Research article

Genomic organization and single-nucleotide polymorphism map of desmuslin, a novel intermediate filament protein on chromosome 15q26.3

Yuji Mizuno[1,2], Annibale A Puca[1], Kristine F O'Brien[1], Alan H Beggs[3] and Louis M Kunkel*[1]

Address: [1]Howard Hughes Medical Institute, Children's Hospital and Harvard Medical School, Boston, Massachusetts 02115, USA, [2]Current address: Department of Neurology, Gunma University School of Medicine, 3-39-22 Showa, Maebashi, Gunma 371-8511, Japan, and [3]Division of Genetics, Children's Hospital and Harvard Medical School, Boston, Massachusetts, 02115,USA

E-mail: Yuji Mizuno - mizunoy@med.gunma-u.ac.jp; Annibale A Puca - apuca@rascal.med.harvard.edu; Kristine F O'Brien - kfobrien@mac.com; Alan H Beggs - beggs@rascal.med.harvard.edu; Louis M Kunkel* - kunkel@rascal.med.harvard.edu
*Corresponding author

Abstract

Background: Desmuslin is an α-dystrobrevin-interacting protein expressed primarily in heart and skeletal muscle. The desmuslin protein interacts with and is closely related to desmin, a protein encoded by a locus mutated in some forms of hereditary distal myopathy. As a muscle-specific intermediate filament protein, desmuslin is also a candidate for myopathies of unknown etiology.

Results: The desmuslin gene was localized to chromosome 15q26.3 by electronic screening of the human DNA sequence database. Primer pairs were designed to amplify the 5 exons of the desmuslin gene in 11 overlapping DNA segments. The desmuslin gene was screened for mutations in 71 patients with various forms of myopathy for which there was no known cause. In this analysis, 10 common and 2 rare amino acid altering single-nucleotide polymorphisms were identified, all of which were seen in a control population of individuals thus making these unlikely causes of the phenotype. Interestingly, one of the single-nucleotide polymorphisms found in a patient resulted in a premature stop codon in the first exon. The nonsense mutation was also detected in the patient's unaffected father and one unaffected control; it was detected in 0.44% (2/454) of unrelated chromosomes and is therefore predicted to have a homozygous frequency of 0.002%.

Conclusion: No causative mutations were found in the desmuslin gene. However, the single-nucleotide polymorphisms mapped in this study represent a well-mapped group that can be used for disequilibrium studies of this region of chromosome 15q26.3.

Background

Dystrophin and its associated proteins are thought to be involved in the anchoring of the muscle cell membrane to the extracellular matrix [1], and the absence of many of these proteins can lead to the phenotype of muscular dystrophy [2]. The dystrophin-associated protein com-plex (DAPC) consists of several subgroups of protein complexes, each associated either directly or indirectly with dystrophin. The sarcoglycans are four transmem-brane proteins [3] that are organized by a fifth protein called sarcospan [4]. This complex is thought to be in-volved in signalling at the cell membrane [5]. A second

subcomplex, known as the dystroglycan complex [6], interacts directly with dystrophin in the cytoplasm and laminin in the extracellular matrix, thus providing a structural link between the inside and the outside of the cell. A third subcomplex involves the dystrobrevins [7, 8] and syntrophins [9] which form a complex at the C-terminal region of dystrophin [10,11,12] and have an as yet unidentified function.

Desmuslin (DMN) was recently identified as an α-dystrobrevin-interacting protein via the yeast two-hybrid method [13]. Both desmuslin mRNA and protein are expressed primarily in cardiac and skeletal muscle and the gene encodes a novel intermediate filament (IF) protein of 1253 amino acids. Electron microscopic analysis shows that desmuslin colocalizes with desmin, another muscle IF protein. Co-immunoprecipitation experiments revealed that the desmuslin and α-dystrobrevin interaction involves the region of protein encoded by exons 8-16 of α-dystrobrevin and domains 1A through 2A of the desmuslin rod domain. Desmuslin is hypothesized to function as a mechanical support to the muscle myofibers by making a previously unrecognized linkage between the extracellular matrix via the DAPC and the Z-discs through desmin and plectin [14].

