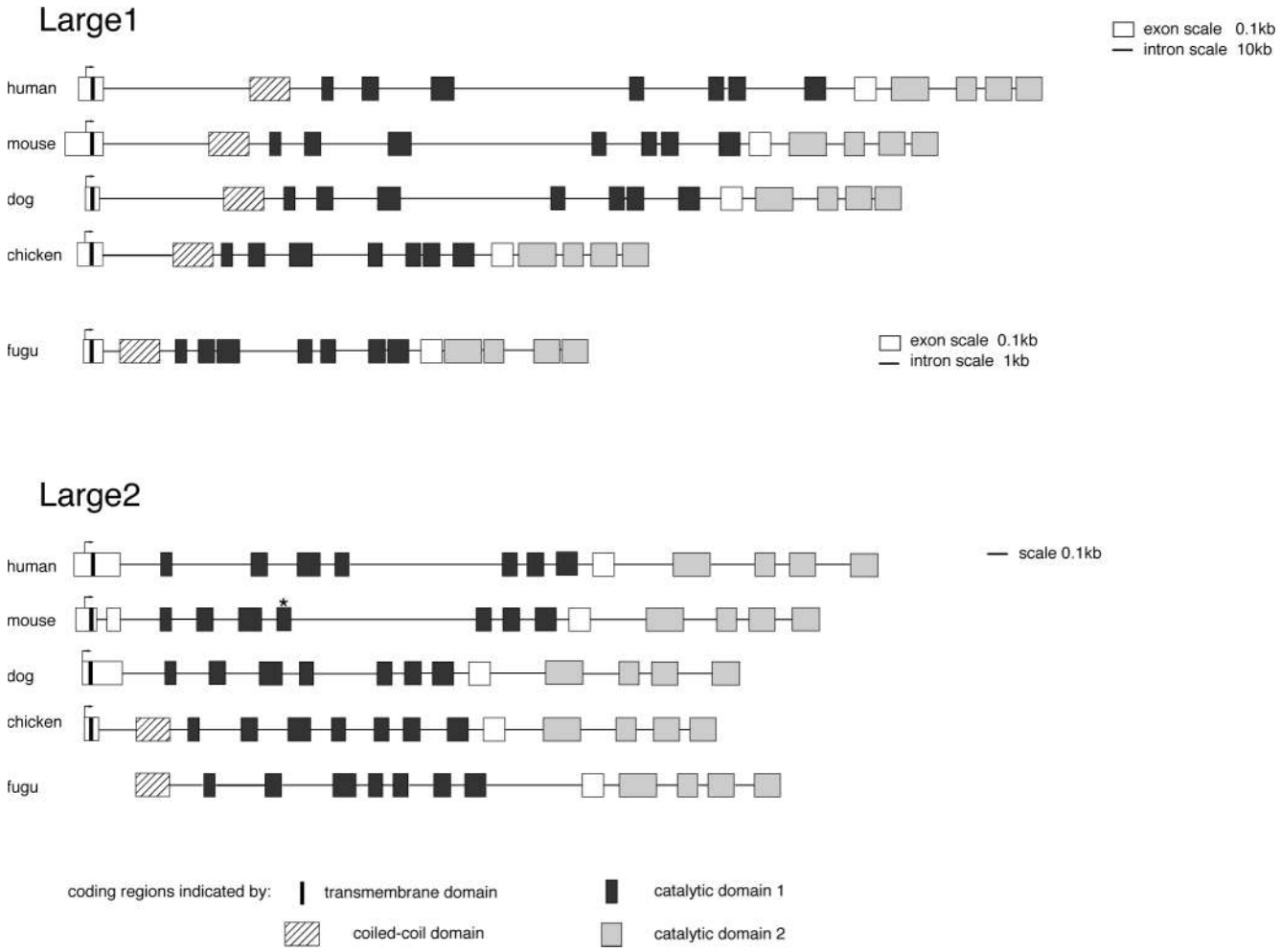


Fig. 3. Gene structures of vertebrate LARGE homologs. A schematic diagram showing the genomic structures of the *LARGE1* and *LARGE2* homologs in human, mouse, dog, chicken, and *Takifugu rubripes*. Exons are drawn to the same scale for all the genes. For human, mouse, dog, and chicken *LARGE1* genes, the introns are drawn at a 100× smaller scale; for Fugu, introns are drawn at a 10× smaller scale. For all the *LARGE2* orthologs, exons and introns are drawn to the same scale. For simplicity, only the coding exons are shown. An arrow indicates the location of the ATG initiation codon. A black bar indicates the coding region for the transmembrane domain. Exons that encode catalytic domain 1 are shaded dark gray and exons that encode catalytic domain 2 are shaded light gray. The exon containing the predicted coiled-coil domain is hatched; this exon is absent in the human, mouse, and dog *LARGE2* genes. The first coding exon of the mouse gene is interrupted by a 75 bp intron. The asterisk indicates the exon that is alternatively spliced in mouse *Large2*.

