Limb-Girdle Muscular Dystrophy Type 2H Associated with Mutation in *TRIM32,* a Putative E3-Ubiquitin–Ligase Gene

Patrick Frosk,[1] Tracey Weiler,[1] Edward Nylen,[1] Thangirala Sudha,[1] Cheryl R. Greenberg,[1,2] Kenneth Morgan,[3] T. Mary Fujiwara,[3] and Klaus Wrogemann[1,2]

[1]Departments of Biochemistry and Medical Genetics and [2]Department of Pediatrics and Child Health, University of Manitoba, Winnipeg; and [3]Departments of Human Genetics and Medicine, McGill University and Research Institute of the McGill University Health Centre, Montreal

Limb-girdle muscular dystrophy type 2H (LGMD2H) is a mild autosomal recessive myopathy that was first described in the Manitoba Hutterite population. Previous studies in our laboratory mapped the causative gene for this disease to a 6.5-Mb region in chromosomal region 9q31-33, flanked by *D9S302* and *D9S1850.* We have now used additional families and a panel of 26 microsatellite markers to construct haplotypes. Twelve recombination events that reduced the size of the candidate region to 560 kb were identified or inferred. This region is flanked by *D9S1126* and *D9S737* and contains at least four genes. Exons of these genes were sequenced in one affected individual, and four sequence variations were identified. The families included in our study and 100 control individuals were tested for these variations. On the basis of our results, the mutation in the tripartite-motif–containing gene (*TRIM32*) that replaces aspartate with asparagine at position 487 appears to be the causative mutation of LGMD2H. All affected individuals were found to be homozygous for D487N, and this mutation was not found in any of the controls. This mutation occurs in an NHL (named after the proteins <u>N</u>CL1, <u>H</u>T2A, and <u>L</u>IN-41) domain at a position that is highly conserved. NHL domains are known to be involved in protein-protein interactions. Although the function of TRIM32 is unknown, current knowledge of the domain structure of this protein suggests that it may be an E3-ubiquitin ligase. If proven, this represents a new pathogenic mechanism leading to muscular dystrophy.

Introduction

The limb-girdle muscular dystrophies (LGMDs) are a heterogeneous group of primary myopathies with both autosomal dominant and autosomal recessive modes of inheritance (Bushby 1999*a;* Kaplan 2001). Recessively inherited forms LGMD2A–LGMD2I map to nine loci. The two most common types are LGMD2A and LGMD2B, which are caused by mutations in the genes encoding calpain 3 and dysferlin, respectively. LGMD2C–LGMD2F are caused by mutations in members of the sarcoglycan complex, and LGMD2G is caused by mutations in the gene for the sarcomeric protein telethonin. The genes responsible for LGMD2H and LGMD2I have not yet been identified.

LGMD2H (MIM 254110) is a mild form of autosomal recessive muscular dystrophy, with a variable clinical presentation (Shokeir and Kobrinsky 1976; Weiler et al.

1997; Bushby 1999*b*). Onset is usually within the 2d or 3d decade of life, and progression is slow. Most patients have remained ambulatory into the 6th decade of life. There is no evidence of cardiac or facial involvement. To date, this disease has been described only in the Hutterite population of North America (Shokeir and Kobrinsky 1976; Shokeir and Rozdilsky 1985; Weiler et al. 1997).

Our group recently mapped the locus for LGMD2H to an interval, flanked by *D9S302* and *D9S1850,* in chromosomal region 9q31-33 (Weiler et al. 1998), which is now estimated to be 6.5 Mb (University of California–Santa Cruz [UCSC] Human Genome Project Working Draft). Herein, we describe the fine mapping and characterization of the LGMD2H candidate region. This was achieved by haplotype analysis and comparison of sequences from an affected individual and sequences generated by either the Human Genome Project (Lander et al. 2001) or Celera Genomics (Celera Publication Site; Venter et al. 2001). We sequenced genes in the candidate region for pathogenic mutations and identified *TRIM32* as the most likely gene for LGMD2H. TRIM32 is a member of the tripartite-motif (TRIM) family of proteins described elsewhere (Reymond et al. 2001). Current views of the domains that are characteristic of this protein family suggest that TRIM32 may be involved in the ubiquitin (Ub)–proteasome pathway (Joazeiro and

Weissman 2000). This pathway has not previously been implicated in muscular dystrophy and therefore represents a novel pathogenic mechanism.

Subjects and Methods

Subjects

We studied 35 clinically affected Hutterite patients and members of their immediate families. Signed informed consent was obtained from each individual, and the study was approved by the University of Manitoba research-ethics board. Genotypes were obtained for 117 individuals, who constitute six interrelated families (two to four generations) with 2–12 family members affected with LGMD2H. Samples of DNA from 100 individuals not known to have LGMD or to be related to the Hutterites were used as controls; no demographic or ethnic information for these individuals was available to us. Eight Centre d'Étude du Polymorphisme Humain (CEPH) families with recombinations in chromosomal region 9q3 (S. Povey, personal communication) were used to help define marker order; CEPH individuals 1331-01, 1331-02, 1347-01, and 1347-02 were used to determine the size of microsatellite alleles (Fondation Jean Dausset–CEPH).

DNA Analysis

Genotypes were obtained as described elsewhere (Weiler et al. 1998). Haplotypes were constructed manually, by minimizing the number of recombinants and assuming no mutation of marker alleles. The candidate region was defined on the basis of observed or inferred ancestral recombination events.

