Am. J. Hum. Genet. 60:1128–1138, 1997

Multiple Independent Molecular Etiology for Limb-Girdle Muscular Dystrophy Type 2A Patients from Various Geographical Origins

I. Richard,[1] L. Brenguier,[1] P. Dinçer,[2] C. Roudaut,[1] B. Bady,[3] J.-M. Burgunder,[4] R. Chemaly,[5] C. A. Garcia,[6] G. Halaby,[4] C. E. Jackson,[7] D. M. Kurnit,[8] G. Lefranc,[9] C. Legum,[10] J. Loiselet,[4] L. Merlini,[11] A. Nivelon-Chevallier,[12] E. Ollagnon-Roman,[3] G. Restagno,[13] H. Topaloglu,[2] and J. S. Beckmann[1]

[1]URA 1922 CNRS, Généthon, Evry, France; [2]Medical Biology Department, Hacettepe University, Ankara; [3]Hôpital Neurologique et Neuro-Chirurgical Pierre Wertheimer, Lyon; [4]Neuromorphologishes Labor, Neurologie Inselspital, Bern; [5]Hotel-Dieu de France et Faculté de Médecine, Université Saint-Joseph, Beirut; [6]Lousiana State University Medical Center, New Orleans; [7]Division of Clinical and Molecular Genetics, Henry Ford Hospital, Detroit; [8]Howard Hughes Medical Institute, University of Michigan Medical Center, Ann Arbor; [9]Institut de Génétique Moléculaire, Laboratoire d'Immunogénétique Moléculaire UMR CNRS 9942, University of Montpellier, Montpellier, France; [10]Elias Souraski Medical Center, Genetic Institute, Ichilov Hospital, Tel Aviv; [11]Neuromuscular Laboratory, Istituto Ortopedico Rizzoli, Bologna; [12]Centre Hospitalier Régional et Universitaire de Dijon, Hôpital d'Enfants, Dijon; and [13]Azienda Ospedaliera San Giovanni Battista, Servizio Universitario Convenzionato di Genetica Medica, Turin

Summary

Limb-girdle muscular dystrophies (LGMDs) are a group of neuromuscular diseases presenting great clinical heterogeneity. Mutations in CANP3, the gene encoding muscle-specific calpain, were used to identify this gene as the genetic site responsible for autosomal recessive LGMD type 2A (LGMD2A; MIM 253600). Analyses of the segregation of markers flanking the LGMD2A locus and a search for CANP3 mutations were performed for 21 LGMD2 pedigrees from various origins. In addition to the 16 mutations described previously, we report 19 novel mutations. These data indicate that muscular dystrophy caused by mutations in CANP3 are found in patients from all countries examined so far and further support the wide heterogeneity of molecular defects in this rare disease.

Introduction

Owing to overlapping clinical features, it was proposed, in 1995 at a European NeuroMuscular Centre workshop in Naarden, The Netherlands (Bushby and Beckmann 1995), to group temporarily the autosomal progressive proximal muscular dystrophies under the denomination of limb-girdle muscular dystrophies (LGMDs) (Bushby and Beckmann 1995). These myopathies represent a group of rare neuromuscular diseases clinically defined by progressive atrophy and weakness of the proximal limb muscles, showing considerable variation in age at onset, evolution, and severity (Walton and Natrass 1954; Bushby and Beckmann 1995).

At least six distinct genetic entities leading to an autosomal recessive LGMD are recognized. LGMD type 2A (LGMD2A; MIM 253600) was the first to be defined by its linkage to chromosome 15q in patients from an isolate from Réunion Island (Beckmann et al. 1991). In families from Brazil, from an Amish community in northern Indiana, and from France, the disease subsequently was found to be linked to this chromosome, although genetic heterogeneity also was evidenced (Passos-Bueno et al. 1993; Allamand et al. 1995). The LGMD2B locus was mapped on chromosome 2p, to a 4-cM interval between D2S291 and D2S286 (Bashir et al. 1994; Passos-Bueno et al. 1995). LGMD2C and LGMD2D (also known as "severe childhood autosomal recessive muscular dystrophies," or "SCARMDs") were characterized by marked deficiency of a 50-kD dystrophin-associated glycoprotein named "α-sarcoglycan," or "adhalin" (Matsumura et al. 1992), and were linked to chromosome 13q and chromosome 17q, respectively (Ben Othmane et al. 1992; Azibi et al. 1993; Roberds et al. 1994). Mutations in the α-sarcoglycan gene, localized to chromosome 17 (McNally et al. 1994; Roberds et al 1994), are the cause of LGMD2D (Roberds et al. 1994; Piccolo et al. 1995), whereas LGMD2C is due to a defect in γ-sarcoglycan (Noguchi et al. 1995). A fifth locus (LGMD2E), mapping on chromosome 4q and encoding for β-sarcoglycan, was identified as being responsible for the disease that was present in a number of Amish pedigrees from southern Indiana (Lim et al. 1995), as well as in one other, unrelated patient (Bönnemann et al. 1995). Finally, the LGMD2F locus was mapped recently on chromosome 5q33-q34 (Passos-Bueno et al. 1996) and identified subsequently

Received August 13, 1996; accepted for publication January 24, 1997.

Address for correspondence and reprints: Dr. Jacques S. Beckmann, Généthon, 1 rue de l'Internationale, 91000 Evry, France. E-mail: beckmann@genethon.fr

0002-9297/97/6005-0014$02.00

as the δ-sarcoglycan gene (Nigro et al. 1996*a*, 1996*b*). In addition, some families still seem to belong to none of the above-mentioned groups.

In 1995, the LGMD2A locus was identified as CANP3, the gene encoding the large subunit of the muscle-specific calcium-activated neutral protease 3 (Richard et al. 1995). To determine the geographic and genomic distributions of calpain 3 mutations, we examined LGMD2 families from different origins, for linkage to chromosome 15 and/or for CANP3 mutations.

Subjects and Methods

Case Ascertainment and Genotype Analysis

Thirty-two LGMD families from seven different geographic areas were referred by the physicians involved in this study. Whereas most of these families had dystrophin immunohistochemistry scoring, only the Italian families also had their sarcoglycan tested. In addition, there was no systematic case assessment, because of the multicentric nature of this study. DNA from members from these families was analyzed with highly polymorphic markers (D15S514, D15S779, D15782, D15S780, D15S778, and D15S222) of chromosome 15q15.1-q15.3 (Allamand et al. 1995). PCR amplification of 50 ng of DNA was performed in a 50-μl final reaction mix, as described in the study by Fougerousse et al. (1994). Two-point and multipoint LOD-score calculations were performed by use of the 5.1 version of the LINKAGE software package (Lathrop et al. 1985). The disease was considered to be a fully penetrant recessive trait with a gene frequency of .001. Haplotypes were constructed by the minimization of the number of recombination events.

