UMD-DYSF database at a glance

- The **UMD-DYSF Locus Specific Database** has been compiled to provide up-to-date information about mutations of the *DYSF* gene. It aims at making the information readily accessible to anyone interested in the genetic variations of the *DYSF* gene, and to provide an easy way for those who investigate these variations to report their most recent findings. UMD-DYSF includes interactive analysis of disease-causing mutation statistics and distribution, and bioinformatics tools for the interpretation of novel sequence variants.

- **Documentation on the *DYSF* gene** (gene characteristics, sequence, exonic organisation and orthologues) can be accessed through 

page.

- **Documentation on the dysferlin protein** (structural organisation, tissue distribution, functional roles, known interacting proteins or mice models for dysferlinopathies) can be accessed through 

page.

- **Clinical characteristics of dysferlinopathies** are summarised on 

page.

- The 

page describes **key UMD tools** and propose a **search module**.

- **Additional UMD tools** are available on the 

page.

- All **references** entered in UMD-DYSF can be accessed from the 

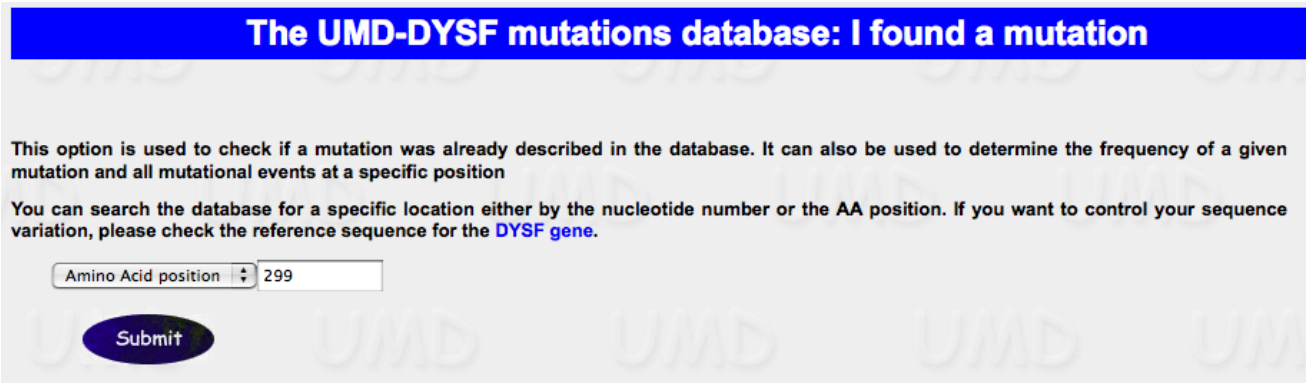
page.

- The next pages will guide you through the UMD-DYSF database with five examples of questions you might want to answer :
  I. I found a sequence variation: how can I check if it has already been reported in the literature?
  II. How can I identify all reported mutations localized in a specific *DYSF* exon?
  III. How can I see at a glance all mutations included in the UMD-DYSF database?
  IV. I found a missense variation: how can I evaluate its pathogenicity?
  V. How can I evaluate a possible deleterious effect of a sequence variant on normal splicing?

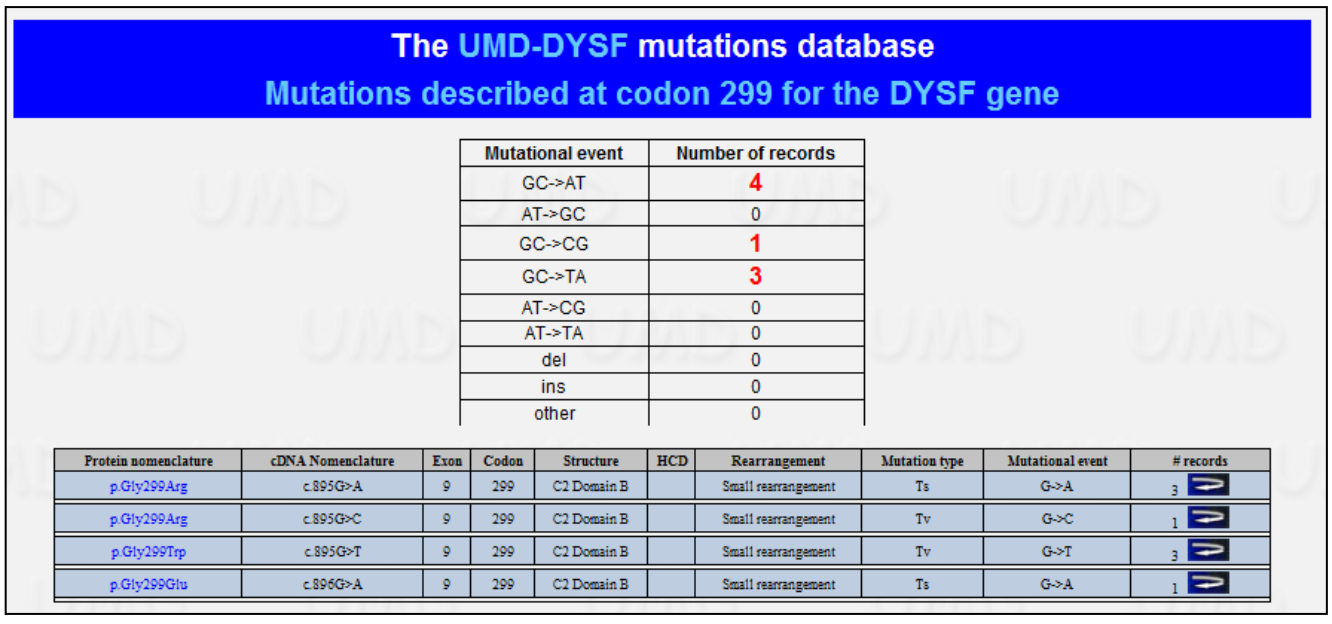
I. I found a sequence variation and want to check if it has already been reported in the literature

1. Go to the Mutations page using the button on the left panel, and select * I Found a Mutation

2. Enter your amino-acid or nucleotide position and submit:



3. The result page for amino acid 299 indicates eight reported records corresponding to four different mutational events.



