Molecular bases of autosomal recessive limb-girdle muscular dystrophies

V. Nigro
Dipartimento di Patologia Generale e Centro di Eccellenza sulle Malattie Cardiovascolari, Seconda Università degli Studi di Napoli, Napoli, Italy
Telethon Institute of Genetics and Medicine (TIGEM), Napoli, Italy

Limb-girdle muscular dystrophies (LGMD) are a heterogeneous group of genetically determined disorders with a primary or predominant involvement of the pelvic or shoulder girdle musculature. The clinical course is characterized by great variability, ranging from severe forms with rapid onset and progression to very mild forms allowing affected people to have fairly normal life spans and activity levels. Sixteen loci have been so far identified, six autosomal dominant and ten autosomal recessive. Linkage analyses indicate that there is further genetic heterogeneity both for dominant as well as for recessive LGMD. The dominant forms (LGMD1) are generally milder and relatively rare, representing less than 10% of all LGMD. The autosomal recessive forms (LGMD2) are much more common, having a cumulative prevalence of 1:15,000 with a number of geographical differences. The product of ten autosomal recessive LGMD genes has so far been identified. They are: calpain-3 (LGMD2A), dysferlin (LGMD2B), alpha-sarcoglycan (LGMD2D), beta-sarcoglycan (LGMD2E), gamma-sarcoglycan (LGMD2C), delta-sarcoglycan (LGMD2F), telethonin (LGMD2G), TRIM32 (LGMD2H), fukutin-related protein (LGMD2I) and titin (LGMD2J). There are, however, at least 25% of families who can be excluded from any known locus. The present review is devoted to outline the present advancements in the molecular bases of autosomal recessive LGMD.

Key words: Limb-girdle muscular dystrophies; sarcoglycan.

Introduction

The term muscular dystrophy refers to a group of degenerative diseases characterized by progressive weakening and deterioration of skeletal muscles and finally the heart and respiratory muscles. Muscular dystrophies are clinically and genetically heterogeneous. The most severe forms result in dramatic physical weakness, so children lose the ability to do things like walk, sit upright, breathe easily, and move their arms and hands. The increasing weakness often leads to other serious complications and, for many, a shortened life span.

The classical Duchenne and Becker muscular dystrophies are X-linked, but the differential diagnosis is with numerous autosomal forms newly identified with similar or milder clinical outcome, defined limb-girdle muscular dystrophies or LGMD. This acronym now defines a group of genetically determined disorders with a primary or predominant involvement of the pelvic or shoulder girdle musculature. The clinical course is characterized by great variability, ranging from severe forms with rapid onset and progression to very mild forms allowing affected people to have fairly normal life spans and activity levels (1-3).

Sixteen LGMD loci have been so far identified, six autosomal dominant and ten autosomal recessive. Linkage analyses indicate that there is further genetic heterogeneity both for dominant as well as for recessive LGMD (Table). The dominant forms (LGMD1) are generally milder and relatively rare representing less than 10% of all LGMD. The autosomal recessive forms (LGMD2) are much more common, having a cumulative prevalence of 1:15,000 with some differences among countries, depending on the carrier distribution and the degree of consanguinity.

Three genes for LGMD1 have been identified: myotilin for LGMD1A (4), lamin A/C for LGMD1B (5) and caveolin-3 for LGMD1C (6). Additional three forms (LGMD1D, LGMD1E and LGMD1F) have been mapped but the genes are still unknown (7-9). At present ten autosomal recessive LGMD genes have been recognized (10). They are: calpain-3 for LGMD2A (11), dysferlin for LGMD2B (12,13), alpha-sarcoglycan (adhalin) for LGMD2D (14), beta-sarcoglycan for LGMD2E (15,16), gamma-sarcoglycan for LGMD2C (17),...
delta-sarcoglycan for LGMD2F (18,19), telethonin for LGMD2G (20), TRIM32 for LGMD2H (21), fukutin-related protein for LGMD2I (22) and titin for LGMD2J, even if this gene has been conclusively recognized as a cause of tibial muscular dystrophy (23). There are, however, at least 25% of families who can be excluded from any known locus and 40% of isolated LGMD cases with no mutation in any known gene (unpublished). The present review is devoted to summarize the molecular bases of autosomal recessive LGMD.