Mutation Analysis of LGMD2A Families

One hundred nanograms of human DNA were used per PCR under buffer and cycle conditions and with the primers described in the study by Richard et al. (1995). For heteroduplex analysis (Keen et al. 1991), 10 μl of PCR products denatured at 94°C for 5 min were allowed to slowly renature for 1 h, at room temperature. Then, they were electrophoresed on 1.5-mm–thick Hydrolink Mutation Detection Enhancement (MDE) gels (AT Biochem) at 500–600 V for 12–15 h, depending on fragment length. The migration profile was visualized under UV illumination after ethidium bromide staining. Illegitimate transcription analyses of the CANP3 cDNA in LGMD2 patients were performed as in the study by Richard et al. (1995), after extraction of total cellular RNA from lymphoblastoid cell lines by RNA-Zol (Bioprobe Systems), according to the method of Chomczynski and Sacchi (1987). For sequence analyses, the PCR products were subjected to dye-dideoxy sequencing, after purification through Microcon devices (Amicon).

Results

After the identification of the CANP3 gene as the LGMD2A locus, a number of families with limb-girdle wasting and elevated serum creatine kinase (CK) were referred to us by several physicians. We chose to investigate those which either were large enough for linkage analysis or presented consanguinity. Two small families (IT7093 and US9351) also were included, since we had access to RNA samples, which allowed for a rapid mutation search. Altogether, 32 families from different origins were examined. Data for 15 families are not presented, since we either could exclude linkage to chromosome 15 or failed to identify disease-associated mutations (table 1). The latter case applies to four families, although the genotyping results are compatible with chromosome 15 linkage. The other 17 kindreds (fig. 1) include two LGMD2 families from Italy (IT5414 and IT7093), one family from Israel (IS1), two families from Lebanon (L5988 and L7111), two families from France (M3390 and M2997), one family from Switzerland (S5393), six families from Turkey (T10, T11, T13, T14, T17, and T18), and three families from the United States (US9351, US9344, and US3). All but four families (US9351, T10, T11, and T13) include two or more affected members (fig. 1). Polymorphic markers from the LGMD2A region were used, allowing the construction of haplotypes and the calculation of LOD scores in families with informative pedigrees (the order of the markers was based on the physical mapping data of the LGMD2A contig; Fougerousse et al. 1994). LOD scores

Table 1

Number of Families Investigated, according to Country of Origin and Linkage Classification

	No. of Families			
Country	Total	Chromosome 15 Linked	Excluded by Linkage	Inconclusive
France	2	2	0	0
Israel	4	1	2	1
Italy[a]	(3)	(2)	(1)	(0)
Lebanon	4	2	2	0
Switzerland	6	1	4	1
Turkey[b]	(6)	(6)	(0)	(0)
United States	7	3	2	2
Total[c]	23	9 (39.1%)	10 (43.5%)	4 (17.4%)

[a] Data for the Italian families are in parentheses because these families were known to be sarcoglycan positive.

[b] Data for the Turkish families are in parentheses because these families were part of a subsequent, enlarged study including 21 families, of which 10 families were shown to be chromosome 15 linked (Dinçer et al., in press).

[c] Data for the Italian and Turkish families are excluded from the totals.