As several IF proteins, including desmin, have been implicated in human genetic disorders such as dominant and recessive congenital and adult onset myopathies [15,16,17,18], desmuslin becomes a candidate to be involved in myopathies as well. Supporting this is the exclusive expression of DMN in skeletal and cardiac muscle. The DMN gene was analyzed for mutations in 71 patients with various forms of muscular dystrophy and myopathy. In addition to 9 single-nucleotide polymorphisms (SNPs) that do not change the protein sequence, we identified 12 SNPs that do alter the residue they encode. Although examination of controls has shown that none of them is likely to cause the phenotype, our results are useful for future disequilibrium studies of this region of chromosome 15q26.3 and for mutation analysis and association studies in other genetic disorders.

Results and Discussion

Using the DMN cDNA sequence (AF359284) as a query to blast the public human genome database, we found a genomic clone (AC018999) that included the entire desmuslin coding sequence. The alignment identified 5 exons which are schematically indicated in Figure 1. Eleven sets of primer pairs were developed to amplify the 5 exons and the flanking intron sequences (Table 1). The primers were designed to amplify overlapping segments

of the desmuslin coding sequence as indicated in Figure 1.

Table 1: Eleven PCR primer sets that allow the amplification of the entire coding sequence and surrounding intronic sequence of the DMN gene.

Primer	Primer Sequence
1F	5' - ATCCACCAAGGCCAGGACACTCGTATGCTCC
1R	5' - TCGTCCTCGTACAGCTGCAC
2F	5' - GGAGAGGACGAGACCGGGACAAGACCA
2R	5' - CGGAAATGCATGGTAAGGCTG
3F	5' - GAGCTGCAGCGCCTGGATGCGGAGGAG
3R	5' - GTTAAAGCAGCTGCCACCTAACCCCATGAC-GAGGA
4F	5' - GGAAACCATCCTGCAAGTGACTGCATG
4R	5' - GAGACACCTGCTCAGTGACCTAAGC
5F	5' - GGTGACACTATGGTCATTGGCCGCATACCAC
5R	5' - CCAGGGGTGTGAAACACAAGTTCAGGTGCAC
6F	5' - CACAAAAACGAGCACCTCTGACCTCTCTGGG
6R	5' - GAATGACGGTCCTCTCATTTGACCG
7F	5' - CCAACGTCAGAACTTTCTCTCCAACC
7R	5' - CCAGGCCAGCTTCATCGGAAACATC
8F	5' - GTCACTTACGTGGACAGGAAAGAGC
8R	5' - GACTGTGGGTTCGGCATGAAATTC
9F	5' - GGGCGAGCAGAGTTATTTTGTGTCC
9R	5' - CTCCATCTCACTGGCCTCATCTTTGC
10F	5' - GAGCGTTTCCGTGGATGTCAAGAAG
10R	5' - GATGACTTCTCTGGACAGTGGACTTTGACCGG
11F	5' - GAGCGTTTCCGTGGATGTCAAGAAG
11R	5' - GTTAAGGCCTTTGGATGTTGGCACG

Table 2: Summary of clinical diagnoses and pathological findings of patient samples studied for alterations in DMN.

Clinical Diagnosis	Description	# of samples
Congenital Myopathy	Nemaline Myopathy	50
	Multicore/minicore Disease	6
	Myotubular Myopathy	2
	Spheroid Body Myopathy	1
	Desmuslin accumulation	2
	Undefined	2
Muscular Dystrophy	Limb-Girdle MD	4
	Undefined	4