The UMD-DYSF mutations database

Mutations described at codon 299 for the DYSF gene

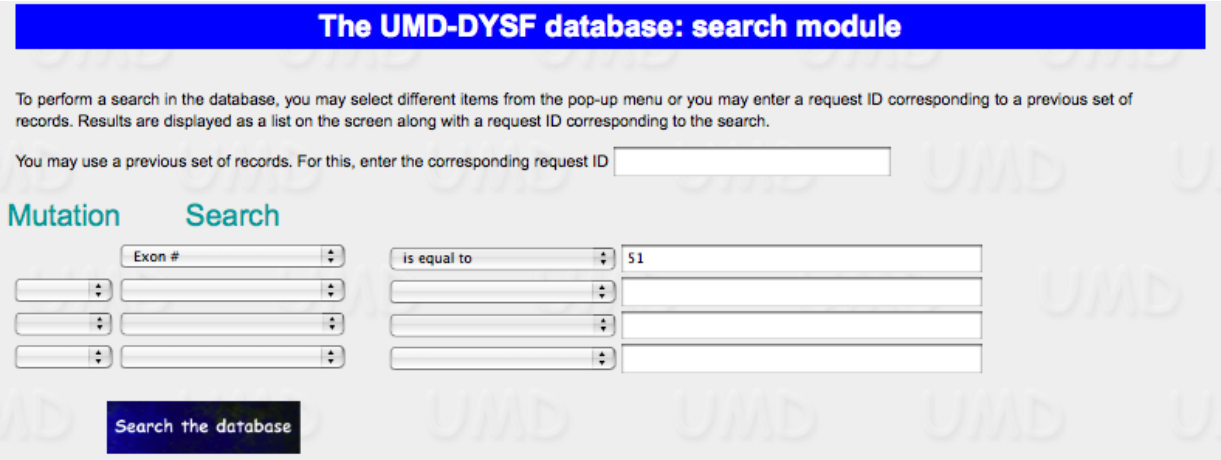
Mutational event	Number of records
GC->AT	4
AT->GC	0
GC->CG	1
GC->TA	3
AT->CG	0
AT->TA	0
del	0
ins	0
other	0

Protein nomenclature	cDNA Nomenclature	Exon	Codon	Structure	HCD	Rearrangement	Mutation type	Mutational event	# records
p.Gly299Arg	c.895G>A	9	299	C2 Domain B		Small rearrangement	Ts	G->A	3
p.Gly299Arg	c.895G>C	9	299	C2 Domain B		Small rearrangement	Tv	G->C	1
p.Gly299Trp	c.895G>T	9	299	C2 Domain B		Small rearrangement	Tv	G->T	3
p.Gly299Glu	c.896G>A	9	299	C2 Domain B		Small rearrangement	Ts	G->A	1

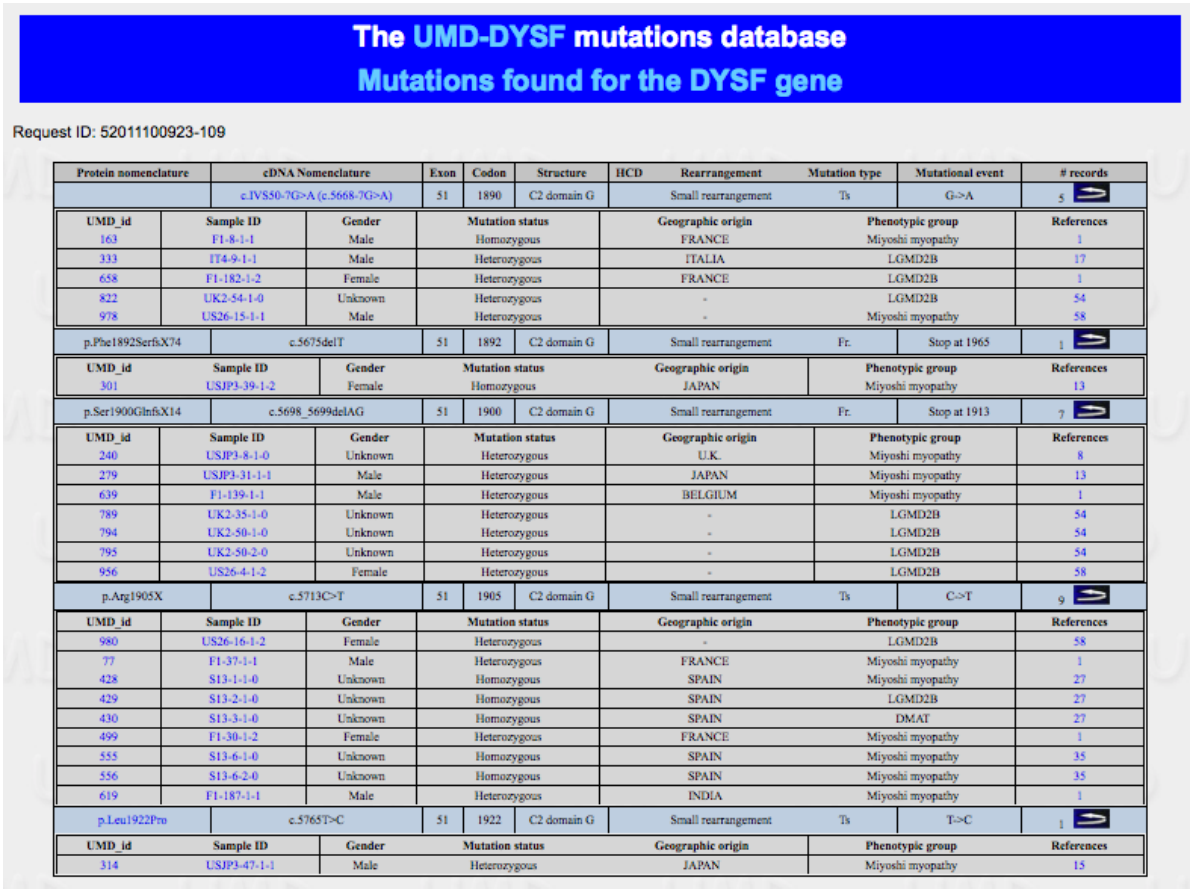
II. I want to identify all mutations localized in exon 51 and reported in the literature