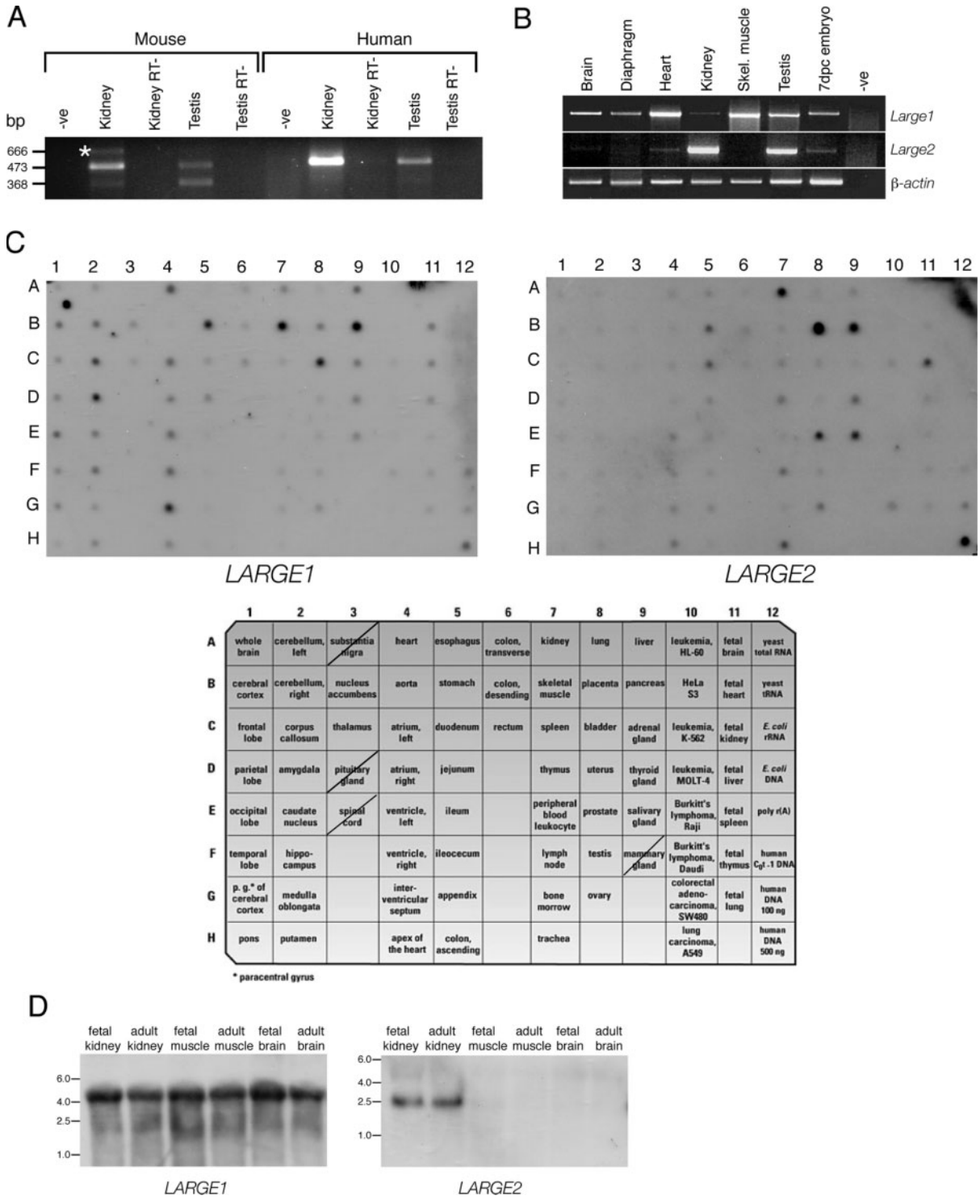
signal in fetal skeletal muscle. No transcripts were detected in adult skeletal muscle, or in the human brain samples by northern blot.

Large2 and LARGE1 proteins both localize to the Golgi apparatus

As no antibodies to the endogenous proteins are currently available, the complete coding regions of mouse *Large2* and human *LARGE1* (clones containing the full-length mouse *Large1* coding sequence were unstable in *Escherichia coli*) were fused in-frame to EGFP or DsRed2. The coding sequence of the mutated transcript of Large^{myd} mice, which would encode a truncated Large protein consisting of the cytoplasmic, transmembrane, and coiled-coil domains, was amplified by RT–PCR from skeletal muscle RNA and was also used to generate fusion proteins. Localization of the epitope-tagged proteins was investigated in normal rat kidney fibroblasts (NRK cells) and the mouse myoblast cell line C2C12 using immunofluorescence (Figure 5). The Golgi markers used were α-mannosidase II (Velasco *et al.*, 1993) in NRK cells and either golgin 97 (Kjer-Nielsen *et al.*, 1999) or GM130 (Nakamura *et al.*, 1995) in C2C12 cells.

In NRK cells, the Large2-EGFP fusion protein is targeted to the Golgi, as shown by co-localization with α-mannosidase II (Figure 5). We saw a similar localization pattern for LARGE1-EGFP (data not shown). In transfected C2C12 myoblasts, both Large2-EGFP and LARGE1-EGFP showed co-localization with golgin 97 (data not shown) and GM130 (Figure 5). Thus, although our expression data suggest that Large2 is not normally expressed in skeletal muscle, the protein is capable of localizing to the Golgi in myoblasts.

The truncated Large^{myd} protein product [Large1(myd), Figure 5] also localized to the Golgi when C-terminally tagged, indicating that the N-terminal 143 residues of the



	1	2	3	4	5	6	7	8	9	10	11	12
A	whole brain	cerebellum, left	substantia nigra	heart	esophagus	colon, transverse	kidney	lung	liver	leukemia, HL-60	fetal brain	yeast total RNA
B	cerebral cortex	cerebellum, right	nucleus accumbens	aorta	stomach	colon, desending	skeletal muscle	placenta	pancreas	HeLa S3	fetal heart	yeast tRNA
C	frontal lobe	corpus callosum	thalamus	atrium, left	duodenum	rectum	spleen	bladder	adrenal gland	leukemia, K-562	fetal kidney	*E. coli* rRNA
D	parietal lobe	amygdala	pituitary gland	atrium, right	jejunum		thymus	uterus	thyroid gland	leukemia, MOLT-4	fetal liver	*E. coli* DNA
E	occipital lobe	caudate nucleus	spinal cord	ventricle, left	ileum		peripheral blood leukocyte	prostate	salivary gland	Burkitt's lymphoma, Raji	fetal spleen	poly r(A)
F	temporal lobe	hippo-campus		ventricle, right	ileocecum		lymph node	testis	mammary gland	Burkitt's lymphoma, Daudi	fetal thymus	human C_0t -1 DNA
G	p. g.* of cerebral cortex	medulla oblongata		inter-ventricular septum	appendix		bone morrow	ovary		colorectal adeno-carcinoma, SW480	fetal lung	human DNA 100 ng
H	pons	putamen		apex of the heart	colon, ascending		trachea			lung carcinoma, A549		human DNA 500 ng

* paracentral gyrus

Fig. 4. *Large1* and *Large2* have different expression profiles. (**A**) Reverse-transcriptase polymerase chain reaction (RT–PCR) of human and mouse *Large2* mRNAs from testis and kidney RNA showing the presence of two splice forms in mouse. The 473-bp product has retained the exon, while the 368-bp product has the exon excluded. The asterisk indicates amplification of incompletely spliced mRNA in mouse kidney that has retained intronic sequence. (**B**) RT–PCR of *Large1* and *Large2* mRNA from mouse tissues (skeletal muscle RNA was isolated from quadriceps femoris). Approximately equal amounts of first-strand cDNA were used in each reaction. β-Actin was used as a control housekeeping gene. The RT–PCR product for mouse *Large2* does not span the alternatively spliced exon. (**C**) Commercial multiple tissue poly A+ dot-blot (Clontech) containing RNA (levels normalized to β-actin) from a wide range of human tissues hybridized with radiolabeled cDNA probes to *LARGE1* and *LARGE2*. The filter was washed under high stringency conditions (0.5× SSC, 65°C). The filter did not contain RNA from substantia nigra, pituitary gland, spinal cord, or mammary gland (indicated by a line though the appropriate grid). (**D**) A northern blot containing 10 μg of each total RNA was hybridized sequentially with radiolabeled cDNA probes to *LARGE2* and *LARGE1* as in panel (**C**).