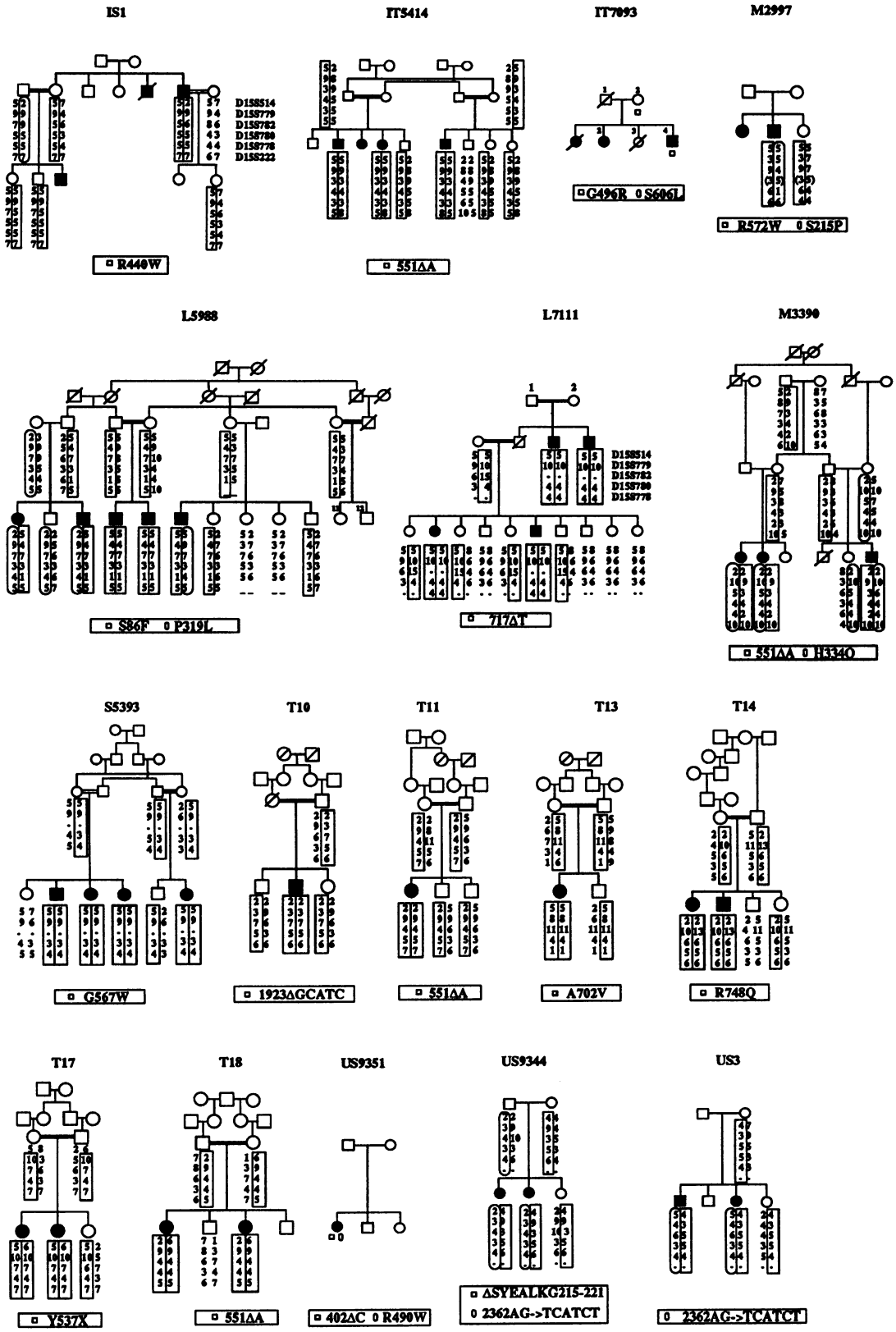

Figure 1 Genealogies and haplotypes. Blackened squares and circles represent affected males and females, respectively, and healthy individuals are represented by unblackened squares (males) and circles (females). Haplotypes bearing the disease locus are boxed, and the corresponding mutations are indicated below each pedigree. Marker loci are listed from centromere to telomere (top to bottom) and are indicated to the right of the pedigrees IS1 and L7111. Ungenotyped markers in the haplotypes are noted with a dash (–).

were not computed for the Turkish families, the Swiss family, and family L7111, since we could infer, by their homozygosity, that they belonged to the LGMD2A group. For the examined families, pairwise LOD scores between the nearest flanking markers and the disease, are indicated in table 2. The overall maximum LOD scores for D15S779 and D15S782 were 5.82 (θ = .05) and 8.07 (θ = .00), respectively. Homozygosity in the area of the LGMD2A gene was present in patients from 11 consanguineous families (IT5414, L5988, L7111, IS1, S5393, T10, T11, T13, T14, T17, and T18; fig. 1).

Table 2

Pariwise LOD Scores between Markers D15S779 and D15S782 and the Disease Locus, in the Families Studied

Marker and Family[a]	LOD Score at θ = .00	.001	.01	.05	.10
D15S779:					
IS1	−.12	−.12	−.12	−.11	−.10
IT5414	2.28	2.27	2.22	2.01	1.73
L5988	2.05	2.07	2.15	2.21	2.06
M2297	.12	.12	.12	.10	.08
M3390	1.46	1.46	1.43	1.28	1.10
US9344	.73	.72	.71	.62	.51
US3	−.98	−.94	−.70	−.29	−.11
Total	5.54	5.58	5.81	5.82	5.27
D15S782:					
IS1	−.03	−.02	−.02	.01	.02
IT5414	2.39	2.39	2.33	2.09	1.78
L5988	2.83	2.82	2.79	2.60	2.31
M2297	.12	.12	.12	.09	.06
M3390	1.60	1.59	1.55	1.37	1.14
US9344	.73	.72	.71	.62	.51
US3	.43	.42	.42	.37	.32
Total	8.07	8.04	7.90	7.15	6.14

[a] The first letter(s) of the family designation refers to geographic origin. M = metropolitan France; US = United States; L = Lebanon; IT = Italy; and IS = Israel.

Mutation Search

We proceeded to search for CANP3 mutations in all of the families included in this study. By use of flanking intronic primers, each of the 24 exons was PCR amplified by use of genomic DNA as a template and was subjected to heteroduplex analysis (Keen et al. 1991; Richard et al. 1995) or to direct sequencing for consanguineous families. The PCR products were sequenced whenever a reduced-mobility heteroduplex band was revealed in ethidium bromide–stained MDE gels (fig. 2). In addition, mRNA analyses were performed on muscle biopsies from patients of families T13 and T17 and on illegitimate transcription products (Chelly et al. 1989) from patients of families L5988, IT7093, and S5393, for which established lymphoblastoid cell lines were available. After reverse transcription, PCR was performed by use of one of four sets of primers, to obtain four overlapping fragments spanning the entire CANP3 coding sequence (Richard et al. 1995). The products were electrophoresed in agarose gels and were visualized after ethidium bromide staining. No aberrant splicing was detected, and thus all PCR fragments were sequenced. This procedure led to the detection of the second mutation present in families IT7093 and L5988.

Mutations were detected in 24 of the 25 chromosomes examined (fig. 2 and table 3). Altogether, there were 19 mutations—12 missense mutations (S86F, S215P, P319L, H334Q, R440W, R490W, G496R, G567W, R572W, S606L, A702V, and R748Q), 1 in-frame deletion (ΔSYEALKG215–221), 5 frameshift mutations (19–23ΔGCATC, 402ΔC, 551ΔA, 717ΔT, and 2362AG→TCATCT), and 1 stop mutation (Y537X). As expected from family history and from haplotype data, 11 mutations were found in the homozygous state. When DNA samples from other members of the family were available, the parental origin and the segregation of altered alleles were determined within the pedigree. None of the missense mutations was found in >100 control chromosomes (data not shown).

We failed to detect the second mutation in one of the LGMD2A chromosomes (detection efficiency 96%). No cell lines were available for the corresponding pedigree (US3). Hence, we cannot take advantage of illegitimate transcription to uncover the second LGMD2A mutation, by, for instance, monitoring heteroduplexes among reverse transcriptase–PCR illegitimate-transcription products (Chelly et al. 1989; Richard and Beckmann 1995). We therefore can only suspect the presence of an additional mutation in this family.

Nature of the Mutations

Among the missense mutations (fig. 3), one (S86F) is localized within domain I of the protein; three (S215P, P319L, and H334Q) are localized within domain II, and one (P319L) of these is in the IS1 sequence; six (R440W, R490W, G496R, G567W, R572W, and S606L) are localized within domain III; and two (A702V and R748Q) are localized within domain IV. Whereas all mutations are located in residues conserved throughout all muscle-specific calpains (see fig. 2 of the study by Richard and Beckmann [1996]), four mutations affect residues conserved throughout all mammalian members of the calpain family. All but three of the point mutations represented here are either C→T or G→A transitions; these mutations represent 9 (47%) of 19 mutations. In this study, 8 (42%) of 19 affected CpG sites (table 3) may represent spontaneous deamination of 5-methyl cytosine at CpG sites, which is one of the main causes of mutations in vertebrate coding sequences. This percentage is slightly higher than the average frequency of 35% that