1. Go to the Mutations page using the button on the left panel, and select * I Want to Search the Database

2. Submit your search « Exon# is equal to 51 »



3. The result page shows five reported mutational events in exon 51 and more details can be obtained by clicking on any item in the #records column (see detailed results below).



The UMD-DYSF mutations database

Mutations found for the DYSF gene

Request ID: 52011100923-109

Protein nomenclature	cDNA Nomenclature	Exon	Codon	Structure	HCD	Rearrangement	Mutation type	Mutational event	# records
	c.IVS50-7G>A (c.5668-7G>A)	51	1890	C2 domain G		Small rearrangement	Ts	G->A	5

UMD_id	Sample ID	Gender	Mutation status	Geographic origin	Phenotypic group	References
163	FI-8-1-1	Male	Homozygous	FRANCE	Miyoshi myopathy	1
333	IT4-9-1-1	Male	Heterozygous	ITALIA	LGMD2B	17
658	FI-182-1-2	Female	Heterozygous	FRANCE	LGMD2B	1
822	UK2-54-1-0	Unknown	Heterozygous	-	LGMD2B	54
978	US26-15-1-1	Male	Heterozygous	-	Miyoshi myopathy	58

Protein nomenclature	cDNA Nomenclature	Exon	Codon	Structure	HCD	Rearrangement	Mutation type	Mutational event	# records
p.Phe1892SerfsX74	c.5675delT	51	1892	C2 domain G		Small rearrangement	Fr.	Stop at 1965	1

UMD_id	Sample ID	Gender	Mutation status	Geographic origin	Phenotypic group	References
301	USJP3-39-1-2	Female	Homozygous	JAPAN	Miyoshi myopathy	13

Protein nomenclature	cDNA Nomenclature	Exon	Codon	Structure	HCD	Rearrangement	Mutation type	Mutational event	# records
p.Ser1900GlnfsX14	c.5698_5699delAG	51	1900	C2 domain G		Small rearrangement	Fr.	Stop at 1913	7

UMD_id	Sample ID	Gender	Mutation status	Geographic origin	Phenotypic group	References
240	USJP3-8-1-0	Unknown	Heterozygous	U.K.	Miyoshi myopathy	8
279	USJP3-31-1-1	Male	Heterozygous	JAPAN	Miyoshi myopathy	13
639	FI-139-1-1	Male	Heterozygous	BELGIUM	Miyoshi myopathy	1
789	UK2-35-1-0	Unknown	Heterozygous	-	LGMD2B	54
794	UK2-50-1-0	Unknown	Heterozygous	-	LGMD2B	54
795	UK2-50-2-0	Unknown	Heterozygous	-	LGMD2B	54
956	US26-4-1-2	Female	Heterozygous	-	LGMD2B	58

Protein nomenclature	cDNA Nomenclature	Exon	Codon	Structure	HCD	Rearrangement	Mutation type	Mutational event	# records
p.Arg1905X	c.5713C>T	51	1905	C2 domain G		Small rearrangement	Ts	C->T	9

UMD_id	Sample ID	Gender	Mutation status	Geographic origin	Phenotypic group	References
980	US26-16-1-2	Female	Heterozygous	-	LGMD2B	58
77	FI-37-1-1	Male	Heterozygous	FRANCE	Miyoshi myopathy	1
428	S13-1-1-0	Unknown	Homozygous	SPAIN	Miyoshi myopathy	27
429	S13-2-1-0	Unknown	Homozygous	SPAIN	LGMD2B	27
430	S13-3-1-0	Unknown	Homozygous	SPAIN	DMAT	27
499	FI-30-1-2	Female	Heterozygous	FRANCE	Miyoshi myopathy	1
555	S13-6-1-0	Unknown	Homozygous	SPAIN	Miyoshi myopathy	35
556	S13-6-2-0	Unknown	Homozygous	SPAIN	Miyoshi myopathy	35
619	FI-187-1-1	Male	Heterozygous	INDIA	Miyoshi myopathy	1

Protein nomenclature	cDNA Nomenclature	Exon	Codon	Structure	HCD	Rearrangement	Mutation type	Mutational event	# records
p.Leu1922Pro	c.5765T>C	51	1922	C2 domain G		Small rearrangement	Ts	T->C	1