**LGMD2A**

The identification of the basis of LGMD2A (11) was obtained six years after the cloning of calpain-3 gene (CAPN3) (24). The CAPN3 gene spans 53kb of genomic sequence at chromosome 15q15.2 (25) and the transcript is composed of 24 exons encoding a 94kDa muscle-specific protein. Calpains are intracellular nonlysosomal cysteine proteases modulated by calcium ions. A typical calpain is a heterodimer composed of two distinct subunits, one large (>80 kDa) and the other small (30kDa). While only one gene encoding the small subunit has been demonstrated, in the last few years several large subunit cDNAs have been cloned from different organisms in addition to m-calpain and micro-calpain. They have ubiquitous or tissue-specific expression. Calpains cleave target proteins to modify their properties, rather than “break down” the substrates (26). Why a defect in an enzymatic rather than a structural protein causes muscular dystrophy? Calpain 3 may have a signalling role by controlling IkappaBalpha turnover and indirectly regulating NF-kappaB dependent expression of survival genes (27). Other data indicate that calpain 3 specifically binds to connectin/titin (28), and its activity is probably suppressed by this binding. This would link LGMD2A to titin partners. Overexpression of a calpastatin transgene in mdx muscle (dystrophin-deficient) reduced dystrophic pathology, this suggest that calpains may play an active role in necrotic processes in dystrophic muscle (29).

LGMD2A is the most frequent form of LGMD: to date, >120 different causative mutations have been reported (30). There is a marked clinical variability in calpain patients that poorly correlate with the mutation type (31). In a large study 53/191 of patients with LGMD phenotype (28%) had partial or total calpain-3 deficiency, while it was observed in 8% of hyperCKemia, 6% of proximal myopathy, and 6% of distal myopathy patients (32). We found CAPN3 mutations in 36% of Italian LGMD patients with severe proximal muscle involvement and major motor disabilities (unpublished results).

| V. Nigro |

**Table 1. LGMD loci.**

<table>
<thead>
<tr>
<th>Autosomal dominant LGMD</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>• LGMD1A 5q31.2 myotilin (4) MIM 159000</td>
<td></td>
</tr>
<tr>
<td>• LGMD1B 1q22 lamin A/C (5) MIM 159001</td>
<td></td>
</tr>
<tr>
<td>• LGMD1C 3p25.3 caveolin 3 (6) MIM 607801</td>
<td></td>
</tr>
<tr>
<td>• LGMD1D 6q22 ? (7)</td>
<td></td>
</tr>
<tr>
<td>• LGMD1E 7q35 ? (8)</td>
<td></td>
</tr>
<tr>
<td>• LGMD1F 7q31.1 ? (9)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Autosomal recessive LGMD</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>• LGMD2A 15q15.2 calpain 3 (11) MIM 253600</td>
<td></td>
</tr>
<tr>
<td>• LGMD2B 2p13.2 dysterlin (12,13) MIM 253601/254130</td>
<td></td>
</tr>
<tr>
<td>• LGMD2C 13q12 gamma-sarcoglycan (17) MIM 253700</td>
<td></td>
</tr>
<tr>
<td>• LGMD2D 17q21.33 alpha-sarcoglycan (14) MIM 600119</td>
<td></td>
</tr>
<tr>
<td>• LGMD2E 4q11 beta-sarcoglycan (15,16) MIM 604286</td>
<td></td>
</tr>
<tr>
<td>• LGMD2F 5q33.3 delta-sarcoglycan (18,19) MIM 601287</td>
<td></td>
</tr>
<tr>
<td>• LGMD2G 17q12 telethonin (20) MIM 601954</td>
<td></td>
</tr>
<tr>
<td>• LGMD2H 9q33.1 TRIM32 (21) MIM 254110</td>
<td></td>
</tr>
<tr>
<td>• LGMD2I 19q13.32 FKRP (22) MIM 607155</td>
<td></td>
</tr>
<tr>
<td>• LGMD2J 2q24.3 titin ?</td>
<td></td>
</tr>
</tbody>
</table>
LGMD2B

