

Étude de la fonction des protéines MBNL au cours du développement à l'aide de cellules souches humaines induites à la pluripotence

*Study of the function of MBNL proteins during
development using human induced pluripotent stem cells*

Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 569

Innovation Thérapeutique : du fondamental à l'appliqué (ITFA)

Spécialité de doctorat : Physiologie / Physiopathologie

Unité de recherche : Inserm, I-Stem, 91100 Corbeil-Essonnes, France.

Référent : Université d'Évry Val d'Essonne

Thèse présentée et soutenue à Corbeil-Essonnes le 24 Mars 2021, par

Antoine MÉRIEN

Composition du Jury

Laurent SCHAEFFER

Professeur des universités, Université Lyon 1 – UMR 5310/U1217

Président

Dider AUBOEUF

Directeur de recherche, ENS Lyon – UMR 5239/U1210

Rapporteur

Marc BITOUN

Directeur de recherche, Sorbonne Université – UMR-S 974

Rapporteur & Examineur

Mario AMENDOLA

Chargé de recherche, Université Évry – UMR-S 951

Examineur

Geneviève GOURDON

Directrice de recherche, Sorbonne Université – UMR-S 974

Examinatrice

Cécile MARTINAT

Directrice de recherche, Université Évry – UMR-S 861

Directrice de thèse

Titre : Étude de la fonction des protéines MBNL au cours du développement à l'aide de cellules souches humaines induites à la pluripotence

Mots clés : Cellules souches pluripotentes humaines, CRISPR/Cas9, Dystrophie myotonique de type 1

Résumé : L'épissage alternatif est apparu comme un mécanisme fondamental non seulement pour la diversification des isoformes des protéines mais aussi pour la régulation spatio-temporelle du développement. Par conséquent, une meilleure compréhension de la manière dont ce mécanisme est régulé permettrait non seulement d'élucider les principes biologiques fondamentaux, mais aussi de déchiffrer les mécanismes pathologiques impliqués dans les maladies dans lesquelles ces processus moléculaires sont dérégulés. Dans le cadre de cette thèse, nous avons utilisé des cellules souches pluripotentes humaines pour déchiffrer, au cours de la myogenèse humaine, le rôle des protéines MBNL, une famille de régulateurs d'épissage spécifiques aux tissus dont la perte de fonction est associée à la dystrophie myotonique de type 1 (DM1), une maladie neuromusculaire héréditaire. Grâce à la technologie CRISPR/Cas9, nous avons généré des cellules souches pluripotentes d'origine humaine (hiPSC) déplétées en protéines MBNL et évalué les conséquences moléculaires et fonctionnelles de cette perte sur la génération de cellules musculaires squelettiques. Nos résultats ont indiqué que les protéines MBNL sont spécifiquement requises pour la maturation myogénique tardive mais pas pour l'engagement myogénique précoce. Par une analyse transcriptomique, nous avons pu mettre en évidence les voies moléculaires régulées par ces protéines durant la myogenèse, ainsi que les phénomènes de compensation entre les paralogues MBNL. Cette étude nous a également permis d'identifier un nouveau défaut d'épissage alternatif dans la DM1, régulé par les protéines MBNL, et qui aboutit à des anomalies structurales du compartiment post-synaptique musculaire. Ensemble, nos résultats révèlent à la fois la temporalité de l'action des protéines MBNL dans la myogenèse humaine et permettent également d'identifier de nouvelles voies moléculaires régulées par ces protéines et pouvant être impliquées dans le développement de la DM1. A plus long terme, les outils développés dans cette étude devraient également faciliter l'identification de nouvelles stratégies thérapeutiques capables de faire face à la perte de fonction de ces protéines.

Title : Study of the function of MBNL proteins during development using human induced pluripotent stem cells

Keywords : Human pluripotent stem cells, CRISPR/Cas9, Myotonic dystrophy type 1

Abstract : Alternative splicing has emerged as a fundamental mechanism not only for the diversification of protein isoforms but also for the spatiotemporal control of development. Therefore, a better understanding of how this mechanism is regulated has the potential not only to elucidate fundamental biological principles, but also to decipher pathological mechanisms involved in diseases where normal splicing networks are mis-regulated. As part of this thesis, we took advantage of human pluripotent stem cells to decipher during human myogenesis the role of MBNL proteins, a family of tissue-specific splicing regulators whose loss of function is associated with Myotonic Dystrophy type 1 (DM1), an inherited neuromuscular disease. Thanks to the CRISPR/Cas9 technology, we generated human-induced pluripotent stem cells (hiPSCs) depleted in MBNL proteins and evaluated the molecular and functional consequences of this loss on the generation of skeletal muscle cells. Our results indicated that MBNL proteins are specifically required for the late myogenic maturation but not for early myogenic commitment. By a transcriptomic analysis, we were able to highlight the molecular pathways regulated by these proteins during myogenesis, as well as the compensatory effects between MBNL paralogs. This study also allowed us to identify a new alternative splicing defect in DM1, regulated by MBNL proteins, which leads to structural abnormalities of the muscular post-synaptic compartment. Together, our results reveal the temporal requirement of MBNL proteins in human myogenesis and allow the identification of new molecular pathways regulated by these proteins that could be involved in the development of DM1. In the longer term, the tools developed in this study should also facilitate the identification of new therapeutic strategies capable to cope with the loss of function of these proteins.

Remerciements

Je tiens en premier lieu à remercier les membres du jury qui ont accepté d'évaluer les travaux réalisés au cours de ce doctorat : les Dr. Didier Auboeuf et Marc Bitoun pour leur contribution en tant que rapporteurs ainsi que les Dr. Geneviève Gourdon, Mario Amendola et le Pr. Laurent Schaeffer qui se sont joints à eux en tant qu'examineurs. J'adresse également un merci tout particulier à Geneviève et Didier pour avoir suivi les avancées de ce projet ainsi que pour leurs conseils.

Je remercie également l'école doctorale "Innovation Thérapeutique du fondamental à l'appliqué" qui m'a permis de bénéficier d'une allocation doctorale sans laquelle je n'aurais pas pu réaliser l'ensemble de mes travaux de thèse.

Je tiens aussi à remercier Marc Peschanski et Raymond Zakhia pour avoir accepté de m'accueillir au sein du laboratoire I-Stem, me permettant d'une part d'avoir réalisé cette thèse dans les meilleures conditions et d'autre part d'avoir développé une expertise dans le domaine des cellules souches.

Je remercie Cécile Martinat pour m'avoir tout d'abord accueilli lors de mon stage de Master 2 et pour m'avoir donné par la suite l'opportunité de poursuivre en thèse au sein de son équipe. Merci de m'avoir permis de travailler sur des projets passionnants ! Merci pour ton soutien dans les moments importants et compliqués ; de la préparation du concours de l'école doctorale à celles des PhD meeting, sans oublier les premières expériences fastidieuses d'obtention des clones CRISPR ! Merci également pour ta confiance et pour l'autonomie que tu m'as laissée tout au long de cette thèse. Enfin, merci pour ton aide dans cette dernière ligne droite pour la correction de ce manuscrit dans un contexte compliqué en raison de la crise sanitaire actuelle.

Je remercie toutes les personnes qui ont pu contribuer de près ou de loin aux avancées de ce projet de recherche. Merci à Nicolas Charlet-Berguerand, Denis Furling, Stéphane Vassilopoulos et Jeanne Lainé pour leur disponibilité et leurs précieux conseils et un grand merci à Gilles Moulay pour le temps consacré et le travail qu'il a fourni dans ce projet. Merci également à Jean-Paul Concordet pour sa bienveillance ainsi que pour la production des protéines Cas9. Je remercie aussi le Pr. Claire Legay pour son regard extérieur et ses remarques toujours pertinentes.

Je remercie sincèrement toutes les personnes d'I-Stem qui ont joué un rôle important dans l'avancée de cette thèse. Un très grand merci à ma binôme de thèse Julie Tahraoui pour toutes ces heures passées devant ces microscopes qui me détestent tant ! Merci pour ton aide précieuse et ton investissement qui m'ont permis de mener à bien ce projet. Merci à Michel Cailleret pour m'avoir tout appris sur les systèmes CRISPR/Cas9 ! Merci pour ton

implication dans ce projet depuis les premiers jours et pour tes précieux conseils. Je n'oublierai jamais ce petit mouvement de "teaspoon" à l'origine de tous ces clones CRISPR et sans lequel nous n'aurions jamais pu battre le record du nombre de codons STOP générés ! Je remercie également Margot Jarrige, Hélène Polvèche et Alexandre Carteron pour tout le temps qu'ils ont consacré aux nombreuses analyses génomiques et transcriptomiques. Merci à Jean-Baptiste Dupont et Christian Pinset pour leur expertise sur les cellules musculaires ainsi que pour leur regard critique et leurs remarques pertinentes. Merci également à Céline Leteur et Jérôme Polentes pour leur patience et leur aide technique au niveau des différentes analyses de microscopie. Je remercie Léa Lesueur pour m'avoir si bien accueilli lors de mon arrivée à I-Stem et pour m'avoir formé à de nombreuses techniques. Merci également pour ta précise aide concernant la cytométrie en flux ! Je remercie aussi Karine Giraud-Triboult et Lina El-Kassar pour leur disponibilité et toutes les analyses caryotypiques des lignées CRISPR obtenues et le génotypage des SNP.

Je tiens également à remercier tous les membres de l'équipe "Maladies du motoneurone". À Julie Tahraoui, Camille Januel et Juliette Duchesne De Lamotte, mes trois acolytes de thèse, merci pour votre soutien quotidien et votre bonne humeur malgré ces années difficiles que nous avons traversées ensemble. Merci à Sandrine Baghdoyan et Jacqueline Gide pour leur bienveillance, leurs conseils et les nombreuses discussions passionnantes qui ont contribué à la bonne ambiance du secteur 7. Merci à Vincent Mouilleau qui m'aura accompagné pendant ces longs trajets de RER D pas toujours très agréables. Merci aussi à Adeline Beuriot, Florine Roussange et Judith Lorant qui prennent la relève de l'équipe et qui participent également à la bonne ambiance générale du laboratoire.

Un merci tout particulier à Julie qui, au-delà de m'avoir formé à la culture cellulaire au cours de mon stage de Master 2 ainsi qu'à de nombreuses analyses de microscopie, m'a surtout supporté pendant 4 ans ! À ma binôme DM1, merci pour ta rigueur, pour la qualité de ton travail et pour ta passion pour la recherche. J'ai énormément appris à tes côtés, autant du point de vue des connaissances scientifiques qu'au niveau des valeurs humaines. Je ne pouvais sincèrement pas mieux tomber. Je te remercie en revanche un peu moins pour tes choix musicaux en secteur 7, qui m'auront toutefois permis de réviser les grands classiques des années 80, sans oublier la seule et l'unique Céline Dion ! Merci pour toutes ces discussions et réflexions autour de nos résultats respectifs qui nous ont permis de résoudre un tas d'interrogations. Je suis très fier de la complicité que nous avons pu développer et sans toi cette thèse aurait eu une tout autre saveur. On pourra dire qu'on se sera bien démené ces dernières années !

Je n'aurai jamais assez de place pour citer tout le monde mais je tiens à remercier tous les membres d'I-Stem que j'ai côtoyé de plus ou moins près. Merci pour les discussions de couloir qui ont pu contribuer à résoudre certains problèmes ou qui m'ont permis au contraire de me changer les idées. Merci à Michel Cailleret pour ton soutien ainsi que pour tous les

bons moments passés en secteur 7 ou en BM à refaire le monde. Merci à Pascal Fragner pour ton humour et ta bonne humeur à toute épreuve ! Merci de m'avoir aidé à savoir prendre du recul. Merci à Emmanuelle Massourides, Virginie Mournetas, Jean-Baptiste Dupont et Christian Pinset pour les discussions passionnantes autant en sciences qu'en gastronomie ! J'ai adoré partager ces moments avec vous. À Christian, merci de votre soutien autant au niveau professionnel que personnel. À Jean-Baptiste, je te souhaite le meilleur pour la nouvelle aventure qui t'attend à Nantes. Merci à Walter Habeler pour ta bienveillance et ton soutien qui m'ont permis d'avoir confiance en moi dans les moments difficiles. Merci à Lina pour ta patience et pour m'avoir accompagné à l'azote un nombre incalculable de fois. Merci à Simona Gribaudo, Karim Ben M'Barek et Margot Jarrige pour vos remarques et vos conseils. J'ai énormément appris à vos côtés.

Je tiens également à remercier tous mes amis qui ont toujours répondu présents au cours de ces dernières années. Merci à Flo, Enzo, Igor, Florent, Marie, Thomas, Adrien, Cédric et Toto pour m'avoir soutenu depuis si longtemps. Merci pour tous les bons moments passés sur la place Georges Moustaki (ou ailleurs) ! Il est évident que sans vous, je n'en serais pas là aujourd'hui. Je mesure la chance que j'ai et j'espère vous avoir à mes côtés encore très longtemps. Merci aussi à Clément, Florent, Luc et Hugo de la Team Biotech qui m'ont énormément encouragé et qui ont toujours été là pour mon remonter le moral. J'ai également beaucoup de chance de vous avoir dans ma vie. On finira bien par créer une start-up ensemble ! Je remercie également toute l'équipe de rugby de l'Amaury Club de Paris et anciens du Stade Paris 6, et j'espère tous vous revoir bientôt sur le pré.

Pour conclure, j'adresse mes remerciements à toute ma famille : à mes grands-parents et à mon arrière-grand-mère, à mes tantes et oncles, à ma cousine Alison, à Fred et au petit Benjamin récemment arrivé dans la famille, je sais que je n'ai pas pu me rendre très disponible ces trois dernières années mais j'espère qu'on pourra rapidement rattraper tout ce temps perdu. Étant le seul scientifique de la famille, suivre et comprendre mes études n'a pas dû être une mince affaire mais je vous rassure, c'est bientôt fini ! Merci à ma grand-mère Mireille et mon grand-père Bernard pour leur confiance, leur tendresse et pour toutes les belles valeurs qu'ils m'ont transmises. Enfin, je remercie énormément mes parents pour m'avoir toujours soutenu et pour avoir toujours eu confiance en moi. Je ne pourrai jamais assez vous remercier pour tout ce que vous avez fait (et que vous faites encore). Sans vous, rien de tout cela n'aurait été possible.

Je dédie ce manuscrit à mon arrière-grand-père Louis ainsi qu'à mon oncle Pascal parti beaucoup trop tôt. Même si vous n'êtes plus là pour le voir, j'espère vous rendre fiers.

* * *

« Parfois il faut savoir saisir les opportunités et ne pas les laisser filer. »

Une inconnue au Jardin du Luxembourg.

* * *

Table des matières

Table des figures	I
Table des tableaux	III
Abréviations	IV
Avant-propos	1
Introduction	3
1. La Dystrophie Myotonique de type 1	4
1.1 Épidémiologie et génétique	6
1.2 Symptômes et formes associées	8
1.2.1 Forme tardive	9
1.2.2 Forme adulte	9
1.2.3 Forme juvénile/infantile	10
1.2.4 Forme congénitale	10
1.3 Mécanismes moléculaires des défauts associés à la DM1	11
1.3.1 Instabilité somatique de la mutation	11
1.3.2 Le gène DMPK et son environnement	13
1.3.3 Gain de fonction toxique à l'ARNm DMPK	17
1.3.3.1 <i>Foci</i> intranucléaires	17
1.3.3.2 Altération des protéines de liaison à l'ARN	19
1.3.3.2.1 Les protéines CELF	20
1.3.3.2.2 Les protéines MBNL	23
1.3.3.3 DM1 : une splicéopathie	24
1.3.3.4 Autres mécanismes pathologiques	28
1.4 Modèles d'étude de la DM1	30
1.5 Stratégies thérapeutiques	34
1.5.1 Cibler l'ADN DMPK	36
1.5.2 Cibler l'ARNm DMPK	37
1.5.3 Cibler les protéines de liaison à l'ARN	39
1.5.4 Cibler les défauts d'épissage	40
2. La famille des protéines MBNL	41
2.1 Les gènes MBNL	41
2.1.1 Organisation génomique	41
2.1.2 Profils d'expression	43

2.2 Le rôle des protéines MBNL dans le développement	45
2.2.1 L'épissage alternatif	45
2.2.1.1 Mécanismes général de l'épissage alternatif	46
2.2.1.2 L'épissage alternatif au cours du développement	48
2.2.2 Les protéines MBNL comme facteurs d'épissage	49
2.2.2.1 Mécanisme d'action des protéines MBNL dans l'épissage	49
2.2.2.2 Rôle développemental des protéines MBNL	51
2.2.2.3 Différentes activités d'épissage selon les isoformes MBNL	53
2.2.3 Autres fonctions des protéines MBNL	55
2.3 Implication des protéines MBNL dans la DM1	57
2.3.1 Séquestration et rôle dans la dynamique de formation des <i>foci</i>	57
2.3.2 Les modèles d'étude des protéines MBNL	59
2.3.2.1 Génération des lignées simple KO et spécialisation fonctionnelle	60
2.3.2.2 Déplétions combinées des protéines MBNL et compensation ..	63
2.3.3 Les défauts d'épissage associés aux protéines MBNL	65
2.3.4 Autres défauts attribués à la séquestration des protéines MBNL	68
3. Les cellules souches pluripotentes humaines	71
3.1 Obtention des cellules souches pluripotentes humaines	71
3.1.1 Les cellules souches embryonnaires humaines	71
3.1.2 Les cellules souches induites à la pluripotence	75
3.2 Caractéristiques des CSP humaines	78
3.2.1 Maintien de la pluripotence in vitro	78
3.2.2 Caractéristiques moléculaires des CSP humaines et contrôle qualité	80
3.2.3 Différenciation ciblée des CSP humaines en cellules musculaires squelettiques	82
3.3 Les CSP humaines et la modélisation pathologique	85
3.3.1 Contribution des CSP humaines dans la DM1	85
3.4 Cellules souches pluripotentes et CRISPR/Cas9	90
3.4.1 CRISPR/Cas9 : origines et généralités	90
3.4.2 CRISPR/Cas9 et édition du génome	93
3.4.2.1 Mécanisme d'action	93
3.4.2.2 Deux mécanismes de réparation de l'ADN double brin	95
3.4.2.3 Limitations	96
3.4.2.4 Avancées et perspectives des systèmes CRISPR/Cas9	97
3.4.3 CRISPR/Cas9 combiné aux iPSC comme outil thérapeutique pour les maladies génétiques	100
3.4.3.1 Thérapie génique et thérapie cellulaire	102
3.4.3.2 Modélisation pathologique	103

Objectifs et démarche expérimentale	106
Travaux de recherche	109
Article 1 : La déplétion des protéines MBNL dans les cellules souches pluripotentes révèle une transition d'épissage anormale de SORBS1 affectant le développement neuromusculaire dans la DM1	110
Article 2 : Altération de la connectivité développementale dépendante des protéines MBNL au niveau de la communication neuromusculaire dans la DM1	168
Discussion générale	216
A. Stratégie de modélisation pathologique pour mieux comprendre les désordres musculaires associés à la DM1	217
B. Rôle des protéines MBNL dans la myogenèse humaine et contribution dans la physiopathologie musculaire de la DM1	220
C. Mise en évidence du défaut d'épissage de l'exon 25 de SORBS1 nécessaire à la formation des clusters des AChR	223
D. Étude du rôle des protéines MBNL dans un autre type cellulaire : les motoneurones	230
Conclusions et perspectives	232
Références bibliographiques	235