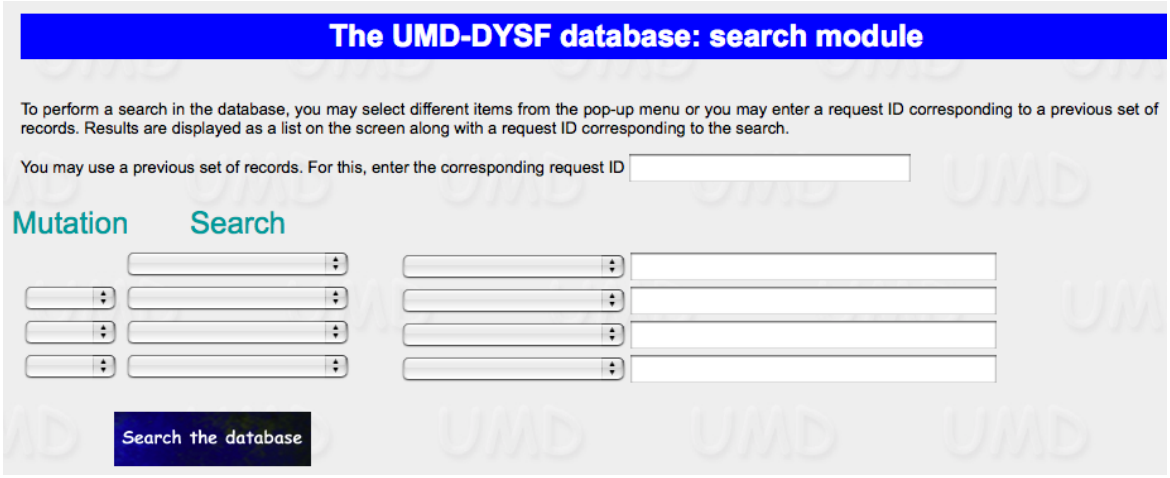
UMD_id	Sample ID	Gender	Mutation status	Geographic origin	Phenotypic group	References
314	USJP3-47-1-1	Male	Heterozygous	JAPAN	Miyoshi myopathy	15

III. I want to see at a glance all mutations included in the UMD-DYSF database

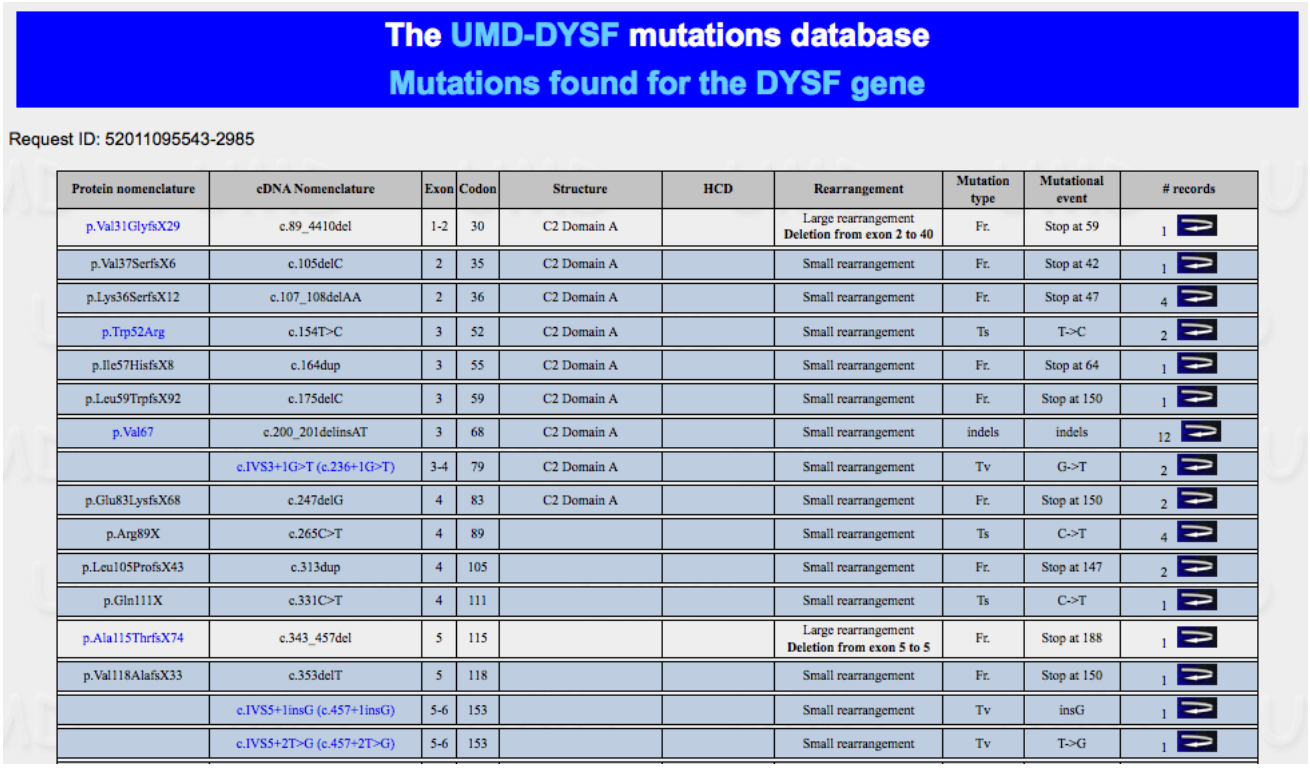
1. Go to the Mutations page using the button on the left panel, and select