In 1998 two groups cloned the gene of LGMD2B. The proposed name ‘dysferlin’ combined the role of the gene in producing muscular dystrophy with its relatedness to the spermatogenesis factor fer-1 of Caenorhabditis elegans (12,13). The dysferlin gene (DYSF) spans 233kb of genomic sequence at chromosome 2p13.2 and the major transcript is composed of 6,911 nt (25). There are additional human fer-1-like genes: myoferlin at 10q24 (33), otoferlin at 2p23 that is mutated in nonsyndromic deafness type9 (DFNB9) (34), FER1L3 at 10q23.3 (35) and Fer-1-like 4 at 20q11.

Dysferlin is a membrane protein of about 230kDa containing six hydrophilic C2 domains. These are present in many membrane-associated proteins and bind calcium and anionic phospholipids (36). In synaptotagmins C2 motifs have been involved in regulating the process of calcium-dependent fast exocytosis (37). Dysferlin expression is not limited to muscle and starts at 5-6 weeks (38). Dysferlin-null mice were generated by homologous recombination deleting the transmembrane domain. Dysferlin-null mice maintain a functional dystrophin-glycoprotein complex, but nevertheless develop a progressive muscular dystrophy. In contrast with mdx mice, where exercise causes sarcolemma damage and increased CPK, dysferlin-deficient mouse sarcolemma is structurally stable. Campbell suggested that dysferlin-deficient muscle fibres are defective in Ca(2+)-dependent sarcolemma resealing. In normal muscle membrane repair is an active process in skeletal muscle fibres for which dysferlin may have an essential role (39).

Recessively inherited mutations in the dysferlin gene cause muscular dystrophies called “dysferlinopathies”, which include limb-girdle muscular dystrophy type 2B (LGMD2B) and Miyoshi myopathy (MM) that is an adult-onset distal muscular dystrophy. LGMD2B affects earlier the proximal muscles of the arms whereas MM affects the posterior muscles of the leg. Weiler et al. (40) reported that in large Canadian kindred with both LGMD2B and MM patients, all affected individuals were homozygous for the same mutation (Pro791Arg) with a similar reduction of dysferlin expression in the two types of patients. Additional identical dysferlin mutations were found in limb-girdle muscular dystrophy type 2B and distal myopathy. Additional factors (e.g., modifier or booster genes) may contribute to the phenotypic expression of causative mutations in dysferlinopathies (41).

Matsuda et al. (42) reported that dysferlin coimmunoprecipitates with caveolin-3 from biopsied normal human skeletal muscles. Caveolin-3 (CAV3) is a skeletal muscle membrane protein which is important in the formation of caveolae and in which mutations cause dominantly inherited limb girdle muscular dystrophy type 1C. A secondary deficiency of Calpain 3 was observed in eight out of 16 LGMD2B families (43). This may indicate an interaction between dysferlin and Calpain-3 in muscle. No reduction of the CAPN3 transcript was detected.

