

**UNIVERSITE DE MONTREAL**  
Faculté et département de Médecine  
Centre de recherche du CHU Sainte Justine

**UNIVERSITE PARIS VI – PIERRE ET MARIE CURIE**  
Ecole doctorale 515 : Complexité Du Vivant  
Unité INSERM 787- Institut de Myologie

**THESE DE DOCTORAT DE RECHERCHE**

En spécialité génétique humaine intitulée

**Identification des bases génétiques des myopathies à multi-  
minicores avec ou sans cardiomyopathie**

Soutenue par

**Claire CHAUVEAU**

le 24 septembre 2012

Thèse réalisée en cotutelle sous la direction des Dr Ana FERREIRO (MD. PhD., Université Paris VI) et Dr Mark SAMUELS (Prof., Université de Montréal)

Thèse présentée à la Faculté PARIS IV-PIERRE ET MARIE CURIE en vue de l'obtention des grades de Docteur et de Philosophiæ Doctor (PhD) en Sciences Biomédicales

Devant le jury composé de :

<b>Dr Ana FERREIRO / Dr Mark SAMUELS</b>	Directeurs de thèse
<b>Pr. Denise BUSSON</b>	Président/Examineur
<b>Dr. Patrick COSSETTE</b>	Rapporteur
<b>Dr. Isabelle RICHARD</b>	Rapporteur
<b>Dr. Andréa RICHTER</b>	Représentante de la doyenne de la faculté de Médecine de l'Université de Montréal

Septembre 2012 © Claire Chauveau

**Université de Montréal Faculté de Médecine**  
**Université Paris IV- ED 515**

Cette thèse intitulée :

**Identification des bases génétiques des myopathies à multi-  
minicores avec ou sans cardiomyopathie**

présentée par :

**Claire Chauveau**

a été évaluée par un jury composé des personnes suivantes :

**Pr Denise Busson, Président-Examineur**

**Dr Ana Ferreiro, Directeur de recherche**

**Dr Mark Samuels, Co-directeur de Recherche**

**Dr Patrick Cossette, Examineur**

**Dr Isabelle Richard, Examineur externe**

**Dr Andréa Richter, Représentant de la doyenne de la faculté de Médecine  
de l'Université de Montréal**

## I. Résumé de la thèse

### Identification des bases génétiques des myopathies à multi-minicores avec ou sans cardiomyopathie

Bien que les bases physiopathologiques de beaucoup de maladies musculaires soient dorénavant connues, les myopathies congénitales à cores (MCs), maladies génétiques qui se présentent dès la naissance avec un retard du développement moteur, une faiblesse musculaire et des complications respiratoires et/ou cardiaques parfois mortelles, demeurent mal comprises. Des mutations dans *RYR1*, *SEPNI*, *TTN*, *ACTA1*, *CFL2* et *MEGF10* ont été associées aux MCs, pourtant, dans plus de 50% des cas, le gène responsable reste à identifier.

L'objectif de ma thèse a été de clarifier les mécanismes physiopathologiques des MCs par l'identification de nouveaux gènes ou de nouvelles mutations. Cette thèse a eu une dimension internationale concrétisée par la mise en place d'une cotutelle UPMC (France) et UdeM (Québec).

J'ai développé deux axes de recherche complémentaires. D'une part j'ai étudié 21 familles informatives avec MC récessive, scoliose et atteinte respiratoire, en combinant clonage positionnel et étude de gènes candidats et en utilisant des outils variés allant du génotypage au séquençage de nouvelle génération (NGS). En parallèle, j'ai étudié 24 familles avec une MC autosomique récessive affectant les muscles cardiaque et squelettiques et dont le phénotype était semblable à celui observé chez des patients avec des délétions dans les 6 derniers exons de *TTN*. Ainsi pour l'analyse de cette deuxième cohorte, nous avons appliqué une stratégie de séquençage de gène candidat ciblée sur ces exons et de NGS pour le reste du gène.

Pendant mon doctorat j'ai identifié les défauts génétiques de 8 des 45 familles étudiées (18 %), et caractérisé 3 nouvelles entités médicales, dont deux MCs dues à des nouvelles mutations de *TTN*. Ces résultats ont servi à l'identification de nouvelles interactions protéiques de la titine et contribuent à définir *TTN* comme une cause majeure de pathologies musculaires cardiaques et/ou squelettiques. Une troisième nouvelle forme de MC est provoquée par une mutation d'un coactivateur transcriptionnel peu connu et jamais associé à une maladie. Ces résultats ont révélé un nouvel acteur clef et une nouvelle voie de signalisation dans la physiopathologie du muscle, ont eu un bénéfice direct en termes de conseil génétique et ouvrent la voie pour le développement de thérapies.

#### Mots clefs:

Myopathie congénitale, Cardiomyopathie, Titine, Analyse de liaison, Séquençage de nouvelle génération

## **II. Thesis abstract**

### **Identification of the genetic basis for multiminicore myopathies with or without cardiomyopathy**

While the pathophysiological bases of many muscular diseases are nowadays well known, congenital core myopathies (CMs) remain poorly understood. CMs are genetic diseases which generally present at birth with delayed motor development, muscle weakness, and sometimes fatal respiratory or cardiological complications. Mutations in RYR1, SEPN1, ACTA1, TTN and MEGF10 have been associated with various CMs, yet for about 50% of CM cases the responsible gene has not been identified.

The objective of my thesis was to clarify the pathophysiological mechanisms of new forms of CM through the identification of new genes or new mutations in known genes. This thesis had an international dimension as manifested by a UPMC (France) and UdeM (Québec) joint direction.

I developed two complementary axes of research. First, I studied 21 informative families with a recessive CM with scoliosis and respiratory failure, for which I combined positional cloning and candidate gene studies, using various tools from genotyping to next generation sequencing (NGS). The second part of this work consisted on the analysis of 24 families with recessive CM affecting both cardiac and skeletal muscles. Their phenotype was similar to that previously observed in cases with deletions in the last 6 exons of the giant gene TTN. Thus we applied a candidate gene strategy through direct Sanger sequencing coupled with NGS for the analysis of this second cohort.

During my PhD work I identified the molecular defect in 8 out of the 45 families included (18%), which led to the identification and characterization of 3 novel medical entities, including two new CMs due to novel defects of TTN. These results served to identify new titin protein interactions, and participate in the definition of TTN defects as a major cause of both cardiac and skeletal muscle conditions. A third new form of CM is due to mutations of a poorly-known transcriptional coactivator whose role in striated muscle physiology was unknown and which had never been associated to a human condition. Globally, these results unveiled a novel important protein and pathway in muscle pathophysiology, have direct health benefits (molecular diagnosis) and open the way for therapeutic investigations.

#### **Key words:**

Congenital myopathy, cardiomyopathy, titin, linkage analysis, next-generation sequencing

### III. Table des matières

<b>I.</b>	<b>Résumé de la thèse</b>	<b>3</b>
<b>II.</b>	<b>Thesis abstract</b>	<b>4</b>
<b>III.</b>	<b>Table des matières</b>	<b>5</b>
<b>IV.</b>	<b>Liste des tableaux</b>	<b>11</b>
<b>V.</b>	<b>Liste des figures</b>	<b>12</b>
<b>VI.</b>	<b>Liste des annexes</b>	<b>13</b>
<b>VII.</b>	<b>Liste des abréviations</b>	<b>14</b>
<b>VIII.</b>	<b>Remerciements</b>	<b>20</b>
<b>INTRODUCTION</b>		<b>23</b>
<b>1.</b>	<b>Le muscle strié</b>	<b>25</b>
<b>1.1.</b>	<b>Physiologie du muscle strié</b>	<b>25</b>
1.1.1.	Structure et organisation du muscle strié	25
1.1.1.1.	<i>Un tissu aux constituants multiples</i>	25
1.1.1.2.	<i>La fibre musculaire : constituant majeur du muscle strié</i>	26
1.1.1.3.	<i>Le sarcomère : unité contractile du muscle strié</i>	27
1.1.1.4.	<i>Les filaments : anatomie moléculaire du muscle strié</i>	28
1.1.2.	Fonctionnement et propriétés du muscle strié	29
1.1.2.1.	<i>La contraction musculaire</i>	29
1.1.2.1.1.	<u>La modulation calcique</u>	29
1.1.2.1.2.	<u>Le couplage excitation-contraction</u>	30
1.1.2.1.3.	<u>Le relâchement musculaire</u>	31
1.1.2.2.	<i>La plasticité musculaire</i>	31
1.1.2.2.1.	<u>Régénération/dégénération</u>	32
1.1.2.2.2.	<u>Hyperplasie/hypertrophie</u>	32
1.1.2.3.	<i>Le myocarde : un muscle strié unique</i>	33
<b>1.2.</b>	<b>Pathologie du muscle strié : aperçu sélectif</b>	<b>33</b>
1.2.1.	Myopathie ou dystrophie ?	33
1.2.1.1.	<i>Les myopathies d'origine génétique</i>	34
1.2.1.2.	<i>Le cas des myopathies à début précoce</i>	34

1.2.2. Les dystrophies musculaires congénitales (DMC)	35
1.2.2.1. <i>DMC avec déficit en laminine-<math>\alpha</math>2 (mérosine)</i>	35
1.2.2.2. <i>Dystroglycanopathies: défaut de glycosylation de l'<math>\alpha</math>-dystroglycane</i>	36
1.2.2.3. <i>DMCs avec rétractions proéminentes : Collagène VI et sélénoprotéine N</i>	37
1.2.2.4. <i>DMCs avec déficience en intégrine <math>\alpha</math>-7</i>	39
1.2.3. Les myopathies congénitales	39
1.2.3.1. <i>Myopathies à cores : myopathie à cores centraux (CCD), myopathie à multi-minicores (MmD)</i>	39
1.2.3.2. <i>Les myopathies myotubulaires et centronucléaires</i>	40
1.2.3.3. <i>Myopathies à bâtonnets</i>	40
<b>2. La titine : un acteur majeur du muscle, un vrai challenge de recherche</b>	<b>41</b>
<b>2.1. Structure et fonction de la titine</b>	<b>41</b>
2.1.1. Le titan du muscle	41
2.1.1.1. Une organisation bien définie	42
2.1.1.1.1. <i>La ligne Z</i>	42
2.1.1.1.2. <i>La bande I</i>	43
2.1.1.1.3. <i>La bande A</i>	44
2.1.1.1.4. <i>La bande M</i>	44
2.1.1.2. Rôles multiples et nombreux partenaires	46
2.1.2. Epissage alternatif et différentes isoformes	48
2.1.2.1. Les quelques séquences décrites	48
2.1.2.2. Un champ infini de possibilités	49
2.1.2.3. Les TTN non musculaires	49
<b>2.2. Mutations de la TTN et pathologies associées</b>	<b>49</b>
2.2.1. Pertinence génétique	49
2.2.1.1. <i>Types de variation</i>	50
2.2.1.2. <i>Mode de transmission</i>	50
2.2.1.3. <i>Répartition sur le gène</i>	51
2.2.1.4. <i>Variants de signification clinique inconnue ou incertaine</i>	52
2.2.2. Pertinence clinique	52
2.2.2.1. <i>Les phénotypes cardiaques</i>	52
2.2.2.1.1. <u>La cardiomyopathie dilatée (DCM)</u>	53

2.2.2.1.2. <u>La cardiomyopathie hypertrophique (HCM)</u>	53
2.2.2.1.3. <u>La cardiomyopathie arythmogène ventriculaire droite (ARVC)</u>	54
2.2.2.2. <i>Les phénotypes musculaires</i>	55
2.2.2.2.1. <u>La dystrophie musculaire tibiale tardive (TMD)</u>	55
2.2.2.2.2. <u>La dystrophie musculaire des ceintures de type 2J (LGMD2J)</u>	55
2.2.2.2.3. <u>La myopathie héréditaire avec atteinte respiratoire précoce (HMERF)</u>	56
2.2.2.3. <i>La myopathie à début précoce avec cardiomyopathie fatale (EOMFC): une titinopathie à phénotype mixte</i>	56
<b>2.3. Les modèles animaux</b>	<b>57</b>
2.3.1. Le zébrafish	57
2.3.1.1. <i>Le mutant pickwick</i>	57
2.3.1.2. <i>Le mutant runzel</i>	57
2.3.2. La souris	58
2.3.2.1. <i>Modèle murin de TMD</i>	58
2.3.2.2. <i>Modèle murin de TMD/LGMD2J (FINmaj mutation)</i>	58
2.3.2.3. <i>Modèle murin de DCM</i>	58
2.3.2.4. <i>Modèle murin d'analyse de bande M</i>	59
<b>3. La génétique des maladies à transmission mendélienne</b>	<b>59</b>
<b>3.1. Identification d'une mutation : méthodologies contemporaines</b>	<b>59</b>
3.1.1. Du ciblage	60
3.1.1.1. <i>Criblage du génome (génotypage) et analyse de liaison</i>	60
3.1.1.2. <i>Clonage positionnel</i>	61
3.1.1.3. <i>La recherche appliquée aux familles consanguines : principe de la cartographie par homozygotie</i>	61
3.1.2. ... au séquençage	62
3.1.2.1. <i>Stratégie directe de gène candidats</i>	62
3.1.2.2. <i>Le séquençage de nouvelle génération</i>	63
<b>3.2. La troisième génération à l'aune d'une révolution</b>	<b>65</b>
<b>SUJETS ET METHODES</b>	<b>67</b>
<b>1. Patients</b>	<b>69</b>
<b>2. Etudes morphologiques</b>	<b>70</b>
<b>2.1. Préparation des biopsies</b>	<b>70</b>
<b>2.2. Colorations histologiques standard et immuno-marquages</b>	<b>70</b>

2.2.1. Colorations histologiques	70
2.2.1.1. <i>La coloration hématoxyline-éosine</i>	70
2.2.1.2. <i>Les colorations oxydatives</i>	71
2.2.2. Immunomarquages de la titine et de l' $\alpha$ -actinine	71
<b>3. Etudes génétiques</b>	<b>73</b>
<b>3.1. Génotypage</b>	<b>73</b>
3.1.1. Approche mixte de cartographie par homozygotie - gène candidat	73
3.1.2. Criblage du génome	73
3.1.3. Analyses de liaison	74
<b>3.2. Séquençage de type Sanger</b>	<b>75</b>
<b>3.3. Séquençage d'exomes</b>	<b>76</b>
3.3.1. Préparation des échantillons	76
3.3.2. Analyses bioinformatiques	78
<b>4. Cultures de cellules primaires</b>	<b>81</b>
<b>4.1. Décongélation des cellules</b>	<b>81</b>
<b>4.2. Amplification des cellules</b>	<b>81</b>
<b>4.3. Induction de la différenciation</b>	<b>82</b>
<b>4.4. Congélation des cellules</b>	<b>82</b>
<b>5. Analyse fonctionnelle des conséquences des mutations</b>	<b>83</b>
<b>5.1. A l'échelle des ARN (famille C)</b>	<b>83</b>
5.1.1. Cultures cellulaires	83
5.1.2. Extraction des ARN totaux	83
5.1.3. Transcription inverse des ARNs en ADN complémentaire	83
5.1.4. Quantification relative par PCR quantitative en temps réel	84
<b>5.2. A l'échelle protéique : Analyse par Western-Blot</b>	<b>85</b>
5.2.1. Analyse des mutations <i>TTN</i> (P5)	85
5.2.2. Analyse de l'expression de la protéine identifiée dans la famille C	85
5.2.3. Analyse du dichroïsme circulaire	86
<b>RESULTATS</b>	<b>91</b>
<b>1. PARTIE I : Etude d'une série de patients avec myopathie congénitale à cores associée à une atteinte cardiaque</b>	<b>93</b>
<b>1.1. Constitution d'une série de patients: critères d'inclusion</b>	<b>93</b>

<b>1.2. Etude génétique</b>	<b>94</b>
1.2.1. Etat des lieux au début du projet	94
1.2.2. Analyse du gène candidat TTN : séquençage traditionnel	95
1.2.3. Analyse d'exome et identification d'une deuxième mutation hétérozygote composite de TTN chez le Patient 2	97
<b>1.3. Phénotype des familles avec mutations TTN identifiées</b>	<b>99</b>
<b>1.4. Etude fonctionnelle</b>	<b>101</b>
1.4.1. Expression des TTN mutées chez la patiente 1	101
<b>2. PARTIE II : Une nouvelle forme de MmD avec non-compactation ventriculaire et arthrogrypose est causée par des mutations TTN hétérozygotes composites conduisant à une perte total de fonction du domaine kinase</b>	<b>105</b>
<b>2.1. Etude clinique : un nouveau phénotype</b>	<b>105</b>
2.1.1. Présentation du patient : phénotype et contexte familial	105
2.1.2. Phénotype morphologique	106
<b>2.2. Etude génétique</b>	<b>107</b>
2.2.1. Etat des lieux au début du projet	107
2.2.2. Analyse du gène candidat TTN par séquençage Sanger	108
2.2.3. NGS, analyse d'exome : identification d'un variant d'épissage	108
2.2.4. Diagnostic prénatal	109
<b>2.3. Etude fonctionnelle</b>	<b>109</b>
2.3.1. Mutation d'épissage	109
2.3.1.1. <i>Analyse in-silico</i>	109
2.3.1.2. <i>Discrimination allélique dans les muscles striés</i>	110
2.3.2. Mutation TK	111
2.3.2.1. <i>Impact structurel</i>	111
2.3.2.2. <i>Impact sur le signalosome</i>	112
2.3.3. Analyses protéiques	113
2.3.3.1. <i>Immunofluorescence in-situ</i>	113
2.3.3.2. <i>Western blot</i>	114
<b>3. PARTIE III : Etude d'une série de patients MmD sans atteinte cardiaque : des mutations d'un coactivateur transcriptionnel provoquent une nouvelle forme grave de myopathie congénitale</b>	<b>117</b>
<b>3.1. Etude clinique</b>	<b>117</b>
3.1.1. Constitution d'une série de patients: hétérogénéité phénotypique	117

3.1.2. La famille C : un nouveau phénotype clinique et histologique	117
<b>3.2. Etude génétique</b>	<b>118</b>
3.2.1. Etat des lieux au début du projet	118
3.2.2. Génotypage et analyse de liaison	118
3.2.2.1. <i>Identification de 2 loci dans la famille C</i>	118
3.2.2.2. <i>Second génotypage et cartographie fine : confirmation d'un locus</i>	119
3.2.3. Séquençage des gènes candidats positionnels	120
<b>3.3. Etude fonctionnelle des variants identifiés dans la famille C</b>	<b>121</b>
3.3.1. Analyse quantitative de l'expression des ARNm chez les patients	121
3.3.2. Analyse de l'expression protéique par western-blot	122
3.3.3. Impact sur les facteurs myogéniques	123
<b>DISCUSSION</b>	<b>127</b>
<b>Première conclusion : L'implication des mutations TTN hétérozygotes composites dans les pathologies musculaires et cardiaques est probablement sous-estimée</b>	<b>129</b>
<b>Seconde conclusion : La liaison de la TTN à une variété de défauts cardiaques est plus importante que celle actuellement estimée</b>	<b>135</b>
<b>Troisième conclusion : Un KO constitutif de TK chez l'homme mène à la description d'une nouvelle entité médicale associant une atteinte cardiaque et squelettique</b>	<b>137</b>
<b>Quatrième conclusion : Une mutation dans une nouvelle protéine potentiellement impliquée dans la myogénèse est responsable d'une nouvelle maladie du muscle chez l'homme</b>	<b>139</b>
<b>Conclusion et perspectives</b>	<b>141</b>
<b>IX. Les sources documentaires</b>	<b>143</b>
<b>X. Les appendices (Annexes)</b>	<b>153</b>
<b>XI. Curriculum Vitæ</b>	<b>171</b>

#### **IV. Liste des tableaux**

<b>Tableau 1</b>	<b>76</b>
<b>Tableau 2</b>	<b>94</b>

## V. Liste des figures

Figure 1	26
Figure 2	27
Figure 3	28
Figure 4	31
Figure 5	37
Figure 6	42
Figure 7	45
Figure 8	47
Figure 9	48
Figure 10	51
Figure 11	64
Figure 12	65
Figure 13	77
Figure 14	79
Figure 15	80
Figure 16	87
Figure 17	87
Figure 18	88
Figure 19	89
Figure 20	96
Figure 21	98
Figure 22	100
Figure 23	102
Figure 24	105
Figure 25	107
Figure 26	109
Figure 27	110
Figure 28	111
Figure 29	113
Figure 30	115
Figure 31	118
Figure 32	121
Figure 33	122
Figure 34	123
Figure 35	133

## **VI. Liste des annexes**

<b>Annexe 1 : Tableau descriptif des principales myopathies congénitales</b>	<b>153</b>
<b>Annexe 2 : Primers utilisés pour le séquençage de l'extrémité N-terminale de TTN</b>	<b>154</b>
<b>Annexe 3 : Mutations <i>TTN</i> répertoriées dans la littérature</b>	<b>155</b>
<b>Annexe 4 : Famille K: illustration de l'importance de l'investigation clinique</b>	<b>157</b>
<b>Annexe 5 : Famille G : Illustration de l'importance des collaborations scientifiques</b>	<b>159</b>
<b>Annexe 6 : Poster congrès annuel de l'American Society of Human Genetics 2012</b>	<b>161</b>
<b>Annexe 7 : Article soumis dans <i>Annals of Neurology</i></b>	<b>162</b>

## VII. Liste des abréviations

A, T, G, C	Bases de l'ADN : Adénine, Thymine, Guanine, Cytosine
ACTA1	$\alpha$ -actine squelettique
ACTC	$\alpha$ -actine cardiaque
ADN	Acide désoxyribonucléique
ADP	Adénosine diphosphate
ARN	Acide ribonucléique
ARVC	Cardiomyopathie arythmogène ventriculaire droite ( <i>Arrhythmogenic right ventricular cardiomyopathy</i> )
ASD	( <i>Atrial Septal Defect</i> ) Anomalie du septum auriculaire
ATP	Adénosine triphosphate
BAG3	(Bcl-2-associated athanogene-3)
BIN1	Amphiphysine 2
C2C12	Lignée de cellules musculaires de souris
Ca <sup>2+</sup>	Calcium
CaM	Calmoduline
CAPN3	Calpaine 3
CD	Dichroïsme circulaire
CFL2	Cofiline-2
CIA	Communication Intra Auriculaire

CK	Créatine kinase
CNV	Variation du nombre de copie ( <i>Copy number variation</i> )
DCM	Cardiomyopathie dilatée ( <i>Dilated cardiomyopathy</i> )
DMC	Dystrophie musculaire congénitale
DNM2	Dynamine 2
dpf	Jour post-fertilisation
ECG	Electrocardiographie
EJC	( <i>Exon Junction Complex</i> )
ELC	Chaîne légère essentielles de la myosine, ( <i>Essential Light Chain</i> )
EOMFC	( <i>Early-Onset Myopathy With Fatal Cardiomyopathy</i> )
Fhl2	( <i>Four and a half LIM domains protein 2</i> )
FINmaj	Mutation fondatrice finnoise de <i>TTN</i>
Fn3	Fibronectine 3
GRCh37/hg19	Génome humain de référence fournit par UCSC à partir du <i>Genome Reference Consortium</i>
HCM	Cardiomyopathie hypertrophique ( <i>Hypertrophic cardiomyopathy</i> )
HGMD	( <i>Human Gene Mutation Database Leiden</i> )
HGVS	( <i>Human Genome Variation Society</i> )
HMERF	Myopathie héréditaire avec atteinte respiratoire précoce ( <i>Hereditary myopathy with early respiratory failure</i> )
IF	Immunofluorescence

Ig	Immunoglobuline
kDa	kilo Dalton, unité de mesure
KO	Technique de biologie moléculaire permettant d'inactiver un gène ( <i>Knock-Out</i> )
LAMP2	( <i>Lysosome-Associated Membrane Protein 2</i> )
LGMD2J	Dystrophie musculaire des ceintures de type 2J ( <i>Muscular dystrophy, limb-girdle, type 2j</i> )
LINE-1	Rétrotransposon ( <i>Long INterspersed (nuclear) Elements</i> )
LMNA	Lamine A/C
LOVD	( <i>Leiden Open Variation Database</i> )
MC	Myopathie congénitale
<i>mdm</i>	Mutant murin de <i>Tin</i> ( <i>Muscular dystrophy with myositis</i> )
Mex	Exon de la titine codant les domaines de la bande M, ( <i>M-line encoding exon</i> )
MmD	Myopathie à multiminicores, ( <i>Multiminicores disease</i> )
MRFs	( <i>myogenic regulatory factors</i> ), facteurs de transcription myogéniques
MTM1	Myotubularine
MURF2	( <i>Muscle specific RING finger protein 2</i> ) équivaut à TRIM55
MyBP-C	Protéin C de liaison à la myosine ( <i>Myosin binding protein C</i> )
MYBPC3	( <i>Myosin-Binding Protein C gene</i> )
MYH7	( <i>Myosin, Heavy Chain 7 gene</i> )
MyHC	Chaîne lourde de la myosine, ( <i>Myosin Heavy Chain</i> )
MYL2	( <i>Myosin, Light Chain 2 gene</i> ) un des deux types de RLC du muscle cardiaque

MYL3	<i>(Myosin, Light Chain 3 gene)</i> un des deux types de ELC du muscle cardiaque et squelettique lent
MyoD	Facteurs de transcription de la famille des MRFs
MyoG	Myogénine
MYOM1	Myomésine 1
MYOM2	Myomésine 2, ( <i>M band protein</i> )
MYOM3	Myomésine 3
N2-A	Isoforme de <i>TTN</i> d'expression ubiquitaire
N2-B	Isoforme de <i>TTN</i> d'expression cardiaque
N2BA	Isoforme cardiaque de <i>TTN</i>
NBR1	<i>(Neighbor of BRCA1 gene 1)</i>
NCBI	National Center for Biotechnology Information
NGS	<i>(Next Generation Sequencing)</i> Séquençage de nouvelle génération
Novex-1	Isoforme de <i>TTN</i>
Novex-2	Isoforme de <i>TTN</i>
Novex-3	Isoforme de <i>TTN</i>
Obscn	Obscurine
Obsl1	Obscurine like 1
OMIM	<i>(Online Mendelian Inheritance in Man)</i>
p62	<i>(Ubiquitin binding protein p62)</i>

PEVK (K)	Domaine protéique de la titine riche en proline (P), glutamine (E), valine (V) et lysine
Pi	Phosphate inorganique
<i>pik<sup>m171</sup></i>	Mutant de <i>TTN</i> chez le poisson rayé
PRKAG2	( <i>Protein Kinase, Amp-Activated, Noncatalytic, Gamma-2</i> )
PTC	Codon stop premature ( <i>Premature termination codon</i> )
RLC	Chaîne légère régulatrices de la myosine, ( <i>Regulatrice Light Chain</i> )
<i>runzel</i>	Mutant de <i>TTN</i> chez le poisson rayé
RyR1	Récepteur à la Ryanodine de type 1
SEPN1	Sélénoprotéine N
SERCA	sarco/endoplasmic reticulum Ca <sup>2+</sup> -ATPase
siRNA	( <i>silencing RNA</i> )
SNP	( <i>Single Nucleotide Polymorphism</i> )
SQSTM1	( <i>Sequestosome-1</i> ) ligand de <i>TTN</i>
SRF	Facteur de transcription, ( <i>Serum response factor</i> )
TCAP	Titine-CAP, gène codant la téléthonine
TK	Domaine kinase de la titine
TMD	Dystrophie musculaire tibiale tardive ( <i>Tibial muscular dystrophy, tardive</i> )
TNNT1	Troponine T, du muscle squelettique lent
TNNT2	Troponine T du muscle cardiaque
TPM1	$\alpha$ -tropomyosine du muscle squelettique

TPM2	$\beta$ -tropomyosine
TPM3	$\alpha$ -tropomyosine du muscle squelettique lent
<i>Ttn</i>	Orthologue de <i>TTN</i> chez la souris
<i>ttn</i>	Orthologue de <i>TTN</i> chez le poisson zébre
TTN	Titine
UTR	Région non traduite de l'ADN, ( <i>Untranslated regions</i> )
VUS	Variant de signification inconnue ( <i>Variant of unknown significance</i> )
WES	( <i>Whole exome sequencing</i> ) Séquencage d'exome
ZASP/LDB3	( <i>Z-disk alternatively spliced PDZ-containing protein</i> )
Zr	Répétition Z de la ligne de <i>TTN</i>

## VIII. Remerciements

Dresser une liste de toutes les personnes que j'aimerais remercier nécessiterait le défrichage de la moitié des forêts québécoises, l'âme même de la belle province... La liste qui suit n'est donc pas exhaustive, mais ceux dont les noms n'apparaîtront pas noir sur blanc savent de toutes façons qu'ils ne sont pas oubliés.

Je remercie en tout premier lieu et très sincèrement le docteur Ana Ferreiro qui m'a accueillie dans son laboratoire et qui a su me donner l'autonomie et l'indépendance dont j'avais besoin pour réaliser ce projet de recherche. Merci pour cette relation de confiance qui s'est installée naturellement. Merci également d'avoir accepté de me laisser partir pour le Québec.

Merci à Mark de m'avoir si bien accueillie sans me connaître, d'avoir partagé de précieux conseils, ses connaissances théoriques et techniques. Merci pour son soutien, sa réactivité et ses anecdotes sur le monde de l'entreprise.

Je tiens également à remercier les membres du jury. Merci au Pr. Denise Busson d'en avoir accepté la présidence; merci au Dr Andréa Richter et aux Dr Isabelle Richard et Dr Patrick Cossette d'avoir accepté de juger ce travail et d'en être rapporteurs. Merci à tous nos collaborateurs. Une mention spéciale au Dr Mathias Gautel pour ses échanges à propos de la titine.

Merci à toutes les personnes des secrétariats des divers départements, facultés et bureaux des relations internationales pour leurs aides administratives relatives à la procédure de cotutelle, et plus particulièrement un énorme merci à Sandy Lalonde.

Merci à tous les résidents du 2<sup>ème</sup> étage du 105 pour leurs sympathies, leurs présences et passions communicatives de la recherche. Un énorme merci, tout particulier aux membres de mon équipe, Charline, John, Sandrine, Laurianne, Caroline. Sans oublier bien sur, Cédric, mon boxeur tatoué préféré. Merci aussi à Alice, Jess et Valou, mes camarades de thèse et de galère administrative!

Merci à la gang du déjeuner, le combo pain-fromage-café-chocolat restera un délicieux moment de nostalgie dans tous les pays par lesquels je passerai... Merci à Nath et Sesti qui ont instauré ce moment! Merci à Nath de m'avoir plongé dans la course à pied et surtout à Daniel d'avoir partagé tous ces kilomètres avec moi. Nath, c'est pas grave si tu nous a lâché, le jeu en valait vraiment la chandelle, ta récompense est magnifique! Piera, grazie mille pour tous le temps passé ensemble et

toute l'aide lors des missions souris avec Charline... Merci à Maria et Susanne entre autre pour ces magnifiques moments sportifs sur la péniche. Merci à Vincent pour l'énergie dispensée dans le labo et toutes les supers blagues qui, même pas forcément drôles, sont toujours appréciées. Merci à Luigi, mi amigo español, peut être qu'un jour hablaremos realmente español! Et merci à Sonia pour la tentative de cours, on a pas été de très bon élèves, encore heureux qu'on a pas essayé de se mettre à l'Euskara!

Enfin merci à Piera et Sonia d'avoir été là à un moment crucial de cette thèse et d'avoir fait en sorte qu'elle puisse se poursuivre malgré tout... Sonia, un énorme merci pour ce mythique comité de thèse post belette, tu mérites une grande part des éloges que j'ai reçu! Merci pour ça et pour tout le reste, tu m'as montré que c'était possible...

Merci à ma famille ; aux relecteurs: Cédric, Christine, Catherine et surtout Hélène ; à mes amis désolée de ne pas avoir fais commerce ou lettres, il fallait bien qu'un d'entre nous choisisse la science. Merci à Laura, c'était sympas de vivre ça ensemble! Merci enfin à Aurélie, Brunelle et mes colocs Christelle et Nico (je laisse ma chambre aux trains!).

Et pour finir merci aux nombreux participants du concours de la best tune to work until dawn, et bravo à Jean-Edouard, grand gagnant, quand tout sera fini il restera toujours la musique et la mélodie...



# **INTRODUCTION**



## 1. Le muscle strié

Le muscle représente l'un des 4 types de tissu animal avec le tissu épithélial, le tissu conjonctif et le tissu nerveux. Il a des **capacités contractiles** et constitue avec le tissu nerveux l'un des seuls tissus excitables. On distingue deux types de muscles : les muscles striés et les muscles lisses. Les muscles striés sont dits **squelettiques** si leur contraction permet de mouvoir des parties du corps. Au nombre avoisinant les 650, ils sont contrôlés par le système nerveux central, volontaire, relient généralement les os entre eux et permettent ainsi la motricité. Le muscle **cardiaque** est également un muscle strié, mais muni d'un système de contraction propre, sensible aux stimulations hormonales, et par conséquent, est le seul muscle strié qui ne peut être volontairement contrôlé. Enfin, les muscles lisses sont sous le contrôle du système nerveux autonome, involontaire, et ont pour fonction d'aider au transport de différents milieux (eau, air, sang...) à l'intérieur de l'organisme.

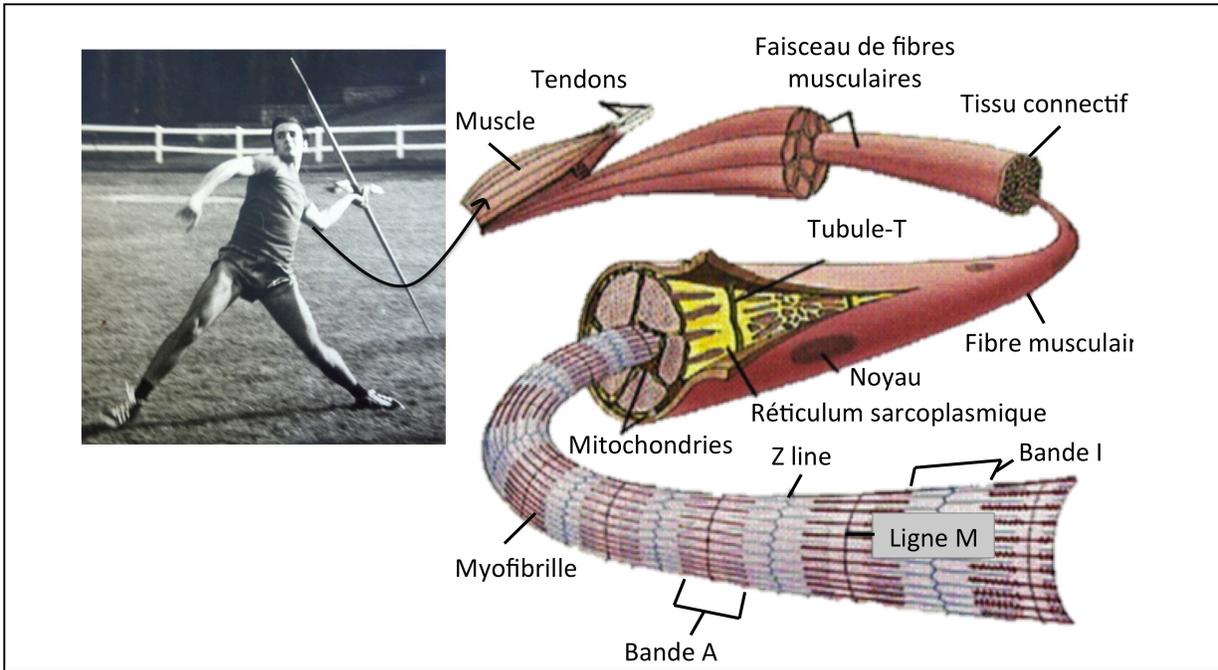
### 1.1. Physiologie du muscle strié

#### 1.1.1. Structure et organisation du muscle strié

##### 1.1.1.1. *Un tissu aux constituants multiples*

Chaque muscle strié squelettique est un organe bien délimité principalement constitué de **fibres musculaires**. Ces organes renferment aussi du tissu conjonctif assurant le maintien de la structure du muscle, des vaisseaux sanguins permettant l'approvisionnement en oxygène et nutriments, des neurofibres régissant l'activité musculaire et un pool de **cellules progénitrices**. Ces cellules, nommées cellules satellites, sont indifférenciées, localisées à la périphérie des fibres et sont caractérisées par leur détermination myogénique. Elles interviennent notamment lors de la **régénération musculaire**.

Le muscle strié squelettique est organisé en faisceaux (ou fascicules) comportant chacun de vingt à quarante fibres circonscrites dans une enveloppe fibreuse externe, l'épimysium, et dont la cohésion est assurée par un tissu conjonctif spécifique, l'endomysium. Ces fibres musculaires se terminent à leurs extrémités par des filaments de collagène, qui, regroupés, forment les tendons et assurent la fixation du muscle sur ses points d'insertion (**Fig. 1**).



**Figure 1 : Organisation du muscle strié**

*(Modifié d'après Purves et al., life;the science of biology, 4<sup>ème</sup> édition)*

#### 1.1.1.2. La fibre musculaire : constituant majeur du muscle strié

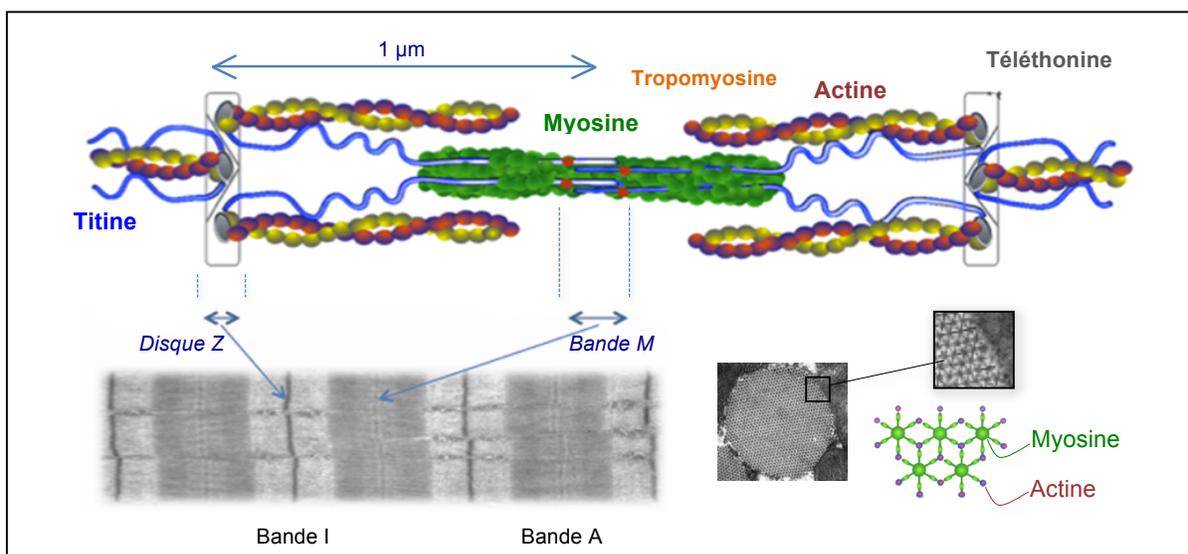
Une fibre provient de la fusion de nombreux précurseurs musculaires appelés **myoblastes** en un syncytium plurinucléé dont le nombre de noyaux varie en fonction du type musculaire. Ces noyaux sont situés à la périphérie du sarcoplasme (cytoplasme de la cellule musculaire) contre la face interne du sarcolemme (membrane plasmique du tissu musculaire). Ce dernier est électriquement excitable et invaginé, permettant ainsi la transmission de l'influx nerveux vers le centre de la fibre. A l'intérieur du sarcoplasme, le réticulum sarcoplasmique (réticulum endoplasmique du muscle) entoure les unités contractiles du muscle, les **myofibrilles** (**Fig. 1**). Celles-ci s'étendent selon une organisation parallèle sur toute la longueur de la fibre en une structure tubulaire allongée, constituée de **myofilaments** dont les deux types majoritaires sont les filaments fins d'actine et épais de myosine.

Les cellules musculaires ne constituent pas une population homogène, il existe deux catégories principales de fibres, les fibres lentes de type 1 intervenant dans les exercices prolongés et les fibres rapides de type 2 plus impliquées dans les efforts de courte durée. Ces deux types se distinguent sur des critères morphologiques, physiologiques et biochimiques. Les **fibres de type 1** présentent un métabolisme essentiellement **oxydatif**, aérobie. Elles sont aussi très riches en mitochondries et en myoglobine, un pigment musculaire capable de fixer plus fortement l'oxygène que l'hémoglobine et

qui donne au muscle sa couleur rouge caractéristique. Comparativement, les **fibres de type 2** ont un métabolisme de nature **glycolytique**, anaérobie, sont plus pauvres en mitochondries et myoglobine mais très riches en glycogène et enzymes glycolitiques qui leur confèrent une couleur plus claire.

### 1.1.1.3. Le sarcomère : unité contractile du muscle strié

La base de l'organisation fonctionnelle des myofibrilles est le **sarcomère** dont la composition protéique des différentes régions est à l'origine de l'aspect strié du muscle. Un sarcomère s'étend entre deux **disques Z**, nom donné aux complexes protéiques permettant l'ancrage des filaments fins d'actine entre lesquelles s'intercalent les filaments épais de myosine. Les sarcomères des différentes myofibrilles sont situés au même niveau selon un axe longitudinal, ceci étant rendu possible par un ensemble de protéines de soutien reliant transversalement les filaments épais d'un même faisceau musculaire qui porte le nom de **bande M** (Fig. 1 et 2). Cet alignement des sarcomères est observable en microscopie électronique comme l'alternance de bandes claires (ou I pour isotropes) et sombres (ou A pour anisotropes), respectivement centrées autour des disques Z et bande M (Fig. 2).



**Figure 2 : Organisation du sarcomère**

Représentation schématique des éléments constitutifs d'une unité sarcomérique et leur agencement structurel (haut, longitudinal ; bas-droite, transversal). Observation des sarcomères en microscopie électronique sur coupe de muscle strié (bas-gauche, longitudinale ; bas-milieu ; transversale). Bas-droite, grossissement.

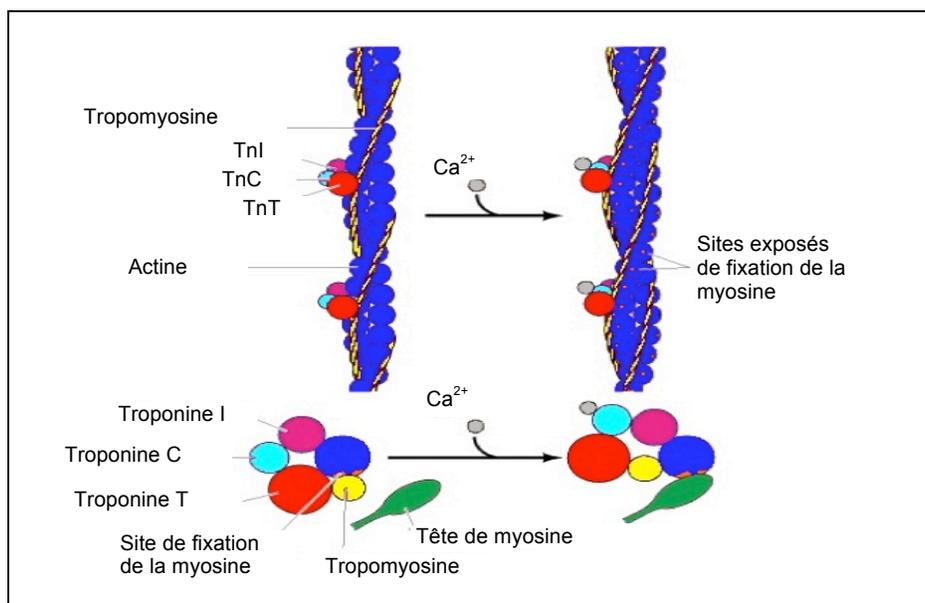
Cette différence d'aspect s'explique par la composition interne des deux régions sarcomériques : les **bandes I** ne comportent que des filaments fins et apparaissent donc plus claires, alors que les **bandes A** plus sombres contiennent les filaments épais et les extrémités des filaments

fins les chevauchant. Ainsi agencées, les fibres musculaires révèlent une **striation** transversale périodique qui donne son aspect particulier au muscle squelettique en coupe longitudinale et lui vaut son nom de muscle strié (**Fig. 2**).

#### 1.1.1.4. Les filaments : anatomie moléculaire du muscle strié

Les myofilaments sont constitués de différentes **protéines contractiles** présentes dans la majorité des types cellulaires, mais plus abondantes dans le muscle.

Les filaments fins résultent de l'assemblage de trois éléments fondamentaux, l'actine G, la tropomyosine et la troponine. L'actine G est une protéine globulaire de 42 kDa, qui se polymérise pour former un homopolymère filamenteux d'actine, appelée actine F (fibrillaire). Deux monomères d'actine F s'enroulent l'un autour de l'autre dans une hélice double pour former le filament d'actine. De son côté, la tropomyosine est une protéine filamenteuse légère, qui se lie à l'actine F en se logeant au creux des sillons formés entre ses deux hélices. Elle assure un rôle stabilisateur sur la structure du filament fin d'actine et empêche également la fixation actine/myosine. Enfin, la troponine, dernier composant des filaments fins, est en réalité un hétérotrimère de troponine C, T et I qui adopte une configuration pseudo-globulaire dans laquelle chaque constituant a une fonction spécifique. La troponine C fixe le calcium ( $\text{Ca}^{2+}$ ), la troponine T fixe la troponine C à la tropomyosine, et la troponine I, sous-unité régulatrice, inhibe la liaison



**Figure 3 : Structure et organisation du filament fin**

Les filaments épais sont constitués de centaines de molécules de myosine de classe II (sarcomériques) formant un squelette rigide. Ces molécules s'organisent en hexamères associant deux chaînes protéiques lourdes (MyHC), de 200 kDa chacune, ayant l'apparence d'une structure coudée composée d'une partie linéaire terminée par une extrémité globulaire, à deux paires de chaînes légères (20 kDa chacune), une dite essentielle (ELC), l'autre dite régulatrice (RLC), qui stabilisent la jonction entre la partie linéaire et la tête de la chaîne lourde (Berg, Powell et al. 2001). C'est au niveau de la partie linéaire C-terminale en hélice  $\alpha$  que se fait la dimérisation de la protéine. Deux monomères forment une hélice par sur-enroulement de leurs parties filamenteuses avec à leurs extrémités les têtes de myosines composées de 190 acides-aminés adoptant une conformation globulaire. Ainsi, les chaînes lourdes s'organisent en faisceaux d'où émergent à intervalles réguliers les têtes globulaires, sièges de l'**activité enzymatique** de la myosine renfermant chacune un site actif fixant l'ATP.

## 1.1.2. Fonctionnement et propriétés du muscle strié

### 1.1.2.1. La contraction musculaire

#### 1.1.2.1.1. La modulation calcique

La base de l'activité contractile du muscle squelettique repose sur la modulation de la concentration cytoplasmique de  $\text{Ca}^{2+}$ . Cette modulation fait intervenir une série d'étapes constituant le **cycle excitation-contraction-relaxation**, événement menant à la contraction musculaire. Au repos, le réticulum sarcoplasmique qui entoure les fibres musculaires est un élément de stockage du calcium où il est retenu par la calsequestrine. Le réticulum émet à intervalles réguliers des protubérances appelées citernes terminales (**Fig. 1**), au niveau desquelles se trouvent également des invaginations de la membrane plasmique, les tubules transverses (tubules T). Les points de jonction entre réticulum sarcoplasmique et tubule T sont appelés triades ou **jonctions triadiques**, un tubule T interagissant avec deux citernes terminales pour former une jonction localisée au niveau de la strie Z des myofibrilles.

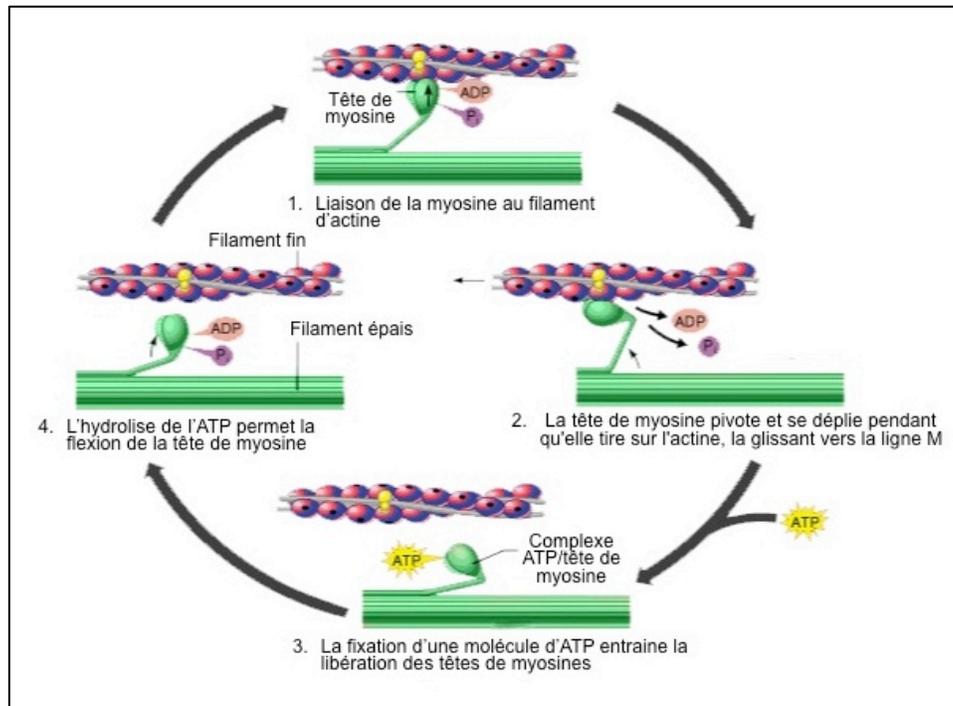
En réponse à la stimulation de son nerf moteur, la fibre musculaire développe un potentiel d'action qui se propage le long du sarcolemme jusque dans les tubules T. L'arrivée d'une vague de dépolarisations au niveau des tubules T entraîne l'ouverture de canaux calciques voltage dépendant situés sur les citernes terminales, ce qui déclenche, au niveau de la jonction triadique, un **relargage massif du  $\text{Ca}^{2+}$  sarcoplasmique** vers le cytoplasme suivant le gradient de concentration. Le calcium

ainsi libéré se fixe à la troponine au niveau de sa sous-unité contractile, la troponine C, et déclenche *in fine* la contraction musculaire.

#### 1.1.2.1.2. Le couplage excitation-contraction

Au repos, la tropomyosine est bloquée sur l'actine par le complexe de troponines empêchant ainsi l'interaction de la myosine avec l'actine. Sous cet état, la tête de la myosine contient une molécule d'ATP et est inclinée de 45°. Il n'y a pas de contraction musculaire. L'arrivée de la vague de calcium qui va être fixée par la troponine C entraîne un **changement conformationnel** de cette dernière, ce qui induit une rotation de la tropomyosine autour de l'actine d'environ 30°. Dans cette position, qui persiste sur toute la durée de présence du  $\text{Ca}^{2+}$ , le site de liaison de la myosine sur l'actine est démasqué. Cet état permet le développement d'une liaison faible, entre les deux protéines, ce qui active le site catalytique de l'enzyme et provoque l'hydrolyse de l'ATP en ADP+Pi. La liaison précédemment formée devient alors forte et, en absence d'ATP, la tête de la myosine opère une flexion de 45° supplémentaires, pour atteindre un angle maximal de 90° avec la partie linéaire. Les têtes de myosines rentrent alors en contact direct avec l'actine, et leurs mouvements sur les filaments fins permettent le **glissement relatif des filaments d'actine et de myosine** (*théorie de Hugh Huxley, 1954*) (Huxley 1969). Ceci a pour conséquence le raccourcissement des sarcomères, d'une taille de 3,4  $\mu\text{m}$  en condition relâchée à près de 1  $\mu\text{m}$ , qui se manifeste par le rapprochement de deux disques Z adjacents (**Fig. 3 et 4**).