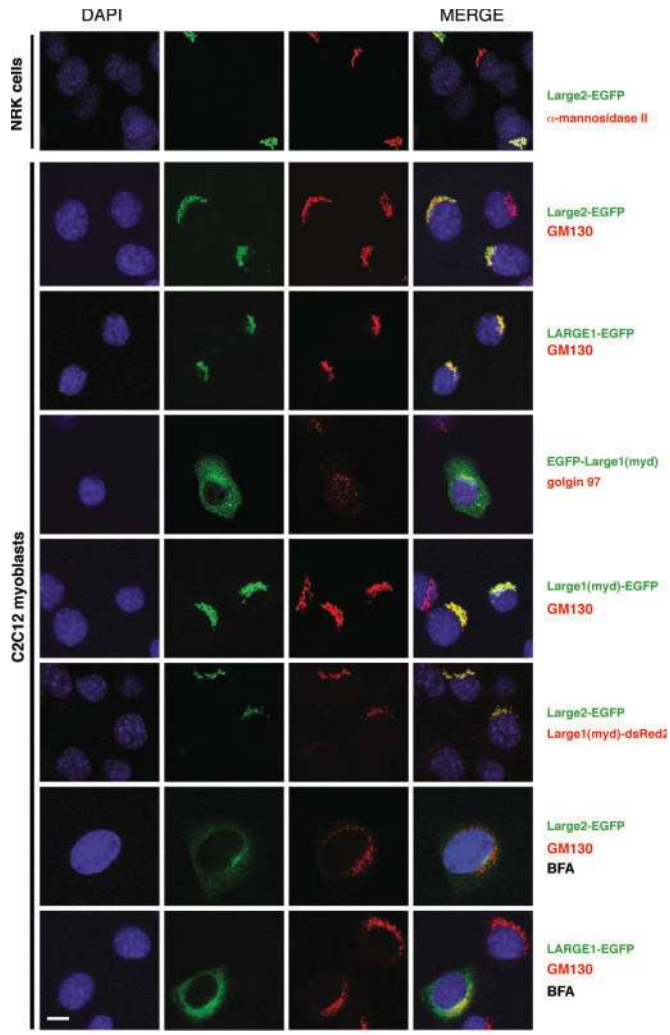


Fig. 5. LARGE proteins localize to the Golgi apparatus. Cells were transfected with constructs as labeled to the right of the figure. All the data shown are using C-terminal tagged constructs except for the N-terminal EGFP-tagged mutant Large1 protein EGFP-Large1(myd). Cells were stained with DAPI to show the location of nuclei. Immunostaining for Golgi marker is as indicated. Three channel confocal images were captured; individual and merged images are shown. Scale bar = 10 μm. BFA indicates cells were treated with brefeldin A.

protein are sufficient for correct targeting and retention. Co-transfection of Large1(myd)-DsRed2 and Large2-EGFP constructs showed complete co-localization of the two proteins, confirming the correct localization of the mutant protein. We found that N-terminal tagged protein did not localize to the Golgi (Figure 5), instead showing a diffuse staining consistent with retention in the endoplasmic reticulum (ER). Presumably, the tag interferes with correct folding and/or targeting of the protein in this configuration.

To confirm the Golgi localization of the tagged proteins, cells were treated with brefeldin A (BFA), which causes many Golgi enzymes to redistribute to the ER (Doms *et al*., 1989; Lippincott-Schwartz *et al*., 1989). After treatment, both LARGE1-EGFP and Large2-EGFP showed a diffuse staining, indicating the proteins had been redistributed to the ER. GM130, a cis-Golgi matrix protein, remains associated with Golgi remnants around the nucleus after BFA treatment and does not redistribute to the ER with Golgi enzymes (Nakamura *et al*., 1995), resulting in a punctate perinuclear staining. Thus, both LARGE1 and Large2 localize to the Golgi apparatus and show a redistribution pattern typical of Golgi enzymes following BFA treatment. Our data are consistent with both proteins having a role in protein glycosylation and/or processing.

Transient expression of Large2 in myoblasts increases α-DG glycosylation and laminin binding

The high level of amino acid similarity between LARGE1 and LARGE2 suggests the latter may also act in the DG glycosylation pathway. To investigate this, we used the C2C12 myoblast cell line and the monoclonal antibody IIH6, which recognizes glycosylated α-DG. C2C12 cells can be readily differentiated into myotubes, and this is accompanied by increased DG expression and generation of glycosylated forms of α-DG that can be detected by IIH6 antibody (Campanelli *et al*., 1994; Gee *et al*., 1994). C2C12 myoblasts and differentiated myotubes express *Large1* but not *Large2* (data not shown). Overexpression of *LARGE1* in myoblasts results in hyperglycosylation of α-DG (Barresi *et al*., 2004). Therefore, to test whether *Large2* was able to alter α-DG glycosylation, we transfected C2C12 myoblasts with the mouse *Large2* coding sequence cloned into the pcDNA3.1V5-His vector.

Figure 6(A) shows western blots of protein extracts from mouse skeletal muscle (positive control) and transfected and non-transfected C2C12 myoblasts. β-DG was used as a loading control. Using IIH6, we saw the typical "smear" of α-DG at 140–200 kD in mouse skeletal muscle and a positive signal of about 120–130 kD in our non-transfected C2C12 myoblast cell line. This positive signal in myoblasts also shows laminin-binding activity by overlay assay (Figure 6A) and probably corresponds to partially glycosylated DG. Myoblasts transfected with *Large2* showed increased IIH6 immunoreactivity of mass 150–250 kD, which we presume corresponds to increased glycosylation of α-DG. This increase in IIH6 immunoreactivity is concomitant with an increase in laminin binding on overlay assay (Figure 6A) and indicates that Large2 is also able to act in the DG glycosylation pathway.