The SJL mouse strain that is a model for susceptibility to many induced autoimmune diseases carries a homozygous 171-bp in-frame deletion in the Dysf gene, removing most of the conserved fourth C2 domain (44). Muscles of eight-month-old SJL/J mice had higher levels of MHC class I expression than muscles of either C57BL/6J (wild-type) or SJL/J B2m (-/-) mice. (45) Campanaro et al (46) performed gene expression profiling using muscle from eight LGMD2B patients. Again MHC class I genes, S100 calcium-binding protein, sarcolipin, and genes involved in protein biosynthesis were up-regulated, while sarcomeric protein such as titin, nebulin and telethonin were down-regulated. In the majority of dysferlin patients, macrophages infiltrates and CD4+>CD8+ cells are found. This is consistent with the activation of a TH1 pathway.

To date, about 70 different variations have been found along the 55 exons of the DYSF gene with no clear mutation hot spot. The incidence of dysferlin mutations in LGMD was assessed by different methods. By WB analysis Fanin et al (32) found a dysferlin deficiency in 6.4% of patients with LGMD and myopathy, while Tagawa et al. (47) found a deficiency of dysferlin protein in 19% of LGMD and 75% of MM patients. There is no general agreement about the fraction of LGMD due to DYSF mutations. We estimate that DYSF mutations account for less than 10% of all LGMD.

LGMD2C-2F

Loss-of-function mutations in any of the genes encoding the four members of the skeletal muscle sarcoglycan complex, alpha, beta, gamma and delta-sarcoglycan cause the LGMD2D, 2E, 2C and 2F, respectively (14-19).
The alpha-sarcoglycan gene spans 10kb of genomic sequence at chromosome 17q21.33 and the major transcript is composed of 10 exons. The protein product of 387 amino acids and 50kDa was originally named adhalin and contains a “dystroglycan-type” cadherin-like domain that is present in metazoan dystroglycans (48).

The beta-sarcoglycan gene spans 15kb of genomic sequence at chromosome 4q11 and the major transcript is composed of 6 exons. The protein is of 318 amino acids and weighs 43kDa.

The gamma-sarcoglycan gene spans 144kb of genomic sequence at chromosome 13q12.12 and the major transcript is composed of 8 exons. The delta-sarcoglycan is by far the largest LGMD gene, spanning 433kb of genomic sequence at chromosome 5q33.3 and the major transcript is composed of 9 exons. Intron 2 alone spans 164kb, being one the largest one of the human genome. Delta and gamma sarcoglycan are homologous and of identical size (35kDa).

Furthermore, two additional members of the sarcoglycan were identified: epsilon (49) and zeta-sarcoglycan (50). Considering that sarcoglycan complex is biochemically a heterotetramer, these epsilon and zeta sarcoglycan can function in different tissues by replacing alpha and gamma or delta sarcoglycan, respectively (figure 1). Recently, epsilon-sarcoglycan mutations were associated with myoclonus dystonia (51). Sarcoglycans are components of the dystrophin/utrophin-complexes. They are all N-glycosylated transmembrane proteins with a short intra-cellular domain, a single transmembrane region and a large extra-cellular domain containing a cluster of conserved cysteines. Their role is performed in close association with the dystrophin complex and “sarcoglycanopathies” are clinically very similar to DMD/BMD, with calf hypertrophy and frequent heart involvement. The heart damage is also common in DMD/BMD patients and carriers and in the various strains of cardiomyopathic hamsters BIO14.6 as a consequence of a promoter deletion of the delta-sarcoglycan gene, but the importance of their requirement, demonstrated by the severity of disease is not yet understood (52). The absence of the sarcoglycan complex may cause a disorganization of the force-transmitting costameric cytoskeleton and disruption of sarcolemmal membrane integrity in muscle. The sarcoglycan could be critical for myofibril organization, in view of the cadherin-like domain of alpha and epsilon sarcoglycan.

Alltogether these forms are defined “sarcoglycanopathies” and account for <10% of all LGMD cases. We identified gamma>alpha>delta beta sarcoglycan mutations in Italian patients. The incidence of sarcoglycanopathies is higher among severe LGMD often with intrafamilial variability (53). The molecular diagnosis is addressed by protein analysis and only a few families with sarcoglycan deficiency have no mutations in any of the genes.