Am. J. Hum. Genet. 60:1128–1138, 1997

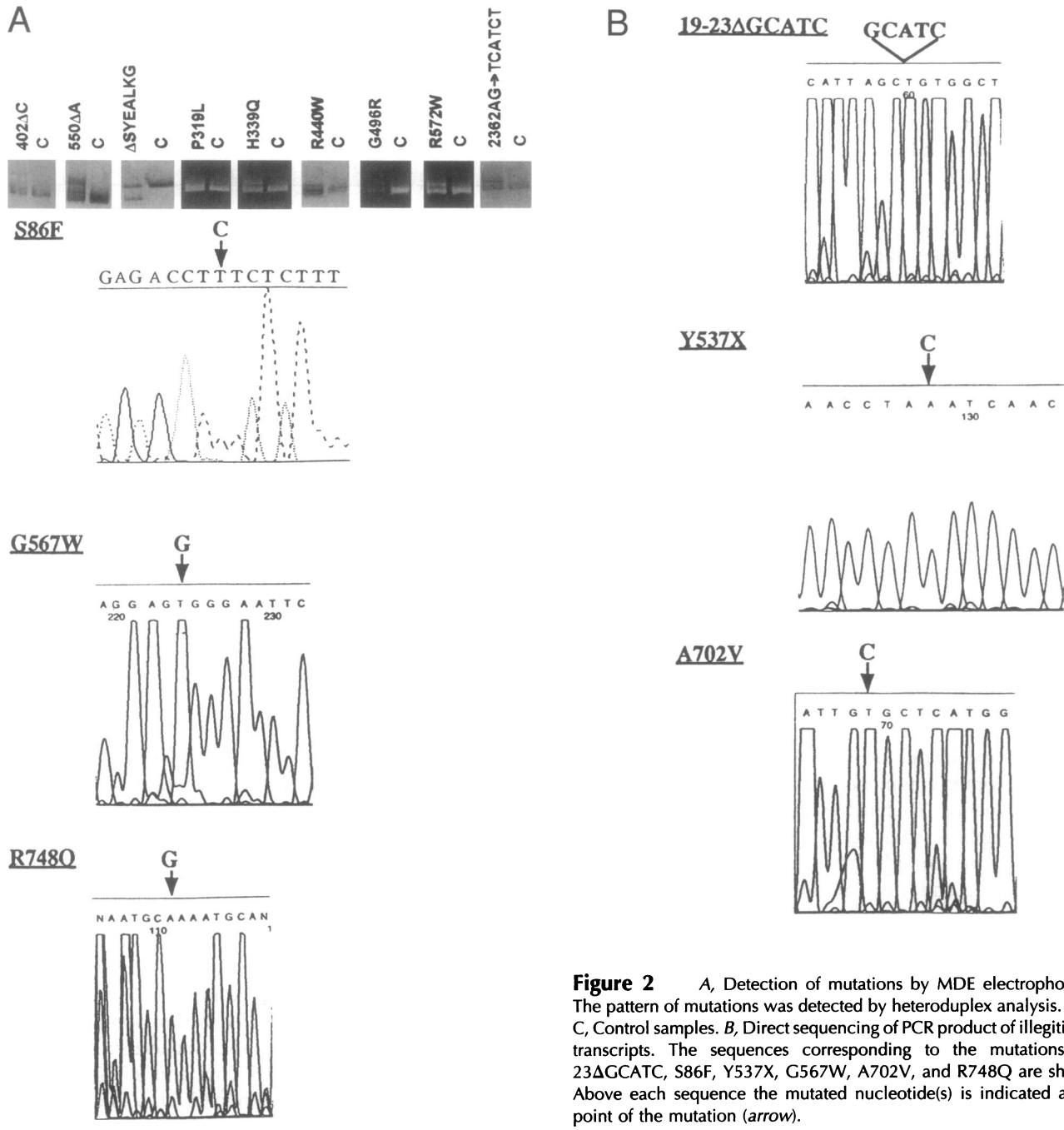

Figure 2 *A,* Detection of mutations by MDE electrophoresis. The pattern of mutations was detected by heteroduplex analysis. Lane C, Control samples. *B,* Direct sequencing of PCR product of illegitimate transcripts. The sequences corresponding to the mutations 19-23ΔGCATC, S86F, Y537X, G567W, A702V, and R748Q are shown. Above each sequence the mutated nucleotide(s) is indicated at the point of the mutation (*arrow*).

was reported by Cooper and Youssoufian (1988). The nucleotide deletions or insertion 19-23ΔGCATC, 402ΔC, 551ΔA, 717ΔT, and 2362AG→TCATCT lead to frameshifts and premature termination of translation and therefore are presumed to be null mutants. A deletion of seven amino acids (ΔSYEALKG in position 215–221) in domain II, preserving the correct protein frame, was detected in patients from one family (US9344). The deleted amino acids are conserved in all muscle-specific calpains, and four of them (Y, E, L, and G) are conserved in all mammalian calpains (Sorimachi and Suzuki 1992).

Among the mutations identified in this study, two frameshift mutations and one missense mutation have been described in a previous study (Richard et al. 1995). In order to determine whether these mutations may have a common origin, haplotypes constructed by use of five microsatellites (D15S514, D15S779, D15S782, D15S780, and D15S778) flanking the LGMD2A locus and spanning <1 cM were examined in these families (table 4), except with regard to the R490W mutation (only the proband in family US9351 was available). Altogether, the 551ΔA in exon 4 was found in one Italian

Table 3

CANP3 Mutations in LGMD2A Families

Exon	Family/Families	Nucleotide Position[a]	Nucleotide Change[b]	Amino Acid Position[c]	Mutation[d]
1	T10[e]	19–23	ΔGCATC	7–8	19–23ΔGCATC
1	L5988[e]	257	TCT→TTT	86	S86F
3	US9351	402	GCC→GC	134	402ΔC
4	M3390, IT5414,[e] T11,[e] T18[e]	551	ACG→CG	184	551ΔA
5	M2888	643	TCC→CCC	215	S215P
5	US9344	643–663	Δ21 bp	215–221	ΔSYEALKG215-221
5	L7111	717	ΔT	239	717ΔT
7	L5988	956	CCG→CTG[f]	319	P319L
7	M3390	1002	CAC→CAG[f]	334	H334Q
10	IS1	1318	CGG→TGG[f]	440	R440W
11	US9351	1468	CGG→TGG[f]	490	R490W
11	IT7093	1486	GGG→AGG	496	G496R
13	T17[e]	1611	TAC→TAA	537	Y537X
13	S5393[e]	1699	GGG→TGG	567	G567W
13	M2997	1714	CGG→TGG[f]	572	R572W
16	IT7093	1817	TCG→TTG[f]	606	S606L
19	T13[e]	2105	GCG→GTG[f]	702	A702V
21	T14[e]	2243	CGA→CAA[f]	748	R748Q
22	US3, US9344	2362–2363	AG→TCATCT	788	2362AG→TCATCT