Ce phénomène se produisant simultanément pour tous les sarcomères des différentes fibres composant les différents faisceaux d'un même muscle, il en résulte un raccourcissement global du muscle constituant la **contraction musculaire**. Selon le type de muscle et le stade de développement de l'individu, il existe une diversité dans la vitesse de contraction musculaire qui est permise par la variabilité de types de chaînes lourdes et légères de la myosine présentes sous différentes isoformes dans l'organisme.



**Figure 4 : Couplage excitation-contraction dépendant de l'ATP**  
 (Modifié d'après Elaine N. Marieb, *Human anatomy & physiology* 5<sup>ème</sup> édition)

#### 1.1.2.1.3. Le relâchement musculaire

La phase de repos, ou relaxation, résulte d'une **dissociation du  $\text{Ca}^{2+}$**  de la troponine C suivie de sa recapture dans le réticulum sarcoplasmique sous l'action de la pompe à calcium SERCA (sarco/endoplasmic reticulum  $\text{Ca}^{2+}$ -ATPase), dépendante de l'ATP. Si le  $\text{Ca}^{2+}$  est toujours présent il y a nouvelle contraction ; des cycles successifs d'attachement, de mouvements latéraux et de séparations des deux types de filaments se succèdent jusqu'à épuisement des réserves en  $\text{Ca}^{2+}$  intrasarcoplasmique, la relaxation totale n'étant permise qu'après le retour final du  $\text{Ca}^{2+}$  sarcoplasmique à sa concentration basale.

#### 1.1.2.2. *La plasticité musculaire*

Le muscle squelettique est le tissu de l'organisme soumis aux plus fortes contraintes physiologiques : croissance musculaire, sollicitation intense lors de la pratique sportive, modulations répétées de sa masse... Il s'est donc adapté à ces demandes variées en devenant un **tissu** hautement **dynamique** aux **propriétés régénératives** et à la **grande plasticité** capable de supporter d'importantes variations anatomiques.

#### 1.1.2.2.1. Régénération/dégénération

Dans le cas d'une lésion musculaire, les fibres du tissu atteint peuvent présenter des signes de souffrance intracellulaire comme une **nécrose**, une dégénérescence vacuolaire... Dans les premières heures qui suivent le traumatisme, un nettoyage de la zone lésée par les cellules inflammatoires (macrophages, polynucléaires) ne laisse des fibres nécrotiques que leurs **membranes basales** (matrice extracellulaire à l'interface de l'épithélium et du tissu musculaire qui contribue à la cohésion structurale de l'épithélium et, par sa perméabilité, régule les échanges de molécules) qui vont alors servir d'ossature à la **régénération musculaire**. Un pool de cellules satellites progénitrices issues des membranes basales des cellules lésées mais migrant également des cellules saines avoisinantes, entre alors en différenciation pour donner des myoblastes qui vont coloniser la niche vacante. L'alignement et la fusion de ces myoblastes en myotubes, évoluant en fibres musculaires, permettent la régénération de la zone blessée.

Ce processus régénératif fait aussi intervenir la fibronectine et les collagènes de type 3, 4 et 5 qui forment un réseau d'ancrage pour les fibroblastes, composants du futur tissu conjonctif. Enfin, pour que la régénération soit possible deux facteurs sont indispensables : la **vascularisation** de la zone blessée qui va apporter l'oxygène nécessaire à la prolifération myoblastique et fibroblastique et l'**innervation** de la jeune fibre musculaire qui va permettre sa maturation. Cette réinnervation qui est permise par une repousse des fibres nerveuses atteintes ou bien se fait à partir des fibres adjacentes intactes, est suivie de l'établissement de nouvelles jonctions neuro-musculaires. En l'absence d'innervation la nouvelle fibre musculaire s'atrophie et régresse au stade de myotube.

#### 1.1.2.2.2. Hyperplasie/hypertrophie

L'**hyperplasie** est un terme médical désignant une augmentation de volume d'un tissu ou d'un organe due à une augmentation du nombre de ses cellules, alors que l'**hypertrophie** définit une augmentation propre du volume des cellules. Le muscle strié est capable des deux phénomènes, il peut réguler positivement et négativement sa masse mais est aussi le siège de modifications qualitatives.

L'hypertrophie est provoquée par l'augmentation de la synthèse protéique en réponse à un exercice musculaire intense. L'accroissement du nombre de filaments d'actine et de myosine favorise le développement de nouvelles myofibrilles, ce qui provoque l'augmentation des surfaces de section des fibres musculaires. Bien qu'il soit généralement admis que ce processus soit la cause principale de l'augmentation de la masse musculaire en réponse à l'entraînement en force, l'hyperplasie, résultant

de la prolifération et différenciation de cellules satellites, semble dans certains cas être également impliquée.

### 1.1.2.3. *Le myocarde : un muscle strié unique*

Le **myocarde** est le tissu musculaire du cœur. C'est un muscle épais et creux se contractant de manière rythmique qui est composé de cellules musculaires cardiaques spécialisées appelées **cardiomyocytes**. Ces cellules diffèrent des cellules des autres muscles striés en cela qu'elles sont à la fois intétanisables (incapables de contraction prolongée), excitables, douées d'automatisme, conductrices (elles transmettent l'excitation) et contractiles. Les cardiomyocytes sont composés de filaments parallèles présentant une striation identique à celle observée dans les muscles squelettiques, mais sont en revanche séparés par de courtes travées de tissus conjonctifs riches en vaisseaux sanguins nécessaires à la forte demande énergétique, ce qui en fait, au contraire des myocytes des autres muscles striés, des cellules isolées.

Une autre caractéristique histologique de ce tissu est la présence de stries très marquées, dites scalariformes, qui permettent un contact étroit entre les extrémités distales des cellules. Ce contact résulte en une diffusion rapide de la contraction à travers tout le myocarde qui se comporte donc comme un syncytium bien que ses cellules soient isolées.

## 1.2. Pathologie du muscle strié : aperçu sélectif

### 1.2.1. Myopathie ou dystrophie ?

Le terme myopathie désigne les maladies musculaires en général, alors que les dystrophies musculaires sont des formes particulières de myopathies caractérisées par une dégénérescence (nécrose) des cellules musculaires. L'évolution clinique et l'aspect morphologique des biopsies musculaires suggèrent des **mécanismes physiopathologiques distincts** pour les différentes entités regroupées sous ces nomenclatures. La pathologie des myopathies distingue deux grands groupes selon que la maladie soit **héréditaire ou acquise**. Les myopathies d'origine génétique, concernant directement ce travail de thèse, sont classées en trois sous-groupes selon l'âge d'apparition des premiers symptômes, la rapidité du développement de la maladie et le mode de transmission.

### 1.2.1.1. *Les myopathies d'origine génétique*

Parmi ces myopathies on distingue entre autres :

- les **dystrophies musculaires**, dans lesquelles une altération primaire des fibres musculaires provoque leur destruction progressive et qui peuvent être des formes infantiles (dystrophie musculaire congénitale, dystrophie musculaire de Duchenne,) ou de début à l'âge adulte (dystrophie myotonique de Steiner, dystrophie facio-scapulo humérale, dystrophie des ceintures). Elles se manifestent du point de vue histopathologique par la présence de fibres en nécrose-régénération et par une augmentation du tissu conjonctif endomysial (fibrose endomysiale).

- les **myopathies métaboliques**, qui sont dues à des déficits enzymatiques au niveau des voies métaboliques permettant au muscle de produire de l'énergie. Ces voies peuvent être celles de la dégradation des sucres, glycolyse et glycogénolyse (maladie de Pompe) ; du métabolisme des graisses (lipidose) ; mais aussi celles intervenant dans le transport et l'oxydation des acides gras dans la mitochondrie (myopathies mitochondriales).

- les **myopathies congénitales**, qui sont attribuées à une anomalie du développement et de la maturation des fibres musculaires lors de la période fœtal. Leur définition repose sur l'observation au sein des fibres musculaires, par techniques histoenzymologiques et ultrastructurales, d'anomalies structurelles spécifiques (**Tableau en annexe 1**).

### 1.2.1.2. *Le cas des myopathies à début précoce*

Ce groupe dans lequel ont été classés tous les patients étudiés dans cette thèse inclut aussi bien les **dystrophies musculaires congénitales** (DMCs), pour lesquelles les protéines responsables de la pathologie appartiennent en général au réseau membranaire permettant l'interaction avec la matrice extracellulaire, que les **myopathies congénitales** (MCs), caractérisées par des anomalies de protéines du sarcomère, du réticulum sarcoplasmique ou des voies de transports cellulaires permettant la maturation des protéines.

Malgré les récentes avancées sur la compréhension des mécanismes physiopathologiques des myopathies à début précoce, il reste de nombreux cas de MCs et de DMCs pour lesquels les bases génétiques et moléculaires restent inconnues.

### 1.2.2. Les dystrophies musculaires congénitales

Les **dystrophies musculaires congénitales** (DMCs) forment un groupe hétérogène de maladies génétiques à transmission autosomique récessive. Elles sont présentes dès la naissance ou les premiers mois de la vie et se manifestent cliniquement par une hypotonie néonatale, une faiblesse musculaire, un retard de l'acquisition de la marche et des rétractions musculaires précoces conduisant parfois à des déformations articulaires. A ces atteintes musculaires peuvent s'associer des malformations du système nerveux central provoquant un retard mental plus ou moins important, des troubles de la déglutition ou des troubles respiratoires. Elles se caractérisent morphologiquement par des phénomènes de nécrose/régénération, par la présence de noyaux internalisés et par la présence d'une fibrose visibles sur les biopsies musculaires (Dubowitz, 1995) Le phénotype clinique varie de maladies sévères et souvent fatales à des atteintes plus modérées compatibles avec la survie à l'âge adulte.

La nosologie distingue 4 catégories majeures de DMC (Lisi and Cohn, 2007) qui sont fonction des défauts génétiques identifiés et classées selon que les protéines codées par les gènes mutés sont exprimés dans la matrice extra-cellulaire, la membrane cellulaire, le golgi ou le réticulum sarcoplasmique. On distingue donc :

#### 1.2.2.1. Les DMC avec déficit en laminine- $\alpha$ 2 (mérosine)

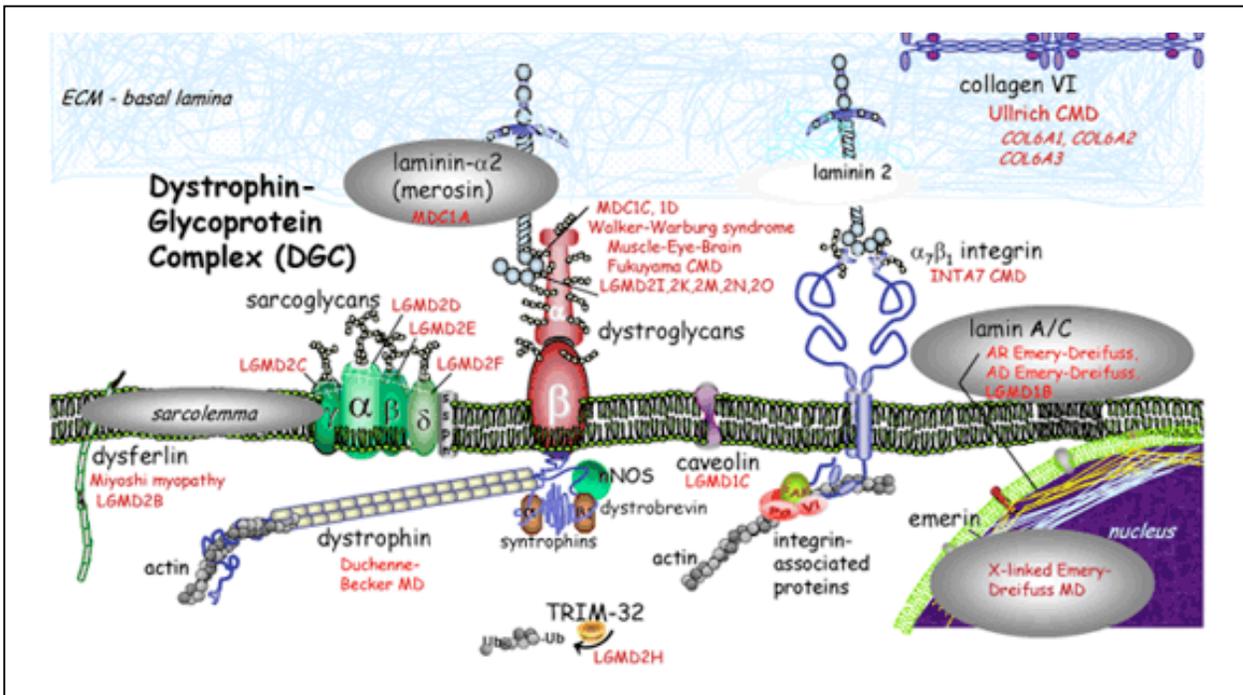
Elles correspondent à la forme la plus répandue des DMC, représentant 40 à 50 % des cas. Les laminines forment un groupe de glycoprotéines de haut poids moléculaire (400-900kDa) localisées dans la matrice extra-cellulaire (Fig. 5). Les protéines matures sont composées de trois chaînes ( $\alpha$ ,  $\beta$  et  $\gamma$ ) formant un hétérotrimère avec plusieurs domaines fonctionnels. Les laminines forment un élément de structure essentiel de la lame basale. Elles se lient à d'autres macromolécules de la lame basale comme le collagène IV, le perlecan ou l'agrine mais également à la surface cellulaire via l' $\alpha$ -dystroglycane et l'intégrine, ses deux principaux récepteurs. Cette liaison provoque une cascade de signalisation agissant dans des processus comme la formation de la lame basale, la reconnaissance cellule-cellule, les mouvements cellulaires, la différenciation ou la survie tissulaire (Vachon et al., 1996), (Cognato and Yurchenco, 1999). Dans le muscle squelettique, les chaînes majoritaires sont  $\alpha$ 2,  $\beta$ 1,  $\gamma$ 1 et  $\beta$ 2, (Sewry et al., 1995). La chaîne  $\alpha$ 2 est codée par le gène LAMA2 (6q22-q23) dont les mutations causent la dystrophie musculaire congénitale 1A (DMC1A), qui représente la plus grande proportion des cas de DMC (Lisi and Cohn, 2007). Ces mutations provoquent une déficience quantitative ou fonctionnelle de la laminine- $\alpha$ 2 et peuvent aboutir à des déficits complets ou partiels

en mérosine (Helbling-Leclerc et al., 1995), (Hillaire et al., 1994). Il en résulte la formation d'une protéine tronquée qui n'est que partiellement fonctionnelle, et ceci à des degrés variables, ce qui explique l'extrême variabilité phénotypique.

#### 1.2.2.2. *Les dystroglycanopathies: défaut de glycosylation de l' $\alpha$ -dystroglycane*

Le dystroglycane est important pour la formation de la lame basale (Henry and Campbell, 1998). La production de souris déficientes en dystroglycane conduit à une létalité embryonnaire due à l'incapacité de former la lame basale séparant l'embryon de la circulation maternelle (Williamson et al., 1997). Cette protéine, codée par le gène DAG1, subit une modification post-traductionnelle produisant l' $\alpha$ - et le  $\beta$ -dystroglycane (Barresi and Campbell, 2006). Ancrée dans le sarcolemme, le  $\beta$ -dystroglycane lie la dystrophine, protéine intracellulaire associée au cytosquelette d'actine (Fig. 5). Du côté extracellulaire, le  $\beta$ -dystroglycane lie l' $\alpha$ -dystroglycane. L' $\alpha$ -dystroglycane, fortement glycosylé, complète le lien du cytosquelette vers la lame basale en interagissant avec les laminines (Ervasti and Campbell, 1993), l'agrine (Campanelli et al., 1994), le perlécan (Talts et al., 1999) et la neurexine (Sugita et al., 2001).

Les DMCs à transmission autosomique récessive, souvent associées à des modifications structurales des yeux et du cerveau, sont dues à des mutations des gènes POMT1, POMT2, POMGnT, FKRP, fukutin et LARGE (Cohn, 2005), (Martin, 2005), (Muntoni, 2004). Ces six gènes codent pour des glycosyltransférases, avérées ou potentielles, impliquées dans la glycosylation de l' $\alpha$ -dystroglycane. Les maladies provoquées sont le syndrome de Walker-Walburg (WWS), le syndrome Muscle-eye-brain (MEB), la DMC1C, la DMC de Fukuyama (FCMD) et la DMC1D (Fig. 5). Une cardiomyopathie dilatée est parfois associée à ces pathologies et en particulier aux mutations du gène FKRP. Dans tous les cas, la glycosylation anormale de l' $\alpha$ -dystroglycane est la caractéristique commune. Ce défaut est visible sur les biopsies musculaires où l'anticorps dirigé contre l' $\alpha$ -dystroglycane ne reconnaît plus l'épitope. Bien que cette déficience apparaisse comme secondaire, elle prend une part importante au mécanisme de pathogénéicité, d'où le terme de dystroglycanopathies.



**Figure 5 :** Schéma représentatif des protéines impliquées dans les dystrophies musculaires congénitales (pathologie indiquées en rouge). D'après University of Iowa, Department of pathology website

### 1.2.2.3. Les DMCs avec rétractions proéminentes : Collagène VI et sélénoprotéine N

**La dystrophie musculaire d'Ullrich :** Le collagène VI est un composant ubiquitaire de la matrice extra-cellulaire (Fig. 5). Cette protéine forme un réseau de microfibrilles dans la matrice extra-cellulaire qui interagit avec de nombreuses protéines et semble avoir un rôle dans la prolifération cellulaire et la protection contre l'apoptose des fibroblastes (Caldwell et al., 1990), (Irwin et al., 2003). La protéine mature est composée de 3 chaînes polypeptidiques ( $\alpha 1$ ,  $\alpha 2$  et  $\alpha 3$ ) respectivement codées par les gènes *COL6A1*, *COL6A2* et *COL6A3*. Avant d'être sécrétées dans l'espace extra-cellulaire, les trois chaînes s'assemblent de manière anti-parallèle puis s'associent latéralement en tétramères (Zhang et al., 2002). Ces tétramères s'associent ensuite bout à bout formant un réseau de microfibrilles qui interagit avec la fibronectine, le biglycane et le collagène IV. Les mutations récessives du collagène VI sont à l'origine de la dystrophie musculaire d'Ullrich (UCMD), qui représente la deuxième DMC la plus fréquente. Cette maladie est caractérisée par une faiblesse musculaire généralisée à début précoce, à progression lente, des rétractions musculaires des articulations proximales, une hyperlaxité distale et une raideur du rachis (Pepe et al., 2002). Certains patients développent une insuffisance respiratoire (Mercuri et al., 2002) mais il n'y a pas d'atteinte cardiaque. Les patients UCMD ne présentent pas de retard mental et les images d'IRM montrent le développement normal du cerveau.

Des déficiences variables en collagène VI sont mises en évidence par immunohistochimie sur les biopsies musculaires des patients; les différents degrés sont corrélés avec la sévérité du phénotype clinique (Mercuri et al., 2002), (Demir et al., 2004). La myopathie de Bethlem est une forme allélique de UCMD (Bethlem and Wijngaarden, 1976).

**La dystrophie musculaire avec rigidité du rachis (RSMD1) :** Une forme rare de DMC autosomique récessive causée par des mutations du gène *SEPNI* (Moghadaszadeh et al., 2001), (Allamand et al., 2006). *SEPNI* code la sélénoprotéine N, une glycoprotéine ubiquitaire de 70kDa localisée dans le réticulum sarcoplasmique (Petit et al., 2003). La particularité des sélénoprotéines est la présence d'une sélénocystéine où un atome de sélénium remplace l'atome de soufre. Cet acide aminé est inséré au niveau d'un codon stop UGA particulier. La plupart des autres sélénoprotéines caractérisées sont des enzymes impliquées dans les réactions d'oxydo-réduction, où la sélénocystéine est située dans le site actif, mais la fonction de la sélénoprotéine N est encore inconnue. Les patients RSMD1 présentent dès les premières années de leur vie des rétractions musculaires conduisant à une rigidité du rachis et de la cage thoracique, une faiblesse des muscles du visage et du cou ainsi qu'une hypotonie prédominante des muscles axiaux. Pendant l'enfance, les patients développent une scoliose et une faiblesse du diaphragme nécessitant une ventilation assistée. Les patients n'ont pas de retard mental et ont une IRM cérébrale normale. Ils ne présentent pas d'hypertrophie musculaire. Les biopsies musculaires montrent des signes myopathiques: une modification de la taille des fibres musculaires, une prédominance des fibres de type I et quelques noyaux internalisés. Il y a peu de phénomènes de nécrose-régénération. Beaucoup de cas présentent des minicores, c'est-à-dire des zones de déplétion en mitochondries (visibles en microscopie optique sur des colorations oxydatives) et de désorganisation sarcomérique (visibles en microscopie électronique). Les biopsies musculaires de certains patients montrent des changements dystrophiques avec une importante fibrose endo- et péri-mysiale et la présence de noyaux internalisés.

#### 1.2.2.4. Les DMCs avec déficience en intégrine $\alpha$ -7

L'intégrine  $\alpha$ 7 $\beta$ 1 est un récepteur hétérodimérique de la laminine- $\alpha$ 2 (mérosine) situé à la surface des myocytes qui permet de connecter la surface cellulaire à la lame basale (Fig. 5) (Cohn et al., 1999), (Vachon et al., 1997). Des mutations du gène *ITAG7*, qui code pour l'intégrine  $\alpha$ -7, ont été décrites chez trois patients présentant une forme de DMC autosomique récessive caractérisée par un retard des acquisitions motrices et parfois un retard mental. Une expression normale de la laminine- $\alpha$ 2 et une déficience en intégrine- $\alpha$ 7 sont observées sur les biopsies musculaires (Hayashi et al., 1998).

#### 1.2.3. Les myopathies congénitales

Les patients atteints de MCs présentent en général des signes cliniques dès la naissance tels qu'une **hypotonie** et une **faiblesse musculaire**. Le taux de créatine kinase sérique (CK) est souvent normal ou peu modifié. Typiquement, il n'y a pas d'aspect dystrophique (fibres en nécrose-régénération ou fibrose endomysiale) observable sur les tissus. Certaines formes sont bien définies cliniquement, morphologiquement et génétiquement et sont la conséquence d'anomalies de protéines sarcomériques appartenant aux filaments fins (nébuline, actine- $\alpha$ , tropomyosine, troponine T...), de protéines impliquées dans le transport membranaire (myotubularine, dynamine 2) ou la tubulation de la membrane (amphyphysine 2) ou de protéines situées dans le réticulum sarcoplasmique (sélénoprotéine N et récepteur à la ryanodine de type I) (Tableau en annexe 1).

##### 1.2.3.1. Myopathies à cores : myopathie à cores centraux (CCD), myopathie à multi-minicores (MmD)

Le diagnostic de myopathies à **cores** est posé suite à l'observation anatomique de lésions caractéristiques appelées cores, qui sont des zones focales de **déplétion en mitochondries** et de **désorganisation sarcomérique** des fibres musculaires.

Dans la myopathie à cores centraux (Central Core Disease, CCD) les cores s'étendent sur la totalité de l'axe longitudinal de la fibre musculaire du patient. Ces myopathies, longtemps considérées comme le paradigme de myopathie dominante, sont associées à des mutations du gène du récepteur à la ryanodine de type I (*RYR1*) (Jungbluth 2007) et de l' $\alpha$ -actine squelettique (*ACTA1*) (Kaindl, Ruschendorf et al. 2004).

Tous les patients inclus dans cette thèse ont fait l'objet d'un diagnostic de myopathie congénitale de type **multi-minicore** (Multi-minicore Disease, MmD). Les MmD forment un groupe de pathologies cliniquement **hétérogène** mais dont le mode de transmission est **autosomique récessif** (**Tableau en annexe 1**). Cette forme est principalement causée par des mutations homozygotes ou hétérozygotes composites de *SEPN1* (Ferreiro, Quijano-Roy et al. 2002), *RYR1* (Ferreiro, Monnier et al. 2002) ou *TTN* (Carmignac, Salih et al. 2007), codant respectivement la sélénoprotéine N, le récepteur à la ryanodine de type I et la titine. Plus récemment des mutations ont été identifiées dans les gènes *CFL2* (Ockeloen, Gilhuis et al. 2012) codant la cofiline ou encore dans *MEGF10*, un gène exprimé dans les cellules satellites (Boyden, Mahoney et al. 2012).

Dans ce cas précis, les cores présents sur les biopsies des patients sont de petites tailles, s'étendant sur quelques sarcomères dans l'axe longitudinal de la fibre, et nombreux dans chaque fibre musculaire, par conséquent ils sont qualifiés de **multi-minicores**. Les caractéristiques générales des MmD sont une **hypotonie néonatale**, un **retard moteur**, une **faiblesse musculaire généralisée** associée à une **amyotrophie pas ou peu progressive** et souvent une **atteinte respiratoire** associée.

#### 1.2.3.2. *Les myopathies myotubulaires et centronucléaires*

Les myopathies **centronucléaire** et **myotubulaire** sont définies morphologiquement par la présence de fibres musculaires ressemblant aux myotubes caractérisés par la présence de **noyaux centraux** parfois entourés d'un halo où les myofilaments sont absents. La myopathie myotubulaire est la forme la plus sévère de ces deux myopathies et est définie par une transmission liée au chromosome X (Laporte, Hu et al. 1996) avec des mutations du gène de la myotubularine (*MTM1*). La myopathie centronucléaire, faiblement évolutive et rarement associée à une cardiomyopathie (Verhiest, Brucher et al. 1976) (Gospe, Armstrong et al. 1987), peut se retrouver chez les patients sous **formes récessives** (dues à des mutations de l'amphiphysine 2, *BINI*) (Nicot, Toussaint et al. 2007) **ou dominante** dans le cas de mutations de *DNM2* codant pour la dynamine 2 (Bohm, Biancalana et al. 2012) ou de *RYR1* (Jungbluth 2007).

#### 1.2.3.3. *Myopathies à bâtonnets*

Les **myopathies à bâtonnets** (ou myopathies à némaline) sont un groupe génétiquement hétérogène présentant un mode de transmission autosomique dominant ou récessif, avec des mutations dans 6 gènes codant pour des protéines sarcomériques du filament fin : l' $\alpha$ -actine squelettique (*ACTA1*), l' $\alpha$ -tropomyosine (*TPM3*), la  $\beta$ -tropomyosine (*TPM2*), la troponine T

(*TNNT1*), et la cofiline-2 (*CFL2*) (**Tableau en annexe 1**). Sous sa forme congénitale, elle fait partie CMs les plus fréquentes

L'anomalie structurale caractéristique est la présence de bâtonnets intra-cytoplasmiques (ou nemaline bodies) correspondant à des anomalies des stries Z associées à l'interruption de l'organisation des filaments fins. Il existe une variabilité dans la sévérité de la maladie que ce soit pour les formes dominantes ou récessives. En général, les patients présentent une faiblesse musculaire du visage, du cou et des muscles proximaux des membres (North, Laing et al. 1997 ; Sanoudou and Beggs 2001).

## **2. La titine : un acteur majeur du muscle, un vrai challenge de recherche**

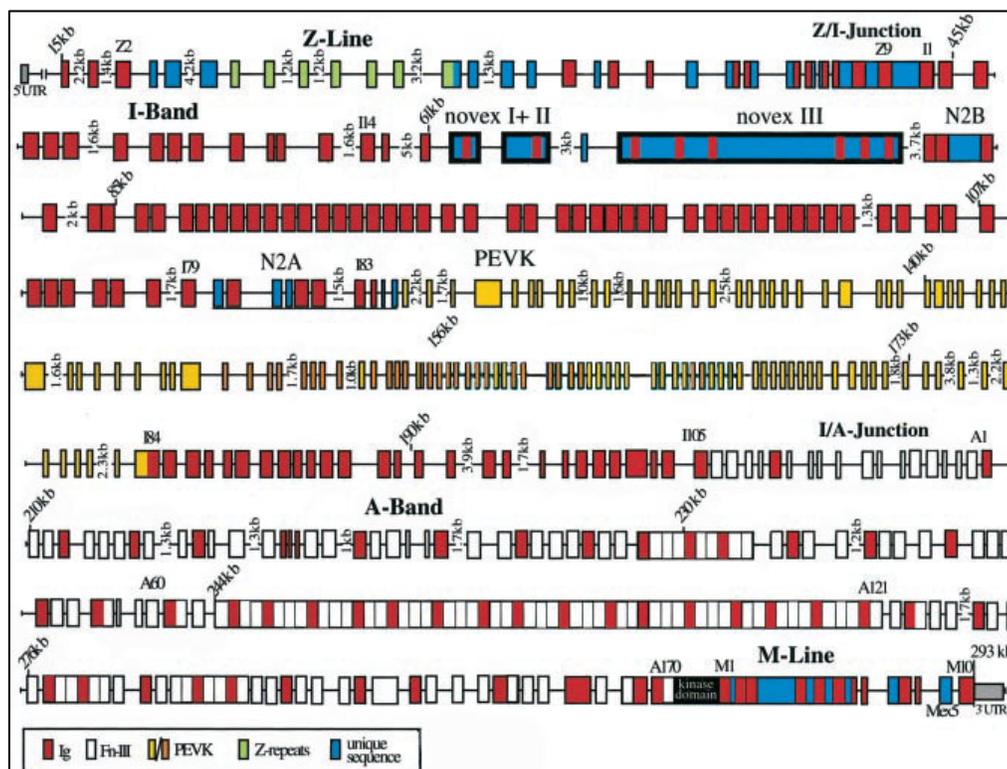
### **2.1. Structure et fonction de la titine**

#### **2.1.1. Le titan du muscle**

La *Titine* (*TTN*, OMIM#188840) est la **plus grande** protéine retrouvée chez les vertébrés. Elle constitue avec l'actine et la myosine le troisième type de filament le plus exprimé dans les muscles striés où elle représente environ 10% de la masse myofibrillaire totale. Un filament de *TTN* s'étend sur une longueur physique d'environ 1,2 µm et couvre la moitié d'un sarcomère du disque Z à la bande M (**Fig.2**) (Labeit, Barlow et al. 1990).

La séquence codante de la *TTN* consiste en **363 exons** localisés sur environ 300 kb du chromosome **2q31** et pouvant transcrire un ARN messager de plus de 100 kb (**Fig. 6**). Cette séquence native pourrait hypothétiquement être traduite en une protéine composée de 38 138 résidus pour un poids moléculaire avoisinant les **4 200 kDa** (Bang, Centner et al. 2001 ; Hackman, Vihola et al. 2002). La titine est une protéine modulaire composée à environ 90% de répétitions de classe I ou II, partageant des homologues avec les motifs fibronectine de type III (Fn-3) ou immunoglobuline C2 (Ig) respectivement (Benian, Kiff et al. 1989).

La structure primaire de la TTN est organisée en 4 régions distinctes : la **ligne Z**, extrémité localisée au niveau du disque Z constituant la portion amino-terminale (N-ter) de la protéine, la **bande I** et la **bande A** se terminant au niveau de la **bande M**, à l'autre extrémité de la protéine, en position carboxy-terminale (C-ter).



**Figure 6 : Structure génomique de TTN** (Modifié d'après Bang et al., 2001)  
La couleur orangée dans la région PEVK symbolise les exons spécifiques des isoformes fœtales.

### 2.1.1.1. Une organisation bien définie

#### 2.1.1.1.1. La ligne Z

La ligne Z de la TTN associe des domaines protéiques de type Ig et 7 structures protéiques nommées **répétitions Z** (Zr1 à Zr7) présentant un épissage alternatif. C'est en partie à travers cette ligne Z que les interactions inter-moléculaires avec les filaments fins d'actine se forment et permettent l'ancrage de la TTN dans le disque Z du sarcomère (Gautel, Goulding et al. 1996; Linke, Ivemeyer et al. 1997).

#### 2.1.1.1.2. La bande I

La bande I est composée d'éléments alternant des domaines Ig et des séquences uniques, de structures et de fonctions variées qui fournissent à la TTN l'élasticité et la tension passive nécessaire à l'élongation sarcomérique ainsi qu'au retour du muscle à sa taille de repos après le relâchement musculaire (Huebsch, Kudryashova et al. 2005). La bande I peut être décrite comme l'alignement séquentiel d'une série de domaines de type Ig, de trois séquences uniques suivis par un domaine N2-B et de deux nouvelles séries de domaines Ig organisées en tandem encadrant une région N2-A directement suivie par une région PEVK (Fig. 6) (Bang, Centner et al. 2001).

- Les séquences uniques

Parmi les derniers éléments structurels identifiés dans la TTN, ces séquences particulières, au nombre de 3, donnent leurs noms aux isoformes **novex-1**, **novex-2** et **novex-3**. Elles correspondent aux **exons 45, 46 et 48** codant respectivement 125, 192 et 216 résidus protéiques situés juste après la jonction ligne Z / bande I. Chaque séquence est spécifique de son isoforme et ne se retrouve dans aucune autre (Bang, Centner et al. 2001).

- La région PEVK

Cette portion de la protéine doit son nom à son importante composition (70%) en résidus **proline -P, glutamate -E, valine -V et lysine -K** (Labeit and Kolmerer 1995; Freiburg, Trombitas et al. 2000; Bang, Centner et al. 2001), et agit tel un **ressort moléculaire** capable d'entraîner les filaments fins grâce à ses domaines de liaison à l'actine (Linke, Kulke et al. 2002). La région PEVK présente une **séquence hautement répétitive**, siège de nombreux **épissages alternatifs** résultant en différentes isoformes aux propriétés élastiques multiples. Certains exons codant PEVK ont une expression purement fœtale (Fig. 6) et cessent d'être exprimés au cours du développement intra-utérin ou durant les premières semaines de vie. La taille de PEVK est donc maximale dans les isoformes d'expression purement embryonnaire, le nombre des exons exprimés composant ce domaine variant pour atteindre un maximum de 114 exons dans la forme adulte la plus grande.

Ces variations de la bande I, et plus particulièrement des exons PEVK (Freiburg, Trombitas et al. 2000; Trombitas, Redkar et al. 2000), qui entraînent un rétrécissement de la TTN rigidifient le muscle, et sont directement corrélées avec la capacité d'élongation des sarcomères et les propriétés

élastiques des différents types musculaires (Wang, McCarter et al. 1991).

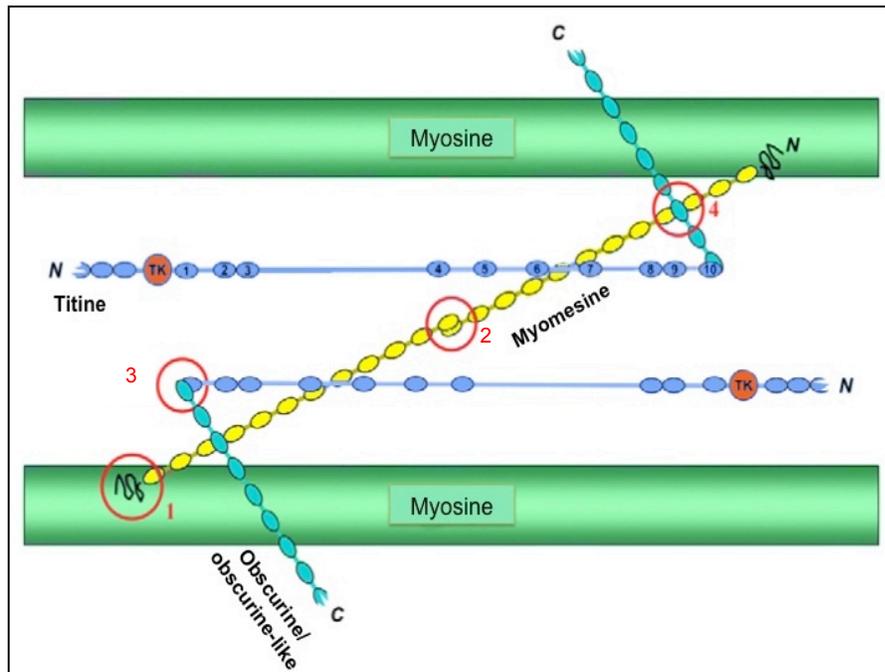
#### 2.1.1.1.3. La bande A

La seconde partie de la TTN, la **bande A**, est caractérisée par l'alternance de motifs **hautement répétés** d'une centaine de résidus, associant des **domaines Ig** et **Fn3** (Bang, Centner et al. 2001). Elle semblerait jouer un rôle de **treillis pour l'assemblage des filaments épais** auxquels elle est liée via ses multiples interactions avec la myosine et sa protéine de liaison, MyBP-C (Maruyama, Hu et al. 1987; Labeit, Gautel et al. 1992; Freiburg and Gautel 1996).

#### 2.1.1.1.4. La bande M

Les 6 derniers exons en 3' de la séquence du gène *TTN* (**Mex1 à Mex6 ou exons 358 à 363**) codent pour une région spécifique de 200 kDa appelée bande M qui comprend l'unique domaine régulateur **sérine-thréonine kinase** (TK) de la protéine, codé par Mex1 (Labeit and Kolmerer 1995; Bang, Centner et al. 2001). Essentiellement composée de domaines Ig, la bande M inclue néanmoins 7 séquences uniques (notées is1 à is 7) dont is4 est un **site de phosphorylation** de la TTN et is7, codée par l'exon Mex5, est absente de certaines isoformes suite à un épissage alternatif (Gautel, Leonard et al. 1993).

La bande M, en permettant l'ancrage des filaments épais de myosine au milieu des sarcomères participe à leur maintien structural. Cette action se fait via ses interactions avec la myomésine et le complexe formé par l'obscurine et l'obscurine-like 1 (Obscn/Obsl1). La myomésine interagit avec la myosine par son extrémité N-terminale (**Fig. 7, 1**), la partie C-terminale permet quand à elle sa dimérisation antiparallèle (**Fig. 7, 2**) et contient également un site de liaison pour le domaine m4 de la TTN (Obermann, Gautel et al. 1997). Obscn, au niveau de ses 3 premiers domaines en N-terminal, interagit avec le dernier domaine C-terminal de TTN (**Fig. 7, 3**) à l'extrémité de la protéine, mais également avec la myomésine. Enfin, TTN et myomésine interagissent avec Obsl1 (Sauer, Vahokoski et al. 2010). Cette organisation spécifique réalise ainsi le pontage entre les filaments de myosine et de titine (**Fig. 7**) (Lange, Xiang et al. 2005) ce qui permet d'assurer le maintien de la bande M lors de contraintes mécaniques.



**Figure 7 : Interactions protéiques au niveau de la bande M (Adapté de Fukuzawa, 2008)**

La myomésine interagit avec la myosine (1) et le complexe Obscn/Obsl1, lui même lié au dernier exon en C-terminal de la titine (3). L'homodimérisation antiparallèle de la myomésine en C-terminal (2) préfigure le centre de symétrie de cette structure pontée qui permet l'ancrage des filaments épais de myosine au milieu des sarcomères.

- Le domaine kinase : TK

Le domaine TK, initialement localisé dans la strie Z pendant la myofibrillogénèse (Furst, Osborn et al. 1988; van der Ven, Schaart et al. 1993; van der Ven and Furst 1997), compte parmi ses substrats la téléthonine (Mayans, van der Ven et al. 1998; Mues, van der Ven et al. 1998) et se positionne lors de la maturation au niveau de la bande M. Une très grande conservation intra- et inter-espèces de même qu'un effet pathologique des mutations répertoriées dans TK suggèrent une importance fonctionnelle de ce domaine.

**TK est associé à la fois à l'expression et au turnover de la TTN (Lange, Xiang et al. 2005).** En immunofluorescence, le domaine TK colocalise avec la périphérie du sarcomère au niveau de la bande M, où les portions C-terminales de deux TTN issues de sarcomères adjacents se superposent suivant une orientation antiparallèle (Furst, Osborn et al. 1988; Sauer, Vahokoski et al. 2010). L'activité kinase de TK est **autoinhibée** par son extrémité C-terminale régulatrice (Kobe, Heierhorst et al. 1996). Cette inhibition est provoquée par la liaison de cette queue régulatrice au site de fixation

de l'ATP, en interaction spécifique avec les résidus impliqués dans la catalyse. Le départ de la queue inhibitrice du site actif, suite à la phosphorylation d'une tyrosine et sa liaison à la calmoduline (Labeit, Gautel et al. 1992; Puchner, Alexandrovich et al. 2008) permet alors l'activation du domaine TK. Cette activation est également rendue possible par un mécanisme physique d'étirement de la TTN ayant pour conséquence la dissociation de la queue inhibitrice du site actif et le démasquage de ce dernier (Grater, Shen et al. 2005).

Enfin le domaine **TK**, dans sa conformation ouverte, est aussi impliqué dans un complexe de signalisation (**signalosome**) faisant intervenir la ligase spécifique du muscle, **MURF2**, et les deux protéines à doigts de Zinc, **nbr1** et **p62** (Lange, Xiang et al. 2005). L'activation du signalosome se fait en réponse au démasquage du site actif de TK permettant son interaction avec nbr1 ce qui a pour conséquence de diriger p62, sa protéine associée, jusqu'aux sarcomères. A son tour, p62 interagit avec MURF2 qui est un ligand du facteur de transcription SRF. En absence d'action mécanique (contraction/étirement), le signalosome est dissocié et p62/MURF2 est transloqué dans le cytoplasme entraînant une diminution du pool de SRF nucléaire et par conséquent induisant une répression de son activité transcriptionnelle (Lange, Xiang et al. 2005).

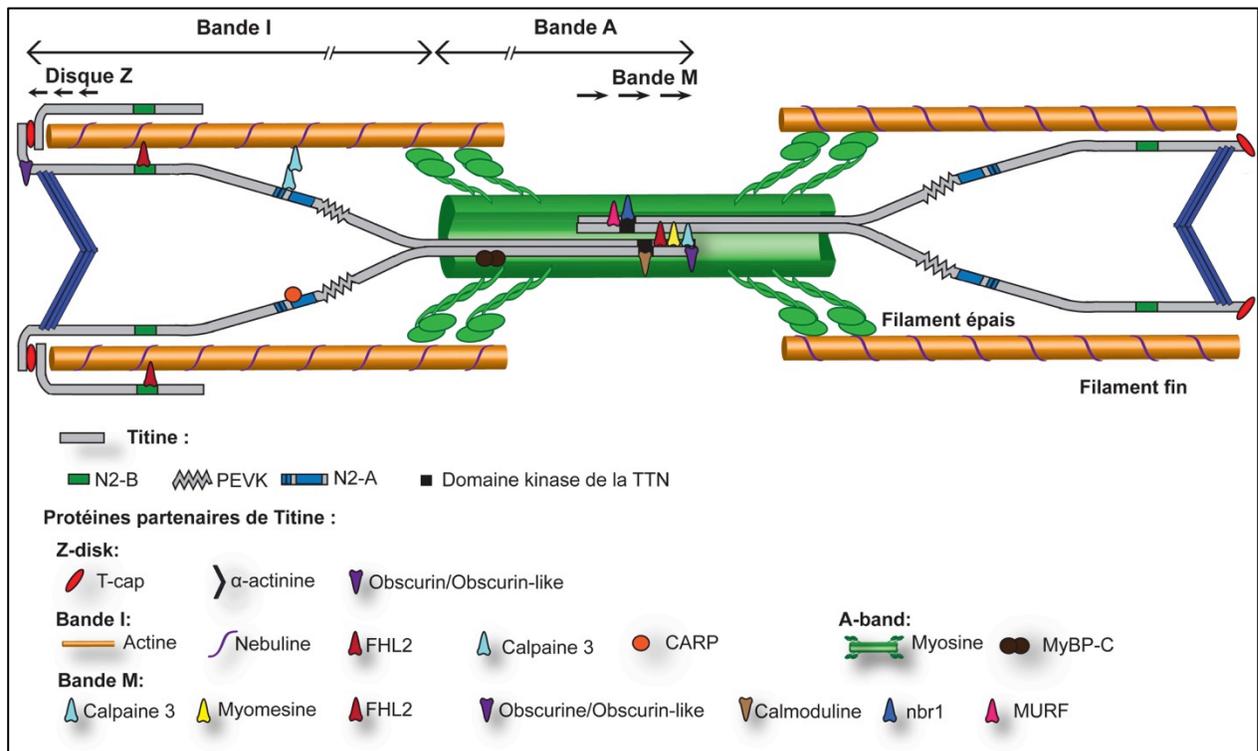
#### *2.1.1.2. Rôles multiples et nombreux partenaires*

Dans les cellules musculaires, la TTN est impliquée dans divers **rôles clefs** incluant le contrôle de l'assemblage des filaments épais de myosine (Labeit, Gautel et al. 1992) et le maintien des composants des sarcomères en place (Freiburg and Gautel 1996). Elle assure également la transmission de la force contractile au disque Z et la régulation de la tension de repos au sein de la bande I (Labeit and Kolmerer 1995; Linke, Ivemeyer et al. 1997; Gutierrez-Cruz, Van Heerden et al. 2001). Enfin, elle assure un rôle de médiateur dans diverses voies de signalisation à travers son **domaine kinase catalytique** TK (Mayans, van der Ven et al. 1998; Trinick and Tskhovrebova 1999; Miller, Musa et al. 2003).

Chez l'homme, une expression ubiquitaire de la TTN a été démontrée dans les tissus connectifs et tissus mous, mais si les rôles régulateurs de la TTN dans les muscles striés commencent à être bien établis, ses fonctions putatives dans les tissus non-musculaires restent mal définis et font encore l'objet de spéculations.

La taille inhabituelle de la TTN et ses multiples fonctions la rendent parfaitement adaptée à son rôle de **protéine architecturale** et de **traceur moléculaire** pour la myofibrillogénèse à travers notamment ses capacités de liaison à une pléthore de protéines essentielles à la fonction musculaire (**Fig. 8**) (Trinick 1994; Houmeida, Holt et al. 1995). Tous les domaines de la TTN sont impliqués dans

ces multiples interactions, du disque Z - avec entre autres la téléthonine et l' $\alpha$ -actine squelettique (codées par les gènes *TCAP* et *ACTA1*) (Gregorio, Trombitas et al. 1998; Mues, van der Ven et al. 1998; Young, Ferguson et al. 1998) - jusqu'à la bande M, où la TTN forme un complexe ternaire avec de multiples partenaires structurels et fonctionnels. Ce complexe inclut l'obscurine (Obscn) par liaison directe avec la protéine obscurine-like 1 (Obsl1) (Pernigo, Fukuzawa et al. 2010), la calmoduline (CaM) (Amodeo, Castiglione Morelli et al. 2001), la famille des myomésines (incluant une myoméline d'expression ubiquitaire, une isoforme spécifique aux fibres rapides appelée protéine-M et une isoforme spécifique aux fibres lentes dénommée myoméline 3 respectivement codées par les gènes *MYOM1*, *MYOM2* et *MYOM3*) (Agarkova, Ehler et al. 2003; Agarkova and Perriard 2005; Schoenauer, Lange et al. 2008) et la calpaïne 3 (CAPN3) (Beckmann and Spencer 2008)

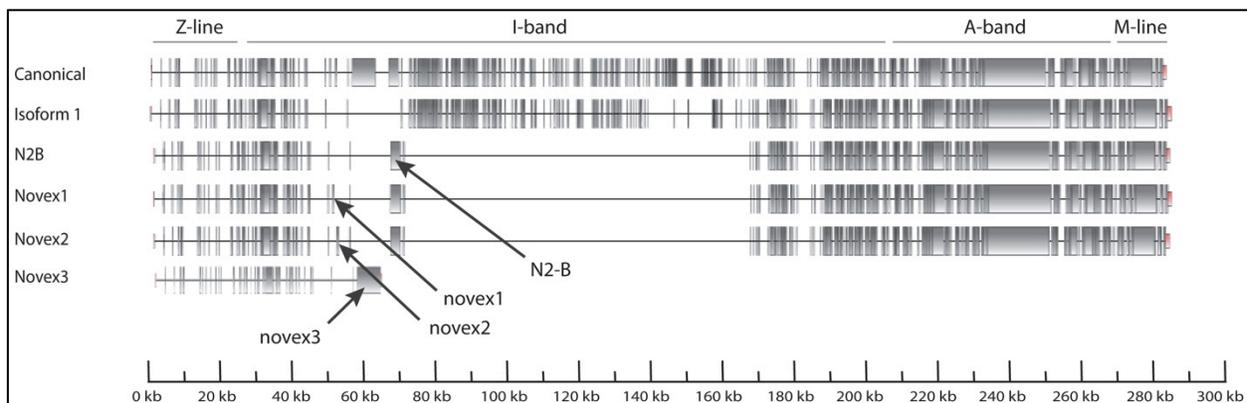


**Figure 8 : Représentation schématique de l'intégration de la titine dans le sarcomère et de ses principales interactions avec des protéines partenaires.**

## 2.1.2. Epissage alternatif et différentes isoformes

### 2.1.2.1. Les quelques séquences décrites

L'existence de l'isoforme canonique de TTN (prédite aux alentours de 4200 kDa) qui exprimerait les 363 exons identifiés n'a jamais été expérimentalement confirmée, en revanche les séquences complètes de 6 isoformes humaines sont disponibles sur le site du NCBI (**novex-1** [NP\_597676.3], **novex-2** [NP\_597681.3], **novex-3** [NP\_596870.2], **N2-A** [NP\_596869.4], **N2-B** [NP\_003310.4] et **TTN isoforme1** [NP\_001243779.1]). La plus petite de ces isoformes exprime l'exon unique **novex-3** qui est reconnu par la machinerie d'épissage comme un **exon terminal alternatif** résultant en la synthèse ubiquitaire d'une TTN tronquée en C-terminal et n'exprimant que les 48 premiers exons des 363 de *TTN* pour un poids moléculaire de **616 kDa**. A l'opposé du spectre, l'**isoforme 1**, d'expression spécifique au soleus, exprime 312 exons de *TTN* pour **3700 kDa** (Freiburg, Trombitas et al. 2000; Bang, Centner et al. 2001). Entre les deux, l'isoforme prédominante du **muscle squelettique**, **N2-A**, présente la plus grande région PEVK mais n'exprime pas l'exon N2-B ; en opposition à l'isoforme purement **cardiaque N2-B** dans laquelle n'est pas traduite une partie des exons codant pour les répétitions PEVK. Les deux dernières isoformes, la purement **cardiaque novex-1** et la **cardiaque et squelettique novex-2** diffèrent seulement de l'isoforme N2-B par l'expression des deux exons, auxquels elles doivent respectivement leurs noms (**Fig. 9**).



**Figure 9 : Différentes isoformes de la titine**

Schématisme des séquences génomiques des isoformes de TTN identifiées dans la base de données NCBI. Les blocs gris représentent les exons, les lignes continues figurent les introns et les blocs roses les séquences UTRs. La position relative des exons au regard des régions principales de la protéine est indiquée en haut de la figure, l'échelle en bas de la figure donne une indication de la taille en kb des isoformes. Les positions de certains exons caractéristiques sont indiquées par des flèches.

### 2.1.2.2. Un champ infini de possibilités

Au vu de sa complexité, le gène *TTN* pourrait en théorie générer plus d'un million de variants d'épissages issus de la seule région PEVK d'un unique individu (Guo, Bharmal et al. 2010). La dichotomie qui a longtemps régné dans le domaine de la TTN, et qui voulait que les régions N2A et N2B soient respectivement restreintes à une expression purement squelettique et cardiaque, a été mise à mal suite à l'observation, dans le myocarde, d'une TTN associant les deux domaines et conséquemment appelée N2BA. De plus, il semblerait que les seules variations de N2A puissent donner plusieurs formes de N2BA (Guo, Bharmal et al. 2010; Linke and Kruger 2010).

Les isoformes les plus grandes de la TTN sont exprimées durant les périodes fœtales et prénatales à la suite desquelles elles sont soumises à des événements d'épissage restreints à la bande I menant à une profusion d'isoformes différenciellement épissées et dont les formes adultes les plus longues se retrouvent dans les muscles squelettiques (Granzier and Irving 1995).