Desmin, a desmuslin-interacting protein, has been implicated in hereditary distal myopathies implying that desmuslin itself may also be involved in myopathies of unknown etiology. DNA samples isolated from 71 dystrophic and myopathic patients (as described in Table 2)
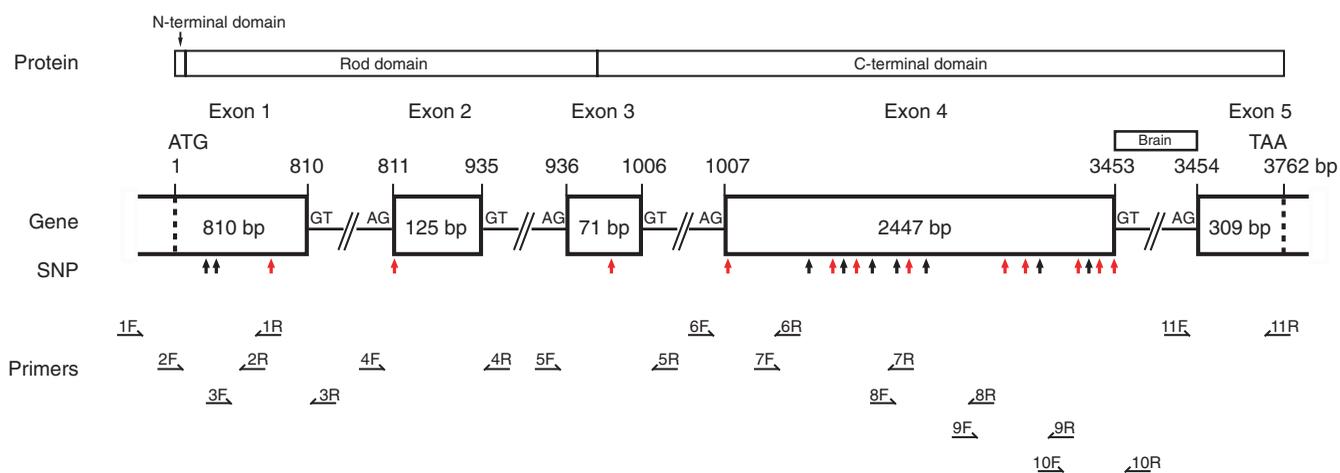
Figure 1
Schematic representation of the desmuslin gene and protein. Desmuslin protein structure, genomic organization, approximate location of SNPs identified in this study and primer pairs used to amplify the entire coding sequence are indicated. A comparison of the *DMN* cDNA sequence with its related genomic DNA led to the identification of 5 exons ranging from 71 bp (exon 3) to 2447 bp (exon 4). The exact start of transcription in exon 1 and the termination of transcription at the end of exon 5 are unknown. Unlike the brain *DMN* cDNA, muscle *DMN* cDNA does not contain the sequence between exons 4 and 5 (indicated by an open box). The locations of the 21 SNPs are indicated by arrows; the 12 SNPs that alter an amino acid codon are shown in red and the 9 silent SNPs are shown in black. The exact position of each of the SNPs is described in Table 3. The positions of the 11 primer pairs are also shown below the respective exons.

were analyzed by direct sequencing to determine whether the phenotype was the result of a desmuslin mutation. All observed variants were subsequently tested in ≤ 156 control individuals. These analyses revealed 10 common and 2 rare SNPs [19, 20] that alter an amino acid codon; 7 result in non-conservative amino acids changes (Table 3).

None of the SNPs seem associated with the patient phenotypes as each was also found in the control population. Nine silent SNPs, which did not alter an amino acid, were also identified in the patient population; several of these SNPs showed a high degree of heterozygosity. Most were tested for and detected in the control population, thus making desmuslin an unlikely candidate to be involved in these myopathies, but this first pass can be followed by others in a larger set of patient samples.

Interestingly, SNP 3, a C598T substitution resulting in a premature stop codon in exon 1, was detected in a single Nemaline myopathy patient. Neither the proband's mother nor any of the other patients was found to carry this mutation (Figure 2A). The C598T mutation removed a Pvu II site, thus permitting detection of the alteration by restriction digest analysis of a PCR product generated from genomic DNA. The assay is outlined schematically in Figure 2B and documented in Figure 2C. Using this Pvu II assay, the C598T change was detected in the unaffected father of the patient, making it unlikely that heter-

ozygosity for the C598T change is the causative mutation in the myopathic proband.

The removal of the Pvu II site allowed rapid screening of additional patients and control samples by DNA amplification of the region and subsequent digestion of the PCR product with Pvu II. One of 312 control chromosomes was found to have lost the Pvu II site, and the heterozygous presence of the C598T mutation was confirmed by DNA sequencing of the PCR product. Thus it is likely that the C598T mutation is not responsible for the myopathic phenotype of the proband. This is a surprising result considering that a single copy of this stop codon is present in 0.44% (2/454) of the population and 0.002% of the population is predicted to carry this nonsense mutation in a homozygous state. The phenotype of the latter individuals is not known.