The high level of sequence similarity between *LARGE1* and *LARGE2*, the different expression profiles, and the ability of both proteins to increase glycosylation of α-DG suggests they may have tissue-specific roles in DG glycosylation. Furthermore, there may be functional redundancy between the two genes. Western analysis of a range of tissues from control and Large^{myd} mice showed correlation between preservation of DG glycosylation (indicated by retention of immunoreactivity with IIH6 and VIA4 antibodies) and high levels of *Large2* expression. Figure 6(B) shows a VIA4 immunoblot; Large^{myd} kidney tissue shows a strong positive signal with this antibody, in contrast to the other tissues where VIA4 immunoreactivity is significantly reduced or absent.

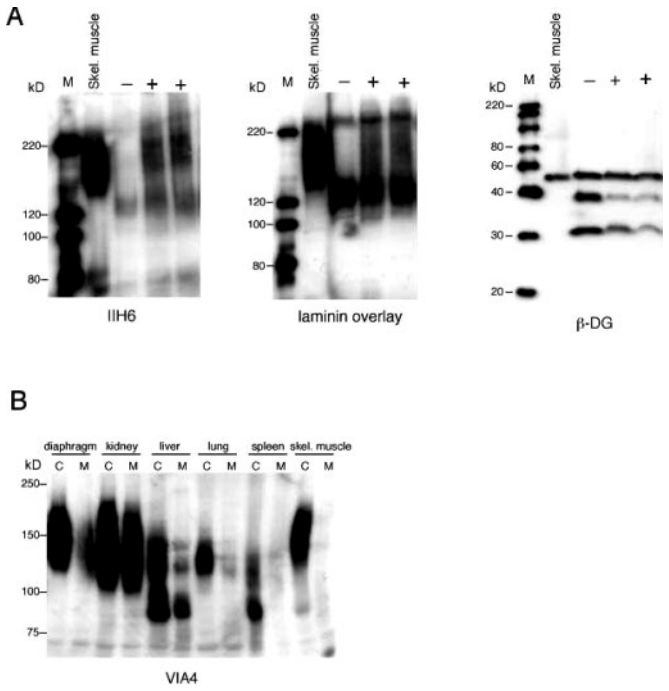


Fig. 6. Expression of *Large2* in C2C12 myoblasts results in hyperglycosylation of α-dystroglycan (α-DG). (**A**) Western blots with IIH6 (α-DG) or 8D5 (β-DG) using total protein extracts from mouse skeletal muscle, non-transfected myoblasts (–), or two separate experiments where myoblasts were transfected with pLarge2V5His (+). β-DG was used as a loading control. With IIH6, α-DG in mouse skeletal muscle is heterogeneous in size, ranging from 140 to 200 kD. There is a positive signal of about 120–130 kD in our non-transfected C2C12 myoblast cell line. This positive signal in myoblasts also shows laminin-binding activity by overlay assay and probably corresponds to partially glycosylated dystroglycan. Myoblasts transfected with *Large2* showed increased IIH6 immunoreactivity of 150–250 kD concomitant with an increase in laminin binding on the overlay assay. (**B**) A VIA4 immunoblot of total protein extracts from a range of tissues (quadriceps femoris was used for skeletal muscle extracts) from Large^{myd} littermates. C, wild type; M, homozygous mutant. Large^{myd} kidney tissue shows a strong positive signal with this antibody, in contrast to the other tissues where VIA4 immunoreactivity is significantly reduced or absent.

Discussion

The similarity in sequence and genomic organization of *LARGE1* and *LARGE2* suggests that that the two genes arose from duplication of an ancestral gene. From examination of genome databases and RT–PCR analysis, we have identified homologs of *LARGE1* and *LARGE2* in mammals, fish, and chicken genomes. This indicates that duplication of the ancestral *LARGE* gene occurred at least 450 million years ago, prior to the divergence of mammals, avians, and ray-finned fishes from their common ancestor (Kumar and Hedges, 1998). We searched extensively for further homologs of *LARGE* in these organisms employing a variety of bioinformatic approaches, but did not find any. Thus, this family of putative glycosyltransferases appears to contain only two members in vertebrates. For the mammalian *LARGE1* and *LARGE2* genes, the amino acid similarity between paralogs is confined to the two predicted catalytic domains. This appears to be due to acquisition of a novel initial coding exon by the mammalian *LARGE2* genes, because in other vertebrate orthologs, the similarity with LARGE1 extends over the whole protein including the N-terminal cytoplasmic, transmembrane, and coiled-coil domains. With the exception of *Fugu*, all the vertebrate *LARGE1* genes are 40- to 100-fold bigger than their *LARGE2* paralogs. This difference is due to the much bigger introns in the *LARGE1* orthologs. There is some evidence that important regulatory regions for *Large1* are widely dispersed: a mutant allele of *Large* caused by a transgene insertion 160 kb downstream of the gene has recently been reported (Levedakou *et al.*, 2005). This insertion results in silencing of *Large*, and homozygous mutant mice have a phenotype closely resembling the Large^{myd} mutation.

For mouse *Large2*, we identified alterative splicing within the part of the gene encoding the first catalytic domain. However, the functional consequences and significance of excluding this exon are unclear as RT–PCR data suggest this alternative splicing is not conserved in human (Figure 4A). Alternative splicing of 5′ UTR exons appears relatively common within glycosyltransferase genes and may be related to translational control (Martinez-Duncker *et al.*, 2004) or use of tissue-specific promoters (Taniguchi *et al.*, 2001; Falkenberg *et al.*, 2003). However, only a few splice variants that produce mRNAs with different protein-encoding capabilities have been described in the literature. For example, the mouse α1,2-galactosyltransferase gene generates four mRNA isoforms that differ within the stem region (Joziasse *et al.*, 1992).

We show that epitope-tagged LARGE1 and Large2 both localize to the Golgi apparatus and that the N-terminal 143 amino acids are sufficient for correct targeting and retention. It is possible that the LARGE1 and LARGE2 proteins act together as a complex, as has been demonstrated for the mammalian and *Drosophila* POMT1 and POMT2 mannosyltransferases (Ichimiya *et al.*, 2004; Manya *et al.*, 2004) and for the EXT1 and EXT2 glycosyltransferases involved in heparan sulfate synthesis (McCormick *et al.*, 2000). However, this seems unlikely since RT–PCR and dot-blot analyses of *LARGE1* and *LARGE2* show very different tissue-expression profiles. In addition, overexpression of each gene independently is sufficient to increase DG glycosylation.