[a] Numbered on the basis of the cDNA, starting from ATG.
[b] The mutated nucleotides are underlined.
[c] Numbered on the basis of the protein sequence, starting from the first methionine residue.
[d] Nomenclature follows the recommendations proposed at the second "Locus-Specific Mutation Databases" meeting, which was held on March 24, 1996, in Heidelberg, Germany.
[e] Family in which patient(s) is homozygous for the mutation.
[f] Modified CpG sites.

family (IT5414 [this study]), in three French families (M3390 [this study] and M1394 and M2888 [Richard et al. 1995]), and in two Turkish families (T11 and T18 [this study]). In these six families, one can see a minor divergence in the haplotypes, particularly at marker D15S782. Clearly, the two Turkish families and one of the French families (M2888) share a common core haplotype, and this also could be true for the other families, to the extent that the discrepant D15S782 alleles derive by mutation from the common ancestral allele. The 2362AG→TCATCT mutation in exon 21 was encountered not only in two American families (US9344 and US3 [this study]) but also in a Brazilian kindred and in a Réunion Island pedigree (Richard et al. 1995). On the chromosomes bearing this mutation, there is conservation of the complete four-marker haplotype in the Brazilian and the Réunion Island families and in one American family but not in the second American family (table 4). This result thus is suggestive of an independent origin of this mutation in the latter family.

Clinical Features

Table 5 summarizes some of the salient clinical features of these patients (note that the assessments were neither systematic nor complete). All patients showed elevated serum CK, with elevations in a range of 7–84 times the upper limit of normal. There was marked intra- and interfamilial phenotypic variability in age at onset (range 2.5–40 years) and loss of ambulation. For example, affected individuals in family IT7093 presented with a very mild phenotype, with onset at ages 30 years and 40 years, and are still ambulatory at ages 54 years and 66 years, respectively. In contrast, two Lebanese sibs had onset at age 6 years and had loss of independent walking at ages 13 years and 15 years.

In families M2997, M3390, and S5393, an intrafamilial variability in age at onset and in evolution can be noted. The proband in family M2997 had onset at age 14 years, and, at his present age (23 years), his walking perimeter is 500 m with severe hyperlordosis, whereas his sibling who was diagnosed by a high CK level at age 12 years but was asymptomatic at that time is still able to participate in some moderate physical exercise at age 20 years. In family M3390, the two sisters had onset at ages 17 years and 26 years. Likewise, in family S5393, one affected sib had onset at age 13 years and was unable to walk unaided for >50 m at age 25 years, whereas two other affected sibs were

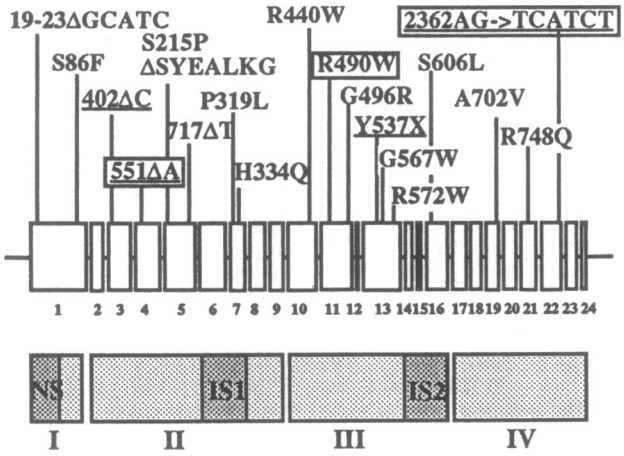

Figure 3 Distribution of mutations along the CANP3 protein. *Top,* Name and position of the mutations are indicated along the 24 exons, which are indicated by unblackened boxes, with the corresponding numbers written below. Null-type mutations (deletions, insertions, and stop codons) are underlined, and recurrent mutations are boxed. *Bottom,* Schematic representation of the CANP3 protein, its four domains (I, II, III, and IV), and the muscle-specific sequences (NS, IS1, and IS2).

still walking at ages 25 years and 26 years, even though their ages at onset (4 years and 6 years, respectively) were much earlier.

The two mutations identified in family L5988 have different clinical consequences. Patients who are homozygous for the S86F mutation had disease onset at age 6–7 years and presented with severe weakness in their lower limbs, leading to loss of mobility <8 years after onset. In contrast, their cousins who are compound heterozygotes for the S86F/P319L mutations are comparatively mildly affected, with disease onset at age 15–17 years, on average, and with loss of mobility at age 32 years for one of them, whereas the two others still are walking at ages 29 years and 28 years. Furthermore, it is interesting to note that all healthy heterozygotes carrying the S86F mutation have unexpectedly mildly elevated CK levels (data not shown). This observation also suggests that this mutation affects muscle cells in a somewhat dominant manner. A similar observation of elevated CK levels among obligate carriers also was seen in family IT5414. No CK data are available for other 551ΔA carriers from the other pedigrees.

In two families, L5988 and US9351, there was difficulty in establishing a clinical diagnosis of muscular

Table 4

Haplotypes for the 551ΔA and the 2362AG→TCATC Mutations

Mutation and Marker[a]	Family[b]					
	IT5414	M1394[c]	M2888[c]	M3390	T11	T18
551ΔA:						
D15S514	5 (.31)	2 (.38)	2	2	2	2
D15S779	9 (.28)	9	9	9	9	9
D15S782	3 (.01)	5 (.35)	4 (.09)	3 (.01)	4	4
D15S780	4 (.16)	4 (.16)	4 (.16)	4 (.16)	5 (.03)	4
D15S778	3 (.08)	3	2 (.13)	2	7 (.02)	5 (.48)
	B505[c]	R14[c]	US3	US9351		
2362AG→TCATC:						
D15S514	2 (.38)	2	5 (.31)	2		
D15S779	3 (.14)	3	4 (.07)	3		
D15S782	4 (.09)	4	6 (.21)	4		
D15S780	3 (.44)	3	3	3		
D15S778	4 (.27)	4	4	4		

[a] Markers are listed according to their linear map order. The disease locus is represented by a horizontal gray bar between markers D15S779 and D15S782.

[b] The corresponding allele frequencies determined in the CEPH reference families are indicated in parentheses. The shared haplotypes are shaded. The identification of the mutation in the families indicated by underlining is from this study.

[c] The identification of the mutation in this family was described in the study by Richard et al. (1995*a*).