Table des figures

Figure 1 : Différentes formes de la DM1 en fonction du nombre de répétitions CTG et de l'âge d'apparition de la maladie et phénomène d'anticipation	7
Figure 2 : Symptômes cliniques associés à la DM1	8
Figure 3 : Hypothèses du rôle physiopathologique de DMPK dans la DM1	14
Figure 4 : Hypothèses mécanistiques pour la pathogénèse moléculaire de la DM1	15
Figure 5 : Formation et dynamique des <i>foci</i> intranucléaires	18
Figure 6 : Altération des protéines de liaison à l'ARN par les transcrits DMPK mutants ..	20
Figure 7 : Antagonisme entre les protéines MBNL et les protéines CELF et conséquences pour l'épissage du canal chlore musculaire CLCN1 dans la DM1	24
Figure 8 : Importance de DMD et BIN1 dans la structure et l'homéostasie du muscle	27
Figure 9 : Récapitulatif des mécanismes pathologiques impliqués dans la DM1	29
Figure 10 : Exemples d'atteintes observées dans les modèles murins de la DM1	34
Figure 11 : Stratégies thérapeutiques pour la DM1	35
Figure 12 : Essais cliniques en cours de développement pour la DM1	39
Figure 13 : Organisation des exons codants des gènes <i>MBNL</i>	42
Figure 14 : Profils d'expression des gènes <i>MBNL</i> dans différents tissus humains	43
Figure 15 : Profils d'épissage des paralogues <i>MBNL</i> au cours du développement	44
Figure 16 : Mécanisme général de l'épissage alternatif et assemblage du spliceosome ..	47
Figure 17 : Liaison des protéines MBNL à l'ARN et épissage alternatif	50
Figure 18 : MBNL1 et MBNL2 inhibent le maintien de la pluripotence	52
Figure 19 : Mbnl1 est responsable de plusieurs transitions d'épissage post-natales musculaires	53
Figure 20 : Boucle de rétrocontrôle de l'épissage de l'exon 5 de MBNL1 et MBNL2	54
Figure 21 : Autres fonctions exercées par les protéines MBNL	55
Figure 22 : Fonctions cytoplasmiques des protéines MBNL et antagonisme avec les protéines CELF	56
Figure 23 : Interactions entre les protéines MBNL et les répétitions CUG toxiques	58
Figure 24 : Contribution tissu-spécifique des protéines Mbnl1 et Mbnl2	62
Figure 25 : Compensation fonctionnelle entre Mbnl1 et Mbnl2 et modélisation des formes sévères de la DM1	64
Figure 26 : Exacerbation des défauts d'épissage chez les souris Mbnl1,2 DKO	67
Figure 27 : L'embryogénèse, l'origine des cellules hES	72
Figure 28 : Représentation schématique de l'obtention d'hiPSC	75
Figure 29 : Revisite du concept de destinée cellulaire et notion de trans-différenciation	76
Figure 30 : Méthodes de reprogrammation des cellules somatiques pour la génération d'hiPSC	77
Figure 31 : Voies de signalisation impliquées dans la pluripotence et la différenciation des CSP	79
Figure 32 : Modèle des réseaux de signalisation régulant la balance pluripotence - différenciation des CSP humaines	80
Figure 33 : Différenciation ciblée des CSP humaines	81

Figure 34 : Obtention de cellules musculaires squelettiques à partir des CSP	82
Figure 35 : Reproduction de phénotypes cellulaires et moléculaires associés à la DM1 à l'aide des CSP humaines et leurs dérivés	86
Figure 36 : Contribution des CSP humaines DM1 dans la recherche de composés potentiellement thérapeutiques	88
Figure 37 : Étapes de la réponse immunitaire adaptative médiée par les systèmes CRISPR/Cas chez les procaryotes	91
Figure 38 : Classification des systèmes CRISPR/Cas et répartition au sein des procaryotes	92
Figure 39 : Représentation schématique du système CRISPR/Cas9 de type II-A	94
Figure 40 : Les mécanismes de réparation de l'ADN double brin	95
Figure 41 : Vue d'ensemble des avancées technologiques du système CRISPR/Cas9	99
Figure 42 : Publications relatives à l'utilisation de CRISPR/Cas9 et des iPSC humaines ..	101
Figure 43 : Représentation du déroulement de la recherche impliquant les iPSC humaines et CRISPR/Cas9 pour de nouvelles alternatives thérapeutiques	101
Figure 44 : Stratégies CRISPR/Cas9 ayant été utilisées dans des modèles DM1.....	104
Figure 45 : Utilisation des hiPSC iCas9 et de leurs dérivés	219
Figure 46 : Stratégie de knock-in pour corriger le défaut d'épissage de l'exon 25 de SORBS1 dans des hiPSC DM1	224
Figure 47 : Potentiels effets de l'exclusion de l'exon 25 de Sorbs1 au niveau des jonctions neuromusculaires chez la souris	226
Figure 48 : Représentation schématique de SORBS1 et prédictions tridimensionnelles .	228
Figure 49 : L'exclusion de l'exon 25 de SORBS1 ne semble pas affecter l'organisation du réseau d'actine	230
Figure 50 : L'exon 25 de SORBS1 est nécessaire pour la formation des clusters des AChR	233

Table des tableaux

Tableau 1 : Exemples de défauts d'épissage observés dans les muscles DM1	26
Tableau 2 : Vue d'ensemble des différents modèles murins pour l'étude de la DM1	33
Tableau 3 : Les modèles murins pour l'étude de la contribution des protéines MBNL dans la DM1	60
Tableau 4 : Principaux épissages alternatifs dérégulés dans les modèles murins knock-out Mbnl	66
Tableau 5 : Les lignées hESC porteuses de mutations génétiques dans le monde	73
Tableau 6 : Exemple de protocoles d'induction myogénique par différenciation dirigée	84
Tableau 7 : Contribution des CSP humaines DM1 à l'identification de nouveaux mécanismes moléculaires	87

Abréviations

AA: Acide Ascorbique	GFP: Green Fluorescent Protein
AAV: Virus Adéno-Associé	GMP: Good Manufacturing Practices
ActD: Actinomycine D	GPCR: G protein-coupled receptor
ADN: Acide DésoxyriboNucléique	gRNA: guide RNA
ARN: Acide RiboNucléique	GSK3-β: Glycogen synthase kinase 3β
ARNm: Acide ribonucléique messenger	HAC: chromosome artificiel humain
ASO: Oligonucléotide Anti-Sens	HDAC(i): Histone deacetylase (inhibitors)
BIN1: Bridging Integrator 1 (Amphiphysine 2)	HDR: Homology Directed Repair
BMP: Bone Morphogenetic Protein	HEK: Human embryonic kidney
CACNA1S: Calcium voltage-gated channel α1S	hESC: human embryonic stem cells
CDK4: Cyclin dependent kinase 4	HGF: Hepatocyte Growth Factor
CDM: Congenital myotonic dystrophy	hiPSC: human induced-pluripotent stem cells
CELF: CUG-BP and ETR-3-like factors	HIS: Homotypic interaction sequence
CLCN1: Chloride voltage-gated channel 1	hnRNP: heterogeneous nuclear ribonucleoprotein
CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats	HSA: Human Skeletal Actin
crRNA: CRISPR RNA	HSA^{LR}: HSA Long Repeats
CSP: Cellules souches pluripotentes	hTERT: human telomerase reverse transcriptase
CUG-BP: CUG-Binding Protein	IDMC: International Myotonic Dystrophy Consortium
dCas9: dead CRISPR associated protein 9	IGF: Insulin-like Growth Factor
DDX: DEAD-box	INSR (IR): Insulin Receptor
DHEA: Déhydroandrosténédione	iPSC: induced-pluripotent stem cells
DHPR: Récepteur à la dihydropyridine	ISE: Intronic Splicing Enhancer
DKO: Double Knock-Out	ISS: Intronic Splicing Silencer
DM: Dystrophie Myotonique	JNM: Jonction neuromusculaire
DMD: Duchenne Muscular Dystrophy	K⁺: (ions) Potassium
DMPK: Myotonic Dystrophy Protein Kinase	KO: Knock-Out
DMWD: DM1 Locus, WD Repeat Containing	KSR: Knockout Serum Replacement
DPI: Diagnostic préimplantatoire	LDB3 (ZASP): LIM domain binding 3
DsRed: Discosoma sp. Red fluorescent protein	LDN: Low-dose Naltrexone
EB: Embryoid body	MBNL: Muscleblind-like
EGF: Epidermal Growth Factor	MCI: Masse cellulaire interne
ERK: Extracellular Receptor Kinase	MEF: Mouse Embryonic Fibroblast
ESC: Cellules souches embryonnaires	miRNA: micro-RNA
ESE: Exonic Splicing Enhancer	MMR: Mismatch Repair
ESS: Exonic Splicing Silencer	MOE: 2'O-methoxyethyl
ETR-3: Elav-type ribonucleoprotein 3	MOR (REM): Mouvement oculaire rapide
FDA: Food and Drug Administration	MSC: Mesenchymal Stem Cell
FGF: Fibroblast Growth Factor	MSH: MutS Homolog
FISH: Fluorescence in situ hybridization	
FIV: Fécondation in vitro	
FOXP1: Forkhead box P1	

mTOR: mammalian Target of Rapamycin
MYHC: Myosin heavy chain
MYOD: Myoblast Determination protein 1
MYOG: Myogenin
MYPT1: Myosin Phosphatase target subunit 1
Na+: (ions Sodium)
NHEJ: Non-homologous end-joining
NLS: Nuclear Localization Sequence
NMDAR1: N-methyl-D-aspartate receptor 1
NOVA: Neuro-Oncological Ventral Antigen
NSC: Neural Stem Cells
OCT4: Octamer-binding transcription factor 4
OKSM: Oct4, KLF4, Sox2, c-Myc
PAA: Poly-Adénylation Alternative
PAM: Protospacer Adjacent Motif
PAX: Paired-box
PBZ: Phenylbutazone
PCR: Polymerase Chain Reaction
PDGF: Platelet-derived Growth Factor
PIN: PiT N-terminus
PKC: Protein Kinase C
PLM: Phospholemman
PLN: Phospholamban
PMA: Procéation médicalement assistée
PMO: Phosphorodiamidate Morpholino
Oligonucleotide
PROMM: Proximal Myotonic Myopathy
PTBP: Polypyrimidine tract-binding protein
RAB3A: Ras-related protein 3A
RAN: Repeat-associated non-AUG
RAR γ : Retinoic acid receptor gamma
RBFox: RNA Binding Fox
RER: Réticulum endoplasmique rugueux
RNP: Ribonucléoprotéine
ROCK: Rho-associated coiled-coil containing
protein kinase
RRM: RNA Recognition Motif
RYR1: Ryanodine Receptor 1
SERCA1: Sarco/endoplasmic reticulum calcium
ATPase1
shRNA: short hairpin RNA
siRNA: small interfering RNA
SNC: Système Nerveux Central
snRNA: small nuclear RNA

snRNP: small nuclear ribonucleoprotein
SP1: Specific Protein 1
SSEA: Stage specific embryonic antigen
STAT: Signal transducer and activators of
transcription
TAL: Transcription activator-like
TALEN: TAL effector nuclease
TGF- β : Transforming Growth Factor β
TKO: Triple Knock-Out
TNNT2 (cTNT): (cardiac) Troponin T2
TPM: Transcripts Per Million
tracrRNA: trans-activating CRISPR RNA
UTR: Untranslated region
WT: Wild-Type
ZFN: Zinc Finger Nuclease
ZNF: Zinc Finger

Avant-propos

Les cellules souches pluripotentes humaines, d'origine embryonnaire (hESC) ou induites à la pluripotence (hiPSC), ont ouvert de nouvelles perspectives thérapeutiques majeures, tant par leur propriété d'auto-renouvellement qui leur confère le statut de source de matériel biologique illimitée, que par leur capacité à se différencier dans tous les types cellulaires de l'organisme. Elles ont ainsi révolutionné le monde scientifique, principalement au niveau de la recherche sur la thérapie cellulaire ainsi qu'au niveau de la modélisation pathologique. De plus, la technologie des hiPSC a rendu ces applications beaucoup plus accessibles. En effet, ces cellules obtenues à partir de cellules somatiques adultes et présentant les mêmes propriétés que les hESC, ont particulièrement permis de s'affranchir de la difficulté d'obtention des hESC ainsi que des questions éthiques qu'elles soulèvent. Elles ont notamment fait l'objet du prix Nobel de médecine en 2012.

La modélisation pathologique repose sur l'utilisation de cellules souches pluripotentes porteuses de la mutation causale d'une maladie génétique ayant pour but de reproduire, dans un modèle cellulaire humain et non transformé, les défauts moléculaires caractéristiques de la pathologie donnée. Les conséquences d'une telle mutation peuvent alors être étudiées au cours des processus de différenciation des cellules souches pluripotentes vers le ou les type(s) cellulaire(s) spécifiquement atteint(s) dans la pathologie. Plus récemment, une nouvelle technologie, également récompensée par un prix Nobel en 2020, s'est peu à peu imposée dans tous les laboratoires de recherche, avec un intérêt particulier concernant les maladies génétiques : les systèmes CRISPR/Cas9. Ce nouvel outil d'édition génomique apporte une fois de plus de nouvelles perspectives thérapeutiques autant du point de vue des thérapies géniques et cellulaires qu'au niveau de la modélisation pathologique, facilitant par exemple la génération de nouvelles lignées cellulaires pour l'étude de maladies rares monogéniques. Ensemble, tous ces modèles cellulaires représentent une opportunité inestimable pour l'amélioration de la compréhension des mécanismes moléculaires et cellulaires impliqués dans le développement de maladies monogéniques ainsi que pour la recherche de nouvelles pistes thérapeutiques, en utilisant par exemple ces derniers comme de réels outils pour des criblages de molécules potentiellement thérapeutiques.

C'est dans ce contexte que j'ai effectué ma thèse au sein de l'institut I-STEM (Institut des cellules Souches pour le Traitement et l'Etude des maladies Monogéniques), laboratoire dédié au développement de nouvelles stratégies thérapeutiques pour les maladies rares d'origine génétique en utilisant les cellules souches pluripotentes humaines. Ce centre de recherche est situé sur le campus du Génopole d'Evry et a été fondé en 2005 par le Dr. Marc Peschanski. C'est lors de mon stage de Master 2 en 2017 que j'ai rejoint I-STEM au sein de

l'équipe "Maladies du motoneurone" dirigée par le Dr. Cécile Martinat et qui s'intéresse particulièrement à l'Amyotrophie spinale infantile (SMA) et à la Dystrophie Myotonique de type 1 (DM1). Dans le but d'étudier les mécanismes pathologiques associés à la DM1, j'ai pu évaluer au cours de ce stage la faisabilité de développer un nouveau modèle cellulaire humain de perte de fonction de la protéine MBNL1, impliquée dans la physiopathologie de la DM1, à l'aide des hiPSC et de la technologie CRISPR/Cas9. Cette preuve de concept m'a donné par la suite l'opportunité de poursuivre ces recherches en tant qu'étudiant en thèse au sein de l'équipe.

L'objectif général de ma thèse a été d'étudier la fonction des protéines MBNL au cours du développement, en particulier dans la spécification myogénique, et ce dans le but de mieux comprendre leur implication dans la physiopathologie humaine de la DM1. La première étape fut de générer plusieurs lignées hiPSC différentiellement déplétées en protéines MBNL, en s'appuyant sur les systèmes d'édition génomique CRISPR/Cas9 pour cibler les différents paralogues *MBNL* afin d'obtenir une lignée *MBNL1* KO, une lignée *MBNL2* KO, ainsi que les lignées *MBNL1,2*DKO et TKO pour s'affranchir des phénomènes de compensation entre ces derniers. Dans un deuxième temps, nous nous sommes intéressés aux conséquences moléculaires et cellulaires de la perte de fonction de ces protéines sur la génération de cellules musculaires squelettiques à l'aide d'un protocole de différenciation myogénique dépourvu de transgène, permettant également d'étudier les étapes développementales qui pourraient être primordiales dans la mise en place des mécanismes physiopathologiques de maladies génétiques neuromusculaires telles que la DM1. Enfin, nous avons pu réaliser des analyses transcriptomiques sur ces cellules musculaires squelettiques obtenues à partir des hiPSC DM1 ainsi que des hiPSC *MBNL* KO générées, dans le but d'identifier de nouveaux mécanismes moléculaires. Cette étude nous a ainsi permis d'identifier un nouveau défaut d'épissage alternatif, régulé par les protéines MBNL1 et MBNL2, et qui semble affecter le compartiment post-synaptique.

Cette thèse m'a offert l'opportunité de pouvoir appréhender ces deux technologies de rupture que sont les hiPSC et CRISPR/Cas9 et d'évaluer la pertinence ainsi que les limitations de leur utilisation pour des applications de modélisation pathologique.

INTRODUCTION

1. La Dystrophie Myotonique de type 1

Les maladies liées aux répétitions de triplets (aussi appelées maladies à triplets) sont des maladies génétiques rares causées par des expansions anormales et instables de répétitions de trinuécléotides dans l'ADN. Ces formes particulières de mutation génétiques ont été mises en évidence au début des années 1990 et sont à l'origine de plus d'une quarantaine de maladies génétiques comme la maladie de Huntington, la sclérose latérale amyotrophique, le syndrome tremblement /ataxie associé à l'X fragile (FXTAS) ou encore la dystrophie myotonique (Paulson 2018). Chez un individu sain, ces répétitions de séquences, également appelées régions microsatellites, sont des séquences nucléotidiques polymorphes réparties tout au long du génome humain. Les expansions anormales proviennent toutes à la base de ces répétitions polymorphiques qui préexistent de manière normale chez les individus. Cela signifie que des mutations *de novo* peuvent apparaître dans des familles jusqu'alors non touchées par la maladie. Ces cas dits sporadiques restent toutefois très rares (Paulson 2018). La Dystrophie Myotonique de type 1 (DM1) appartient à cette famille de maladies et est associée à une répétition anormale de triplet CTG dans la région 3' non codante du gène codant la DMPK (Brook et al. 1992). Dans le cadre de la DM1, il a été reporté l'existence de « pré-mutations » portées par des individus asymptomatiques ou très peu affectés et qui favoriseraient l'apparition de la maladie au cours des générations (Kumar et al. 2014). En effet, ces expansions anormales de répétitions sont caractérisées par une grande instabilité qui va la plupart du temps conduire à l'augmentation de leur taille avec les générations suivantes. C'est ce qu'on appelle le phénomène d'anticipation. De manière générale, plus la longueur des répétitions sera importante, plus l'âge d'apparition des symptômes sera précoce et plus ces symptômes seront sévères. Enfin, la dernière caractéristique commune des maladies à triplet consiste en une grande variabilité des phénotypes, y compris au sein d'une même famille, reflétant principalement les différences de taille des répétitions (Paulson 2018). Les patients DM1 présentent ainsi un nombre de répétitions allant de 50 à plus de 2000, comparé à moins de 37 répétitions chez les patients sains. En raison du phénomène d'anticipation, les symptômes de la maladie vont alors s'aggraver de génération en génération et apparaître de plus en plus tôt. Plusieurs formes de la maladie ont ainsi été décrites allant de formes très sévères et précoces, les formes congénitales, à des formes tardives et plus légères qui se manifestent après l'âge adulte (Turner and Hilton-Jones 2010).