Currently, the biochemical activities of LARGE and LARGE2 are still unknown. However, there is now a substantial body of evidence for a functional role for LARGE1 in DG glycosylation (Grewal *et al.*, 2001; Michele *et al.*, 2002; Barresi *et al.*, 2004; Kanagawa *et al.*, 2004; Patnaik and Stanley, 2005). Of the other genes associated with dystroglycanopathies, *POMT1* and *POMGnT1* are probably involved in O-mannosylation (Zhang *et al.*, 2002; Manya *et al.*, 2004), but the biochemical activities of fukutin and FKRP are unknown. In cells deficient in POMT1, POMGnT1, or fukutin, forced expression of *LARGE* still results in hyperglycosylation of α-DG (Barresi *et al.*, 2004), indicating that, at least in this experimental overexpression system, LARGE may act in a different glycosylation pathway. Overexpression studies in CHO cells defective for a number of steps in glycan synthesis suggest that LARGE1

can modify both *N*- and *O*-glycans (Patnaik and Stanley, 2005). On the basis of homologies of the catalytic domains, both LARGE1 and LARGE2 might add a repeated glycan structure such as poly-*N*-acetyllactosamine, which can be attached to *N*-glycans, *O*-glycans, and glycolipids (Ujita and Fukuda, 2001). It is clear that determination of the biochemical activity of LARGE will be a key step toward determining its relationship to the other glycosylation proteins associated with dystroglycanopathies.

DG is a widely expressed, essential glycoprotein (Henry and Campbell, 1999). Null mutations of DG in the mouse are embryonic lethal at very early stages of development (Williamson *et al*., 1997), as are engineered *fukutin* (Takeda *et al*., 2003) and *POMT1* (Willer *et al*., 2004) null mice; yet homozygous Large^{myd} mice are viable and only a restricted number of issues are affected (Holzfeind *et al*., 2002). We speculate that partial functional redundancy between *Large1* and *Large2* might account for the relatively mild phenotype of the Large^{myd} mouse. For example, glycosylation of DG is preserved in the kidney, a tissue that normally expressed high levels of Large2 mRNA in homozygous mutant mice. Overlapping functions have been postulated for other glycosylation genes such as the UDP-*N*-acetylgalactosamine:polypeptide *N*-acetylgalactosaminyl-transferases (Ten Hagen *et al*., 2003) and could have implications for development of therapeutic strategies for muscular dystrophy. As the overexpression of *LARGE1* can restore IIH6 immunoreactivity and laminin-binding activity to α-DG in cells from patients with mutations in distinct genes, up-regulation of this gene might be applicable to different dystroglycanopathies (Barresi *et al*., 2004). Our data suggest that *LARGE2* should be considered an additional potential therapeutic target because it is also able to increase DG glycosylation in myoblast cells.

Materials and methods

Database and bioinformatics analysis

Blast searches of Ensembl genomic assemblies (http://www.ensembl.org) were used to locate genes. Genomic and EST sequences were assembled using Sequencher software (Gene Codes Corporation, Ann Arbor, MI). Gene structures were then predicted with a combination of Ensembl gene predictions, BLAST searches, and Nucleotide Identity of X analysis (http://www.rfcgr.mrc.ac.uk). A ClustalX multiple protein sequence alignment (bios.u-strasb.fr) was generated using predicted amino acid sequences for the LARGE homologs. The alignment is displayed using Mac Boxshade (iubio.bio.indiana.edu). Transmembrane, coiled-coil, and catalytic domain predictions were made using Protein Identity of X analysis (http://www.rfcgr.mrc.ac.uk). The maximum likelihood tree was generated using the ProML program within the Phylip suite of programs (evolution.gs.washington.edu/phylip/) and displayed using Dendromaker (http://www.cib.nig.ac.jp/dda/timanish/dendromaker/ home.html).

Confirmation of coding sequences and expression analysis

For RT–PCR amplification, RNA was isolated from tissues or embryos using TRIZOL (Invitrogen, Paisley, Scotland) or purchased from Ambion (Huntingdon, Cambs, UK). First-strand cDNA was synthesized using oligodT primers and an ImPROM-II cDNA synthesis kit (Promega, Madison, WI). Gene-specific primers were then used in PCR to generate overlapping products (sequences supplied on request). PCR products were then subcloned into pT-EASY (Promega) and sequenced, allowing assembly of full coding sequences. For RT–PCR expression analysis in mouse, primer sequences used were: LargeF2 5′-TTCCTGGCTGCCTCTTTGAC-3′ and LargeR2 5′-TGGACGTGTGTTTTTCCAGA-3′; Large2F2 5′-CCAGGCTCTCACACCCTACG-3′ and Large2R2 5′-GGACACCCTGGTCAGCCTCT-3′. For both primer pairs, an annealing temperature of 60°C was used with 30 cycles of amplification. To test for alternative splicing of *Large2*, we used primers AltspliceF 5′-ACTCTTCCGAACATGGATGG-3′ and AltspliceR 5′-GACCGCATTGAAGATGTCCT-3′. In mouse, these primers amplify a product of 472 bp if exon 7 is included in the mRNA and a product of 367 bp if it is omitted. For human, we used primers HLarge2f 5′-GCCAGAAACATCCTGGAGAC-3′ and HLarge2r 5′-TCACAGCGTTGAAGATGTCC-3′. An annealing temperature of 64°C was used with 30 amplification cycles.

The tissue distributions of human *LARGE1* and *LARGE2* mRNAs were investigated by hybridizing multiple-tissue RNA dot-blots (Clontech, San Diego, CA) with radiolabeled gene-specific cDNA probes. The probes corresponded to 679–1433 bp of GenBank Accession number NM_004737 for *LARGE1* or to 582–1385 bp of GenBank Accession number NM_152312 for *LARGE2*. Filters were washed in 0.2× SSC at 65°C and exposed for 72 h. A northern blot containing 10 μg of total human RNA from several human and fetal tissues (RNA was purchased from Stratagene, La Jolla, CA) was hybridized using the same probes and conditions.

Construct generation and transient transfections

PCR (using Proofstart Taq polymerase, Qiagen, Crawley, Sussex, UK) was used to amplify the complete coding regions of human *LARGE* and mouse *Large2* from cDNA clones. The PCR products were then subcloned into the *Sal*I/*Sac*II sites of pEGFP-N1 and pDsRed2-N1 (Clontech) to make C-terminal fusion proteins. The complete open-reading frame of the mutant mRNA product in the Large^{myd} mouse was amplified by RT–PCR and cloned into the EcoRI/BamHI sites of pEGFP-N1 or pDsRed2-N1. This Large^{myd} product was also cloned into pEGFP-C1 to make an N-terminal fusion. For the transient expression of Large2 protein in C2C12 cells, the full-length coding region was amplified and cloned in-frame into the SalI/SacII sites of pcDNAV5His (Invitrogen). All primer sequences will be supplied on request. All constructs were fully sequenced on both strands to ensure that they did not contain polymerase errors.