Oligonucleotide primers (see table S1, available online only) were designed to amplify exons of the candidate genes from genomic DNA, for single-stranded conformational analysis (SSCA) and/or sequencing. For most exons of >350 bp, overlapping primers were designed. SSCA was performed on PCR products from three individuals: one patient and two siblings (one of whom has one disease-associated chromosome and the other of whom has no disease-associated chromosomes). The analysis was performed under three conditions: (*a*) 6% acrylamide (19:1) with 10% glycerol at 20°C, with 8 mA, for 20 h; (*b*) 6% acrylamide (19:1) with 20% glycerol at 4°C, with 10 W, for 20 h; and (*c*) MDE gel per manufacturer's instructions (FMC BioProducts). PCR products to be sequenced were separated by electrophoresis in 1% agarose gels. DNA fragments were purified from gel slices by GenElute (Sigma-Aldrich). The fragments were then sequenced, with the same primers used to generate the PCR products, on an ABI 377 DNA sequencer (Applied Biosystems) at the Centre for Applied Genomics, the Hospital for Sick Children, Toronto. Pu-

tative mutations were identified by comparing the sequence obtained from the patient and the published sequences (GenBank; Celera Publication Site). DNA from an unaffected sibling was then sequenced to confirm the existence of two different alleles.

PCR assays based on either length variation or restriction-enzyme digestion were developed to screen the variations that were identified by DNA sequencing. The six Hutterite families included in this study and 100 control individuals were tested for the presence of each variation. All fragments were subjected to electrophoresis in an 8% acrylamide gel and were visualized by ethidium bromide staining unless otherwise specified. We named the mutations according to the recommendations of the Nomenclature Working Group (Antonarakis and Nomenclature Working Group 1998). The four variations were studied as follows (for description of each sequence variation, see table 1):

V1. Genomic DNA was amplified using primers PAPPA_ex.3F (5′-GCTCTAAATTATTTGGAGAGG-3′) and PAPPA_ex.3R (5′-GCACTAAGTGTATTACTG-TCC-3′). The 11-bp deletion was readily visible as a shift in fragment size (from 291 to 280 bp) in 12% polyacrylamide gels.

V2. Genomic DNA was amplified using primers designed with a mismatch (denoted by a lowercase letter), P14_BsmAIF (5′-GCTGGCTGTGGAGAATGCgT-3′) and P14_BsmAIR (5′-AGCACCTGCACATACCT-GtC-3′). P14_BsmAIR creates a *Bsm*AI restriction site in all fragments, whereas P14_BsmAIF creates a *Bsm*AI site only in the V2(Ser) allele. Restriction-enzyme digestion of the V2(Ser) allele by *Bsm*AI yields three fragments (94, 24, and 27 bp), whereas the V2(Tyr) allele has two fragments (118 and 27 bp).

V3. Genomic DNA was amplified using primers PAPPA_ex.18F (5′-CTCTCTTGGTCCTAACTCTG-3′) and PAPPA_ex.18R (5′-GAATTGCCCCTCCTACTGC-3′). Restriction-enzyme digestion of the V3(C) allele yields two fragments (162 and 63 bp), whereas the V3(T) allele migrates as a single fragment (225 bp).

V4. Genomic DNA was amplified using primers HT2A_ex.2(5)F (5′-GGTGTGACTGACAGCTATG-3′) and H2_HpyR (5′-AGCTTTCCACCTTCCACgT-3′). H2_HpyR was designed with a mismatch (denoted by a lowercase letter) that creates an *Hpy*99I restriction site only in the V4(Asp) allele. Digestion of the V4(Asp) allele yields two fragments (136 and 16 bp), whereas the V4(Asn) allele migrates as a single fragment (152 bp).

Expression Analysis

A 12-tissue northern blot obtained from Origene Technologies was used to determine the normal distribution of the *TRIM32* transcript. A *TRIM32* probe was

Table 1

Four Sequence Variations Found in Exons within the *LGMD2H* Candidate Region

MUTATION	GENE (LOCATION)	SEQUENCE VARIATION(S)[a]	VARIATION TYPE(S)	CONTROL INDIVIDUALS	
				Genotype[b]	No.
V1	*PAPPA* (intron 2)	c.(1479-59)–(1479-49)del	Deletion	+/+	27
				+/−	50
				−/−	23
V2	*PAPPA* (exon 14)	c.3671C→A, p.S1224Y	Transversion, missense mutation	Ser/Ser	8
				Ser/Tyr	33
				Tyr/Tyr	59
V3	*PAPPA* (exon 18)	c.4374C→T, p.C1458C	Transition, silent mutation	C/C	96
				C/T	4
				T/T	0
V4	*TRIM32* (exon 2)	c.1459G→A, p.D487N	Transition, missense mutation	Asp/Asp	100
				Asp/Asn	0
				Asn/Asn	0

[a] Numbering of nucleotides and amino acids is based on GenBank (accession number NM_002581.1 for *PAPPA* and accession number NM_012210.2 for *TRIM32*); the A of the initiator Met codon is considered to be nucleotide 1.

[b] For V1, the presence or deletion of 11 bp is denoted by + or −, respectively; for V2 and V4, alleles are denoted by amino acid codes; and, for V3, alleles are denoted by nucleotides.

prepared by amplifying a 733-bp fragment of the gene by use of the primers HT2A_ex.2(5)F (5′-GGTGTGAC-TGACAGCTATG-3′) and HT2A_ex.2(6)R (5′-CATAG-GCTGAGTCTATTCTGC-3′) from the genomic DNA of an unaffected individual. The fragment was subjected to electrophoresis in a 1% agarose gel, was purified as described in the "DNA Analysis" subsection, and was then labeled using the Random Primers Labeling kit (Invitrogen). A product with a specific activity of 3×10^9 cpm/μg was obtained. The blot was hybridized using Ultra-hyb hybridization solution (Ambion) according to the manufacturer's instructions. The hybridized membrane was visualized by autoradiography with Biomax MS film (Kodak). Film exposure was performed at −80°C, with intensifying screens, for 100 h.