* I Want to Search the Database

2. Submit the search « Exon# is not equal to» followed by an empty field.



3. Your search result will appear as follows:



The UMD-DYSF mutations database

Mutations found for the DYSF gene

Request ID: 52011095543-2985

Protein nomenclature	cDNA Nomenclature	Exon	Codon	Structure	HCD	Rearrangement	Mutation type	Mutational event	# records
p.Val31GlyfsX29	c.89_4410del	1-2	30	C2 Domain A		Large rearrangement Deletion from exon 2 to 40	Fr.	Stop at 59	1
p.Val37SerfsX6	c.105delC	2	35	C2 Domain A		Small rearrangement	Fr.	Stop at 42	1
p.Lys36SerfsX12	c.107_108delAA	2	36	C2 Domain A		Small rearrangement	Fr.	Stop at 47	4
p.Trp52Arg	c.154T>C	3	52	C2 Domain A		Small rearrangement	Ts	T->C	2
p.Ile57HisfsX8	c.164dup	3	55	C2 Domain A		Small rearrangement	Fr.	Stop at 64	1
p.Leu59TrpfsX92	c.175delC	3	59	C2 Domain A		Small rearrangement	Fr.	Stop at 150	1
p.Val67	c.200_201delinsAT	3	68	C2 Domain A		Small rearrangement	indels	indels	12
	c.IVS3+1G>T (c.236+1G>T)	3-4	79	C2 Domain A		Small rearrangement	Tv	G->T	2
p.Glu83LysfsX68	c.247delG	4	83	C2 Domain A		Small rearrangement	Fr.	Stop at 150	2
p.Arg89X	c.265C>T	4	89			Small rearrangement	Ts	C->T	4
p.Leu105ProfsX43	c.313dup	4	105			Small rearrangement	Fr.	Stop at 147	2
p.Gln111X	c.331C>T	4	111			Small rearrangement	Ts	C->T	1
p.Ala115ThrfsX74	c.343_457del	5	115			Large rearrangement Deletion from exon 5 to 5	Fr.	Stop at 188	1
p.Val118AlafsX33	c.353delT	5	118			Small rearrangement	Fr.	Stop at 150	1
	c.IVS5+1insG (c.457+1insG)	5-6	153			Small rearrangement	Tv	insG	1
	c.IVS5+2T>G (c.457+2T>G)	5-6	153			Small rearrangement	Tv	T->G	1

IV. I found a missense variation and want to evaluate its pathogenicity using the bio-informatics tool UMDpredictor®

1. Go to the Mutations page using the button on the left panel, and select

* I Want to Analyze the Impact of a Missense Variant

2. The result will appear as follows. The user can access to its region of interest by selecting the appropriate nucleotide range at the top of the page. The last two columns indicate the UMD-predictor pathogenicity score accompanied by the predicted impact of the mutation. In the example selected below, the user can evaluate the pathogenicity of missense variants affecting codons coding for Gly299, Glu300 and Phe301.

The UMD-DYSF mutations database
All possible missense and synonymous variations

Nucleotides : 1 to 500 | 501 to 1000 | 1001 to 1500 | 1501 to 2000 | 2001 to 2500 | 2501 to 3000 | 3001 to 3500 | 3501 to 4000 | 4001 to 4500 | 4501 to 5000 | 5001 to 5500 | 5501 to 6000 | 6001 to 6240

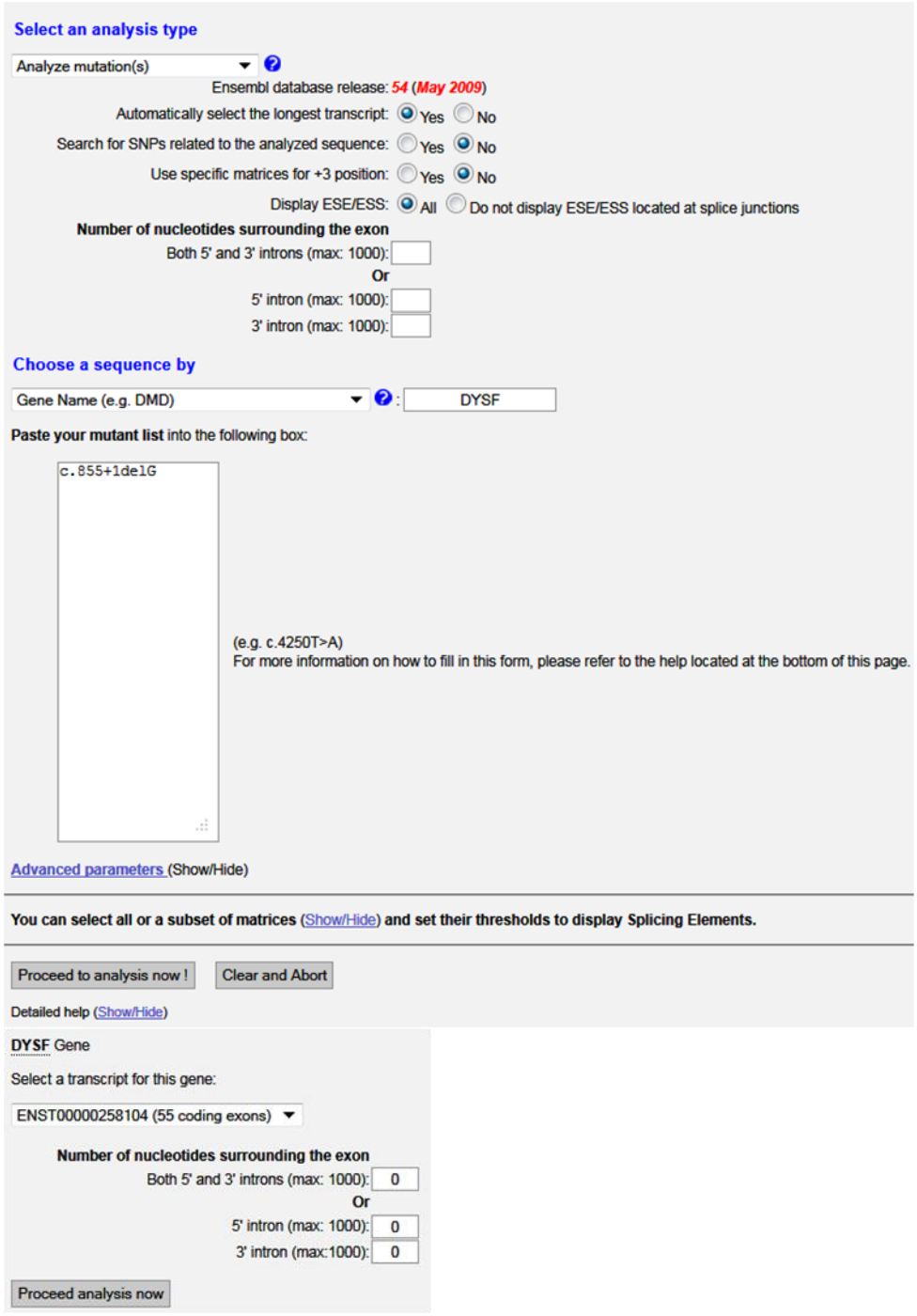
The UMD-Predictor® algorithm was used to predict the pathogenicity of all possible non-synonymous or synonymous mutations from the *DYSF* gene