Table 5

Clinical Features of LGMD2A Patients included in This Study

Family	Mutation	Peak CK[b] (IU/liter or × Normal)	Age at Onset (Years)	Walking Ability (Present Age, in Years)	Comments[c]
IS1	R440W[a]	13×	21	Loss at age 55 years	Calf hypertrophy; pelvic girdle more affected
IT5414	551ΔA[a]	27×	10–13	Loss at age 24–28 years	Scapulohumeral distribution at the beginning; extensor digitorum brevis muscle hypertrophy; Achille's tendon contractures
IT7093	G496R/S606L	7×	40	Still walking (66)	Scapulohumeral distribution at the beginning
	G496R/S606L	10×	30	Still walking (54)	Scapular winging
L5988	S86F[a]	7×	7	Loss at age 15 years	First diagnosis = polymyositis
	S86F[a]	20×	7	Loss at age 13 years	Calf hypertrophy
	S86F/P319L	10×	15–17	Still walking (31–32)	First diagnosis = polymyositis; pelvic-girdle distribution
L7111	717ΔT[a]	8×	19	Still walking (23)	nd
M2997	R572W/S215P	4,500	18	Still walking (20)	Pelvic girdle more affected
	R572W/S215P	4,830	14	Still walking (23)	nd
M3390	551ΔA/H334Q	1,094	17	Still walking (33)	Retraction; distal muscle also affected
S5393	G567W[a]	na	13	Loss at age 25 years	nd
	G567W[a]	2,000	4	Still walking (25)	nd
	G567W[a]	6×	13	Still walking (31)	nd
	G567W[a]	84×	6	Loss at age 26 years	nd
T10	19-23ΔGCATC[a]	2,980	3	Still walking (11)	Scapular winging
T11	551ΔA[a]	7,493	12	Still walking (14)	Calf hypertrophy
T13	A702V[a]	4,140	11	Still walking (13)	nd
T14	R748Q[a]	5,396	2.5	Still walking (15)	Calf hypertrophy
T17	Y537X[a]	5,427	11	Still walking (14)	nd
T18	551ΔA[a]	1,961	10	Still walking (22)	Scapular winging
US9351	402ΔC/R490W	2,040	11	Loss at age 42 years	Beginning with weakness of the calves; first diagnosis muscular dystrophy or inflammatory myopathy
US9344	ΔSYEALG/2362AG→TCATCT	5,000	10–14	Loss at age 28–29 years	nd
US3	2362AG→TCATCT/?	1,300	12	Loss at age 32 years	nd
	2362AG→TCATCT/?	1,600	13·	Loss at age 34 years	nd

[a] Homozygous mutation.
[b] na = not available.
[c] nd = no data.

dystrophy, because of the nonspecific features of their presentation. Two patients belonging to different branches of family L5988 were diagnosed independently with polymyositis. A muscle biopsy of one of them showed minimal abnormalities at the limits of significance. Both of these patients were treated with steroids for an extended period. Eventually, family history led to a consideration of the diagnosis of muscular dystrophy. In pedigree US9344, the proband first presented at age 11 years with calf weakness following periods of high fever. The diagnosis at that time was poliomyelitis. Examination of muscle biopsies performed at age 25 years showed a severe inflammatory necrotizing myopathy. A similar picture could be seen in either polymyositis or dystrophy. The patient then was treated with steroids for 3–4 mo. Since there was never a definitive response to steroid therapy, the patient has been considered since to have a dystrophy. In these two families, it is only the identification of the pathogenic mutation that allowed the definitive LGMD2A diagnosis.

Discussion

The current report is the first attempt to establish a systematic molecular nosology of LGMD2A. Thirty-two families from seven different countries were examined. There was no a priori bias in the initial ascertainment of these LGMD2 families, since, for all but two, there was no knowledge regarding the sarcoglycan status. Linkage analysis and mutation screening identified 17 of them as LGMD2A. The identification of LGMD2A families from an increasing number of countries supports the notion that CANP3 mutations are a cause of LGMD in patients from a variety of ethnic and geographic backgrounds. On the basis of our data and with the consideration of the unknowns as non-LGMD2A (table 1), calpain-deficient families are estimated to represent ⩾39% of all cases with an LGMD diagnosis that have been referred to us. The Italian families were disregarded in the calculation of this percentage, because of the ascertainment bias due to the fact that they were known to be sarcoglycan positive. In this calculation,

we also omitted the Turkish families since they were part of a subsequent and enlarged study of 21 families, of which 10 families were shown to be chromosome 15 linked (Dinçer et al., in press), 6 of which were included here. If we consider the Turkish set, the fraction of LGMD2A cases is $\geq .45$. On the basis of these data, the calpainopathies may represent a significant proportion of all progressive autosomal recessive muscular dystrophies, although studies of a higher number of independent cases are needed to allow precise estimation of this percentage.

Among the 19 mutations identified here, 3 mutations—551ΔA, R490W, and 2362AG→TCATCT—were published previously. These mutations were found in six, two, and four unrelated kindreds, respectively (Richard et al. 1995; this study). The marker haplotypes flanking the disease locus in three French metropolitan, one Sicilian, and two Turkish families carrying the 551ΔA mutation suggest a common mutational event rather than a recurrent coincidental mutation. Considering the geographic dispersion of these families, we concluded that 551ΔA may be an old mutation. Along the same line, it should be noted that a Brazilian, a Réunion Island (Richard et al. 1995), and an American family all share the 2362AG→TCATCT mutation and the flanking haplotype. In contrast, a fourth family (US3) carrying this mutation presents a divergent haplotype and could, therefore, reflect the occurrence of an independent mutational event. Comparison of haplotypes carrying R490W was not possible because the proband of family US9351 has not been genotyped, since the phase could not be determined in this small pedigree.

Whereas, in some instances, it is possible—on the basis of genetic and molecular data—to infer a common ancestral mutation, the same type of information also can be used to challenge hypothesized founder effects. In this context, it is interesting to note that the Swiss family (S5393) presents a different mutation than is presented by the chromosome 15–linked Amish families, despite the suspected common geographic origin (Jackson and Carey 1961).