There is the possibility that either the heterozygous C598T *DMN* mutation combined with another mutation in a different gene or the C598T in a homozygous state causes muscular dystrophy or myopathy of unknown genetic pathogenesis. The C598T *DMN* mutation is a good candidate to be a modifier for muscular and/or cardiac phenotypes, considering: (i) the frequency of the mutation in the normal population in the heterozygous as well as homozygous state; and (ii) the key role of desmuslin in connecting the Z-discs to the extracellular matrix. The desmuslin SNPs should be examined in a larger set of
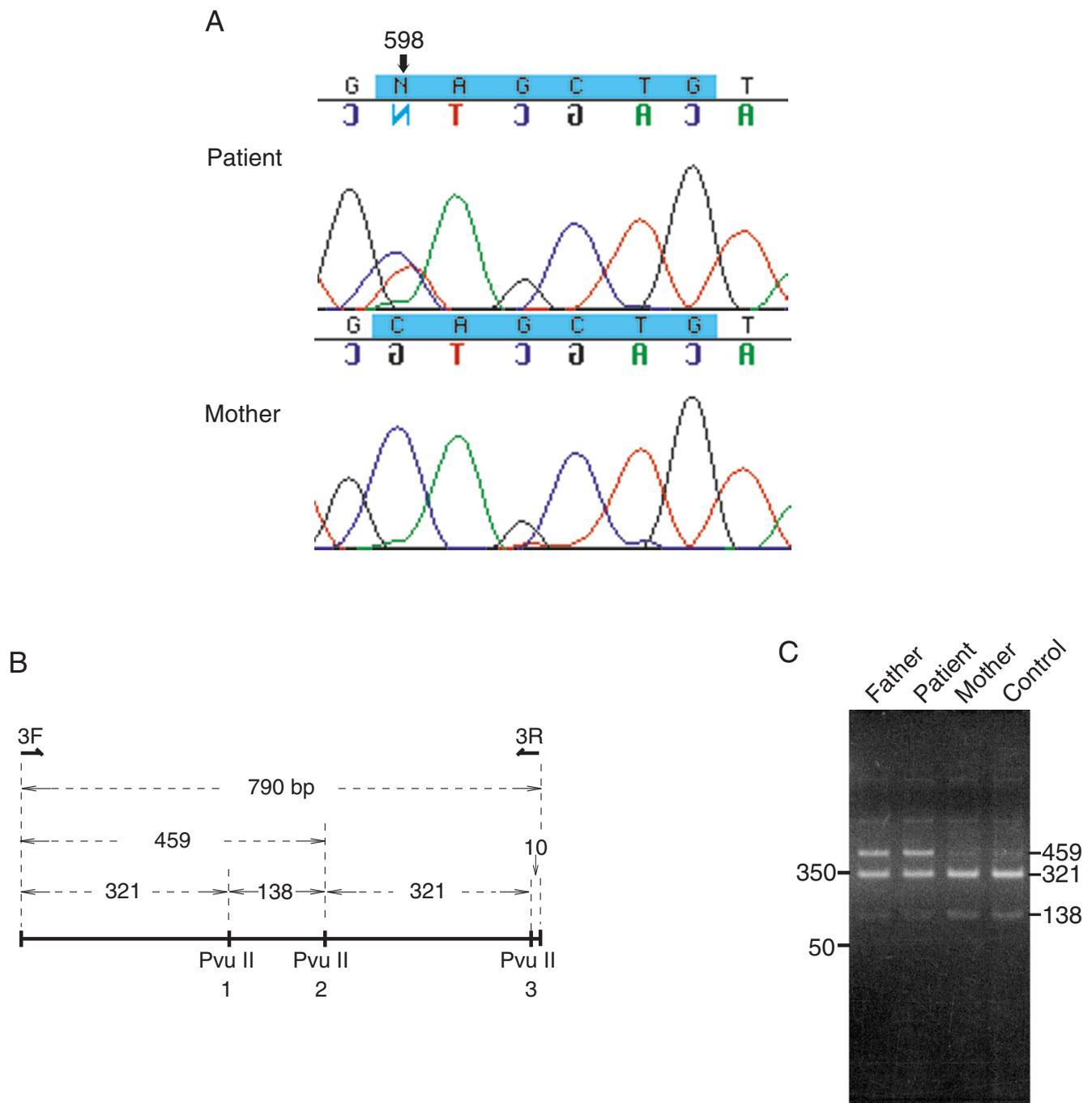
Figure 2
Analysis of SNP3 (C598T) in patient and control samples. (A) Sequence results of the C598T patient and his mother. At nucleotide number 598, the patient has a C and T, indicated by N. The patient's unaffected mother is homozygous for C. (B) Pvu II site amplified with 3F and 3R. The size of PCR product amplified with 3F and 3R is 790 bp and contains three Pvu II restriction sites. When the PCR product from a control subject is digested with Pvu II, four products of 321 bp, two of 138 bp and two of 10 bp are generated. Since the C598T removes the first Pvu II restriction site, digestion of the amplification product will produce a fragment of 459 bp instead of 321 bp and 138 bp. (C) Electrophoresis of PCR product digested with Pvu II shows that the patient and his unaffected father both have a band of 459 bp. The 138 bp fragment in the mother and control is twice as strong as those of the patient and the father. Similarly the 321 bp band in the mother and control is stronger than in the patient and his father.