Alors que dans la grande majorité des maladies monogéniques, la mutation causale de la pathologie engendre une perte de fonction partielle ou totale de la protéine encodée par le gène muté, les mécanismes physiopathologiques associés à ces maladies à triplet semblent beaucoup plus complexes. A titre d'exemple, la DM1 apparaît comme une maladie atypique. En effet, la DM1 n'est pas une maladie à perte de fonction mais une maladie à gain de fonction toxique, c'est-à-dire que la mutation va engendrer l'acquisition

de produits toxiques pour la cellule. De plus, dans la DM1, ces produits toxiques ont la particularité de ne pas être établis au niveau protéique mais au niveau de l'ARN messager (ARNm) du gène muté, ce qui renforce d'autant plus ce caractère atypique (Gatchel and Zoghbi 2005).

La DM1 a été décrite pour la première fois en 1909 par le docteur Hans Gustav Wilhelm Steinert en Allemagne suite à l'observation de patients atteints d'une myotonie (retard de la décontraction musculaire après stimulation nerveuse) associée à une faiblesse ou atrophie musculaire. La même année, les docteurs Batten et Gibb décrivent des cas similaires en Angleterre. Dans ses travaux s'intitulant *"About the clinical and anatomical appearance of muscular atrophy in patients with myotonia"*, Hans Steinert fut le premier à décrire ce désordre complexe auquel il donna le nom de *"Dystrophien Myotoniker"*, soit dystrophie myotonique (DM) mais également appelée « maladie de Steinert » (Mishra et al. 2018). Quelques années plus tard, une fréquence élevée de cataractes a été observée chez les patients atteints de DM, suggérant que la DM ne serait pas une maladie purement musculaire. En effet, d'autres études ont révélé par la suite un tableau clinique beaucoup plus complexe de la DM avec des symptômes gastro-intestinaux, des altérations du système nerveux central et des déficiences cognitives, des perturbations du rythme cardiaque ou encore des troubles endocriniens (Bird 2020; Schoser and Timchenko 2010; Harper 2001). Ce n'est qu'en 1992, grâce aux avancées dans le domaine de la génétique et du séquençage, que la mutation causale de la DM a été identifiée. Il s'agit d'une expansion anormale de répétitions du triplet de nucléotides CTG dans la partie non traduite 3'UTR du gène codant pour la *Dystrophy Myotonica Protein Kinase* (DMPK), situé sur le chromosome 19 (locus 19q13.3) (Brook et al. 1992; Buxton et al. 1992; Fu et al. 1992; Mahadevan et al. 1992). Cependant, plusieurs patients présentant des symptômes s'appariant à la DM ne portaient pas cette mutation. Ceci a permis la description d'une seconde forme de la maladie, caractérisée également par des expansions anormales mais cette fois-ci de répétitions CCTG, dans l'intron 1 du gène *ZNF9*, au niveau du chromosome 3 (locus 3q21) (Liquori et al. 2001; Ranum et al. 1998; Thornton et al. 1994). Cette nouvelle forme de la maladie présente des symptômes très similaires à ceux de la DM tels que la myotonie associée à une atrophie musculaire, la cataracte, des atteintes cardiaques et cognitives, ou encore une insulino-résistance (Schoser and Timchenko 2010; Ricker et al. 1994). En revanche, il existe à l'inverse des différences importantes entre les deux formes de la pathologie, avec notamment des atteintes moins sévères et l'absence de formes congénitales pour la forme mutée sur le gène *ZNF9*, ou encore au niveau de la nature des muscles atteints. En effet, alors que la première forme de la maladie touche essentiellement les muscles distaux, la deuxième forme de la maladie affecte à l'inverse les muscles proximaux, raison pour laquelle elle a été nommée *"Proximal Myotonic Myopathy"* (PROMM) (Schoser and Timchenko 2010). De plus, contrairement à la première forme, cette PROMM ne semble pas présenter de phénomène

d'anticipation et aucune corrélation entre le nombre de répétitions CCTG et la gravité des symptômes n'a été observée (Wenninger et al. 2018). Par la suite, le Consortium International de la Dystrophie Myotonique (IDMC) a instauré, lors de l'édition 2000, une nouvelle nomenclature pour distinguer ces deux formes. Désormais, la forme classique de la maladie, la maladie de Steinert portant la mutation sur le gène *DMPK* porte le nom de dystrophie myotonique de type 1 (DM1), et la seconde forme de la maladie celui de dystrophie myotonique de type 2 (DM2) (Meola and Cardani 2015). Dans le cadre de cette thèse, nous nous intéresserons uniquement à la DM1.

1.1. Épidémiologie et génétique

Selon le Ministère des Solidarités et de la Santé*, il existe à ce jour plus de 7 000 maladies rares identifiées dans le monde. Une maladie est considérée comme rare si sa prévalence est égale ou inférieure à 0,05%, c'est-à-dire si elle touche 1 personne ou moins sur 2 000. Cela représente plus de 3 millions de personnes en France, ce qui constitue un véritable enjeu de santé publique. La grande majorité (80%) de ces maladies rares sont d'origine génétique et la dystrophie myotonique de type 1 en fait partie. Bien que considérée comme maladie génétique rare, la DM1 correspond toutefois à l'une des formes de dystrophie musculaire la plus fréquente chez l'adulte avec une prévalence estimée à 10 personnes sur 100 000 avec un mode de transmission autosomique dominant touchant autant les filles que les garçons (Wood et al. 2018; Wenninger et al. 2018; Theadom et al. 2014). Il s'agit d'une prévalence moyenne car celle-ci peut varier selon les régions ou pays. En effet, elle est très faible en Afrique, avec moins d'une personne sur 100 000, mais est à l'inverse plus importante en Croatie par exemple avec 18 personnes atteintes en moyenne sur 100 000. De même, cette prévalence peut varier dans certaines régions très précises comme c'est le cas avec certaines régions du Québec (la région de Charlevoix et le Saguenay-Lac-Saint-Jean) où il est estimé que 150 personnes sont atteintes sur 100 000. Ce chiffre très élevé peut s'expliquer entre autres par un phénomène de dérive génétique appelé effet fondateur qui se produit lorsqu'un nombre restreint d'individus colonise un nouveau territoire, diminuant ainsi la variabilité génétique. Ainsi, la combinaison d'une forte natalité, d'un flux migratoire limité et du caractère autosomique dominant de la DM1 a contribué à l'augmentation dramatique du nombre de personnes atteintes par cette pathologie dans ces régions (Pratte et al. 2015; Yotova et al. 2005).

Comme mentionné plus haut, les personnes non affectées par cette maladie présentent entre 5 et 37 répétitions CTG dans la région 3'UTR du gène *DMPK* alors que les patients atteints de la DM1 possèdent un nombre de répétitions compris entre 50 jusqu'à plus de 2 000 pour les cas les plus sévères. Différentes études ont montré une certaine corrélation

* <https://solidarites-sante.gouv.fr/soins-et-maladies/prises-en-charge-specialisees/maladies-rares/article/les-maladies-rares>

entre le nombre de répétitions CTG et l'âge d'apparition des symptômes ainsi que de leur sévérité ce qui a amené à une classification des formes cliniques de la DM1 selon ces différents critères. On distinguera ainsi la forme tardive, la forme adulte, les formes juvéniles et infantiles ainsi que la forme congénitale (**Figure 1-A**) (Gourdon and Meola 2017; Turner and Hilton-Jones 2014). Il existe également une autre catégorie de patients asymptomatiques qui possèdent entre 38 et 49 répétitions CTG, portant ce qu'on appelle la pré-mutation. La probabilité que ces derniers développent certains symptômes de la DM1 au cours de leur vie est très faible. En revanche, ils présentent un risque certain d'avoir par la suite des enfants possédant un plus grand nombre de répétitions CTG en raison du phénomène d'anticipation dû à l'instabilité de celles-ci (**Figure 1-B et C**).

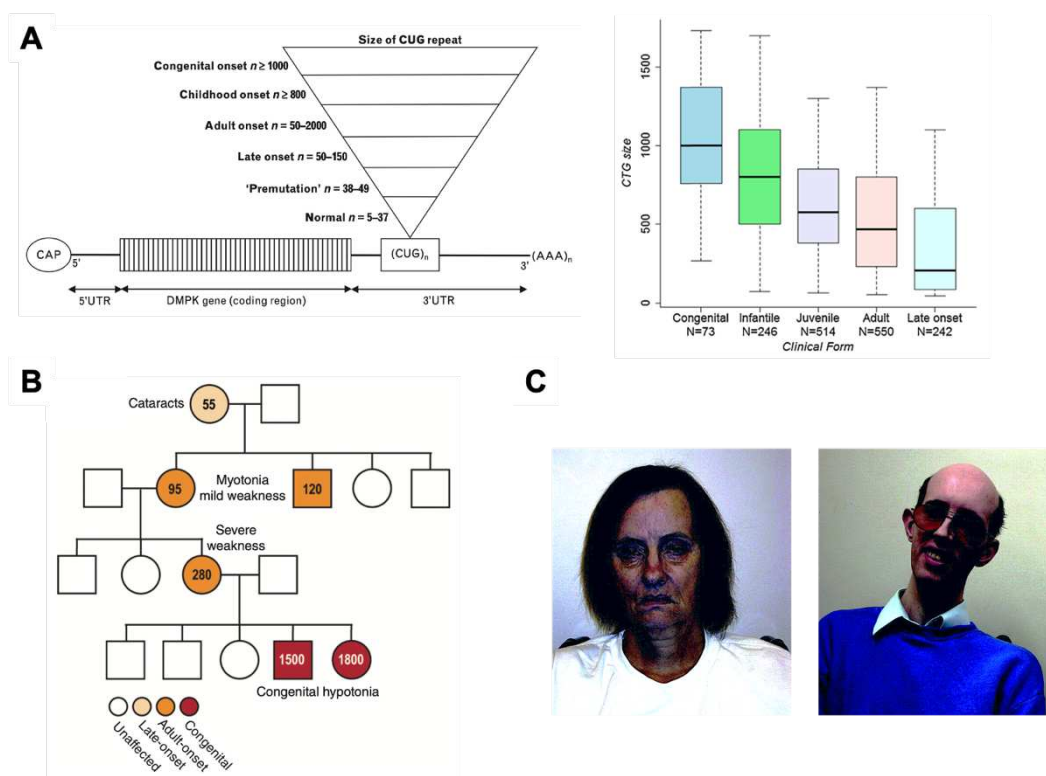


Figure 1 : Différentes formes de la DM1 en fonction du nombre de répétitions CTG et de l'âge d'apparition de la maladie et représentation du phénomène d'anticipation

(A) ARN messager de DMPK et relation entre la taille des répétitions CTG et les différentes formes de la DM1 en fonction de l'âge d'apparition de la maladie. La distribution de la taille des répétitions CTG entre les différentes formes cliniques (à droite) a été effectuée à partir d'une étude sur 2167 patients DM1, via le registre du DM-Scope. (B) Généalogie représentative d'une famille DM1. Les hommes et les femmes sont représentés par des carrés et des cercles respectivement. Le nombre de répétitions CTG est indiqué chez les individus atteints. (C) Exemple d'une mère atteinte d'une forme adulte de la DM1 (environ 530 répétitions CTG) et de son fils atteint d'une forme congénitale (environ 1300 répétitions CTG) illustrant l'augmentation du nombre de répétitions CTG au cours des générations. Adapté de (Thomas et al. 2018; De Antonio et al. 2016; Turner and Hilton-Jones 2010, 2014).

Ce phénomène d'anticipation peut varier selon deux facteurs principaux : le sexe du parent porteur de la mutation et la taille des répétitions CTG portées par celui-ci. Il a d'abord été observé que les formes les plus sévères de la maladie provenaient dans la très grande majorité des cas de mères atteintes. Une étude publiée en 2016 sur plus de 1400 patients DM1 a montré que seulement 9% des cas d'apparition de formes congénitales de la DM1 étaient associées à l'héritage paternel. À l'inverse, un nombre important de pères asymptomatiques ou très légèrement atteints (possédant moins d'une centaine de répétitions CTG) transmettent la maladie sous une forme clinique reconnaissable (Dogan et al. 2016; Savić Pavićević et al. 2013). L'une de ces études a révélé une autre implication du sexe qui n'avait jamais été reconnue auparavant et qui concerne le profil clinique et la gravité de la DM1, avec notamment une morbidité et une mortalité plus élevées chez les hommes (Dogan et al. 2016).

1.2. Symptômes et formes associées

C'est donc sur la base de l'âge d'apparition de la maladie, des symptômes cliniques et de leur degré de sévérité que les quatre grandes formes cliniques de la DM1 ont été définies (**Figure 2**).

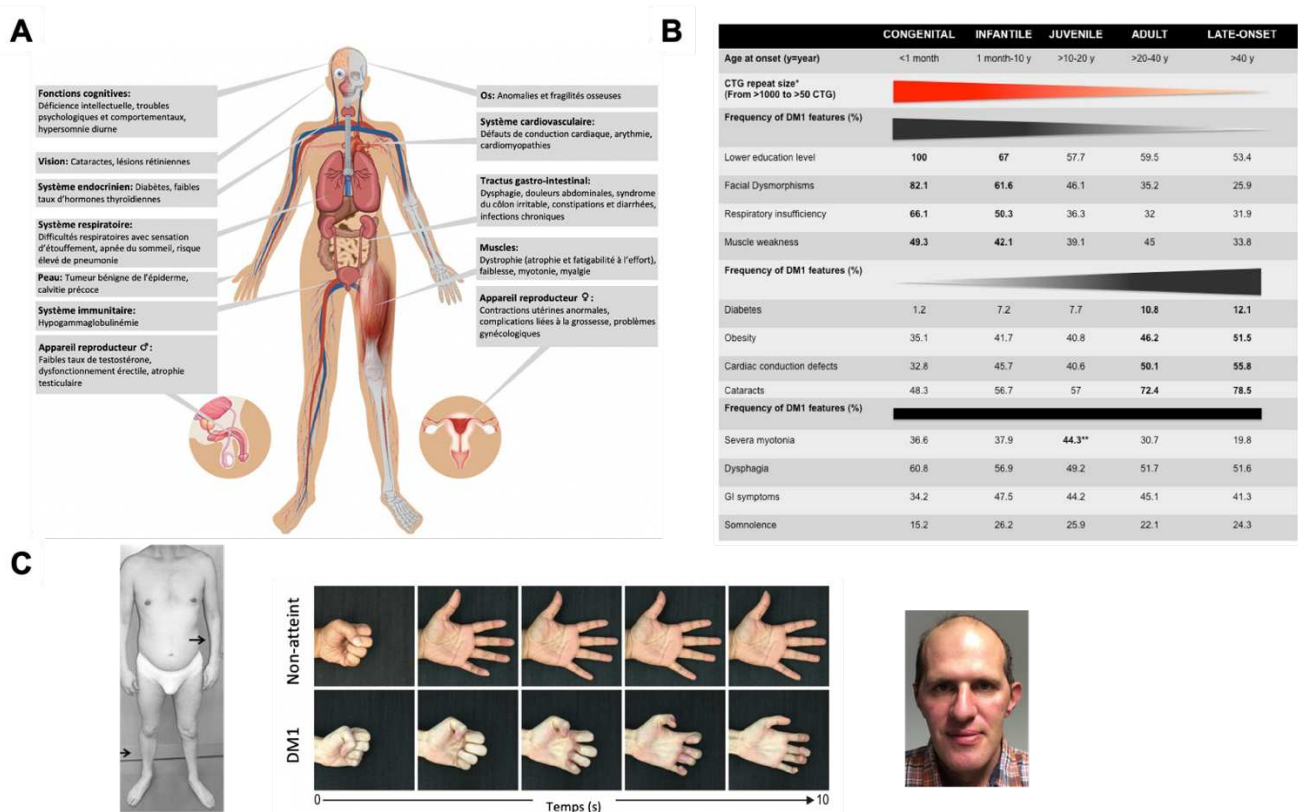


Figure 2 : Symptômes cliniques associés à la DM1

(A) Vue d'ensemble des symptômes associés à la DM1. **(B)** Fréquences des principaux symptômes en fonction des différentes formes de la maladie (à partir de l'étude sur les 2167 patients DM1 de De Antonio et al., 2016). **(C)** Exemple de trois symptômes fréquents chez les patients DM1 : dystrophie des muscles distaux (à gauche), myotonie (au centre) et dysmorphie faciale (à droite). *Adapté de (Tomé and Gourdon 2020; Thomas et al. 2018; Wenninger et al. 2018; Schoser and Timchenko 2010) et du site <http://myotonicdystrophy.com>.*

1.2.1. Forme tardive

Il s'agit de la forme la moins sévère de la maladie. Les patients atteints par cette forme peuvent développer des symptômes « légers » apparaissant au-delà de 40 ans et qui incluent une cataracte ou encore une calvitie frontale précoce. Ils présentent peu d'atteintes musculaires (parfois une myotonie légère) et possèdent entre 50 et 150 répétitions CTG. La forme tardive de la DM1 est souvent compliquée à diagnostiquer car les signes cliniques, s'ils sont présents, peuvent passer inaperçus. Le diagnostic de la cataracte s'avère souvent très utile pour identifier ces patients pouvant transmettre la maladie sous des formes plus sévères en raison de l'instabilité de ces répétitions (le plus souvent via une transmission paternelle) (Meola and Cardani 2015). Enfin, diagnostiquer les patients présentant une forme tardive permet aussi d'identifier d'autres membres de la familles porteurs de la mutation et ainsi leur prise en charge avant l'apparition des symptômes (Turner and Hilton-Jones 2010).

1.2.2. Forme adulte

C'est la forme la plus classique et la plus répandue de la DM1. Celle-ci se déclare entre 20 et 40 ans, la plupart du temps après 30 ans. Les atteintes sont multi-systémiques car elles concernent aussi bien les muscles que les fonctions cognitives, les systèmes cardio-vasculaire, respiratoire, visuel endocrinien, ou encore l'appareil reproducteur et le tractus gastro-intestinal (**Figure 2-A**). Les patients atteints d'une forme adulte présentent une grande variabilité du nombre de répétitions, allant le plus fréquemment de 100 jusqu'à 1000 répétitions. Parmi tous les différents symptômes pouvant se manifester chez ces patients, les principaux sont avant tout musculaires avec par exemple une dystrophie progressive des muscles distaux accompagnée d'une myotonie touchant aussi bien les muscles striés squelettiques que les muscles lisses (**Figure 2-C**). Ils présentent également une faiblesse musculaire au niveau des muscles de la face, du diaphragme, des muscles intercostaux ou encore des muscles oculaires externes. Toutes ces atteintes engendrent alors des difficultés pour parler, se nourrir et respirer. Les troubles cardiaques, avec notamment des anomalies de conduction et des arythmies, contribuent de manière importante à la mortalité de ces patients (Wenninger et al. 2018; Meola and Cardani 2015).

1.2.3. Forme juvénile/infantile

La forme juvénile et la forme infantile sont souvent associées car elles présentent globalement le même tableau clinique. Les symptômes peuvent apparaître entre 1 et 10 ans pour les formes infantiles et entre 10 et 20 ans pour les formes juvéniles. Ces formes sont généralement observées chez les individus portant plus de 800 répétitions CTG. Les premiers symptômes à se développer sont plutôt d'ordre cognitif avec des difficultés d'apprentissage, d'élocution et des déficiences intellectuelles. Les atteintes musculaires apparaissent majoritairement à l'adolescence avec une faiblesse des muscles de la face et des muscles distaux, une myotonie et même dans certains cas des symptômes cardiovasculaires pouvant entraîner des complications importantes (Wenninger et al. 2018). Ces symptômes musculaires apparaissant plus tardivement, le diagnostic des formes infantiles est plus compliqué que celui des formes juvéniles et sera plutôt établi à la suite de la déclaration de la maladie chez l'un des parents (Meola and Cardani 2015).