### 2.1.2.3. Les TTN non musculaires

Une longue isoforme de la TTN, originellement nommée **smitin** ou encore sm-titin pour « *smooth-titin* » (Kim and Keller 2002; Labeit, Lahmers et al. 2006) a également été identifiée dans l'appareil contractile du **muscle lisse**. Des expériences *in-vitro* ont montré que la sm-titin interagissait comme son homologue du muscle strié avec la myosine et l' $\alpha$ -actinine de muscles lisses (Chi, Olenych et al. 2005; Chi, Simon et al. 2008). Ces résultats laissent à supposer que ces interactions jouent un rôle capital dans l'établissement, le maintien ainsi que l'organisation dynamique de l'appareil contractile du muscle lisse. La TTN possède également un signal de localisation nucléaire fonctionnel à son extrémité N-terminale (Qi, Chi et al. 2008) en accord avec l'observation expérimentalement démontrée d'une **expression nucléaire** de la protéine. Sous cette forme, la TTN semble prendre part à la condensation des chromosomes durant le cycle mitotique et pourrait lier la chromatine et/ou l'actine nucléaire à l'enveloppe nucléaire durant l'interphase, assumant ainsi une fonction dans le maintien de la structure des chromosomes (Machado and Andrew 2000; King and Jhou 2010).

## 2.2. Mutations de la TTN et pathologies associées

### 2.2.1. Pertinence génétique

Au début de ma thèse en octobre 2008, seulement 15 mutations de *TTN* avaient été identifiées

chez des patients présentant diverses myopathies impliquant seulement les muscles squelettiques ou seulement le cœur. Avec l'avènement du séquençage de nouvelle génération le nombre de mutations *TTN* identifiées a rapidement augmenté et il existe actuellement environ une centaine de mutations de *TTN* décrites comme pathologiques dans les 3 bases de données majeures que sont **OMIM** (<http://omim.org/entry/188840>), **HGMD** (<http://www.hgmd.cf.ac.uk/ac/index.php>) et **LOVD** (<http://www.dmd.nl/>). Dans un souci de clarté les variants mentionnés dans cette thèse ont été renommés en utilisant une séquence de référence incluant les 363 exons codants identifiés dans le gène *TTN*, et ce suivant les recommandations de l'HGVS (Human Genome Variation Society) (pour plus de détails sur la nomenclature voir (den Dunnen and Antonarakis 2000)). Cette séquence de référence a été construite par les curateurs *TTN* de la base de données LOVD et peut être directement téléchargée sur leur site internet ([http://www.DMD.nl/seqs/TTN\\_codingDNA.html](http://www.DMD.nl/seqs/TTN_codingDNA.html)). Dans le tableau 2 qui regroupe les variants les plus notables de *TTN*, une colonne a été ajoutée qui fournit également la nomenclature utilisée lors de la première publication de chaque mutation ou bien la nomenclature la plus usitée lorsque celles-ci diffèrent.

#### 2.2.1.1. *Types de variation*

La centaine de variations de séquence de *TTN* rapportée ici (**Tableau, en annexe 3**) comprend **toutes les catégories de mutation existantes**: 24 faux-sens, 25 décalages du cadre de lecture (18 délétions, 5 insertions, 1 indel, 1 duplication), une duplication majeure de 18kb menant à une variabilité du nombre de copies (CNV), 29 mutations non-sens et 19 variants de site d'épissage. Il y a également une mutation particulière appelée FINmaj, qui consiste en un événement d'insertion-délétion dans lequel deux blocs de 11 nucléotides sont interchangeés (c.100076\_100086del11ins11) résultant en une substitution de 4 résidus sans décalage du cadre de lecture (Hackman, Vihola et al. 2002).

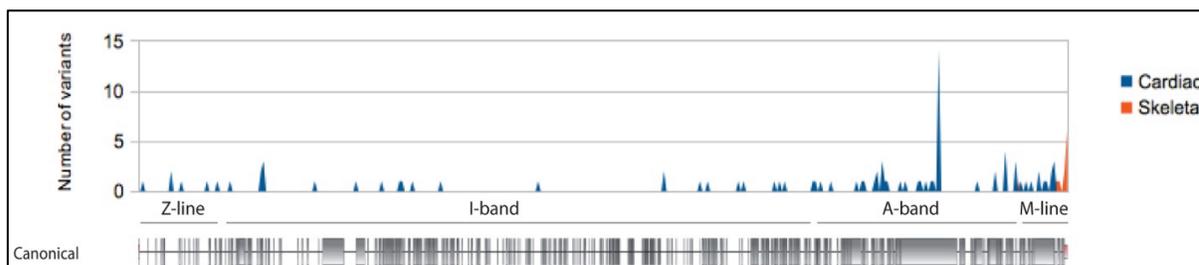
#### 2.2.1.2. *Mode de transmission*

La majorité des changements de séquences identifiés suivent une transmission de type autosomique dominante à l'exception de 3 combinaisons de variants hétérozygotes composites retrouvées chez 3 patients différents, ainsi que de 2 délétions identifiées dans la bande M (c.98963delA, c.97820\_97827delACCAAGTG) (Carmignac, Salih et al. 2007) présentant une transmission récessive. **La plupart des variants sont privés**, n'occurant que dans des cas sporadiques isolés ou au sein de patients d'une même famille. Malgré cela, une délétion identifiée dans Mex6 chez deux familles espagnoles non-apparentées (Hackman, Marchand et al. 2008), un faux-sens

dans Mex1 chez deux familles suédoises (Lange, Xiang et al. 2005) et la mutation FINmaj illustrent la possibilité d'un partage de certaines mutations *TTN* par **effet fondateur**. La mutation FINmaj s'érige une fois encore en exception, elle existe sous un double état allélique associé à deux phénotypes d'apparition et de sévérité différentes. Lorsque l'allèle muté est présent en une seule copie, à l'état hétérozygote, chez un patient ce dernier développe une dystrophie musculaire tibiale tardive (TMD), alors que les patients présentant l'allèle muté à l'état homozygote sont atteints par une dystrophie musculaire des ceintures de type 2J (LGMD2J, OMIM#608807) débutant dans l'enfance.

### 2.2.1.3. Répartition sur le gène

Parmi les 363 exons de *TTN*, 118 codant différentes parties de la protéine arborent un ou plusieurs variants décrits dans la littérature. Quatre de ces variants interviennent dans les domaines localisés physiquement au niveau du **disque Z**, 28 sont présents dans la **bande I** (3, 1 et 4 polymorphismes étant spécifiquement associés aux régions N2-B, N2-A, et PEVK respectivement). Outre les 55 changements recensés dans la **bande A**, 13 variations supplémentaires sont associées à la **bande M** (Fig. 10). Ce nombre élevé de mutations au sein de la bande M (13%) au regard de sa taille (seulement 6, 5% du gène) pourrait être indicateur d'un « hotspot » de mutations, mais il faut prendre en compte le fait que cette région a historiquement fait l'objet de plus d'études que toutes les autres parties de *TTN*. En réalité avant l'avènement du séquençage de nouvelle génération et le multiplexage des analyses, séquencer l'intégrité du gène *TTN* était un travail de « titan » que très peu d'équipes de recherche avaient entrepris. Il est donc prématuré de qualifier la bande M de site préférentiel de mutation tant que le reste du gène n'aura pas fait l'objet d'un plus grand nombre d'études génétiques.



**Figure 10 : Répartition des mutations de *TTN* relativement à l'isoforme canonique du gène**

Ce graphique incorpore les 101 variants à effet pathogène répertoriés dans le tableau 2 (cf annexe). Les phénotypes cardiaque ou squelettique associés à ces variants sont représentés en utilisant les couleurs, bleu et orange respectivement.

#### 2.2.1.4. *Variants de signification clinique inconnue ou incertaine*

En plus des variants dont la causalité a été expérimentalement prouvée, beaucoup de variants dont la signification clinique reste incertaine (**VUS** pour variants of unknown significance) ont été rapportés dans la littérature. Ces variations dans la séquence codante du gène, bien que publiées comme étant des mutations pathologiques, ne sont pas toujours clairement liées aux phénotypes observés chez les patients. Certains VUS n'ont pas fait l'objet de recherches dans une population contrôlée et/ou d'interprétation clinique de leur impact fonctionnel, et requièrent ainsi des analyses plus approfondies avant d'être qualifiés de réelles mutations. C'est le cas de beaucoup des variations de séquence récemment rapportées dans les travaux publiés par **Herman et al.** (Herman, Lam et al. 2012).

Suite à l'analyse par séquençage de nouvelle génération du gène *TTN* dans une cohorte de 543 cas de cardiomyopathies et 249 contrôles, les auteurs de l'article mentionné ci-dessus ont rapporté 72 variants tronquants (nonsense, décalage du cadre de lecture ou variant d'épissage) ; tous sont hétérozygotes à l'exception d'un changement homozygote retrouvé dans une famille consanguine et de deux combinaisons hétérozygotes composites chez deux patients. Dans ces trois derniers cas, les polymorphismes ségrègent parfaitement avec les phénotypes cliniques observés, en revanche pour la majorité des familles analysées restantes, la ségrégation des variants avec le phénotype ne peut pas être clairement affirmée, soulevant la probabilité de faire d'eux des VUS. Cette conjecture réfuterait l'hypothèse des auteurs associant les pathologies des patients à un effet dominant négatif dû à l'expression d'une isoforme tronquée de *TTN* aux propriétés toxiques, et rejoindrait nos observations précédemment décrites d'une absence d'effet pathogénique d'une mutation tronquante de *TTN* à l'état hétérozygote (Carmignac, Salih et al. 2007).

### 2.2.2. Pertinence clinique

Les mutations de *TTN* décrites affectent généralement un seul type musculaire, cardiaque ou squelettique. Jusqu'à ce jour, seulement deux mutations ont été identifiées par notre équipe chez des patients présentant une atteinte conjointe des deux types de muscles striés (Carmignac, Salih et al. 2007).

#### 2.2.2.1. *Les phénotypes cardiaques*

Sur la centaine de mutations mentionnées dans ce travail de thèse (**Tableau, en annexe 3**), 84 sont associées à un phénotype exclusivement cardiaque qui peut être une cardiomyopathie dilatée

(**DCM**, OMIM#604145) ou hypertrophique (**HCM**, OMIM#613765), ou encore une cardiomyopathie ventriculaire droite arythmogène (**ARVC4**, OMIM#602087).

#### 2.2.2.1.1. La cardiomyopathie dilatée (DCM)

Les **DCM** sont un groupe **hétérogène** de maladies cardiaques caractérisées par une dilatation ventriculaire accompagnée d'une réduction de la fonction systolique (Richardson, McKenna et al. 1996; Itoh-Satoh, Hayashi et al. 2002). Avec une prévalence avoisinant les 36,5 cas pour 100 000 personnes et une transmission héréditaire dans près de 30% des cas, les DCM sont les formes les plus fréquentes de cardiomyopathie (Kamisago, Sharma et al. 2000). La plupart des cas familiaux de DCM présente un mode de transmission autosomique dominant, une apparition des premiers symptômes dans la seconde ou troisième décennie de vie associée à une pénétrance qui augmente avec l'âge, et une manifestation clinique qui consiste en général en une insuffisance cardiaque progressive (Richardson, McKenna et al. 1996). Près de 40 gènes différents codant notamment des protéines sarcomériques ont été associés aux DCM, dont l'étiologie définit 4 sous-groupes en fonction de l'impact des mutations sur : la cytoarchitecture cardiaque, l'intégrité des sarcomères cardiaques, l'enveloppe nucléaire ou encore les fonctions mitochondriales (Dellefave and McNally 2010).

Parmi les 70 mutations de *TTN* associées à une DCM (**Tableau 2, en annexe**), 6 sont des faux-sens retrouvés indépendamment sur les différentes parties du gène (3 dans la ligne Z, 2 dans la bande-I, 1 dans l'exon Mex1 de la ligne-M). Il y a également 26 non-sens décrits ; 3 sont dans la bande-I dont 2 sur l'exon cardiaque N2-B, les 23 variants restants étant situés dans la bande A. Enfin, 19 variants changent le cadre de lecture (5 dans la bande I dont 3 dans la région PEVK et 14 dans la bande A) et 18 changements sont prédits comme ayant un impact sur l'épissage de *TTN*. Le dernier événement mutationnel est l'unique CNV décrit à ce jour dans ce gène (Herman, Lam et al. 2012).

#### 2.2.2.1.2. La cardiomyopathie hypertrophique (HCM)

L'**HCM** est la cause la plus commune de mort cardiaque subite chez le jeune adulte (< 35 ans) (Frey, Luedde et al. 2012). Elle est caractérisée par une hypertrophie ventriculaire gauche et/ou droite, le plus souvent asymétrique, et engageant principalement le septum inter-ventriculaire (Richardson, McKenna et al. 1996). On observe chez le patient un dysfonctionnement diastolique des ventricules cardiaques accompagné par la désorganisation des fibres du myocarde, ainsi que la prolifération du tissu amorphe situé entre les fibres musculaires (fibrose) augmentant progressivement la rigidité du

muscle cardiaque (Arimura, Bos et al. 2009).

L'étiologie des HCM n'est pas entièrement établie, cependant un contexte familial est noté chez environ 60% des patients avec une transmission compatible avec un trait génétique **autosomique dominant**. Des analyses génétiques ont révélé que des mutations dans des gènes codant pour les myofilaments ou les protéines du disque Z des sarcomères cardiaques étaient responsables de certaines formes de la maladie. Ainsi, plus de 20 gènes ont été associés à l'HCM, avec la majorité des mutations identifiées dans les gènes *MYH7* et *MYBPC3* codant respectivement la myosine et sa protéine C de liaison (Arimura, Bos et al. 2009; Dellefave and McNally 2010; Frey, Luedde et al. 2012).

Six variants *TTN* ont été identifiés chez des patients présentant une HCM (**Tableau, en annexe 3**), 4 sont des mutations faux-sens autosomiques dominantes présentes dans la ligne-Z (c.2219G>T; p.R740L), dans le domaine N2-A (c.28280G>A; p.Arg9427His et c.28592G>A; p.Arg9531Gln) et dans l'exon 49 codant N2-B (c.11397C>A; p.S3799Y). Les 2 derniers variants ont été retrouvés chacun dans des patients en association avec une mutation dans un gène déjà établi dans l'étiologie des HCM et ont donc été classifiés dans la littérature comme VUS (variant of unknown significance) (Herman, Lam et al. 2012). Ces variants correspondent à une délétion dans la bande I (c.23798\_23810delGTCAAGATATCTG; p.Gly8250Valfs\*7) prédisant l'apparition d'un codon stop prématuré (**PTC**) ainsi qu'un variant d'épissage (c.35263+1G>T) sur le site 5' de l'exon 157 prédisant la synthèse d'une protéine tronquée dans la région PEVK.

#### 2.2.2.1.3. La cardiomyopathie arythmogène ventriculaire droite (ARVC)

L'**ARVC** également appelée dysplasie ventriculaire droite arythmogène, est une cardiomyopathie héréditaire génétiquement hétérogène. Elle est caractérisée par la perte du myocarde ventriculaire droit qui est remplacé progressivement par du tissu fibro-graisseux menant aux arythmies cardiaques puis à la mort cardiaque subite (Azaouagh, Churzidse et al. 2011). La moitié des cas d'ARVC est familiale et héritée selon un mode de transmission **autosomique dominant** avec pénétrance réduite.

En raison de sa proximité à l'un des 12 loci d'ARVC connus, *TTN* a été séquencé chez 38 patients présentant une ARVC familiale. **Huit variants** (**Tableau, en annexe 3**) non synonymes de *TTN* ont ainsi été identifiés dans 7 patients (2 étant associés selon une conformation hétérozygote composite chez un même patient). Les 8 mutations sont toutes des faux-sens, localisées dans la **bande I** (4 variants), la **bande A** (3 variants) et la **bande M** pour une des deux hétérozygotes composites.

### 2.2.2.2. Les phénotypes musculaires

Douze mutations contribuent à 3 maladies du muscle squelettique sans affectation cardiaque : la dystrophie musculaire tibiale tardive (**TMD**, OMIM#600334), la dystrophie musculaire des ceintures de type 2J (**LGMD2J**, OMIM#608807) et la myopathie proximale dominante avec atteinte respiratoire précoce d'Edström (**HMERF**, OMIM#603689).

#### 2.2.2.2.1. La dystrophie musculaire tibiale tardive (TMD)

La **TMD** est une myopathie distale de transmission **autosomique dominante** et d'apparition tardive, se manifestant autour de la quarantaine par l'affaiblissement et l'atrophie des muscles de la loge antérieure de la jambe (plus précisément au niveau des muscles tibiaux ou muscles releveurs du pied) (Udd 2012). Les patients TMD ne présentent jamais de cardiomyopathie, et la structure des sarcomères apparaît préservée lors de l'analyse par microscopie électronique (Udd, Vihola et al. 2005). La TMD a été majoritairement décrite en Finlande où sa prévalence approche les 20/100 000 (Udd 2012). Toutes les mutations TMD de *TTN* se trouvent dans les exons **Mex5** (1 délétion) ou **Mex6** (3 faux-sens, 1 non-sens, 1 délétion, 1 indel) de la bande M. Les patients TMD finlandais sont tous porteurs d'un haplotype commun indicatif d'une mutation ancestral unique transmise par un effet fondateur. Cette mutation nommée **FINmaj** consiste en un événement d'insertion délétion de 11 nucléotides localisé dans le dernier exon de *TTN* (Mex 6-363) (g.293269\_293279delinsTGAAAGAAAAA, c.102857\_102867delAAGTAACATGGinsTGAAAGAAAAA, p.Glu34286\_Trp34289delinsValLysGluLys) (Hackman, Vihola et al. 2002) (Tableau, en annexe 3).

#### 2.2.2.2.2. La dystrophie musculaire des ceintures de type 2J (LGMD2J)

La mutation **FINmaj**, qui provoque une TMD lorsqu'elle est présente à l'état hétérozygote chez les patients, est en revanche associée, à l'état **homozygote**, à la myopathie **LGMD2J** au phénotype **particulièrement sévère**.

Les premiers symptômes apparaissent durant l'enfance et concernent d'abord et surtout la ceinture pelvienne ou moins fréquemment la ceinture scapulaire avec, dans ce cas, une asymétrie dans la perte de la masse et la force musculaire. Le phénotype LGMD2J n'existe pas chez les parents porteurs hétérozygotes de la mutation FINmaj, qui eux développent le phénotype TMD plus commun (Udd, Vihola et al. 2005; Udd 2012). Une des hypothèses pouvant expliquer la différence de sévérité entre TMD et LGMD2J fait intervenir un mécanisme de perte de fonction incluant une déplétion

secondaire en CAPN3 (Hackman, Marchand et al. 2008).

#### 2.2.2.2.3. La myopathie héréditaire avec atteinte respiratoire précoce (HMERF)

La myopathie d'Edström (**HMERF**) est une maladie **autosomique dominante** décrite chez des patients suédois dont le diagnostic repose sur la présence d'une faiblesse musculaire proximale des membres supérieurs et inférieurs associée à une insuffisance respiratoire dont l'apparition survient chez le jeune adulte ainsi que l'identification de nombreuses inclusions cytoplasmiques sur les biopsies musculaires (Edstrom, Thornell et al. 1990). L'insuffisance respiratoire représentant une cause fréquente de mortalité parmi les patients HMERF (Lange, Xiang et al. 2005).

Jusqu'à très récemment une seule mutation *TTN* avait été associée à HMERF, une transition (c.97348C>T) dans le premier exon de la bande M (*Mex1*) qui entraîne la substitution d'une arginine par un tryptophane dans le **domaine TK** (publiée sous p.R279W suivant l'utilisation d'une nomenclature relative au domaine TK). Cette mutation empêche l'interaction de la *TTN* avec une protéine cargo, *nbr1*, ce qui perturbe une voie de signalisation impliquant aussi *p62* et *MuRF2* (Lange, Xiang et al. 2005). Récemment *Ohlsson et al.*, ont identifié la première mutation de la **bande A** de *TTN* associée à HMERF, (g.274375T>C; p.Cys30071Arg) (**Tableau, en annexe 3**) dans trois familles suédoises partageant un même haplotype hérité d'un ancêtre commun (Ohlsson, Hedberg et al. 2012).

#### 2.2.2.3. *La myopathie à début précoce avec cardiomyopathie fatale (EOMFC): une titinopathie à phénotype mixte*

Deux variants *TTN* ont été identifiés par notre équipe dans deux familles consanguines présentant une **cardiomyopathie dilatée fatale associée à une myopathie à multiminicores**, phénotype qui a été nommé **EOMFC** (OMIM#611705). Cette condition particulière est **la seule titinopathie congénitale purement récessive** et impliquant à la fois les muscles **cardiaque et squelettique**. Elle se manifeste dès la naissance ou la petite enfance avec une hypotonie, une faiblesse musculaire et un retard du développement moteur (Carmignac, Salih et al. 2007).

Les deux délétions identifiées dans **Mex1** (g.289385\_289392delACCAAGTG) et **Mex3** (g.291394) de la bande M, sont transmises sous un mode récessif et les porteurs hétérozygotes sont parfaitement sains. Chacune de ces délétions provoque un décalage du cadre de lecture (*Mex1*: p.His33534Hisfs\*9, *Mex3*: p.Lys33915Asnfs\*5) avec apparition d'un PTC et synthèse de protéines

tronquées. Les polypeptides ainsi traduits sont correctement intégrés dans la structure des sarcomères chez les patients porteurs de la mutation Mex3, en revanche il leur manque le domaine de liaison à CAPN3 localisé dans la portion C-terminale tronquée de la bande A (Carmignac, Salih et al. 2007).

### 2.3. Les modèles animaux

Plusieurs modèles animaux ont été créés dans les deux organismes modèles *danio rerio* (poisson zébre) et *mus musculus* (souris) dans le but de mimer des pathologies humaines ou bien d'examiner la fonction physiopathologique des différentes parties de la TTN. Au même titre que dans l'espèce humaine, les orthologues *ttn* du poisson et *Ttn* de la souris intègrent un exon N2-B d'expression spécifique au muscle cardiaque et une bande I aux propriétés élastiques.

#### 2.3.1. Le zébrafish

Parmi les poissons zèbres mutants de *TTN* utilisés en laboratoire, deux, *pik*<sup>m171</sup> et *runzel* sont majoritairement utilisés comme modèle d'étude de cardiomyopathie dilatée et d'affection du muscle squelettique.

##### 2.3.1.1. Le mutant *pickwick*

Le mutant *pickwick*<sup>m171</sup> (*pik*<sup>m171</sup>) cible l'exon cardiaque N2-B de *ttn* en exprimant une mutation récessive létale qui résulte en son épissage alternatif (Xu, Meiler et al. 2002). Le cœur du poisson mutant a une capacité contractile réduite, et ce dès les premiers battements cardiaques, la formation des myofibrilles est initiée mais la structure des sarcomères est anormale. Ce phénotype est dû au blocage de l'assemblage des unités sarcomériques pendant la myogenèse primaire qui a pour conséquence l'installation d'un désordre fonctionnel ressemblant à la cardiomyopathie dilatée observée chez l'homme.

##### 2.3.1.2. Le mutant *runzel*

C'est suite à une **mutagénèse chimique** qu'un laboratoire de l'Université de Tuebingen a obtenu le mutant *runzel*. Il présente une structure et une fonction musculaire normale à deux jours post-fertilisation (dpf), mais est un mauvais nageur et montre une perte progressive de l'organisation du tissu musculaire à partir de 5 dpf. Lors des premières étapes du développement embryonnaire, la *ttn* subit plusieurs variations qui mènent à l'expression d'isoformes variées dans le muscle adulte. Des analyses protéiques ont décelé au moins 9 *ttn* de tailles distinctes dans les muscles adultes de poissons

sauvages alors que dans les mutants *runzel* un nombre restreint d'isoformes a été isolé. Bien que ce nombre réduit d'isoformes exprimées semble suffire à la sarcomérogenèse, il ne permet pas le maintien de l'organisation myofibrillaire. Aucune anomalie cardiaque n'a été décrite chez les mutants *runzel*, confirmant un phénotype propre au muscle squelettique. A l'état homozygote la mutation *runzel* est létale entre 10 et 12 dpf, alors qu'aucune réduction de la durée de vie n'est observée chez les hétérozygotes (Steffen, Guyon et al. 2007).

### 2.3.2. La souris

#### 2.3.2.1. Modèle murin de TMD

Il existe plusieurs modèles de souris avec des mutations *Ttn*, mais un seul, le mutant *mdm* (muscular dystrophy with myositis) est un **modèle spontané** présentant un phénotype atrophique. Dans la souris *mdm*, *Ttn* a subi plusieurs réarrangements complexes incluant une délétion de 779 nucléotides et une insertion (~2,4 kb), respectant le cadre de lecture, d'un rétrotranspon LINE-1 tronqué en 5'. Ce modèle présente également une délétion de 83 acides-aminés initialement localisés dans la région N2-A, incluant 21 résidus d'un des deux sites de liaison de la CAPN3 sur la *Ttn*. Cette dernière mutation entraîne une réduction de 50 à 60 % du niveau d'expression de CAPN3 dans les muscles affectés, mimant ainsi une des caractéristiques des TMD (Garvey, Rajan et al. 2002).

#### 2.3.2.2. Modèle murin de TMD/LGMD2J (FINmaj mutation)

Une souris mutante intégrant la mutation **FINmaj** a été créée en remplaçant une portion de la séquence de Mex6 -5'GAA ATA ACA TGG 3'- par -5'GTG AAA GAA AAA 3'-, résultant en la substitution des acides-aminés natifs EITW par le motif VKEK caractéristique de la **mutation humaine**. Les souris mutantes hétérozygotes développent un phénotype dystrophique d'apparition tardive (~9 mois) impliquant les muscles distaux, alors que chez les animaux homozygotes les premiers signes surviennent au niveau du soleus un mois après la naissance puis s'étendent à la plupart des muscles vers le sixième mois. Étonnamment, chez ces derniers, au contraire des sujets humains, le cœur est également sévèrement affecté (Charton, Daniele et al. 2010).

#### 2.3.2.3. Modèle murin de DCM

L'équipe de M.Gotthardt (Radke, Peng et al. 2007) a généré un modèle de **souris DCM** en empêchant par une technique de **knock-out** (KO) l'expression de l'exon codant pour le domaine cardiaque **N2-B**. La technique utilisée n'induit pas de PTC et permet la synthèse d'une protéine TTN

qui est correctement intégrée dans le sarcomère. Les souris mutantes atteignent l'âge adulte, sont fertiles, mais présentent des cœurs sous-développés et ont une expression significativement réduite de Fhl2 (Four and a half LIM domains protein 2), consistante avec la taille réduite des myocardiocytes. La longueur des sarcomères est réduite chez les animaux mutants, mais ils gardent cependant leurs capacités d'assemblage.

#### 2.3.2.4. *Modèle murin d'analyse de bande M*

Afin d'évaluer l'importance et d'examiner le rôle de la bande M de *Ttn* dans la physiologie musculaire, une souris mutante **KO conditionnel** pour MEx1 et MEx2 (**TK et séquences flanquantes**) a été générée (Gotthardt, Hammer et al. 2003). Ils ont examiné le rôle de ce domaine pendant l'embryogenèse et ont démontré son importance dans le développement cardiaque tant dans les stades précoces que postnataux. L'excision de MEx1/MEx2 tôt dans le développement embryonnaire provoque la mort *in-utero* des souris mutantes ; au contraire, l'excision aux stades tardifs du développement embryonnaire permet la survie des souris mais cause une myopathie progressive qui aboutit à la mort autour de la cinquième semaine après la naissance. Une réduction de la taille des fibres musculaires sans changement de leur composition est observée chez les animaux mutants (Peng, Raddatz et al. 2005), qui finissent par développer une hypertrophie cardiaque menant à une insuffisance cardiaque terminale (Peng, Raddatz et al. 2007).

### 3. La génétique des maladies à transmission mendélienne

#### 3.1. Identification d'une mutation : méthodologies contemporaines

Les termes transmission d'**hérédité mendélienne**, **monogénisme** ou encore **monofactorielle** sont indifféremment utilisés pour caractériser la transmission de maladies dues à la mutation d'un seul gène et transmis des parents à la descendance via le matériel génétique. Ce mécanisme s'oppose à la transmission non mendélienne, qui concerne l'héritabilité d'un caractère à travers les générations, mais cette fois sans altération de la séquence nucléotidique. Cette transmission concerne entre autres les changements épigénétiques et de génomes d'organites comme les plasmides ou les mitochondries.

L'identification d'un **variant génétique** chez un individu malade nécessite la comparaison de l'ordonnement des quatre bases d'oligonucléotide, adénine (A), thymine (T), cytosine (C) et guanine (G) de son ADN avec une séquence de référence. L'apparition des séquenceurs automatisés

d'ADN, la détermination complète des séquences d'ADN d'organismes de référence (animaux, plantes et génomes microbiens) ainsi que la rapidité et le volume de séquences atteint par les technologies modernes ont significativement accéléré la recherche en génétique. L'évolution de la discipline depuis les premières séquences d'ADN obtenues au début des années 1970 en utilisant des méthodes laborieuses a été fulgurante et laisse présager l'avènement de nouvelles technologies aux possibilités toujours plus excitantes.

### 3.1.1. Du ciblage...

#### 3.1.1.1. *Criblage du génome (génotypage) et analyse de liaison*

Le génome humain est réparti sur 23 paires de chromosomes (22 paires d'autosomes plus une paire de chromosomes dits sexuels permettant de déterminer le genre de l'individu et appelés gonosomes), composant le **génomme nucléaire** auquel vient s'ajouter le génome mitochondrial. Ces chromosomes intègrent **3.10<sup>9</sup> paires de bases** dont la quasi intégralité est non codante. Toutefois, même si seulement 1,5% du génome est constitutif du nombre estimé de **23 000 gènes** codant le protéome humain, l'identification d'un variant de séquence responsable d'une maladie reste un réel challenge de recherche. Il a donc fallu développer des stratégies spécifiques d'analyses permettant de balayer le génome (génotypage ou screening) à la recherche de mutations.

La méthodologie permettant d'effectuer ces screens passe par l'analyse d'outils moléculaires appelés **marqueurs génétiques** et qui consistent en des **séquences polymorphes d'ADN** transmissibles à travers les générations selon les **lois de Mendel** et dont la **haute variabilité** (également observable entre les allèles d'un même sujet) permet une discrimination entre individus. Depuis les débuts des analyses génétiques, de nombreux types de marqueurs moléculaires aux propriétés distinctes ont été identifiés ; je ne m'arrêterai dans cette thèse que sur les deux dont j'ai directement fait usage. Ces marqueurs sont les **microsatellites**, séquences d'ADN formées par la répétition continue en tandem de motifs composés de 2 à 10 nucléotides (le plus commun étant le motif CA), et les **SNPs** qui sont des polymorphismes de bases nucléotidiques isolées. Les microsatellites sont abondants, ils représentent environ 1,6% du génome humain où ils sont répartis aussi bien sur les séquences codantes que non codantes, et où leur localisation présente une bonne conservation dans les espèces phylogéniquement proches. La grande variabilité des microsatellites est due à leur fréquence de mutation plus élevée que d'autres régions neutres du génome, le type de mutation concerné étant une duplication du motif répété par glissement de la polymérase lors de la réplication de l'ADN. L'utilisation des marqueurs microsatellites dans les analyses génétiques à des

fins de **génotypage** s'est démocratisée dans le milieu des années 90 ; l'utilisation des SNPs est quand à elle plus récente. Les SNPs, de par leur nature, sont beaucoup plus fréquents que les microsatellites, ils concernent environ une base sur mille et représentent près de 90% de toutes les variations génétiques humaines. Tout comme les microsatellites, les SNPs se retrouvent aussi bien dans les régions exoniques et introniques qu'intergéniques.

La comparaison des marqueurs moléculaires entre différentes espèces permet d'établir la carte de leurs similitudes et différences génétiques, et a constitué la base de la phylogénie moderne. Dans le domaine médical les marqueurs génétiques permettent, par comparaison entre patients et contrôles, d'identifier des loci d'intérêt ségrégeant (co-hérités) avec le trait pathologique, et sont utilisés dans les analyses de liaison des maladies monogéniques et d'association des maladies multifactorielles.

#### *3.1.1.2. Clonage positionnel*

L'approche historiquement utilisée pour **identifier le gène responsable d'une pathologie** consiste en une stratégie de **clonage positionnel** permettant la recherche dans le génome d'une séquence nucléotidique (gène, séquence régulatrice ou xénogénique...) responsable d'un trait observable. Cette stratégie se divise en deux étapes, l'identification de la zone génomique d'intérêt, dans laquelle est supposée être la séquence nucléotidique liée à la maladie, suivie de son séquençage permettant l'identification du gène responsable du phénotype observé.

La première phase du clonage positionnel consiste en un **balayage généralisé** du génome permettant d'identifier par localisation primaire une grande zone génomique d'intérêt en utilisant des marqueurs génétiques espacés, directement suivi par une **cartographie fine** de la région d'intérêt. Cette seconde étape, qui a pour but la réduction maximale de la zone initialement identifiée, s'effectue progressivement en augmentant séquentiellement le nombre de marqueurs spécifiques de la région étudiée. Une fois cet objectif atteint, une dernière étape de détermination de la séquence nucléotidique de la région candidate par séquençage est réalisée, permettant ainsi par comparaison avec une séquence de référence d'identifier les défauts moléculaires contenus dans cette région.

#### *3.1.1.3. La recherche appliquée aux familles consanguines : principe de la cartographie par homozygotie*

Cette approche repose sur le principe que d'une génération à une autre, les gènes d'un même chromosome ne se transmettent pas de manière indépendante mais par blocs (**loci**). Ceci est le corollaire direct du fait que les recombinaisons chromosomiques, permettant un **brassage génétique**

indispensable au maintien de la **diversité** ainsi qu'à l'adaptabilité d'une espèce, ne surviennent qu'en quelques copies par chromosome et par génération. Par conséquent, si un individu reçoit d'un de ses parents une mutation délétère, il recevra également de ce même parent les segments chromosomiques encadrant la mutation.

A partir de cette observation, la réflexion sous-jacente à la **cartographie par homozygotie** devient évidente. En effet, si l'individu malade vient d'une famille **consanguine** (impliquant que ses géniteurs partagent un **ancêtre commun**) et que l'on pose le postulat que sa pathologie est liée à la consanguinité, alors le locus contenant le gène muté serait transmis via les différentes générations, descendant du fondateur au patient. Ce dernier reçoit de son père et de sa mère le locus portant la mutation qu'ils ont eux-mêmes indépendamment reçu de leur ancêtre commun. Le patient sera donc **homozygote** pour le bloc chromosomique (d'autant plus long que le nombre de générations le séparant du fondateur est réduit) entourant cette mutation. Dans les familles consanguines, l'une des façons de localiser le gène responsable d'une pathologie sur le génome est donc de rechercher des **locus homozygotes partagés** par les seuls atteints de la famille, et hétérozygotes ou absents chez les non atteints. C'est l'application de la cartographie par homozygotie.

### 3.1.2. ... au séquençage

Une Analyse génétique basée sur une approche par clonage positionnel pure est une stratégie longue et coûteuse ; elle est donc généralement couplée à d'autres approches, comme la recherche de gènes candidats.

#### 3.1.2.1. *Stratégie directe de gènes candidats*

Cette approche consiste en l'analyse chez les patients de gènes particuliers, par étude des marqueurs microsatellites adjacents, ce qui permet de déterminer leur liaison avec le phénotype pathologique, ou bien par **séquençage direct**. La sélection des gènes se fait selon leur implication dans des pathologies pour lesquelles le phénotype est proche de celui observé chez les patients, ou encore selon le profil d'expression ou bien la pertinence fonctionnelle des protéines qu'ils codent. Ces gènes sont alors qualifiés de **gènes candidats fonctionnels** en opposition aux **gènes candidats positionnels** identifiés par analyse de liaison.

### 3.1.2.2. *Le séquençage de nouvelle génération*

L'amélioration constante des technologies de séquençage a permis l'avènement d'une forme d'analyse haut débit qui a été nommée **séquençage de nouvelle génération (NGS)** et plus récemment qualifiée de **seconde génération**. Cette technologie relativement récente, dont les débuts remontent à 2005, est caractérisée par l'utilisation d'**approches massivement parallèles** permettant de séquencer des centaines de milliers de fragments simultanément, et ce beaucoup plus rapidement et à moindre coût que par des techniques de séquençage traditionnelle de type Sanger. Ainsi, il est maintenant possible de réaliser ce qui est longtemps resté impensable, à savoir le séquençage de l'exome complet (~180 000 exons) d'un individu et même de la quasi intégralité de son génome (exons + introns).

Trois acteurs majeurs ont dominé, dès ses débuts, le marché du séquençage de nouvelle génération, **Roche**, **Illumina** et **Life Technologies** (Applied Biosystem) dont les plateformes respectives sont les séquenceurs 454, Genome Analyser et SOLiD4. Depuis, le marché s'est étoffé proposant un panel de technologies au principe et caractéristiques propres (**Fig. 11**). Comme chaque plateforme possède ses avantages et inconvénients, l'utilisation de l'une ou l'autre technologie dépendra des champs d'applications souhaités ainsi que du type d'organisme étudié. Plusieurs paramètres qualitatifs permettent de cibler son choix tel : la capacité de séquençage en nombre de base par run, la taille moyenne des reads ou amplicons et la profondeur moyenne de séquençage qui équivaut au nombre de fois dont est séquencée chaque base du génome. A noter que parmi ce panel, le PGM d'Ion torrent est le seul à connaître une évolution constante en terme de capacités de séquençage (10Mb – reads de 100b – Juin 2011 / 100Mb – reads de 100b – Sept 2011 / 100Mb – reads 200b – Nov 2011 / 1Gb – Jan 2012).

Séquenceurs 2 <sup>ème</sup> génération											
Société	Roche			Illumina				Life Technologies			
Plateforme											
Technologie	Titanium	FLX Titanium FLX +					Chip 314 Chip 316 Chip 318				
Acides nucléiques (matrice)											
Ligation adaptateurs											
Méthode d'amplification	PCR en émulsion		« Bridge PCR »				PCR en émulsion				
Méthode de séquençage	Synthèse (Pyroséquençage)		Synthèse				Ligation				
Durée de séquençage/run	10h	10h 20h	26h	8jrs	8jrs	14jrs	2h	12jrs	8jrs	8jrs	
Capacité (Mb) séquençage/run	50	500 900	1500	100000	200000	95000	>10 >100 >1000	70000	80000	150000	
Taille moyenne des reads	400	400 700	150+150	100+100	100+100	150+150	100 >100 >100	50+35	75+35	75+35	
Coût (\$) /run	1100	6200	750	10000	20000	11500	500 750 950	8150	6100	10500	
Coût machine + annexes ((K\$))	110+25	500+30	125	560	690	250	50+20	480+55	350+55	600+55	
Exactitude de séquençage (%)	99	99	99,9	99,9	99,9	99,9	99	99,95	99,95	99,99	

Figure 11 : Comparatif des différentes plateformes de NGS. D'après biorigami.com

Concernant le **séquençage de novo qui concerne mon projet de thèse**, un résultat optimal peut être obtenu par l'association de deux technologies générant à la fois des reads longs (type 454, Roche) et une profondeur conséquente (type GAIIx, Illumina) ce qui permet de pallier aux problèmes liés aux homopolymères et erreurs de séquençage (Fig. 12). Toutefois, concernant les analyses d'exome réalisées pour ce travail de thèse, elles ont été faites exclusivement avec la technologie Sequencing by Oligo Ligation and Detection, SOLiD4, dont nous disposons au centre de recherche. Les avantages de ce système est qu'il génère beaucoup de amplicons et présente une précision de séquençage élevée pour un coût relativement faible par rapport aux autres technologies. En revanche le principal inconvénient du SOLiD est qu'il génère des amplicons de taille très petites qui rendent leur alignement incertain. De plus, le séquençage est fait sous un format spécifique appelé color-space dont l'analyse, complexe, n'est possible que par couple de bases consécutives.

Enfin, le paramètre de profondeur est à prendre en considération pour les champs



d'ADN en amont du séquençage) et se concentrent autour de deux stratégies : le séquençage en temps réel et le séquençage de molécule unique. Libérées des étapes de clonage et de constitution de banques génomiques, elles sont encore plus rapides (séquençage d'un génome complet en 6h) que les meilleurs rendements actuels et permettront peut être de franchir la barrière symbolique de coût du séquençage d'un génome pour 100 dollars.

Néanmoins, nous n'en sommes qu'aux balbutiements de ces nouvelles technologies et leur utilisation est loin d'être démocratisée, d'autant plus que l'analyse et l'interprétation de l'énorme quantité de données générées est très difficile en l'absence d'autres informations. De fait, les approches classiques précédemment décrites ne sont pas encore obsolètes, et leurs applications conjointes consistent même en une stratégie efficace ayant fait ses preuves dans la recherche de mutations pathogènes pour le diagnostic moléculaire des maladies monogéniques. C'est cette stratégie que j'ai appliquée à des formes non caractérisées de myopathies congénitales à multiminicores et que je décris dans cette thèse.

## **SUJETS ET METHODES**



## 1. Patients

Les deux études génétiques réalisées portent respectivement sur les patients de 24 familles (14 cas sporadiques, 6 cas familiaux et une famille consanguine) et 21 familles (3 familiaux et 18 consanguins) ainsi que sur les membres asymptomatiques de leur famille. Les familles sont d'origines variées (Europe, Afrique, Eurasie). Les parents sont asymptomatiques. Les patients présentaient les caractéristiques cliniques et morphologiques suivantes:

- a) une faiblesse musculaire généralisée, à début précoce
- b) une cardiomyopathie primaire pour le premier groupe de patient
- c) la présence sur les biopsies de muscle squelettique ou cardiaque de lésions de type dystrophique et/ou de type minicore
- d) pas d'altération des diverses protéines modifiées dans d'autres myopathies à début précoce

Le phénotype clinique des patients a été évalué à partir de différents documents incluant selon disponibilité les photos des patients, les dossiers cliniques, les comptes-rendus hospitaliers, les clichés d'IRM musculaire et les échocardiographies. Les parents sains des patients 1, 2 et 5 ont eu un examen clinique complet pour cette étude, incluant une échographie cardiaque.

Les études de cette thèse ont été approuvées par le comité consultatif national d'éthique du groupe Hospitalier Pitié-Salpêtrière (CPP-Ile de France-VI). Les consentements écrits, libres et éclairés ont été obtenus des patients, de leurs parents (si le patient est mineur), ainsi que des sujets contrôles selon le protocole NIH 12-N-0095. Les investigations ont été menées en accord avec la Déclaration d'Helsinki (2000). Les échantillons de contrôles cardiaques ont été obtenus de la Banque de Tissus pour la Recherche (Paris).

## 2. Etudes morphologiques

Sept biopsies musculaires de quadriceps ont été étudiées: deux biopsies pour la patiente P1 (14 et 23 ans), une pour les patients P2 (12 ans), P3 (16 ans), P5 (15 mois), V3 et VI3. Une biopsie cardiaque de ventricule gauche provenant de l'explant lors de la transplantation cardiaque de la patiente P5 (4 ans) a également été analysée.

### 2.1. Préparation des biopsies

Pour les colorations histologiques et les marquages immuno-histochimiques, chaque échantillon a été monté sur bouchon de liège et congelé dans de l'isopentane rapidement après le prélèvement. Les colorations et les marquages immuno-histochimiques ont été réalisés sur des coupes transversales et/ou longitudinales des biopsies coupées au cryostat (8  $\mu\text{m}$  d'épaisseur).

Quatre fragments de muscles squelettiques des patients P1, P5, V3 et VI3 et de cœur de la patiente P5 ont été fixés et préparés pour la microscopie électronique (Centre d'Histopathologie de Risler, Hôpital Pitié-Salpêtrière ou au Children's Hospital of Philadelphia (P5)).

### 2.2. Colorations histologiques standards et immuno-marquages

#### 2.2.1. Colorations histologiques

A l'exception de la coloration NADH pour la patiente 1 (Fig.10, f) que j'ai réalisée, les colorations histologiques ont été systématiquement réalisées dans le cadre du diagnostic au Centre d'Histopathologie de Risler (Hôpital Pitié-Salpêtrière), au Children's Hospital of Philadelphia ou au CHU de Rangueil. Les colorations suivantes réalisées selon des protocoles standards (Ceuterick and Martin 1996) ont été utilisées: Hémateine-Eosine (HE), B-nicotinamide adénine dinucléotide-tétrazolium réductase (NADH-TR), cytochrome oxydase (COX).

##### 2.2.1.1. *La coloration hémateine-éosine*

La coloration hémateine-éosine (HE) est une coloration bichromique alliant une laque nucléaire (l'hématoxyline de Harris de nature basique) et un colorant cytoplasmique (l'éosine de

nature acide). Cette coloration permet de visualiser la structure générale des fibres musculaires, l'aspect et la position des noyaux, la présence de fibrose et de cellules inflammatoires. En HE, le cytoplasme des fibres musculaires apparaît en rose et les noyaux en violet. Des zones éosinophiles peuvent être visualisées en rouge et correspondent à des accumulations de matériel myofibrillaire ou des corps cytoplasmiques. Enfin les fibres basophiles ont un cytoplasme coloré en bleu.

#### *2.2.1.2. Les colorations oxydatives*

Les colorations NADH-TR et COX permettent de révéler l'activité oxydative des fibres musculaires. Selon les types de fibres, le contenu mitochondrial varie. Ainsi, ces deux réactions permettent de distinguer les fibres de type I (ou oxydatives) où le marquage est de forte intensité, des fibres de type 2A à activité et marquage intermédiaire, et 2B (non-oxydatives) très faiblement marquées. Dans le muscle normal, ces colorations se présentent sous la forme d'un réseau cytoplasmique homogène.

La coloration NADH-TR emploie un sel de tétrazolium (incolore et soluble) comme accepteur d'électrons. Une fois réduit, il devient un composé coloré et insoluble situé au site actif de l'enzyme. Le sel de tétrazolium couramment utilisé est le nitrobleu tétrazolium (NBT) [chlorure de 2,2'-di-p-nitrophényl-5,5'-diphényl-3,3'-(3,3'-diméthoxy-4,4'-biphénylène) di-tétrazolium] qui donne un produit final de couleur bleue.

La réaction NADH-TR détecte une réductase présente dans les mitochondries et le réticulum sarcoplasmique. Cette coloration est également utile pour révéler les agrégats tubulaires accumulés dans le réticulum sarcoplasmique.

Enfin, la coloration COX permet de révéler en brun la membrane interne des mitochondries. La méthode utilisée pour la détection de l'activité COX utilise la diaminobenzidine comme donneur d'électron produisant un produit de coloration marron. Ce marquage est plus intense dans les fibres de type I que dans les fibres de type II, reflétant directement les quantités variables de mitochondries dans ces types de fibres.

#### *2.2.2. Immunomarquages de la titine et de l' $\alpha$ -actinine*

Pour les patients P1 et P5, l'expression des différentes portions de la titine a été étudiée par des expériences d'immunofluorescences sur des biopsies musculaires squelettiques (P1, P5) et cardiaques (P5) avec les anticorps suivants: anti-titine monoclonal de souris T51 (m9) (1/100) (donné

par M.Gautel) (Obermann, Gautel et al. 1997), anti-titine Zt12 polyclonal de lapin (1/200) (donné par Pr T. Arimura) (Hackman, Vihola et al. 2002).

Pour la patiente P5, l'immunomarquage dirigé contre l' $\alpha$ -actinine a été également réalisé sur la biopsie cardiaque, par l'équipe du Dr Mathias Gautel.

Pour chacun de ces anticorps, le protocole d'immunofluorescence suivant a été utilisé. Après séchage des coupes pendant 15 minutes, celles-ci ont été fixées au paraformaldéhyde (4%) pendant 10 minutes à température ambiante. Les coupes ont été rincées dans une solution PBS-1X / Glycine 0,1M (3 fois 5 minutes). Elles ont ensuite été incubées dans une solution de blocage (PBS 1X, sérum albumine bovine 3%) à température ambiante pendant 40 minutes, afin de saturer les sites non-spécifiques. Toutes les incubations ont été réalisées dans une chambre humide et à température ambiante. Les lames ont été incubées ensuite pendant deux heures à température ambiante avec les anticorps primaires (dilués dans la solution de blocage). Après 3 rinçages de 10 minutes au PBS-1X, les anticorps secondaires dilués dans la solution de blocage (1/400), ont été mis en contact des coupes pendant 45 minutes (lapin polyclonal anti-souris Ig/FITC Dako ou Alexa 488 chèvre anti-lapin, Molecular probes). Les lames ont été rincées trois fois 5 minutes au PBS-1X puis montées dans du milieu de montage Vectashield- DAPI (Vector) afin de colorer les noyaux en bleu. Les signaux fluorescents ont été observés au microscope confocal à balayage laser (Leica) ou au microscope confocal inversé (Zeiss LSM 510 Meta, Mathias Gautel) et photographiés avec une caméra Photometrics Cool Snap fx (Roper Scientific) contrôlée par le logiciel Metavue (Roper Scientific). Les images obtenues ont été traitées avec le logiciel Photoshop.CS5 (Adobe, San Jose CA).

### 3. Etudes génétiques

L'ADN génomique des patients a été extrait à partir de prélèvements sanguins ou de lignées lymphoblastoïdes par la banque d'ADN et de tissus de Généthon. Dans la famille de la patiente 5, l'ADN et l'ARN foetal ont été obtenus à partir d'une biopsie musculaire prélevée suite à l'interruption de grossesse.

#### 3.1. Génotypage

##### 3.1.1. Approche mixte de cartographie par homozygotie - gène candidat

Pour toutes les familles du groupe sans atteinte cardiaque, une seule génération est atteinte suggérant une transmission autosomique récessive de la maladie. De plus, une partie conséquente (86%) des familles étudiées sont consanguines, c'est-à-dire qu'il existe un ancêtre commun aux parents. Pour cette raison, nous avons utilisé le principe de cartographie par homozygotie (Lander and Botstein 1987). En parallèle de cette étude, l'approche dite par gène candidat a été utilisée, c'est-à-dire que des gènes particuliers ont été étudiés par séquençage direct. Ces gènes ont été choisis selon leurs implications dans des pathologies pour lesquelles certaines données phénotypiques étaient concordantes avec le phénotype des patients étudiés.

##### 3.1.2. Criblage du génome

Initialement, un criblage du génome utilisant 330 marqueurs microsatellites fluorescents (Applied Biosystems), répartis sur tous les chromosomes et distants en moyenne de 10 à 25 cM, a été effectué pour tous les membres (sains et atteints) de 15 familles du groupe sans atteinte cardiaque (les plus informatives), à Généthon par le Dr Ana Ferreiro. Les marqueurs microsatellites ont été amplifiés par la réaction de polymérisation en chaîne de l'ADN (PCR). La PCR permet d'amplifier de manière exponentielle une séquence d'ADN particulière, ici des microsatellites, à l'aide d'oligonucléotides situés de part et d'autre de cette séquence, grâce à une Taq polymérase thermostable. Pour chaque couple d'amorces, une des amorces est fluorescente couplée en 5' au fluorochrome 6-carboxyfluoresceine- (FAM), 6-carboxy-2, 4, 7, 7- hexachlorofluoresceine- (HEX) ou 4, 7, 2', 7'- tetrachloro-6-carboxyfluoresceine- (TET) phosphoramidites. Les couples d'amorces sont choisis dans des régions uniques de part et d'autre des microsatellites afin d'être spécifique de chaque marqueur. Les migrations sur séquenceur ABIPrism377 permettent de séparer les produits selon leur taille après

détection par un laser des fragments fluorescents. Les résultats bruts (tailles des fragments amplifiés pour chaque marqueur microsatellite et pour chaque individu) nous ont ensuite été retournés par le Généthon afin de procéder aux analyses de liaison.

Dans un second temps un génotypage de 15 familles du groupe de patients sans atteinte cardiaque a été réalisé par le centre national de séquençage du Généthon en utilisant le panel de marqueurs Affymetrix de la puce 250k Nsp, incluant environ 262 000 marqueurs répartis de façon homogène sur le génome. Enfin suite à des problèmes techniques lors du génotypage de la famille C, un second génotypage de cette famille a été réalisé avec le human panel V (Illumina) intégrant 6 056 SNPs.