myopathy and muscular dystrophy patients to test this hypothesis. More specifically, it should be determined whether there is a statistically significant difference in the frequency of individual or clustered SNP haplotypes between the patient and the control populations.

The desmuslin coding sequence is highly polymorphic with 9 SNPs that do not alter an amino acid and 12 that do. Ten of these variations have minor allele frequencies that are represented in over 10% of the chromosomes

sampled. These 10 are spaced within a genomic interval of ≤ 28 kilobases, or approximately 1 per 2800 bases. These highly informative markers are thus very useful for genetic analysis. The presence of so many SNPs within the coding sequence of a protein is somewhat unexpected and is much greater than that documented for other genes. This degree of variability may reflect a redundancy of desmuslin at the Z-line where other IF proteins with potentially overlapping functions are located [21].

Table 3: Summary of SNPs in the *DMN* gene.

SNP	Exon	Domain	Nucleotide Position	Nucleotide Change	Amino Acid Position		Amino Acid Change	# of chromosomes	
								Total	%
1	1	Rod	408	CG**G**	136		none	91	64.1
				CG**A**				51	35.9
2	1	Rod	457	**A**GG	153		none	91	64.1
				CGG				51	35.9
3	1	Rod	598	**C**AG	200	Gln	non-conservative	452	99.6
				TAG		STOP		2	0.4
4	2	Rod	815	GC**G**	272	Ala	conservative	103	57.9
				GT**G**		Val		75	42.1
5	3	C-terminal	988	**G**TT	330	Val	conservative	165	90.7
				ATT		Ile		17	9.3
6	4	C-terminal	1063	**A**GG	335	Arg	non-conservative	79	87.8
				TGG		Trp		11	12.2
7	4	C-terminal	1596	TC**C**	532		none	128	91.4
				TC**T**				12	8.6
8	4	C-terminal	1700	CC**G**	567	Pro	conservative	149	81.9
				C**T**G		Leu		33	18.1
9	4	C-terminal	1713	AG**C**	571		none	162	89.0
				AG**T**				20	11.0
10	4	C-terminal	1835	G**A**A	612	Glu	non-conservative	180	97.8
				G**C**A		Ala		4	2.2
11	4	C-terminal	1953	AC**T**	651		none	131	72.8
				AC**C**				49	27.2