1.2.4. Forme congénitale

Les formes congénitales de la DM1 (CDM) sont les formes les plus sévères de la pathologie et sont associées aux plus longues tailles de répétitions CTG (plus de 1000). Cependant, la CDM présente un tableau clinique légèrement distinct de celui des autres formes de la maladie et ne doit pas être considérée comme une simple forme adulte de la DM1 en plus précoce et plus sévère. Les formes congénitales sont majoritairement héritées de la mère et se manifestent dès la période intra-utérine avec un excès de liquide amniotique et une réduction des mouvements fœtaux. Après l'accouchement, les principaux symptômes sont une grande faiblesse musculaire généralisée (hypotonie), une dysmorphie faciale et une importante détresse respiratoire (Meola and Cardani 2015). L'insuffisance respiratoire est à l'origine d'une importante mortalité (environ 30 à 40%). La majorité des enfants atteints de CDM sont capables de marcher de manière autonome. La myotonie n'est pas présente dans les premières années mais d'autres symptômes d'ordre cognitifs et psychologiques vont se manifester, souvent de manière plus sévère que pour les formes infantiles, avec un important retard mental et dans certains cas l'apparition de troubles autistiques (Wenninger et al. 2018). Les patients CDM ont également une espérance de vie réduite, qui est d'environ 45 ans en excluant les morts néonatales (Bird 2020).

1.3. Mécanismes moléculaires des défauts associés à la DM1

La base génétique de la DM1 a été révélée en 1992 avec l'identification de la mutation : une expansion anormale et instable des répétitions des triplets nucléotidiques CTG dans la région non codante 3'UTR du gène DMPK, localisée sur le chromosome 19 au niveau du locus 19q13.3 (Harper 2001; Brook et al. 1992). Il a par la suite fallu comprendre comment cette mutation pouvait être responsable de défauts aussi nombreux et complexes, impliquant autant de systèmes différents. L'hypothèse d'un seul et unique mécanisme pathologique fut donc rapidement exclue, d'autant plus que le rôle de la protéine DMPK n'est toujours pas encore bien établi. De nombreuses études ont ainsi pu mettre en évidence des hypothèses mécanistiques autres que la seule haplo-insuffisance de la protéine DMPK, comme par exemple l'implication des gènes adjacents au gène *DMPK* ou encore la dérégulation de protéines de liaison à l'ARN impliquées dans l'épissage alternatif ou la stabilité des ARNm (Sicot et al. 2011).

D'autre part, la grande instabilité des répétitions CTG évoquée précédemment constitue en elle-même un mécanisme jouant un rôle important dans la physiopathologie de la DM1, notamment au niveau de la grande variabilité observée entre les patients DM1 (Tomé and Gourdon 2020).

1.3.1. Instabilité somatique de la mutation

L'expansion anormale et dynamique des répétitions CTG est la base moléculaire du phénomène d'anticipation observé au sein des familles DM1. Quelques études se sont intéressées à la distribution des tailles des répétitions CTG dans les cellules germinales ou encore dans des embryons et ont montré que cette instabilité intergénérationnelle intervenait au cours des premières étapes de la gamétogénèse et des premiers stades post-zygotiques (Tomé and Gourdon 2020; Dean et al. 2006; De Temmerman et al. 2004). Par ailleurs, des études ont également révélé la présence de variations importantes au niveau de la taille de ces répétitions dans les cellules somatiques de patients DM1. Ce phénomène, appelé mosaïcisme, met ainsi en évidence l'instabilité somatique des répétitions CTG au sein d'un même patient atteint de la DM1. Le degré d'instabilité somatique va varier en fonction de trois critères différents : le tissu cellulaire concerné, la taille initiale des répétitions CTG et l'âge du patient. Cette instabilité se met en place dans les tissus entre la 13^{ème} et la 16^{ème} semaine de gestation. De manière générale, les plus grandes répétitions présentent l'hétérogénéité somatique la plus élevée et continuent à s'agrandir avec le temps (Tomé and Gourdon 2020; Martorell et al. 1997). Il existe peu d'études ayant analysé les différences de tailles des répétitions CTG en fonction des tissus, particulièrement en raison des difficultés d'accès aux tissus humains. Cependant, il a été montré que les plus grandes expansions de

répétitions de CTG étaient retrouvées au niveau de tissu musculaire, notamment le cœur et le diaphragme, de la peau et des testicules, comparés à d'autres tissus comme le cervelet, la rate ou encore le thymus (Tomé and Gourdon 2020; Kinoshita et al. 1996; Wohrle et al. 1995).

Différents mécanismes influençant la dynamique de l'instabilité des répétitions CTG ont été mis en évidence comme par exemple la réparation de l'ADN ou encore certaines modifications épigénétiques. En effet, des études ont montré dans des modèles murins de la pathologie, l'implication de la voie de réparation des mésappariements (MMR) avec en particulier l'implication des protéines MSH (MutS homologs), principaux acteurs de ce système de réparation de l'ADN. En croisant un modèle murin de la DM1 portant plus de 300 répétitions CTG avec des souris dans lesquelles le gène *Msh2* a été inactivé, il a été observé que l'absence de *Msh2* avait un effet sur l'instabilité somatique et intergénérationnelle des répétitions, non pas en induisant des expansions mais au contraire des contractions (Savouret et al. 2003). De plus, une récente étude a révélé que l'inactivation du gène *MSH2* dans des cellules souches pluripotentes humaines DM1 avait pour effet de stabiliser les répétitions CTG au cours des passages (Franck et al. 2020). D'autres études ont également montré que *Msh2* devait s'associer avec un partenaire, *Msh3* ou *Msh6*, afin de former un complexe fonctionnel. Les protéines *Msh3* et *Msh6* ont un rôle important en raison de leurs fonctions antagonistes par rapport à la modulation de la taille des répétitions CTG, soit en les allongeant dans le cas d'un complexe *Msh2-Msh3*, soit en les réduisant si *Msh2* interagit au contraire avec *Msh6*. Ainsi, l'inactivation du gène *Msh6* dans un modèle murin de la DM1 entraîne une augmentation significative de la fréquence des expansions des répétitions CTG alors que l'inactivation du gène *Msh3* est suffisante pour diminuer ces expansions (Foiry et al. 2006; van den Broek et al. 2002). Plus récemment, des analyses effectuées sur plus de 400 patients DM1 ont montré que des polymorphismes au sein du gène *MSH3*, ayant un effet sur l'expression et/ou l'activité protéique de *MSH3*, étaient associés au degré d'instabilité somatique des répétitions (Flower et al. 2019).

Enfin, l'effet de modifications épigénétiques sur l'instabilité des répétitions CTG ainsi que sur la variabilité des symptômes chez les patients DM1 a également été étudié. Il a été observé que les niveaux de méthylation de l'ADN en amont des expansions de répétitions CTG du gène *DMPK* étaient corrélés avec la longueur celles-ci. De plus, des analyses par régression linéaire multiple ont montré que les niveaux de méthylation de l'ADN au sein du gène *DMPK* pourraient également contribuer à la variabilité phénotypique des patients, en particulier au niveau musculaire et respiratoire, indépendamment de la longueur des répétitions CTG (Tomé and Gourdon 2020; Légaré et al. 2019).

1.3.2. Le gène DMPK et son environnement

Les premières études mécanistiques se sont concentrées sur les effets de la mutation sur le gène *DMPK* lui-même avec l'hypothèse que la présence de ces répétitions résulte en l'haplo-insuffisance de DMPK. En effet, la transcription du gène *DMPK* muté engendre la production d'un ARNm présentant une structure en épingle à cheveux au niveau de la région 3'UTR causée par les répétitions CUG. Cette structure va être capable d'interagir avec des protéines de liaison à l'ARN formant ainsi des complexes ribonucléoprotéiques qui vont s'agréger sous la forme de *foci* se retrouvant séquestrés à l'intérieur du noyau de la cellule. Il a ainsi été montré que ces *foci* résultaient en une diminution de la quantité des protéines DMPK de moitié dans les cellules musculaires dérivées de patients DM1 (Furling et al. 2001), dans la mesure où, seul l'ARNm provenant de l'allèle sain du gène *DMPK* est exporté vers le cytoplasme de la cellule pour être traduit en protéine.

Le gène *DMPK* code une sérine/thréonine kinase présentant une forte homologie de séquences avec les *myotonic dystrophy kinase-related Cdc42-binding kinases* (MRCK), les *Rho-associated kinases* (ROCK) ainsi qu'avec la *citron kinase* (CRIK). La protéine DMPK est exprimée de manière ubiquitaire, majoritairement dans les muscles squelettiques, le cœur, les muscles lisses et le cerveau (Mulders et al. 2011; Kaliman and Llagostera 2008). D'un point de vue structurel, la protéine DMPK est constituée d'un domaine à activité kinase dans sa partie N-terminale et d'un domaine hélice alpha formant des motifs en super-hélice (coiled-coil) dans sa partie C-terminale. Elle possède également un domaine d'adressage membranaire et des sites d'autophosphorylation (Elkins et al. 2009). Il existe six isoformes majeures de la protéine DMPK qui résultent de l'épissage alternatif. Celles-ci diffèrent principalement au niveau de la partie C-terminale, affectant leur localisation intracellulaire : les quatre isoformes contenant une longue queue C-terminale vont être préférentiellement localisées au niveau des membranes du réticulum endoplasmique ou des mitochondries alors que les deux autres formes n'ayant pas de queue C-terminale seront localisées dans le cytoplasme sous forme soluble. L'autre particularité entre les différentes isoformes est la présence ou non d'un motif VSGGG unique qui a été décrit pour moduler l'activité d'autophosphorylation de la DMPK. L'autophosphorylation de la DMPK va induire un changement conformationnel de la protéine stabilisant son domaine kinase dans une conformation ouverte favorisant ainsi la fixation de substrats (Wansink et al. 2003).

Différentes protéines impliquées dans des mécanismes cellulaires pouvant être corrélés à certains symptômes de la DM1 ont été identifiées comme cible de l'activité kinase de protéine DMPK (**Figure 3**). Parmi celles-ci, la protéine Phospholemman (PLM), qui peut induire des courants chlorure (Cl⁻) via l'activation d'un canal chlore, voit sa fonction modulée par sa phosphorylation par la DMPK (Mounsey et al. 2000). Dans le cadre d'une haplo-insuffisance en protéines DMPK, la dérégulation du PLM pourrait être corrélée avec le

dysfonctionnement des canaux chlorure responsables de la myotonie chez les patients DM1 (Kaliman and Llagostera 2008). D'autres exemples avec la phosphorylation inhibitrice de la *Myosin phosphatase 1* (MYPT1) augmentant la sensibilisation au calcium des muscles lisses, ou encore avec la phosphorylation du Phospholamban (PLN) impliqué dans la régulation du calcium intracellulaire, semblent indiquer un rôle de la protéine DMPK dans l'homéostasie calcique (Kaliman and Llagostera 2008). Cependant la contribution de ces différentes protéines n'a jamais été validée d'un point de vue physiopathologique. De par ses fonctions de régulateur de l'épissage alternatif et de la stabilité de certains ARNm, la protéine CUG-BP est considérée comme l'un des acteurs majeurs de la DM1. La surexpression d'une forme hyperphosphorylée de CUG-BP a été observée dans un modèle de souris dans lequel le gène *Dmpk* a été inactivé ainsi que chez des patients DM1. Cependant, sa potentielle phosphorylation par la protéine DMPK n'a jamais été démontrée (Kaliman and Llagostera 2008). Il a également été suggéré que la protéine DMPK jouerait un rôle dans l'adressage membranaire du récepteur à l'insuline (IR) et du récepteur à l'IGF1 (*Insulin Growth Factor 1*) (Kaliman and Llagostera 2008; Llagostera et al. 2007).

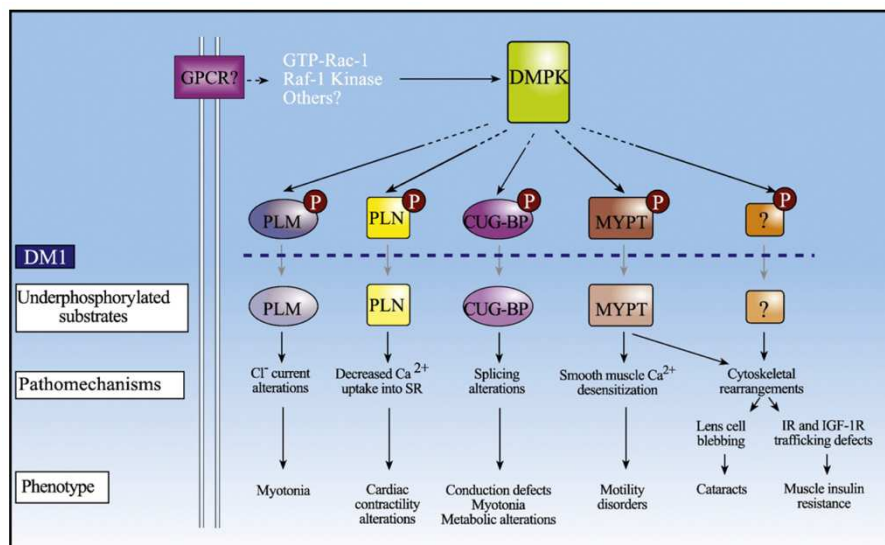


Figure 3 : Hypothèses du rôle physiopathologique de DMPK dans la DM1

Plusieurs cibles phosphorylées par la DMPK sont impliquées dans des fonctions cellulaires corrélées à différentes altérations observées dans la DM1 : Phospholemman (PLM), Phospholamban (PLN), CUG-binding protein (CUG-BP), et Myosin phosphatase 1 (MYPT1). GPCR: Récepteur couplé aux protéines-G ; SR: Réticulum sarcoplasmique ; IR: Récepteur à l'insuline ; IGF-1R: Récepteur de l'IGF1 (Kaliman and Llagostera 2008).

En parallèle, des études ont montré que la protéine DMPK jouait également un rôle au niveau de la différenciation musculaire. En effet, la déplétion de la protéine *Dmpk* dans des cellules musculaires murines entraîne une diminution de l'expression de plusieurs marqueurs de la différenciation musculaire comme la myogénine (MyoG) et MyoD, accompagnée d'un

défaut de fusion des cellules musculaires et donc d'une formation diminuée de myotubes (Harmon et al. 2008). Afin d'avancer dans la compréhension de la fonction de la protéine Dmpk, un modèle de souris a été généré dans lequel le gène *Dmpk* a été inactivé (souris *Dmpk*^{-/-}). Ces souris développent une myopathie légère et progressive qui pourrait être apparentée à une forme tardive de la DM1. Ils ont notamment pu observer chez les souris plus âgées une variation de la taille des fibres musculaires avec une dégénérescence accrue, ainsi qu'une fibrose. Cependant, ce modèle murin ne présente pas les symptômes les plus fréquents de la maladie comme l'atrophie musculaire des muscles distaux, la myotonie ou encore une cataracte (Jansen et al. 1996; Reddy et al. 1996).

Ainsi, il semble que l'haplo-insuffisance de la protéine DMPK puisse être impliquée dans la mise en place de certaines atteintes pathologiques de la DM1. En revanche, la grande majorité de ces études n'ayant pas été plus détaillées par la suite, il en ressort que l'hypothèse de l'haplo-insuffisance de DMPK ne permet pas d'expliquer à elle seule toute l'étendue des atteintes de la maladie ni leur degré de sévérité. Cela suggère donc qu'il existe d'autres mécanismes responsables de la pathogenèse de la DM1 (**Figure 4**). En raison de sa nature instable et de sa taille pouvant être importante, la mutation au sein du gène DMPK altère la structure de la chromatine au niveau du locus. Il a alors été envisagé que l'expansion des répétitions CTG pourrait également impacter l'expression des gènes voisins *DMWD* et *SIX5* (Llamusi and Artero 2008).

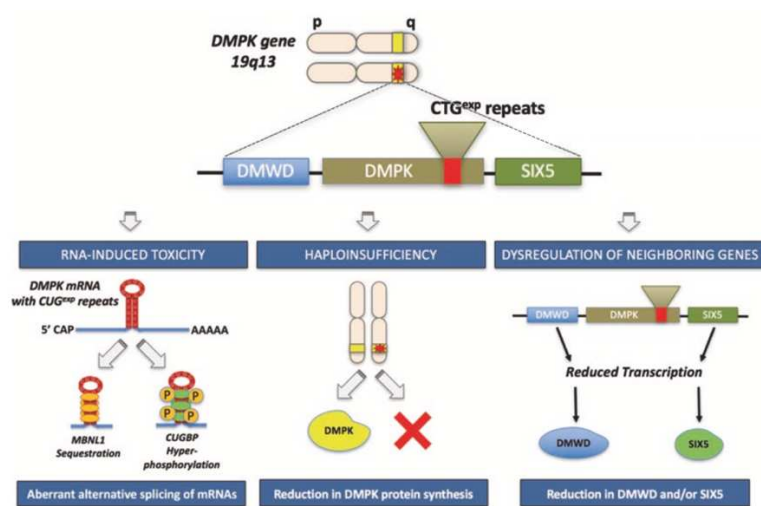


Figure 4 : Hypothèses mécanistiques pour la pathogénèse moléculaire de la DM1

L'expansion des répétitions CTG entraîne un gain de fonction toxique au niveau de l'ARNm *DMPK* qui va déréguler la fonction de protéines de liaison à l'ARN induisant des défauts d'épissage alternatifs (à gauche). La mutation entraîne une haplo-insuffisance responsable de la réduction de la synthèse de la protéine DMPK (au centre) ainsi qu'une dérégulation de l'expression des gènes adjacents (à droite). *Adapté de (Chuah et al. 2019).*

Cette hypothèse a été principalement renforcée par des études observant une diminution du niveau d'expression de DMWD dans des fibroblastes issus de biopsies de patients DM1 (Alwazzan et al. 1999; Gennarelli et al. 1999). Le gène *DMWD* (*dystrophia myotonica-containing WD repeats*) se situe juste en amont du gène *DMPK* et code une protéine possédant quatre motifs WD. D'autres protéines contenant ces motifs WD ont été décrites comme jouant un rôle dans la transduction du signal ou dans des fonctions de régulation cellulaire (Johnson and Jarjour 2006). Ce gène pourrait être impliqué dans certains aspects cliniques de la DM1 car il est préférentiellement exprimé dans des tissus comme le cerveau et les testicules. Ainsi, la diminution de l'expression de DMWD pourrait être impliquée dans la mise en place des atteintes cognitives et de l'appareil reproducteur masculin. Un rôle potentiel de DMWD a même été proposé au niveau de la régulation transcriptionnelle de la croissance et de la maturation du système nerveux (Westerlaken et al. 2003). Néanmoins, aucune étude n'a permis à ce jour de démontrer les véritables fonctions de DMWD ainsi que son implication dans la DM1.