Reverse-transcriptase PCR (RT-PCR) was performed on RNA that was isolated from cultured skin fibroblasts obtained from both an affected individual and an unaffected individual. RNA was isolated using an Oligotex Direct mRNA Isolation kit (Qiagen), and first-strand synthesis was performed using the Expand RT system (Roche) according to the manufacturers' instructions. Primers used for amplification of the *PAPPA* gene were PAPPA16/17 (5′-AACGGGCCTTCAAGAC-TCAG-3′) and PAPPA19/20 (5′-ACAACTCTCTCTAC-CCGTG-3′).

Results

Haplotype Analysis

Haplotype analysis was performed both to define the order of markers used and to fine map the candidate region. Genotypes were obtained for the population that we studied and for selected CEPH individuals, and then haplotypes were constructed (complete data not shown). Three recombination events were identified in individuals C-06, C-07, and E-11, thereby reducing the candidate region from 6.5 Mb to an interval of 713 kb flanked by *D9S241* and *D9S931* (fig. 1). The region was further reduced to a 560-kb region flanked by *D9S1126* and *D9S737,* on the basis of two inferred ancestral recombination events (in individuals B-14 and E-23; fig. 1).

Physical Mapping

Our early physical-mapping efforts used a combination of Human Genome Project sequences and radiation-hybrid mapping. Subsequent efforts focused on contiguous sequence obtained from the Celera Publication Site (GA_x54KRCDB2PK: 3500001–4500000) (see fig. 2). Four genes are present in the region, which spans 560 kb of DNA:

1. *PAPPA* (GenBank accession number NM_002581.1) encodes pregnancy-associated plasma protein A (PAPPA, also known as "insulin-like growth factor [IGF]–binding protein 4 [IGFBP4] protease"). A high concentration of this protein is found in the serum of pregnant women (Silahtaroglu et al. 1993; Conover et al. 1999; Lawrence et al. 1999). The IGFBP4-protease substrate, IGFBP4, is known to be a potent inhibitor of the IGF-signaling system, and PAPPA appears to prevent this inhibition.

2. *Astrotactin 2* (*ASTN2* [GenBank accession number AF116574.1]) encodes a protein highly paralogous to *ASTN1,* which is a mediator of neuronal-glial contact (Edmondson et al. 1988).
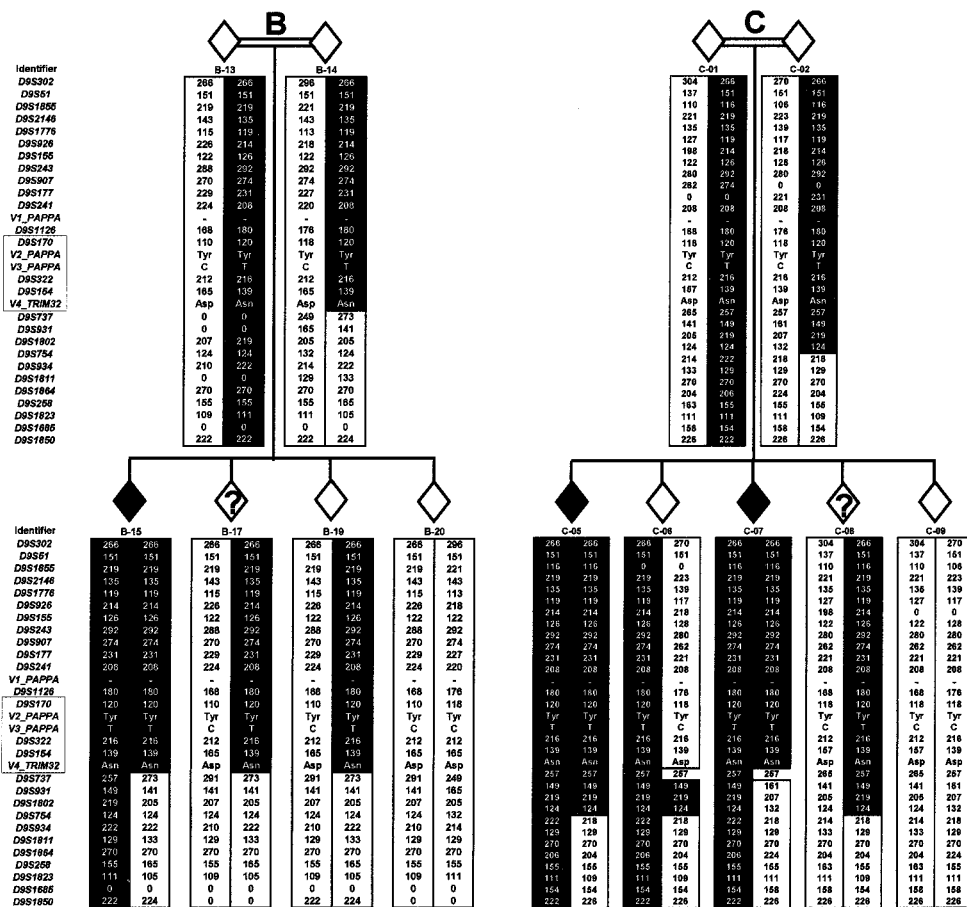
Figure 1 Chromosomal region 9q31-33 haplotypes, showing recombinations (in individuals C-06, C-07, and E-11) that reduce the 6.5-Mb *LGMD2H* candidate region, flanked by *D9S302* and *D9S1850,* to a 713-kb region, flanked by *D9S241* and *D9S931.* Inferred ancestral recombinants in individuals B-14 and E-23 further reduce the region to an interval of 560 kb, flanked by *D9S1126* and *D9S737* (*boxed regions in the columns of markers*). Individuals in family C (individuals C-01, C-02, and C-05–C-09) are represented as C-5, C-6, and C-18–C-22, respectively, by Weiler et al. (1998). Individuals in families B and E were not included in the study by Weiler et al. (1998). Blackened diamonds represent affected individuals, unblackened diamonds represent unaffected individuals, and unblackened diamonds with a question mark (?) represent individuals with unknown phenotype. Alleles are given (in bp), and "0" indicates that no data are available. Blackened backgrounds behind haplotypes indicate consensus disease-associated alleles, square brackets indicate inferred alleles, and the absence of an outer vertical border indicates that an allele is uninformative for a crossover. See table 1 for the alleles of V1–V4. Alleles of *D9S1850* were reevaluated for this study, and some are 2 bp larger than those reported by Weiler et al. (1998).