**LGMD2G**

The second category of mechanisms includes mutations affecting sarcomere. Myotilin mutations cause LGMD1A (4), Titin-cap (telethonin) mutations cause LGMD2G (20), while titin defects are claimed to cause LGMD2J and distal myopathy (23). The telethonin gene (TCAP) spans 1.2kb of genomic sequence at chromosome 17q12 and the transcript is composed of 2 exons. The protein product is a 19kDa protein TCAP encodes a protein found in striated and cardiac muscle that binds to the titin Z1-Z2 domains and is a substrate of titin kinase, interactions thought to be critical to sarcomere assembly (55). Only two different mutations have been described in the TCAP gene in Brazilian patients. A mutation (R87Q) was found in a patient with dilated cardiomyopathy resulting in a marked defect in T-cap interaction/localization. Further, a human muscle LIM protein (MLP) mutation (W4R) associated with dilated cardiomyopathy (DCM) results in a marked defect in Telethonin interaction/localization (56).
LGMD2H

The Tripartite-motif-containing gene 32 (TRIM32) gene spans 14kb of genomic sequence at chromosome 9q33.1 and the transcript is composed of 2 exons, with the first noncoding and the second that encode a 673 aa protein of 72kDa. The TRIM genes are a large family of proteins that share a common function: by means of homo-multimerization in specific cell compartments (57). The TRIM motif includes three zinc-binding domains, a RING, a B-box type 1 and a B-box type 2, and a coiled-coil region. The protein localizes to cytoplasmic bodies. The protein has also been localized to the nucleus, where it interacts with the activation domain of the HIV-1 Tat protein. The Tat protein activates transcription of HIV-1 genes (58).

LGMD2H is very rare. No other group has identified additional cases outside the Manitoba Hutterite group. A homozygous missense mutations Asp487Asn of the TRIM32 gene cause LGMD2H in 41 patients from Hutterite families (21). Although the function of TRIM32 is unknown, analysis of the domain structure of this protein suggests that it may be an E3-ubiquitin ligase. Ubiquitin ligases mediate the transfer of a small molecule called ubiquitin to the target protein, a process called “ubiquitylation”. This is a key process to limit the abundance of selected proteins in cells, by selectively marking unwanted or damaged proteins for degradation within cells. When unwanted proteins are not properly removed, abnormalities in cellular functions may occur and cause diseases such as cancer and inflammation. There are three classes of ubiquitin ligases (E1, E2 and E3), of which the E3 ubiquitin ligases have the most substrate specificity. If proven, this represents an additional pathogenic mechanism leading to muscular dystrophy.

LGMD2I

The extracellular part of the dystrophin/utrophin-associated complex is also involved in congenital muscular dystrophies, as well as in LGMD2I. Fukuyama-type congenital muscular dystrophy (FCMD), is one of the most common autosomal recessive disorders in Japan characterized by congenital muscular dystrophy associated with brain malformation (micropolygria) due to a defect in the migration of neurons caused by mutations in the fukutin gene at 9q31 (59). Mutations in the fukutin-related protein gene (FKRP) at 19q13 cause a form of congenital muscular dystrophy with secondary laminin alpha2 deficiency and abnormal glycosylation of alpha-dystroglycan (60). The same gene is also involved in LGMD2I (22). The fukutin-related protein gene spans 12kb of genomic sequence at chromosome 19q13.32 and the transcript is composed of 4 exons, with the first three are noncoding.

All of these diseases are associated with changes in alpha-dystroglycan expression due to a glycosylation defect of alpha-dystroglycan. Dystroglycan is normally expressed and recognized by polyclonal antibodies, but it is normally glycosylated and not recognized by monoclonal antibodies directed against certain epitopes. FKRP is resident in the Golgi apparatus. The P448L mutation that result in the CMD1C cause a complete mislocalization of the protein and the alpha-dystroglycan is not processed, while LGMD2I mutations affect the putative active site of the protein or cause inefficient Golgi localization (61).