Comme pour le gène *DMWD*, une diminution de l'expression du gène *SIX5* (*sine oculis homeobox homolog 5*) a également été révélée dans différents tissus, particulièrement dans des biopsies cérébrales post-mortem, dans des fibroblastes ainsi que dans des cellules musculaires issues de patients DM1 (Frisch et al. 2001; Alwazzan et al. 1999; Thornton et al. 1997). Cette diminution semblerait associée à l'inhibition de la transcription de l'allèle muté (Alwazzan et al. 1999). Préalablement connu sous le nom de DMAHP pour *dystrophia myotonica adjacent homeobox protein*, le gène *SIX5* est situé directement en aval du gène *DMPK*, partageant ainsi des séquences régulatrices au niveau de son promoteur avec ce dernier. Ainsi, plus la taille de l'expansion des répétitions CTG sera importante, plus ces éléments régulateurs seront éloignés de la partie codante du gène *SIX5* (Johnson and Jarjour 2006). Ce gène code un facteur de transcription à homéodomaine exprimé principalement dans le muscle squelettique, le cœur, le cerveau, les lymphocytes. De par son rôle de facteur de transcription, plusieurs cibles de *SIX5* ont été identifiées, chacune pouvant potentiellement contribuer à différentes atteintes de la DM1. Par exemple, l'IGF1 et l'IGF2 sont deux cibles de *SIX5* dont la dérégulation pourrait être à l'origine de la résistance à l'insuline observée chez les patients DM1. L'expression de la myogénine est également une cible du facteur de transcription *SIX5*, ce qui pourrait expliquer certaines atteintes musculaires (Johnson and Jarjour 2006; Fougousse et al. 2002). Un modèle de drosophile dans lequel le gène homologue à *SIX5* (*D-Six4*) a été inactivé révèle que celui-ci semble nécessaire pour le développement du muscle et des gonades, suggérant ainsi une implication pour *SIX5* au niveau des atrophies musculaires et testiculaires dans la DM1. Un modèle de souris *Six5* (-/-) a d'ailleurs pu mettre en évidence cette atrophie testiculaire accompagnée de défauts de maturation de la spermatogénèse. La présence de cataractes a également été observée dans ce modèle mais celles-ci ne présentent pas les opacités

iridescentes caractéristiques des cataractes des patients DM1. De plus, ce modèle murin n'a montré aucune atteinte des muscles squelettiques (Llamusi and Artero 2008; Sarkar et al. 2004; Wakimoto et al. 2002).

Bien que la majorité des études présentées ici soient de simples corrélations entre les atteintes observées dans la DM1 et les cibles des trois protéines DMPK, DMWD et SIX5, il semble toutefois que ces trois protéines pourraient être impliquées dans la mise en place de certains phénotypes caractéristiques de la DM1, d'autant plus chez les patients atteints de formes congénitales ou infantiles, dont la taille des répétitions CTG est très importante. Il est indéniable en revanche que ces hypothèses mécanistiques n'expliquent pas à elles seules l'ensemble du tableau clinique des patients DM1. Une autre hypothèse mécanistique, sur laquelle nous allons nous attarder à présent, est basée sur le gain de fonction toxique de l'ARNm mutant, causé par l'expansion anormale des répétitions CTG et la formation de *foci* intranucléaires au niveau des transcrits *DMPK* (**Figure 4**).

1.3.3. Gain de fonction toxique de l'ARN messenger DMPK

1.3.3.1. *Foci* intranucléaires

L'une des caractéristiques fondamentales de la physiopathologie de la DM1 est l'accumulation des ARNm *DMPK* mutés contenant les expansions de répétitions CUG, sous la forme d'agrégats nucléaires distincts également appelés *foci* intranucléaires (Pettersson et al. 2015). La preuve de la formation de ces *foci* a été démontrée la première fois en 1995 dans des fibroblastes et des biopsies musculaires de patients DM1 à l'aide de la technique d'hybridation *in situ* (RNA-FISH) combinée à une sonde fluorescente de 10 répétitions CAG complémentaires aux répétitions CUG (**Figure 5-A**) (Taneja et al. 1995). Par la suite, de nombreuses études ont également détecté la présence nucléaire de ces *foci* dans différents tissus de patients DM1 tels que les muscles squelettiques ainsi que dans des neurones issus de biopsies de cerveau post-mortem (Jiang et al. 2004; Mankodi et al. 2003; Davis et al. 1997). Une autre étude a révélé que les *foci* se mettaient en place très tôt au cours du développement en les détectant à la fois dans des tissus fœtaux de patients DM1 ainsi que dans des tissus embryonnaires et néonataux dans un modèle de souris DM1 (Michel et al. 2015).

Concernant la structure de ces *foci* intranucléaires, les répétitions CUG forment une structure secondaire en forme d'épingle à cheveux (aussi appelée tige-boucle) qui peut être plus ou moins longue en fonction du nombre de répétitions CUG. Il a été montré que onze triplets CUG étaient suffisants pour former une telle structure (Napierała and Krzyzosiak 1997). Ces structures en épingle à cheveux sont stabilisées par l'appariement des bases

nucléotidiques C et G. Cependant, ces liaisons C-G et G-C sont séparées par des mésappariements U-U qui contribuent à un équilibre entre stabilisation et déstabilisation, raison pour laquelle ces structures sont souvent appelées épingles à cheveux "glissantes" (**Figure 5-B**). Ces mésappariements U-U forment ainsi des renflements au sein de la structure en épingle à cheveux lui donnant cet aspect si particulier. Ce sont donc en réalité des structures dynamiques au sein desquelles les triplets CUG se déplacent et se réalignent continuellement (Hale et al. 2019). Une étude a également observé des changements dynamiques au cours du cycle cellulaire dans des cellules neuronales dérivées de cellules souches pluripotentes de patients DM1, avec des cycles de dissociation / reformation des *foci* intranucléaires (Xia and Ashizawa 2015).

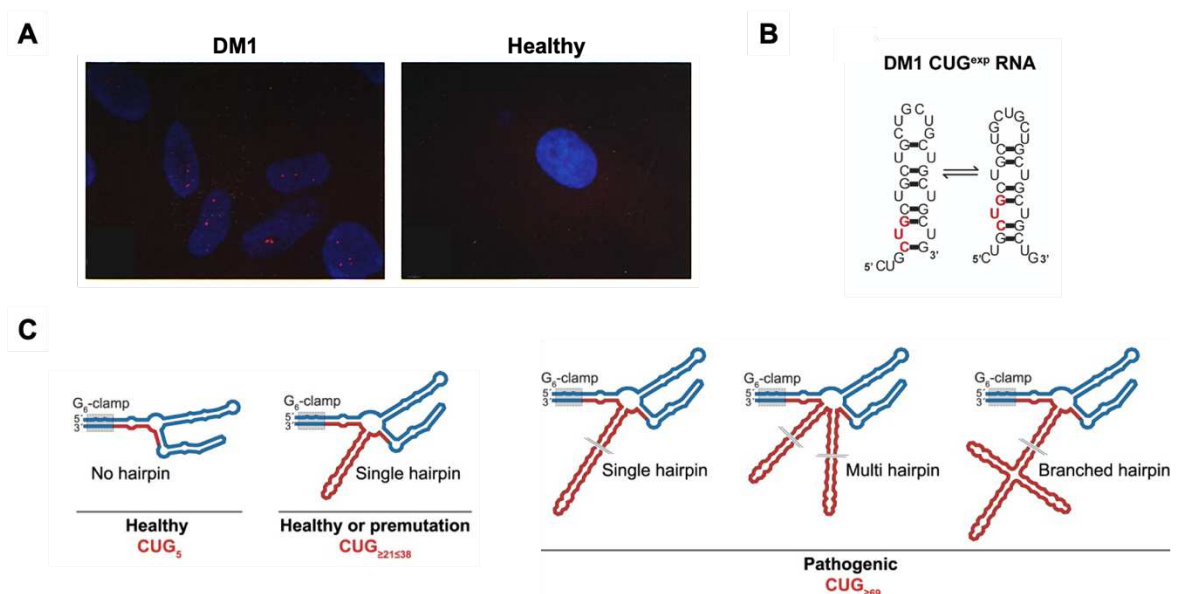


Figure 5 : Formation et dynamique des *foci* intranucléaires

(A) Première observation de *foci* intranucléaires par RNA-FISH dans les noyaux de fibroblastes issus de patients DM1. **(B)** La structure secondaire des répétitions CUG en épingle à cheveux est une structure dynamique. Cette dynamique est représentée ici en suivant un seul triplet CUG (en rouge). **(C)** Modèle proposant différentes structures secondaires des répétitions CUG (en rouge) en fonction de leur taille. Ces structures deviennent pathogènes à partir de 69 répétitions CUG et peuvent adopter différentes conformations : simples, multiples ou ramifiées. L'expansion des répétitions n'entraîne cependant pas de modification au niveau des séquences ARN adjacentes (en bleu). *Adapté de (Hale et al. 2019; van Cruchten et al. 2019; Taneja et al. 1995).*

Récemment, un modèle a été proposé concernant la formation de ces structures en épingle à cheveux (van Cruchten et al. 2019). En comparant différentes tailles de répétitions CUG (5, 21, 38, 69, 147 et 197 répétitions), cette étude démontre que la taille des répétitions module la structure de l'épingle à cheveux, lui conférant ainsi une plasticité conformationnelle (**Figure 5-C**). Les ARNm composés de seulement 5 répétitions ne forment pas de structure en épingle à cheveux. Ces structures sont observées avec 21 ou 38

répétitions CUG (correspondant à des patients sains ou asymptomatiques) mais sont très courtes et uniquement sous la forme d'une simple tige. Au-delà de 69 répétitions, les structures deviennent pathogènes et adoptent différentes conformations longues et glissantes, sous la forme soit d'une longue tige simple, d'épingles à cheveux multiples ou encore de tiges ramifiées entre elles. Cette étude a également révélé que l'expansion anormale des répétitions n'entraînait pas de changement de structure dans les séquences ARN adjacentes. En effet, les épingles à cheveux sont isolées structurellement du reste de l'ARNm DMPK par deux hélices stables riches en GC et par une "pince" formée par six appariements G-C consécutifs (G_6 -clamp) (van Cruchten et al. 2019). Les structures plus complexes observées dans le cadre d'expansions importantes des répétitions CUG sont encore plus propices à la séquestration de protéines de liaison à l'ARN et pourraient ainsi expliquer le degré de sévérité entre les différentes formes de la DM1.

1.3.3.2. Altération des protéines de liaison à l'ARN

De par leur structure spécifique, les ARNm mutants induisent un gain de fonction toxique dans la cellule en modulant l'expression ou la disponibilité de différents acteurs moléculaires. Plusieurs protéines de liaison à l'ARN ont ainsi été décrites pour interagir avec les répétitions CUG (**Figure 6**). Par exemple, des facteurs de transcription tels que SP1 (Specificity factor 1), STAT1, STAT3 et RAR γ (Récepteur gamma à l'Acide Rétinoïque) ont été retrouvés co-localisés avec les répétitions CUG par immunofluorescence dans des cellules musculaires de patients. Leur séquestration pourrait ainsi réduire leur disponibilité au niveau de la chromatine active entraînant une perturbation des profils d'expression de leurs gènes cibles (Magana and Cisneros 2012). L'expression d'un autre facteur de transcription, NKX2-5, a été retrouvée au contraire augmentée dans des cellules musculaires squelettiques et cardiaques de patients DM1, induisant des défauts d'expression de gènes. Cependant sa potentielle interaction avec les répétitions CUG n'a jamais été démontrée (Yadava et al. 2008). En effet, plusieurs facteurs n'étant pas directement séquestrés par les *foci* peuvent tout de même être dérégulés dans la DM1. C'est par exemple le cas du facteur de transcription SPEN/SHARP dont la fonction est altérée dans la DM1 en raison d'un défaut de localisation de la protéine. Une étude a révélé une corrélation entre l'expression des répétitions CTG et la mauvaise localisation cytoplasmique de ce facteur de transcription (Dansithong et al. 2011). La dérégulation de ces facteurs de transcription pourrait ainsi être à l'origine de nombreux défauts d'expression génique chez les patients DM1, et ces derniers ne sont pas les seuls facteurs dérégulés. Cependant cette hypothèse n'a jamais été validée de façon formelle dans le cadre de la DM1.

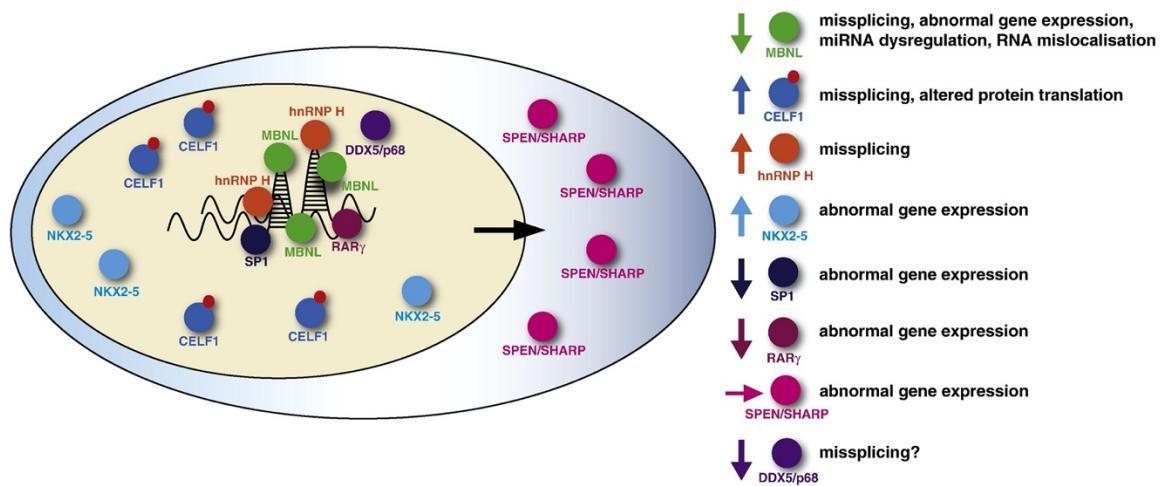


Figure 6 : Altération des protéines de liaison à l'ARN par les transcrits *DMPK* mutants

En s'accumulant dans le noyau de la cellule sous la forme de *foci*, les ARNm *DMPK* mutants vont affecter l'expression, la distribution ou encore la localisation de multiples protéines de liaison à l'ARN, perturbant ainsi leur fonction (Sicot and Gomes-Pereira 2013).

En effet, des facteurs de l'épissage alternatif comme Staufen1 sont aussi retrouvés co-localisés avec les *foci* (Ravel-Chapuis et al. 2012). Plusieurs ribonucléoprotéines hnRNP, notamment l'hétérodimère hnRNP-H/hnRNP-F ainsi que hnRNP-A1, ont également été retrouvées co-localisées au niveau des *foci*. De plus, ces facteurs d'épissage sont retrouvés surexprimés dans les cellules de patients DM1, ce qui pourrait ainsi contribuer à des défauts d'épissage d'ARNm cibles. Cependant les mécanismes par lesquels leur séquestration par les *foci* aboutirait à leur surexpression n'ont jamais été démontrés (Jiang et al. 2004; Sicot and Gomes-Pereira 2013; Talukdar et al. 2011).

Parmi toutes les protéines de liaison à l'ARN dérégulées dans la DM1, deux familles de protéines ont fait plus particulièrement l'objet de nombreuses études : la famille des protéines CELF et la famille des protéines MBNL.

1.3.3.2.1. Les protéines CELF

Dans le but de mieux comprendre les mécanismes moléculaires associés à la DM1, des expériences ont été réalisées visant à identifier les protéines capables de se lier spécifiquement aux motifs CUG *in vitro*. Ces expériences de précipitation (pull-down) utilisaient une séquence de 8 répétitions CUG dans des cellules HeLa. C'est ainsi qu'a été initialement découverte la protéine CUG-BP1 (pour *CUG-binding protein 1*) (Timchenko et al. 1996). CUGBP1 appartient à la famille des protéines CELF (*CUG-BP- and ETR-3-like factors*) qui est constituée de six gènes paralogues codant les protéines CUG-BP1 (ou CELF1), ETR-3 (*ELAV-type RNA binding protein 3*) également appelée CUG-BP2 ou CELF2, ainsi que les

protéines CELF3 à CELF6. Cette famille de protéines très conservées au cours de l'évolution est impliquée dans différents processus de la biologie des ARN comme par exemple l'épissage alternatif de certains pré-ARNm ainsi que la traduction ou encore la dégradation des ARNm (mRNA decay) (Dasgupta and Ladd 2012). Chacune des protéines CELF contient trois domaines RRM (*RNA recognition motif*) de liaison à l'ARN, deux se situant à l'extrémité N-terminale et un autre à l'extrémité C-terminale. Ce dernier, le domaine RRM3, est également nécessaire à la fonction d'épissage des protéines CELF (Singh et al. 2004). La région qui sépare le 2^{ème} domaine RRM du 3^{ème} est appelée "domaine divergent" car les similitudes de séquences au niveau de cette région divisent les protéines CELF en deux sous-groupes distincts : CUG-BP1 et ETR-3 d'un côté et les CELF3 à 6 de l'autre (Ladd et al. 2001). L'expression des protéines CELF est ubiquitaire mais varie tout de même en fonction des différents paralogues. En effet, de manière générale, la première sous-famille (CELF1-2) est plus largement exprimée que la sous-famille CELF3-6. CUG-BP1 et ETR-3 sont majoritairement exprimés dans les muscles squelettiques, le cœur et le système nerveux central (Dasgupta and Ladd 2012). Il a été montré que leur profil d'expression variait également au cours du développement humain, plus particulièrement lors de la transition du stade fœtal au stade adulte. En effet, l'expression de CUG-BP1 diminue significativement au cours du développement des muscles squelettiques pour atteindre un faible niveau d'expression dans le muscle adulte. Au niveau du système nerveux central, en particulier cérébral, son expression reste constante lors des premiers stades développementaux mais chute drastiquement jusqu'à être à peine détectable dans le cerveau adulte (Ladd et al. 2001). Concernant ETR-3, son niveau d'expression augmente dans le muscle et le cerveau mais une transition a été observée avec un changement d'isoforme de la protéine entre les stades embryonnaires et adultes (Ladd et al. 2001). Cette étude a également révélé que cette diminution de l'expression de CUG-BP1 et ce changement d'isoforme d'ETR-3 étaient corrélés avec des modifications développementales au niveau de l'épissage alternatif de certains ANRm, permettant de passer de l'expression d'isoformes fœtales à celles d'adulte. C'est notamment le cas de la Troponine T cardiaque (cTNT) par exemple, qui subit un changement au cours du développement en passant d'une forme fœtale contenant l'exon 5 dans le muscle embryonnaire à une forme sans l'exon 5 dans le muscle adulte. Ainsi, de par leur rôle de facteur d'épissage et en raison des changements de leur expression au cours du développement, CUG-BP1 et ETR-3 seraient fortement impliqués dans ce processus d'épissage alternatif pour la transition fœtale-adulte de nombreux ARNm (Chau and Kalsotra 2015; Ladd et al. 2001).