3. *TRIM32* (GenBank accession number NM_012210.2), previously known as "*HT2A*" (Fridell et al. 1995), encodes a member of the TRIM family of proteins. It was originally discovered using a yeast two-hybrid assay with the human immunodeficiency virus type 1 Tat acting as bait (Fridell et al. 1995; Reymond et al. 2001). The RING-finger domain suggests that this protein is an E3-Ub ligase.
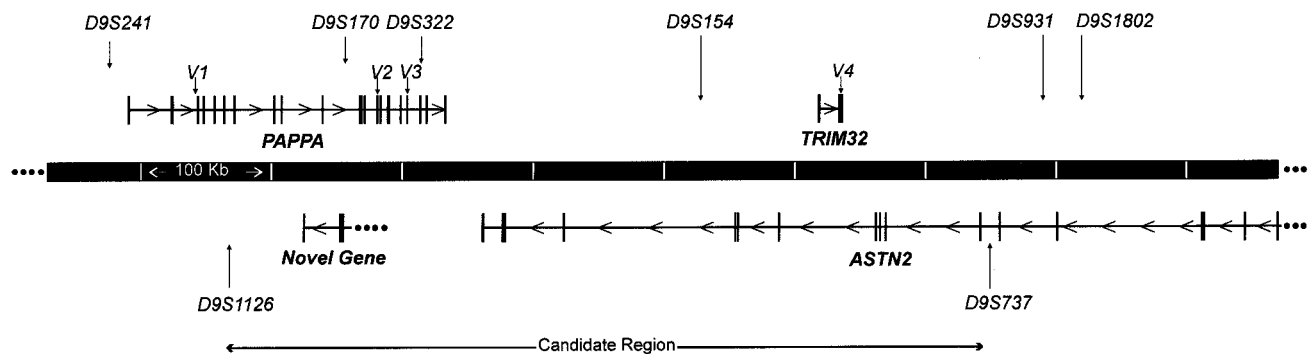
4. Sequencing of orphan expressed-sequence tag (EST) clones revealed a putative fourth gene in the region (data not shown). The 5' end of this gene is still undefined, but the 3' end consists of at least four exons found in the cDNA clone IMAGE:2591170 (GenBank acces-

sion number AW090658.2). Attempts to amplify and sequence the 5' end of this gene have been unsuccessful owing to the apparent rarity of the transcript in all tissues that we have studied, including muscle (data not shown).

Mutation Screening

All four aforementioned genes were considered as potential candidates for *LGMD2H,* on the basis of their chromosomal position. Both SSCA and DNA sequencing were used to search for disease-causing mutations. SSCA of exons located between the flanking microsatellite mark-

ers *D9S1126* and *D9S737* revealed only one silent substitution (i.e., V3) in *PAPPA*. DNA sequencing of the exons of the candidate gene shown in fig. 2 revealed three more sequence variations (V1, V2, and V4). We were unable to detect V2 and V4 by SSCA in our initial screen; however, subsequent amplification in a smaller-sized PCR product allowed visualization of both alleles of V2 and of V4, by SSCA. The four sequence variations—an 11-bp deletion and three point mutations—were identified on the basis of differences between sequences from an affected individual and published sequences (table 1). V1, an 11-bp deletion in intron 2 of *PAPPA*, can be excluded as the disease-causing mutation, because it does not segregate with *LGMD2H* and is outside the candidate region

Figure 2 Physical map of the *LGMD2H* candidate region, flanked by *D9S1126* and *D9S737*. The four sequence variations identified in samples from patients are labeled "V1"–"V4", as in table 1. The map shows 1 Mb and is based on contiguous sequences obtained from GenBank and the Celera Publication Site.

as defined by an inferred ancestral recombination event (in individual E-23; fig. 2). Both affected and unaffected family members were homozygous for the deletion, and it was common among controls (fig. 1 and table 1). V2, a missense mutation in codon 1224 of *PAPPA,* can also be excluded as a disease-causing mutation. All affecteds and all but one of the obligate carriers (i.e., individual E-02; fig. 1) are homozygous for the V2(Tyr) allele. Both alleles of V1 and of V2 are common in Hutterite families and controls.