Careful examination of precisely identified LGMD2 patients will allow for the establishment of more accurate and specific diagnostic criteria. It should be emphasized that, whereas this multicentric study highlights the wide distribution of LGMD2A, the fact that these patients were reported to us by different clinicians, that the clinical data were not obtained prospectively, and that the assessments were neither systematic nor complete prevent, at this time, the presentation of a normalized clinical overall picture. Thus, the phenotypic data are to be considered here as indicative rather than representative. The phenotypic features of the ascertained LGMD2A patients, with regard to reported age at onset and loss of walking ability, are highly variable (table 5), even more variable than reported in previous studies (Bushby and Beckmann 1995; Fardeau et al. 1996). The mean age at onset is 14.2 years (range 2.5–40 years). Muscle wasting is first seen indifferently in a pelvic or a scapular-girdle distribution. In 3 of 16 cases, calf hypertrophy can be noted. In agreement with previous reports, disease evolution is never as severe as in Duchenne muscular dystrophy and sometimes can be very mild (table 5).

Knowledge of the mutations permits us, in principle, to assess possible correlations between the nature or site of the mutation and the resulting phenotype. These correlations, however, may be difficult to establish, because the phenotypic effect of a particular allele can be modified by the nature of the second mutated allele, by genetic factors in the vicinity or at other loci, or by nongenetic factors. Intrafamilial variability can be noted in M2997 and in M3390, emphasizing the difficulties in drawing such correlations. The fact that this cohort regroups patients examined by different physicians further contributes to render these cross-comparisons difficult. Furthermore, the follow-up duration was sometimes too short for us to have a correct idea of the disease's evolution.

In view of these caveats, the identification of homozygous LGMD2A patients in a number of families (IS1, IT5414, L5988, L7111, S5393, T10, T11, T13, T14, T17, and T18; fig. 1), each patient singularizing a specific CANP3 mutation, is of great interest: these patients are worth further thorough clinical studies. In particular, the S86F mutation represents the first case of LGMD2A in which severity is associated with a missense mutation (as a rule of thumb, the latter type was shown to have a less severe prognosis than was seen for null mutations; Fardeau et al. 1996; M. Fardeau, B. Eymard, F. M. S. Tomé, A. Lopez de Munain, I. Richard, and J. S. Beckmann, unpublished data). So far, S86F is the only missense mutation identified that is located in the first protein domain, which is presumed to play a role in regulating calpain 3's catalytic activity (Sorimachi et al. 1989). It is not impossible that the presence of such a mutation could alter the function of this protein by disturbing its regulation and, hence, could lead to a severe disease in the homozygous state and to asymptomatic CK elevation (1.5–6 times normal) in all seven healthy heterozygous carriers (data not shown). Finally, the observation of elevated CK among obligate carriers could have important implications for genetic and family counseling. It needs, however, to be confirmed in a larger sample, to determine if this observation is symptomatic or anecdotal.

The identification of CANP3 as the gene responsible for LGMD2A was the first demonstration of an enzymatic—rather than a structural protein—defect causing muscular dystrophy. This identification also was marked

by an unexpected finding—the discovery of several independent mutations in the small inbred population of Réunion Island, which led to the proposal of a digenic model of inheritance for this disease (Richard et al. 1995). In all families examined so far, segregation of the mutations correlates with the clinical manifestations (no asymptomatic carrier of two mutant calpain genes was identified) and, therefore, does not allow validation of this model. It should be emphasized, however, that, because of both the low prevalence of the disease and the difficulties of ascertainment, the number of healthy siblings in these LGMD2A families is too small for us to reach a significant conclusion.

In this study, 12 missense, 1 nonsense, and 5 frameshift mutations and 1 in-frame deletion were identified. These findings confirm the association of LGMD2A with mutations distributed all over the coding sequence of the CANP3 gene, as well as over all protein domains (fig. 3). Thus, altogether, 35 distinct CANP3 mutations have been reported thus far (Richard and Beckmann 1995; Richard et al. 1995; this study). Mutations of all types have been observed, with the exception of large deletions, large insertions, and duplications. The identified missense mutations could represent a useful source for establishment of structure/function relationships, thereby providing useful tools for a better comprehension of the associated physiopathology of calpainopathies.

In a diagnostic perspective, it should be stressed that none of these mutations can account for a substantial fraction of LGMD2A patients, although the same mutations were observed occasionally in independent families. The absence of predominant mutations or the clustering in specific exons does not support the establishment of a simple molecular-genetic diagnostic scheme, as for the ΔF508 mutation in cystic fibrosis. But the recognition of a recurrent haplotype could serve as a useful hint to circumvent the necessity to scan the entire gene, by targeting the screening to specific mutations.

Finally, the identification of CANP3 mutation(s) provides the opportunity to address, as illustrated in this study, the prevalence and the frequency of calpainopathies among the progressive muscular dystrophies; eventually to establish phenotype/genotype relationships; and, finally, to classify small families as LGMD2A. The examples of pedigrees L5988 and US9351 unambiguously show that it indeed is possible now to make a genetic diagnosis by direct mutational analysis: a correct diagnosis is of crucial importance for the patient's care, treatment, and well-being.

Acknowledgments

We gratefully acknowledge the technical expertise of Corinne Cruaud and Nabiha Salem, as well as the support of Dr. K. Azibi and Prof. M. Fardeau. We thank Prof. H. Aydenian and Dr. I. Touitou for the determination of CK levels. We also are thankful to the reviewers for their critical suggestions, which resulted in a much-improved manuscript. This research was supported by grants from the Association Française contre les Myopathies, the Actions Concertées Coordonnées des Sciences du Vivant, and the Groupement de Recherche et d'Études sur les Génomes.

References

Allamand V, Broux O, Richard I, Fougerousse F, Chiannilkulchai N, Bourg N, Brenguier L, et al (1995) Preferential localization of the limb-girdle muscular dystrophy type 2A gene in the proximal part of a 1-cM 15q151-q153 interval. Am J Hum Genet 56:1417–1430

Azibi K, Bachner L, Beckmann JS, Matsumura K, Hamouda E, Chaouch M, Chaouch A, et al (1993) Severe childhood autosomal recessive muscular dystrophy with the deficiency of the 50 kDa dystrophin-associated glycoprotein maps to chromosome 13q12. Hum Mol Genet 2:1423–1428

Bashir R, Strachan T, Keers S, Stephenson A, Mahjneh I, Marconi G, Nashef L, et al (1994) A gene for autosomal recessive limb-girdle muscular dystrophy maps to chromosome 2p. Hum Mol Genet 3:455–457

Beckmann JS, Richard I, Hillaire D, Broux O, Antignac C, Bois E, Cann H, et al (1991) A gene for limb-girdle muscular dystrophy maps to chromosome 15 by linkage analysis. C R Acad Sci III 312:141–148