Pour chaque famille les échantillons d'ADN de tous les enfants atteints, des parents ainsi que des frères et sœurs non atteints, si disponibles, ont été systématiquement génotypés.

### 3.1.3. Analyses de liaison

Pour déterminer si deux loci sont liés, il faut étudier la fréquence de recombinaison entre ces deux loci, appelée fraction de recombinaison ( $\Theta$ ). En analyse de liaison paramétrique, le paramètre d'intérêt est donc  $\Theta_{AB}$ , qui est la fraction de recombinaison entre deux loci A et B. Ici, les 2 loci A et B correspondent au gène de la maladie et au marqueur génétique.

La valeur de  $\Theta$  se situe dans l'intervalle  $0 < \Theta < 1/2$ . Ainsi,  $\Theta = 0$  représente une liaison parfaite (les 2 loci sont co-hérités ensemble 100% du temps),  $\Theta = 1/2$  représente l'indépendance des 2 loci en termes de ségrégation (seconde loi de Mendel) et  $\Theta < 1/2$  représente l'hypothèse de liaison. **Le test du LOD-score consiste à comparer l'hypothèse H0 d'indépendance génétique ( $\Theta=1/2$ ) à l'hypothèse alternative de liaison génétique H1 (liaison à  $\Theta < 1/2$ ).** Ce test permet donc de déterminer s'il existe une liaison significative entre une région chromosomique et la maladie. Les étapes de l'analyse de liaison sont les suivantes:

1/ Etablir un modèle de transmission du trait (récessif, dominant ou autre).

Ce modèle s'établit en spécifiant 3 paramètres  $f_0$ ,  $f_1$  et  $f_2$  appelés pénétrances et correspondant respectivement aux probabilités d'être atteint si les deux allèles sont sauvages ( $f_0$ ), l'un des deux allèles est muté ( $f_1$ ) ou les deux allèles sont mutés ( $f_2$ ). Les deux cas particuliers étant le modèle de transmission récessif ( $f_0 = f_1 = 0$  et  $f_2 = 1$ ) ou dominant ( $f_0 = 0$  et  $f_1 = f_2 = 1$ ).

2/ Calculer la fonction de vraisemblance des pedigrees.

Cette fonction de vraisemblance va dépendre de  $\Theta$ , de la composition familiale, de  $\Phi = (f_0; f_1; f_2)$  et des fréquences des allèles du gène de la maladie ainsi que du marqueur (polymorphisme). Cette fonction de vraisemblance est notée  $L(\Theta; \Phi; \dots)$ .

3/ Tester l'hypothèse  $H_0 : \Theta = 1/2$  (indépendance des loci) versus  $H_1 : \Theta < 1/2$  (liaison) à l'aide d'un test de rapport des vraisemblances.

Pour ce faire, on calcule  $L_0 = L(\Theta = 1/2; \Phi; \dots)$ , la vraisemblance sous  $H_0$ . On calcule également  $L_1 = L(\Theta < 1/2; \Phi; \dots)$ , la vraisemblance sous  $H_1$ . La statistique de test est appelée le LOD-score et est défini de la façon suivante:  $Z(\Theta) = \log_{10}(L_1 / L_0)$ . En pratique, si  $Z(\Theta) > 3$ , on rejette l'hypothèse nulle donc on conclut que le gène de la maladie est lié au marqueur (la probabilité d'observer l'échantillon est 1000 fois plus grande sous  $H_1$  que sous  $H_0$ ). Si  $Z(\Theta) < -2$ , on exclu la liaison de la maladie au locus. Dans les cas  $-2 < Z(\Theta) < 3$ , on ne peut pas conclure.

Les analyses de liaison et calcul de LOD score ont été effectuées avec le logiciel Merlin (Abecasis, Cherny et al. 2002) en utilisant les critères suivants:

- transmission autosomique récessive de la maladie
- fréquence de recombinaison identique pour les deux sexes
- fréquence de la maladie estimée à 0.0001
- fréquence isoallélique de 0.1
- une seule classe de liability
- pénétrance de 0.98

### 3.2. Séquençage de type Sanger

Les six derniers exons de la titine codant les domaines de la bande M (358 à 363 ou Mex1 à Mex6 respectivement pour M-line encoding exons 1 to 6) ont été amplifiés chez les patients du groupe avec atteinte cardiaque avec des amorces et selon une méthode décrites précédemment (Marchand, Hajdari et al. 2003). Cette première étape permet d'amplifier les fragments d'intérêts avec des oligonucléotides spécifiques (Tableau 1). L'ADN génomique (30 ng) a été placé dans 50  $\mu$ l final de mélange PCR contenant 5  $\mu$ l de tampon 10X avec MgCl<sub>2</sub> (Invitrogen), 2 mM de chaque amorce, 0.2

mM de dNTPs (Invitrogen) et 1.5 unité de Taq Polymérase (Invitrogen). Après une dénaturation initiale à 95°C pendant 5 minutes, 10 étapes d'amplifications ont été réalisées à des températures d'hybridation décroissantes (de 65 à 55°C). Ces étapes sont constituées de 3 cycles (dénaturation (95°C), hybridation (72°C), élongation (72°C) de 5 secondes, 10 secondes et 1 minute respectivement). Une élongation finale a été effectuée à 72°C pendant 7 minutes. Les produits PCR ont été contrôlés par fluorescence après migration sur gel d'agarose 2%. Les fragments amplifiés ont ensuite été purifiés sur colonne NucleoSpin, ExtractII (Macherey-Nagel).

<b>Bande M</b>	Noms des amorces	Séquences amorces sens	Séquences amorces antisens	Taille du produit PCR (pb)
Exon 358	MEX 1,2	TTATTAAGTGGGGGTGGA	TGCTTCAAATGATTCATG	979
	MEX 1,3	TGGATTGTCCATCGTTGTG	CCTCCTTCTCACCACACTGC	963
	MEX 1,4	CGAATTCGATCTCAGAAGGG	GGACAGTGGCTGACCATCTT	867
	MEX 1,5	TACTGGCAAATGCAGAATGC	CTCTTCTCAAGACGCAGCC	956
	MEX 1,6	GAGATAGTGAGACCAGCCC	TGAAAGGCTGCTGACTCAAA	922
	MEX 1,7	ACTCCAGAGAGAACTCGCCC	ACGCTGTAATTGCCCTCATC	971
	MEX 1,8	TAAGTACTTCTGCCCGCCAC	TGGCCTGTAGAATGCAAATG	997
Exon 359	MEX 2	TCTTCTAAATTCAGCTTCCCA	TGTGTGTTTCTGCTTTGGTG	390
Exon 360	MEX 3	AAAAGGTGGGGGTCTCTTT	TCTTCAGATGTGGAAGACAT	811
Exon 361	MEX 4	ATCCCTGAAATCGAATGGT	ACATCAGTTGGCTGTCCCTC	422
Exon 362-363	MEX 5/6	GGGTTATGCTGCTGTGTGTG	AAAGATTAGTCCGTGTGAAA	922

**Tableau 1 : Séquences des amorces utilisées lors du séquençage de la portion du gène de la titine codant pour les domaines de la bande M (exons 358 à 363)**

Dans le cadre de l'étude de ce groupe de patient j'ai également réalisé le séquençage des 28 exons en positions N-terminal (3 patients) de la titine ainsi que de 102 exons codant le domaine PEVK de titine. (3 patients) (**Tableau en annexe 2**).

Dans le cadre de l'étude du groupe de patients sans atteinte cardiaque nous avons séqué environ une centaine de gènes candidats positionnels

Les réactions de séquences ont été réalisées à partir de 1 à 5 µl de produit de PCR purifié, à la plateforme génomique de l'hôpital Cochin.

### 3.3. Séquençage d'exome

#### 3.3.1. Préparation des échantillons

La capture et l'enrichissement de l'exome de 8 patients (incluant P2 et P5) ont été réalisés en utilisant le kit SureSelect Human All Exon 50Mb (Agilent technologies) selon le protocole

recommandé par le fournisseur (Fig. 13). Les produits d'amplification ont été séquencés sur un séquenceur de nouvelle génération « Sequencing by Oligo Ligation and Detection » SOLiD™ 4 System (Applied Biosystems).

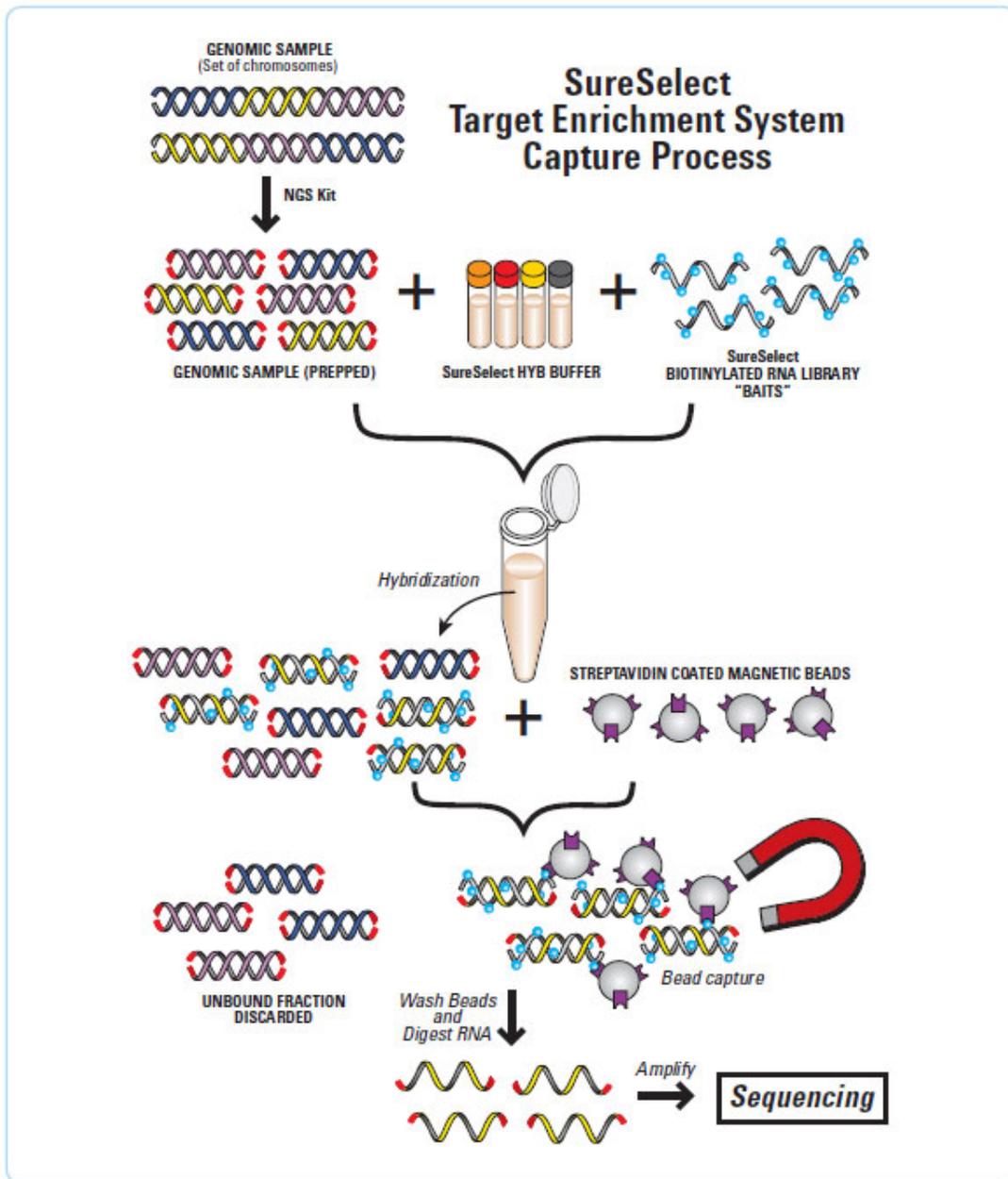


Figure 13 : Protocole de préparation des ADN génomiques des patients.

Le fractionnement des ADNg des patients se fait en utilisant le processus Covaris Adaptive Focus Acoustic (AFA) qui est réalisé dans des conditions isothermes, assurant le maintien de l'intégrité de l'échantillon d'acide nucléique et en garantissant une conservation élevée de la structure de l'ADN double brin. La capture des régions ciblées est fonction du design de la librairie d'oligo ARN utilisée. L'hybridation entre la librairie d'ARN et les fragments d'ADNg se fait sur la nuit. L'amplification est réalisée par ePCR, ou PCR en émulsion en 5 étapes : La fixation des matrices se fait avec des billes.

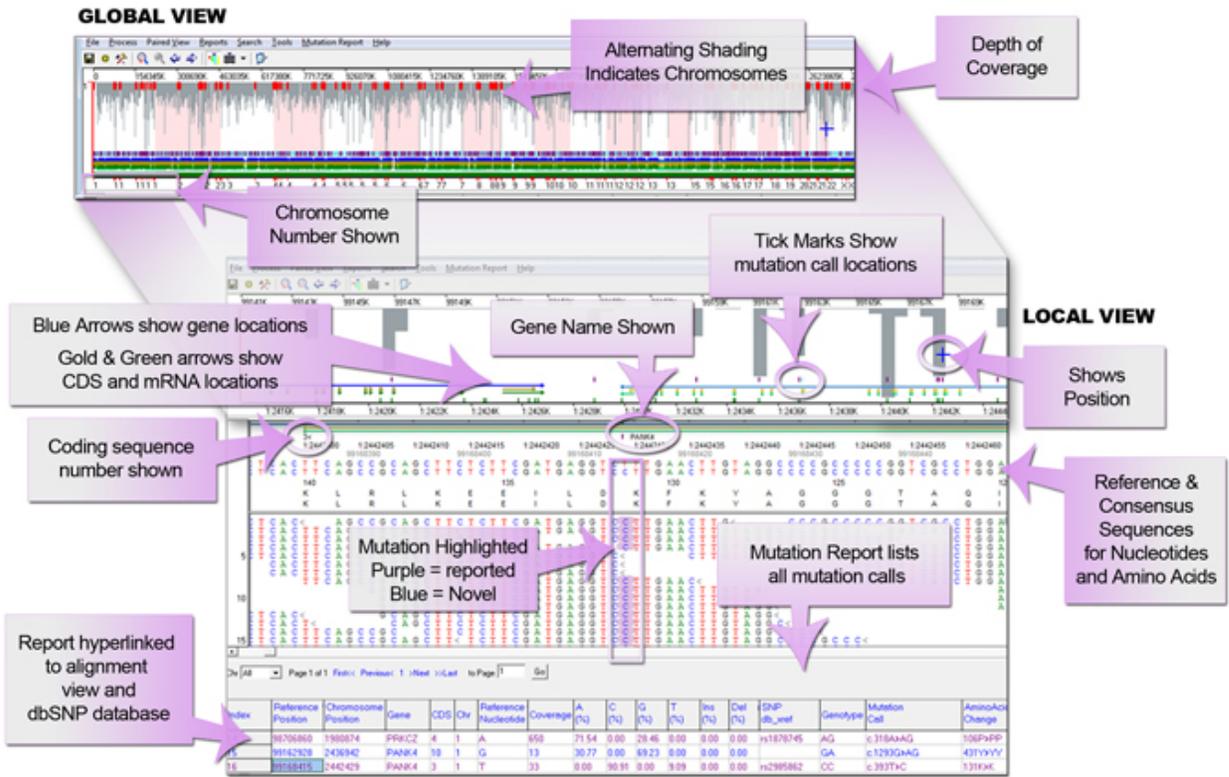
1 - Obtention d'une émulsion : chaque goutte contenant une seule bille et un seul fragment d'ADN. 2 - Emulsion après amplification : chaque bille est liée à plus de 30.000 copies du fragment d'ADN de départ. 3 - Préparation de 125 millions à 1 milliards de billes/run suivant les projets. 4 - Cassage de l'émulsion et lavage des billes ; séparation des billes amplifiées des billes non amplifiées. 5 - Terminal transférase.

Les amplicons obtenus sont alors séquencés sur le séquenceur SOLiD4.

### 3.3.2. Analyses bioinformatiques

Les données brutes de séquences en format « color-space » (csfasta) ont été analysées avec le logiciel NextGENe (Softgenetics, Inc.) dont l'algorithme d'alignement utilise un processus en 3 étapes pour aligner les amplicons, obtenus suite à l'étape d'amplification par PCR des fragments capturés d'ADN génomique des patients, au génome de référence -hg19, GRCh37- de la base de données de University of California Santa Cruz (UCSC) (**Fig. 14**). La première étape fait correspondre par itération aléatoire des fragments de la séquences de référence plus petits que la longueur des amplicons avec ces derniers, puis NextGENe génère *in silico* l'extension des différents alignements obtenues afin, en dernière étape, de trouver la meilleure position correspondant aux alignements les plus précis des amplicons contenant les mutations ponctuelles (SNP) et les indels. L'assurance que les amplicons ont bien été cartographiés est permise par la profondeur du séquençage.

## NextGENe Viewer/Browser



**Figure 14 :** Ensemble des informations prodiguées par le visionneur du logiciel d'analyse NextGENe. Les paramètres du logiciel permettent d'intégrer au génome de référence les positions pour lesquelles les bases de données public ou des bases de données privées signalent des polymorphismes. Ce qui permet de distinguer dans le tableau de variants final les polymorphismes connus et non répertoriés.

Finalement, les variants privés non répertoriés ont été identifiés dans les échantillons des patients par comparaison avec les bases de données dbSNP, 1000 Génomes, et d'autres données d'exomes obtenus de projets indépendants en cours au Centre de Recherche du CHU Sainte Justine (Montréal). Puis les analyses des conséquences fonctionnelles des variants ont été réalisées en utilisant des logiciels libres de prédiction d'effet pathogène de mutations (SIFT, PolyPhen, ESE finder...) (Fig. 15).

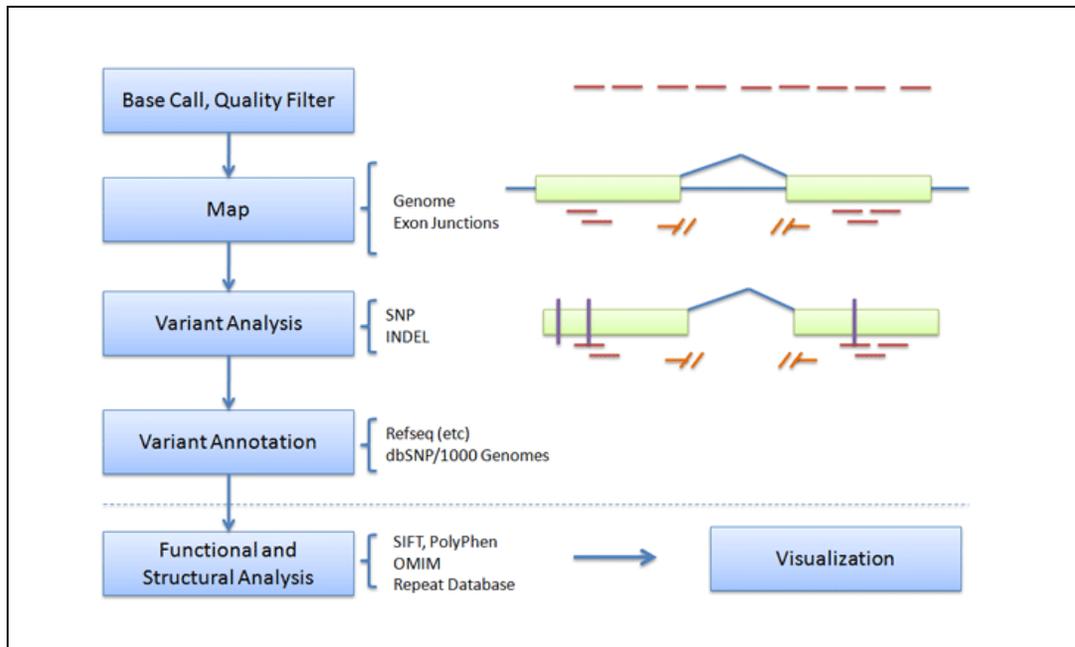


Figure 15 : Stratégie d'analyse d'exome d'après [www.EdgeBio.com](http://www.EdgeBio.com)

## **4. Cultures de cellules primaires**

Pour étudier les conséquences des mutations identifiées, nous avons utilisé des cultures de cellules primaires de fibroblastes des patients et de la mère hétérozygote asymptomatique (IV4) de la famille C et de sujets contrôles ainsi que de myoblastes de la patiente VI3. Après le prélèvement, les échantillons musculaires ont été envoyés au Généthon (Banque d'ADN et de cellules) pour procéder à l'amplification des cellules. Une fois amplifiées, les cellules ont été congelées dans au moins 3 ampoules. Les ampoules nous ont ensuite été envoyées pour procéder aux différentes expériences.

### **4.1. Décongélation des cellules**

La décongélation d'une ampoule de cellules stock de  $10^6$  cellules se fait rapidement dans un bain-marie à 37°C, les cellules sont remises en suspension, dans une flasque de culture de 75 cm<sup>2</sup>, dans un volume final de 15 ml de milieu de croissance F10-Ham (Gibco) contenant 20% de sérum de veau fœtal (Gibco). L'ensemble est maintenu dans une étuve à 37°C, sous atmosphère humide et 5% de CO<sub>2</sub>. Le milieu des cellules est changé tous les deux jours.

### **4.2. Amplification des cellules**

Une fois les cellules arrivées à 70% de confluence, (au-delà elles commencent à se différencier) le milieu de culture est aspiré. La boîte est rincée avec 10 ml de PBS-1X sans calcium (Gibco). Après aspiration, 2 ml de trypsine-EDTA (Gibco) sont ajoutés et laissés en contact pendant 5 min à 37°C jusqu'à ce que les cellules se détachent. L'action de la trypsine est ensuite inhibée par l'ajout de 8 ml de milieu de croissance. Les cellules sont homogénéisées et individualisées par aspirations et refoulements successifs. Dans une nouvelle flasque de 175 cm<sup>2</sup>, les 10 ml de cellules trypsinisées sont alors ajoutés à 10 ml du même milieu puis l'ensemble est incubé dans une étuve à 37°C, sous atmosphère humide et 5% de CO<sub>2</sub>. Ces étapes sont reproduites jusqu'à ce que la quantité de cellule soit suffisante. Une fois la quantité de cellules nécessaires obtenue, une quantité fixe de 100 000 cellules est directementensemencée dans des puits de plaques 6 puits (études pour ARN et protéines).

### **4.3. Induction de la différenciation**

Lorsque les cellules atteignent 70% de confluence, la différenciation des myoblastes en myotubes peut alors être induite par changement de milieu ; le milieu dit « de fusion » (DMEM, Gibco) est composé de 2% de sérum de cheval (HyClone),  $10^{-6}$  M d'insuline (Sigma) et  $2.5 \cdot 10^{-6}$  M de dexaméthasone (Sigma). Les cellules commencent à fusionner au bout de deux jours (présence de myotubes avec deux noyaux et alignement des cellules les unes par rapport aux autres). Selon le type d'étude, les cellules sont laissées en différenciation plusieurs jours avec changement de milieu tous les deux jours. Pour les études des ARNs et des protéines les cellules sont récoltées 5 jours après induction de la différenciation.

### **4.4. Congélation des cellules**

Pour la congélation, les cellules sont trypsinisées et reprises dans un volume final de 10 ml contenant le milieu de croissance selon la méthode décrite précédemment. Les cellules sont ensuite dénombrées à l'aide d'une cellule de Mallassez (Preciss France) et la suspension cellulaire est centrifugée à 1500 rpm pendant 10 minutes à température ambiante. Après aspiration du surnageant, le culot de cellules est repris dans du sérum de veau fœtal (Gibco) et 10% de DMSO (Sigma), à une concentration finale de 1 million de cellules par ml. Les cellules sont aliquotées par fractions de 1 ml dans des tubes de congélation Eppendorf, puis placées dans une boîte de congélation contenant le volume adéquat d'isopropanol. La boîte est ensuite placée au moins 48 heures à  $-80^{\circ}\text{C}$ , puis les ampoules de cellules sont conservées dans l'azote liquide.

## 5. Analyse fonctionnelle des conséquences des mutations

### 5.1. A l'échelle des ARN (famille C)

Les ARN totaux ont été extraits à partir de cultures primaires de fibroblastes des patients (V3, V4, VI3), de la mère hétérozygote (IV4) et d'un groupe de contrôles ; ainsi que de myotubes de la patiente VI3 et de contrôles.

#### 5.1.1. Cultures cellulaires

Les myotubes à 10 jours de différenciation ont été lavés dans du PBS-1X sans calcium et trypsinés (200 µl trypsine/puits, Invitrogen) pendant 5 minutes à 37°C. Les cellules décollées ont été remises en suspension dans 2 ml de milieu de fusion et centrifugées à 1500 rpm pendant 10 minutes. Le surnageant a été ôté sans toucher le culot de cellules. Il est possible de congeler les culots secs de cellules à -80°C à ce stade ou de procéder directement à l'extraction des ARNs.

#### 5.1.2. Extraction des ARN totaux

L'extraction d'ARN a été effectuée selon le protocole fourni avec le kit d'extraction d'ARN RNeasy mini kit (Qiagen) à partir d'un minimum de 20 coupes de 8 µm d'épaisseur pour les muscles congelés ou d'un puits d'une plaque de culture cellulaire de 6 puits, où les myotubes sont confluent et différenciés. Pour chaque type de cellules, les extractions ont été réalisées en parallèle à partir de deux puits de cultures différents.

Les culots de cellules ont été remis en suspension par pipetage dans 350 µl de tampon de lyse RLT (fourni dans le kit) contenant 1% de b-mercaptoéthanol. Le lysat a été homogénéisé par un minimum de 10 aspirations et refoulements successifs au travers d'une seringue avec une aiguille de 0,9 mm de diamètre. L'éthanol 70% (350 µl) a été ajouté au lysat et mélangé longuement par pipetage pour provoquer une précipitation des acides nucléiques. Les 700 µl obtenus ont été transférés sur une colonne d'affinité (fournie avec le kit) qui lie les molécules d'ARN lors de la centrifugation pendant 15 secondes à 10 000 rpm. La colonne a été lavée par ajout de 700 µl de tampon RW1 fourni par le kit et centrifugée pendant 15 secondes à 10 000 rpm. Une étape de digestion à la DNase 1 a été réalisée afin de dégrader toute trace d'ADN génomique (incubation à température ambiante, 15 minutes). La colonne a ensuite été lavée par ajout de 700 µl de tampon RW1 suivi d'une centrifugation pendant 15 secondes à 10 000 rpm. Sur chaque colonne, 500 µl de tampon RPE (dilué au 1/5 dans de l'éthanol) ont été ajoutés puis centrifugés 15 secondes à 10 000 rpm. De nouveau, 500 µl de tampon RPE (dilué

au 1/5 dans de l'éthanol) ont été ajoutés sur la colonne et centrifugés 2 minutes à 10 000 rpm. Ces étapes ont permis d'enlever les résidus d'éthanol pouvant interférer avec la suite du protocole. La colonne a ensuite été séchée par centrifugation, 1 minute à 10 000 rpm, puis l'ARN a été élué par ajout de 30 µl d'eau sans RNase et centrifugation 1 minute à 10 000 rpm.

Les ARNs totaux ont été quantifiés par mesure d'absorbance au spectrophotomètre Nanodrop ND-1000 (Nanodrop technologies) et stockés à -20°C.

### 5.1.3. Transcription inverse des ARNs en ADN complémentaire

L'ADN complémentaire (ADNc) a été obtenu par transcription inverse à partir de 1 µg d'ARNs totaux en utilisant le kit SuperScript III First-strand Synthesis SuperMix (Invitrogen). Selon la nature des ARNs, deux protocoles sont utilisés. L'ADNc correspondant aux ARNm est spécifiquement synthétisé à partir d'amorces polyT s'appariant avec la queue polyA des messagers. Les ADNc correspondant aux ARN pré-messagers sont amplifiés par transcription inverse en utilisant des hexamères de séquences aléatoires (random hexamers) pouvant s'apparier à tous les ARNs. Le kit fournit les deux types d'amorces, les expériences ont été réalisées en suivant les protocoles du fournisseur (Invitrogen).

L'ARN (1 µg) est mis en présence du mix contenant les dNTPs, les d'oligos (dT)<sub>20</sub>, du MgCl<sub>2</sub> et des random hexamers, puis est incubé à 25°C pendant 10 minutes, puis à 50°C pendant 30 minutes. La synthèse se termine par l'inactivation de l'enzyme à 85°C pendant 5 minutes, suivie de la dégradation des ARNs en ajoutant 2 unités de RNase H incubée à 37°C pendant 20 minutes. Les ADNc ont été ensuite stockés à -20°C.

### 5.1.4. Quantification relative par PCR quantitative en temps réel

Les ADNc ont été quantifiés par PCR quantitative en temps réel en utilisant le kit QPCR SYBR green 480 Master I (Light Cycler, Roche). La PCR quantitative est une technique destinée à quantifier la quantité d'un type d'ARN initialement présent dans un échantillon. Pour cela, un intercalant de l'ADN (SYBR Green) est utilisé. Cette molécule fluoresce lorsqu'elle est complexée avec de l'ADN double-brin. Ainsi au fur et à mesure des cycles de la PCR, une quantité croissante de fluorescence est détectée. Le cycle seuil, ou C<sub>p</sub>, est le cycle à partir duquel le produit est détectable et où le signal fluorescent est directement proportionnel à la quantité de matrice initiale. La détermination des C<sub>p</sub> pour les ARN contrôles et mutés permet de réaliser les quantifications.

Le risque d'erreurs de cette expérience est la détection des dimères d'oligonucléotides et d'un produit de PCR non spécifique. Pour vérifier l'absence de ces produits, les courbes de fusion

(variation de fluorescence en fonction de la température  $dF/dT = f(T)$ ) ont été obtenues. Chaque produit y est représenté sous forme d'un pic. Des courbes de standardisation (avec des triplicats de dilutions 1/10, 1/50 et 1/250 de l'ADNc) chez un échantillon contrôle ont été réalisées préalablement aux expériences de quantification afin de pouvoir standardiser la quantité de matrice initiale aux Cp.

Les ADNc ont été ajoutés à un mélange contenant les primers et le SYBR Green en suivant le protocole fournit. Les oligonucléotides ont été choisis à l'intérieur de la séquence de l'ADN complémentaire du gène étudié avec pour critères de sélection une taille optimale des amorces de 20 bases, une température d'hybridation de 60°C et un pourcentage optimal de bases GC de 50%.

Une dénaturation initiale à 95°C pendant 15 minutes est nécessaire à l'activation de l'enzyme du kit. Elle a été effectuée préalablement à toute amplification et a été directement suivie par les étapes d'amplification constituées de 50 cycles de dénaturation (95°C, 15 secondes), hybridation (58°C, 20 secondes), élongation (72°C, 20 secondes). Enfin, une étape de fusion (augmentation de 0.1°C par seconde jusqu'à 95°C) a été réalisée.

Lors de la quantification des ARNm du gène, les niveaux d'expression de ce gène ont été rapportés à ceux du gène codant la sous unité ribosomique 18S.

## **5.2. A l'échelle protéique : Analyse par Western-Blot**

### **5.2.1. Analyse des mutations TTN (P5)**

Pour l'analyse de l'expression protéique de TTN chez la patiente 5, des sections congelées de muscle cardiaque de la patiente ainsi que d'un individu contrôle ont été lysés dans un tampon Laemmli contenant 6 M d'urée, chauffés à 90°C et conservés à -80°C. Les lysats ont été fractionnés par électrophorèse sur un gel 10-15% de type SDS-PAGE et transférés sur une membrane PVDF pour une détection avec des anticorps polyclonal anti-lapin m10-1 dilués au 1/500, monoclonal anti-TTN T12 au 1/500 ou Z1Z2 au 1/1000 (x112-x113) pour détecter la partie N-terminal de la protéine située en amont de la mutation nonsense. La quantité de protéine chargée a été normalisée par rapport aux marquages de l'actine ou de GAPDH. Les analyses western-blots ont été menés par l'équipe du Dr Bjarne Udd en suivant les procédures standard.

### **5.2.2. Analyse de l'expression de la protéine identifiée dans la famille C**

Les protéines ont été extraites à partir de culture de cellules de C2C12. Les échantillons ont été lysés dans un tampon de lyse contenant 75 mM TRIS HCl pH 6.8, 15% SDS, 20% glycérol, 5%

bleu de bromophénol saturé, 5% de b-mercaptoéthanol et 1/2 pastille d'antiprotéases (Roche). Les différentes fractions ont été dénaturées pendant 5 minutes à 95°C. Les protéines (40µg) ont été fractionnées par électrophorèse sur un gel dénaturant de polyacrylamide 4-15% de type SDS-PAGE (BIORAD precast). La migration a été effectuée pendant 1 heure à 200 V constants dans un tampon de migration Tris-glycine 1X, pH 8 (0.25 M Tris; 1.92 M glycine; 0.1% SDS).

Après fractionnement sur SDS-PAGE, les protéines ont été transférées sur une membrane nitrocellulose 0.45µm (Invitrogen) dans un tampon de transfert pH 8 (0.25 M Tris; 1.92 M glycine; 20% isopropanol). Le gel et la membrane ont été disposés entre 2 fois 2 feuilles de papier Whatman et des éponges dans une cuve de transfert humide (XCell Blot Module, Invitrogen). Le transfert s'est effectué dans ce tampon pendant 45 à 50V. La membrane, sur laquelle les protéines ont été transférées, a été saturée par incubation pendant 1 heure dans le tampon de blocage (PBS; 3% lait en poudre; 0.2 % Tween-20, 0.05% SDS), puis hybridée sous agitation pendant la nuit dans le même tampon en présence d'un anticorps spécifique de la protéine d'intérêt dilué au 1/1000. La membrane a été ensuite lavée pendant 3 fois 5 minutes dans le tampon de blocage et l'hybridation de l'anticorps secondaire couplé à un fluorochrome (dilué au 1/7500) a été effectuée pendant 45 minutes dans le tampon de blocage à l'obscurité. La membrane a alors été lavée pendant 3 fois 5 minutes dans le tampon de blocage, puis 1 fois 5 minutes dans du PBS. La réaction antigène/anticorps a été visualisée par fluorescence sur un scanner infrarouge LI-COR Odyssey (Bioscience).

### 5.2.3. Analyse du dichroïsme circulaire

L'analyse des conséquences de la mutation du domaine kinase de la TTN sur la structure secondaire de la protéine a été réalisée par l'équipe du Dr Mathias Gautel. Les spectres d'UV lointains ont été acquis sur un spectrophotomètre de dichroïsme circulaire Chirascan de Applied Photophysics (Leatherhead, UK) dans des longueurs d'ondes comprises entre 200 et 260 nm. Le dichroïsme circulaire définit la capacité d'un matériau à absorber différemment la lumière selon que sa polarisation est circulaire droite ou circulaire gauche, l'une des deux composantes étant absorbée plus rapidement que l'autre. De cette manière, les structures en hélice- $\alpha$  et feuillet- $\beta$  des protéines et en double hélice des acides nucléiques présentent des dichroïsmes circulaires caractéristiques, et la mesure de leurs variations permet d'évaluer les effets des mutations sur la structure des protéines.

Selon la théorie ondulatoire de la lumière. Une onde lumineuse d'une certaine amplitude voyage dans une direction donnée avec une forme sinusoïdale. L'angle des pics de l'onde par rapport à

l'axe de sa propagation peut avoir n'importe quelle valeur, mais il est possible de sélectionner un angle particulier grâce à un filtre polarisant (polaroïd). Consistant en de longues molécules alignées dans une direction précise, un tel filtre ne laisse passer que les oscillations d'une onde électromagnétique orientées dans son axe de transmission. La résultante de deux rayons lumineux polarisés à  $90^\circ$  l'un de l'autre mais se propageant dans la même direction est une autre onde sinusoïdale située entre les deux premières (Fig. 16).

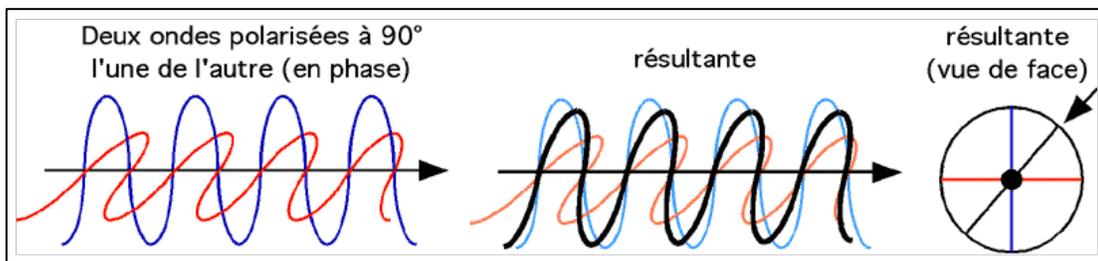


Figure 16 : Polarisation de deux ondes en phase

L'utilisation d'un filtre spécifique (un *modulateur photoélastique*) qui en plus de laisser passer les deux ondes de départ, les déphase d'une valeur de  $\pi/2$ , génère comme résultante une spirale. Si une des deux ondes est retardée par rapport à l'autre d'un quart de longueur d'onde ( $\lambda/4$ ) la résultante sera une hélice de pas droit. Si l'onde est avancée de  $\pi/2$  par rapport à l'autre, l'hélice obtenue est de pas gauche (Fig. 17).

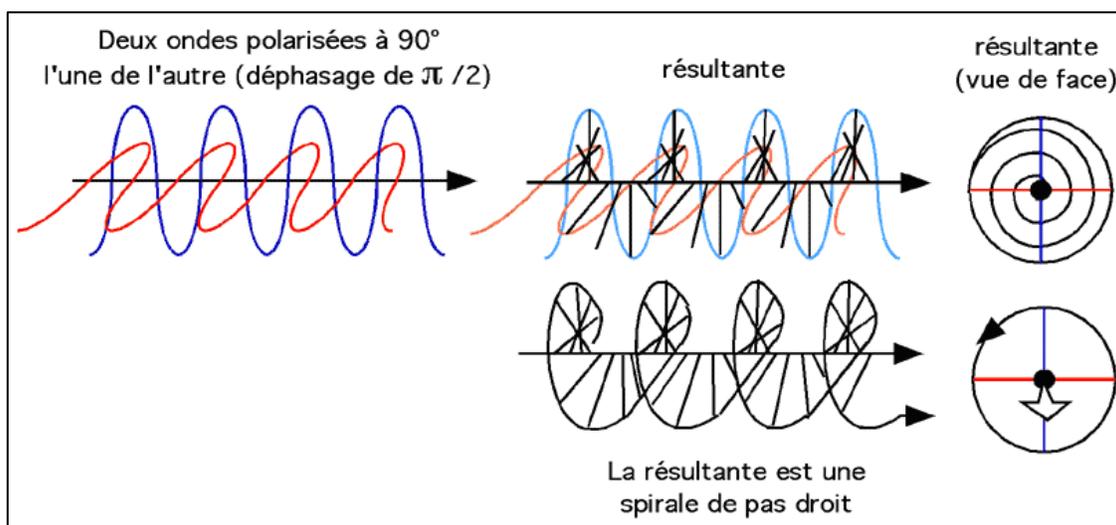
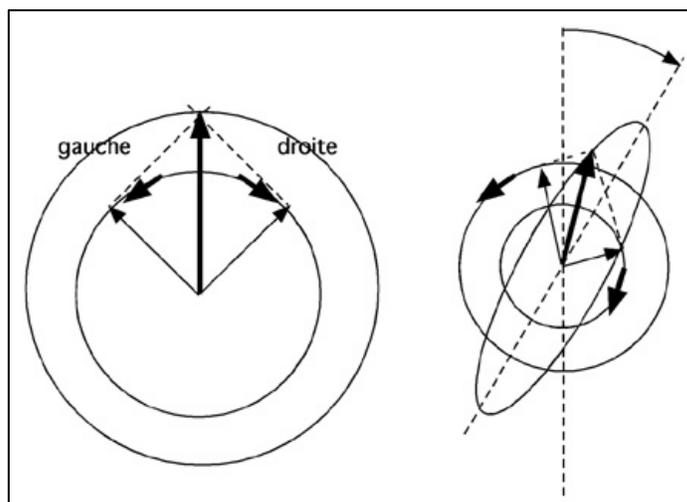


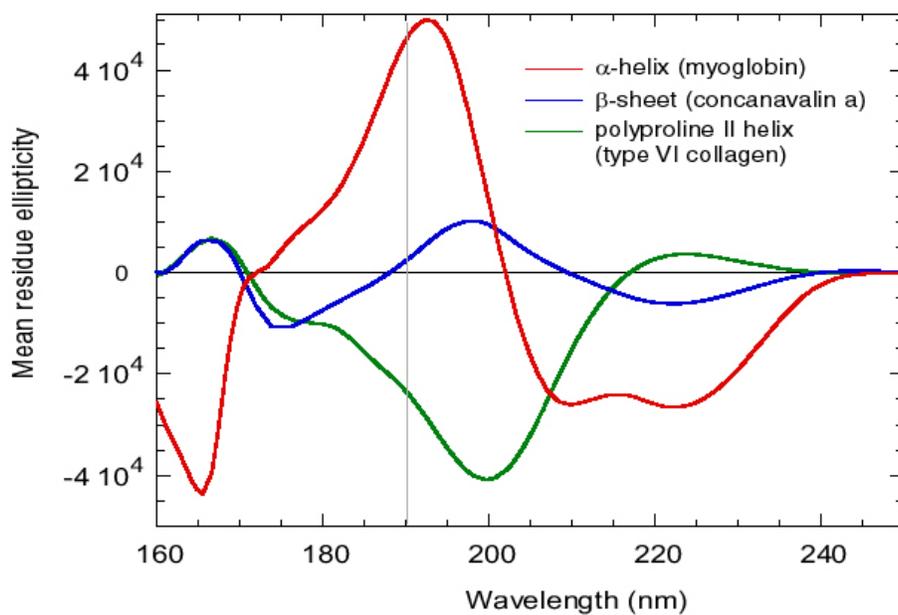
Figure 17 : Polarisation à droite de deux ondes déphasées

Les structures secondaires n'absorbent pas de façon égale la lumière polarisée circulairement vers la droite et la lumière polarisée circulairement vers la gauche. L'absorption préférentielle de l'une de ces deux polarisations résultera en une déviation de la résultante appelée dichroïsme circulaire (au lieu d'un cercle, le tracé de la résultante entre la spirale tournant à droite et la spirale tournant à gauche donnera une ellipse) (**Fig. 18**).



**Figure 18 : Représentation du dichroïsme circulaire.** Gauche absorption équivalente de la lumière polarisée droite et gauche.  
Droite : Absorption préférentielle de la lumière polarisée droite

Elle permet d'établir un spectre de forme caractéristique pour trois structures secondaires retrouvées dans les protéines: l'hélice, le feuillet et la forme aléatoire (ou random coil), qui chacune présentent des asymétries structurales (**Fig. 19**).



**Figure 19 : Spectre de DC pour les 3 structures protéiques secondaires majeures**  
**L'ellipticité moyenne des structures tertiaires majoritaires des protéines est donnée en fonction de la longueur d'onde**

rouge : spectre de la myoglobine, protéine essentiellement composée d'hélices  $\alpha$

bleu : spectre de la concanavaline, protéine essentiellement composée de feuillets  $\beta$

jaune : spectre du collagène de type VI , protéine essentiellement composée d'hélices polyproline II



## **RESULTATS**



## 1. PARTIE I : Etude d'une série de patients avec myopathie congénitale à core associée à une atteinte cardiaque

Les **cardiomyopathies** sont des affections relativement **fréquentes**, sans distinction ethnique, dans lesquelles la **fonction du myocarde est altérée** et qui peuvent être associées à une atteinte du muscle squelettique. Dans cette première partie des résultats de ma thèse, je décrirai l'analyse centrée sur le gène de la titine (*TTN*) d'une série de patients avec cardiomyopathie primaire associée à une myopathie congénitale à multiminicores. Parmi les différents patients de cette cohorte, 5 se sont révélés porteurs de nouvelles mutations *TTN* non répertoriées, les résultats d'analyse de 4 d'entre eux constituent le développement de cette première partie (**Fig. 20**). L'étude du 5ème patient est rapportée en partie 2.

### 1.1. Constitution d'une série de patients: critères d'inclusion

Mon équipe d'accueil est un laboratoire de référence pour l'étude de MmD. Suite à la description par l'équipe d'une forme de MmD avec cardiomyopathie dilatée fatale due à des délétions homozygotes de la *TTN* (Carmignac, Salih et al. 2007), nous avons collecté grâce à des collaborations internationales une série de **24 familles** (23 cas sporadiques plus 1 famille consanguine d'origine turque) présentant un phénotype peu décrit dans la littérature. Les 23 cas sporadiques sont tous issus de parents non-consanguins d'origine ethniques et/ou géographiques variées, rendant l'hypothèse d'un **effet fondateur peu probable**. Les critères d'inclusion étaient : 1) présence d'une **faiblesse musculaire congénitale ou infantile** (avant l'âge de 2 ans) ; 2) une **cardiomyopathie primaire** ; et 3) présence de **lésions de type minicore** sur la biopsie musculaire (**Tableau 2**). La décision d'analyser le gène candidat *TTN* dans cette série repose sur deux observations : la vérification en amont d'absence de mutations dans les principaux gènes connus retrouvés associés à des cardiomyopathies et la similarité phénotypique de la série de patients avec les patients précédemment décrits avec les mutations homozygote dans la ligne M.

Suite à la plus grande facilité d'accès et la démocratisation du séquençage de nouvelle génération les patients pour lesquels des mutations de *TTN* n'ont pas encore été identifiées seront analysés en NGS afin d'exclure rigoureusement tous les exons codants les isoformes adultes de la *TTN*. Dans le même temps et si toujours aucune mutation n'est identifiée dans *TTN*, les résultats du NGS donneront accès à un panel de SNP qui permettra de réaliser une cartographie fine pour une analyse de liaison à la recherche d'autres gènes candidats positionnels dans ces familles.

	P1	P2	P3	P4
âge des premiers symptômes (mois)	15		24	24
premiers symptômes	hypotonie	Faiblesse des muscles proximaux (plus importante) et distaux, ptosis		
atteinte respiratoire (âge du diagnostic, ans)	Oui (12)	non		
rétraction	coudes, genoux	non	coudes, chevilles	coudes, chevilles
histologie	core centraux, fiber type 1 prédominance des fibres de type 1	multiminicore, noyaux centraux, prédominance des fibres de type 1	variation de la taille des fibres, centralisation nucléaire importante, légère fibrose endomysiale, augmentation du tissu adipeux interstitiel, multiminicore	multiminicore
atteinte cardiaque (âge du diagnostic, ans)	cardiomyopathie hypertrophique (16/17)	cardiomyopathie dilatée	cardiomyopathie hypertrophique (vers 20 ans)	non
complication	alteration majeure de la fonction systolique	non		
scoliose	cyphoscoliose, lordose	non		

**Tableau 2 : Données cliniques de 4 patients du groupe avec atteinte cardiaque**

## 1.2. Etude génétique

### 1.2.1. Etat des lieux au début du projet

Chez tous les cas sporadiques l'exclusion, réalisée au laboratoire, de mutations dans les gènes majeurs liés aux cardiomyopathies associées à une atteinte musculaire squelettique, notamment *FHL1*, *DES* et *LMNA* a précédé leur inclusion dans la cohorte. En parallèle, dans le cadre d'une recherche de diagnostic l'exclusion des gènes *ACTC*, *LAMP2*, *MYBPC3*, *MHY7*, *MYL2*, *MYL3*, *PRKAG2*, *TNNT2*, *TPM1*, *RYR1* a été réalisée par les équipes des cliniciens nous ayant adressé les échantillons

des patients.

En revanche, pour la famille consanguine d'origine turque (**famille 1, Fig. 20**), l'inclusion s'est faite suite aux résultats d'**analyses de liaison** du génome des parents et des deux enfants atteints. Le criblage génomique a été réalisé en deux étapes, une première avec un panel de **450 marqueurs microsatellites**, suivi d'un génotypage avec **250 000 SNPs** (Affimetrix, Nsp-250K) (*cf. partie III pour plus de détails*). La faible informativité *per se* de la famille, du fait de sa petite taille, n'a pas permis l'identification d'un locus avec LODscore significatif d'une liaison génétique avec le phénotype observé. Néanmoins, **3 loci** sur les chromosomes 2 (41Mb), 4 (11Mb) et 9 (18Mb) ont présenté un LODscore équivalent au score maximum simulé pouvant être obtenu dans cette famille ( $Z=1,2037$ ). Une hiérarchisation intégrant le profil d'expression ainsi que la fonction et l'implication dans la physiologie musculaire des protéines codées par les gènes candidats positionnels présents dans ces loci, nous a permis d'identifier **TTN** comme candidat fonctionnel le plus probable. Par conséquent, nous avons décidé d'intégrer cette famille à notre cohorte et ce malgré une apparition tardive du phénotype cardiaque.

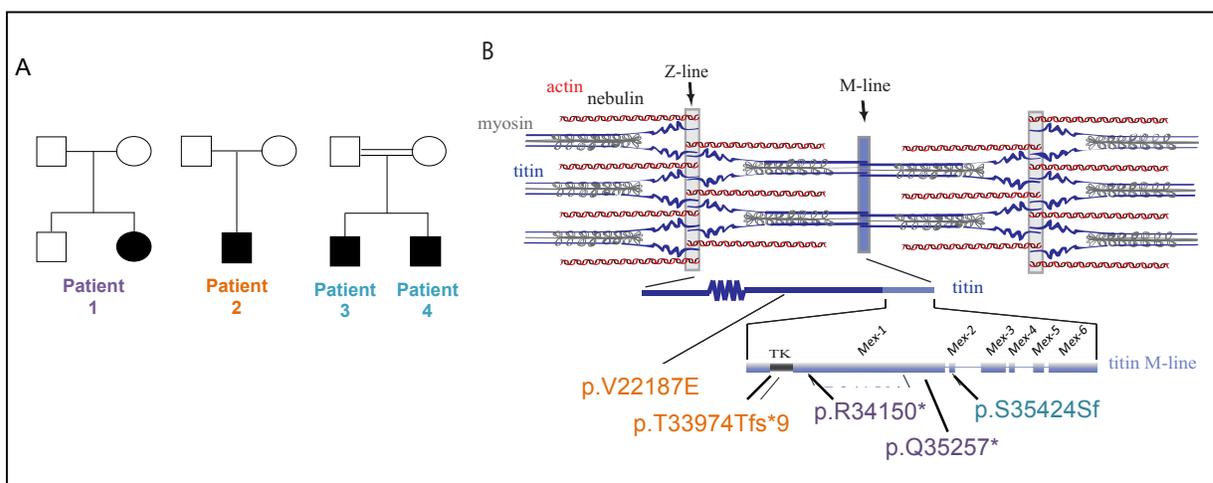
### 1.2.2. Analyse du gène candidat TTN : séquençage traditionnel

La séquence codante du gène *TTN* s'étend sur plus de 100 kb, répartie en **363 exons** majoritairement composés de motifs nucléotidiques répétés ; aussi son séquençage intégral représente une tâche extrêmement difficile. Convaincus de l'importance de la TTN dans les atteintes du muscle strié, nous avons néanmoins décidé de relever ce défi, et j'ai commencé au début de cette thèse le séquençage manuel de *TTN* par méthode traditionnelle de type Sanger, le NGS n'étant pas encore disponible. Etant donné que les mutations *TTN* préalablement décrites, identifiées par l'équipe dans un phénotype congénital musculaire et cardiaque, se situent dans les exons 358 et 360, nous avons choisi de débiter l'analyse par les **6 derniers exons** (358 à 363) de la partie 3' du gène, codant pour le domaine protéique localisé au niveau de la **bande M et représentant 7% de la séquence codante**.