LGMD2I mutations appear to be a relatively common cause of LGMD, accounting for at least 10% of all LGMD with either severe or mild phenotypes (62,63).

LGMD2J

TTN is one of the most complex human genes. The titin gene spans 294,442 bp of genomic sequence at chromosome 2q31 and the major transcript is composed of 363 exons that encode the largest protein of the human genome composed of 38,138 amino acids with a physical length of 2 microns. A 11-bp indel mutation in the last titin exon cause tibial muscular dystrophy and Gerull et al. (64) showed that a 2-bp insertion in exon 326 of the TTN gene cause autosomal dominant dilated cardiomyopathy (CMD1G; 604145). Titin is a good candidate for LGMD2J and recent data support the direct causative role of titin mutations.

Conclusions

LGMD present a higher degree of interfamilial and intrafamilial variability of severity than DMD/BMD. Among the ten AR LGMD genes so far identified, five address to the dystrophin/utrophin-complexes, two point to the sarcomere, while other three could disclose alternative mechanisms of damage.
In families with alpha (LGMD2D) and gamma sarcoglycan (LGMD2C) mutations, there are marked differences with patients who are wheelchair-bound and age-matched sibs with moderate weakness. In the different cardiomyopathic hamster models the genetic background has a strong influence on the nature of the cardiomyopathy. In addition, diet and certain drugs can improve the life expectancy of these animals. Even more striking differences are observed with dysferlin defects. LGMD2B and Miyoshi myopathy are distinct disorders that affect proximal and distal muscles, respectively. These two forms have been also found in the same families with identical dysferlin gene mutations. In addition, in a LGMD2f family we identified a non affected woman that carries a pathogenic homozygous FKRP mutation. These facts immediately indicate that the penetrance and expressivity of LGMD alleles may be influenced by additional factors (i.e. life style) or yet unknown variations in other genes (multiallelic inheritance). These data may have an impact on prognosis as well as on differentiated treatment of LGMD.

There are at least 25% of families who are unlinked to any known locus and 40% of isolated cases with severe or intermediate LGMD phenotype with no mutation in any known gene (unpublished results). This figure may be overestimated, since a number of cases may be misdiagnosed. But considering that the percentage of these patients is very similar independently from the centre of first diagnosis, additional LGMD genes are expected. Considering that recessive mutations generally implicate loss-of-function of distinct components of the same machinery, while dominant mutations are generally more heterogeneous, the question arises whether these further recessive LGMD defects produce common molecular mechanisms of muscle wasting.

The second question is whether there is a multitude of genes each accounting for a limited number of cases, or there are one or two major disease genes that cause all these forms. We consider the first possibility more realistic and similar to the widespread genetic heterogeneity observed in other human diseases (i.e. retinitis pigmentosa or mental retardation). This would imply that if there are 5-50 yet unknown LGMD genes, each one will be responsible for a very small fraction of cases. To have some chances to identify these genes, the starting point is the identification of a pool of best candidate genes among many thousands of muscle transcripts (65-69). To this purpose, huge collection of DNA samples and cost-effective high-throughput screening methods are needed.

Acknowledgements
This study was supported by grants from Telethon (TIGEM-11 and TIGEM-C20), Ministero dell’Istruzione dell’Università e della Ricerca (MIUR: PRIN 2001 and PRIN 2002), Ministero della Salute (d.lgs 502/92).

References
Molecular bases of autosomal recessive limb-girdle muscular dystrophies


25. Human Genome Browser Gateway at http://genome.ucsc.edu/cgi-bin/hgGateway?org=human

26. Human Genome Browser Gateway at http://www.dmd.nl


30. McNally EM, Ly CT, Kunkel LM. Human epsilon-sarcoglycan is highly related to alpha-sarcoglycan (adhalin),