V3 results in a silent substitution in codon 1458 of *PAPPA.* The V3(T) allele was found to have a frequency of 2% in controls. Since the V3 mutation does not change the amino acid sequence of *PAPPA,* it could be disease causing if it leads to alternate splicing or if the codon usage of the V3(T) allele is significantly different. However, codon usage does not appear to be a factor, since the two coding triplets are used at similar frequencies in humans; UGC accounts for 55% of all cysteine codons, and UGU accounts for 45% of all cysteine codons (Codon Usage Database). Alternative splicing is unlikely because *in silico* analysis of the splice consensus sequences shows no predicted change with the introduction of V3(T) (NIX package; U.K. Human Genome Mapping Project Resource Centre). RT-PCR was used to confirm this prediction by amplifying the region containing V3 (*PAPPA* exons 17–19), from both affected and unaffected RNA isolated from cultured skin fibroblasts. No differences between the affected and unaffected samples were apparent (data not shown).

V4 results in a nonconservative substitution of an acidic aspartate by a neutral asparagine at amino acid 487 of *TRIM32.* The V4(Asn) and V3(T) alleles are in complete linkage disequilibrium in the Hutterite families. The V4(Asn) allele was not observed in any of the controls. All 35 affected individuals in whom LGMD2H had been diagnosed are homozygous for V4(Asn) and V3(T). There were six asymptomatic individuals who

were also homozygous for these alleles and the disease-associated chromosome. Four of these individuals are <20 years of age, which is consistent with our previous findings that the onset of LGMD2H is age dependent (Weiler et al. 1998). The other two individuals are in their 30s and do not meet the strict criteria for affected individuals that have been established by Weiler et al. (1997).

The V4 amino acid substitution occurs within the third NHL (named after the proteins N̲CL1, H̲T2A [TRIM32], and L̲IN-41 [Slack and Ruvkun 1998]) domain of TRIM32. NHL domains are conserved motifs found in many prokaryotic and eukaryotic proteins. This missense mutation occurs at one of the two most highly conserved positions in this domain (table 2) (Slack and Ruvkun 1998).

Northern Blot Analysis

LGMD2H is a disease that is mainly restricted to skeletal muscle. Evidence from ESTs (UniGene accession number Hs.236218) suggests that the *TRIM32* transcript is expressed in a wide variety of tissues, including skeletal muscle. To confirm the presence of the transcript in skeletal muscle, we probed a multiple-tissue northern blot (Origene Technologies) with a 733-bp fragment obtained from the coding sequence of *TRIM32.* The results (shown in fig. 3) indicate that the transcript is indeed widely expressed and, in particular, is expressed in skeletal muscle.

Discussion

On the basis of observed recombinants, the candidate gene for LGMD2H was previously defined to be in a 3.9–6.2-cM interval flanked by *D9S302* and *D9S1850* (Weiler et al. 1998). The region could be reduced to a 0.3-cM interval flanked by *D9S302* and *D9S934* by as-

Table 2

**Alignment of Selected NHL Domains from Nine Eukaryotic Species**

	AMINO ACID	
SOURCE	Position	Sequence
Consensus sequence	1–27	FDYPRGVAVDSDGO-IVVADSE--N
Species and protein[a] (GenBank accession number):		
Homo sapiens:		
TRIM32 (NP_036342.1)	371–398	FNLPVSLYVTSQGE-VLVADRG--N
	428–455	NLTPLSVAMNCQGL-IGVTDSY--D
	469–496	LSKPWGITALPSGQ-FVVTDVE--G
	520–550	TCDAEGTVYFTQGLGLNL-ENRQNE
	575–602	FRCIAGMCVDARGD-LIVADSS--R
	616–643	LTCPVGIALTPKGQ-LLVLDCW--D
PAM (NP_000910.1)	683–711	FTVPHSLALVPLLGQLCVADRE--N
TRIM2 (NP_056086.1)	534–561	FTNLQGVAASTNGK-ILIADSN--N
TRIM3 (NP_006449.1)	533–560	LQRPTGVAVDTNGD-IIVADYD--N
Caenorhabditis elegans:		
NCL1 (AAC14263.1)	586–613	FTEPSGVAVNGQGD-IVVADTN--N
	633–662	LLYPNRVAVNRTTGDFVVTERS--P
	677–704	LQHPRGVCVDSKGR-IIVVECK--V
	720–747	LEFPNGVCTNDKNE-ILISDNR--A
LIN-41B (AAF15530.1)	763–791	TNYPIGVGINSLGE-VVVADNH-NN
	845–872	LCRPWGICVDQRGR-VIVADRS--N
	892–919	FDRPAGITTNSLNN-IVVADKD--N
	939–966	FNYPWGVATNSHNA-IAVSDTR--N
	987–1014	LDSPRGLCYLPDGQ-LLITDFN--N
	1035–1062	FVRPQGVVIDPEGH-ILVCDSR--N
	1120–1147	LDRPTDLAVGPDGR-IYVVDFG--N
YMB4 (CAA79562.1)	712–739	LNWPRGICALSGGL-VATCDSS--N
F26F4.6 (AAA91222.1)	761–788	LHCPSGFCLSDTDD-ILIADTN--N
F21F3.1 (AAB42278.1)	279–307	FVVPHSLSLIEDMNIVCVADRE--N
PAM (AAB37637.1)	424–451	FYLPHGIYVDKDGF-VYTTDVG--S
Mus musculus:		
Pam (NP_038654.1)	581–608	FYLPHGLSIDTDGN-YWVTDVA--L
Trim3 (BAA83343.1)	713–740	LYGPQGLALTSDGH-VVVADAG--N
Rattus norvegicus:		
Pam (NP_037132.1)	581–608	FYLPHGLSIDTDGN-YWVTDVA--L
Trim3 (AAC17997.1)	486–513	FTNLHPLSAASSGR-IVVADSN--N
Bos taurus PAM (AAA30683.1)	779–806	FDMPHDIAASEDGT-VYVGDAH--T
Equus caballus PAM (BAA06104.1)	682–710	FRVPHSLALVPHLGQLCVADRE--N
Xenopus laevis XELCAM (AAA49667.1)	582–610	FRIPHSLTMISDQGQLCVADRE--N
Drosophila melanogaster:		
brat (AAF53771.1)	780–807	FTEPSGVAVNAQND-IIVADTN--N
dpld (AAF59245.1)	537–564	VSRPWGLCVDKMGH-VLVSDRR--N
cg12130 (AAF58870.1)	387–415	FQNPHDVAVTADGNEIYVAELN--P
Lymnaea stagnalis LPAM (AAD42258.1)	1578–1605	YFMPHGIEVDNQGN-LWLTDVA--L