Le séquençage de ces exons chez le propositus de la famille turque (Patient 3) a ainsi mis à jour une **délétion homozygote** de deux nucléotides dans **Mex2** (exon359). La présence de cette mutation **c.106272-106273delTG (p.Ser35424Serfs\*11)** a ensuite été confirmée à l'état homozygote dans l'ADN génomique du frère atteint (Patient 4) et à l'état hétérozygote chez chacun des parents, témoignant d'une pathologie autosomique récessive liée à la consanguinité (**Fig. 20**).

D'autre part, nous avons identifié chez la Patient 1 deux mutations non-sens **hétérozygotes**

**composites** dans Mex1 (exon 358), en aval du domaine kinase TK de TTN. Les études de ségrégation dans la famille ont révélé des origines respectivement paternelle et maternelle des variants **c.102388C>T (p.Arg34150\*)** et **c.105697C>T (p.Gln35257\*)**, chaque changement étant présent à l'état hétérozygote dans les ADNs génomiques des parents. Le frère non atteint de la patiente était homozygote sauvage pour les deux positions. Ces deux mutations prédisent la synthèse d'une TTN tronquée dans la bande M de respectivement 1 796 et 689 résidus (**Fig. 20**).



**Figure 20 : Pedigree (gauche) et mutations identifiées chez les patients 1 à 4 de la cohorte**

Mex1 était également impliqué pour le patient 2, chez lequel le séquençage de la bande M a révélé une délétion hétérozygote d'un nucléotide au niveau de TK prédisant un décalage du cadre de lecture et l'apparition d'un PTC après 9 triplets résultant en la disruption du domaine kinase. Cette mutation **c.101922delT (p.Thr33974Thrfs\*9)** avait été transmise par le père du patient, qui ne présentait pas de phénotype (**Fig.10**). Des investigations cliniques et cardiologiques (électrocardiogramme, échocardiogramme) complémentaires ont alors confirmé l'**absence complète chez le père d'anomalies** aussi bien cardiaques que musculaires. La mère du patient n'était pas porteuse de la mutation. Afin d'expliquer la différence de phénotype entre le patient et son père non atteint, nous avons émis l'hypothèse d'une hétérozygotie composite similaire à celle observée chez la patiente 1, avec une seconde mutation *de novo* ou d'origine maternelle affectant uniquement le deuxième allèle du patient. Par conséquent, nous avons poursuivi le séquençage de *TTN* à la recherche de cette deuxième mutation. Une autre région, la **ligne Z**, a suscité notre intérêt. Contrairement au reste du gène, elle est constituée de beaucoup de séquences uniques et présentait donc l'avantage de pouvoir être facilement analysée. De plus, elle code une région de TTN impliquée dans de nombreuses

interactions protéiques, ce qui augmente la probabilité d'un effet délétère des mutations la concernant. Nous avons donc séquencé les **28 premiers exons de TTN** codant la ligne Z et la jonction ligne Z / bande I, sans trouver de mutation. Finalement nous nous sommes tournés vers une autre région particulièrement importante dans la fonction de TTN : le domaine **PEVK**. Séquencer **101 exons** de la région PEVK chez le patient 2 fut un exercice long, difficile et qui s'est avéré infructueux. Arrivés à ce stade, nous avons séquencé **135 exons de TTN** sans trouver la deuxième mutation permettant de confirmer le diagnostic moléculaire du patient. Les limites techniques qui s'imposaient à nous et l'évolution des techniques de séquençage nous ont alors conduit à envisager une stratégie alternative qui nous permettrait de finaliser l'analyse de *TTN*. Le développement rapide du NGS ouvrait à la fois de nouvelles possibilités tout en posant la question de l'analyse d'une grande quantité de données. Cette expertise n'était pas disponible à l'époque à Paris mais était l'un des axes de recherche du Dr Mark Samuels à Montréal. L'idée d'intégrer le patient 2 dans un projet de **séquencage d'exome** haut débit (WES, *whole exome sequencing*) que je mènerais à Montréal dans le cadre d'une thèse en cotutelle s'est donc tout naturellement imposée.

### 1.2.3. Analyse d'exome et identification d'une deuxième mutation hétérozygote composite de *TTN* chez le patient 2

Pour réaliser le séquençage de l'exome, nous avons opté pour l'utilisation de la librairie d'oligonucléotides **Agilent SureSelect All exon kit 50Mb** (*Agilent, Santa Clara, CA*), qui à l'époque (avril 2009) était la seule à couvrir l'intégralité de l'exome humain (180 000 exons) en ciblant un total de 50 Mb de séquence qui incluait la **totalité de l'isoforme N2A** de *TTN* répertoriée dans la base de données du NCBI (312 exons, NM\_133378.4). Le Whole Exome Sequencing (WES) a généré un total de 78,8 millions de séquences, dont 61% ont pu être alignées sur les séquences de référence du génome humain (GRCh37/hg19) par le logiciel d'analyse NextGene (SoftGenetics). Parmi ces séquences cartographiables, **27,6 millions** (57%) avaient été correctement ciblées par la banque d'oligos ARN. Les séquences avaient une taille moyenne de 48 nucléotides et une couverture moyenne de 30x. Leur analyse filtrée contre les banques de données **dbSNP, 1000 génomes** ainsi qu'une **banque interne** au centre de recherche de l'hôpital Ste Justine (Montréal, QC, Canada) a confirmé la présence de la délétion de TK d'origine paternelle précédemment identifiée (**Fig. 21**), et a mis à jour 31 changements de séquences additionnels. Un seul de ces polymorphismes a été confirmé par séquençage traditionnel, les 30 autres étant dus à des artefacts du WES. En effet afin d'identifier des possibles mutations nous avons interrogé avec le logiciel NextGENe chaque position/base de

l'exome de sorte à générer une liste de positions avec une différence allélique d'au moins 25%. En d'autres termes, les positions pour lesquelles ils existaient minimum deux et jusqu'à 4 valeurs différentes (A, T, G, C) et avec au moins 25% des amplicons ne présentant pas la même valeur. Par conséquent si la couverture à la position interrogée est faible alors une erreur de séquençage sur un ou deux amplicons peut générer un taux de variations supérieure ou égal à 25 % et donc des faux positifs. Le nouveau variant identifié résultait en la substitution d'une valine par un glutamate dans un domaine Fn3 conservé, localisé à la fin de la bande I et codé par l'exon 316 (c.66560G>A ; p.V22187E). La mère du patient était hétérozygote pour ce variant, qui n'était pas présent chez le père ; de plus, l'analyse du reste de l'exome nous a permis d'exclure des mutations potentielles dans d'autres gènes associés à une atteinte musculaire squelettique et cardiaque. Ces résultats combinés prouvent la présence, chez le patient 2, d'une pathologie autosomique récessive associée à deux nouvelles mutations hétérozygotes composites de *TTN*.

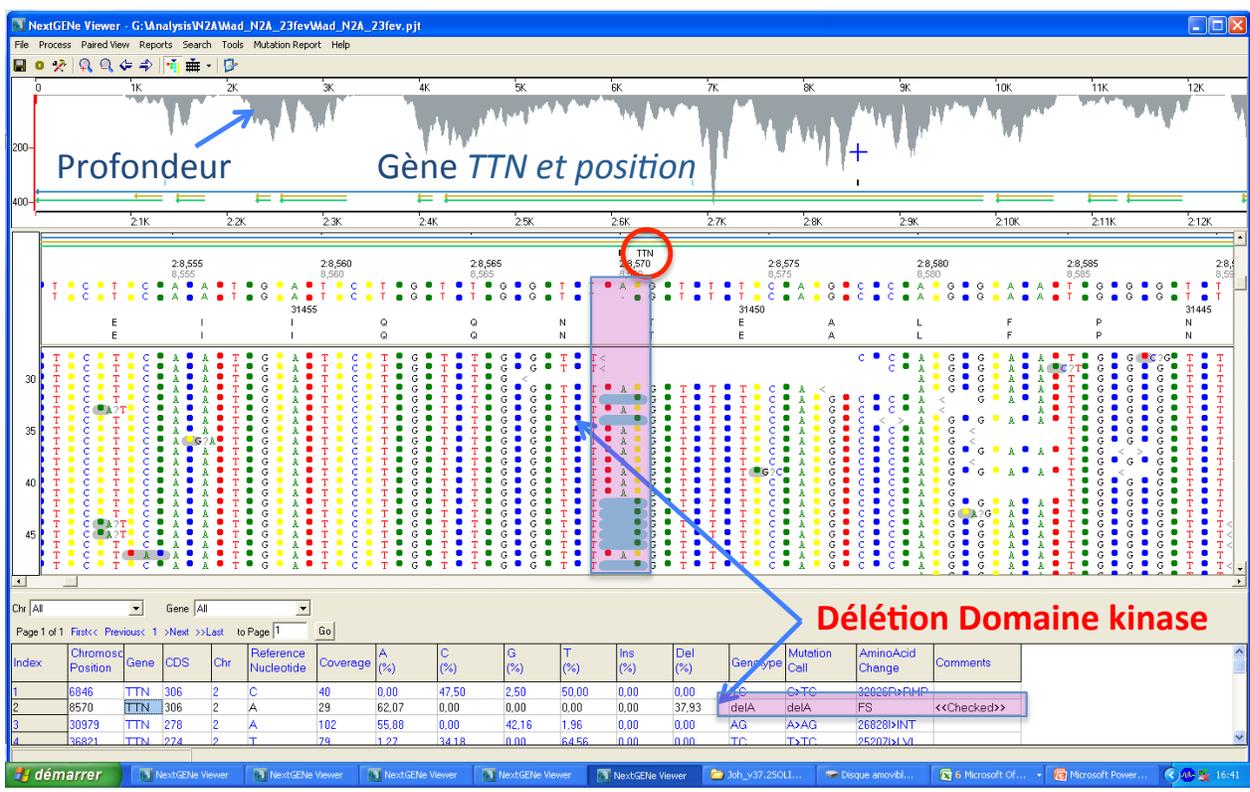


Figure 21 : Capture d'écran du logiciel NextGENe figurant la délétion en Mex1 identifiée chez le patient 3. La position nucléotidique de l'exon Mex1 présentant la délétion est couverte par 29 reads. Cette délétion est présente dans 37,93% des 29 reads.

### 1.3. Phénotype des familles avec mutations *TTN* identifiées

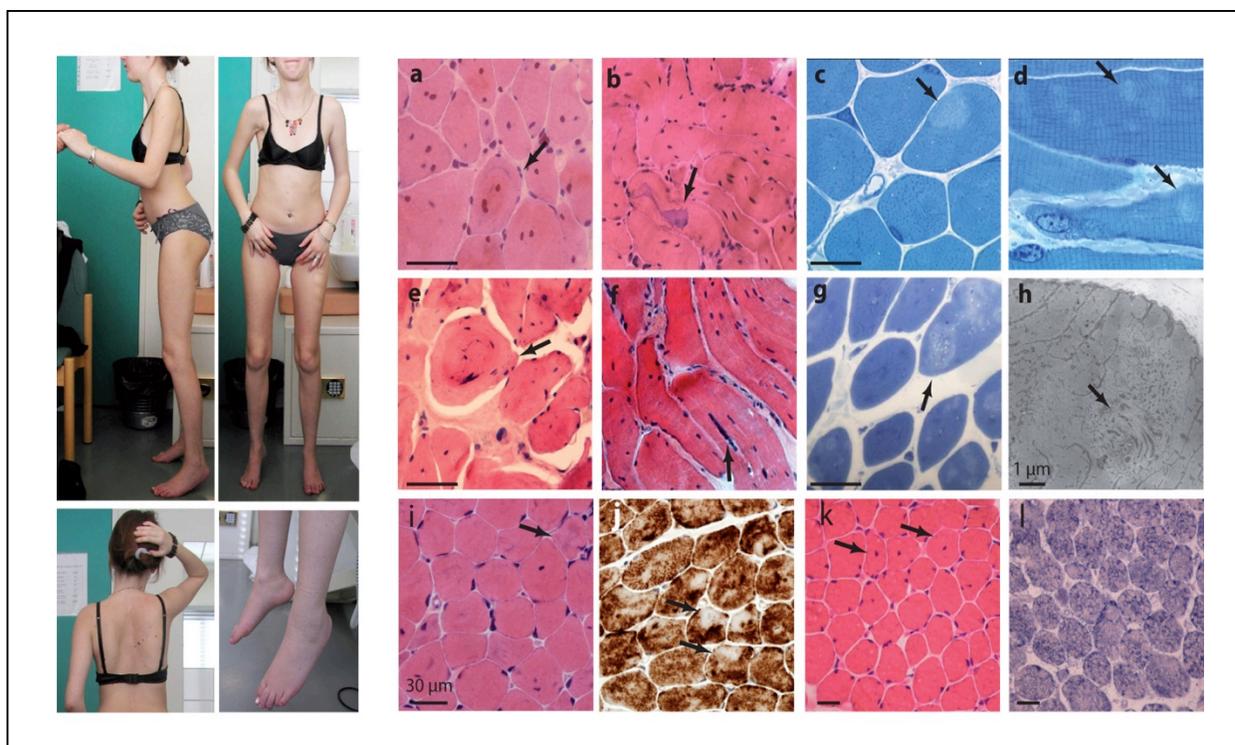
Les 3 familles (4 patients) avec de nouvelles mutations de la *TTN* associées à un phénotype autosomique récessif pur montraient un phénotype clinique et morphologique commun et reconnaissable, bien qu'avec une sévérité variable notamment en ce qui concerne l'atteinte cardiaque.

La **Patiente 1**, est née à terme sans complications de parents d'origine française. Elle a présenté une hypotonie néonatale attribuée initialement à une communication inter-auriculaire (défaut de fermeture du foramen ovale dans le septum inter-auriculaire du cœur) qui s'est refermée spontanément. Elle a présenté un retard dans l'acquisition du soutien de sa tête et de la marche (acquise à 2 ans). Elle n'a jamais été capable de courir ou de sauter et a toujours eu des difficultés pour monter les escaliers. A 7 ans, un steppage (difficulté pour la marche due à l'impossibilité de relever correctement les pieds, provoquée par une faiblesse des muscles tibiaux antérieurs) a été remarqué, ainsi que des rétractions tendineuses des articulations des coudes et des chevilles. Son taux de CK était élevé aux alentours de 6 fois la normale. L'apparition d'une **scoliose thoracique** a entraîné des complications respiratoires à l'âge de 12 ans à la suite desquelles la patiente a subi une chirurgie corrective à 13 ans. En plus d'une **amyotrophie globale diffuse**, la patiente présentait une **faiblesse musculaire d'évolution lente**. A partir de l'âge de 16 ans, elle a développé une cardiomyopathie dilatée touchant les ventricules droit et gauche (fractions de raccourcissement et d'éjection de 21% et 40% respectivement à 23 ans) ainsi que le septum, sans hypertension pulmonaire. L'atteinte cardiaque reste stable sous traitement.

Le **patient 2**, un garçon, présentait une **faiblesse congénitale bénigne et non progressive** ; il n'a jamais été capable de courir et avait des difficultés pour monter les escaliers depuis l'âge de 5 ans. La faiblesse touchait aussi bien les **muscles proximaux** que distaux. Il avait également des rétractions discrètes des tendons d'Achilles et du rachis, ainsi qu'une scoliose discrète. Son niveau de créatine kinase (CK) était normal. La faiblesse musculaire reste stable mais il a développé à l'âge de 15 ans une cardiomyopathie dilatée très rapidement progressive entraînant une défaillance cardiaque terminale et une greffe cardiaque quelques mois après. Enfin, les **patients 3 et 4**, deux frères nés de parents **consanguins** au premier degré (cousins germains) d'origine turque, présentaient une myopathie congénitale avec des **rétractions articulaires**, localisées au niveau des coudes et chevilles, au premier plan de la clinique. L'âge d'apparition des premiers symptômes, qui consistaient en une faiblesse musculaire d'évolution lente associée aux rétractions, était de 2 ans pour les deux enfants. Le niveau de CK chez ces patients était élevé (facteur de 4 par rapport à la normale) et des lésions

compatibles avec des **multiminicores** étaient observables sur les biopsies de muscle squelettique. Contrairement aux autres patients de la série, l'**atteinte cardiaque** détectée par échocardiographie a été **tardive et bénigne** ; une cardiomyopathie dilatée, débutante asymptomatique, a été détectée suite à une échocardiographie systématique à l'âge de 19 ans chez le frère aîné, alors que le cadet n'avait aucune anomalie échocardiographique à 16 ans.

Du point de vue histologique, tous les patients présentaient des lésions de type multi-minicore en l'absence de signe de nécrose, de fibrose endomysiale ou de régénération, ce qui était l'un des critères d'inclusion dans l'étude. Plusieurs autres caractéristiques pathologiques ont été observées associées aux minicores, notamment une prédominance des fibres de type 1 et un degré variable de noyaux anormalement positionnés à l'intérieur de la fibre musculaire. Bien que les lésions histologiques étaient similaires et reconnaissables chez la majorité des patients, la sévérité était variable. La patiente 1 notamment présentait un très grand nombre de **noyaux internalisés**, parfois organisés en chaînes (**Fig. 22 ; f**), et ceci plus particulièrement sur sa deuxième biopsie, suggérant une évolution de la pathologie. Comparativement les patients 2 (**Fig. 22 ; i-j**) et 3 (**Fig. 22 ; k-l**) présentaient peu de ces noyaux centraux.



**Figure 22 : Planche résumant les données cliniques et morphologiques des 4 patients avec titinopathie récessive (MmD et cardiomyopathie dilatée) de la série**

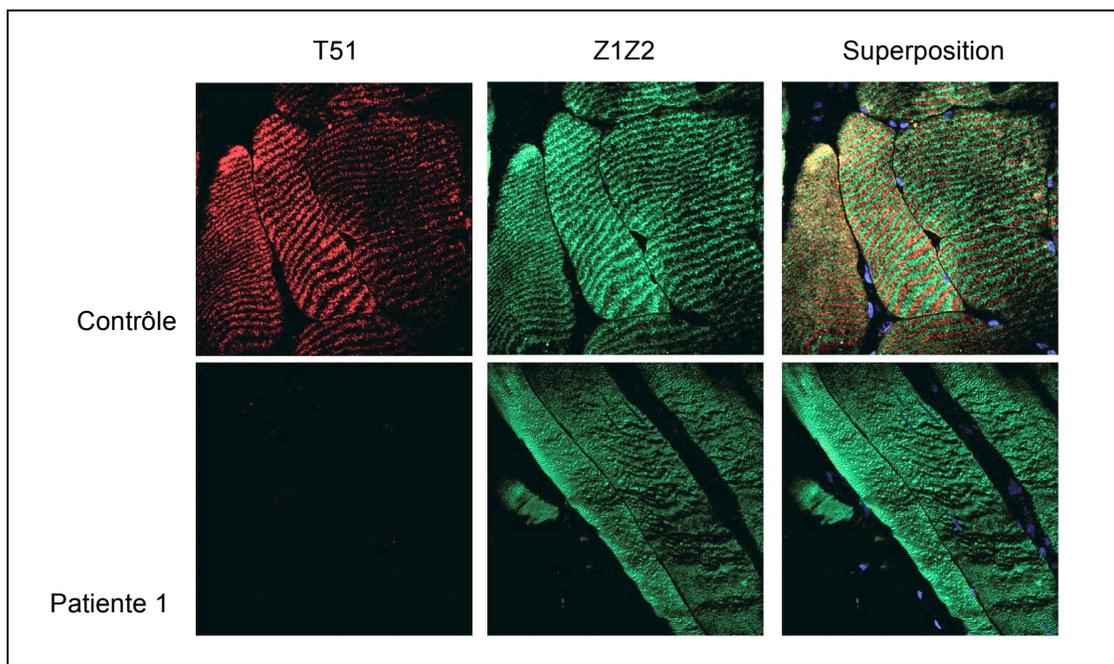
Les photos cliniques (gauche) de la patiente 1 montrent des rétractions importantes des tendons d'Achille (chevilles, profil) et des coudes (face) ; un thorax plat ; une amyotrophie globale (profil et face) ; une scoliose chirurgicalement corrigée ainsi que des pieds tombants. Du fait de la faiblesse des fléchisseurs du cou et d'une raideur du rachis, la patiente est obligée de tenir sa tête lors de la marche ou la station debout. Les analyses histologiques de biopsies de quadriceps de la patiente 1 à 14 ans (a-d) et 23 ans (e-h), du patient 2 à 12 ans (i-j) et 3 à 16 ans (k-l) sont présentées à droite. Le marquage hématoxyline-éosine de sections transversales (a, e, i, k) et semi-longitudinale (b, f) montre des centralisations nucléaires abondantes, des fibres « annulaires » présentant deux zones concentriques (a et e, flèches), ainsi que des zones basophiles caractéristiques localisées au centre des fibres (b et k, flèches). Des minicores, zones focales de déplétion en mitochondries (j, cox et l, NADH) et désorganisation des sarcomères (transverse (c, g), et section longitudinale semi-fine (d), microscopie électronique (h)), étaient présents chez tous les patients.

## 1.4. Etude fonctionnelle

### 1.4.1. Expression des TTN mutées chez la patiente 1

Nous disposions d'une biopsie de muscle squelettique uniquement de la patiente 1, sur laquelle nous avons réalisé des analyses d'immunofluorescences avec des anticorps dirigés contre les deux extrémités de TTN, ligne Z et bande M ; pour cette dernière nous avons utilisé un anticorps dirigé contre un épitope codé par Mex3, en aval des deux mutations Mex1 de la patiente. Nous avons ainsi pu observer une absence totale de marquage de la bande M dans l'échantillon de la patiente, au contraire du tissu contrôle où une striation préservée caractéristique de la structure des sarcomères, avec une alternance des deux anticorps, a été observée. En revanche, le marquage avec l'anticorps dirigé contre la ligne Z présentait une striation d'aspect similaire chez la patiente et le contrôle (**Fig. 23**).

Ainsi, malgré une troncation en C-terminal au niveau de la bande M, la TTN s'avère être **correctement intégrée** à la structure du sarcomère chez la patiente 1 ; de plus la striation observée avec l'anticorps spécifique de la ligne Z semble indiquer que les mutations retrouvées chez cette patiente n'entraînent pas de défaut dans la formation des sarcomères.



**Figure 23 : Analyse de la TTN par immunofluorescence (IF)**

L'IF a été réalisée sur des cryosections de quadriceps d'un contrôle (haut) et de la patiente 1 (bas) avec des anticorps dirigés contre des épitopes localisés en amont, ligne Z (Z1Z2, vert) et en aval, bande M (T51, rouge) des mutations de la patiente. Une absence de marquage avec l'anticorps T51 associée à une striation observée avec l'anticorps Z1Z2, sont compatibles avec une expression et une bonne intégration de TTN tronquées dans les sarcomères. La présence des noyaux est observée par un marquage au dapi (bleu). Magnification : 63X.

En Conclusion de cette première partie de ma thèse, mon travail a permis de :

**Confirmer l'existence d'une forme de MmD autosomique récessive typique et reconnaissable associée aux mutations de TTN impliquant la bande M.** Ce phénotype est marqué, du point de vue clinique, par une faiblesse musculaire débutant dans les 2 premières années de la vie, par une cardiomyopathie dilatée et par une atteinte rétractile importante qui peut être dans certains cas comparable à une dystrophie musculaire d'Emery-Dreifuss.

**Elargir le spectre de sévérité de la maladie**, notamment en ce qui concerne la cardiomyopathie. En effet, dans les deux familles précédemment décrites avec cette forme de titinopathie, l'atteinte cardiaque débutait dans l'enfance et a conduit au décès des 5 patients avant l'âge adulte. L'étude des patients 3 et 4 montre que l'atteinte cardiaque peut être beaucoup plus tardive et bénigne, et que son absence à l'adolescence n'exclut pas une implication de la TTN.

**Identifier 5 nouvelles mutations TTN privatives, dont pour la première fois deux familles hétérozygotes composites.** Etant donné que la prévalence de mutations tronquantes de TTN a été

calculée à 3% de la population générale (Herman, Lam et al. 2012), nous proposons que cette situation d'hétérozygotie composite peut être fréquente et à l'origine d'un spectre large de maladies musculaires et/ou cardiaques.

**Confirmer l'importance de la bande M de la TTN dans la pathologie musculaire squelettique, avec ou sans atteinte cardiaque.** Néanmoins, des mutations de ce domaine ont été exclues chez 20 patients de notre série (80%), ce qui suggère l'implication d'autres domaines et/ou protéines.

**Valider l'apport important du NGS pour l'étude des gènes géants, dont *TTN* est le paradigme.**



## 2. PARTIE II : Une nouvelle forme de MmD avec non-compaction ventriculaire et arthrogrypose est causée par des mutations *TTN* hétérozygotes composites conduisant à une perte total de fonction du domaine kinase.

La patiente sur laquelle va porter la deuxième partie des résultats de ma thèse a été recrutée dans le cadre de l'étude de la série de cas de myopathie à multiminicores associées à une atteinte cardiaque (*cf. partie I*). Ses particularités phénotypiques la distinguant des autres patients de la cohorte, son cas va être traité séparément.

### 2.1. Etude clinique : un nouveau phénotype

#### 2.1.1. Présentation du patient : phénotype et contexte familial

Au moment de son inclusion dans l'étude, la patiente représentait un **cas sporadique** issu de parents **caucasiens non consanguins** d'origine germano-américaine. La famille comprenait également un frère aîné cliniquement sain. Durant ma dernière année de thèse nous avons été contactés par le clinicien de la patiente pour un diagnostic prénatal suite à une nouvelle grossesse dans la famille.

Concernant la patiente 5, outre des mouvements fœtaux réduits, la mère n'a pas déclaré d'autres problèmes anténataux. Une **hypotonie** discernable dès la naissance était accompagnée d'un retard du développement moteur et d'une **arthrogrypose** touchant les épaules, coudes, poignets, mains et articulations des membres inférieurs. L'arthrogrypose congénitale multiple est un syndrome, dont l'incidence est estimée à 1/3000 naissances, qui est caractérisé par une fixation ou raideur de certaines articulations (Bamshad, Van Heest et al. 2009), le défaut de mobilité étant dû à une insuffisance musculaire par atrophie ou par fibrose et à une rétraction des ligaments articulaires. L'hypotonie néonatale s'est accompagnée d'un retard du développement psychomoteur.



**Figure 24 : Phénotype clinique**

Arbre généalogique (A). Rétractions touchant les membres inférieurs et supérieurs, rétrognathisme (B, 3 mois) et cyphoscoliose (radiographie du rachis à l'âge de 6 ans, C) observable chez la patiente à l'âge de 6 ans. Non compaction ventriculaire révélée par échocardiographie (D).

Entre autres caractéristiques observables, la patiente présentait également un cou court, des oreilles d'implantation basse, un rétrognathisme (mâchoire projetée en arrière) et une cyphoscoliose (double déformation de la colonne vertébrale, associant une déviation latérale [scoliose] et une déviation à convexité postérieure [cyphose]) observable dès les premiers mois et qui a progressé pendant l'enfance (**Fig. 24**).

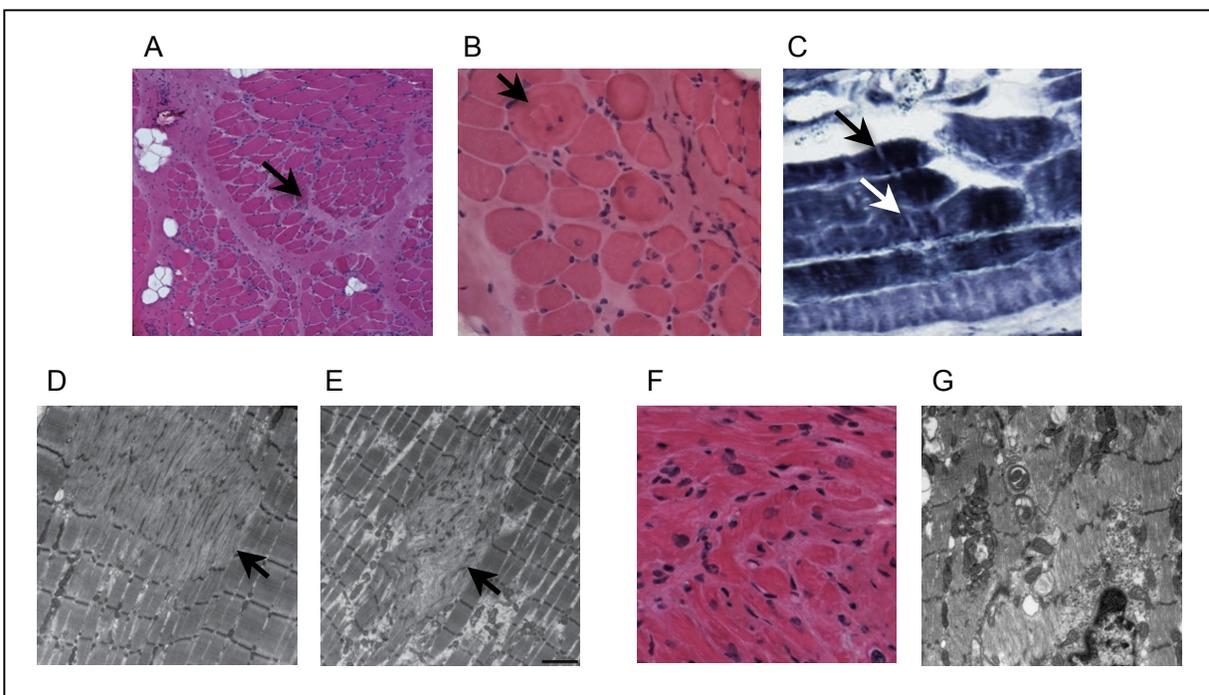
Sur le plan cardiaque, un syndrome de **non compaction du ventricule gauche (LVNC)** a été diagnostiqué à la naissance et a conduit à une insuffisance cardiaque terminale nécessitant une greffe de cœur à l'âge de 5 ans. La LVNC est une cardiomyopathie congénitale liée à un arrêt prématuré de développement du tissu cardiaque, conduisant à une apparence du myocarde en double couche : une externe compacte mince et une interne non compactée à l'aspect spongieux et qui présente des profondes trabéculations dans la paroi ventriculaire. La couche sous-endocardique spongieuse prédomine dans les régions apicales et médianes du ventricule gauche, mais épargne généralement le septum. Les conséquences de la LVNC sont un défaut de contraction conduisant à une hyposystolie (diminution de la contraction du cœur et dilatation excessive de ses cavités) pouvant entraîner une tachycardie ventriculaire et une mort subite.

Les examens cliniques et cardiologiques (échocardiographie, ECG) des parents aux âges de 38 ans (père) et 42 ans (mère) n'ont révélé aucune anomalie significative, et il n'y avait pas d'histoire de cardiopathie dans la famille.

### 2.1.2. Phénotype morphologique

L'étude histologique d'une biopsie endomyocardique (ventricule gauche) a montré des **noyaux hypertrophiques** et une **fibrose** compatibles avec une cardiomyopathie. L'anatomie pathologique de l'explant après greffe cardiaque a montré un myocarde extrêmement spongieux et confirmé **l'importante non-compaction**. L'analyse de la biopsie de muscles squelettiques de la patiente a montré un phénotype histologique compatible avec une myopathie congénitale (MC) avec des lésions de type **minicore** au sein des fibres musculaires et également de nombreux **noyaux internalisés** ainsi qu'une très discrète **fibrose endomysiale**. Les études en microscopie électronique ont révélé des sarcomères d'aspect normal jouxtant des zones de fibres musculaires avec une **désorganisation complète** présentant des bavures et des **agrégats** de matériel électron-dense au niveau des disques Z (**Fig. 25**). Ces observations rappellent celles associées à des mutations des protéines constitutives des disques Z comme la myotiline, ZASP/LDB3 (Z-disk alternatively spliced PDZ-containing protein) ou encore la protéine chaperonne impliquée dans l'autophagie BAG3 (Bcl-2-

associated athanogene-3). Des mutations de ces protéines ont été liées à un turn-over défectueux des protéines et à des cas de cardiomyopathies (Selcen 2011).



### Figure 25 : Phénotype morphologique.

Les cryosections transversales du tissu musculaire squelettique (HE, 20x (A) et 60x (B)) révèlent une fibrose endomyofibrillaire discrète (A, flèche), des noyaux centralisés (A, chevron) et quelques fibres occasionnelles d'aspect annulaire avec des noyaux multiples non-périphériques (B, flèche). L'analyse des sections longitudinales (NADH-TR, flèches (C) et en microscopie électronique (D-E)) témoignent de la présence de zones focales de déplétion en mitochondries et de désorganisation du sarcomère (minicores) concernant préférentiellement la ligne Z (D) ou bien l'intégralité du sarcomère (E). Les analyses microscopiques (HE, F) et optique (G) du ventricule gauche montrent des noyaux hypertrophiés et une fibrose du tissu.

## 2.2. Etude génétique

### 2.2.1. Etat des lieux au début du projet

Les manifestations cliniques observées chez la patiente n'avaient jamais été décrites conjointement dans un même patient et constituent donc un **nouveau phénotype**.

Le dossier de la patiente nous a été adressé après des études infructueuses des principaux gènes candidats liés aux cardiopathies et arthrogyposes congénitales : *ACTA1*, *LAMP2*, *MYBPC3*,

**MHY7, MYL2, MYL3, PRKAG2, TNNT2, TPM1, RYR1** et **LMNA** (séquençage des régions codantes et jonctions intron-exon pour analyse des sites d'épissage).

Sur la base d'un phénotype associant une atteinte cardiaque et squelettique, ainsi que devant l'aspect des lésions histologiques (minicores et noyaux centraux compatibles avec une titinopathie), la patiente a été incluse dans la série de patients avec atteinte cardiaque décrite dans la première partie des résultats.

### 2.2.2. Analyse du gène candidat *TTN* par séquençage Sanger: identification d'un variant TK

L'analyse de la **bande M de TTN** (*cf. partie 1 pour plus de détails*) a permis d'identifier un changement hétérozygote d'une thymine vers une cytosine en Mex1 (**c.T94734C**) qui prédit la substitution d'un acide-aminé tryptophane hautement conservé situé dans le **domaine kinase** (TK) de la TTN vers une arginine (p.W31504R, ou **p.W260R** dans la nomenclature spécifique de la région TK). Ce variant faux-sens n'est pas répertorié dans les bases de données et n'a pas été retrouvé dans 120 ADNs de contrôles d'origine Caucasienne. Les études de ségrégations ont montré une **origine paternelle** du variant, le père de la patiente étant hétérozygote pour ce changement de séquence. En revanche, la mère et le frère aîné avaient un génotype identique à la séquence de référence pour les deux allèles de la position (*homozygote wild-type*).

Afin d'expliquer la différence entre le phénotype extrêmement sévère de la patiente et son père parfaitement sain, nous avons émis l'hypothèse d'une récessivité de la maladie et de la présence d'une deuxième mutation hétérozygote de *TTN*, c'est-à-dire, d'une **hétérozygotie composite**. Par conséquent nous avons intégré la patiente dans notre projet de séquençage d'exomes au même titre que le patient 2 (*cf. partie 1*).

### 2.2.3. NGS, analyse d'exome : identification d'un variant d'épissage

L'analyse des exons restants de la *TTN* (*cf. Partie 1 pour les détails*) par séquençage d'exomes a permis l'identification de deux variants privés de *TTN* : celui précédemment identifié dans TK et un changement inconnu présent au niveau du site donneur d'épissage 5' de l'intron 39 au niveau de la jonction strie Z-bande I (**c.G9388+1C**). Ce variant, absent chez 120 contrôles, a été confirmé par **reséquençage** en méthode traditionnelle **Sanger** dans l'échantillon de la patiente et a été retrouvé seulement dans l'ADN génomique de la mère ; le père et le frère ne présentaient pas le variant. La ségrégation de cette nouvelle mutation corroborait donc l'hypothèse d'une transmission **hétérozygote composite** de mutations *TTN* responsables du phénotype autosomique récessif chez la patiente.

## 2.2.4. Diagnostic prénatal

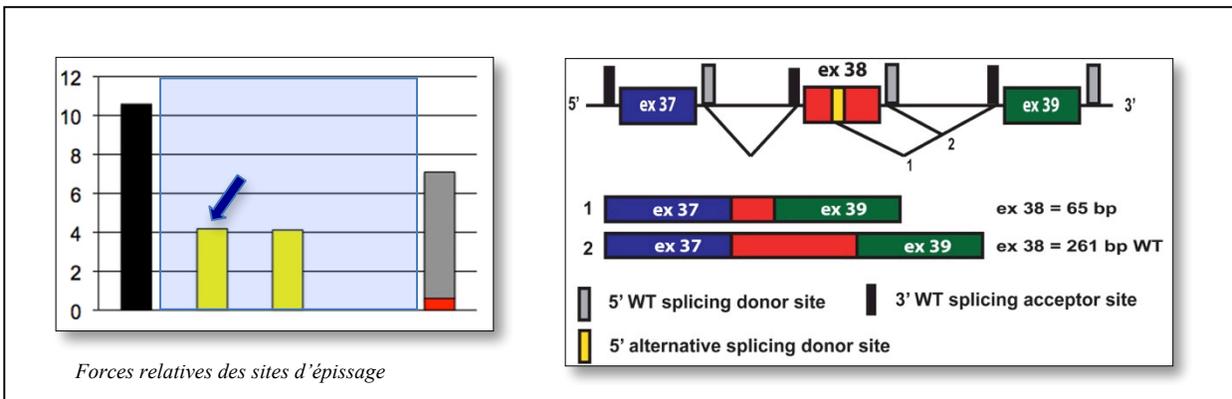
L'identification de ces deux mutations chez la patiente a permis à la famille de bénéficier d'un diagnostic prénatal pour une grossesse déjà en cours chez la mère. Le laboratoire de diagnostic génétique du Children's Hospital Philadelphia, où la famille est suivie par le **Dr Carsten Bonnemann**, a confirmé la présence des mutations et l'étude de ségrégation familiale que nous avons établie, et identifié par séquençage d'un échantillon d'ADN génomique fœtal que le fœtus était hétérozygote composite pour les deux mutations, c'est-à-dire, qu'il avait un génotype identique à celui de sa sœur atteinte. Suite à ce résultat, la décision d'un arrêt de grossesse a été prise par la famille.

## 2.3. Etude fonctionnelle

### 2.3.1. Mutation d'épissage

#### 2.3.1.1. Analyse *in-silico*

Comme la mutation **c.G9388+1C** affecte un **site consensus d'épissage**, nous avons procédé à des analyses *in-silico* (*ESE-finder et mutation Taster*) qui ont prédit la perte du site d'épissage au profit de deux **sites alternatifs** cryptiques pouvant être reconnus et utilisés à la place du site muté par la machinerie d'épissage (**Fig. 26**). L'utilisation de l'un ou l'autre de ces sites provoque un décalage du cadre de lecture et la génération d'un PTC, entraînant la synthèse d'une protéine tronquée au début de l'exon 39 (**Fig. 26**).



**Figure 26 : Analyse prédictive du variant d'épissage**

Gauche : Graphique représentant les valeurs relatives de force en unités arbitraires associées aux différents sites d'épissage (noir, site accepteur ; jaune, site donneur alternatif ; gris, site donneur sauvage ; rouge, site donneur

muté) associés au bornes de l'exon 38 (zone bleutée). Le site donneur alternatif ayant la plus forte probabilité de reconnaissance par le complexe d'épissage est indiqué par une flèche. Droite : Comparaison des tailles estimées de l'exon 38 en cas d'utilisation du site donneur alternatif majeur (1) et sauvage (2) d'épissage.

### 2.3.1.2. Discrimination allélique dans les muscles striés

Pour ces analyses, nous disposons de biopsies cardiaque (ventricule gauche) et de muscle squelettique (paravertébrale) de la patiente, et également de muscle squelettique du fœtus porteur des deux mutations. Afin de vérifier les anomalies d'épissage prédites par les analyses bioinformatiques, nous avons cherché à identifier les ARN messagers (ARNm) synthétisés dans ces différents tissus. Pour ce faire nous avons amplifié et séquencé des ADN complémentaires (ADNc) obtenus à partir d'ARNm extraits des différentes biopsies. Nous avons alors pu confirmer la présence de 3 différentes formes alléliques dans le tissu cardiaque ainsi que dans le tissu fœtal. La sauvage et celles correspondant à l'utilisation des deux sites alternatifs prédit *in-silico*. En revanche, nous avons seulement identifié l'isoforme de plus petite taille dans la biopsie du muscle squelettique de la patiente (Fig. 27).

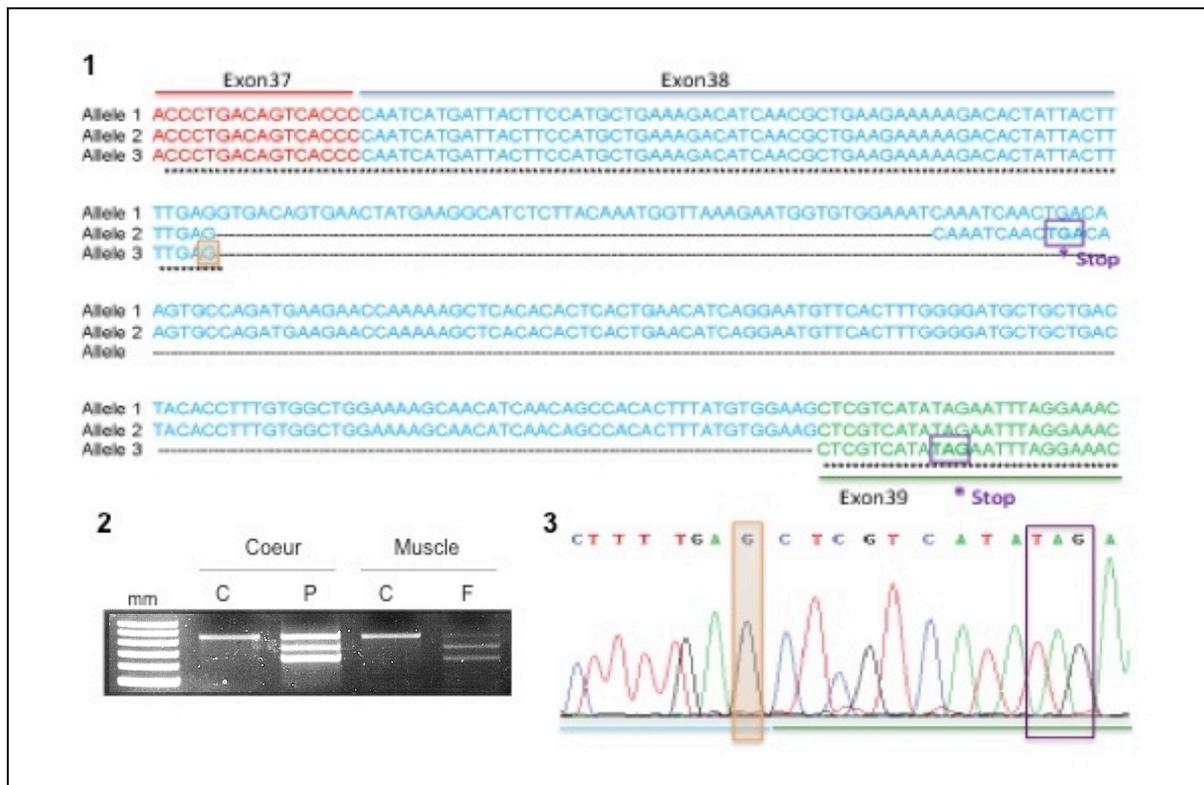


Figure 27 : Analyse de l'expression des ARNm de *TTN* chez la patiente

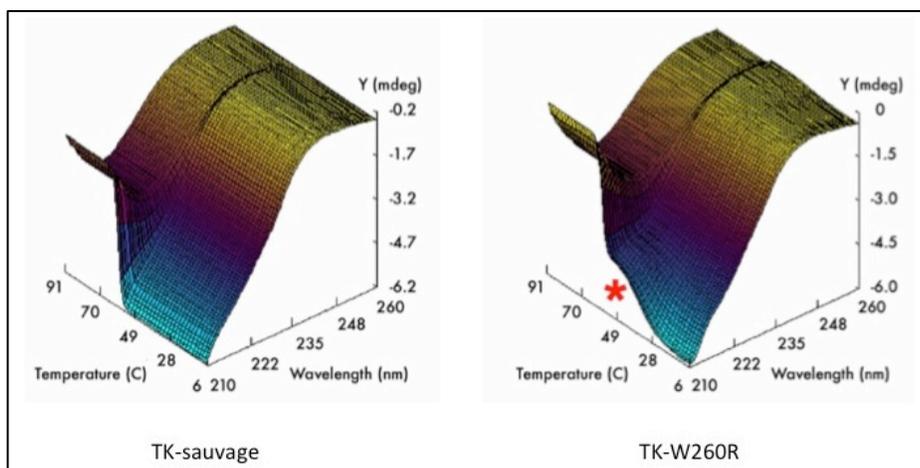
Alignement des séquences des 3 isoformes d'ADNc obtenus à partir des ARNm extraits des biopsies. Les PTCs résultant du décalage du cadre de lecture dans les deux séquences mutées sont encadrés en violet (1). Migration sur gel d'agarose des trois isoformes identifiées (WT, issue de la reconnaissance du site donneur alternatif majeur et issue de la reconnaissance du site donneur alternatif secondaire) dans le tissu cardiaque du patient P et musculaire squelettique du fœtus F ; les contrôles cardiaques (ligne 1) et muscles (ligne 3) ne présentent que la bande de plus grande taille correspondant à l'isoforme sauvage (2). Chromatographe de la seule isoforme mutée, correspondant à l'allèle 3, retrouvée dans le tissu de muscle squelettique de la patiente.

### 2.3.2. Mutation TK

La mutation de TK, p.W260R, touche un résidu hydrophobe fortement conservé et localisé dans le cœur du domaine catalytique : par conséquent, nous avons présumé qu'elle pourrait affecter des interactions requises pour la myofibrillogénèse dans laquelle est impliqué TK.

#### 2.3.2.1. Impact structurel

Dans le but de caractériser de façon approfondie les effets de la mutation identifiée chez la patiente, Mathias Gautel (King's College, London), un de nos collaborateurs sur ce projet et spécialiste du domaine kinase de TTN, a réalisé une analyse de la stabilité thermique du mutant W260R en spectroscopie par mesure du dichroïsme circulaire (CD). Les analyses de la stabilité thermique par spectroscopie de dichroïsme circulaire ont montré une seule transition structurelle majeure avec perte de la structure hélicoïdale à plus de 60 ° C pour le TK-WT, alors que le TK-W260R montre une perte partielle de la structure secondaire dès 35 ° C (Fig. 28).



**Figure 28 : Analyse du dichroïsme circulaire CD**

Les spectres CD enregistrés aux températures de 6 °C à 94 °C pour des longueurs d'onde allant de 200 à 260 nanomètres montrent une transition unique de structure secondaire du TK sauvage (gauche), avec une perte de l'ellipticité négative (y) à 209 nm pour une température proche de 59 °C. En comparaison, TK-W260R montre deux transitions dans l'ellipticité négative, la première survenant vers les 42 °C (droite, astérisque) suivie d'une seconde vers 57 °C.

Nous avons ensuite présumé que la mutation, p.W260R, du domaine kinase TK, visant un résidu hydrophobe fortement conservé au coeur du domaine catalytique, pourrait affecter des interactions protéiques de TTN avec ses partenaires. Nous avons donc analysé les propriétés biochimiques de la mutation TK p.W260R par rapport à TK sauvage. La mutation W260R peut être exprimée dans des cellules d'insectes, est complètement soluble et ne provoque aucune agrégation protéique, comme en témoignent les profils d'éluion normales de chromatographie de filtration sur gel par rapport au TK sauvage. Par conséquent, nous avons exécuté un crible génétique double hybride chez la levure en utilisant une construction comportant le domaine auto-inhibiteur de TK et son domaine fibronectine adjacent (Fn-TK). Nous avons ainsi pu tester l'effet de la mutation p.W260R sur les interactions de TK avec ses ligands connus que sont Nbr1 et la queue auto-inhibitrice de TK et comparer ceci aux effets de la mutation TK dominante HMERF (p.R279W).

Le TK sauvage interagit à la fois avec sa propre queue auto-inhibitrice en *trans* ainsi qu'avec Nbr1. En revanche, les deux interactions ont été complètement abrogées avec la construction TK W260R, en accord avec la position centrale de ce résidu fortement conservé dans le domaine catalytique (figure 2A). Par comparaison, le mutant kinase HMERF R279W maintient une capacité d'interaction avec sa queue auto-inhibitrice tout en perdant les interactions avec Nbr1.

De plus, aucune activité catalytique de l'enzyme mutante n'a été détectée. Ces résultats suggèrent que la mutation p.W260R provoque une perte complète de fonction des TK sauvages qui abroge les voies de signalisations reliant les bande M et disque Z via NBR1, impliquée à la fois dans la myofibrillogénèse et dans les processus autophagiques.

### *2.3.2.2. Impact sur le signalosome*

L'équipe de M. Gautel a ensuite réalisé une étude d'interactions par **criblage double hybride** en exprimant chez la levure une construction comportant TK avec son domaine auto-inhibiteur ainsi que le domaine FN3 directement en amont de ce dernier.

les conséquences de la mutation W260R sur les ligands connus de TK que sont Nbr1 et la queue autoinhibitrice de TK ont été étudiées par comparaison avec les effets de la seule autre mutation connue de TK : la mutation R279W responsable du phénotype dominant HMERF. La construction avec un TK sauvage montre une liaison en *trans* avec sa propre queue inhibitrice ainsi qu'avec Nbr1. La construction HMERF, en revanche, ne montre qu'une liaison avec le domaine autoinhibiteur mais perd la liaison avec Nbr1. Enfin, l'effet le plus marqué se retrouve avec la construction W260R pour laquelle une **perte des deux interactions** a été notée.

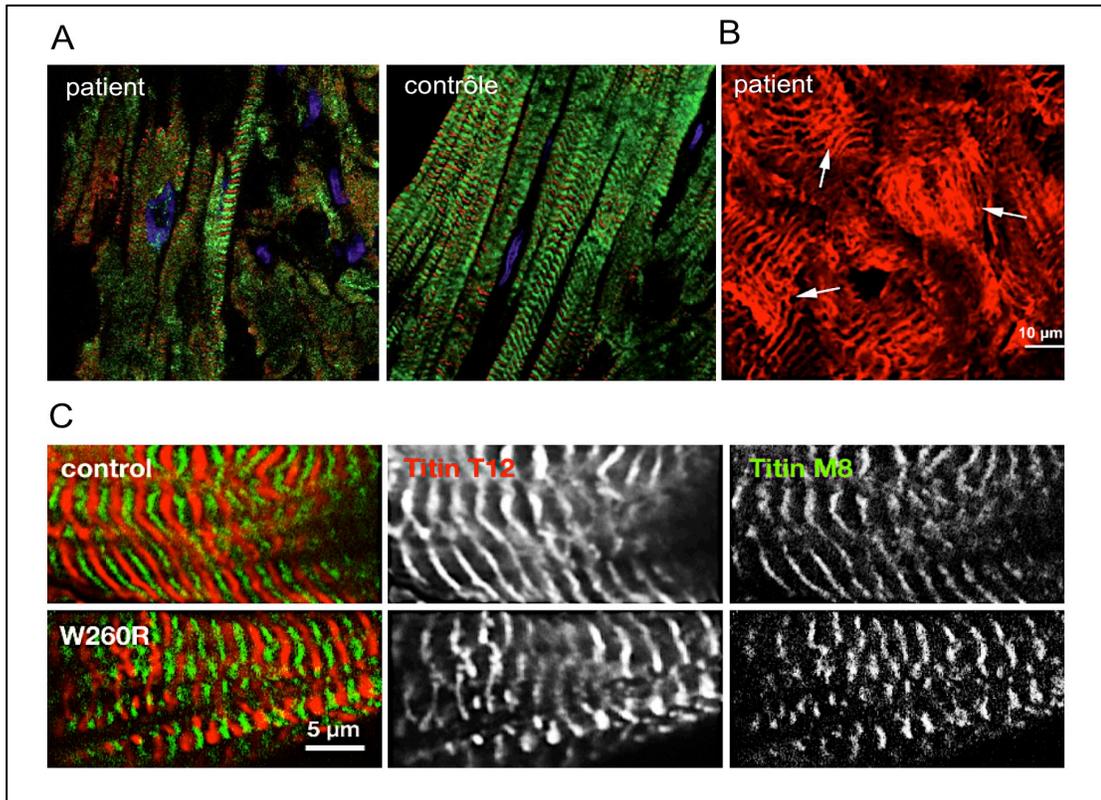
L'ensemble de ces résultats suggère donc que la mutation de **W260R** entraîne une **perte de fonction complète de TK**, déstabilisant sa structure, annulant son activité enzymatique et abrogeant complètement la voie de signalisation commune aux lignes M et Z via Nbr1 impliqué dans le turnover des protéines et l'autophagie. Par conséquent, cette absence de domaine kinase fonctionnel, jamais décrite auparavant chez l'humain, révèle une nouvelle voie et mécanisme physiopathologique dans les titinopathies.