La protéine CUG-BP1 existe sous une forme nucléaire ainsi que sous une forme cytoplasmique, lui permettant d'intervenir dans différents processus de la biogénèse des ARNm. La forme cytoplasmique est impliquée au niveau de l'initiation de la traduction des ARNm, de leur stabilité ainsi que dans le processus de dégradation des ARNm, alors que la

forme nucléaire joue le rôle de facteur d'épissage. Afin d'effectuer cette fonction de facteur d'épissage, ces protéines interagissent avec des régions introniques riches en nucléotides U et G, et adjacentes aux exons sujets à l'épissage alternatif. CUG-BP1 interagit préférentiellement avec des motifs UGU alors que ETR-3 a une plus grande affinité pour les motifs UGUU (Teplova et al. 2010). Bien que décrite comme se liant aux motifs CUG (Timchenko et al. 1996), il a été découvert que la protéine CUG-BP1 préférerait se lier à la base de la structure en épingle à cheveux et que sa fixation n'était pas proportionnelle à la taille des répétitions CUG. En effet, la protéine CELF ne se lie pas à ces structures d'épingle à cheveux sous forme double brin, mais uniquement aux répétitions sous forme simple brin. Ainsi, dans un contexte DM1, les protéines CUG-BP1 ne sont pas séquestrées par les *foci* intranucléaires (Chau and Kalsotra 2015; Michalowski et al. 1999). En revanche, plusieurs études ont démontré que le niveau d'expression de la protéine CUG-BP1 était anormalement élevé, jusqu'à deux à quatre fois plus dans des tissus de patients DM1 par rapport à des tissus sains. Ces études ont montré que cette surexpression serait la conséquence de l'activation anormale de la protéine kinase C (PKC) qui va alors agir de manière post-traductionnelle et hyper-phosphoryler la protéine CUG-BP1, entraînant ainsi une augmentation de sa stabilité dans la DM1 (Kuyumcu-Martinez et al. 2007; Savkur et al. 2001). Cependant la relation entre l'expansion des répétitions CUG et la voie de signalisation à l'origine de l'hyper-phosphorylation de CUG-BP1 par la PKC reste encore méconnue. Des études ont toutefois observé cette surexpression de la forme phosphorylée de CUG-BP1 dans différents modèles murins d'étude de la DM1 tels que des souris dans lesquelles le gène *DMPK* a été inactivé (*Dmpk*^{-/-}) ou encore dans des muscles de souris exprimant 960 répétitions CTG de manière tissu-spécifique (EpA960). Ces résultats suggèrent que la PKC pourrait être une cible indirecte de la protéine DMPK et que l'haplo-insuffisance de celle-ci dans la DM1 entraînerait une hyper-activation de la PKC, aboutissant à une hyper-phosphorylation de CUG-BP1 (Orengo et al. 2008; Wang et al. 2007; Roberts et al. 1997). Par ailleurs, il a aussi été montré dans le tissu cardiaque de patients DM1 que des micro-ARN (miR-23a/b), régulant normalement dans un contexte non-pathologique la traduction de CUG-BP1, étaient sous-exprimés, ce qui pourrait aussi contribuer à sa surexpression (Kalsotra et al. 2014).

Enfin, d'autres souris transgéniques ont été générées afin de confirmer l'implication de CUG-BP1 dans la pathogenèse de la DM1, en récapitulant le gain de fonction de la protéine pour tenter de modéliser la maladie. Ainsi, la surexpression de CUG-BP1 dans les muscles squelettiques et dans le cœur de souris reproduit de nombreux phénotypes associés à la DM1 comme par exemple des défauts de conduction cardiaque ou encore une atrophie des fibres musculaires corrélée à une faiblesse musculaire. Une forte mortalité post-natale a également été observée chez ces modèles murins de surexpression de CUG-BP1, majoritairement liée à des faiblesses respiratoires. D'autre part, ces modèles ont également

permis l'observation de nombreux défauts d'ordre moléculaire comme des défauts d'épissage alternatif ou des défauts au niveau de la stabilité d'ARNm. En revanche, aucun signe de myotonie n'a été observé dans ces différents modèles (Chau and Kalsotra 2015; Koshelev et al. 2010; Ward et al. 2010; Ho et al. 2005). Récemment, une étude a montré que la surexpression de la forme nucléaire seule de CUG-BP1 suffisait pour induire une atrophie musculaire chez les souris (Cox et al. 2020).

1.3.3.2.2. Les protéines MBNL

L'autre famille de protéines majoritairement impliquée dans la physiopathologie de la DM1 est la famille des protéines Muscleblind, ou MBNL, qui est composée de trois paralogues : MBNL1, MBNL2 et MBNL3. Contrairement aux protéines CELF, les protéines MBNL ont été retrouvées directement séquestrées par les *foci* intranucléaires dans des tissus de patients DM1 (Miller et al. 2000). Ces protéines ont été identifiées pour leur capacité à se lier spécifiquement aux structures en épingle à cheveux sous forme double brin de l'ARN mutant DMPK, et ce proportionnellement à la taille des répétitions CUG (Miller et al. 2000). D'un point de vue fonctionnel, les protéines MBNL interviennent dans différents processus de la biogénèse des ARNm tels que leur transport et leur localisation ou encore leur traduction. Cependant, leur fonction principale reste la régulation de l'épissage alternatif de certains ARN pré-messagers. De façon similaire aux protéines CELF, les protéines MBNL sont également essentielles à la transition foetale-adulte du profil d'épissage des ARNm au cours du développement. Mais contrairement aux protéines CELF, les niveaux d'expression des protéines MBNL vont augmenter au cours du développement, permettant ainsi la transition des épissages d'un grand nombre d'ARNm du stade foetal vers l'acquisition du stade adulte. On parle alors de fonctions d'épissage antagonistes entre les protéines MBNL et les protéines CELF (Ho et al. 2004). Ainsi, dans un contexte non pathologique, les protéines MBNL semblent faiblement exprimées pendant les stades embryonnaires contrairement aux protéines CELF, régulant l'épissage alternatif de nombreux ARNm cibles en faveur de profils d'épissage foetaux. Au cours du développement, il y a donc une transition entre l'expression des protéines CELF qui diminue au cours du développement et celle des protéines MBNL qui à l'inverse augmente, permettant alors l'acquisition de profils d'épissage adultes. Dans le cadre de la DM1, la balance entre l'expression des protéines MBNL et des protéines CELF est totalement déséquilibrée, induisant de nombreux défauts d'épissage entraînant le maintien de l'expression d'isoformes foetales de protéines dans des tissus adultes. C'est notamment le cas de la protéine CLCN1, le canal chlore musculaire (**Figure 7**). Au stade adulte, le profil d'épissage de l'ARNm codant le CLCN1 est dépourvu de l'exon 7a et conduit à la production d'une protéine fonctionnelle. Chez les patients DM1, on observe un défaut d'exclusion de l'exon 7a au niveau de l'ARNm de CLCN1, maintenant ainsi l'expression de l'isoforme foetale. En effet, l'inclusion de l'exon 7a entraîne l'apparition d'un codon STOP prématuré

aboutissant à la production d'une protéine tronquée non fonctionnelle. Ce défaut d'épissage de CLCN1 conduisant à la perte de fonction de la protéine serait tout au moins en partie, responsable de la myotonie observée chez les patients DM1 (Thomas et al. 2018; Mankodi et al. 2002).

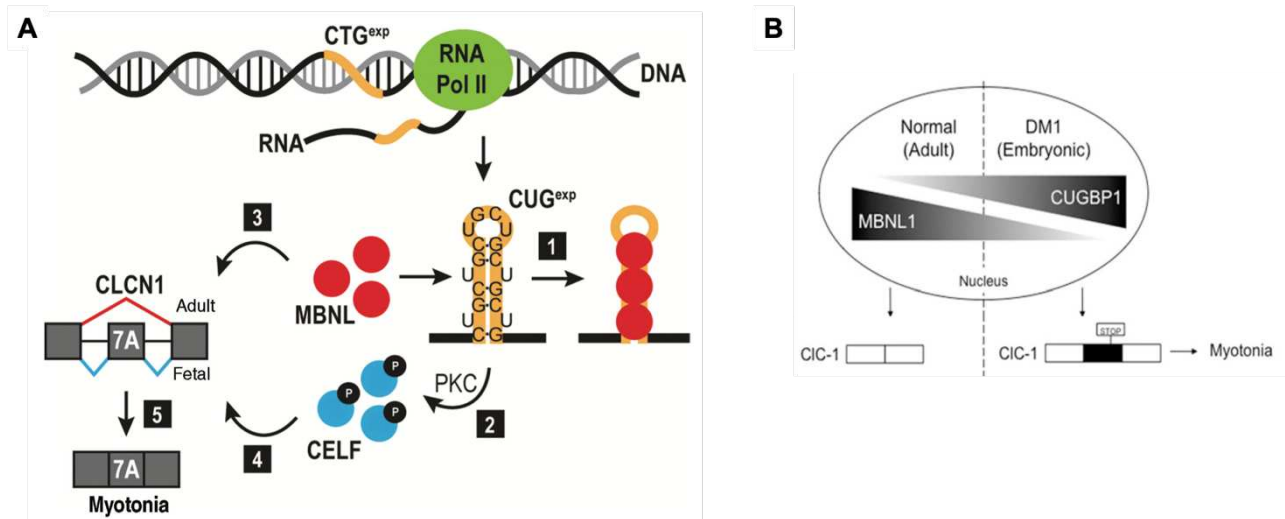


Figure 7 : Antagonisme entre les protéines MBNL et les protéines CELF et conséquences pour l'épissage du canal chlore musculaire CLCN1 dans la DM1

(A) L'expansion anormale des répétitions CUG (orange) forment des *foci* intranucléaires qui vont séquestrer les protéines MBNL (rouge) et induire l'hyper-activation de CELF1. Ces deux altérations vont être responsables de nombreux défauts d'épissage, entraînant l'émergence d'isoformes fœtales dans les tissus adultes. (B) Cette cascade conduit notamment à l'inclusion de l'exon 7a dans l'ARNm de CLCN1, à l'origine de la myotonie observée chez les patients DM1. *Adapté de (Thomas et al. 2018; Lee and Cooper 2009).*

En raison de l'importance de la contribution des protéines MBNL dans la mise en place des défauts moléculaires et cellulaires de la DM1, cette famille de protéines fera l'objet d'un chapitre entier dans la deuxième partie de cette introduction.

De nombreuses études ont effectivement révélé que la fonction d'épissage liée aux protéines MBNL était pour grande partie responsable des défauts d'épissage caractéristiques de la DM1, qui lui vaut ainsi le nom de splicéopathie.

1.3.3.3. DM1 : une splicéopathie

En raison de la fonction altérée de plusieurs familles de facteurs d'épissage, la DM1 est caractérisée par de nombreux défauts d'épissage. Plus d'une centaine ont été répertoriés dans différents tissus et cellules issus de patients DM1, majoritairement au niveau de tissus musculaires et du système nerveux central (López-Martínez et al. 2020).

Dans le cadre de ce manuscrit, et pour cibler plus spécifiquement les résultats qui y sont présentés, nous nous intéresserons principalement aux défauts d'épissage touchant le muscle squelettique. Dans le but d'expliquer certains phénotypes musculaires de la DM1, de nombreuses études se sont intéressées à la caractérisation des défauts d'épissage alternatif présents dans les muscles squelettiques et cardiaques de patients atteints de la DM1 (**Tableau 1**). En 2013, Masayuki Nakamori et ses collègues ont ainsi analysé les ARNm transcrits à partir de biopsies réalisées dans le muscle tibial antérieur d'une cohorte de 50 patients DM1. Cette étude a révélé en particulier 42 épissages alternatifs altérés spécifiquement chez ces patients DM1, et dont certains avaient déjà été décrits dans des études précédentes. En mesurant la force de flexion dorsale de la cheville chez ces patients, ils ont pu associer une vingtaine de ces défauts d'épissage à la sévérité de la faiblesse musculaire (Nakamori et al. 2013). Cependant, dans cette étude, il n'était question que d'une simple corrélation. Peu des anomalies d'épissage alternatif identifiées à ce jour dans la DM1 ont pu être associées à des phénotypes cellulaires en particulier. Parmi ceux-ci, l'inclusion anormale de l'exon 7a de *CLCN1* décrite ci-dessus fut le premier défaut d'épissage dont l'effet au niveau fonctionnel a été démontré et directement associé à une atteinte caractéristique de la DM1 : la myotonie (Mankodi et al. 2002).

Par la suite, quelques défauts d'épissage ont été étudiés de manière plus approfondie. C'est notamment le cas du gène *DMD*, le plus grand gène humain codant la dystrophine. Cette protéine est un composant majeur du complexe dystrophine-glycoprotéines dont le rôle est de relier les protéines du cytosquelette à la membrane sarcomérique. C'est un facteur clé de la maintenance de la structure des fibres musculaires (**Figure 8-A**). Plusieurs études ont observé un défaut d'inclusion des exons 71 et 78 de *DMD* dans des biopsies musculaires de patients DM1 mais les effets fonctionnels de la perte de ces exons n'avaient pas été décrits (Nakamori et al. 2013; Yamashita et al. 2012). En 2015, l'équipe de Denis Furling a décrit le rôle de l'exon 78 de *DMD* ainsi que les conséquences de son défaut d'épissage. Son équipe a ainsi montré que l'exclusion de l'exon 78 modifiait le cadre de lecture du dernier exon (exon 79) de *DMD*, induisant un codon STOP plus tardif. Cette modification engendre au niveau protéique une queue C-terminale de 31 acides aminés au lieu de 13 lorsque l'exon 78 est présent. Ceci entraîne un changement conformationnel au niveau de l'extrémité C-terminale de la dystrophine, passant d'une conformation en feuillet β à une structure en hélice α . Ce changement de conformation de la dystrophine a pour conséquences un remodelage des fibres musculaires associées à des anomalies structurales caractéristiques des muscles squelettiques de patients DM1, ce qui aboutit, tout au moins en partie, à une certaine faiblesse musculaire (Rau et al. 2015).

Tableau 1 : Exemples de défauts d'épissage observés dans les muscles DM1

Gène	Exon/intron	Altération dans la DM1	Système biologique	Implications dans la DM1	Référence(s)
<i>ALPK3</i>	Exon 2	Inclusion	Biopsie muscle squelettique patient DM1/CDM	Non décrites (n.d.)	Nakamori (2017-2013)
<i>ATP2A1 (SERCA1)</i>	Exon 22	Exclusion	Biopsie et myotubes dérivés de patient DM1/CDM	Altération de l'homéostasie calcique intracellulaire	Nakamori (2017-2013) et Kimura (2005)
<i>BIN1</i>	Exon 11	Exclusion	Biopsie muscle squelettique patient DM1/CDM	Faiblesse musculaire due à l'altération des tubules T	Nakamori (2017-2013), Fugier (2011)
<i>CACNA1S</i>	Exon 29	Exclusion	Biopsie muscle squelettique patient DM1/CDM	Faiblesse musculaire due à l'altération du couplage excitation/contraction	Nakamori (2017-2013) et Tang (2012)
<i>CAPN3</i>	Exon 16	Exclusion	Biopsie muscle squelettique patient DM1	n.d.	Nakamori (2013), Yamashita (2012) et Lin (2006)
<i>CLCN1</i>	Intron 2	Inclusion	Biopsie muscle squelettique patient DM1/CDM	Myotonie	Nakamori (2017-2013), Lueck (2007), Charlet (2002) et Mankodi (2002)
	Exon 7a	Inclusion			
<i>DMD</i>	Exon 71	Exclusion	Biopsie et myotubes dérivés de patient DM1/CDM	Faiblesse musculaire due à l'altération de l'intégrité membranaire	Nakamori (2017-2013), Yamashita (2012) et Rau (2015)
	Exon 78	Exclusion			
<i>DTNA</i>	Exon 11a	Exclusion	Biopsie et myotubes dérivés de patient DM1/CDM	Altération du signal de transduction du complexe dystrophine-glycoprotéines	Nakamori (2017-2013-2008) et Yamashita (2012)
	Exon 12	Exclusion			
<i>INSR</i>	Exon 11	Exclusion	Biopsie et myotubes dérivés de patient DM1/CDM	Insulino-résistance ? (diminution de la réponse métabolique à l'insuline)	Nakamori (2017-2013) et Savkur (2001)
<i>LDB3 (ZASP)</i>	Exon 11	Inclusion	Biopsie muscle squelettique patient DM1/CDM	n.d.	Nakamori (2017-2013) et Lin (2006)
<i>MBNL1</i>	Exon 5	Inclusion	Biopsie muscle squelettique patient DM1/CDM	Défauts d'épissage	Nakamori (2017-2013) et Yamashita (2012)
	Exon 7	Inclusion			
	Exon 10	Inclusion			
<i>MBNL2</i>	Exon 7	Inclusion	Biopsie muscle squelettique patient DM1/CDM	Défauts d'épissage	Nakamori (2017-2013) et Lin (2006)
<i>MYOM1</i>	Exon 17a	Inclusion	Biopsie muscle squelettique patient DM1/CDM	Instabilité de la bande M sarcomérique	Nakamori (2017) et Koebis (2011)
<i>NCOR2</i>	Exon 10	Inclusion	Biopsie muscle squelettique patient DM1/CDM	n.d.	Nakamori (2017) et Yamashita (2012)
<i>NFIX</i>	Exon 7	Inclusion	Biopsie muscle squelettique patient DM1/CDM	n.d.	Nakamori (2017-2013) et Yamashita (2012)
<i>NRAP</i>	Exon 12	Exclusion	Biopsie muscle squelettique patient DM1/CDM	n.d.	Nakamori (2017-2013), Yamashita (2012) et Lin (2006)
<i>PDLIM3</i>	Exon 5	Exclusion	Biopsie muscle squelettique patient DM1/CDM	n.d.	Nakamori (2017-2013), Yamashita (2012) et Lin (2006)
<i>PKM</i>	Exon 10	Inclusion	Biopsie muscle squelettique patient DM1	Altération de l'homéostasie du glucose et atrophie des fibres musculaires de type 1	Gao & Cooper (2013)
<i>RyR1</i>	Exon 70	Exclusion	Biopsie et myotubes dérivés de patient DM1	Faiblesse musculaire liée à une diminution des contractions	Nakamori (2013) Kimura (2005)
<i>SCN5A</i>	Exon 6a/6b	Inclusion/exclusion	Biopsie muscle cardiaque patient DM1	Arythmie et défaut de conduction cardiaque	Freyermuth (2016)
<i>TNNT2 (cTNT)</i>	Exon 5	Inclusion	Biopsie muscle squelettique patient DM1	Défauts cardiaques ?	Bosè (2019), Ho (2004), Philips (1998)
<i>TTN</i>	Exon M5	Inclusion	Biopsie muscle squelettique patient DM1/CDM	n.d.	Nakamori (2017-2013) et Lin (2006)

Le défaut d'épissage du gène *BIN1* (ou Amphiphysine 2) a également été associé à la faiblesse musculaire observée chez les patients DM1. Ce gène code une protéine ayant un rôle majeur dans la biogenèse des tubules transverses (tubules T) qui sont des invaginations de la membrane des cellules musculaires. Ces structures sont essentielles pour le couplage excitation-contraction (**Figure 8-B**). L'exon 11 est très important pour l'activité de la protéine car il code un domaine de liaison à des phospholipides membranaires (les phosphoinositides) permettant la formation de ces invaginations (Lee et al. 2002). Dans le cadre de la DM1, un défaut d'inclusion de l'exon 11 de BIN1 a été observé chez les patients et provoque ainsi la désorganisation de ces structures, participant ainsi à la faiblesse musculaire caractéristique de la DM1 (Fugier et al. 2011).