Table 3: Summary of SNPs in the*DMN* gene. *(Continued)*

12	4	C-terminal	2010	TA**C**	670		none	176	97.8
				TA**T**				4	2.2
13	4	C-terminal	2282	C**C**G	761	Pro	conservative	114	91.9
				C**T**G		Leu		10	8.1
14	4	C-terminal	2424	CC**C**	808		none	63	75.0
				CC**T**				21	25.0
15	4	C-terminal	2836	**C**GG	946	Arg	non-conservative	324	99.4
				TGG		Trp		2	0.6
16	4	C-terminal	2927	C**A**G	976	Gln	non-conservative	187	98.4
				C**G**G		Arg		3	1.6
17	4	C-terminal	2937	CC**G**	979		none	187	99.5
				CC**A**				1	0.5
18	4	C-terminal	3176	C**C**G	1059	Pro	conservative	126	98.4
				C**T**G		Leu		2	1.6
19	4	C-terminal	3195	TT**C**	1065		none	95	76.6
				TT**T**				29	23.4
20	4	C-terminal	3200	C**G**T	1067	Arg	non-conservative	114	91.9
				C**C**T		Pro		10	8.1
21	4	C-terminal	3230	T**C**A	1077	Ser	non-conservative	115	89.8
				T**T**A		Leu		13	10.2

The details of 21 SNPs are shown. The exon and domain location as well as the nucleotide and amino acid positions of SNP are given. Also listed are the major and minor alleles, both as nucleotide and amino acid, and the corresponding frequencies, in terms of numbers of chromosomes and percentage. The nucleotide substitution is highlighted in bold. The amino acid change indicates the degree of alteration in the encoded residue; the change is described as none if there was no change, conservative if the alteration does not change the residue's polarity or charge, and non-conservative if the polarity and/or charge of the encoded residues differ.

Conclusions

The characterization of the individual components that form protein complexes in skeletal muscle and heart, especially information concerning genotype-phenotype correlations, will increase our understanding of the pathophysiology of human muscle diseases. We encourage other groups to test for the presence of the C598T *DMN* mutation in their human patient samples affected by muscular and cardiac diseases. The generation of des-muslin null animal models will also contribute to the understanding of the role of this protein in muscle and cardiac disease.

Materials and Methods

Blood samples were obtained from patients and control subjects under IRB-approved informed consent. The genomic DNA extracted from these blood samples was used as the template for the PCR assay. Using an Advan-

tage-GC cDNA PCR polymerase (Clontech) and *Pfu* DNA polymerase (Stratagene), PCR was carried out with primers specific to the exon and/or intron of *DMN* sequence (Table 1) under the following conditions: 94°C for 1 min followed by 35 cycles of 94°C for 1 min, 57-59 °C for 1 min, 72°C for 3 min and a final extension time of 10 min at 72°C. The amplification products were purified using the QIAGEN PCR purification kit and analyzed on either an ABI 373 or 377 automated sequencer with fluorescent dye terminator chemistry (Applied Biosystems).

The PCR products were also digested with Pvu II and Msp I in order to screen for the C598T and C2836T SNPs, respectively.

Abbreviations
DAPC: Dystrophin-associated protein complex; DMN: Desmuslin, SNP: Single-nucleotide polymorphism; IF: Intermediate filament