### 2.3.3. Analyses protéiques

#### 2.3.3.1. Immunofluorescence in-situ

Outre les lésions de type minicore, les précédentes analyses ne montraient pas d'anomalie majeure de la structure du tissu musculaire chez la patiente, à l'exception d'agrégats protéiques de la strie Z. Conséquemment des analyses par **immunofluorescence** (IF) ont été réalisées, d'abord, contre les extrémités (ligne Z et bande M) de TTN pour une observation générale de l'intégration de la protéine dans les tissus musculaires ; puis contre les partenaires protéiques connus de ces régions pour une idée plus précise des conséquences fonctionnelles des mutations identifiées.

L'immunofluorescence dirigée contre un épitope de la **strie Z** de TTN d'une part, ainsi que de la **téléthonine** (ligand spécifique de la ligne Z de TTN et substrat de TK) d'autre part, a révélé une colocalisation du marquage avec une **striation préservée**. Un résultat similaire a été obtenu avec des anticorps dirigés contre des épitopes de la bande M de TTN et la myoméline, avec cependant quelques interruptions occasionnelles du marquage. Ceci est compatible avec une expression et une intégration sarcomérique préservées de la TTN de taille complète, a priori codée par l'allèle paternel, et incluant la mutation TK. En revanche, le marquage de l' $\alpha$ -actinine, protéine reliant les disques Z, a montré beaucoup de zones anormales, similaires aux **bavures de strie Z** observées dans les cas de myopathies myofibrillaires, qui peuvent être assimilées aux zones électron-dense observées en EM. Ces zones présentaient aussi un marquage positif pour différentes protéines de liaison à l' $\alpha$ -actinine incluant la myotiline, la filamine-C et ZASP/LDB3 (**Fig. 29**).



**Figure 29 : Analyses de TTN par immunofluorescence (IF)**

Le marquage par IF des extrémités de TTN (ligne Z, vert et bande M, rouge) sur cryosections longitudinales de myocarde du ventricule gauche de la patiente (A, gauche) et d'un contrôle (A, droite) montrent chez la patiente des zones structurales préservées jouxtant des aires totalement désorganisées. Cette désorganisation est également observable au niveau de la strie Z sous forme de bavures du marquage de l' $\alpha$ -actinine (patiente, B). Une striation similaire à celle du contrôle est tout de même observée dans une grande proportion du tissu de la patiente (ligne Z en rouge et bande M en vert, C), ce qui indique que la TTN de longueur complète, a priori codée par l'allèle paternel et incluant la mutation TK, est bien exprimée et intégrée dans les sarcomères.

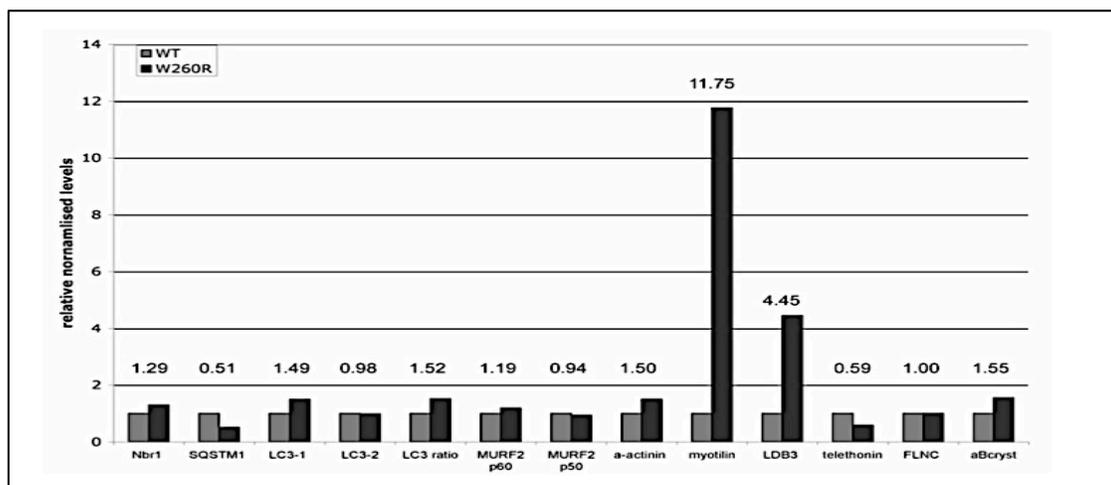
### 2.3.3.2. Western blot

Du fait de sa taille, les analyses électrophorétiques de la TTN issue d'extraits protéiques des biopsies des patients sont extrêmement complexes et ne donnent pas de résultats totalement concluants. Néanmoins, en collaboration avec l'équipe du **Dr Bjarne Udd** en Finlande (ayant une grande expertise dans les WB TTN), nous avons tenté de quantifier l'expression de TTN dans les échantillons cardiaques de la patiente et également de confirmer l'expression protéique de l'allèle maternel par WB. Par comparaison avec un contrôle sain ainsi qu'un contrôle TMD, les analyses ont montré une possible réduction du niveau globale d'expression de TTN chez la patiente ; en revanche aucune bande de taille correspondante à celle attendue dans le cas d'une synthèse de l'allèle maternel

n'a été révélée. Ceci pourrait être dû à une dégradation de l'ARNm par nonsense-mediated decay (NMD) ou bien aux limitations techniques du WB de la titine.

Concernant les expressions d'autres protéines impliquées dans la physiologie générale du muscle, une augmentation modérée de l' $\alpha$ -actinine - constituant essentiel du cytosquelette des cellules eucaryotes, ainsi que des fibres musculaires - (facteur de 1,4), et de la protéine de choc thermique  $\alpha$ - $\beta$  cristalline (facteur 1,5), drastique de ZASP/LDB3 (protéine du disque Z dont les mutations sont reliées à une désintégration des myofibrilles, facteur de 4,5) et de la myotiline (facteur de 11,8) ainsi qu'une légère diminution de la téléthonine (facteur de 0,6) ont été observées (**Fig. 30**). Des analyses complémentaires des protéines interagissant avec TK (de tailles normales donc analysables) ont révélé des changements modérés dans les niveaux de Nbr1 (augmentation selon un facteur de 1,3) et de son ligand SQSTM1 (diminution d'un facteur de 0,5), ainsi que de LC3-1 (augmentation selon un facteur de 1,5). Ces 3 protéines sont impliquées dans l'autophagie, confirmant ainsi l'implication physiopathologique de cette voie de dégradation protéique dans la mutation W260R. Enfin, aucun changement n'a pu être démontré dans le niveau de MURF2. Il est à noter que ces analyses ne peuvent être faites que sur un seul patient et par conséquent nous ne pouvons pas réaliser d'analyses statistiques précises et obtenir des barres d'erreurs.

L'ensemble de ces résultats suggèrent un effet significatif du variant TK-W260R dans le **processus autophagique** impliqué dans le contrôle qualité des protéines, ainsi que dans l'**homéostasie** des protéines sarcomériques associées au disque Z. Ceci confirme l'implication, inconnue auparavant de la titine dans cette voie et un nouveau mécanisme physiopathologique dans les titinopathies.



**Figure 30 : Quantification relative de l'expression de différentes protéines musculaires dans le tissu cardiaque de la patiente (W260R) et un échantillon contrôle (WT).**

En conclusion, l'étude présentée dans cette deuxième partie de ma thèse m'a conduit à :

**Identifier et caractériser phénotypiquement et génétiquement une nouvelle maladie congénitale** due à deux nouvelles mutations de TTN.

**Confirmer l'importance des mutations hétérozygotes composites de la TTN**, associant une mutation tronquante et une mutation faux-sens, comme mécanisme causal d'un nombre croissant de maladies musculaires et cardiaques.

Identifier la **TTN comme un nouveau gène impliqué dans l'arthrogrypose et dans la LVNC**.

**Caractériser pour la première fois une absence complète de domaine kinase fonctionnel chez l'humain**, et ses conséquences au niveau de l'architecture du sarcomère et des interactions protéine-protéine.

Grâce à la collaboration avec l'équipe de M. Gautel, **identifier une nouvelle implication de la TTN dans l'homéostasie des protéines sarcomériques et l'autophagie**, qui représente un nouveau mécanisme physiopathologique dans les titinopathies.

### **3. PARTIE III : Etude d'une série de patients MmD sans atteinte cardiaque : des mutations d'un coactivateur transcriptionnel provoquent une nouvelle forme grave de myopathie congénitale.**

#### **3.1. Etude clinique**

##### **3.1.1. Constitution d'une série de patients: hétérogénéité phénotypique**

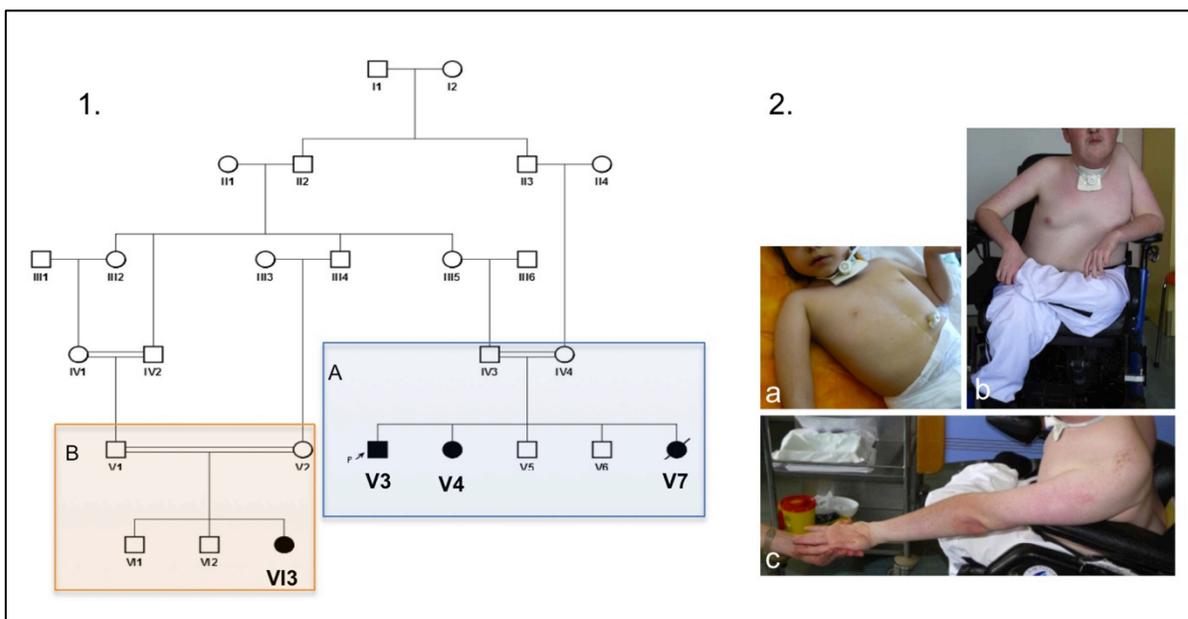
Suite à une collaboration internationale, nous avons collecté une série de 21 familles informatives dont la caractéristique commune était un diagnostic de myopathie congénitale de transmission autosomique récessive, avec présence de minicores sur la biopsie musculaire, pouvant être associée à une atteinte respiratoire et/ou une scoliose (phénotype de type MmD classique ou « SEPN1-like ». Des mutations dans les gènes connus associés aux myopathies à cores avaient été exclues dans tous les cas. Néanmoins, ces familles présentaient une hétérogénéité phénotypique assez marquée reflétant une probable hétérogénéité génétique. Par conséquent, nous avons classé les patients en sous-groupes phénotypiques homogènes afin de faciliter l'identification des mutations causales. Nous avons effectué des études de génotypage que j'ai analysé de façon globale, par sous-groupes phénotypiquement homogènes et famille par famille. Néanmoins, cette étude a révélé une hétérogénéité génétique plus importante que prévu et n'a pas permis d'identifier de gène commun aux familles de même sous-groupe. Par conséquent, j'ai choisi de centrer la dernière partie des résultats de ma thèse uniquement sur l'analyse de la famille pour laquelle nous avons été en mesure d'identifier un nouveau gène responsable d'une nouvelle forme de myopathie congénitale.

##### **3.1.2. La famille C : un nouveau phénotype clinique et histologique**

Une famille consanguine particulièrement informative de gens du voyage d'origine française incluse dans cette cohorte présentait un phénotype atypique, particulièrement grave, marqué par une faiblesse majeure (les enfants atteints n'ont jamais marché), une scoliose nécessitant une correction chirurgicale et une atteinte respiratoire sévère entraînant une ventilation assistée de jour comme de nuit. Les atteints présentaient aussi une hyperextensibilité des articulations ainsi qu'un phénotype cutané (**Fig. 31**), qui rappellent certaines caractéristiques des collagénopathies (myopathies dues à des mutations de gènes codant le collagène VI). La famille était constituée de 2 branches et incluait plusieurs boucles de consanguinité (**Fig. 31**). Une variabilité modérée dans la sévérité des symptômes a été observée parmi les 3 enfants atteints. En plus d'une myopathie à cores, le frère aîné (V3 sur le

pedigree) a présenté un retard de puberté absent chez sa sœur (V4). Un surpoids qui n'a pas pu être clairement lié au phénotype musculaire a aussi été noté chez les deux patients les plus âgés de la famille (V3, V4). Une jeune sœur également atteinte (V7), décédée prématurément à l'âge de 2 ans à cause d'une insuffisance respiratoire, présentait les mêmes signes et sévérité que son frère aîné.

Les analyses histologiques de muscle squelettique chez les patients V3, V4 et V13 ont montré, en plus de multiples minicores, des anomalies structurales particulières des fibres musculaires de type cap, des agrégats protéiques majoritairement localisés à la périphérie des fibres et adoptant une conformation en "casquette".



**Figure 31 : Phénotype clinique de la famille C**

Arbre généalogique présentant les deux branches (A et B) de la famille (1). Une assistance respiratoire par trachéotomie a été mise en place dès le plus jeune âge chez les patients V3 et V13 (Photo clinique, 2.b et 2.a respectivement). Phénotype cutané et hyperlaxité observée chez le patient V3 (2.c).

## 3.2. Etude génétique

### 3.2.1. Etat des lieux au début du projet

Une implication de la plupart des gènes connus provoquant ce type de pathologie musculaire, notamment *RYR1* et *SEPN1*, avait été exclue chez tous les patients de la cohorte.

### 3.2.2. Génomage et analyse de liaison

Les 15 familles les plus informatives incluant 12 familles consanguines ont fait l'objet d'un screening du génome avec un panel d'environ 450 marqueurs microsatellites et/ou près de 262 000 SNPs (Affymetrix 250 K, Nsp array). J'ai réalisé une analyse combinée de ces résultats suivant une approche mixte associant clonage positionnel et recherche de gènes candidats fonctionnels permettant d'associer les avantages de chaque technique et d'obtenir un gain d'informativité substantiel pour l'analyse de liaison. Pour les familles consanguines, l'analyse génétique a suivi le principe de cartographie par homozygotie dont la puissance est spécialement appropriée pour l'étude de maladies récessives rares. Les résultats obtenus ont montrés de multiples loci contenant de nombreux gènes dans l'ensemble des familles, sans toutefois révéler de régions candidates plus probables de contenir le gène muté responsable du phénotype clinique dans l'ensemble de la série ou pour chaque sous-groupe phénotypiquement homogène.

#### 3.2.2.1. Identification de 2 loci dans la famille C

Initialement nous n'avions connaissance que de la branche A de la famille C, nous avons donc effectué le génotypage du génome avec 262 000 SNPs sur tous les ADNs de cette branche, à l'exception de la jeune sœur décédée pour laquelle nous n'avions pas d'échantillon. Suite à des calculs de **lod score** réalisés avec le logiciel **Merlin** (Abecasis, Cherny et al. 2002), j'ai identifié dans cette famille 2 régions sur les chromosomes **15q22** et **22q11**, **homozygotes par descendance** chez les seuls atteints, qui étaient statistiquement liées au phénotype observé chez les patients. Pour ces deux loci, les parents étaient porteurs hétérozygotes d'un haplotype identique issu d'un ancêtre commun.

#### 3.2.2.2. Second génotypage et cartographie fine : confirmation d'un locus

Nous avons ensuite appris l'existence dans cette famille d'une jeune cousine issue d'une seconde branche de la famille (branche B) et présentant un phénotype myopathique à début précoce pratiquement identique avec celui décrit chez les patients V3 et V4 (**Fig. 31**). Nous avons vérifié que les défauts structurels observés sur les biopsies de ces différents patients étaient aussi similaires, ce qui laissait penser qu'ils étaient porteurs de la même anomalie génétique. Par conséquent, nous avons décidé de réaliser un génotypage Illumina (human panel V, 6 056 SNPs) sur les échantillons obtenus à

partir de cette nouvelle patiente et de ses parents et frères et sœurs non atteints. Ce second génotypage ayant pour objectif d'analyser spécifiquement dans les membres de la branche B les deux loci, 15q22 et 22q11, précédemment mis à jour dans la branche A, la résolution de ce panel de marqueurs Illumina, bien qu'inférieure à celle du panel Affymetrix utilisé lors du premier génotypage, était donc amplement suffisante. L'analyse comparée des deux génotypes m'a permis d'exclure le locus du chromosome 22 pour lequel la patiente 3 présentait un haplotype hétérozygote. En revanche, j'ai pu réduire le locus du chromosome 15 à un locus commun (associé à un **LODscore de 4,26**) aux 3 enfants atteints s'étendant sur 3 Mb et incluant **32 gènes candidats**.

### 3.2.3. Séquençage des gènes candidats positionnels

Parmi ces 32 candidats positionnels, aucun ne présentait une expression ou une fonction spécifique au muscle strié. Par conséquent la priorisation concernant leur séquençage s'est faite sur des critères techniques prenant en compte la complexité et la taille de leurs séquences codantes. Suite au séquençage des 7 premiers gènes, j'ai identifié une mutation sur un site d'épissage d'un gène codant une protéine peu documentée dont la caractéristique principale était de posséder un domaine conformationnel de type pleckstrine. Cette protéine d'expression ubiquitaire n'avait aucune implication connue dans la physiologie musculaire et ne paraissait pas être un candidat évident par conséquent nous avons décidé de poursuivre l'analyse des autres gènes du locus en parallèle de l'étude des conséquences de ce nouveau variant. Etant donné la présence de lésions de type cap, qui avaient été préalablement associées à des mutations de **TPM2** (codant la  $\beta$ -tropomyosine), j'ai aussi exclu par séquençage des mutations de ce gène.

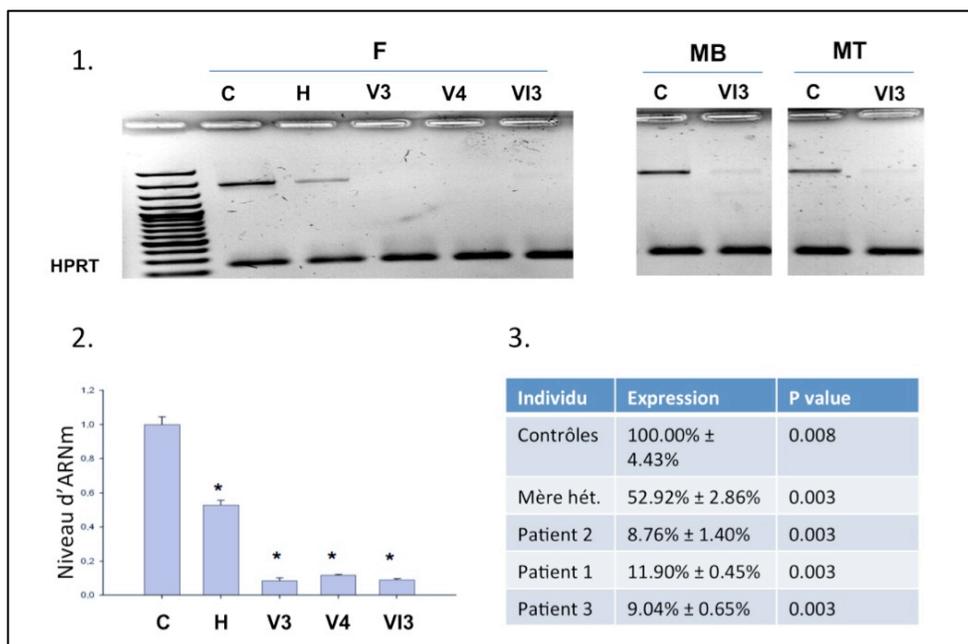
Finalement, le séquençage du 10<sup>ème</sup> candidat positionnel a révélé la présence d'une mutation non-sense dans un **coactivateur de transcription des récepteurs nucléaires**, là encore, très peu documenté et n'ayant **jamais été associé à une maladie humaine**. Ce variant, introduit dans la séquence protéique un PTC, **p.W297X**, prédisant la synthèse d'un polypeptide tronqué de 49% de sa taille. Ce changement coségrégait avec le phénotype dans les deux branches de la famille et n'a pas été retrouvé chez 120 contrôles. Par conséquent nous avons donc concentré nos analyses fonctionnelles sur cette nouvelle protéine.

### 3.3. Etude fonctionnelle des variants identifiés dans la famille C

Pour réaliser ces analyses nous avons obtenu des cultures de fibroblastes des 3 patients et de la mère de la branche A de la famille, ainsi que d'une culture de myoblastes et d'une biopsie musculaire de la jeune cousine (VI3).

#### 3.3.1. Analyse quantitative de l'expression des ARNm chez les patients

Nous avons alors cherché à caractériser les conséquences de la mutation stop sur l'expression des ARNm relativement à l'état haplotypique (hétérozygote ou homozygote) des tissus. Des études par RT-PCR quantitative ont montré une expression des ADNc obtenus à partir des ARNm du coactivateur transcriptionnel réduite de 52 % chez la mère hétérozygote, alors que les patients affichaient entre 8 et 12% d'expression résiduelle (**Fig. 32**).



**Figure 32 : Analyse de l'expression des ARNm du second gène muté dans la famille C**

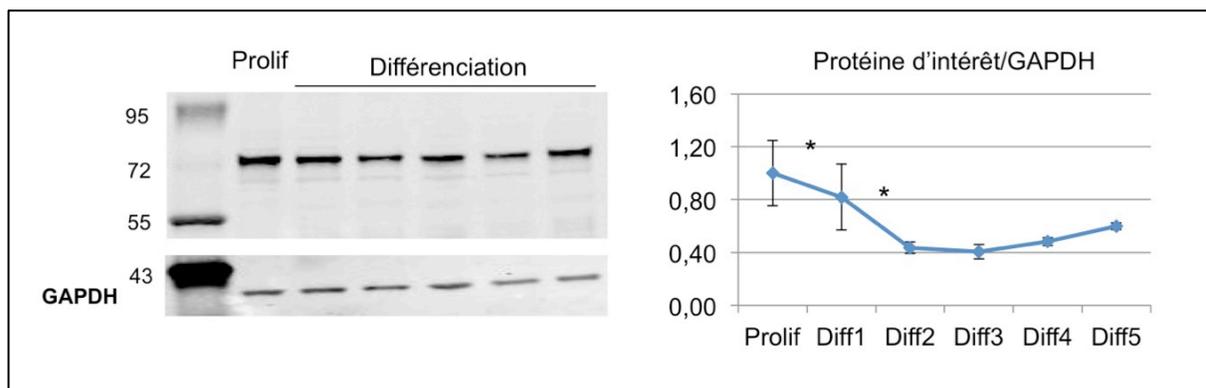
La migration électrophorétique sur gel des ADNc issus des ARNm extraits de culture de fibroblastes (F), révèle une bande spécifique de l'amplification du gène d'intérêt d'intensité moindre dans l'échantillon maternel hétérozygote (H) par rapport au pool de contrôles (C), et une expression résiduelle minimale dans les échantillons V3 et V4 et VI3 (1), en accord avec la quantification du niveau d'expression des ARNm mesurés (2) obtenue par PCR quantitative (3). HPRT a été utilisé comme gène de référence. Une bande d'intérêt d'intensité moindre est également observée avec les échantillons de myoblastes (MB) ou de myotubes (MT) (1).

Cette observation pourrait s'expliquer par le fait que l'expression d'une isoforme de taille réduite à 51% de la séquence protéique originelle du gène nouvellement identifié a de fortes chances d'être associée à un effet toxique dominant négatif et donc d'entraîner une prise en charge des messagers mutés par le système NMD. L'activation potentielle de ce système de protection de la cellule aurait alors pour conséquence leur dégradation, observable dans les tissus des patients.

### 3.3.2. Analyse de l'expression protéique par western-blot

Nous avons par la suite étudié l'expression protéique de ce nouveau gène et montré qu'elle suivait le profil d'expression des ARNm avec une déplétion complète en protéine dans les patients homozygotes et un niveau d'expression réduit de moitié dans les cellules de la mère hétérozygote.

Dans un second temps, nous avons étudié le profil d'expression protéique du coactivateur transcriptionnel au cours de la différenciation musculaire de myoblastes contrôles en myotubes. Ces résultats préliminaires semblent montrer une expression maximale dans les myoblastes en prolifération et une diminution de l'expression de la protéine d'intérêt lors de la différenciation myogénique (**Fig. 33**).

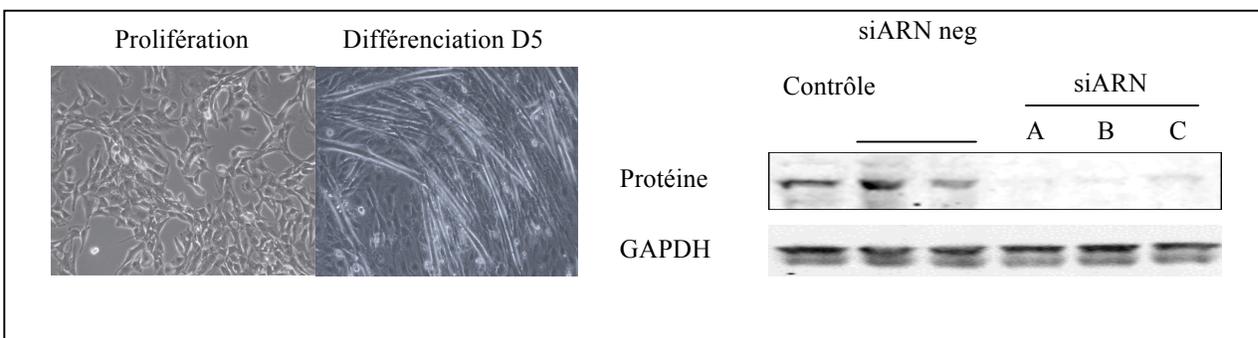


**Figure 33 : Analyse de l'expression protéique lors de la différenciation musculaire du gène identifié dans la famille C**

L'expérience a été réalisée sur des myoblastes C2C12 en état prolifératif et pendant les 5 premiers jours de leur différenciation en myotubes. La protéine d'expression ubiquitaire GAPDH a servi de contrôle de référence pour la normalisation de l'expression protéique de la protéine mutée chez nos patients.

### 3.3.3. Impact sur les facteurs myogéniques

Enfin, les dernières expériences menées sur ce nouveau gène ont consisté en des analyses par immunofluorescence de l'expression des facteurs myogéniques (MRFs, myogenic regulatory factors) MyoD, myogénine et de la chaîne lourde de la myosine dans des conditions physiologiques et en absence d'expression de la protéine d'intérêt. En traitant des cellules de type C2C12 (lignée musculaire de souris) avec des ARN interférences (siRNA) spécifiques du messenger d'intérêt, ce qui empêche leur prise en charge par le système de traduction de la cellule, nous avons été en mesure d'éteindre quasiment complètement l'expression du coactivateur transcriptionnel (réduction de près de 97% d'expression) mimant ainsi le phénotype des patients (**Fig. 34**). L'analyse de ces résultats suggère une diminution d'expression des facteurs myogéniques myogénine et MF20 (introduire) au cours de la différenciation dans les cellules knocked-down pour le nouveau gène. Ces résultats sont préliminaires et doivent être répétés, en effet nous n'avons pas pu exclure un effet possible de la transfection sur l'expression de la myogénine au regard de la légère diminution observée au 3<sup>ème</sup> jour sur la population scramble (contrôle de transfection). Cet effet n'est en revanche pas observable sur les myoblastes exprimant MF20, les populations scramble et contrôles présentant un profil similaire.



#### Figure 34 : Validation du modèle *in vitro* de diminution d'expression de la protéine d'intérêt

Des myoblastes C2C12 contrôles ont été transfectés avec des siRNA scramble (contrôle de transfection) et spécifiques du gène muté (A, B, C) dans la famille C, puis cultivés en prolifération pour permettre leur synchronisation. Des marquages de la myogénine et de la myosine (chaîne lourde, MF20) ont ensuite été réalisés par IF pendant respectivement 3 et 4 jours post initiation de la différenciation.

L'analyse de ces résultats, va pour l'instant dans le sens d'une diminution d'expression des facteurs myogéniques au cours de la différenciation corrélée à l'absence d'expression de la protéine identifiée chez les patients dans les cellules traitées avec les siRNA.

L'ensemble de ces résultats semble indiquer une implication dans le processus de différenciation musculaire de ce coactivateur transcriptionnel nouvellement identifié comme muté chez les patients de la famille C. Très peu d'informations sont disponibles sur cette protéine, il a néanmoins été démontré qu'elle forme un complexe protéique avec 3 autres partenaires. Des études complémentaires sont en cours.





## **DISCUSSION**



L'objectif de cette thèse consistait en l'identification et la caractérisation des bases moléculaires de nouvelles formes de myopathies congénitales. Elle devait permettre une meilleure compréhension des mécanismes physiologiques sous-jacentes à ces entités médicales et ainsi apporter des bénéfices directs aux patients mais surtout participer à l'approfondissement de notre compréhension de la physiologie et la physiopathologie du muscle strié.

Ce travail s'est fait en explorant deux séries de patients, l'une présentant un phénotype associant une myopathie congénitale et une cardiomyopathie ; et l'autre d'atteinte strictement restreinte aux muscles squelettiques. Nous avons ainsi mis en évidence une implication du gène *TTN* dans une variété de phénotypes beaucoup plus étendue et avec une incidence probablement plus élevée que celles estimées jusqu'à ce jour. Nous avons relié pour la première fois la *TTN* à un défaut cardiaque fréquent dans les cas pédiatriques et dont l'association avec d'autres symptômes nous a permis de décrire une nouvelle entité médicale. Et finalement, nous avons montré l'implication d'une nouvelle protéine dans la mise en place du tissu musculaire, mutée dans une famille au phénotype particulier identifiant une nouvelle maladie du muscle chez l'homme.

### **Première conclusion : L'implication des mutations *TTN* hétérozygotes composites dans les pathologies musculaires et cardiaques est probablement sous-estimée**

L'étude génétique d'une série de cas avec atteinte cardiaque regroupant 24 familles qui présentaient un phénotype clinique et morphologique particulier, marqué par une atteinte des muscles squelettiques de type myopathie congénitale (CM) et cardiaque (majoritairement des cardiomyopathies dilatées), a révélé des mutations du gène de la *TTN* dans 5 d'entre elles (20% des patients). Une de ces familles présentait une mutation précédemment identifiée dans les myopathies tibiales tardives (TMD), pour les 4 autres familles les mutations n'avaient jamais été décrites.

Une caractéristique commune chez tous ces patients concerne le type de mutation dont ils font l'objet. En effet, nous avons identifié chez les 4 familles (5 patients) des mutations associées à un mode de transmission récessif qui induisent l'apparition directe ou via un décalage du cadre de lecture d'un codon stop prématuré (PTC) et par conséquent une troncation de la *TTN* en C-terminale. L'absence de phénotype chez les parents, porteurs hétérozygotes, associée au mode de transmission

récessif et le début précoce de la maladie sont originaux. A ce jour, exception faite de deux mutations tronquantes homozygotes retrouvées dans la bande M par notre équipe (Carmignac, Salih et al. 2007) et d'un évènement d'insertion/délétion d'origine finnoise qui induit un phénotype pédiatrique lorsqu'il est hérité conjointement sur les deux allèles et un phénotype tardif à l'état hétérozygote (Udd, Vihola et al. 2005), la plupart des mutations liées à un phénotype musculaire identifiées dans le gène de la titine ont été retrouvées dans des maladies à transmission dominante et à début plus tardif.

Fait intéressant, nous décrivons ici des cas cliniques associant une atteinte des muscles squelettiques et cardiaque, combinaison rare en proportion des phénotypes associés à des mutations *TTN* précédemment publiés. En effet, mises à part les deux mutations récessives de la bande M (Carmignac, Salih et al. 2007), la centaine de mutations décrites (pour la plupart très récemment) dans la littérature touchent de façon isolée soit le cœur, soit le muscle squelettique, les mutations propres à chacun de ces types musculaires se répartissant selon un schéma spécifique. Etonnamment, bien que la titine contienne des exons d'expression spécifique au cœur (N2B) ou au muscle squelettique, seules 3 des 84 mutations conduisant à un phénotype exclusivement cardiaque sont situées dans l'exon N2B, les autres étant localisées dans des exons communs aux isoformes cardiaques et squelettiques ; de plus les mutations de type cardiaque se répartissent uniformément sur le gène, alors que la quasi totalité de celles induisant un phénotype purement musculaire sont situées uniquement dans certains exons (Mex1, 5 et 6) qui codent la bande M de la titine (Mex1 à 6 ou 358 à 363, C-terminale). Ces données peuvent suggérer que ce domaine C-terminal est capital pour le fonctionnement du muscle squelettique mais pas du muscle cardiaque, cependant il faut prendre en compte le biais d'analyse dont *TTN* a fait l'objet pendant de nombreuses années. En effet, l'analyse de *TTN* pour les diagnostics génétiques de patients musculaires s'est presque exclusivement concentrée sur la bande M, jusqu'à l'avènement des technologies de séquençage de nouvelle génération ; ainsi le premier variant de *TTN* lié à une myopathie sans atteinte cardiaque a été récemment identifié par NGS chez un patient HMERF et localisé dans la bande A (Ohlsson, Hedberg et al. 2012). Il est donc possible que les technologies de séquençage récentes mènent dorénavant à l'identification de mutations d'expression musculaire aussi bien dans les exons codant pour la bande M que pour les autres régions de la *TTN*.

Les 7 mutations que nous avons identifiées consistent en une délétion homozygotes dans la famille consanguine d'origine turque (c.106272-106273delTG [p.Ser35424Serfs\*11]) ; l'association de deux mutations stop dans une patiente au phénotype squelettique et cardiaque modéré (c.102388C>T [p.Arg34150\*] et c.105697C>T [p.Gln35257\*]) ; finalement l'association chez les deux patients au phénotype cardiaque le plus grave, ayant subi une transplantation durant l'enfance,

d'une mutation faux sens associée soit à une délétion causant un décalage du cadre de lecture (c.66560G>A [p.V22187E] et c.101922delT [p.Thr33974Thrfs\*9]) soit à un variant d'épissage (c.T94734C [p.31504 ou p.W260R] et c.G9388+1C) (**Fig. 35**). Du fait de l'absence de phénotype chez tous les parents ou apparentés porteurs hétérozygotes, ces 7 mutations peuvent toutes être considérées comme provoquant un phénotype purement récessif. Elles sont localisées dans les exons 38, 316, 358 (Mex1) et 359 (Mex2) et au moins une mutation de chaque patient se retrouve dans la bande M, avec un effet qui prédit la troncation et perte partielle de cette région, compatible avec la composante musculaire du phénotype des patients. En revanche, la localisation des mutations identifiées explique plus difficilement le phénotype cardiaque et les mécanismes physiopathologiques menant aux cardiopathies observées chez les patients restent encore à déterminer. Néanmoins, leur nature est en accord avec les données précédemment publiées puisque toutes les mutations provoquant la troncation de la TTN sont associées à des phénotypes cardiaques (**Tableau, en annexe 3**).

Comme évoqué précédemment, dans nos familles les porteurs hétérozygotes des mutations sont asymptomatiques. Une publication récente dans le *New England Journal of Medicine* (**Herman, Lam et al. 2012**) décrivant 72 mutations tronquantes identifiées par NGS dans une cohorte de cardiomyopathies dilatées proposait une transmission autosomique dominante, ce qui nous a incité à vérifier les résultats des investigations cliniques des porteurs sains dans les familles de notre étude. La confirmation d'absence totale de signes cardiaques chez les apparentés après des études cardiologiques complètes soulève deux réflexions : la remise en question de l'hypothèse d'un effet dominant négatif expliquant le phénotype des patients émise par *Herman et al.*, et l'existence probable d'un mécanisme cellulaire protégeant les porteurs de mutations tronquantes à l'état hétérozygote les empêchant de développer la maladie observée chez les homozygotes.

Selon les auteurs de l'article du NEJM ayant décrit la cohorte cardiomyopathique, une toxicité des TTN tronquées sur les formes sauvages expliquerait le phénotype des patients ; or cette réflexion supposerait une pathologie associée à une transmission de nature autosomique dominante qui n'est pas confirmée dans la plupart des familles incluses dans leur étude. Aussi, une partie des pedigrees présentés dans l'article ne montre pas de co-ségrégation nette de la mutation avec le phénotype, ce qui a été attribué à une pénétrance incomplète. De plus, les auteurs ont volontairement écarté de leur analyse tous les variants faux-sens entraînant des substitutions dans la séquence protéique de la TTN sans troncation du polypeptide, dû au grand nombre de changements de ce type identifié par patient. Ces remarques et l'identification dans nos familles de variants faux-sens et tronquant en association hétérozygote composite, ainsi que l'assurance de la nature asymptomatique des hétérozygotes nous

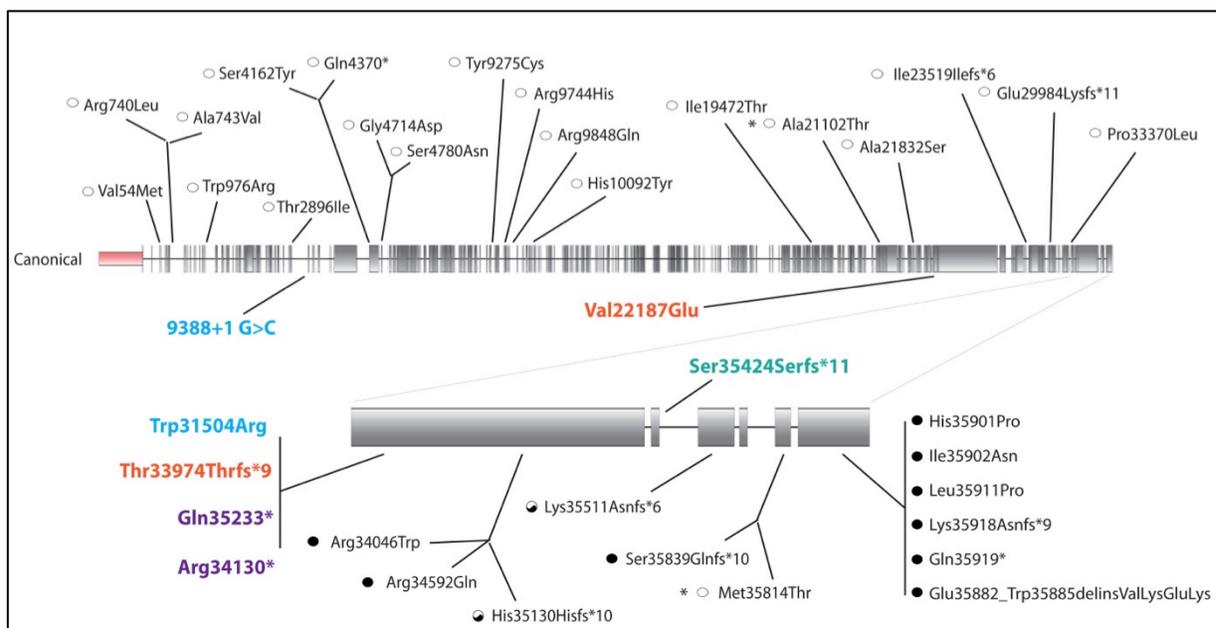
laissent sceptiques quant aux conclusions de cette étude. Nous proposons que l'hétérozygotie composite d'une mutation tronquante et une faux-sens est un nouveau paradigme qui peut expliquer un bon nombre de cas de titinopathie. Par conséquent, une deuxième mutation devrait être systématiquement recherchée après identification d'une mutation tronquante par NGS, et cette technique ne dispense pas d'une analyse familiale approfondie. Une analyse de ségrégation incluant les variants substitutifs identifiés dans la cohorte de cardiomyopathies publiée permettrait de tester l'hypothèse d'un phénotype lié à des mutations hétérozygotes de *TTN* et de déterminer rigoureusement le mode de transmission de la maladie dans les différentes familles. Bien évidemment on ne peut rigoureusement omettre également l'hypothèse de digénisme où des mutations dans des gènes différents expliqueraient la pathologie observée chez les patients.

Quant au phénomène de protection des porteurs hétérozygotes de mutations tronquantes, il pourrait être assuré par un mécanisme de saut d'exon similaire à celui testé dans certains essais thérapeutiques de la dystrophie musculaire de Duchenne. Cette pathologie est provoquée par des mutations de la dystrophine, protéine essentielle à la physiologie musculaire, localisée sous la membrane cellulaire de toutes les fibres musculaires (lisses, striés, ou cardiaques) où elle assure le maintien de l'architecture cellulaire. Des stratégies ciblées de saut d'exon permettent chez la souris de supprimer de la séquence protéique les exons porteurs de mutations et de synthétiser des quasi-dystrophines qui restent fonctionnelles. Cette hypothèse d'un saut d'exon pourrait être testée par analyse des transcrits *TTN* des porteurs asymptomatiques de nos familles, néanmoins l'observation de ce mécanisme est rare dans la nature et a peu de probabilités d'être impliqué dans la protection de nos apparentés sains.

Une autre explication de l'absence de phénotype chez les hétérozygotes pourrait alors être la dégradation post-transcriptionnelle des ARNm mutés qui, en empêchant leur prise en charge par la machinerie cellulaire de traduction, permettrait de réduire de façon drastique la synthèse de la protéine mutée, favorisant l'incorporation de la titine normale dans les sarcomères. Les individus hétérozygotes seraient ainsi protégés contre l'expression de la maladie. L'absence ou la diminution d'expression des ARNm dans les cellules des hétérozygotes est mesurable par des analyses quantitatives par qPCR. Le mécanisme pouvant mener à cette diminution est qualifié de NMD (nonsense-mediated mRNA decay) et se déroule en deux étapes successives, la première permet la reconnaissance de complexes protéiques EJC (exon junction complex) rémanents, positionnés pendant l'épissage à une vingtaine de nucléotides en amont des jonctions exoniques. En absence de PTC ces complexes sont déplacés par les ribosomes ; la présence d'un PTC empêche cette délocalisation et sert de signal à l'initiation du NMD.

La formation d'un complexe de surveillance intervient alors en seconde étape, l'action finale aboutissant à la dégradation des ARNm mutés (Maquat 2004). Le NMD peut être inhibé par traitement aux antibiotiques, ce qui permet par conséquent de valider son implication dans l'observation d'une absence de messagers. L'un des critères nécessaire à l'activation du NMD est une localisation des PTC à plus de 55 nucléotides en amont d'une jonction exonique, permettant de s'affranchir des restrictions physiques liées à l'encombrement stérique des EJC et des ribosomes (Nagy and Maquat 1998). Les 5 mutations générant des PTCs que nous avons mises en évidence dans nos familles respectent ces conditions et sont donc compatibles avec une protection des hétérozygotes par NMD.

Ces deux remarques concernant un phénotype pathologique provoqué par des mutations hétérozygotes composites de *TTN* et une protection des porteurs hétérozygotes sont importantes de par leur impact sur le diagnostic moléculaire des familles. En effet, une assurance certaine du mode de transmission et de l'absence de signes pathologiques chez les porteurs hétérozygotes peut impacter sur les décisions médicales consécutives à l'établissement d'un diagnostic génétique.



**Figure 35 : Mutations récessives de *TTN* identifiées pendant la thèse**

Les mutations *TTN* trouvées chez nos patients sont indiquées par des couleurs, les mutations associées en conformation hétérozygotes composites sont indiquées par des couleurs similaires. Les mutations écrites en noir représentent les mutations décrites dans la littérature avant l'avènement des techniques de séquençage haut débit. Les phénotypes associés aux mutations sont différenciés par des cercles blanc (cardiaque), noir (musculaire) et bicolore (mixte). La portion C-terminale (bande M) de *TTN* fait l'objet d'un grossissement, les exons sont représentés par des rectangles gris et les introns par une ligne continue.



## **Seconde conclusion : La liaison de la TTN à une variété de défauts cardiaques est plus importante que celle actuellement estimée**

Une autre caractéristique commune chez tous les patients avec mutations de la *TTN* identifiées pendant ce travail concerne l'expression clinique des variants. A l'exception des deux frères issus d'une famille consanguine d'origine turque, les 3 autres patients étaient sporadiques et présentaient une atteinte cardiaque précoce et pour deux d'entre eux, extrêmement sévère, ayant nécessité une transplantation cardiaque à 5 et 15 ans. En contraste, l'atteinte musculaire était beaucoup moins grave, peu voire pas évolutive et a permis le maintien de la marche.

La communication inter-auriculaire (CIA) (ou défaut du septum, ASD : *atrial septal defect*) est une malformation cardiaque congénitale fréquente, qui correspond à la persistance d'une ouverture d'origine embryonnaire dans le septum inter-auriculaire (cloison qui sépare les deux oreillettes cardiaques), normalement hermétique après la naissance. Cette communication anormale est responsable de l'apparition d'un « shunt » (dérivation sanguine) gauche-droit: une partie du sang revenant des poumons dans l'oreillette gauche regagne directement les poumons (via la CIA et l'oreillette droite) en court-circuitant la grande circulation. Les CIA représentent près de dix pour cent des cardiopathies congénitales de l'enfant, rapprochant leur incidence à près de 0,70 pour mille naissances.

La titine, de par sa structure et sa fonction est une protéine essentielle au sarcomère, elle participe à son assemblage pendant la myogenèse et contrôle la tension passive et l'élasticité du muscle. De plus, elle participe à des voies de signalisation par ses interactions protéiques dans la strie Z, la bande I et la bande M via son domaine kinase TK (Trinick 1994; Labeit and Kolmerer 1995; Gregorio, Trombitas et al. 1998; Lange, Xiang et al. 2005). Il est donc aisé d'admettre que ses mutations puissent être liées au phénotype musculaire chez nos patients, l'observation de zones focales de désorganisation sarcomérique (lésions minicores) dans leurs fibres musculaires, révélatrices d'une fragilité structurale de la bande M, s'expliquant alors par l'absence de la partie C-terminale de la titine due aux mutations tronquantes identifiées. En revanche, nous ne pouvons pas assurer avec certitude que la mise en place de ces lésions chez nos patients soit effectivement provoquée par les différentes mutations de *TTN*. Etonnamment, la suppression homozygote de la quasi intégralité de la bande M (à partir de Mex2) que nous avons identifiée dans la famille consanguine turque est associée à un phénotype cardiaque modéré chez un des deux patients et même absent chez son frère à l'âge de 16 ans, alors qu'une mutation tronquante en aval (Mex3) a été décrite chez des patients présentant une

cardiomyopathie dilatée très sévère ayant provoqué leur décès avant l'âge adulte (Carmignac, Salih et al. 2007). Cette observation témoigne d'une variabilité dans l'expression de la sévérité des phénotypes associés aux mutations de *TTN*. Plusieurs explications peuvent être apportées pour expliquer cette variabilité, notamment le fait que ces mutations concernent la bande M de titine qui est une région impliquée dans de nombreux complexes protéiques. Chaque troncation de cette portion de la protéine peut avoir des conséquences différentes au niveau conformationnel et de la structure tertiaire, entraînant par la même des modifications dans l'accessibilité des partenaires protéiques de la titine et leur capacité de liaison et de formation des différents complexes. Il serait donc intéressant d'analyser comparativement par immunofluorescence l'expression des partenaires connus de titine chez différents patients présentant des troncations plus ou moins importantes.

Finalement, ces nouvelles mutations identifiées dans notre cohorte et leur taux élevé (20% des patients mutés alors que pour 22 des 25 nous n'avons étudié que la bande M, correspondant à 7.5 % du gène) illustrent le rôle clé de *TTN* dans les cardiomyopathies associées à une atteinte musculaire congénitale, soulignent la probable sous-estimation de leur implication dans les défauts cardiaques primaires et justifient la nécessité de considérer une étude systématique de *TTN* pour tous les patients présentant une atteinte musculaire ou cardiaque sans diagnostic génétique (*cf article soumis, présenté en annexe*).

### **Troisième conclusion : Un KO constitutif de TK chez l'homme mène à la description d'une nouvelle entité médicale associant une atteinte cardiaque et squelettique**

Le défaut cardiaque diagnostiqué chez une de nos patientes consistait en un syndrome de non compaction du ventricule gauche (LVNC). Cette cardiopathie congénitale doit son origine à une réminiscence de caractéristiques anatomiques du myocarde fœtal. Durant la période intra-utérine, le tissu cardiaque immature est souple et d'aspect spongieux. Rapidement après la naissance il se rigidifie et devient capable de supporter les continuels étirements des systoles et diastoles auquel il est soumis lors des contractions cardiaques. L'altération de la structure du myocarde dans les LVNC est due à une embryogénèse incomplète dans laquelle le passage d'un état à l'autre de compaction du tissu cardiaque n'est pas intégralement observé. Chez la patiente, la LVNC était associée à une arthrogrypose congénitale multiple, caractérisée par de nombreuses rétractions des articulations. Ni la LVNC, ni l'arthrogrypose n'avaient été indépendamment associées à des mutations de *TTN*. De plus leur association conjointe chez un même patient n'a jamais été rapportée.

Chez cette patiente les variations de séquence de *TTN* consistent en l'association d'une substitution du domaine sérine-thréonine kinase (TK) localisé dans la portion C-terminale de *TTN* au niveau de la bande M et d'une mutation du site 5' donneur d'épissage de l'intron 39. La mutation du domaine kinase, d'origine paternelle, entraîne un changement drastique de conformation tertiaire du site catalytique de l'enzyme qui a pour conséquence de supprimer les liaisons avec les ligands connus du domaine (nbr1 et la queue autoh inhibitrice de TK) mais également d'induire une perte complète de l'activité enzymatique. La seconde mutation, d'origine maternelle, prédit la troncation de *TTN* après l'exon 38 sur les 363 contenus dans la séquence codante ce qui correspond à l'expression d'une protéine d'une taille de 2 950 résidus sur les 34 350 qui pourraient être théoriquement codés par la séquence complète du gène.

Les isoformes observées sur gel à partir des ADNc devraient correspondre à l'isoforme paternelle pour celle correspondant à la bande de plus grande taille et maternelle pour les deux autres bandes (**Fig.13**). La seule isoforme d'ADNc retrouvée dans les deux types de tissu de la patiente coderait une *TTN* tronquée après 2 950 résidus protéiques sur un peu plus de 33 000 pour les isoformes sauvages les plus grandes répertoriées dans la littérature. Les analyses électrophorétiques de la *TTN* sont extrêmement complexes et nous ne sommes pas parvenus à montrer l'expression dans les tissus de la patiente de cette isoforme constituée des 8,6% de résidus les plus N-terminaux de *TTN*. Au

vu de l'importance de la TTN dans la physiologie du muscle strié, l'expression d'une telle protéine tronquée de près de 92% et son intégration dans les sarcomères entrainerait probablement une toxicité délétère pour la myofibrillogénèse. Il y a donc de fortes probabilités pour qu'un système de surveillance de la cellule, de type NMD (nonsense mediated decay) comme mentionné précédemment, soit mis en place et conduise à la dégradation des ARNm provenant de l'allèle maternel afin d'empêcher une synthèse protéique associée à un effet dominant négatif.