NRK cells or C2C12 myoblasts were grown and maintained (5% CO_2, 37°C) in DMEM supplemented with 10% fetal calf serum and transfected using Effectene transfection reagent (Qiagen). For immunocytochemistry, cells were seeded onto glass coverslips. Twenty hours after transfection, the cells were washed and fresh media added. Forty-eight hours after transfection, cells were either treated for

immunocytochemistry or harvested for western blot analysis. Protein extraction from cell lines or tissues and western blotting was carried out as described previously (Grewal *et al*., 2001).

Immunocytochemistry

Transfected cells were fixed with 4% paraformaldehyde in PBS. To determine the effect of BFA treatment, cells were incubated with BFA (10 μg/ml) for 30 min at 37°C immediately prior to washing and fixing with paraformaldehyde. The cells were then treated for 10 min on ice with 0.5% Triton X-100 in PBS, washed with PBS, and then incubated for 30 min at room temperature with 2% goat serum and 10 mg/ml BSA in PBS. Following incubation with primary antibody for 1 h, cells were washed with PBS, then incubated for 1 h with 1:500 dilution of rhodamine Red-X-labeled goat anti-mouse IgG antibody (Molecular Probes, Leiden, The Netherlands). Finally, cells were mounted in VectaShield mounting medium with DAPI. Primary antibodies were used at the following dilutions: α-mannosidase II (53FC3, Covance, Cambridge, Cambs, UK) at 1:10,000, golgin97 (Molecular Probes) at 1:25, and GM130 (BD Biosciences, San Diego, CA) at 1:500.

Images were examined using a Zeiss LSM510uv Combi META confocal attached to a Zeiss Axiovert 100 microscope. Image acquisition and processing were performed using Zeiss LSM software. Images were imported into Photoshop (Adobe) for generation of figures.

Acknowledgments

The authors thank Tim Self for assistance with confocal microscopy. We acknowledge the funding from The Wellcome Trust, the Muscular Dystrophy Association, USA, and the Biotechnology and Biological Sciences Research Council (BBSRC) to JEH. JEH holds a BBSRC Research Development Fellowship.

Abbreviations

BFA, brefeldin A; DG, dystroglycan; ER, endoplasmic reticulum; EST, expressed sequence tag; LGMD, limb-girdle muscular dystrophy; MDC, congenital muscular dystrophy; MEB, muscle-eye-brain disease; RT–PCR, reverse-transcriptase polymerase chain reaction; WWS, Walker–Warburg syndrome.

Note added in proof

While this manuscript was under revision, two papers were published showing that LARGE proteins localize to the Golgi (Brockington *et al*. 2005) and that *LARGE2* expression induces α-DG hyperglycosylation and laminin binding in CHO and HEK cells (Brockington *et al*. 2005; Fujimura *et al.* 2005).

References

Barresi, R., Michele, D.E., Kanagawa, M., Harper, H.A., Dovico, S.A., Satz, J.S., Moore, S.A., Zhang, W., Schachter, H., Dumansk, J.P., and others. (2004) LARGE can functionally bypass α-dystroglycan glycosylation defect in distinct congenital muscular dystrophy. *Nat. Med.*, **10**, 696–703.

Beltrán-Valero de Bernabé, D., Currier, S., Steinbrecher, A., Celli, J., van Beusekom, E., van der Zwaag, B., Kayserili, H., Merlini, L., Chitayat, D., Dobyns, W.B., and others. (2002) Mutations in the *O*-mannosyltransferase gene POMT1 give rise to the severe neuronal migration disorder Walker–Warburg syndrome. *Am. J. Hum. Genet.*, **71**, 1033–1043.

Beltrán-Valero de Bernabé, D., Voit, T., Longman, C., Steinbrecher, A., Straub, V., Yuva, Y., Herrmann, R., Sperner, J., Korenke, C., Diesen, C., and others. (2004) Mutations in the FKRP gene can cause muscle-eye-brain disease and Walker–Warburg syndrome. *J. Med. Genet*., **41**, e61.

Brancaccio, A., Schulthess, T., Gesemann, M., and Engel, J. (1995) Electron microscopic evidence for a mucin-like region in chick muscle α-dystroglycan. *FEBS Lett.*, **368**, 139–142.

Brockington, M., Torelli, S., Prandini, P., Boito, C., Dolatshad, N.F., Longman, C., Brown, S.C., and Muntoni, F. (2005) Localization and functional analysis of the LARGE family of glycosyltransferases: significance for muscular dystrophy. *Hum. Mol. Genet.*, **15**, 657–665.

Campanelli, J., Roberds, S., Campbell, K., and Scheller, R. (1994) A role for dystrophin-associated glycoproteins and utrophin in agrin-induced AchR clustering. *Cell*, **77**, 663–674.

Cohen, M.W., Jacobsen, C., Godfrey, E.W., Campbell, K.P., and Carbonetto, S. (1995) Distribution of α-dystroglycan during embryonic nerve-muscle synaptogenesis. *J. Cell Biol.*, **129**, 1093–1101.

Côté, P.D., Moukhles, H., Lindebaum, M., and Carbonetto, S. (1999) Chimaeric mice deficient in dystroglycans develop muscular dystrophy and have disrupted myoneural synapses. *Nat. Genet.*, **23**, 338–342.

Coutinho, P.M., Deleury, E., Davies, G.J., and Henrissat, B. (2003) An evolving hierarchical family classification for glycosyltransferases. *J. Mol. Biol.*, **328**, 307–317.

Doms, R.W., Russ, G., and Yewdell, J.W. (1989) Brefeldin A redistributes resident and itinerant Golgi proteins to the endoplasmic reticulum. *J. Cell Biol.*, **109**, 61–72.

Ervasti, J.M. and Campbell, K.P. (1991) Membrane organization of the dystrophin–glycoprotein complex. *Cell*, **66**, 1121–1131.

Falkenberg, V.R., Alvarez, K., Roman, C., and Fregien, N. (2003) Multiple transcription initiation and alternative splicing in the, 5′ untranslated region of the core 2 beta 1-6 *N*-acetylglucosaminyltransferase I gene. *Glycobiology*, **13**, 411–418.

Fujimura, K., Sawaki, H., Sakai, T., Hiruma, T., Nakanishi, N., Sato, T., Ohkura, T., and Narimatsu, H. (2005) LARGE2 facilitates the maturation of alpha-dystroglycan more effectively than LARGE. *Biochem. Biophys. Res. Commun.*, **329**, 1162–1171.