References
1. Yoshida M, Suzuki A, Yamamoto H, Noguchi S, Mizuno Y, Ozawa E: **Dissociation of the complex of dystrophin and its associated proteins into several unique groups by n-octyl beta-D-glucoside.** *Eur J Biochem* 1994, **222**:1055-1061
2. Ozawa E, Nishino I, Nonaka I: **Sarcolemmopathy: Muscular Dystrophies with Cell Membrane Defects.** *Brain Pathol* 2001, **11**:218-230
3. Ozawa E, Noguchi S, Mizuno Y, Hagiwara Y, Yoshida M: **Invited review: From dystrophinopathy to sarcoglycanopathy: Evolution of a concept of muscular dystrophy.** *Muscle Nerve* 1998, **21**:421-438
4. Crosbie RH, Heighway J, Venzke DP, Lee JC, Campbell KP: **Sarcospan, the 25-kDa transmembrane component of the dystrophin-glycoprote in complex.** *J Biol Chem* 1997, **272**:31221-31224
5. Thompson TG, Chan YM, Hack AA, Brosius M, Rajala M, Lidov HG, McNally EM, Watkins S, Kunkel LM: **Filamin 2 (FLN2): A muscle-specific sarcoglycan interacting protein.** *J Cell Biol* 2000, **148**:115-126
6. Ibraghimov-Beskrovnaya O, Ervasti JM, Leveille CJ, Slaughter CA, Sernett SW, Campbell KP: **Primary structure of dystrophin-associated glycoproteins linking dystrophin to the extracellular matrix.** *Nature* 1992, **355**:696-702
7. Sadoulet-Puccio HM, Khurana TS, Cohen JB, Kunkel LM: **Cloning and characterization of the human homologue of a dystrophin related phosphoprotein found at the Torpedo electric organ post-synaptic membrane.** *Hum Mol Genet* 1996, **5**:489-496
8. Sadoulet-Puccio HM, Kunkel LM: **Dystrophin and its isoforms.** *Brain Pathol* 1996, **6**:25-35
9. Ahn AH, Feener CA, Gussoni E, Yoshida M, Ozawa E, Kunkel LM: **The three human syntrophin genes are expressed in diverse tissues, have distinct chromosomal location and each bind to dystrophin and its relatives.** *J Biol Chem* 1996, **271**:2724-2730
10. Ahn A, Kunkel L: **Syntrophin binds to an alternatively spliced exon of dystrophin.** *J Cell Biol* 1995, **128**:363-371
11. Sadoulet-Puccio HM, Rajala M, Kunkel LM: **Dystrobrevin and dystrophin: An interaction through coiled-coil motifs.** *Proc Natl Acad Sci U S A* 1997, **94**:12413-12418
12. Suzuki A, Yoshida M, Ozawa E: **Mammalian alpha 1- and beta 1-syntrophin bind to the alternative splice- prone region of the dystrophin COOH terminus.** *J Cell Biol* 1995, **128**:373-381
13. Mizuno Y, Thompson TG, Guyon JR, Lidov HG, Brosius M, Imamura M, Ozawa E, Watkins SC, Kunkel LM: **Desmuslin, an intermediate filament protein that interacts with alpha - dystrobrevin and desmin.** *Proc Natl Acad Sci U S A* 2001, **98**:6156-6161
14. Hijikata T, Murakami T, Imamura M, Fujimaki N, Ishikawa H: **Plectin is a linker of intermediate filaments to Z-discs in skeletal muscle fibers.** *J Cell Sci* 1999, **112**:867-876
15. Dalakas MC, Park KY, Semino-Mora C, Lee HS, Sivakumar K, Goldfarb LG: **Desmin myopathy, a skeletal myopathy with cardiomyopathy caused by mutations in the desmin gene.** *N Engl J Med* 2000, **342**:770-780
16. Goldfarb LG, Park KY, Cervenakova L, Gorokhova S, Lee HS, Vasconcelos O, Nagle JW, Semino-Mora C, Sivakumar K, Dalakas MC: **Missense mutations in desmin associated with familial cardiac and skeletal myopathy.** *Nat Genet* 1998, **19**:402-403
17. Munoz-Marmol AM, Strasser G, Isamat M, Coulombe PA, Yang Y, Roca X, Vela E, Mate JL, Coll J, Fernandez-Figueras MT, Navas-Palacios JJ, Ariza A, Fuchs E: **A dysfunctional desmin mutation in a patient with severe generalized myopathy.** *Proc Natl Acad Sci U S A* 1998, **95**:11312-11317
18. Sjoberg G, Saavedra-Matiz CA, Rosen DR, Wijsman EM, Borg K, Horowitz SH, Sejersen T: **A missense mutation in the desmin rod domain is associated with autosomal dominant distal myopathy, and exerts a dominant negative effect on filament formation.** *Hum Mol Genet* 1999, **8**:2191-2198
19. Cargill M, Altshuler D, Ireland J, Sklar P, Ardlie K, Patil N, Shaw N, Lane CR, Lim EP, Kalyanaraman N, Nemesh J, Ziaugra L, Friedland L, Rolfe A, Warrington J, Lipshutz R, Daley GQ, Lander ES: **Characterization of single-nucleotide polymorphisms in coding regions of human genes.** *Nat Genet* 1999, **22**:231-238
20. Halushka MK, Fan JB, Bentley K, Hsie L, Shen N, Weder A, Cooper R, Lipshutz R, Chakravarti A: **Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis.** *Nat Genet* 1999, **22**:239-247
21. North KN, Yang N, Wattanasirichaigoon D, Mills M, Easteal S, Beggs AH: **A common nonsense mutation results in alpha-actinin-3 deficiency in the general population.** *Nat Genet* 1999, **21**:353-354