Une mutation du domaine kinase a précédemment été décrite chez un patient présentant une myopathie héréditaire avec atteinte respiratoire précoce (également appelé myopathie d'Edström) (Lang, 2005). Chez ce patient, la mutation du domaine kinase n'abrogeait pas les interactions avec tous les partenaires connus mais en préservait certaines, le domaine kinase était en parti fonctionnel. L'implication de *TTN* dans ce groupe de myopathie a été récemment confirmée (Ohsson, 2012) avec l'identification chez plusieurs patients de mutations *TTN*, mais non localisées dans le domaine kinase. Chez tous les patients, la transmission est invariablement autosomique dominante et le phénotype respiratoire relativement homogène. Le déficit musculaire peut être distal, proximal ou mixte. En revanche, chez notre patiente le phénotype est très sévère, et peut s'expliquer par les conséquences des mutations portées par les deux allèles de *TTN* qui mènent à une probable absence de TK d'origine maternelle associée à une perte de fonction de TK d'origine paternelle. Notre patiente représentant le premier cas équivalent à un KO du domaine kinase de TTN chez l'homme. Etant donnée l'importance fonctionnelle de ce domaine dans la physiologie du muscle ainsi que la létalité observée chez les souris exprimant un KO conditionnel de TK, nous pensons que le phénotype très sévère de la patiente peut s'expliquer par cette absence de domaine kinase fonctionnel dans ses tissus. En conclusion, nous décrivons ici une nouvelle maladie associant une myopathie congénitale, une arthrogrypose et une LVNC que nous avons liées à des défauts de la titine menant à la première perte totale décrite de domaine kinase de TTN chez l'homme.

## **Quatrième conclusion : Une mutation dans une nouvelle protéine potentiellement impliquée dans la myogénèse est responsable d'une nouvelle maladie du muscle chez l'homme**

D'autre part, l'analyse de la seconde série de 21 familles, recrutées dans le cadre d'un projet d'identification de nouveaux gènes responsables de la forme classique de myopathie à multi-minicores (MmD), par analyse de liaison combinée à une étude de gènes candidats, a permis de conclure à une hétérogénéité génétique des patients. Néanmoins, ces familles partageaient des caractéristiques communes ; elles présentaient toutes une atteinte restreinte aux muscles squelettiques, et des similarités physiopathologiques illustrées par la présence de lésions cores au sein des fibres musculaires chez chacune d'entre elles.

Nous avons jusqu'à présent été en mesure d'identifier le gène responsable du phénotype observé chez une seule famille de la série, la famille consanguine C de gens du voyage. Cependant, les trois patients de cette famille montraient certains signes cliniques qui n'étaient pas tout à fait typiques de la forme classique de myopathie à multi-minicores, en particulier la présence dans les fibres musculaires d'agrégats protéiques de type cap, et également d'un phénotype cutané associé à une hyperlaxité rapprochant ces patients des collagénopathies (myopathies avec mutations du collagène VI). Les résultats des analyses génétiques dans cette famille ont conduit à l'identification d'une mutation stop, entraînant la troncation d'une nouvelle protéine n'ayant jamais été impliquée dans une pathologie humaine auparavant et dont les résultats d'études fonctionnelles préliminaires font présager de son implication dans la myogénèse. Cette protéine est un coactivateur de la transcription des récepteurs nucléaires. Elle forme un complexe avec 3 autres protéines partenaires et semble interagir en tant que répresseur avec certaines voies de signalisation impliquées dans les processus de différenciation, de prolifération et d'apoptose cellulaire. Ainsi, nous avons défini une nouvelle forme de myopathie à début précoce chez l'homme que nous situons à l'interface des MmD et collagénopathies. Afin de déterminer les conséquences d'une absence d'expression du coactivateur transcriptionnel nouvellement identifié sur ses interactions avec différents partenaires, le recours à une puce d'expression protéomique paraît être la solution la plus adaptée, et permettrait de clarifier la nature des voies de signalisation dans lesquelles cette nouvelle protéine intervient.

Dans le but de déterminer si ces observations cliniques pouvaient constituer une nouvelle entité homogène nous avons alors recherché d'autres familles présentant des similarités phénotypiques. Au fur et à mesure du déroulement du projet, d'autres familles ont été recrutées et

nous sommes maintenant arrivés à collecter une série de patients avec un phénotype similaire incluant 32 familles et pour laquelle les analyses préliminaires par séquençage direct de ce nouveau gène candidat semblent indiquer la présence d'une mutation chez potentiellement deux autres familles.

Ces résultats constituent à la fois, l'identification d'un nouvel acteur de la physiologie musculaire chez l'homme, son implication dans une nouvelle maladie musculaire et sa liaison potentielle à plusieurs familles conduisant à l'émergence d'un nouveau sous groupe de MCs.

## Clôture et Perspectives

Au cours de cette thèse j'ai pu assister à l'évolution ahurissante des technologies d'analyse génétique. Durant les 4 années sur lesquelles s'est déroulée cette aventure scientifique, j'ai vu défiler les progrès techniques à un rythme impressionnant. Il a suffi de 3 années seulement pour passer de ce qui me paraissait être à l'époque des premiers mois de ma thèse une révolution vertigineuse, comme la puce de génotypage de 250 000 SNPs à partir de laquelle j'ai effectué les analyses de liaison de notre série de familles informatives, au séquençage de génome actuel qui génère des millions de données et qui ne sera qu'une étape avant des techniques toujours plus poussées comme le séquençage en temps réel de molécules uniques. Le fait que j'ai pu prendre part à cette émulation constante de la discipline en réalisant moi-même le séquençage d'exome de nos patients à Montréal, dans le cadre de ma cotutelle, au moment même où la technique se démocratisait a été extrêmement motivant. Cela m'a apporté une expertise précieuse dans l'analyse des résultats et a participé au plaisir que j'ai eu de concrétiser ce dernier cycle universitaire.

Ainsi ces nouvelles technologies d'analyse sont des atouts majeurs pour l'étude de gènes de l'acabit de la titine, et le nombre de mutations de *titine* associées à des maladies humaines est passé de 15 il y a quatre ans à un peu plus de cent mutations aujourd'hui. Nous ne sommes pas en reste dans l'analyse de ce titan, puisque ce travail de thèse a participé à une meilleure caractérisation de l'implication de variants hétérozygotes composites dans les phénotypes cardiaques et musculaires liés à la titine ; à l'association de la titine avec des défauts cardiaques parmi les plus fréquents dans les cas pédiatriques et pour lesquels la titine n'était pas un candidat prioritaire ; à l'identification d'une nouvelle maladie musculaire et cardiaque associée à la première description d'un KO du domaine kinase de titine chez l'homme et enfin à l'identification d'un nouveau rôle dans le turn-over des protéines du domaine kinase de la titine.

Les derniers résultats de cette thèse identifient pour la première fois des mutations dans un coactivateur de la transcription très peu documenté, relie cette nouvelle protéine à la physiologie musculaire et décrivent une nouvelle myopathie congénitale chez l'homme. Finalement, il est amusant de voir qu'ils constituent la partie la moins finalisée de l'ensemble de nos découvertes là où la titine aurait dû être une histoire "rapide de quelques mois"...



## IX. Les sources documentaires

Abecasis, G. R., S. S. Cherny, et al. (2002). "Merlin--rapid analysis of dense genetic maps using sparse gene flow trees." Nat Genet **30**(1): 97-101.

Agarkova, I., E. Ehler, et al. (2003). "M-band: a safeguard for sarcomere stability?" J Muscle Res Cell Motil **24**(2-3): 191-203.

Agarkova, I. and J. C. Perriard (2005). "The M-band: an elastic web that crosslinks thick filaments in the center of the sarcomere." Trends Cell Biol **15**(9): 477-485.

Allamand, V., Richard, P., Lescure, A., Ledeuil, C., Desjardin, D., Petit, N., Gartioux, C., Ferreiro, A., Krol, A., Pellegrini, N., Urtizberea, J. A. and Guicheney, P. "A single homozygous point mutation in a 3'untranslated region motif of selenoprotein N mRNA causes SEPNI-related myopathy." EMBO Rep, 2006, 7(4), 450-4.

Amodeo, P., M. A. Castiglione Morelli, et al. (2001). "Kinase recognition by calmodulin: modeling the interaction with the autoinhibitory region of human cardiac titin kinase." J Mol Biol **306**(1): 81-95.

Arimura, T., J. M. Bos, et al. (2009). "Cardiac ankyrin repeat protein gene (ANKRD1) mutations in hypertrophic cardiomyopathy." J Am Coll Cardiol **54**(4): 334-342.

Azaouagh, A., S. Churzidse, et al. (2011). "Arrhythmogenic right ventricular cardiomyopathy/dysplasia: a review and update." Clin Res Cardiol **100**(5): 383-394.

Bamshad, M., A. E. Van Heest, et al. (2009). "Arthrogryposis: a review and update." J Bone Joint Surg Am **91 Suppl 4**: 40-46.

Bang, M. L., T. Centner, et al. (2001). "The complete gene sequence of titin, expression of an unusual approximately 700-kDa titin isoform, and its interaction with obscurin identify a novel Z-line to I-band linking system." Circ Res **89**(11): 1065-1072.

Barresi, R. and Campbell, K. P. "Dystroglycan: from biosynthesis to pathogenesis of human disease." J Cell Sci, 2006, 119(Pt 2), 199-207.

Beckmann, J. S. and M. Spencer (2008). "Calpain 3, the "gatekeeper" of proper sarcomere assembly, turnover and maintenance." Neuromuscul Disord **18**(12): 913-921.

Benian, G. M., J. E. Kiff, et al. (1989). "Sequence of an unusually large protein implicated in regulation of myosin activity in *C. elegans*." Nature **342**(6245): 45-50.

Berg, J. S., B. C. Powell, et al. (2001). "A millennial myosin census." Mol Biol Cell **12**(4): 780-794.

Bethlem, J. and Wijngaarden, G. K. "Benign myopathy, with autosomal dominant inheritance. A report on three pedigrees." Brain, 1976, 99(1), 91-100.

Bohm, J., V. Biancalana, et al. (2012). "Mutation spectrum in the large GTPase dynamin 2, and genotype-phenotype correlation in autosomal dominant centronuclear myopathy." Hum Mutat **33**(6): 949-959.

Boyden, S. E., L. J. Mahoney, et al. (2012). "Mutations in the satellite cell gene MEGF10 cause a recessive congenital myopathy with minicores." Neurogenetics **13**(2): 115-124.

Caldwell, C. J., Matthey, D. L. and Weller, R. O. "Role of the basement membrane in the regeneration of skeletal muscle." Neuropathol Appl Neurobiol, 1990, 16(3), 225-38.

Campanelli, J. T., Roberds, S. L., Campbell, K. P. and Scheller, R. H. "A role for dystrophin-associated glycoproteins and utrophin in agrin-induced AChR clustering." Cell, 1994, 77(5), 663-74.

**Carmignac, V., M. A. Salih, et al. (2007). "C-terminal titin deletions cause a novel early-onset myopathy with fatal cardiomyopathy." Ann Neurol **61**(4): 340-351.**

Ceuterick, C. and J. J. Martin (1996). "Sporadic early adult-onset distal myopathy with rimmed vacuoles: immunohistochemistry and electron microscopy." J Neurol Sci **139**(2): 190-196.

Charton, K., N. Daniele, et al. (2010). "Removal of the calpain 3 protease reverses the myopathology in a mouse model for titinopathies." Hum Mol Genet **19**(23): 4608-4624.

Chi, R. J., S. G. Olenych, et al. (2005). "Smooth muscle alpha-actinin interaction with smitin." Int J Biochem Cell Biol **37**(7): 1470-1482.

Chi, R. J., A. R. Simon, et al. (2008). "Smooth muscle titin Zq domain interaction with the smooth muscle alpha-actinin central rod." J Biol Chem **283**(30): 20959-20967.

Cohn, R. D. "Dystroglycan: important player in skeletal muscle and beyond." Neuromuscul Disord, 2005, 15(3), 207-17.

Cohn, R. D., Mayer, U., Saher, G., Herrmann, R., van der Flier, A., Sonnenberg, A., Sorokin, L. and Voit, T. "Secondary reduction of alpha7B integrin in laminin alpha2 deficient congenital muscular dystrophy supports an additional transmembrane link in skeletal muscle." J Neurol Sci, 1999, 163(2), 140-52.

Colognato, H. and Yurchenco, P. D. (1999). "The laminin alpha2 expressed by dystrophic dy(2J) mice is defective in its ability to form polymers." Curr Biol, 9(22), 1327-30.

Dellefave, L. and E. M. McNally (2010). "The genetics of dilated cardiomyopathy." Curr Opin Cardiol **25**(3): 198-204.

Demir, E., Ferreira, A., Sabatelli, P., Allamand, V., Makri, S., Echenne, B., Maraldi, M., Merlini, L., Topaloglu, H. and Guicheney, P. "Collagen VI status and clinical severity in Ullrich congenital muscular dystrophy: phenotype analysis of 11 families linked to the COL6 loci." Neuropediatrics, 2004, 35(2), 103-12.

Dubowitz, V. (1995). "Muscle Disorders in Childhood." London, WB Saunders.

den Dunnen, J. T. and S. E. Antonarakis (2000). "Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion." Hum Mutat **15**(1): 7-12.

Edstrom, L., L. E. Thornell, et al. (1990). "Myopathy with respiratory failure and typical myofibrillar lesions." J Neurol Sci **96**(2-3): 211-228.

Ervasti, J. M. and Campbell, K. P. "A role for the dystrophin-glycoprotein complex as a transmembrane linker between laminin and actin." J Cell Biol, 1993, 122(4), 809-23.

**Ferreiro, A., N. Monnier, et al. (2002). "A recessive form of central core disease, transiently presenting as multi-minicore disease, is associated with a homozygous mutation in the ryanodine receptor type 1 gene." Ann Neurol 51(6): 750-759.**

**Ferreiro, A., S. Quijano-Roy, et al. (2002). "Mutations of the selenoprotein N gene, which is implicated in rigid spine muscular dystrophy, cause the classical phenotype of multiminicore disease: reassessing the nosology of early-onset myopathies." Am J Hum Genet 71(4): 739-749.**

Freiburg, A. and M. Gautel (1996). "A molecular map of the interactions between titin and myosin-binding protein C. Implications for sarcomeric assembly in familial hypertrophic cardiomyopathy." Eur J Biochem 235(1-2): 317-323.

Freiburg, A., K. Trombitas, et al. (2000). "Series of exon-skipping events in the elastic spring region of titin as the structural basis for myofibrillar elastic diversity." Circ Res 86(11): 1114-1121.

Frey, N., M. Luedde, et al. (2012). "Mechanisms of disease: hypertrophic cardiomyopathy." Nat Rev Cardiol 9(2): 91-100.

Furst, D. O., M. Osborn, et al. (1988). "The organization of titin filaments in the half-sarcomere revealed by monoclonal antibodies in immunoelectron microscopy: a map of ten nonrepetitive epitopes starting at the Z line extends close to the M line." J Cell Biol 106(5): 1563-1572.

Garvey, S. M., C. Rajan, et al. (2002). "The muscular dystrophy with myositis (mdm) mouse mutation disrupts a skeletal muscle-specific domain of titin." Genomics 79(2): 146-149.

Gautel, M., D. Goulding, et al. (1996). "The central Z-disk region of titin is assembled from a novel repeat in variable copy numbers." J Cell Sci 109 ( Pt 11): 2747-2754.

Gautel, M., K. Leonard, et al. (1993). "Phosphorylation of KSP motifs in the C-terminal region of titin in differentiating myoblasts." EMBO J 12(10): 3827-3834.

Gospe, S. M., Jr., D. L. Armstrong, et al. (1987). "Life-threatening congestive heart failure as the presentation of centronuclear myopathy." Pediatr Neurol 3(2): 117-120.

Gotthardt, M., R. E. Hammer, et al. (2003). "Conditional expression of mutant M-line titins results in cardiomyopathy with altered sarcomere structure." J Biol Chem 278(8): 6059-6065.

Granzier, H. L. and T. C. Irving (1995). "Passive tension in cardiac muscle: contribution of collagen, titin, microtubules, and intermediate filaments." Biophys J 68(3): 1027-1044.

Grater, F., J. Shen, et al. (2005). "Mechanically induced titin kinase activation studied by force-probe molecular dynamics simulations." Biophys J 88(2): 790-804.

Gregorio, C. C., K. Trombitas, et al. (1998). "The NH2 terminus of titin spans the Z-disc: its interaction with a novel 19-kD ligand (T-cap) is required for sarcomeric integrity." J Cell Biol 143(4): 1013-1027.

Guo, W., S. J. Bharmal, et al. (2010). "Titin diversity--alternative splicing gone wild." J Biomed Biotechnol **2010**: 753675.

Gutierrez-Cruz, G., A. H. Van Heerden, et al. (2001). "Modular motif, structural folds and affinity profiles of the PEVK segment of human fetal skeletal muscle titin." J Biol Chem **276**(10): 7442-7449.

Hackman, P., S. Marchand, et al. (2008). "Truncating mutations in C-terminal titin may cause more severe tibial muscular dystrophy (TMD)." Neuromuscul Disord **18**(12): 922-928.

Hackman, P., A. Vihola, et al. (2002). "Tibial muscular dystrophy is a titinopathy caused by mutations in TTN, the gene encoding the giant skeletal-muscle protein titin." Am J Hum Genet **71**(3): 492-500.

Hayashi, Y. K., Chou, F. L., Engvall, E., Ogawa, M., Matsuda, C., Hirabayashi, S., Yokochi, K., Ziober, B. L., Kramer, R. H., Kaufman, S. J., Ozawa, E., Goto, Y., Nonaka, I., Tsukahara, T., Wang, J. Z., Hoffman, E. P. and Arahata, K. "Mutations in the integrin alpha7 gene cause congenital myopathy." Nat Genet, 1998, 19(1), 94-7.

Helbling-Leclerc, A., Zhang, X., Topaloglu, H., Cruaud, C., Tesson, F., Weissenbach, J., Tome, F. M., Schwartz, K., Fardeau, M., Tryggvason, K. and et al. "Mutations in the laminin alpha 2-chain gene (LAMA2) cause merosin- deficient congenital muscular dystrophy." Nat Genet, 1995, 11(2), 216-8.

Henry, M. D. and Campbell, K. P. "A role for dystroglycan in basement membrane assembly." Cell, 1998, 95(6), 859-70.

Herman, D. S., L. Lam, et al. (2012). "Truncations of titin causing dilated cardiomyopathy." N Engl J Med **366**(7): 619-628.

Hillaire, D., Leclerc, A., Faure, S., Topaloglu, H., Chiannikulchai, N., Guicheney, P., Grinas, L., Legos, P., Philpot, J., Evangelista, T. and et al. "Localization of merosin-negative congenital muscular dystrophy to chromosome 6q2 by homozygosity mapping." Hum Mol Genet, 1994, 3(9), 1657-61.

Houmeida, A., J. Holt, et al. (1995). "Studies of the interaction between titin and myosin." J Cell Biol **131**(6 Pt 1): 1471-1481.

Huebsch, K. A., E. Kudryashova, et al. (2005). "Mdm muscular dystrophy: interactions with calpain 3 and a novel functional role for titin's N2A domain." Hum Mol Genet **14**(19): 2801-2811.

Huxley, H. E. (1969). "The mechanism of muscular contraction." Science **164**(3886): 1356-1365.

Irwin, W. A., Bergamin, N., Sabatelli, P., Reggiani, C., Megighian, A., Merlini, L., Braghetta, P., Columbaro, M., Volpin, D., Bressan, G. M., Bernardi, P. and Bonaldo, P. "Mitochondrial dysfunction and apoptosis in myopathic mice with collagen VI deficiency." Nat Genet, 2003, 35(4), 367-71.

Itoh-Satoh, M., T. Hayashi, et al. (2002). "Titin mutations as the molecular basis for dilated cardiomyopathy." Biochem Biophys Res Commun **291**(2): 385-393.

Jungbluth, H. (2007). "Central core disease." Orphanet J Rare Dis **2**: 25.

- Kaindl, A. M., F. Ruschendorf, et al. (2004). "Missense mutations of ACTA1 cause dominant congenital myopathy with cores." *J Med Genet* **41**(11): 842-848.
- Kamisago, M., S. D. Sharma, et al. (2000). "Mutations in sarcomere protein genes as a cause of dilated cardiomyopathy." *N Engl J Med* **343**(23): 1688-1696.
- Kim, K. and T. C. Keller, 3rd (2002). "Smitin, a novel smooth muscle titin-like protein, interacts with myosin filaments in vivo and in vitro." *J Cell Biol* **156**(1): 101-111.
- King, L. and C. R. Jhou (2010). "Nuclear titin interacts with histones." *Chang Gung Med J* **33**(2): 201-210.
- Kobe, B., J. Heierhorst, et al. (1996). "Giant protein kinases: domain interactions and structural basis of autoregulation." *EMBO J* **15**(24): 6810-6821.
- Labeit, S., D. P. Barlow, et al. (1990). "A regular pattern of two types of 100-residue motif in the sequence of titin." *Nature* **345**(6272): 273-276.
- Labeit, S., M. Gautel, et al. (1992). "Towards a molecular understanding of titin." *EMBO J* **11**(5): 1711-1716.
- Labeit, S. and B. Kolmerer (1995). "Titins: giant proteins in charge of muscle ultrastructure and elasticity." *Science* **270**(5234): 293-296.
- Labeit, S., S. Lahmers, et al. (2006). "Expression of distinct classes of titin isoforms in striated and smooth muscles by alternative splicing, and their conserved interaction with filamins." *J Mol Biol* **362**(4): 664-681.
- Lander, E. S. and D. Botstein (1987). "Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children." *Science* **236**(4808): 1567-1570.
- Lange, S., F. Xiang, et al. (2005). "The kinase domain of titin controls muscle gene expression and protein turnover." *Science* **308**(5728): 1599-1603.
- Laporte, J., L. J. Hu, et al. (1996). "A gene mutated in X-linked myotubular myopathy defines a new putative tyrosine phosphatase family conserved in yeast." *Nat Genet* **13**(2): 175-182.
- Linke, W. A., M. Ivemeyer, et al. (1997). "Actin-titin interaction in cardiac myofibrils: probing a physiological role." *Biophys J* **73**(2): 905-919.
- Linke, W. A. and M. Kruger (2010). "The giant protein titin as an integrator of myocyte signaling pathways." *Physiology (Bethesda)* **25**(3): 186-198.
- Linke, W. A., M. Kulke, et al. (2002). "PEVK domain of titin: an entropic spring with actin-binding properties." *J Struct Biol* **137**(1-2): 194-205.
- Lisi, M. T. and Cohn, R. D. (2007) "Congenital muscular dystrophies: New aspects of an expanding group of disorders." *Biochim Biophys Acta*, **1772**(2), 159-72.

- Machado, C. and D. J. Andrew (2000). "D-Titin: a giant protein with dual roles in chromosomes and muscles." J Cell Biol **151**(3): 639-652.
- Maquat, L. E. (2004). "Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics." Nat Rev Mol Cell Biol **5**(2): 89-99.
- Marchand, S., P. Hajdari, et al. (2003). "Touch-down method for high-performance sequencing of polymerase chain reaction products." Anal Biochem **315**(2): 270-272.
- Martin, P. T. "The dystroglycanopathies: the new disorders of O-linked glycosylation." Semin Pediatr Neurol, 2005, 12(3), 152-8.
- Maruyama, K., D. H. Hu, et al. (1987). "Binding of actin filaments to connectin." J Biochem **101**(6): 1339-1346.
- Mayans, O., P. F. van der Ven, et al. (1998). "Structural basis for activation of the titin kinase domain during myofibrillogenesis." Nature **395**(6705): 863-869.
- Mercuri, E., Yuva, Y., Brown, S. C., Brockington, M., Kinali, M., Jungbluth, H., Feng, L., Sewry, C. A. and Muntoni, F. "Collagen VI involvement in Ullrich syndrome: a clinical, genetic, and immunohistochemical study." Neurology, 2002, 58(9), 1354-9.
- Miller, G., H. Musa, et al. (2003). "A targeted deletion of the C-terminal end of titin, including the titin kinase domain, impairs myofibrillogenesis." J Cell Sci **116**(Pt 23): 4811-4819.
- Moghadaszadeh, B., Petit, N., Jaillard, C., Brockington, M., Roy, S. Q., Merlini, L., Romero, N., Estournet, B., Desguerre, I., Chaigne, D., Muntoni, F., Topaloglu, H. and Guicheney, P. "Mutations in SEPN1 cause congenital muscular dystrophy with spinal rigidity and restrictive respiratory syndrome." Nat Genet, 2001, 29(1), 17-8.
- Mues, A., P. F. van der Ven, et al. (1998). "Two immunoglobulin-like domains of the Z-disc portion of titin interact in a conformation-dependent way with telethonin." FEBS Lett **428**(1-2): 111-114.
- Muntoni, F. "Journey into muscular dystrophies caused by abnormal glycosylation." Acta Myol, 2004, 23(2), 79-84.
- Nagy, E. and L. E. Maquat (1998). "A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance." Trends Biochem Sci **23**(6): 198-199.
- Nicot, A. S., A. Toussaint, et al. (2007). "Mutations in amphiphysin 2 (BIN1) disrupt interaction with dynamin 2 and cause autosomal recessive centronuclear myopathy." Nat Genet **39**(9): 1134-1139.
- North, K. N., N. G. Laing, et al. (1997). "Nemaline myopathy: current concepts. The ENMC International Consortium and Nemaline Myopathy." J Med Genet **34**(9): 705-713.
- Obermann, W. M., M. Gautel, et al. (1997). "Molecular structure of the sarcomeric M band: mapping of titin and myosin binding domains in myomesin and the identification of a potential regulatory phosphorylation site in myomesin." EMBO J **16**(2): 211-220.

- Ockeloen, C. W., H. J. Gilhuis, et al. (2012). "Congenital myopathy caused by a novel missense mutation in the CFL2 gene." Neuromuscul Disord **22**(7): 632-639.
- Ohlsson, M., C. Hedberg, et al. (2012). "Hereditary myopathy with early respiratory failure associated with a mutation in A-band titin." Brain **135**(Pt 6): 1682-1694.
- Peng, J., K. Raddatz, et al. (2005). "Muscle atrophy in titin M-line deficient mice." J Muscle Res Cell Motil **26**(6-8): 381-388.
- Peng, J., K. Raddatz, et al. (2007). "Cardiac hypertrophy and reduced contractility in hearts deficient in the titin kinase region." Circulation **115**(6): 743-751.
- Pepe, G., Bertini, E., Bonaldo, P., Bushby, K., Giusti, B., de Visser, M., Guicheney, P., Lattanzi, G., Merlini, L., Muntoni, F., Nishino, I., Nonaka, I., Yaou, R. B., Sabatelli, P., Sewry, C., Topaloglu, H. and van der Kooi, A. Bethlem myopathy (BETHLEM) and Ullrich scleroatonic muscular dystrophy: 100th ENMC international workshop, 23-24 November 2001, Naarden, The Netherlands. Neuromuscul Disord, 2002, 12(10), 984-93.
- Pernigo, S., A. Fukuzawa, et al. (2010). "Structural insight into M-band assembly and mechanics from the titin-obscurin-like-1 complex." Proc Natl Acad Sci U S A **107**(7): 2908-2913.
- Petit, N., Lescure, A., Rederstorff, M., Krol, A., Moghadaszadeh, B., Wewer, U. M. and Guicheney, P. "Selenoprotein N: an endoplasmic reticulum glycoprotein with an early developmental expression pattern." Hum Mol Genet, 2003, 12(9), 1045-53.
- Puchner, E. M., A. Alexandrovich, et al. (2008). "Mechanoenzymatics of titin kinase." Proc Natl Acad Sci U S A **105**(36): 13385-13390.
- Qi, J., L. Chi, et al. (2008). "Nuclear localization of the titin Z1Z2Zr domain and role in regulating cell proliferation." Am J Physiol Cell Physiol **295**(4): C975-985.
- Radke, M. H., J. Peng, et al. (2007). "Targeted deletion of titin N2B region leads to diastolic dysfunction and cardiac atrophy." Proc Natl Acad Sci U S A **104**(9): 3444-3449.
- Richardson, P., W. McKenna, et al. (1996). "Report of the 1995 World Health Organization/International Society and Federation of Cardiology Task Force on the Definition and Classification of cardiomyopathies." Circulation **93**(5): 841-842.
- Sanoudou, D. and A. H. Beggs (2001). "Clinical and genetic heterogeneity in nemaline myopathy--a disease of skeletal muscle thin filaments." Trends Mol Med **7**(8): 362-368.
- Sauer, F., J. Vahokoski, et al. (2010). "Molecular basis of the head-to-tail assembly of giant muscle proteins obscurin-like 1 and titin." EMBO Rep **11**(7): 534-540.
- Schoenauer, R., S. Lange, et al. (2008). "Myomesin 3, a novel structural component of the M-band in striated muscle." J Mol Biol **376**(2): 338-351.
- Selcen, D. (2011). "Myofibrillar myopathies." Neuromuscul Disord **21**(3): 161-171.
- Steffen, L. S., J. R. Guyon, et al. (2007). "The zebrafish runzel muscular dystrophy is linked to the titin gene." Dev Biol **309**(2): 180-192.

- Sewry, C. A., Philpot, J., Mahony, D., Wilson, L. A., Muntoni, F. and Dubowitz, V. "Expression of laminin subunits in congenital muscular dystrophy." Neuromuscul Disord, 1995, 5(4), 307-16.
- Sugita, S., Saito, F., Tang, J., Satz, J., Campbell, K. and Sudhof, T. C. "A stoichiometric complex of neuexins and dystroglycan in brain." J Cell Biol, 2001, 154(2), 435-45.
- Talts, J. F., Andac, Z., Gohring, W., Brancaccio, A. and Timpl, R. "Binding of the G domains of laminin alpha1 and alpha2 chains and perlecan to heparin, sulfatides, alpha-dystroglycan and several extracellular matrix proteins." Embo J, 1999, 18(4), 863-70.
- Trinick, J. (1994). "Titin and nebulin: protein rulers in muscle?" Trends Biochem Sci **19**(10): 405-409.
- Trinick, J. and L. Tskhovrebova (1999). "Titin: a molecular control freak." Trends Cell Biol **9**(10): 377-380.
- Trombitas, K., A. Redkar, et al. (2000). "Extensibility of isoforms of cardiac titin: variation in contour length of molecular subsegments provides a basis for cellular passive stiffness diversity." Biophys J **79**(6): 3226-3234.
- Udd, B. (2012). "Distal myopathies--new genetic entities expand diagnostic challenge." Neuromuscul Disord **22**(1): 5-12.
- Udd, B., A. Vihola, et al. (2005). "Titinopathies and extension of the M-line mutation phenotype beyond distal myopathy and LGMD2J." Neurology **64**(4): 636-642.
- Vachon, P. H., Loechel, F., Xu, H., Wewer, U. M. and Engvall, E. (1996). "Merosin and laminin in myogenesis; specific requirement for merosin in myotube stability and survival". J Cell Biol, 134(6), 1483-97.
- Vachon, P. H., Xu, H., Liu, L., Loechel, F., Hayashi, Y., Arahata, K., Reed, J. C., Wewer, U. M. and Engvall, E. "Integrins (alpha7beta1) in muscle function and survival. Disrupted expression in merosin-deficient congenital muscular dystrophy." J Clin Invest, 1997, 100(7), 1870-81.
- van der Ven, P. F. and D. O. Furst (1997). "Assembly of titin, myomesin and M-protein into the sarcomeric M band in differentiating human skeletal muscle cells in vitro." Cell Struct Funct **22**(1): 163-171.
- van der Ven, P. F., G. Schaart, et al. (1993). "Titin aggregates associated with intermediate filaments align along stress fiber-like structures during human skeletal muscle cell differentiation." J Cell Sci **106 ( Pt 3)**: 749-759.
- Verhies, W., J. M. Brucher, et al. (1976). "Familial centronuclear myopathy associated with 'cardiomyopathy'." Br Heart J **38**(5): 504-509.
- Wang, K., R. McCarter, et al. (1991). "Regulation of skeletal muscle stiffness and elasticity by titin isoforms: a test of the segmental extension model of resting tension." Proc Natl Acad Sci U S A **88**(16): 7101-7105.

Williamson, R. A., Henry, M. D., Daniels, K. J., Hrstka, R. F., Lee, J. C., Sunada, Y., Ibraghimov-Beskrovnaya, O. and Campbell, K. P. "Dystroglycan is essential for early embryonic development: disruption of Reichert's membrane in *Dag1*-null mice." Hum Mol Genet, 1997, 6(6), 831-41.

Xu, X., S. E. Meiler, et al. (2002). "Cardiomyopathy in zebrafish due to mutation in an alternatively spliced exon of titin." Nat Genet **30**(2): 205-209.

Young, P., C. Ferguson, et al. (1998). "Molecular structure of the sarcomeric Z-disk: two types of titin interactions lead to an asymmetrical sorting of alpha-actinin." EMBO J **17**(6): 1614-1624.

Zhang, R. Z., Sabatelli, P., Pan, T. C., Squarzoni, S., Mattioli, E., Bertini, E., Pepe, G. and Chu, M. L. "Effects on collagen VI mRNA stability and microfibrillar assembly of three COL6A2 mutations in two families with Ullrich congenital muscular dystrophy." **J Biol Chem**, 2002, 277(46), 43557-64.



## X. Les appendices (annexes)

### Annexe 1 : Descriptif des principales myopathies congénitales

MC	Transmission	Gènes	Phénotype clinique	Anomalies morphologiques	Références
<b>Cores centraux</b>	AR	<i>RYR1</i>	<ul style="list-style-type: none"> <li>Faiblesse musculaire modérée et peu évolutive</li> <li>Scoliose modérée à sévère</li> </ul>	<ul style="list-style-type: none"> <li>Variabilité taille des fibres</li> <li>prédominance des fibres de type I               <ul style="list-style-type: none"> <li>Noyaux centraux</li> <li>Cores centraux</li> </ul> </li> </ul>	Romero, 2003
	AD	<i>RYR1</i>	<ul style="list-style-type: none"> <li>Faiblesse musculaire               <ul style="list-style-type: none"> <li>Hypotonie</li> <li>Retard moteur</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>Cores centraux</li> <li>Typage de fibres types I ( 99%)</li> </ul>	Jungbluth 2011
		<i>ACTA1</i>	<ul style="list-style-type: none"> <li>Infantile</li> <li>Faiblesse musculaire peu sévère peu évolutive</li> </ul>	Cores centraux	Kaindl, 2004
<b>Myotubulaire</b>	Lié à l'X	<i>MTM1</i>	<ul style="list-style-type: none"> <li>Congénitale</li> <li>Insuffisance respiratoire grave</li> <li>Lethalité</li> </ul>		
<b>Centronucléaire</b>	AR	<i>BIN1</i>	<ul style="list-style-type: none"> <li>Formes infantiles</li> <li>faiblesse musculaire proximale</li> <li>faiblement progressive</li> </ul>	<ul style="list-style-type: none"> <li>Nombreux petits noyaux centraux</li> <li>Typage et hypotrophie des fibres de type I</li> </ul>	Bitoun, 2011
	AD	<i>DMN2</i> <i>RYR1</i>	<ul style="list-style-type: none"> <li>Formes adultes</li> <li>Ptosis bilatéral</li> <li>Anomalies squelettiques</li> <li>Peu évolutive</li> </ul>		
<b>Multi- minicores</b>	AR	<i>SEPN1</i>	<ul style="list-style-type: none"> <li>Hypotonie néonatale</li> <li>Retard moteur</li> <li>Faiblesse musculaire généralisée (4 formes)</li> </ul>	<ul style="list-style-type: none"> <li>Typage et hypotrophie des fibres de type I</li> <li>Minicores &gt;60% des fibres</li> </ul>	Ferreiro, 2002
		<i>RYR1</i>			Monnier, 2003
		<i>TTN</i>			Carmignac, 2007
<b>Némaline</b>	AR	<i>NEB</i>	<ul style="list-style-type: none"> <li>Phénotype variables</li> <li>Faiblesse musculaire du visage, cou et muscles proximaux des membres</li> </ul>	Bâtonnets intra-cytoplasmiques (nemalin bodies)	Laing, 2011
		<i>TNNT1</i>			
		<i>CFL2</i>			
	AR/AD	<i>ACTA1</i>			
	<i>TMP3</i>				
AD	<i>TMP2</i> <i>KBTD13</i>				

**Annexe 2 : Séquences des primers utilisées pour le séquençage de l'extrémité N-terminale de TTN**

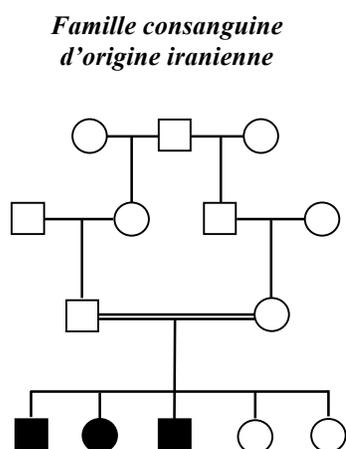
<b>TTN-Nter</b>	<b>Séquences amorces sens</b>	<b>Séquences amorces antisens</b>	<b>Taille du produit de PCR (pb)</b>
ex1	GCAAAGCAGCTCCAGAGTG	AAGTCCTGCAGCAACGTTAAC	247
ex2	GGACAAGGCAGTGAACATGA	GAGCAGAGGATTGGCAGGT	323
ex3	CCTGGCTCTCACTCTTGAC	TCAGCTCACAGCATTCTTCC	443
ex4/5	CTCCATTGGGATATCCAGAA	CTCGTTTCAAACCTAGTTCC	597
ex6	CTCATGATTCACATCCAGTTACC	CAGACTGAGTTTCATGGCAGAA	478
ex7	GATACAGTCTCTGTATTTCTTGG	TGGATGGTGGTTGAGTTAATGT	372
ex8	CACCATGTTGGCGAGGCT	AGTATGCTAACCAACCATGTGGC	334
ex9	GTGACATCTGCAGAAAAGGG	TATAAGCTACCTGCAGCTGG	320
ex10	GTTCACTACTAAGCATCCTGT	GATCCATGATGAAAATGTAGGTG	319
ex11	GCCAGTGAAACAGTATGGATG	CTCTAGATGAGTCCCTGTGT	593
ex12	GTGTGTTTGAAGCCCTGTTCA	GTGAAGAGATAAGTGGAGAAAG	414
ex13	ACTGCTGTATAATGTGGACGC	AGCAGTGAGGCTTGCTTTAC	455
ex14	GTAAAGCAAGCCTCACTGCT	GATGCTCTGTTTCACAGGTTAG	262
ex15	GACTCAGTCCAAGAGGGCT	TACAACTAGCCAACCACCTG	426
ex16	GCATGAGCTAAGAGTTTATGAGA	CATCTGTAAGCTCACTAATGTCA	259
ex17	GTGGCAGTTGAAATGCATAAGTT	GTGCACAGAAACCATATTGTGG	447
ex18/19	CCACAATATGGTTTCTGTGCAC	TCTACTCTAGGCTTCATGCAC	599
ex20	CCACAGATCCACTCACAGAT	CATACAACTGAGGCAAAGACAC	371
ex21	GAATGCATAGTGGTATGACTGT	ACTTGGACCTTCTAATAGCTGT	394
ex22	CAACAGTCTAGCTCAGATACAT	CCAATTTGTATCTTTACTGTGGC	373
ex23	GCCACAGTAAAGATACAAATTGG	CCAATGGTAAGTTTCTGTGCC	449
ex24	CTGAATGAGTTTGAGGTGCCT	CATGAGCAAAAACCTTATCACGC	439
ex25/26	GCGTGATAAGTTTTTGCTCATG	CTTCCCTGGTTATTGGATTTGT	607
ex27.1	ACAAATCCAATAACCAGGGAAG	CTTCAAGTGGCTTCCATCAT	490
ex27.2	CTTCCTTAAGACTTAAGCGCTT	CTGTCCATTGAGGTACCAGT	509
ex27.3	TTGTCTTGTAACCAGAGCCA	CTCTTCTGTTCTGCGCTTGA	494
ex27.4	CTGTTGACACCACTGAAACCA	CAGTCACATCTCTTATGACCAA	528
ex27.5	CAAGGATCTGATGCACACTT	ACACACTCATGGATATTCTTGAA	301
ex28	TTCAAGAATATCCATGAGTGTGT	GATTCTGGCTATGAACTTATAG	530

### Annexe 3 : Mutations TTN répertoriées dans la littérature

Portion de TTN		All exon TTN		Prédiction de l'effet sur la protéine		Nomenclature première publication		Phénotype		Type de mutation		Génotype		Référence	
Z-line	Exon	Changement ADNc	ADNC	ADNC	ADNC	ADNC	ADNC	ADNC	ADNC	ADNC	ADNC	ADNC	ADNC	ADNC	ADNC
	3	c.160G>A		p.Val154Met		p.Val154Met		DCM		missense		het	Itoh-Satoh (2002)		
	14	c.2219G>T		p.Arg740Leu		p.Arg740Leu		DCM		missense		het	Satoh (1999)		
	14	c.2228C>T		p.Ala743Val		p.Ala743Val		DCM		missense		het	Itoh-Satoh (2002)		
	18	c.2926T>C		p.Trp976Arg		p.Trp976Arg		DCM		missense		het	Gerull (2002)		
	18	c.3100G>A		-		-		MD		splicing (5')		comp het <sup>1</sup>	Vasli (2012)		
	28	c.6248delG		p.Arg2083Lysfs*55		p.Arg2083Lysfs*55		DCM		frameshift		het	Herman (2012)		
	32	c.7450C>T		p.Gln2484*		p.Gln2484*		cancer		nonsense (VUS)		het	Greenman (2007)		
	37	c.8687C>T		p.Thr2896Ile		p.Thr2896Ile		ARVC		missense		het	Taylor (2011)		
	49 (N2B)	c.12485C>A		p.Ser4162Tyr		p.Ser3799Tyr (N2B; NP_003310.4)		DCM		missense		het	Itoh-Satoh (2002)		
	49 (N2B)	c.13108C>T		p.Gln4370*		p.Gln4370*		DCM		nonsense		het	Itoh-Satoh (2002)		
	50 (N2B)	c.13696C>T		p.Gln4566*		p.Gln4566*		DCM		nonsense		het	Herman (2012)		
	50	c.14141G>A		p.Gly4714Asp		p.Gly4714Asp		DCM		missense		het	Liu (2008)		
	50	c.14339G>A		p.Ser4780Asn		p.Ser4465Asp		DCM		missense		het	Herman (2012)		
	70	c.20134delG		p.Asp6712Ilefs*4		p.Asp6712Ilefs*4		DCM		frameshift		het	Herman (2012)		
	86	c.24749_24751delGTCAGATACTCG		p.Gln8250Valfs*7		p.Gln7933Val*		DCM		frameshift (VUS)		het	Herman (2012)		
	97	c.27824A>G		p.Tyr9275Cys		p.Tyr9275Cys		ARVC		missense		het	Taylor (2011)		
	103	c.29231G>A		p.Arg9744His		p.Arg8500His		DCM		missense		het	Arimura (2009)		
	104 (N2A)	c.29543G>A		p.Arg948Gln		p.Arg604Gln or p.Arg8606Gln		DCM		missense		het	Arimura (2009)		
	108	c.30274C>T		p.His10092Tyr		p.His8848Cys		ARVC		missense		het	Taylor (2011)		
	119 (PEVK)	c.31427-1G>A		-		c.30476-1G>A		DCM		splicing (3')		het	Herman (2012)		
	126	c.35263-1G>T		-		c.34186+1G>T		DCM		splicing (5')		het	Herman (2012)		
	219 (PEVK)	c.40423G>C		-		c.35635G>C		DCM		splicing (5')		comp het	Herman (2012)		
	219 (PEVK)	c.40423+1G>C		-		c.35635+1G>A		DCM		splicing (5')		het	Herman (2012)		
	233	c.38024-2A>G		-		c.38024-2A>G		DCM		splicing (3')		het	Herman (2012)		
	236	c.43409_43410insA		p.Val14471Aspfs*8		p.Val14471Aspfs*8		DCM		frameshift		het	Herman (2012)		
	240-296	28 kb insertion		p.Pro14894_Thr19238dup		p.Pro13298_Thr17642dup		DCM		duplication		het	Herman (2012)		
	248	c.46101C>A		p.Cys15367*		p.Cys13771*		DCM		nonsense		het	Herman (2012)		
	297	c.58415T>C		p.Ile119472IThr		p.Ile119472IThr		ARVC		missense		het	Taylor (2011)		
	262	c.49124delA		p.Glu116375Glyfs*22		p.Glu116375Glyfs*22		DCM		frameshift		het	Herman (2012)		
	264	c.49513-2delT		-		c.44725+2delT		DCM		splicing (5')		hom	Herman (2012)		
	266	c.50110delT		p.Phe16704Leufs*14		p.Phe15108fs		DCM		frameshift		het	Herman (2012)		
	278	c.53152-1G>T		-		c.48364+1G>T		DCM		splicing (5')		het	Herman (2012)		
	278	c.53353G>T		p.Gly17785*		p.Gly16189*		DCM		nonsense		het	Herman (2012)		
	280	c.53865G>A		p.Trp17955*		p.Trp16359*		DCM		nonsense		het	Herman (2012)		
	284	c.55134+3A>G		-		c.50346+3A>G		DCM		splicing (5')		het	Herman (2012)		
	291	c.56812G>A		p.Ala18938Thr		p.Ala16415Thr		MD		missense		comp het <sup>1</sup>	Vasli (2012)		
	294	c.57196C>T		p.Arg19066*		p.Arg17470*		DCM		nonsense		het	Herman (2012)		
	296	c.57993_57994insG		p.Glu19311Glyfs*26		p.Glu17715fs		DCM		frameshift		het	Herman (2012)		
	301	c.59210-5T>A		-		p.Glu17783*		DCM		nonsense		het	Herman (2012)		
	302	c.59492-1G>A		-		c.54422-5T>A		DCM		splicing (3')		comp het	Herman (2012)		
	302	c.55003+1G>A		-		c.55003+1G>A		DCM		splicing (5')		het	Herman (2012)		
	304	c.61155T>A		p.Cys20385*		p.Cys18789*		DCM		nonsense		het	Herman (2012)		
	304	c.61741C>T		p.Arg20581*		p.Arg18858*		DCM		nonsense		het	Herman (2012)		
	305	c.63304G>A		p.Ala21102Thr		p.Ala18579Thr		DCM		nonsense		het	Herman (2012)		
	306	c.63466C>T		p.Arg21156*		p.Arg21156*		DCM		missense		comp het <sup>3</sup>	Taylor (2011)		
	311	c.64935delC		p.Pro21645Profs*5		p.Pro20049Pro*		DCM		frameshift		comp het	Herman (2012)		
	313	c.65494G>T		p.Ala21832Ser		p.Ala19309Ser		ARVC		missense		het	Taylor (2011)		
	318	c.67213+5G>A		-		c.62425+5G>A		DCM		splicing (5')		hom	Herman (2012)		
	321	c.67360C>T		p.Arg22454*		p.Arg20858*		DCM		nonsense		het	Herman (2012)		
	321	c.68193A>G		-		c.63405A>G		DCM		splicing (5')		het	Herman (2012)		
	323	c.68689-5G>C		-		c.63901+5G>C		DCM		splicing (5')		het	Herman (2012)		
	324	c.69277+1G>A		-		c.64489+1G>A		DCM		splicing (5')		het	Herman (2012)		
	326	c.70555_70556dupAT		p.Ile23519Ilefs*6		c.43628_43629dupAT		DCM		frameshift		het	Gerull (2002)		

Portion de TTN		All exon TTN		Nomenclature première publication		Phénotype		Type de mutation		Génotype		Référence	
Portion de TTN	Exon	Changement ADNc	Prédiction de l'effet sur la protéine	Nomenclature première publication	Phénotype	Type de mutation	Génotype	Référence					
A-band	326	c.70655delA	p.Glu23552Gluufs*8	p.Glu21956fs*	DCM	frameshift	het	Herman (2012)					
A-band	326	c.71845_71851delG	p.Ala23949*	p.Ala22335fs	DCM	non sense	het	Herman (2012)					
A-band	326	c.72533delT	p.Pro24178Profs*8	p.Pro22582fs*	DCM	frameshift	het	Herman (2012)					
A-band	326	c.77511_77527delinsAGA	p.Ser25837Serfs*3	p.Ser24241fs	DCM	frameshift	het	Herman (2012)					
A-band	326	c.81853C>T	p.Gln27285*	p.Gln25689*	DCM	non sense	het	Herman (2012)					
A-band	326	c.84684G>A	p.Trp28228*	p.Trp26632*	DCM	non sense	het	Herman (2012)					
A-band	326	c.85633C>T	p.Arg28545*	p.Arg265949*	DCM	non sense	het	Herman (2012)					
A-band	326	c.85834A>T	p.Lys28612*	p.Lys27016*	DCM	non sense	het	Herman (2012)					
A-band	326	c.85942insA	p.Ser28648Ileifs*1	p.Ser19628ifs*1	DCM	frameshift	het	Yoskovitz (2012)					
A-band	326	c.86324_86325delCT	p.Trp28743*	p.Trp27147*	DCM	non sense	het	Herman (2012)					
A-band	326	c.86686-2T>A	p.Ser28775Trpfs*49	p.Ser27179fs	DCM	frameshift	het	Herman (2012)					
A-band	326	c.87489C>A	-	c.81898-2T>A	DCM	splicing (5')	het	Herman (2012)					
A-band	328	c.87489C>A	p.Tyr29163*	p.Tyr27567*	DCM	non sense	het	Herman (2012)					
A-band	335	c.89765_89768delA	p.Asn2922Metfs*26	p.Tyr28326fs	DCM	frameshift	het	Herman (2012)					
A-band	335	c.89504delG	p.Glu2984Lysfs*11	p.Glu20963Lysfs*10	DCM	frameshift	het	Gerull (2006)					
A-band	339	c.92741G>A	p.Trp30914*	p.Trp29318*	DCM	frameshift	het	Herman (2012)					
A-band	339	c.93031C>T	p.Arg31011*	c.88242C>T	DCM	non sense	het	Herman (2012)					
A-band	339	c.93316G>T	p.Glu31106*	p.Glu29510*	DCM	non sense	het	Herman (2012)					
A-band	339	c.93968_93972delTTAAA	p.Ile29725Serfs*33	p.Trp29725fs	DCM	frameshift	het	Herman (2012)					
A-band	343	c.94999C>T	p.Cys31667Arg	p.Trp274375T>C	HMERF	non sense	het	Ohison (2012)					
A-band	343	c.95029C>T	p.Gln31677*	p.Gln30081*	DCM	non sense	het	Herman (2012)					
A-band	343	c.95281+2insCCT	-	c.90493+2insCCT	DCM	splicing (5')	het	Herman (2012)					
A-band	345	c.95831delA	p.Asn31944Asnfs*1	p.Asn30348fs	DCM	frameshift	het	Herman (2012)					
A-band	347	c.96325_96326insA	p.Trp32109Asnfs*38	p.Trp30513fs	DCM	frameshift	het	Herman (2012)					
A-band	349	c.97357+1G>C	-	c.92569+1G>C	DCM	splicing (5')	het	Herman (2012)					
A-band	352	c.98164_98165delA	p.Arg32722Glyfs*1	p.Arg31126fs	DCM	frameshift	het	Herman (2012)					
A-band	352	c.98371C>T	p.Pro33370Leu	p.Arg31195*	DCM	non sense	het	Herman (2012)					
A-band	354	c.98899A>T	p.Arg32791*	p.Lys31371*	DCM	non sense	het	Herman (2012)					
A-band	355	c.99730-2T>C	-	c.94942-2T>C	DCM	splicing (5')	het	Herman (2012)					
A-band	357	c.100109C>T	p.Pro33370Leu	p.Pro30847Leu	ARVC	missense	het	Taylor (2011)					
A-band	357	c.100310C>A	p.Ser33437*	p.Ser31841*	DCM	non sense	het	Herman (2012)					
M-line	358 (Mex1)	c.102136C>T	p.Arg34046Trp	p.Arg2791Trp (relatif to Tr regulatory tail)	HMERF	missense	het	Lange (2005)					
M-line	358 (Mex1)	c.103835G>A	p.Arg34592Gln	p.Arg25618Gln	DCM	missense	het	Matsumoto (2005)					
M-line	358 (Mex1)	c.105389_105396delACCAAGTG	p.His35130Hisfs*9	p.Arg385-289392delACCAAGTG	EOMFC	frameshift	hom	Carmignac (2007)					
M-line	360 (Mex3)	c.106532delA	p.Lys35511Asnfs*5	p.291394delA	EOMFC	frameshift	hom	Carmignac (2007)					
M-line	362 (Mex5)	c.107441T>C	p.Met35814Thr	p.Met33291Thr	ARVC	missense	comp het*	Taylor (2011)					
M-line	362 (Mex5)	c.107512delT	p.Ser35839Glnfs*10	p.292998delT	TMD	frameshift	het	Hackman (2008)					
M-line	363 (Mex6)	AAA	p.Glu3585delinsValLysGluLys	p.293269_293279delinsTGAAGAAGAAAA	TMD	indel	het	Hackman (2002)					
M-line	363 (Mex6)	c.107702A>C	p.His35901Pro	p.293269_293279delinsTGAAGAAGAAAA	LGMD2J	indel	hom	Udd (2005)					
M-line	363 (Mex6)	c.107705T>A	p.His35901Pro	p.His33378Pro	TMD	missense	het	Pollazon (2010)					
M-line	363 (Mex6)	c.107732T>C	p.Ile35902Asn	p.2933329T>A	TMD	missense	het	Van den Bergh (2003)					
M-line	363 (Mex6)	c.107732T>C	p.Ile35911Pro	p.293356T>C	TMD	missense	het	Hackman (2002)					
M-line	363 (Mex6)	c.107754delA	p.Lys35918Asnfs*9	p.293376delA	TMD	frameshift	het	Hackman (2008)					
M-line	363 (Mex6)	c.107755C>T	p.Gln35919*	p.293379C>T	TMD	non sense	het	Hackman (2008)					

## Annexe 4 : Famille K: illustration de l'importance de l'investigation clinique



### Présentation clinique :

Retard moteur

Rétractions multiples et sévères

Scoliose

Rigidité du rachis

Faiblesse des muscles axiaux

Evolution lente

Atteinte respiratoire : Capacité Vitale à 24%

(P1) et 55% (P3)

Pas d'atteinte cardiaque : ECG normal

### Analyses génétiques :

- Etude de liaison
  - 450 marqueurs microsatellites
  - Affymetrix Genechip® 250K, Nsp Array
- Stratégie de séquençage
  - Séquençage de gènes candidats positionnels
  - Agilent SureSelect All exon kit 50Mb, human whole exome

### Résultats :

Une analyse de liaison par cartographie par homozygotie, en utilisant le principe d'identité par descendance a été réalisée sur les résultats du panel de 450 marqueurs microsatellites. Cette analyse avait permis l'identification d'une alternance de loci potentiellement liés sur le chromosome 21 ainsi que la mise en évidence d'un locus majeur sur le chromosome 17. Ce locus s'étendait sur 4 Mb, la

cartographie fine réalisée par la suite avec le panel de 250 000 SNPs a révélée une recombinaison au sein du locus qui le sépare en deux loci de 2,4 Mb et de 1,6 Mb.