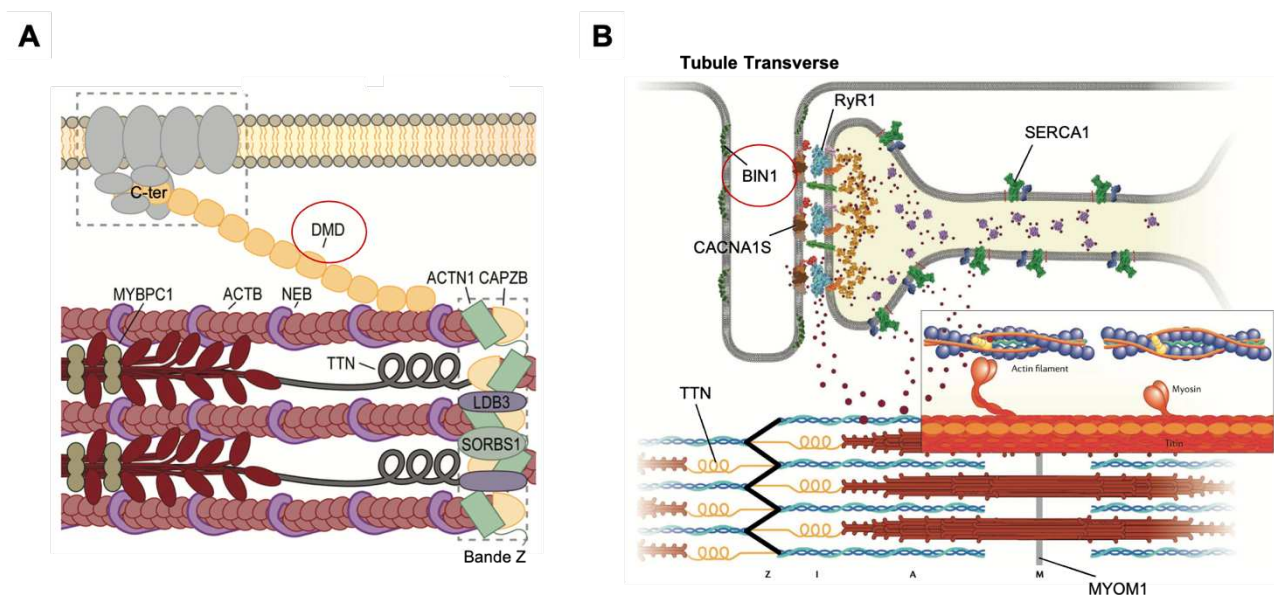


Figure 8 : Importance de DMD et BIN1 dans la structure et l'homéostasie du muscle

(A) La dystrophine, encodée par le gène *DMD* permet le maintien structurel des fibres musculaires en reliant les protéines du cytosquelette à la membrane sarcoplasmique via le complexe dystrophine-glycoprotéines (les glycoprotéines sont encadrées en haut à gauche). Dans le cadre de la DM1, l'exclusion de l'exon 78 va entraîner la modification conformationnelle de la queue C-terminale, déstabilisant ainsi la structure des fibres musculaires avec notamment la désorganisation de la bande Z (encadrée en bas à droite). (B) BIN1 permet la formation des tubules T au niveau de la membrane des cellules musculaires, structures essentielles pour le couplage excitation-contraction. L'exclusion de l'exon 11 chez les patients DM1 va induire une isoforme inactive de BIN1 aboutissant à l'altération des tubules T et par conséquent à une faiblesse musculaire. D'autres protéines présentant également un défaut d'épissage dans les muscles de patients DM1 sont indiquées. *Adapté de (Thomas et al. 2018; Jungbluth et al. 2018).*

D'autres protéines du cytosquelette ou encore des canaux ioniques membranaires faisant partie de ces structures, sont affectées également par des défauts d'épissage alternatif au niveau de biopsies musculaires de patients DM1. A titre d'exemple, une sous-unité du récepteur à la dihydropyridine DHPR (CACNA1S), du récepteur à la ryanodine (RyR1) ou

encore de la pompe Ca^{2+} -ATPase SERCA1 ont été identifiés comme présentant un défaut d'inclusion de l'exon 29, 70 et 11 respectivement. CACNA1S est situé sur la membrane des cellules musculaires au niveau des tubules T tandis que les deux autres sont situés sur la membrane du réticulum sarcoplasmique. Ces trois canaux calciques jouent un rôle essentiel dans le couplage excitation-contraction et dans l'homéostasie calcique du muscle (**Figure 8-B** et **Tableau 1**). Une étude a montré que l'exclusion de l'exon 29 de CACNA1S entraînait une augmentation du flux calcique transitoire à l'origine de défauts de contraction et d'une faiblesse musculaire (Tang et al. 2012). En revanche, concernant RyR1, seule une légère diminution de l'activité du canal a pu être observée. De la même manière, aucune étude n'a pu démontrer à ce jour la corrélation entre le défaut d'épissage de l'exon 11 de SERCA1 et les atteintes musculaires des patients DM1 (Kimura et al. 2005).

Par ailleurs, le défaut d'épissage de l'exon 11 du récepteur à l'insuline (INSR) observé chez les patients DM1 est souvent associé avec l'insulino-résistance caractéristique de la maladie. L'exclusion de l'exon 11 du INSR conduit à l'expression de l'isoforme fœtale du récepteur, INSR-A (contrairement à l'isoforme adulte INSR-B). Cette isoforme fœtale possède une affinité et une réponse à l'insuline plus faibles que l'isoforme adulte. Il a en effet été montré que cette isoforme était majoritaire dans les cellules musculaires dérivées de patients DM1 et que ces cellules DM1 présentaient une réponse métabolique à l'insuline plus faible que des cellules musculaires saines (Savkur et al. 2001). Néanmoins, aucune expérience induisant l'exclusion spécifique de l'exon 11 du INSR dans des cellules saines n'a été effectuée afin de démontrer l'effet fonctionnel de ce défaut d'épissage. Pourtant, celui-ci est encore considéré aujourd'hui comme le défaut d'épissage responsable de l'insulino-résistance des patients DM1, et compte parmi les défauts d'épissage majeurs de la pathologie.

De manière générale, la splicéopathie causée par l'altération de nombreuses protéines de liaison à l'ARN joue un rôle majeur dans la physiopathologie de la DM1. Cependant, ces dernières années, d'autres mécanismes moléculaires ont été décrits comme pouvant également contribuer à la mise en place de la pathologie.

1.3.3.4. Autres mécanismes pathologiques

Des études ont révélé l'existence d'une transcription bidirectionnelle du gène *DMPK* en observant la présence de transcrits *DMPK* anti-sens constitués de longues répétitions CAG. De plus, un mécanisme moléculaire appelé "RAN-Translation" (*Repeat-Associated Non-ATG Translation*) a été décrit comme étant capable d'initier la traduction des répétitions trinuécléotidiques dans tous les cadres de lecture possibles et indépendamment de la présence d'un codon START (ATG). Ainsi, les répétitions CUG présentes sur les transcrits *DMPK* sens vont pouvoir être traduites en longues chaînes polypeptidiques de différentes

natures : polyleucine (polyL), polycystéines (polyC) et polyalanines (polyA) (Pearson 2011). De la même façon, les répétitions CAG des transcrits DMPK anti-sens pourraient également être traduites, produisant alors des chaînes polyglutamine (polyQ), polysérine (polyS) et polyA (**Figure 9**). Ces chaînes polypeptidiques vont s'accumuler entre elles et former des agrégats nucléaires et cytoplasmiques dont certains ont pu être observés dans quelques biopsies musculaires de patients DM1 ainsi que dans des leucocytes circulants issus de patients. Ces observations ont pu être confirmées dans des myoblastes dérivés de patients ainsi que dans des modèles murins de la maladie. Dès lors, l'hypothèse mécanistique serait que ces agrégats protéiques soient à l'origine d'une protéotoxicité chez les patients DM1, hypothèse renforcée par la co-localisation de ces polypeptides issus de la RAN-translation avec des marqueurs apoptotiques (Sicot et al. 2011; Zu et al. 2011).

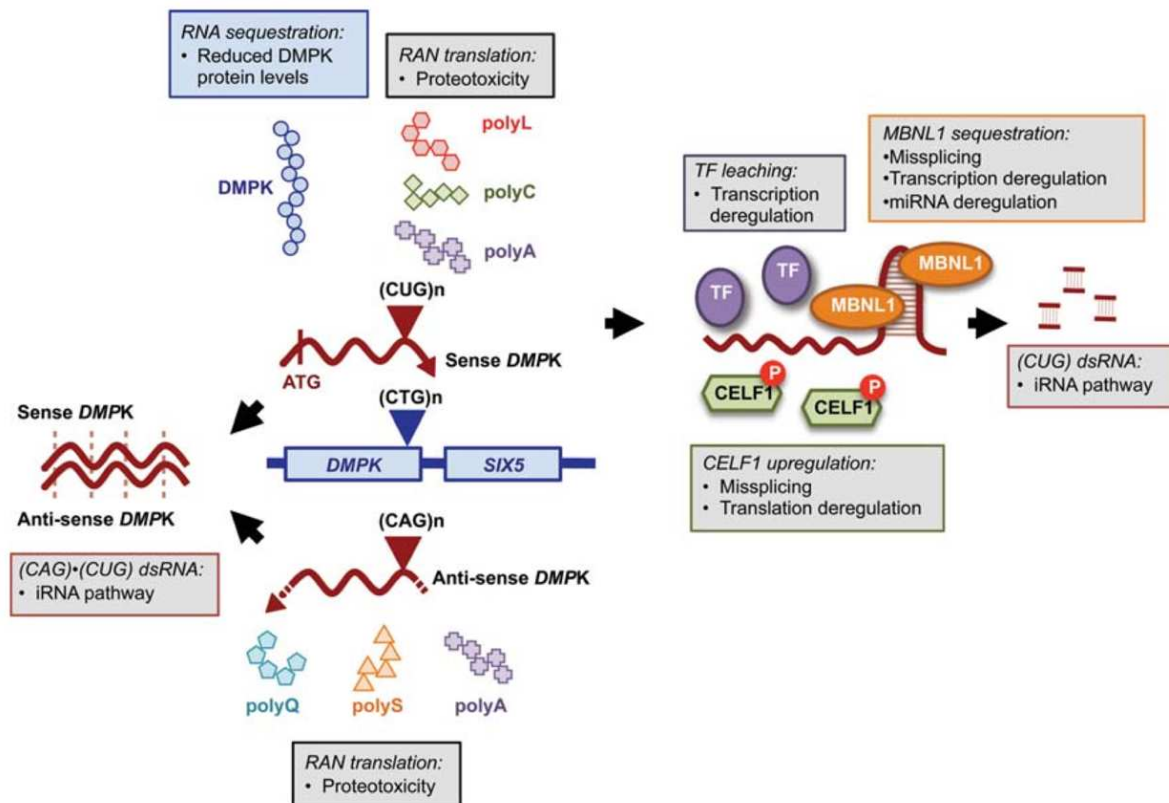


Figure 9 : Récapitulatif des mécanismes pathologiques impliqués dans la DM1

L'expansion des répétitions CTG va être à l'origine de l'haplo-insuffisance de la protéine DMPK ainsi que nombreux défauts d'expression géniques et d'épissage alternatif via l'altération de protéines de liaison à l'ARN. À ces mécanismes s'ajoutent la RAN-translation et l'ARN interférence. *Adapté de (Sicot et al. 2011).*

Parallèlement, une étude a montré que les structures d'épingles à cheveux formées par les répétitions CUG (et CAG dans le cas des transcrits anti-sens) pouvaient être clivées par la ribonucléase Dicer, impliquée dans le processus d'ARN interférence (*iRNA pathway*), pour

ainsi générer de petits ARN interférents (siRNA) (Krol et al. 2007). Ces siRNA sont capables de se lier à des ARNm cibles contenant des motifs de séquences complémentaires et pourraient alors interférer avec leur expression et contribuer à la physiopathologie de la DM1 (**Figure 9**). Cependant, les cibles potentiellement pathologiques de ces siRNA n'ont pas encore été identifiées (Sicot et al. 2011).

1.4. Modèles d'étude de la DM1

L'ensemble des hypothèses mécanistiques présentées ci-dessus découlent de l'utilisation de différents modèles animaux et cellulaires, qui comme nous le verrons plus tard sont également déterminants pour l'identification de nouvelles pistes thérapeutiques. Les modèles cellulaires sont importants pour décrypter les différents mécanismes de la physiopathologie humaine de la DM1 ainsi que pour tester diverses approches thérapeutiques avant de les valider dans des modèles animaux *in vivo*. Dans le cadre de l'étude de la DM1, l'expression exogène de différentes tailles de répétitions CTG de façon stable ou transitoire dans des lignées cellulaires humaines, telles que des cellules HeLa ou HEK, mais également des cellules murines, a été grandement utilisée comme modèle de la pathologie. Cependant, ces modèles de surexpression présentent l'inconvénient de niveaux d'expression des répétitions CTG très variables, n'étant pas sous le contrôle du promoteur *DMPK* endogène (Matloka et al. 2018).

Pour pallier ces limitations, des cellules primaires isolées directement à partir de tissus de patients sont également beaucoup utilisées. Ce sont la plupart du temps des fibroblastes issus du derme ou des myoblastes issus de tissus musculaires. Les fibroblastes ont l'avantage d'être plus accessibles et sont plus simples à manipuler en culture cellulaire. Néanmoins, l'utilisation des myoblastes est aussi d'un grand intérêt car ils permettent de reproduire différentes atteintes musculaires caractéristiques des patients DM1 telles que des défauts de différenciation myogénique ou d'épissage alternatif (Nakamori et al. 2007; Furling et al. 2001). Cependant, travailler avec des cellules primaires comporte également certaines contraintes non négligeables. En particulier, elles ont une capacité proliférative limitée car elles vont entrer en sénescence répllicative après un certain nombre défini de divisions cellulaires, phénomène connu aussi sous le nom d'horloge mitotique (Renault et al. 2000). Récemment, une étude a montré qu'il était possible de contourner cette limitation en générant des lignées immortalisées à partir de fibroblastes et de myoblastes primaires issus de patients DM1. La génération de ces lignées immortalisées est possible grâce à la surexpression de façon stable la télomérase humaine (hTERT) associée à une surexpression de CDK4 pour à la fois empêcher le raccourcissement des télomères responsable de la sénescence répllicative et pour inhiber la voie p16 afin d'augmenter leur capacité proliférative

(Arandel et al. 2017). Ces nouveaux modèles cellulaires représentent désormais une ressource illimitée tout en conservant la plupart des caractéristiques spécifiques des tissus DM1. Il n'a en revanche pas encore été déterminé si ce processus d'immortalisation entraînait des conséquences au niveau du comportement de ces cellules (Matloka et al. 2018). Enfin, les cellules souches pluripotentes humaines sont également une alternative très intéressante pour la modélisation de la DM1 car, en plus de représenter une ressource illimitée de matériel, elles ont la possibilité de se différencier dans tous les types cellulaires via leur capacité de pluripotence. Ceci représente un avantage considérable, notamment pour étudier des tissus très peu accessibles tels que le système nerveux central, et l'intérêt pour la DM1 est d'autant plus important en raison du caractère multi-systémique de la maladie. Les cellules souches pluripotentes humaines feront d'ailleurs l'objet de la troisième partie de l'introduction de cette thèse.

En parallèle de ces modèles cellulaires, plusieurs modèles animaux ont été développés au cours des 25 dernières années. Des organismes simples comme la drosophile (*Drosophila melanogaster*), le ver (*Caenorhabditis elegans*) ou encore le poisson-zèbre (*Danio rerio*) présentent les avantages d'être faciles à manipuler en laboratoire et permettent la génération rapide de nombreuses progénitures. Plusieurs modèles de drosophiles DM1 ont été générés en exprimant différentes tailles de répétitions CTG (de 11 jusqu'à 480 CTG) dans la région 3'UTR d'un gène rapporteur comme la GFP ou la DsRed. Ces études ont montré notamment que la sévérité des phénotypes observés dans ces modèles de drosophile DM1 était corrélée avec la longueur des répétitions CTG, de la même manière que chez les patients DM1. Les drosophiles dans lesquelles 480 répétitions ont été exprimées ont développé plusieurs phénotypes au niveau des muscles, des yeux et du système nerveux central (SNC) ainsi que des défauts d'épissage, concomittant avec l'observation de *foci* intranucléaires en association avec la séquestration des protéines muscleblind (Souidi et al. 2018). L'utilisation des drosophiles a également permis de mettre en évidence le mécanisme de toxicité des répétitions par ARN interférence (Yu et al. 2011). L'importance de ce mécanisme dans la pathologie a plus récemment été appuyée par une étude utilisant le ver *C. elegans* comme modèle de la DM1. En effet, la surexpression de 123 répétitions CTG couplées à la GFP chez *C. elegans* entraîne des défauts de motilité (Qawasmi et al. 2019). Une étude a observé des défauts au cours des étapes précoces du développement, au niveau des muscles et du SNC, en injectant directement 91 répétitions CUG dans les embryons de poissons-zèbres (Todd et al. 2014). Les résultats de ces différentes études suggèrent ainsi une importante conservation entre les espèces des mécanismes pathologiques liés à la toxicité des répétitions CTG. Ces modèles ont également permis d'identifier des composés potentiellement thérapeutiques capables par exemple de corriger certains défauts d'épissage (García-Alcover et al. 2014). Bien que l'utilité de ces modèles de la DM1 ne soit

plus à démontrer, ces petits organismes restent très différents des humains d'un point de vue physiologique, et les corrélations doivent être prudemment établies.

Parmi les modèles animaux développés pour l'étude de la DM1, les modèles murins ont largement contribué à l'état actuel des connaissances sur cette pathologie. En effet, différents modèles de souris ont été générés, soit via l'expression des répétitions CTG, soit en ciblant plus particulièrement des acteurs impliqués dans la physiopathologie de la DM1 tels que les gènes du locus *Dmpk* ou encore des gènes codant des protéines de liaison à l'ARN (**Tableau 2**). L'un des premiers modèles murins de la DM1 fut généré au début des années 2000 via l'expression du gène humain de l'Actine musculaire dans lequel 250 répétitions CTG ont été introduites dans la partie 3'UTR. Ces souris appelées HSA^{LR} (*Human Skeletal Actin Long Repeat*) développent ainsi essentiellement des atteintes musculaires telles que la myotonie ou une légère myopathie. Cependant, aucune atrophie ni faiblesse musculaire n'ont été décrites bien que des *foci* intranucléaires et des défauts d'épissage furent observés dans les cellules musculaires de ces souris (**Figure 10-A**) (Lin et al. 2006; Mankodi et al. 2000).

Une deuxième approche a reposé sur l'expression cette fois-ci systémique du locus humain de *DMPK* contenant différentes tailles de répétitions CTG (de 20 à plus de 1000 répétitions) aboutissant notamment à la génération des souris DM300 et DMSXL. L'expression des transcrits sous le contrôle du locus *DMPK* humain constitue le réel avantage de cette approche. Les souris DM300 présentent différentes atteintes au niveau des muscles et du cerveau. Ce modèle a surtout permis l'étude de l'instabilité des répétitions CTG, notamment en donnant naissance aux souris DMSXL contenant plus de 1000 répétitions (Gomes-Pereira et al. 2007). Malgré une importante mortalité néonatale, les souris DMSXL représentent un des meilleurs modèles pour la forme congénitale de la DM1 car elles développent des phénotypes plus sévères, en particulier au niveau du système nerveux central avec notamment des troubles comportementaux, des défauts de plasticité synaptique ainsi que de nombreux défauts d'épissage (Hernández-Hernández et al. 2013). D'autres atteintes ont également été observées au niveau de la morphologie des jonctions neuromusculaires (JNM) par exemple (**Figure 10-B**) (Panaite et al. 2013). De manière intéressante ces souris sont plus petites que la normale et ce dès la naissance mais aucune explication à cette observation n'a jusqu'à aujourd'hui été décrite (Braz et al. 2018; Gomes-Pereira et al. 2007). Un troisième modèle consiste en l'expression de 960 répétitions CTG de façon inductible et conditionnelle, permettant une étude plus ciblée au sein de tissus spécifiques, principalement les muscles squelettiques et le cœur. Les souris EpA960 ont ainsi reproduit plusieurs atteintes musculaires et cardiaques caractéristiques de la DM1 (**Figure 10-C**) (Orengo et al. 2008).