   [a] Each protein has up to six NHL domains (InterPro accession number IPR001258); however, only one domain from each protein is shown, except for the three after which the NHL domain was named (i.e., NCL1, HT2A [TRIM32], and LIN-41). There are six domains from TRIM32, and the mutation (i.e., D487N) is in the third NHL domain, at position 20, denoted by the underlined letter in each sequence. This aspartate is one of the most highly conserved amino acids among the NHL domains. There are currently 110 eukaryotic NHL domains in the InterPro database; of these, at position 20, 100 have an aspartate, 9 have a glutamate, and 1 has a lysine.

suming that an ancestral recombination had occurred (Weiler et al. 1998). This interval is now estimated to be 4.1 Mb (UCSC Human Genome Project Working Draft). Through the study of additional families with LGMD2H and the use of 22 additional microsatellite markers, we have reduced the candidate region to a 560-kb interval flanked by *D9S1126* and *D9S737,* on the

basis of haplotype analysis. The region contains three annotated genes—*PAPPA, ASTN2,* and *TRIM32*—and one unannotated gene. *TRIM32* is located within intron 12 of *ASTN2,* in the opposite orientation, and part of the unannotated gene is within introns 9 and 10 of *PAPPA,* in the opposite orientation (fig. 2).

All exons within the candidate region were screened
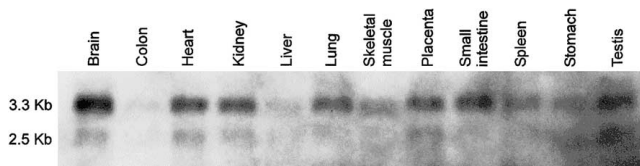
Figure 3   Northern blot analysis using a fragment of the *TRIM32* coding sequence. *TRIM32* shows two transcript sizes, 3.3 and 2.5 kb. Expression is evident in skeletal muscle, but levels appear much higher in brain and heart tissue.

for mutations. Of the four sequence variations detected, V1 and V2 could readily be excluded as disease causing because they do not consistently segregate with disease. V3(T) appears to be an unlikely candidate because it leads neither to an amino acid change, to a codon with significantly altered usage, nor, apparently, to alternative splicing. The latter was confirmed by RT-PCR of mRNA isolated from skin fibroblasts that were obtained from one patient and from a control. However, at this time, we cannot exclude the possibility that the V3(T) allele either disrupts an exonic splicing enhancer that functions only in muscle (Liu et al. 1998, 2001) or alters the secondary structure of the mRNA (Shen et al. 1999).

The best candidate mutation for LGMD2H is the missense mutation V4, in *TRIM32,* which changes codon 487 of TRIM32 from aspartate to asparagine (D487N). There is strong evidence that this D487N mutation causes LGMD2H, because (1) all affected individuals in our study are homozygous for V4(Asn); (2) the D487N mutation was not observed in 100 controls; (3) the D487N mutation occurs in a protein domain that is considered to be an NHL domain (Slack and Ruvkun 1998) and is found in many proteins, and the aspartate in question is one of the two most conserved amino acids in NHL domains (table 2); and (4) the mutation from aspartate to asparagine is a nonconservative change that could result in a significant alteration in the structure and function of TRIM32. Thus, we consider D487N to be the disease-causing mutation for LGMD2H in the patients whom we studied. The information that we currently possess does not exclude the possibility that V3 may contribute additional effects to the phenotype.

TRIM32 has, from N- to C-terminal ends, a RING-finger domain, a B1 box, a coiled-coil domain, and six NHL domains (InterPro accession numbers IPR001841, IPR002991, IPR003649, and IPR001258, respectively) (Apweiler et al. 2001). TRIM32 is a member of a growing family of TRIM, or RING–B-box–coiled-coil, proteins (Reymond et al. 2001; Torok and Etkin 2001). Current knowledge of the RING-finger domain indicates that it may have intrinsic Ub-ligase activity (Barinaga 1999), and it has been suggested that all RING-

finger proteins are E3-Ub ligases (Freemont 2000). This information suggests that TRIM32 is an E3-Ub ligase that functions in the Ub-proteasome pathway.

The Ub-proteasome pathway is a specialized mechanism for the posttranslational regulation of protein levels. The initial step in the Ub-proteasome pathway is the activation of Ub by an E1-Ub–activating enzyme. The activated Ub is transferred to the E2-Ub–conjugating enzyme. The E3 ligase interacts with E2-Ub complex and recruits a protein to be targeted to the 26S proteasome. The E3 ligase catalyzes the transfer of Ub to the target protein. The ubiquitinated protein is released from the complex and is recognized by factors associated with the 26S proteasome. The 26S proteasome digests the ubiquitinated protein into small peptides. The E3 enzymes are responsible for the selection of proteins for ubiquitination by this pathway, and it is thought that a large number of these enzymes exist (Freemont 2000). It has been shown that the RING-finger domain of an E3 ligase mediates interaction with the E2-conjugating enzyme and is the source of the ligase activity (Joazeiro and Weissman 2000). This suggests that recognition of target proteins by E3 ligases occurs through protein-protein interactions with areas of the protein other than the RING-finger domain. In the case of TRIM32, this function may be fulfilled by a series of NHL domains.