c.895G>A	p.Gly299Arg	C2 Domain B	C2 Domain B	0.86	Potential acceptor splice site [91.47]	100	Pathogenous
c.895G>T	p.Gly299Trp	C2 Domain B	C2 Domain B	0.86	No impact	94	Pathogenous
c.895G>C	p.Gly299Arg	C2 Domain B	C2 Domain B	0.86	Potential acceptor splice site [78.23]	100	Pathogenous
c.896G>A	p.Gly299Glu	C2 Domain B	C2 Domain B	0.86	Potential acceptor splice site [81.74]	99	Pathogenous
c.896G>T	p.Gly299Val	C2 Domain B	C2 Domain B	0.86	No impact	82	Pathogenous
c.896G>C	p.Gly299Ala	C2 Domain B	C2 Domain B	0.86	Potential acceptor splice site [78.38]	63	Probable polymorphism
c.897G>A	p.Gly299Gly	C2 Domain B	C2 Domain B	0.86	Potential acceptor splice site [81.88]	29	Polymorphism
c.897G>T	p.Gly299Gly	C2 Domain B	C2 Domain B	0.86	Potential donor splice site [89.64]	18	Polymorphism
c.897G>C	p.Gly299Gly	C2 Domain B	C2 Domain B	0.86	No impact	18	Polymorphism
c.898G>A	p.Glu300Lys	C2 Domain B	C2 Domain B	0.86	No impact	41	Polymorphism
c.898G>C	p.Glu300Gln	C2 Domain B	C2 Domain B	0.86	Potential acceptor splice site [86.04]	41	Polymorphism
c.899A>T	p.Glu300Val	C2 Domain B	C2 Domain B	0.86	No impact	76	Pathogenous
c.899A>G	p.Glu300Gly	C2 Domain B	C2 Domain B	0.86	No impact	76	Pathogenous
c.899A>C	p.Glu300Ala	C2 Domain B	C2 Domain B	0.86	No impact	71	Probably pathogenous
c.900G>A	p.Glu300Glu	C2 Domain B	C2 Domain B	0.86	No impact	18	Polymorphism
c.900G>T	p.Glu300Asp	C2 Domain B	C2 Domain B	0.86	No impact	29	Polymorphism
c.900G>C	p.Glu300Asp	C2 Domain B	C2 Domain B	0.86	No impact	29	Polymorphism
c.901T>A	p.Phe301Ile	C2 Domain B	C2 Domain B	0.86	No impact	47	Polymorphism
c.901T>G	p.Phe301Val	C2 Domain B	C2 Domain B	0.86	Potential donor splice site [74.18]	53	Probable polymorphism
c.901T>C	p.Phe301Leu	C2 Domain B	C2 Domain B	0.86	No impact	41	Polymorphism
c.902T>A	p.Phe301Tyr	C2 Domain B	C2 Domain B	0.86	No impact	29	Polymorphism
c.902T>G	p.Phe301Cys	C2 Domain B	C2 Domain B	0.86	No impact	71	Probably pathogenous
c.902T>C	p.Phe301Ser	C2 Domain B	C2 Domain B	0.86	No impact	76	Pathogenous
c.903C>A	p.Phe301Leu	C2 Domain B	C2 Domain B	0.86	No impact	41	Polymorphism
c.903C>T	p.Phe301Phe	C2 Domain B	C2 Domain B	0.86	No impact	18	Polymorphism
c.903C>G	p.Phe301Leu	C2 Domain B	C2 Domain B	0.86	No impact	41	Polymorphism