Ce dernier locus est en position télomérique et présente un LODscore proche de la liaison génétique ( $Z=2,92$ ) et qui équivaut au LODscore simulé de la famille (LODscore maximum qui peut être atteint dans une famille et qui est fonction de l'informativité de la famille). Ce dernier locus comporte 10 gènes dont le séquençage n'a révélé aucun polymorphisme.

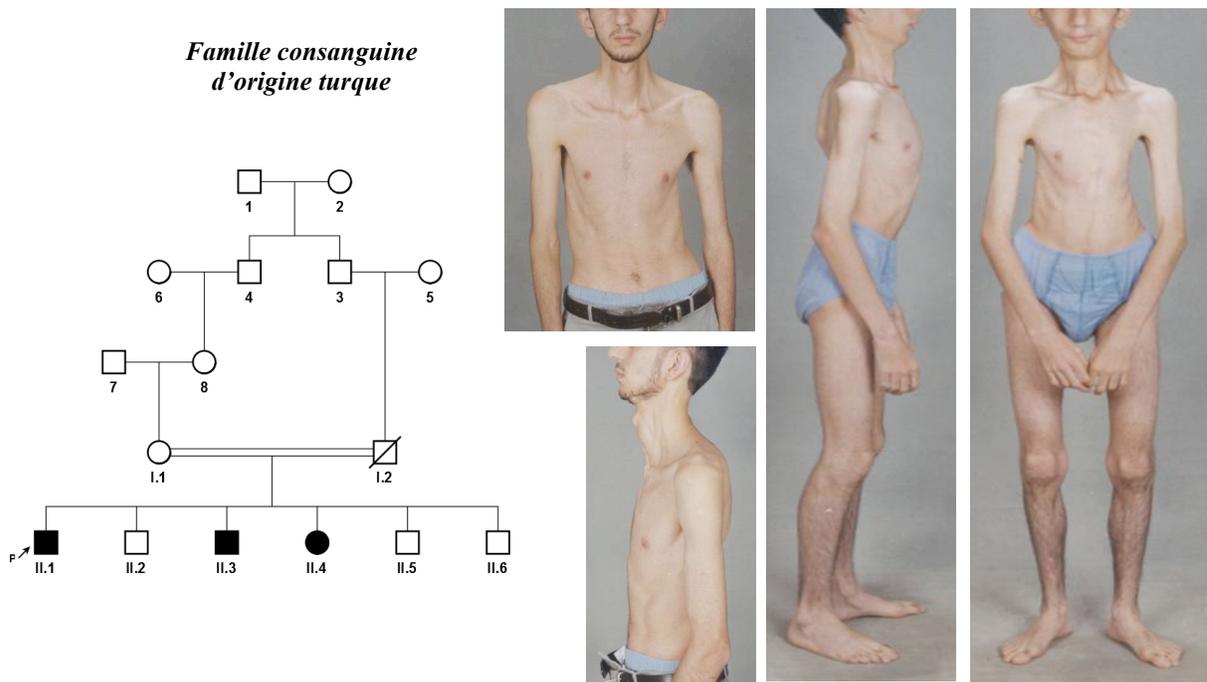
Chromosome	Bornes du locus	Taille du locus	Lod Score
Chr2	2.991.954 - 3.451.672	459.718	2.6
Chr4	186.059.928 - 186.109.172	49.244	2.96
Chr10	59.901.580 - 60.030.194	128.614	1.3
Chr17	126.531.427-128.942.336	2.401.909	1.8
Chr17	129.230.194-130.833.897	1.603.703	2.92
Chr 21	25.045.816 - 25.587.429	541.613	1.9
Chr 21	27.666.227 - 28.128.757	462.530	1.1
Chr 21	28.435.920 - 28.742.200	306.280	2

**Loci candidats potentiels identifiés chez la famille K (taille des loci données en kb)**

Les résultats de liaison sont inattendus pour cette famille, très informative. Les loci identifiés devraient être de plus grande taille. Il est à noter que parmi les trois enfants atteints, la sœur présente un phénotype à priori plus modéré que ses frères, mais que pour des raisons culturelles elle n'a pas pu être rigoureusement examinée. De plus, nous n'avons pas pu confirmer avec certitude l'exactitude du pedigree.

**En conclusion, cette famille illustre parfaitement l'importance d'une bonne investigation clinique en préalable des études génétiques et spécialement dans le cas d'un phénotype atypique ou d'une nouvelle manifestation clinique.**

## Annexe 5 : Famille G : Illustration de l'importance des collaborations scientifiques



### Présentation clinique :

Développement moteur normal

Rétractions

Scoliose

Légère dysphagie

Atteinte respiratoire

Pas d'atteinte cardiaque : ECG normal

### Analyses génétiques :

- Etude de liaison
  - 450 marqueurs microsatellites
- Affymetrix Genechip® 250K, Nsp Array
- Stratégie de séquençage
  - Séquençage du gène *TTN*
  - Agilent SureSelect All exon kit 50Mb, human whole exome

### Liaison majeure à un locus du chromosome 2 (22 Mb, 69 gènes)

Parmi les 69 gènes du locus, *TTN* était le gène candidat fonctionnel le plus évident. J'ai donc séquencé les produits de PCR d'amplification des exons de *TTN* codant la bande M ainsi que la ligne Z et la jonction ligne Z bande I, sans trouver de mutations. PEVK étant un domaine capital à la fonction du muscle strié j'ai poursuivi l'analyse du gène par le séquençage de 101 exons du domaine élastique PEVK, ce qui m'a permis d'identifier une substitution (NM\_133378.3 c.G31563C) touchant

un résidu glutamate très conservé durant l'évolution au niveau inter et intra espèces (NP\_596863.3 p.E10447Q) et absent chez 120 individus contrôles caucasiens non apparentés et 98 individus contrôles d'origine turque.

L'analyse de l'exome par séquençage de nouvelle génération a confirmé l'absence de mutations dans les 68 autres gènes du locus initial. Validant ainsi l'implication de *TTN* dans la pathologie des patients. En revanche cette analyse a également identifié 2 nouvelles substitutions (C>Y) de *TTN* présent dans des domaines Ig des exons 83 et 305, non retrouvées chez 95 témoins caucasiens et 96 témoins turques. Ces deux derniers variants sont prédit non pathogènes par analyses *in-silico* (polyphen, SIFT...), le variant PEVK en revanche est associé à un effet délétère. Cependant, la validation de la pathogénicité de la mutation identifiée nécessite l'utilisation d'outils spécifique, qui relèvent plus des techniques de biophysique, afin de mesurer l'effet de cette substitution sur la capacité contractile du domaine PEVK. Ce travail d'analyse sera donc mené en collaboration avec le Dr Michael Gotthardt (Charité University, Berlin) qui possède une expertise dans le domaine.

**En conclusion nous avons associé la TTN à un nouveau phénotype myopathique congénital avec rétractions, qui est la première description phénotypique de mutation du domaine PEVK chez l'homme. Cette famille illustre donc l'importance du travail collaboratif entre équipe pour accéder aux différentes expertises permettant de mener à bien les travaux de recherche**

## Homozygous or compound heterozygous truncating mutations of *TTN*, from exome analysis, define novel forms of cardiomyopathy with skeletal myopathy

C. Chauveau<sup>1,2</sup>, C. Julien<sup>2</sup>, H. Marks<sup>3</sup>, R. Foley<sup>4</sup>, A. L. Kho<sup>5</sup>, B. Talim<sup>6</sup>, MC. Arne-Bes<sup>7</sup>, E. Uro-Coste<sup>7</sup>, P. Maury<sup>7</sup>, A. Vihola<sup>8</sup>, B. Udd<sup>8</sup>, H. Topaloglu<sup>6</sup>, S. Moore<sup>9</sup>, M. Gautel<sup>5</sup>, C. Bonnemann<sup>4</sup>, ME. Samuels<sup>2</sup>, A. Ferreiro<sup>1,10</sup>

<sup>1</sup>Groupe Myologie, U787 INSERM-UPMC, Paris; <sup>2</sup>Centre de Recherche de l'Hôpital Ste-Justine, UdeM, Montréal; <sup>3</sup>St. Christopher's Hospital for Children, Drexel University College of Medicine, Philadelphia; <sup>4</sup>National Institutes of Health, Bethesda; <sup>5</sup>Cardiovascular Division, King's College London, London; <sup>6</sup>Faculty of Medicine, Hacettepe University, Hacettepe; <sup>7</sup>CHU Rangueil, Toulouse; <sup>8</sup>Folkhälsan Institute, University of Helsinki; <sup>9</sup>University of Iowa; <sup>10</sup>Centre de Référence des maladies neuromusculaires, GH Pitié-Salpêtrière, Paris

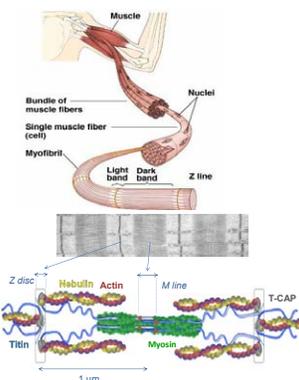
### ABSTRACT

Cardiomyopathies are potentially lethal conditions in which the muscular function of the myocardium is altered. They are frequent and can be associated with skeletal muscle disease. We describe the genetic analysis focusing on the *titin* gene of 24 families with primary cardiomyopathy associated with congenital core myopathy. Titin plays a key structural and functional role in striated muscles. However, *TTN*'s size and complexity preclude full screening by Sanger sequencing, and until recently only a small number of mutations had been found. Our results expand the spectrum of M-line recessive truncating mutations and the spectrum of cardiac phenotypes. Since mutations in the titin M-line domain have been associated with cardiac and skeletal muscle phenotypes, we started with Sanger sequencing of the M line-encoding region encompassed by the last 6 exons of *TTN* (6.5% of the gene). We identified 5 novel mutations in 4 families (16.6% of our series). 2 heterozygous changes were inherited from healthy fathers. Subsequently we performed whole exome sequencing (WES) on those last two families and identified in each one new heterozygous mutation (missense and truncating). All 8 parents were heterozygous and healthy, suggesting that not all truncating *TTN* mutations, particularly if in the M-line, manifest unless associated with a second mutation. The range of cardiological manifestations in our cohort was large and included previously unreported phenotypes. Our results expand the spectrum of *TTN*-related conditions and confirm that WES opens a new era for the study of giant and complex genes, of which *TTN* is the paradigm.

### INTRODUCTION

#### TITIN IN SKELETAL AND CARDIAC MUSCLE CONDITIONS

##### • Muscle structure



##### • Titin = largest protein in nature

Gene: 363 exons  
mRNA: 101 519 bp  
Protein: 33 423 aa, 3 676 kD

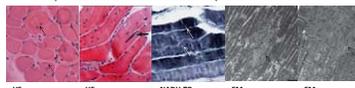
##### • *TTN* variants repartition



##### • Only one phenotype combining cardiac and muscular involvement

Early onset myopathy with lethal cardiomyopathy → 2 recessive mutations in the M-line: 1-BP DEL 291297A, 8-BP DEL 289385 ACCAAGT (Carmignac et al, AnnNeur, 2007)

##### • Congenital myopathies with characteristic anatomopathologic description: area of mitochondrial depletion = cores



### ANALYSIS

#### EXPERIMENT

##### Patients: 24, mostly sporadic cases

- Congenital myopathy with multi-minicores
- Various cardiac defects

##### *TTN* screen: Sanger sequencing and / or NGS

- Traditional sequencing of the last 6 C-ter exons of *TTN* encoding the M-line
- Whole exome sequencing (WES) in 2 patients with only one mutation (families 2 and 4)

##### WES: Agilent SureSelect Human All Exon 50 Mb (Illumina) / SOLiD4

- One pair ended sequencing
- Analysis with NextGene (Softgenetics)
- Alignment on GRCh37/hg19 assembly
- Filtering against 1000 genomes, NCBI database, CHU Ste Justine internal database

#### RESULTS

##### WES statistics

- Mean reads length: 48 nucleotides
- Mean coverage: 26X

##### *TTN* Analysis parameters

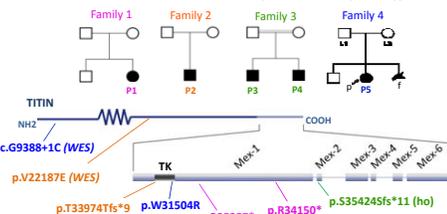
- Search for heterozygous variants
- Minimum coverage = 6X
- Variant minimum frequency = 25%

##### WES Results

- New unreported *TTN* variants per patient = 9
- New unreported variants in the whole exome ≈ 1500

M-line *TTN* (7% of the gene) → mutations in 4 families (16.6% of our series)

##### + Mutations in new *TTN* domains, expansion of the phenotypes



All mutations were carried at the heterozygous state by parents who were healthy after expansive pathological and morphological examination

### CONCLUSION

We identified new mutations in *TTN* with autosomic recessive transmission in patients with: Core myopathy + contractures + large spectrum of heart disease

MmD + Defects of atrial and ventricular septation (P1)

MmD + late/mild dilated (P2, P3) or childhood-onset severe dilated cardiomyopathy

Arthrogryposis + MmD+ Left Ventricular Noncompaction (P5)



M-line



M-line +/- Fn3 domain (I band)



Titin kinase + Ig Domain (I band)

Finally we define a new medical entity :

**Myopathy with Cores, Autosomic Recessive with Heart Disease (MCARHD)**

We suggest that compound heterozygosity of non-symptomatic *TTN* truncated alleles and missense mutations are likely to underly an unexpectedly large spectrum and high percentage of unexplained early- and late-onset myopathies and cardiac conditions. We thus propose that *TTN* should be considered and tested in any uncharacterized cardiac and/or skeletal muscle condition.



**Annexe 7 : Core myopathy, arthrogryposis and ventricular noncompaction: a novel titinopathy**  
(Soumis dans *Annals of Neurology*)

**Core myopathy, arthrogryposis and ventricular non-compaction: a novel titinopathy.**

Claire Chauveau, PhD<sup>§,1,2,3</sup>, Ay Lin Kho, PhD<sup>§,4</sup>, Carsten Bonnemann, MD<sup>5, 6</sup>, Alexander Alexandrovich, PhD<sup>4</sup>, Anna Vihola, PhD<sup>7</sup>, Cedric Julien PhD<sup>#1,2</sup>, Beth Kaufmann, MD<sup>8</sup>, Livija Medne, MS<sup>8,9</sup>, A. Reghan Foley, MD<sup>#6</sup>, Mariarita Santi, MD<sup>10</sup>, Bjarne Udd, MD, PhD<sup>7,11,12</sup>, Mark Samuels, PhD<sup>3,13</sup>, Mathias Gautel, MD, PhD<sup>4</sup> and Ana Ferreira, MD, PhD<sup>1,2,14,15</sup>

1. Inserm, U787 Myology group, Institut de Myologie, Groupe Hospitalier Pitié-Salpêtrière, Paris, France
2. UPMC, UMR787, Paris, France
3. Centre de Recherche de l'Hôpital Ste-Justine, Université de Montréal, Montréal, Canada
4. King's College London, BHF Centre of Research Excellence, Cardiovascular Division and Randall Division for Cell and Molecular Biophysics, London, United Kingdom
5. National Institutes of Health, Bethesda, MD, USA
6. Division of Neurology, The Children's Hospital of Philadelphia, University of Pennsylvania, Philadelphia, USA
7. Folkhälsan Institute of Genetics and Department of Medical Genetics, Haartman Institute, University of Helsinki, Helsinki, Finland
8. Division of Cardiology, The Children's Hospital of Philadelphia, University of Pennsylvania, Philadelphia, USA
9. Division of Human Genetics, The Children's Hospital of Philadelphia, University of Pennsylvania, Philadelphia, USA
10. Division of Neuropathology, The Children's Hospital of Philadelphia, University of Pennsylvania, Philadelphia, USA
11. Neuromuscular Research Center, Tampere University and University Hospital, Tampere, Helsinki, Finland
12. Department of Neurology, Vaasa Central Hospital, Vaasa, Finland
13. Department of Medicine, University of Montreal, Montreal, Quebec, Canada
14. AP-HP, Centre de Référence Maladies Neuromusculaires Paris-Est, Groupe Hospitalier Pitié-Salpêtrière, Paris, F-75013, France
15. AP-HP, Service de Pédiatrie, Centre de Référence Maladies Neuromusculaires (GNMH) Hôpital Raymond Poincaré, 92380 Garches, France

§ Both authors contributed equally to this work

# Current affiliations: Centre de Recherche de l'Hôpital Ste-Justine, Université de Montréal, Montréal, Canada; and Dubowitz Neuromuscular Centre, University College London Institute of Child Health and Great Ormond Street Hospital for Children, London, UK

Running title: A new titinopathy (character count: 17)  
Title character count: 81 (spaces included)  
Word count: abstract 97, body of the manuscript 1496  
Number of figures: 3, color figures: 3, tables: 0

## ABSTRACT

Multi-minicore disease (MmD) is a heterogeneous congenital myopathy. Recessive mutations in *SEPN1*, *RYR1* or *MEGF1* account for only a fraction of MmD cases. Here we describe a new, antenatal-onset form of MmD presenting with arthrogryposis and left ventricular non-compaction. Using whole exome sequencing, we establish that this phenotype is due to novel compoundheterozygous mutations of *TTN* (truncating and missense) resulting in the first-reported absence of a functional titin kinase (TK) in humans. Moreover, we demonstrate a key role of TK in protein turnover of human muscle, pointing to a new pathway and disease mechanism in titinopathies.

## INTRODUCTION

Congenital myopathies are inherited muscle disorders defined by particular architectural changes of the muscle fibers. They typically present with delayed motor development, generalized muscle weakness and impaired ambulation, sometimes associated with scoliosis, respiratory failure and early mortality. Rare CM patients develop cardiomyopathy in childhood or adulthood, but no congenital heart defect has so-far been reported.

Multi-minicore disease (MmD) is a recessive CM histopathologically defined by small areas of mitochondria depletion and sarcomere disorganization (minicores) in muscle fibers. MmD is clinically heterogeneous, with 4 distinct phenotypes described<sup>1</sup>. Mutations in *RYR1*<sup>2,3</sup>, *SEPN1*<sup>4</sup> and *MEGF10*<sup>5</sup> have been associated with MmD, but for many cases the responsible gene remains unknown.

Titin (TTN), encoded by the 363-exon *TTN* gene, is a giant sarcomeric protein. Due to its size, analysis of the whole gene was unfeasible prior to next-generation sequencing (NGS). Therefore, relatively few *TTN* mutations have been reported, mostly heterozygous and associated with adult-onset conditions involving either cardiac or skeletal muscles<sup>6-8</sup>. Homozygous truncating mutations cause the only described congenital titinopathy, a myopathy (with minicores, central nuclei and dystrophic lesions) with childhood-onset dilated cardiomyopathy<sup>9</sup>. TTN's multiple domains play crucial roles in muscle development, structure, elasticity and signaling, many of which remain to be characterized<sup>10</sup>. The TTN kinase domain (TK) is emerging as a pivotal player in the control of muscle gene expression and protein turnover<sup>11</sup>. The only disease associated with TK is an adult-onset dominant myopathy with early respiratory weakness but no cardiac involvement (HMERF)<sup>11, 12</sup>.

## SUBJECTS AND METHODS

Full details are provided in Supplementary Methods.

### *Patient and controls*

Written informed consent was obtained from the patient's parents as well as from the control subjects under NIH protocol 12-N-0095.

### *Genetic studies*

Standard procedures were used to extract genomic DNA from blood samples and mRNA from muscle samples and from fetal tissues obtained after pregnancy termination following detection of two *TTN* mutations in the fetus. *TTN* sequencing was performed on patient genomic DNA using PCR-based sequencing by Genoscreen (Mex 1 to 6) followed by Whole Exome Sequencing using the SureSelect Human All Exon 50Mb Kit (Agilent technologies) and a SOLID4 Next Generation Sequencer (Applied Biosystems). Variants were filtered by comparison with dbSNP, the 1000Genomes dataset, and other exome data.

### *Histological and immunohistochemical studies*

Samples were obtained from the proband at 15 months (diagnostic skeletal muscle biopsy) and 4 years (post-explantation myocardial sample) and were frozen and processed for standard histochemical stainings. Immunostaining and confocal microscopy were done as reported<sup>9,13</sup>.

### *Protein and in vitro studies*

*TTN* Western blot was performed on frozen heart muscle sections from patient and a control as reported<sup>14</sup>. Expression of recombinant TK, circular dichroism spectroscopy and yeast genetic interaction screen were performed as described using a construct comprising the autoinhibited TK domain and its preceding fibronectin-like domain (Fn-TK)<sup>11,15</sup>.

## **RESULTS**

### *A novel MmD phenotype*

We ascertained a female pediatric patient presenting with a previously-unreported phenotype, associating arthrogryposis multiplex congenita (AMC), kyphosis, congenital muscle weakness and cardiac failure. Muscle biopsy was consistent with MmD (Figure 1). Echocardiography revealed a ventricular septal defect and a trabeculated myocardial appearance typical of left ventricular non-compaction (LVNC). Terminal heart failure required transplantation at 4 years of age, with surgical complications resulting in spinal cord infarction and paraplegia. Arthrogryposis and muscle weakness remain otherwise stable.

Clinical and cardiological examinations (echocardiography, ECG) of the healthy, non-consanguineous Caucasian parents at ages 38 and 42 revealed no significant abnormalities. There was no history of cardiomyopathy or sudden cardiac death in either parental family.

### *Classic and NGS genetic analysis: a recessive titinopathy*

Sequencing of potentially-causative genes including *ACTA1*, *LAMP2*, *MYBPC3*, *MHY7*, *MYL2*, *MYL3*, *PRKAG2*, *TNNT2*, *TPM1*, *RYR1* and *LMNA* was negative. Based on the skeletal and heart muscle involvement and the presence of multi-minicores<sup>9</sup>, we sequenced the last 6 *TTN* exons, encoding the C-terminal M-line region (Mex1-Mex6), in the patient's DNA and identified a heterozygous variant in Mex1 (c.T94734C) causing a missense change in a highly conserved residue of the TK domain (p.W31504R, or p.W260R in the structure of TK)<sup>16</sup>. This variant was absent from 120 control samples, but was present at the heterozygous state in the healthy father. Suspecting recessiveness, we performed whole exome sequencing of the patient's DNA. This revealed two private

compound heterozygous *TTN* mutations, the previously identified p.W260R change and a single nucleotide deletion (c.G9388+1C) in the splice donor site at 5' of exon 38, also carried by the healthy mother. *In silico* analysis predicted abnormal splicing and generation of a premature termination codon in exon 39 (Figure 1j,l). cDNA sequencing from the patient's myocardium confirmed the coexistence of a wild-type and a shorter transcript corresponding to truncation of the protein encoded by the maternal allele after 2950 N-terminal amino acids (Figure 1k,m), therefore devoid of TK.

#### *Protein studies: Z-disk and autophagy changes*

Western-blots showed no major reduction of the global amount of TTN in the patient's samples. Immunolabelling of myocardium cryosections using antibodies directed against the TTN N-terminus (Z-disk) and C-terminus (M-band) showed in most cells an alternating striated pattern, demonstrating the expression and integration of a full-length protein presumably encoded by the paternal TK-mutant allele into the sarcomere (Figure 1i).

Electron microscopy also revealed areas with normal sarcomeres, suggesting that the mutations do not impair their ordered assembly. However, many subcellular areas showed complete sarcomere disorganization (Figure 1e) with aggregating electron-dense Z-disk material, reminiscent of those in myofibrillar myopathies (MFM) caused by mutations in Z-disk proteins linked to defective protein turnover<sup>17</sup>. Immunohistochemical studies of M-band and Z-disk proteins showed normal striated patterns for M-band titin and its ligand myomesin, with only occasional areas of disruption. In contrast, staining for the Z-disk alpha-actinin revealed many abnormal areas similar to the Z-disk streaming in MFM (Figure 2). Western blots showed moderate changes in the levels of the autophagy-adaptor and TK-ligand Nbr1 and the Nbr1-ligand SQSTM1<sup>11</sup>, while the levels of LC3-1<sup>18</sup> were increased, indicating a perturbation of normal autophagy (Figure 3d-e). These changes suggest a significant disruption of components of the autophagy protein quality-control machinery and the homeostasis of Z-disk associated proteins.

#### *In vitro studies: the first TK KO*

Biochemical studies revealed that p.W260R TK could be expressed in insect cells, was fully soluble and showed no aggregation, as demonstrated by normal elution profiles on gel filtration chromatography (not shown). Analysis of thermal stability by circular dichroism spectroscopy showed a single major structural transition with loss of helical structure over 60°C in WT TK, while TK p.W260R already showed partial secondary structure loss above 35°C (Figure 3B). We could also not detect catalytic activity of the mutant enzyme.

Genetic interaction assays showed that, while wild-type (WT) TK bound to Nbr1 and to its own autoinhibitory tail in *trans*, both interactions were completely abrogated in the TK p.W260R mutant. In contrast, the previously-documented HMERF p.R279W mutant<sup>11</sup> interacted with its autoinhibitory tail while losing the Nbr1 interactions (Figure 3C).

These results together suggest that p.W260R mutation causes a complete loss-of-function of TK that abrogates M-band and Z-disk linked pathways via Nbr1, both of which are implicated in myofibril turnover and involve the autophagy pathway.

## **DISCUSSION**

Left-ventricular non compaction-syndrome (LVNC) is a recently-described heart defect due to intrauterine arrest of myocardial development and abnormal compaction of the cardiac tissue<sup>19</sup>. LVNC is often X-linked or dominantly inherited<sup>20</sup>; recessive inheritance has been suggested but never

confirmed<sup>21</sup>. Similar to AMC, most LVNC cases remain genetically uncharacterized.

Here we describe a new form of MmD associating AMC and LVNC, and show that it is caused by compound heterozygous mutations of *TTN*. This enlarges the mutational and phenotypic spectrum of *TTN* to include antenatal-onset conditions. It also represents the first recessive form of LVNC and adds *TTN* as a novel candidate gene for LVNC and AMC. Moreover, we report the first complete loss of a functional TK domain in humans and reveal a novel role of the TK in communication between Z-disk, M-band and the autophagy machinery. Our findings thus propose a new pathway and mechanism for disease in titinopathies and suggest that TK might be indispensable for normal cardiac morphogenesis and skeletal muscle maintenance during human development.

This is also the first description of *TTN* compound heterozygosity leading to phenotypical expression of a recessive missense mutation on the background of a recessive truncating mutation. Absence of phenotype in the heterozygous mother confirms that not all truncating *TTN* mutations manifest unless associated with a second mutation<sup>9</sup>. This can potentially explain the recently reported evidence that carriers of truncating *TTN* mutations may amount up to 3% of the population<sup>6</sup> and stresses the importance of careful family studies to establish pathogenicity of nonsense mutations identified by NGS and to exclude compound heterozygosity.

NGS will probably bring about a dramatic expansion of titinopathies. We propose that *TTN*, particularly compound heterozygosity of non-symptomatic *TTN* truncating and missense mutations, is likely to underlie a large spectrum and proportion of currently undiagnosed myopathies and cardiac conditions, this report representing the first of what we think will be a plethora of novel *TTN*-associated phenotypes.

### **Supplemental data**

Supplemental Methods and References.

### **Acknowledgments**

We are grateful to the family for participation in this study. MG holds the BHF Chair of Molecular Cardiology. This study was supported by the MRC (AYK, AA, MG), the INSERM, the Leducq Foundation, the AFM and the AP-HP (Contrat d'Interface to AF). CC is the recipient of a Frontenac grant. We are grateful to Tam T.T. Bui and Alex Drake at the Biomolecular Spectroscopy Centre, King's College London, for assistance with CD spectroscopy.

### **Web Resources**

The URLs for data presented here are:  
ESE Finder, <http://rulai.cshl.edu/tools/ESE2/>

### **Accession Numbers**

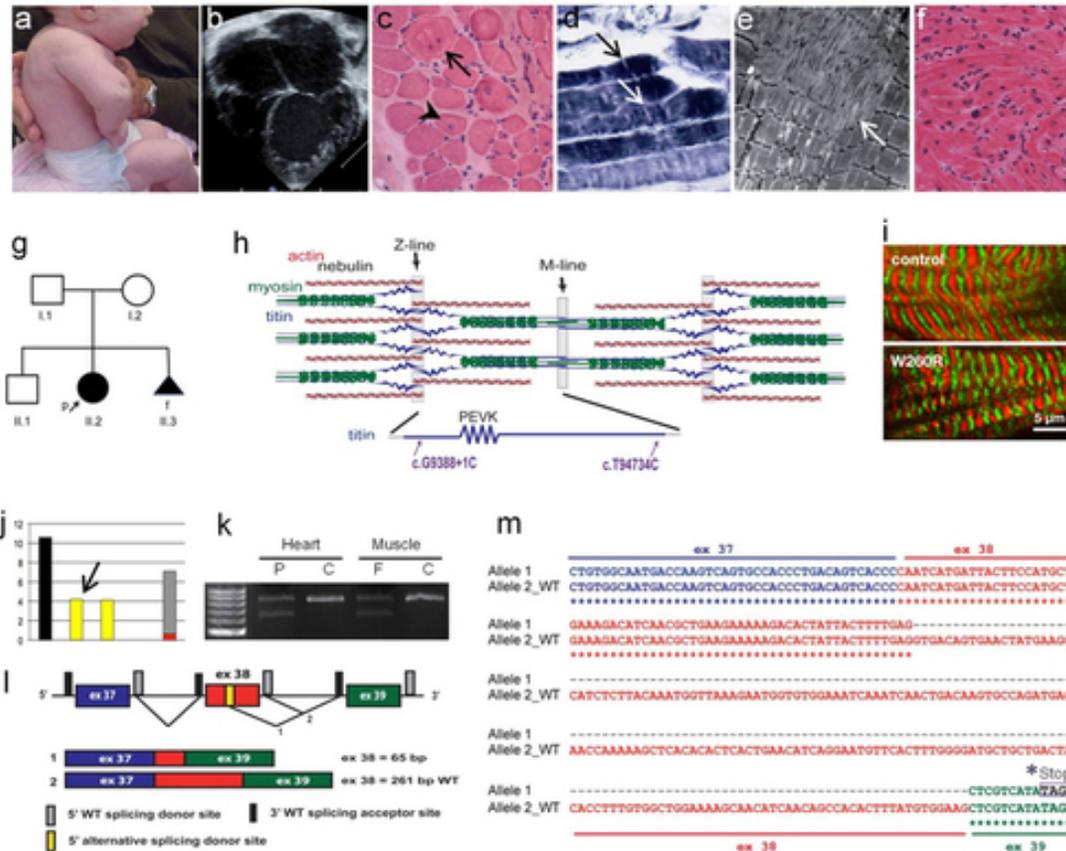
The NCBI accession numbers for the *TTN* sequences reported in this paper are NM\_133378 and NP\_596869

### **References**

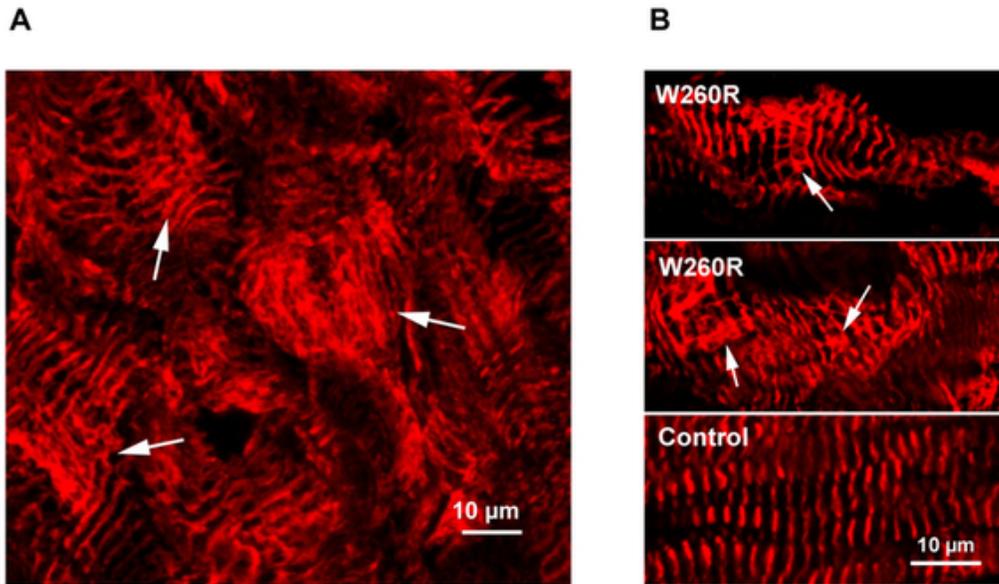
1. Ferreiro A, Estournet B, Chateau D et al. Multi-minicore disease--searching for boundaries: phenotype analysis of 38 cases. *Ann Neurol.* 2000;48:745-757
2. Ferreiro A, Monnier N, Romero NB et al. A recessive form of central core disease, transiently presenting as multi-minicore disease, is associated with a homozygous mutation in the ryanodine receptor type 1 gene. *Ann Neurol.* 2002;51:750-759

3. Jungbluth H, Müller C, Brockington M et al. Autosomal-recessive inheritance of RYR1 mutations in a congenital myopathy with cores. *Neurology*. 2002;59:284-287 [PubMed](#)
4. Ferreiro A, Quijano-Roy S, Pichereau C et al. Mutations of the selenoprotein N gene, which is implicated in rigid spine muscular dystrophy, cause the classical phenotype of multimincore disease: reassessing the nosology of early-onset myopathies. *Am J Hum Genet*. 2002;71:739-749 [PubMed](#)
5. Boyden SE, Mahoney LJ, Kawahara G et al. Mutations in the satellite cell gene MEGF10 cause a recessive congenital myopathy with minicores. *Neurogenetics*. 2012;13:115-124
6. Herman DS, Lam L, Taylor MR et al. Truncations of titin causing dilated cardiomyopathy. *N Engl J Med*. 2012;366:619-628 [PubMed](#)
7. Hackman P, Marchand S, Sarparanta J et al. Truncating mutations in C-terminal titin may cause more severe tibial muscular dystrophy (TMD). *Neuromuscul Disord*. 2008;18:922-928
8. Nowak KJ. Trusting new age weapons to tackle titin. *Brain*. 2012;135:1665-1667 [PubMed](#)
9. Carmignac V, Salih MA, Quijano-Roy S et al. C-terminal titin deletions cause a novel early-onset myopathy with fatal cardiomyopathy. *Ann Neurol*. 2007;61:340-351
10. Kruger M, Linke WA. The giant protein titin: a regulatory node that integrates myocyte signaling pathways. *J Biol Chem*. 2011;286:9905-9912 [PubMed](#)
11. Lange S, Xiang F, Yakovenko A et al. The kinase domain of titin controls muscle gene expression and protein turnover. *Science*. 2005;308:1599-1603 [PubMed](#)
12. Ohlsson M, Hedberg C, Bradvik B et al. Hereditary myopathy with early respiratory failure associated with a mutation in A-band titin. *Brain*. 2012;135:1682-1694 [PubMed](#)
13. Fukuzawa A, Lange S, Holt M et al. Interactions with titin and myomesin target obscurin and obscurin-like 1 to the M-band: implications for hereditary myopathies. *J Cell Sci*. 2008;121:1841-1851 [PubMed](#)
14. Greaser ML, Warren CM. Efficient electroblotting of very large proteins using a vertical agarose electrophoresis system. *Methods Mol Biol*. 2009;536:221-227 [PubMed](#)
15. Politou AS, Gautel M, Pfuhl M et al. Immunoglobulin-type domains of titin: same fold, different stability? *Biochemistry*. 1994;33:4730-4737 [PubMed](#)
16. Mayans O, van der Ven PF, Wilm M et al. Structural basis for activation of the titin kinase domain during myofibrillogenesis. *Nature*. 1998;395:863-869
17. Selcen D, Bromberg MB, Chin SS, Engel AG. Reducing bodies and myofibrillar myopathy features in FHL1 muscular dystrophy. *Neurology*. 2011;77:1951-1959 [PubMed](#)
18. Kliensky DJ, Abdalla FC, Abeliovich H et al. Guidelines for the use and interpretation of assays for monitoring autophagy. *Autophagy*. 2012;8:445-544 [PubMed](#)
19. Oechslin E, Jenni R. Left ventricular non-compaction revisited: a distinct phenotype with genetic heterogeneity? *Eur Heart J*. 2011;32:1446-1456 [PubMed](#)
20. Sasse-Klaassen S, Gerull B, Oechslin E et al. Isolated noncompaction of the left ventricular myocardium in the adult is an autosomal dominant disorder in the majority of patients. *Am J Med Genet A*. 2003;119A:162-167
21. Digilio MC, Marino B, Bevilacqua M et al. Genetic heterogeneity of isolated noncompaction of the left ventricular myocardium. *Am J Med Genet*. 1999;85:90-91 [PubMed](#)

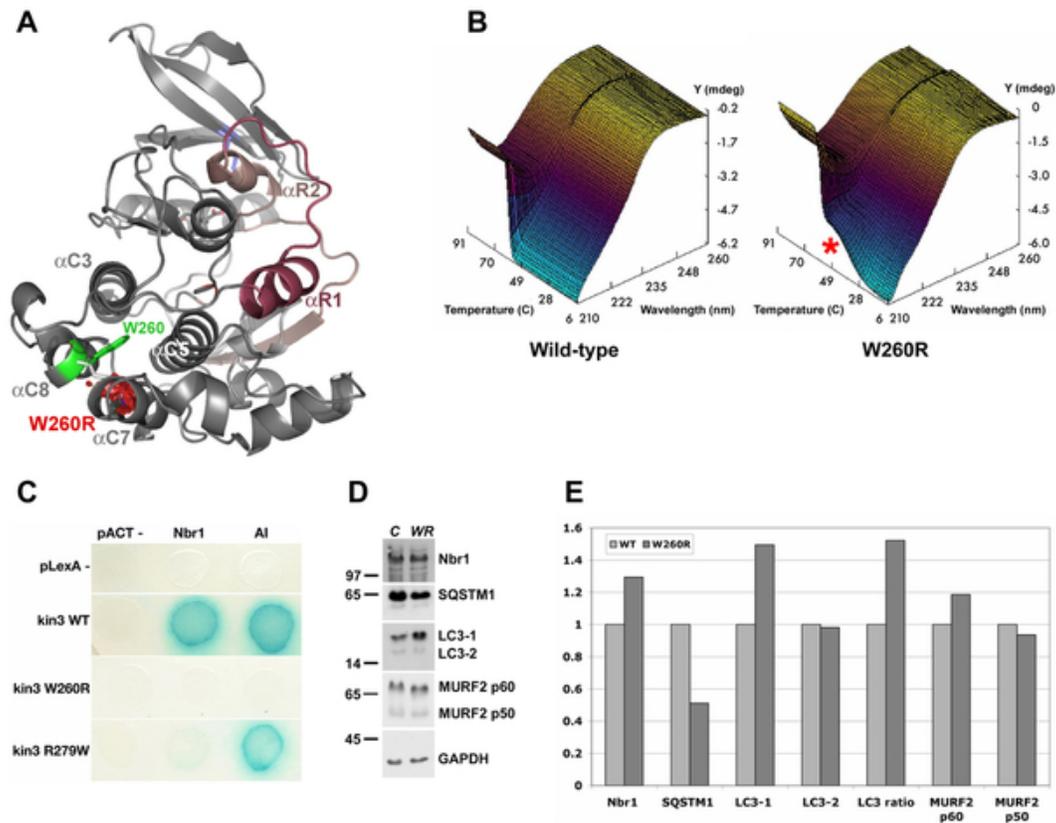
## FIGURES



**Figure 1: Phenotype and *TTN* mutations.** (a): Shoulder, elbow, wrist, hand and lower limb contractures, short neck, low-set ears, retrognathia and kyphoscoliosis at 3 months of age. (b): Echocardiography (four-chamber view) showed severe LVNC. (c-e): Skeletal muscle histology. Transverse cryosections stained with hematoxylin-eosin (HE, c, 60x) showed endomysial fibrosis, nuclear centralisations (arrowhead) and occasional ring-like fibers with multiple non-peripheral nuclei (arrow). There was type 1 fiber predominance, and NADH-TR-stained longitudinal sections (d) and electron microscopy (e) revealed multiple circumscribed areas of mitochondria depletion and sarcomere disorganization (minicores, arrows). (f): HE-stained cryosections of the left ventricle apex showed myocyte hypertrophy and fibrosis (40x). (g): Family pedigree (p: proband; f: female; triangular symbol: terminated pregnancy). (h): Titin protein (bottom) and sarcomere position (top). Mutations are indicated by arrows. (i): Titin immunolabeling on longitudinal cryosections from control (top) and patient (bottom) myocardium with antibodies recognizing titin epitopes upstream (T12, red) and downstream (M8, green) the TK domain. (j-m): Splice site mutation analysis. ESE-finder (j) predicted a major loss of strength of the mutant donor site (down to 7,75%, in red) and presence of two putative alternative cryptic donor sites in exon 38 (yellow). Of these, one (j, arrow and l) was actually used in the patient, leading to the coexistence of wild-type mRNA and a shorter transcript, as demonstrated by cDNA amplification (k) and sequencing (m) of *TTN* exon 38. P, C and F: patient, control and affected fetus respectively.



**Figure 2: Cellular characterization of the p.W260R mutant.** Z-disk disruption and formation of alpha-actinin rich aggregates and “streaming” in patient striated muscle. Confocal images of patient cardiac explant sections (W260R) and control tissue immunolabelled for sarcomeric alpha-actinin. **(a)** Note the abundant aggregates and **(b)** strands of alpha-actinin linking adjacent Z-disks (arrows) resembling Z-disc streaming, and areas where regular Z-disk striation patterns are completely substituted with large areas of aggregated alpha-actinin positive deposits. In contrast, control muscles show normal regular striations with  $\approx 2 \mu\text{m}$  sarcomere spacing. Scale bars:  $10 \mu\text{m}$



**Figure 3: Biochemical characterization of the p.W260R mutant.** (a) Molecular model of TK (based on PDB entry 1TKI) with the regulatory tail shown in red. Important secondary structure elements are labeled. Note that the mutated tryptophane residue, W260, occupies a hydrophobic pocket formed by alpha-helices C3, C5, C7 and C8. Mutation to arginine replaces this conserved hydrophobic residue with a large hydrophilic residue and results also in a predicted steric clash with alphaC7 (red volume), likely destabilizing part of the catalytic domain. (b) The TK p.W260R mutant shows reduced stability. CD spectra recorded from 200-260 nm over the temperature range from 6°C to 94°C reveal a single sharp secondary structure transition in wild-type TK with loss of negative ellipticity (Y) at 209 nm (Z) around 59°C (X). In contrast, the p.W260R mutant shows partial melting (red asterisk) around 42°C followed by a second transition at 57°C. (c) Mutations in TK abrogate Nbr1 binding. Results of yeast-two-hybrid reporter gene assay (growth on histidine-free media and activation of beta-galactosidase activity, blue color reaction) are shown. Wild-type TK interacts with its own autoinhibitory tail (AI) *in trans*, as well as with Nbr1. These interactions are abrogated in the p.W260R mutant, while p.R279W mutant kinase retains the ability to interact with its own AI *in trans*. This suggests that the recessive p.W260R mutant is a complete loss-of-function mutation. (d) Western blot analysis of p.W260R patient heart muscle reveals changes in proteins of the autophagy machinery with an increase of Nbr1, a decrease of SQSTM1, and increase of unlipidated LC3-1 levels. No significant change was found for MURF2, a SQSTM1 and TTN ligand linked to proteasomal protein turnover. Numbers indicate positions of relevant molecular weight markers (kDa) GAPDH: glyceraldehyde 3-phosphate dehydrogenase as internal standard. (e) Quantitation of (d) normalized to GAPDH-levels.

## XI. Curriculum vitæ

### **PhD in Human Genetics and Biomedical Science with strong background in molecular /cellular biology and biochemistry**

- Expertise in high throughput sequencing and genotyping data analysis
- Proficient in rigorous analysis of experimental data
- Competent in using web-open source applications to analyze nucleotidic sequences
- Efficient in adapting new methods and biological tools required for complex data analysis
- Acquired 5 years laboratory experiences in cellular and molecular biology research

### **Strong communication and collaboration skills**

- Presentation in international meetings,
- Collaboration with colleagues to produce peer-reviewed publications
- Successful management of multiple independent research projects
- Mentoring undergraduate students
- Proven critical thinking and ability to work with bioinformaticians and computer science engineers
- Interaction with MD and clinical staffs to coordinate patient information collection and research

### Education

---

- **PhD in Human Genetics** (Oct 2008-Sep 2012), UPMC-Paris 6 University, Paris, France
- **PhD in Biomedical Science** (Oct 2009-Sep 2012), Montréal University, Québec, Canada
- **MSc in Molecular and Cellular Biology**, specialized in Complex trait genetics (Oct 2006-June 2008), UPMC-Paris 6 University, Paris, France
- **BSc in Life Science**, specialized in biology and biochemistry, (Oct 2003-June 2006)

### Research Experience

---

*Sainte Justine Hospital Research Center, Montréal University, Canada*  
*INSERM U787, UPMC-Paris 6 University, France*

*Doctoral Research*  
*Oct 2008-Dec*  
*2012*

**Project:** Identification and characterization of molecular defect underlying congenital myopathies

#### **Responsibilities: Autonomous management of scientific projects**

- Designed and performed research experiments using next generation sequencing analysis
- Analyzed genomic linkage data
- Performed histology, biochemistry and confocal microscopy analysis on human and mouse samples
- Supervised the research projects of graduate students, and trained/supervised technicians
- Participated in grant writing and manuscript submissions
- Participated in series of patient information/sample collections

#### **Methodologies acquired:**

- Genetics: Microsatellites markers and SNPs based linkage studies analysis (*Affymetrix 250K, Illumina Human Panel V*) using statistical softwares (*MERLIN, PLINK, R...*). Whole exome sequencing (*SureSelect, Agilent 50Mb kit, SOLiD4*) data analysis using NextGene software. Nucleotidic sequences analysis using open-web sources: UCSC, Ensembl, dbSNP, Polyphen, ESEfinder...

- Molecular biology: qPCR, DNA and RNA extraction from human muscle biopsies, siRNA transfection
- Cellular biology: Cell culture of human fibroblast and human and mouse muscle cells
- Histology: Biopsies cryosections, *in situ* immunofluorescence
- Microscopy: Confocal analysis, Image analysis using Photoshop, quantification using ImageJ

**Funding obtained:**

- French Ministry of Research grant for PhD Student (3 year)
- Frontenac grant from Québec (FQRNT) and France (Consulat in Québec) governments for thesis in cotutelle  
*Institut Curie, UPMC-Paris 6 University* *Predoctoral Research, Jan-June 2007*

**Project:** Identification of an erythroleukemia oncogene's target involved in the erythropoiesis blocking

**Responsibilities:**

- Devised strategy for the FACS sorting of specific population of pre-erythroid cells
- Maintained the erythroid cell culture maintaining for the team
- Performed molecular biology and biochemistry experiments

**Methodologies acquired:**

- Molecular biology (PCR, RT-PCR, RNA extraction and purification)
- Cellular biology (FACS sorting, Immunofluorescence analysis, cell culture)
- Biochemistry (Western-Blot)

**Awards and Fellowships**

---

- **Best oral presentation** Elsevier award in the World Muscle Society congress (Perth, Australia, 2012)
- **Young researcher fellowship** from the World Muscle Society (2012)
- French genetics society **travel grant** for participation to ASHG meeting in San Francisco (2012)
- **Academic prize of excellence** from the Montréal University post-doctoral faculty (2011)
- Frontenac **Grant for a thesis in cotutelle** between Québec and France (2009-2012)

**Scientific Communications in congress**

---

Short Talk:

- World Muscle Society 2012 annual meeting (Perth, Australia)
- Human Mutation 2011 Symposium (Santorini Greece)

Poster:

- ASHG 62 annual meeting (San Francisco, 2012)
- Inserm IFR14 annual day (Paris, 2011)
- French Society of Myology 6th meeting (Lausanne, 2009)

**Publications**

---

- **Core myopathy, arthrogryposis and ventricular noncompaction: a novel titinopathy**, C.Chauveau *et al.*, *Annals of Neurology* (Under review)
- **TTN : Two hundred variants for one giant**, C.Chauveau, J.Rowell, A.Ferreiro, *Human Mutation* (Review invited, in preparation)
- Two additional first author manuscript in preparation