NHL domains have been shown to be important for the function of a number of proteins—including LIN-41, in *Caenorhabditis elegans,* and brat, in *Drosophila melanogaster,* both of which act as posttranscriptional repressors. Of the seven apparent loss-of-function *lin-41* mutations that are discussed by Slack et al. (2000), five lead to amino acid substitutions found in the NHL domains (Slack et al. 2000). Loss-of-function *brat* mutations include both deletions of the NHL domains and missense mutations (Arama et al. 2000; Sonoda and Wharton 2001). All of these mutations abolish an interaction with another protein, thereby leading to functional consequences.

This information suggests interesting possibilities as to the role of TRIM32 in LGMD. If TRIM32 is indeed an E3-Ub ligase, then the NHL domains could mediate interaction either with other members of an E3-Ub–ligase complex or with the proteins that the complex targets to the proteasome. Because of both the limited distribution of symptoms in patients with LGMD2H and the wide tissue distribution of TRIM32, it is tantalizing to postulate that the third NHL repeat is involved in the interaction with—and, therefore, the targeting of—one or a group of proteins to the proteasome to be degraded. A mutation in the NHL domain would prevent this process, and these target proteins would accumulate to higher concentrations. The restriction of symptoms to muscle tissue may have two explanations: (1) the third

NHL domain may be interacting with muscle-specific proteins; and (2) although the increased level of proteins is not restricted to muscle, it may have a toxic effect only in muscle. Increased levels of some proteins have been reported to cause myopathic symptoms, as has been shown in work involving transgenic mouse lines that overexpress either caveolin-3 (Galbiati et al. 2000) or γ-sarcoglycan (Zhu et al. 2001). Furthermore, the gene mutated in rare forms of Parkinson disease, *PARK2,* is thought to work via a similar pathogenic mechanism. Parkin, the product of *PARK2,* is known to be a RING-finger–containing E3-Ub ligase (Shimura et al. 2000). The neuronal death seen in Parkinson disease is thought to be a direct result of the accumulation of proteins that are no longer ubiquitinated by mutant forms of parkin and that, therefore, are not degraded by the proteasome (Imai et al. 2001; Shimura et al. 2001). Studies to determine whether TRIM32 is indeed an E3-Ub ligase and to identify a similar pathogenic mechanism in parkin are currently under way.

Acknowledgments

Electronic-Database Information

Accession numbers and URLs for data in this article are as follows:

Celera Publication Site, http://public.celera.com/cds/login .cfm (for Genomic Scaffold GA_x54KRCDB2PK: 3500001–4500000)
Codon Usage Database, http://www.kazusa.or.jp/codon/
Fondation Jean Dausset–CEPH, http://www.cephb.fr/
GenBank, http://www.ncbi.nlm.nih.gov/Genbank/ (for *PAPPA* cDNA [accession number NM_002581.1], *PAPPA* genomic DNA [accession number AB020878], *ASTN2* cDNA [accession number AF116574.1], *TRIM32* cDNA [accession number NM_012210.2], and IMAGE:2591170 [accession number AW090658.2])
InterPro, http://www.ebi.ac.uk/interpro/ (for RING-finger domain [accession number IPR001841], B1 box [accession number IPR002991], coiled-coil domain [accession number IPR003649], and NHL domains [accession number IPR001258])

Online Mendelian Inheritance in Man (OMIM), http://www .ncbi.nlm.nih.gov/Omim/ (for LGMD2H [MIM 254110])
UCSC Human Genome Project Working Draft ("Golden Path"), http://genome.ucsc.edu/ (for April 2001 freeze)
UK Human Genome Mapping Project Resource Centre, http://www.hgmp.mrc.ac.uk/ (for NIX package, which includes Grail, Genefinder, Genemark, Fex, Hexon, and Fgene)
UniGene, http://www.ncbi.nlm.nih.gov/UniGene/ (for TRIM32 [accession number Hs.236218])

References

Antonarakis SE, Nomenclature Working Group (1998) Recommendations for a nomenclature system for human gene mutations. Hum Mutat 11:1–3
Apweiler R, Attwood TK, Bairoch A, Bateman A, Birney E, Biswas M, Bucher P, et al (2001) The InterPro database, an integrated documentation resource for protein families, domains and functional sites. Nucleic Acids Res 29:37–40
Arama E, Dickman D, Kimchie Z, Shearn A, Lev Z (2000) Mutations in the β-propeller domain of the *Drosophila brain tumor* (*brat*) protein induce neoplasm in the larval brain. Oncogene 19:3706–3716
Barinaga M (1999) A new finger on the protein destruction button. Science 286:223–225
Bushby KM (1999*a*) The limb-girdle muscular dystrophies: multiple genes, multiple mechanisms. Hum Mol Genet 8: 1875–1882
——— (1999*b*) Making sense of the limb-girdle muscular dystrophies. Brain 122:1403–1420
Conover CA, Oxvig C, Overgaard MT, Christiansen M, Giudice LC (1999) Evidence that the insulin-like growth factor binding protein-4 protease in human ovarian follicular fluid is pregnancy associated plasma protein-A. J Clin Endocrinol Metab 84:4742–4745
Edmondson JC, Liem RK, Kuster JE, Hatten ME (1988) Astrotactin: a novel neuronal cell surface antigen that mediates neuron-astroglial interactions in cerebellar microcultures. J Cell Biol 106:505–517
Freemont PS (2000) RING for destruction? Curr Biol 10: R84–R87
Fridell RA, Harding LS, Bogerd HP, Cullen BR (1995) Identification of a novel human zinc finger protein that specifically interacts with the activation domain of lentiviral Tat proteins. Virology 209:347–357
Galbiati F, Volonte D, Chu JB, Li M, Fine SW, Fu M, Bermudez J, Pedemonte M, Weidenheim KM, Pestell RG, Minetti C, Lisanti MP (2000) Transgenic overexpression of caveolin-3 in skeletal muscle fibers induces a Duchenne-like muscular dystrophy phenotype. Proc Natl Acad Sci USA 97:9689–9694
Imai Y, Soda M, Inoue H, Hattori N, Mizuno Y, Takahashi R (2001) An unfolded putative transmembrane polypeptide, which can lead to endoplasmic reticulum stress, is a substrate of parkin. Cell 105:891–902
Joazeiro CA, Weissman AM (2000) RING finger proteins: mediators of ubiquitin ligase activity. Cell 102:549–552
Kaplan J (2001) Neuromuscular disorders: gene location. Neuromuscul Disord 11:323–331
Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC,