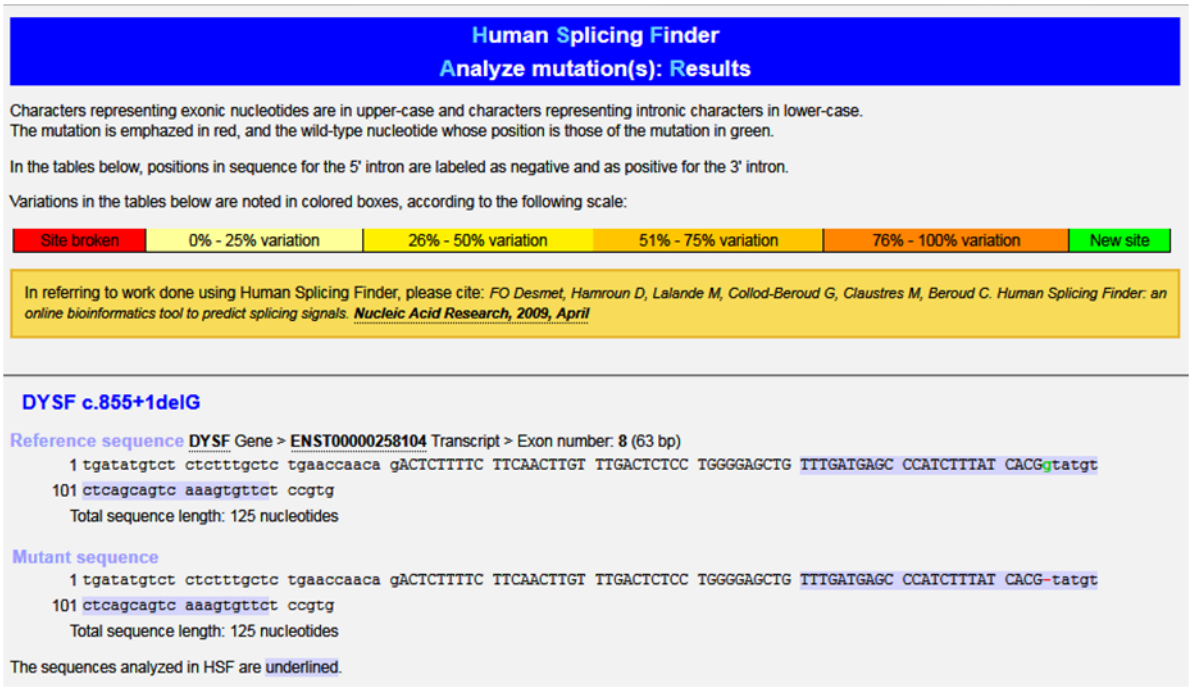
V. I want to evaluate a possible deleterious effect of a sequence variant on normal splicing using the bio-informatics tool Human Splicing Finder

1. Go to the Mutations page using the button on the left panel, and select
* I Want to Analyze an Intronic Variant that will bring you to the Human Splicing Finder page.

2. In the example shown below, the user wants to evaluate the effect of the 5' splice site variant c.855+1delG. The mutation involves a G deletion at the first nucleotide of *DYSF* intron 8. In this case, position 855 corresponds to the last exon 8 nucleotide.



3. The schema below shows the sequence context of the mutation site (the first nucleotide of intron 8) in the reference and mutant sequences.



Human Splicing Finder
Analyze mutation(s): Results

Characters representing exonic nucleotides are in upper-case and characters representing intronic characters in lower-case.
The mutation is emphazed in red, and the wild-type nucleotide whose position is those of the mutation in green.

In the tables below, positions in sequence for the 5' intron are labeled as negative and as positive for the 3' intron.

Variations in the tables below are noted in colored boxes, according to the following scale:

Site broken | 0% - 25% variation | 26% - 50% variation | 51% - 75% variation | 76% - 100% variation | New site

In referring to work done using Human Splicing Finder, please cite: *FO Desmet, Hamroun D, Lalande M, Collod-Beroud G, Claustres M, Beroud C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals.* ***Nucleic Acid Research, 2009, April***

DYSF c.855+1delG

Reference sequence **DYSF** Gene > **ENST00000258104** Transcript > Exon number: **8** (63 bp)

```
      1 tgatatgtct ctctttgctc tgaaccaaca gACTCTTTTC TTCAACTTGT TTGACTCTCC TGGGGAGCTG TTTGATGAGC CCATCTTTAT CACGgtatgt
    101 ctcagcagtc aaagtgttct ccgtg
```

Total sequence length: 125 nucleotides

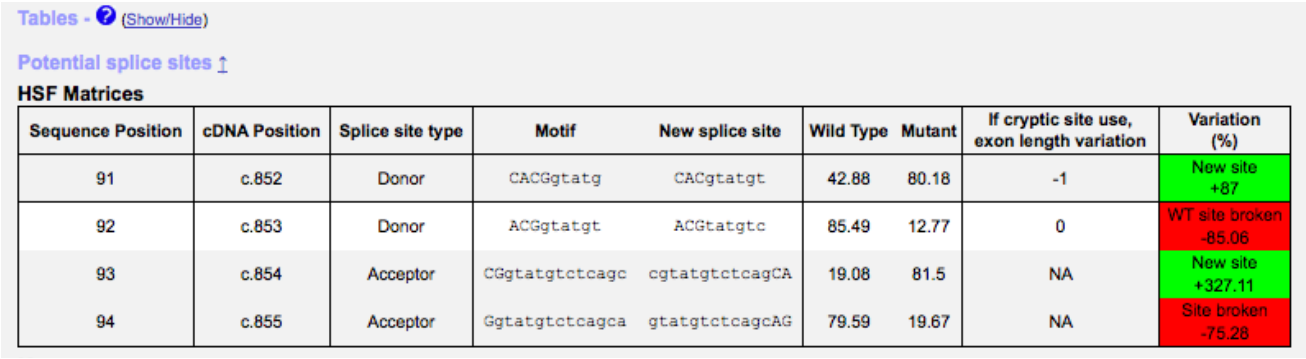
Mutant sequence

```
      1 tgatatgtct ctctttgctc tgaaccaaca gACTCTTTTC TTCAACTTGT TTGACTCTCC TGGGGAGCTG TTTGATGAGC CCATCTTTAT CACG-tatgt
    101 ctcagcagtc aaagtgttct ccgtg
```

Total sequence length: 125 nucleotides

The sequences analyzed in HSF are underlined.