Gee, S.H., Montanaro, F., Lindenbaum, M.H., and Carbonetto, S. (1994) Dystroglycan-α, a dystrophin-associated glycoprotein, is a functional agrin receptor. *Cell*, **77**, 675–686.

Greener, M.J. and Roberts, R.G. (2000) Conservation of components of the dystrophin complex in *Drosophila*. *FEBS Lett.*, **482**, 13–18.

Grewal, P.K. and Hewitt, J.E. (2002) Mutation of *Large*, which encodes a putative glycosyltransferase, in an animal model of muscular dystrophy. *Biochim. Biophys. Acta*, **1573**, 216–224.

Grewal, P.K., Holzfeind, P.J., Bittner, R.E., and Hewitt, J.E. (2001) Mutant glycosyltransferase and altered glycosylation of α-dystroglycan in the myodystrophy mouse. *Nat. Genet.*, **28**, 151–154.

Grisoni, K., Martin, E., Gieseler, K., Mariol, M.C., and Segalat, L. (2002) Genetic evidence for a dystrophin–glycoprotein complex (DGC) in *Caenorhabditis elegans*. *Gene*, **294**, 77–86.

Heinrichs, D.E., Yethon, J.A., and Whitfield, C. (1998) Molecular basis for structural diversity in the core region of the lipopolysaccharides of *Escherichia coli* and *Salmonella enterica*. *Mol. Micro.*, **30**, 221–232.

Henry, M.D. and Campbell, K.P. (1999) Dystroglycan inside and out. *Curr. Opin. Cell Biol.*, **11**, 602–607.

Hewitt, J.E. and Grewal, P.K. (2003) Glycosylation defects: a new mechanism for muscular dystrophy? *Hum. Mol. Genet.*, **12**, R259–R264.

Holzfeind, P.J., Grewal, P.K., Reitsamer, H.A., Kechvar, J., Lassmann, H., Hoeger, H., Hewitt, J.E., and Bittner, R.E. (2002) Skeletal, cardiac and tongue muscle pathology, defective retinal transmission, and neuronal migration defects in the Large^{myd} mouse defines a natural model for glycosylation-deficient muscle-eye-brain disorders. *Hum. Mol. Genet.*, **11**, 2673–2687.

Ibraghimov-Beskrovnaya, O., Ervasti, J.M., Leveille, C.J., Slaughter, C.A., Sernett, S.W., and Campbell, K.P. (1992) Primary structure of

dystrophin-associated glycoproteins linking dystrophin to the extracellular matrix. *Nature*, **355**, 696–702.

Ichimiya, T., Manya, H., Ohmae, Y., Yoshida, H., Takahashi, K., Ueda, R., Endo, T., and Nishihara, S. (2004) The twisted-abdomen phenotype of *Drosophila POMT1* and *POMT2* mutants coincides with their heterophilic protein *O*-mannosyltransferase activity. *J. Biol. Chem.*, **279**, 42638–42647.

Joziasse, D.H., Shaper, N.L., Kim, D., van den Eijnden, D.H., and Shaper, J.H. (1992) Murine α1,3-galactosyltransferase. *J. Biol. Chem.*, **267**, 5534–5541.

Kanagawa, M., Saito, F., Kunz, S., Yoshida-Moriguchi, T., Barresi, R., Kobayashi, Y.M., Muschler, J., Dumanski, J.P., Michele, D.E., Oldstone, M.B., and Campbell, K.P. (2004) Molecular recognition by LARGE is essential for expression of functional dystroglycan. *Cell*, **117**, 953–964.

Kjer-Nielsen, L., Teasdale, R.D., van Vliet, C., and Gleeson, P.A. (1999) A novel Golgi-localisation domain shared by a class of coiled-coil peripheral membrane proteins. *Curr. Biol.*, **9**, 385–388.

Kobayashi, K., Nakahori, Y., Miyake, M., Matsumura, K., Kondo-Iida, E., Nomura, Y., Segawa, M., Yoshioka, M., Saito, K., Osawa, M., and others. (1998) An ancient retrotransposal insertion causes Fukuyama-type congenital muscular dystrophy. *Nature*, **394**, 388–392.

Kumar, S. and Hedges, S.B. (1998) A molecular timescale for vertebrate evolution. *Nature*, **392**, 917–920.

Levedakou, E.N., Chen, X.-J., Soliven, B., and Popko, B. (2005) Disruption of the mouse *Large* gene in the *enr* and *myd* mutants results in nerve, muscle and neuromuscular defects. *Mol. Cell. Neurosci.*, **28**, 757–769.

Lippincott-Schwartz, J., Yuan, L.C., Bonifacino, J.S., and Klausner, R.D. (1989) Rapid redistribution of Golgi proteins into the ER in cells treated with brefeldin A: evidence for membrane cycling from Golgi to ER. *Cell*, **56**, 801–813.

Longman, C., Brockington, M., Torelli, S., Jimenez-Mallebrera, C., Kennedy, C., Khalil, N., Feng, L., Saran, R.K., Voit, T., Merlini, L., and others. (2003) 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.*, **12**, 2853–2861.

Manya, H., Chiba, A., Yoshida, A., Wang, X., Chiba, Y., Jigami, Y., Margolis, R.U., and Endo, T. (2004) Demonstration of mammalian protein *O*-mannosyltransferase activity: coexpression of POMT1 and POMT2 required for enzymatic activity. *Proc. Natl. Acad. Sci. USA*, **101**, 500–505.

Martin, P.T. (2003) Dystroglycan glycosylation and its role in matrix binding in skeletal muscle. *Glycobiology*, **13**, 55R–66R.

Martin, P.T. and Freeze, H.H. (2003) Glycobiology of neuromuscular disorders. *Glycobiology*, **13**, 67R–75R.

Martinez-Duncker, I., Michalski, J.-C., Bauvy, C., Candelier, J.-J., Mennesson, B., Codogno, P., Oriol, R., and Mollicone, R. (2004) Activity and tissue distribution of splice variants of α6-fucosyltransferase in human embryogenesis. *Glycobiology*, **14**, 13–25.

McCormick, C., Duncan, G., Goutsos, K.T., and Tufaro, F. (2000) The putative tumor suppressors EXT1 and EXT2 form a stable complex that accumulates in the Golgi apparatus and catalyzes the synthesis of heparan sulfate. *Proc. Natl. Acad. Sci. USA*, **97**, 668–673.