Ben Othmane K, Ben Hamida M, Pericak-Vance MA, Ben Hamida C, Blel S, Carter SC, Bowcock AM, et al (1992) Linkage of Tunisian autosomal recessive Duchenne-like muscular dystrophy to the pericentromeric region of chromosome 13q. Nat Genet 2:315–317

Bönnemann CG, Modi R, Noguchi S, Mizuno Y, Yoshida M, Gussoni E, McNally EM, et al (1995) β-sarcoglycan (A3b) mutations cause autosomal recessive muscular dystrophy with loss of the sarcoglycan complex. Nat Genet 11:266–273

Bushby KM, Beckmann JS (1995) Report of the 30th and 31st ENMC international workshop on the limb-girdle muscular dystrophies: proposal for a new nomenclature. Neuromuscul Disord 5:337–343

Chelly J, Concordet JP, Kaplan JC, Kahn A (1989) Illegitimate transcription: transcription of any gene in any cell type. Proc Natl Acad Sci USA 86:2617–2621

Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem 162:156–159

Cooper DN, Youssoufian H (1988) The CpG dinucleotide and human genetic disease. Hum Genet 78:151–155

Dinçer P, Leturcq F, Richard I, Piccolo F, Yalnizoğlu D, de Toma C, Akçören Z, et al. A biochemical, genetic and clinical survey of autosomal recessive limb girdle muscular dystrophies in Turkey. Ann Neurol (in press)

Fardeau M, Hillaire D, Mignard C, Feingold N, Feingold J, Mignard D, de Ubeda B, et al (1996) Juvenile limb-girdle muscular dystrophy: clinical histopathological and genetic data on a small community living in the Réunion Island. Brain 119:295–308

Fougerousse F, Broux O, Richard I, Allamand V, Pereira de Souza A, Bourg N, Brenguier L, et al (1994) Mapping of a chromosome 15 region involved in limb-girdle muscular dystrophy. Hum Mol Genet 3:285–293

Jackson CE, Carey JH (1961) Progressive muscular dystrophy: autosomal recessive type. Pediatrics 28:77–84

Keen J, Lester D, Inglehearn C, Curtis A, Bhattacharya S (1991) Rapid detection of single base mismatches as heteroduplexes on Hydrolink gels. Trends Genet 7:5

Lathrop GM, Lalouel JM, Julier C, Ott J (1985) Multilocus linkage analysis in humans: detection of linkage and estimation of recombination. Am J Hum Genet 37:482–498

Lim LE, Duclos F, Broux O, Bourg N, Sunada Y, Allamand V, Meyer J, et al (1995) β-sarcoglycan: characterization and role in limb-girdle muscular dystrophy linked to 4q12. Nat Genet 11:257–265

McNally EM, Yoshida M, Mizuno Y, Ozawa E, Kunkel LM (1994) Human adhalin is alternatively spliced and the gene is located on chromosome 17q21. Proc Natl Acad Sci USA 91:9690–9694

Matsumura K, Tomé FMS, Collin H, Azibi K, Chaouch M, Kaplan JK, Fardeau M, et al (1992) Deficiency of the 50K dystrophin-associated glycoprotein in severe childhood autosomal recessive muscular dystrophy. Nature 359:320–322

Nigro V, de Sà Moreira E, Piluso G, Vainzof M, Belsito A, Politano L, Puca AA, et al (1996*a*) Autosomal recessive limb-girdle muscular dystrophy, LGMD2F, is caused by a mutation in the δ-sarcoglycan gene. Nat Genet 14:195–198

Nigro V, Piluso G, Besito A, Politano L, Puca AA, Papparella S, Rossi E, et al (1996*b*) Identification of a novel sarcoglycan gene at 5q33 encoding a sarcolemmal 35 kDa glycoprotein. Hum Mol Genet 5:1179–1186

Noguchi S, McNally EM, Ben Othmane K, Hagiwara Y, Mizuno Y, Yoshida M, Yamamoto H, et al (1995) Mutations in the dystrophin-associated protein γ-sarcoglycan in chromosome 13 muscular dystrophy. Science 270:819–822

Passos-Bueno MR, Bashir R, Moreira ES, Vainzof M, Marie SK, Vasquez L, Iughetti P, et al (1995) Confirmation of the 2p locus for the mild autosomal recessive limb-girdle muscular dystrophy gene (LGMD2B) in three families allows refinement of the candidate region. Genomics 27: 192–195

Passos-Bueno MR, Moreira ES, Vainzof M, Marie SK, Zatz M (1996) Linkage analysis in autosomal recessive limb-girdle muscular dystrophy (AR LGMD) maps a sixth form to 5q33-34 (LGMD2F) and indicates that there is at least one more subtype of AR LGMD. Hum Mol Genet 5:815–820

Passos-Bueno MR, Richard I, Vainzof M, Fougerousse F, Weissenbach J, Broux O, Cohen D, et al (1993) Evidence of genetic heterogeneity for the autosomal recessive adult forms of limb-girdle muscular dystrophy following linkage analysis with 15q probes in Brazilian families. J Med Genet 30:385–387

Piccolo F, Roberds SL, Jeanpierre M, Leturcq F, Azibi K, Beldjord C, Carrié A, et al (1995) Primary adhalinopathy: a common cause of autosomal recessive muscular dystrophy of variable severity. Nat Genet 10:243–245

Richard I, Beckmann JS (1995) How neutral are synonymous codon mutations? Nat Genet 10:259

——— (1996) Molecular cloning of mouse canp3, the gene associated with limb-girdle muscular dystrophy 2A in human. Mamm Genome 7:377–379

Richard I, Broux O, Allamand V, Fougerousse F, Chiannilkulchai N, Bourg N, Brenguier L, et al (1995) Mutations in the proteolytic enzyme calpain 3 cause limb-girdle muscular dystrophy type 2A. Cell 81:27–40

Roberds SL, Leturcq F, Allamand V, Piccolo F, Jeanpierre M, Anderson RD, Lim LE, et al (1994) Missense mutations in the adhalin gene linked to autosomal recessive muscular dystrophy. Cell 78:625–633

Sorimachi H, Imajoh-Ohmi S, Emori Y, Kawasaki H, Ohno S, Minami Y, Suzuki K (1989) Molecular cloning of a novel mammalian calcium-dependant protease distinct from both m- and mu-type: specific expression of the mRNA in skeletal muscle. J Biol Chem 264:20106–20111

Sorimachi H, Suzuki K (1992) Sequence comparison among muscle-specific calpain P94 and calpain subunits. Biochim Biophys Acta 1160:55–62

Walton JN, Natrass FJ (1954) On the classification, natural history and treatment of the myopathies. Brain 77:169–231