4. Additional tables present results of HSF predictions indicating a new potential splice site in the mutant sequence that uses the last nucleotide of wild type exon 8 as the first nucleotide of the intron.



Tables - (Show/Hide)

Potential splice sites

HSF Matrices

Sequence Position	cDNA Position	Splice site type	Motif	New splice site	Wild Type	Mutant	If cryptic site use, exon length variation	Variation (%)
91	c.852	Donor	CACGgtatg	CACgtatgt	42.88	80.18	-1	New site +87
92	c.853	Donor	ACGgtatgt	ACGtatgtc	85.49	12.77	0	WT site broken -85.06
93	c.854	Acceptor	CGgtatgtctcagc	cgtatgtctcagCA	19.08	81.5	NA	New site +327.11
94	c.855	Acceptor	Ggtatgtctcagca	gtatgtctcagcAG	79.59	19.67	NA	Site broken -75.28

This new splicing event results in a premature stop codon and causes probable degradation of the mRNA by a nonsense-mediated decay mechanism (Wenzel et al., Human Mutation, Mutation in Brief #895 (2006) Online).

APPENDIX - Search module: description of the UMD-DYSF fields

On the 

page, the user can select * I want to Search the Database to query the database. The different fields below can be used to select a specific subset of mutational entries.

Amino acid position. For exonic mutations, indicates the position of the first codon affected by the mutation. For intronic mutations, indicates the position of the codon closest to the mutation.

Exon #. Exon numbers are based on the reference sequence used in UMD-DYSF. For exonic mutations, indicates the number of the exon affected by the mutation. For intronic mutations indicates the two exons flanking the first mutated nucleotide.

Gender. Female / Male / Unknown.

Geographic origin. When available, information was extracted from publications.

Highly Conserved Residues (HCD). Dysferlin charged or polar residues conserved with aspartic residues implicated in Ca2+ coordination in synaptotagmin III (Therrien et al. 2006) are labelled “Calcium binding”.

Mutant codon. For a substitutive change in an exon, indicates the mutated codon. For deletion, insertion or insertion/deletion changes in an exon, displays del* ins* or indels, respectively. For intronic mutations, displays spl* (spl+1 for mutations affecting the first nucleotide of the intron, spl+2 for mutations affecting the second nucleotide of the intron, spl-1 for mutations affecting the last nucleotide of the intron, etc).

Mutation name. Mutation names in UMD-DYSF are given according to nomenclature guidelines from the Human Genome Variation Society (den Dunnen and Antonarakis 2003, http://www.hgvs.org/mutnonem/) and numbered with respect to the DYSF gene cDNA reference sequence (+1=A of ATG). For example, c.6124C>T, c.5698_5699delAG, c.IVS49+3A>G (c.5525+3A>G).

Mutation status. Heterozygous / Homozygous.

Mutation type (Ts/Tv). For single base substitutions, indicates when the mutation is a transition (A<->G) or a transversion (T<->C). For insertions or deletions, indicates whether the change is predicted to promote a frameshift (Fr.) or to maintain the reading frame (InF).

Mutational event. For exonic point mutations that are predicted to respect the reading frame), describes the mutational event (A->G, T>A, etc). For in frame exonic insertions or deletions, displays “in frame del” and “in frame ins”, respectively. For exonic insertions or deletions that are predicted to introduce a frameshift in the coding sequence displays, “Stop at …” followed by the position of the new in-frame stop codon. For intronic insertions or deletions, displays del* or ins*, respectively.

Nomenclature protein. Follows hgvs guidelines and corresponds to automatically predicted protein sequence based on genomic mutation. This field is automatically calculated and filled out, independently of any available experimental data on RNA or protein sequence.

Proband/Relative (Y/N). To search for index cases, select “is equal to” “Y”. To search for relatives, select “is equal to” “N”.

Protein level. Normal / Present / Reduced. When available, information was extracted from publications.

Reference #. Select the reference number of the publication recorded in UMD-DYSF (between 1 and 60, see list of publications referenced in UMD-DYSF on the References page).

Reference: Medline ID. Select the PubMed identifier (PMID) for the publication you are interested in.

Sample ID. An anonymous ID is attributed to each patient data included in the database (for example AUS19-1-1-0).

Structure. The user can select mutations affecting the structural domains annotated for dysferlin. (C2 and Ferlin domains were annotated with Pfam 25.0 predictions; DysF and TM domains were annotated with SMART 6 predictions). Selecting the “all” option will retrieve all mutations (within or outside annotated domains).

Wild type codon. For exonic point mutations and deletions, indicates the codon of the first mutated nucleotide. For exonic insertions, indicates the codon of the first nucleotide localized after the insertion point. For intronic mutations, indicates the codon corresponding to the nucleotide indicated in the mutation name.