Mercuri, E., Brockington, M., Straub, V., Quijano-Roy, S., Yuva, Y., Herrmann, R., Brown, S.C., Torelli, S., Dubowitz, V., Blake, D.J., and others. (2003) Phenotypic spectrum associated with mutations in the fukutin-related protein gene. *Ann. Neurol.*, **53**, 537–542.

Michele, D.E. and Campbell, K.P. (2003) Dystrophin–glycoprotein complex: post-translational processing and dystroglycan function. *J. Biol. Chem.*, **278**, 15457–15460.

Michele, D.E., Barresi, R., Kanagawa, M., Saito, F., Cohn, R.D., Satz, J.S., Dollar, J., Nishino, I., Kelley, R.I., Somer, H., and others. (2002) Post-translational disruption of dystroglycan–ligand interactions in congenital muscular dystrophies. *Nature*, **418**, 417–422.

Moore, S.A., Saito, F., Chen, J., Michele, D.E., Henry, M.D., Messing, A., Cohn, R.D., Ross-Barta, S.E., Westra, S., Williamson, R.A., and others. (2002) Deletion of brain dystroglycan recapitulates aspects of congenital muscular dystrophy. *Nature*, **418**, 422–425.

Nakamura, N., Rabouille, C., Watson, R., Nilsson, T., Hui, N., Slusarewicz, P., Kreis, T.E., and Warren, G. (1995) Characterization of a *cis*-Golgi matrix protein, GM130. *J. Cell Biol.*, **131**, 1715–1726.

Parsons, M.J., Campos, I., Hirst, E.M., and Stemple, D.L. (2002) Removal of dystroglycan causes severe muscular dystrophy in zebrafish embryos. *Development*, **129**, 3505–3512.

Patnaik, S.K. and Stanley, P. (2005) Mouse Large can modify complex *N*-and mucin *O*-glycans on α-dystroglycan to induce laminin binding. *J. Biol. Chem.*, **280**, 20851–20859.

Peyrard, M., Seroussi, E., Sandberg-Nordqvist, A.-C., Xie, Y.-G., Han, F.-Y., Fransson, I., Collins, J., Dunham, I., Kost-Alimova, M., and Imreh, S. (1999) The human *LARGE* gene from 22q12.3-q13.1 is a new distinct member of the glycosyltransferase gene family. *Proc. Natl. Acad. Sci. USA*, **96**, 598–603.

Sasaki, K., Kurata-Miura, K., Ujita, M., Angata, K., Nakagawa, S., Sekine, S., Nishi, T., and Fukuda, M. (1997) Expression cloning of cDNA encoding a human β-1,3-*N*-acetylglucosaminyltransferase that is essential for poly-*N*-acetyllactosamine synthesis. *Proc. Natl. Acad. Sci. USA*, **94**, 14294–14299.

Singh, J., Itahana, Y., Knight-Krajewski, S., Kanagawa, M., Campbell, K.P., Bissell, M.J., and Muschler, J. (2004) Proteolytic enzymes and altered glycosylation modulate dystroglycan function in carcinoma cells. *Cancer Res.*, **64**, 6152–6159.

Takeda, S., Kondo, M., Sasaki, J., Kurahasi, H., Kano, H., Arai, K., Misaki, K., Fukui, T., Kobayashi, K., Tachikawa, M., and others. (2003) Fukutin is required for maintenance of muscle integrity, cortical histiogenesis and normal eye development. *Hum. Mol. Genet.*, **12**, 1449–1459.

Taniguchi, A., Kaneta, R., Morishita, K., and Matsumoto, K. (2001) Gene structure and transcriptional regulation of human gal beta, 1,4(3) GlcNAc alpha 2,3-sialyltransferase VI (hST3Gal VI) gene in prostate cancer cell line. *Biochem. Biophys. Res. Commun.*, **287**, 1148–1156.

Ten Hagen, K.G., Fritz, T.A., and Tabak, L.A. (2003) All in the family: the UDP-GalNAc: polypeptide *N*-acetylgalactosaminyltransferases. *Glycobiology*, **13**, 1R–16R.

Ujita, M. and Fukuda, M. (2001) Regulation of poly-*N*-acetyllactosamine biosynthesis in *O*-glycans. *Trends Glycosci. Glycotechnol.*, **13**, 177–191.

Velasco, A., Hendricks, L., Moremen, K.W., Tulsiani, D.R., Touster, O., and Farquhar, M.G. (1993) Cell type-dependent variations in the subcellular distribution of alpha-mannosidase I and II. *J. Cell Biol.*, **122**, 39–51.

Wiggins, C.A.R. and Munro, S. (1998) Activity of the yeast MNN1 alpha-1,3-mannosyltransferase requires a motif conserved in many other families of glycosyltransferases. *Proc. Natl. Acad. Sci., USA*, **95**, 7945–7950.

Willer, T., Prados, B., Falcón-Pérez, J.M., Renner-Müller, I., Przemeck, G.K.H., Lommel, M., Coloma, A., Valero, M.C., Hrabé de Angelis, M., Tanner, W., and others. (2004) Targeted disruption of the Walker–Warburg syndrome gene *Pomt1* in mouse results in embryonic lethality. *Proc. Natl. Acad. Sci. USA*, **101**, 14126–14131.

Williamson, R.A., Henry, M.D., Daniels, K.J., Hrstka, R.F., Lee, C.J., Sunada, Y., Ibraghimov-Beskrovnaya, O., and Campbell, K.P. (1997) Dystroglycan is essential for early embryonic development: disruption of Reichart's membrane in Dag1-null mice. *Hum. Mol. Genet.*, **6**, 831–841.

Winder, S.J. (2001) The complexities of dystroglycan. *Trends Biochem. Sci.*, **26**, 118–124.

Yoshida, A., Kobayashi, K., Manya, H., Taniguchi, K., Kano, H., Mizuno, M., Inazu, T., Mitsuhashi, H., Takahashi, S., Takeuchi, M., and others. (2001) Muscular dystrophy and neuronal migration disorder caused by mutations in a glycosyltransferase, POMGnT1. *Dev. Cell*, **1**, 717–724.

Zhang, W.L., Betel, C., and Schachter, H. (2002) Cloning and expression of a novel UDP-GlcNAc: alpha-D-mannoside beta, 1,2-*N*-acetylglucosaminyltransferase homologous to UDP-GlcNAc: alpha-3-D-mannoside beta **1**,2-*N*-acetylglucosaminyltransferase I. *Biochem. J.*, **361**, 153–162.