Baldwin J, Devon K, et al (2001) Initial sequencing and analysis of the human genome. Nature 409:860–921

Lawrence JB, Oxvig C, Overgaard MT, Sottrup-Jensen L, Gleich GJ, Hays LG, Yates JR III, Conover CA (1999) The insulin-like growth factor (IGF)-dependent IGF binding protein-4 protease secreted by human fibroblasts is pregnancy-associated plasma protein-A. Proc Natl Acad Sci USA 96: 3149–3153

Liu HX, Cartegni L, Zhang MQ, Krainer AR (2001) A mechanism for exon skipping caused by nonsense or missense mutations in *BRCA1* and other genes. Nat Genet 27:55–58

Liu HX, Zhang M, Krainer AR (1998) Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins. Genes Dev 12:1998–2012

Reymond A, Meroni G, Fantozzi A, Merla G, Cairo S, Luzi L, Riganelli D, Zanaria E, Messali S, Cainarca S, Guffanti A, Minucci S, Pelicci PG, Ballabio A (2001) The tripartite motif family identifies cell compartments. EMBO J 20:2140–2151

Shen LX, Basilion JP, Stanton VP Jr (1999) Single-nucleotide polymorphisms can cause different structural folds of mRNA. Proc Natl Acad Sci USA 96:7871–7876

Shimura H, Hattori N, Kubo S, Mizuno Y, Asakawa S, Minoshima S, Shimizu N, Iwai K, Chiba T, Tanaka K, Suzuki T (2000) Familial Parkinson disease gene product, parkin, is a ubiquitin-protein ligase. Nat Genet 25:302–305

Shimura H, Schlossmacher MG, Hattori N, Frosch MP, Trockenbacher A, Schneider R, Mizuno Y, Kosik KS, Selkoe DJ (2001) Ubiquitination of a new form of α-synuclein by parkin from human brain: implications for Parkinson's disease. Science 293:263–269

Shokeir MH, Kobrinsky NL (1976) Autosomal recessive muscular dystrophy in Manitoba Hutterites. Clin Genet 9:197–202

Shokeir MH, Rozdilsky B (1985) Muscular dystrophy in Saskatchewan Hutterites. Am J Med Genet 22:487–493

Silahtaroglu AN, Tumer Z, Kristensen T, Sottrup-Jensen L, Tommerup N (1993) Assignment of the human gene for pregnancy-associated plasma protein A (PAPPA) to 9q33.1 by fluorescence in situ hybridization to mitotic and meiotic chromosomes. Cytogenet Cell Genet 62:214–216

Slack FJ, Basson M, Liu Z, Ambros V, Horvitz HR, Ruvkun G (2000) The *lin-41* RBCC gene acts in the *C. elegans* heterochronic pathway between the *let-7* regulatory RNA and the LIN-29 transcription factor. Mol Cell 5:659–669

Slack FJ, Ruvkun G (1998) A novel repeat domain that is often associated with RING finger and B-box motifs. Trends Biochem Sci 23:474–475

Sonoda J, Wharton RP (2001) *Drosophila* brain tumor is a translational repressor. Genes Dev 15:762–773

Torok M, Etkin LD (2001) Two B or not two B? overview of the rapidly expanding B-box family of proteins. Differentiation 67:63–71

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, et al (2001) The sequence of the human genome. Science 291:1304–1351

Weiler T, Greenberg CR, Nylen E, Morgan K, Fujiwara TM, Crumley MJ, Zelinski T, Halliday W, Nickel B, Triggs-Raine B, Wrogemann K (1997) Limb girdle muscular dystrophy in Manitoba Hutterites does not map to any of the known LGMD loci. Am J Med Genet 72:363–368

Weiler T, Greenberg CR, Zelinski T, Nylen E, Coghlan G, Crumley MJ, Fujiwara TM, Morgan K, Wrogemann K (1998) A gene for autosomal recessive limb-girdle muscular dystrophy in Manitoba Hutterites maps to chromosome region 9q31-q33: evidence for another limb-girdle muscular dystrophy locus. Am J Hum Genet 63:140–147

Zhu X, Hadhazy M, Groh ME, Wheeler MT, Wollmann R, McNally EM (2001) Overexpression of γ-sarcoglycan induces severe muscular dystrophy: implications for the regulation of sarcoglycan assembly. J Biol Chem 276:21785–21790