Tableau 2 : Vue d'ensemble des différents modèles murins pour l'étude de la DM1

Modèle murin	Description	Caractéristiques associées à la DM1	Limitations	Références
Modèles d'expression de l'ARNm toxique (répétitions CUG)				
Tg26	26 copies du transgène <i>DMPK</i> humain avec 11 répétitions CTG.	Myotonie, myopathie (muscles sq., lisses et cardiaque), défaut d'expression de <i>Clcn1</i> .	Absence de <i>foci</i> intranucléaires.	O'Coirlain (2004) et Jansen (1996)
HSA ^{LR}	250 CTG via le promoteur du gène de l'actine sq. humaine.	Myotonie, myopathie, <i>foci</i> intranucléaires, défauts d'épissage.	Pas de faiblesse ni d'atrophie musculaires, pas de surexpression de CUG-BP1.	Mankodi (2000) et Lin (2006)
DM300	Locus du gène <i>DMPK</i> humain avec 300 CTG.	Myotonie, faiblesse musculaire, histologie musculaire DM1, <i>foci</i> intranucléaires, défauts d'épissage, instabilité somatique et intergénérationnelle des répétitions, anomalies du cerveau.	Nécessité de générer des souris homozygotes.	Vignaud (2010) et Seznec (2000-2001)
DMSXL	Lignée transgénique dérivée du modèle DM300 avec plus de 1000 CTG.	Idem que le modèle DM300 avec une altération des jonctions neuromusculaires, des atteintes du système nerveux central (SNC), de nombreux défauts d'épissage dans le muscle sq. et le SNC, et surexpression de CUG-BP1 et ETR-3 dans le cortex frontal.	Forte mortalité néonatale et diminution importante de la taille des souris.	Hernandez-Hernandez (2013), Panaite (2013), Huguet (2012) et Gomes-Pereira (2007)
EpA960	Surexpression inducible tissu-spécifique (promoteur CMV) de 960 CTG dans l'exon 15 du gène <i>DMPK</i> humain.	Myotonie, myopathie (muscles sq. et cardiaque), faiblesse et atrophie musculaires, histologie musculaire DM1, arythmie, défaut de conduction cardiaque, <i>foci</i> intranucléaires, défauts d'épissage, surexpression de CUG-BP1 et ETR-3.	Fuite d'expression du transgène, forte mortalité due à des complications cardiaques.	Orengo (2008) et Wang (2007)
TRE-EGFP-CTG ⁵ / CTG ²⁰⁰	Surexpression inducible tissu-spécifique (Tétracycline) de la région 3'UTR du gène <i>DMPK</i> humain avec 5 ou 200 CTG.	Myotonie, myopathie, histologie musculaire DM1, défaut de conduction cardiaque, <i>foci</i> intranucléaires (200 CTG), défauts d'épissage, surexpression de Nkx2-5 (cœur) et CUG-BP1 (muscle sq.).	La surexpression de 5 répétitions CTG non pathologiques entraîne des défauts associés à la DM1.	Kim (2013) et Mahadevan (2006)
Modèles concernant les gènes du locus <i>Dmpk</i>				
<i>Dmpk</i> (-/-)	Perte de fonction <i>Dmpk</i> .	Myopathie légère, dégénérescence musculaire et fibrose chez les souris les plus âgées. Surexpression de CUG-BP1.	Pas d'atrophie musculaire, de myotonie, ni de cataracte.	Jansen (1996) et Reddy (1996)
<i>Six5</i> (-/-)	Perte de fonction <i>Six5</i> .	Atrophie testiculaire, défaut de maturation de la spermatogenèse et défaut de conduction cardiaque.	Aucune atteinte des muscles sq. et cataracte différente de celle associée à la DM1.	Sarkar (2004) et Wakimoto (2002)
Modèles concernant les protéines de liaison à l'ARN				
Tg <i>Celf</i>	Surexpression de <i>Celf1</i> dans le muscle sq. ou le cœur.	Atrophie et faiblesse musculaires, défaut de conduction cardiaque, faiblesses respiratoires, défauts d'épissage.	Pas de myotonie, forte mortalité.	Koshelev (2010), Ward (2010) et Timchenko (2004)
<i>Mbnl1</i> (-/-)		Myotonie, histologie musculaire DM1, cataracte sous capsulaire, défauts d'épissage.	Ne récapitulent que les défauts associés aux protéines <i>Mbnl</i> . La double déplétion de <i>Mbnl1</i> et <i>Mbnl2</i> induit une très forte mortalité embryonnaire.	Kanadia (2003) et Lee (2013)
<i>Mbnl2</i> (-/-)	Perte de fonction des protéines <i>Mbnl</i> .	Défauts cognitifs d'apprentissage et troubles du sommeil.		Charizanis (2012)
<i>Mbnl3</i> (-/-)		L'absence de l'isoforme nucléaire altérerait la régénération musculaire.		Poulos (2013)

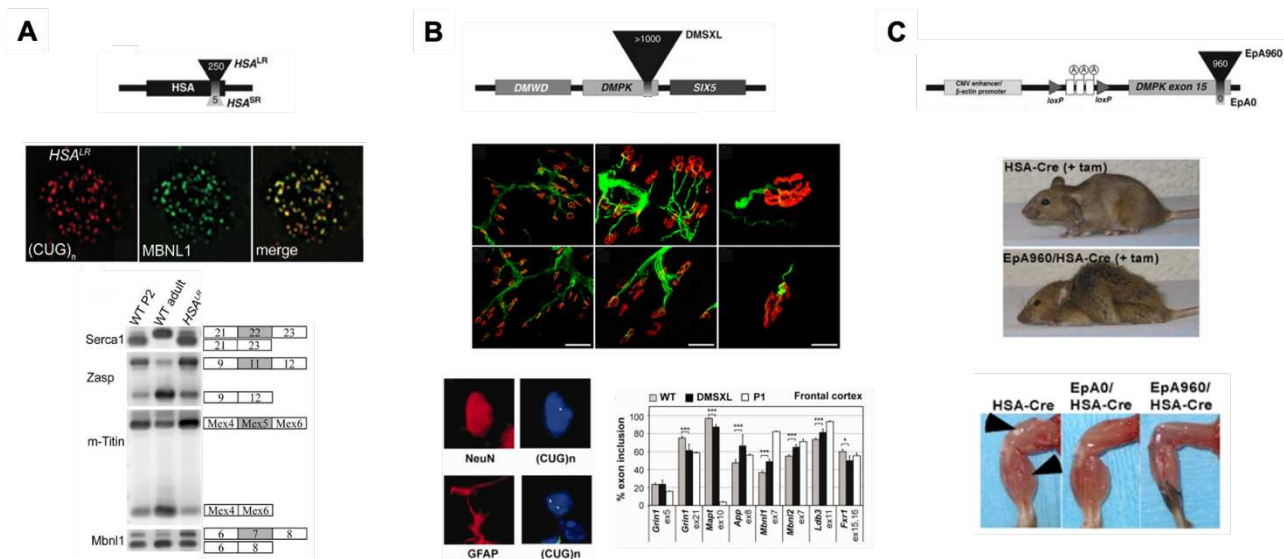


Figure 10 : Exemples d'atteintes observées dans les modèles murins de la DM1

(A) Le modèle murin HSA^{LR} a permis de reproduire certaines atteintes caractéristiques de la DM1 comme la séquestration des protéines MBNL par les *foci* dans les noyaux des cellules musculaires, contribuant à plusieurs défauts d'épissage. **(B)** Le modèle multi-systémique DMSXL présente des phénotypes plus sévères tels que des modifications morphologiques des JNM ainsi que des atteintes du SNC avec la présence de *foci* intranucléaires dans des cellules neuronales et des défauts d'épissage dans le cortex frontal par exemple. **(C)** Après induction au tamoxifène (+tam) dans le muscle squelettique, les souris EpA960 développent une atrophie musculaire sévère, notamment au niveau des membres postérieurs. *Adapté de (Hernández-Hernández et al. 2013; Panaite et al. 2013; Orengo et al. 2008; Gomes-Pereira et al. 2007; Lin et al. 2006).*

Enfin, d'autres modèles murins *Dmpk*(-/-) et *Six5*(-/-) ont été générés en inactivant le gène *Dmpk* ou le gène *Six5* respectivement. Comme discuté précédemment, ces souris ne développent que des phénotypes légers et ne reproduisent pas la grande majorité des atteintes caractéristiques de la DM1. Par la suite, des souris récapitulant la surexpression de Cug-bp1 ou encore la perte de fonction des protéines Mbnl ont vu le jour dans le but d'étudier plus précisément la contribution de ces protéines de liaison à l'ARN dans la physiopathologie de la DM1. Ces modèles ont également beaucoup été utilisés pour l'investigation des défauts d'épissage spécifiques de la DM1. Les modèles murins de perte de fonction des protéines Mbnl seront plus amplement décrits dans la deuxième partie de cette introduction.

1.5. Stratégies thérapeutiques

À l'heure actuelle, il n'existe aucun traitement curatif pour la DM1. Les seuls traitements dont les malades peuvent actuellement bénéficier sont essentiellement symptomatiques. De manière générale, les symptômes les plus fréquemment pris en charge sont les atteintes

musculaires. Par exemple, la Mexiletine et la Lamotrigine sont prescrits aux patients DM1 comme agents anti-myotoniques. Ce sont deux inhibiteurs des canaux sodiques qui sont déjà utilisés en cardiologie (Hoffman and Kaminski 2012; Andersen et al. 2017). Plusieurs recherches ont tenté de mettre au point un traitement contre la faiblesse et l'atrophie musculaire. Malheureusement, ni la DHEA (Déhydroandrosténéidione) ni l'utilisation du rhIGF (*recombinant human Insulin Growth Factor 1*) n'ont abouti à des résultats concluants après avoir atteint la phase II des études cliniques (Heatwole et al. 2011; Pénisson-Besnier et al. 2008). Les patients sont également pris en charge pour leurs atteintes cognitives, en particulier concernant les troubles du sommeil et les troubles de l'humeur, principalement avec la prescription de méthylphényhydrate et d'antidépresseurs.

C'est grâce aux différentes études, menées au cours des deux dernières décennies et ayant permis d'identifier et de mieux comprendre les mécanismes pathologiques impliqués dans la DM1, que de nouvelles stratégies thérapeutiques ont pu émerger, permettant le développement clinique de candidats-médicaments. Ces nouvelles approches consistent à cibler différents acteurs impliqués dans la pathologie. Ainsi, plusieurs possibilités sont envisageables comme par exemple cibler la mutation elle-même au niveau de l'ADN ou au niveau des ARNm *DMPK* toxiques, ou encore cibler les mécanismes dérégulés par la toxicité des transcrits CUG (**Figure 11**). De telles approches thérapeutiques ont pu être mises en place également grâce aux grandes avancées en matière de composés thérapeutiques, notamment avec le développement de petites molécules pharmacologiques ou encore avec l'utilisation d'oligonucléotides anti-sens (ASO).

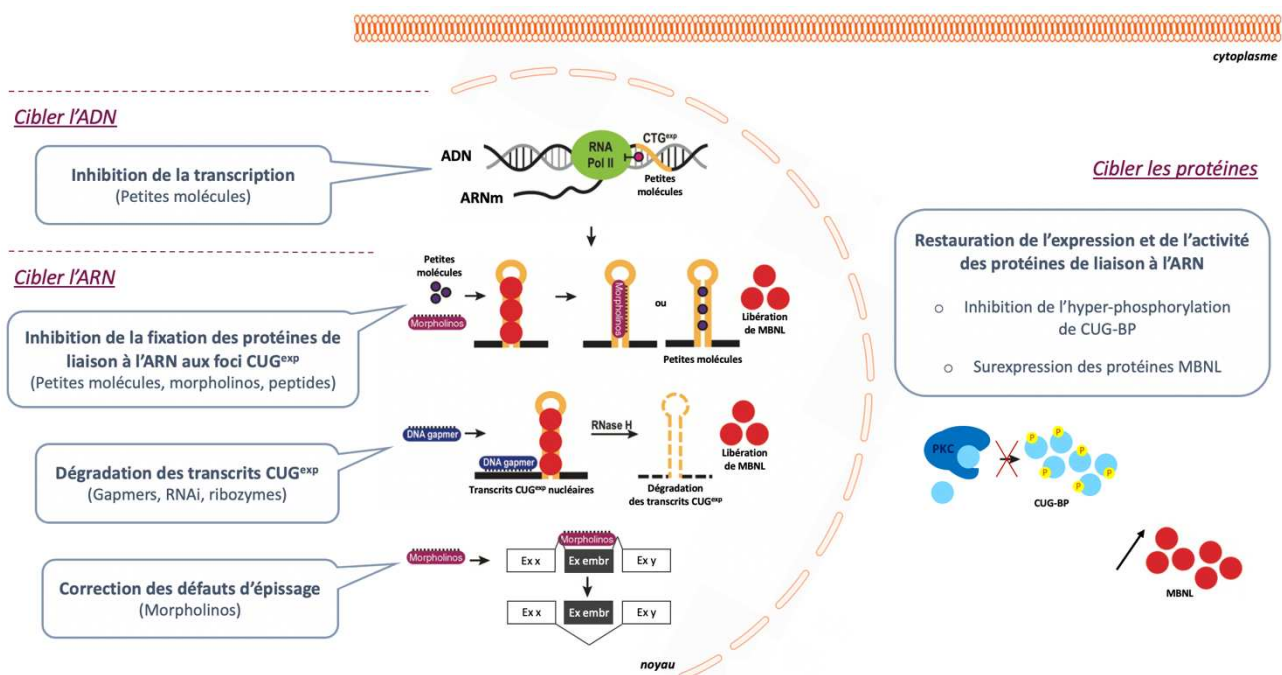


Figure 11 : Stratégies thérapeutiques pour la DM1

Adapté de (Thomas et al. 2018; Klein et al. 2015; Chau and Kalsotra 2015).

1.5.1. Cibler l'ADN *DMPK*

Une première approche consiste à agir au niveau de l'ADN muté, afin d'inhiber la transcription des répétitions CTG. Le blocage de la transcription a l'avantage de prévenir la production des ARNm toxiques ainsi que tous les effets qui en découlent, y compris certains mécanismes qui n'auraient peut-être pas encore été identifiés. De plus, la DM1 étant une maladie autosomale dominante, la mutation n'est donc présente que sur un seul des deux allèles, ce qui laisse supposer que de potentiels composés thérapeutiques puissent être administrés à des doses plus faibles (Reddy et al. 2019). Plusieurs petites molécules ont été identifiées pour inhiber la transcription de *DMPK*, telles que l'Actinomycine D (ActD) ainsi que la pentamidine et ses analogues (López-Morató et al. 2018). L'ActD est un antibiotique polypeptidique déjà utilisé comme médicament anticancéreux pour son activité anti-transcriptionnelle globale, sous le nom de Cosmegen. L'ActD s'intercale dans l'ADN au niveau des séquences riches en GC et bloque alors la progression de l'ARN polymérase II. L'hypothèse a ainsi été émise quant au potentiel thérapeutique de l'ActD dans la DM1 pour agir au niveau des expansions de répétitions CTG. Il a été montré dans le modèle murin HSA^{LR} traité avec l'ActD que les niveaux d'expression des transcrits CUG toxiques étaient significativement diminués et que certains défauts d'épissage étaient corrigés (Siboni et al. 2015). La pentamidine est elle-aussi approuvée par la FDA (agence américaine des produits alimentaires et médicamenteux) et est actuellement utilisée pour plusieurs types d'infections. Les souris HSA^{LR} ont également été traitées avec la pentamidine ou certains dérivés comme l'heptamidine. Globalement ces traitements ont entraîné une diminution des transcrits et des *foci* intranucléaires avec ici aussi la correction de certains défauts d'épissage. En revanche, en raison d'une toxicité non négligeable, aucun essai clinique n'est actuellement en cours. De plus, le mécanisme d'action de la pentamidine est encore inconnu. En effet, bien que de nombreux travaux suggèrent qu'elle agisse au niveau de l'ADN en inhibant ou en réduisant la stabilité de la transcription, d'autres études émettent l'hypothèse d'une potentielle action supplémentaire au niveau traductionnel (López-Morató et al. 2018). Plus récemment, avec les importants développements des systèmes d'édition génomique CRISPR/Cas9, de nouvelles stratégies thérapeutiques peuvent être envisagées. Concernant la DM1, l'excision des répétitions CTG du locus *DMPK* a récemment fait l'objet d'une preuve de concept *in vivo*, chez les souris DMSXL, en injectant des vecteurs AAV exprimant le système CRISPR/Cas9 au niveau du muscle tibial antérieur. Cette récente étude a notamment permis d'observer une diminution des *foci* dans les muscles ayant été injectés (Scrudato et al. 2019).

1.5.2. Cibler l'ARNm *DMPK*

La grande majorité des approches thérapeutiques développées récemment se sont focalisées sur les ARNm *DMPK* toxiques au travers de deux stratégies principales : soit la dégradation des répétitions CUG, soit le blocage stérique au niveau des structures en épingle à cheveux pour inhiber la liaison délétère des protéines séquestrées par ces répétitions. Différents outils ont été générés pour dégrader les répétitions CUG tels que les stratégies anti-sens, l'induction du mécanisme d'ARN interférence ou encore l'utilisation des ribozymes. Les premières expériences ont utilisé des shRNA (*short hairpin RNA*) et des siRNA de séquences complémentaires à celles des transcrits *DMPK* pour induire le mécanisme d'ARN interférence. Ces approches permettent d'empêcher la formation des *foci* et de restaurer certains défauts d'épissage (Sobczak et al. 2013; Langlois et al. 2005). Cependant elles nécessitent d'importants développements pour pouvoir être utilisées pour la thérapie humaine. L'utilisation de ribozymes, ARN pouvant agir en tant qu'enzymes, a également été envisagée, avec par exemple le ribozyme en tête de marteau capable de cliver l'ARNm *DMPK*, n'aboutissant malheureusement à aucun essai clinique (Langlois et al. 2003). D'autres alternatives ont par la suite été étudiées, notamment avec les hU7-snrRNA (*human U7 small nuclear RNA*) contenant une séquence poly-CAG et permettant d'augmenter la destruction spécifique des transcrits *DMPK* mutants, sans affecter les produits de l'allèle *DMPK* non-muté (François et al. 2011).

Plus récemment, les oligonucléotides anti-sens (ASO) ont fait l'objet de nombreuses études et représentent l'un des potentiels thérapeutiques les plus prometteurs. Ces derniers sont de courtes séquences d'ADN simple brin ayant une longueur allant de 18 à 30 nucléotides et capables de se lier spécifiquement à un ARNm cible, via la complémentarité des séquences. Ils vont ainsi former des complexes hybrides ADN/ARN qui seront reconnus par la RNase H qui va alors les dégrader. Il existe différents types d'ASO selon leur chimie (modification ou ajout de groupements chimiques) tels que les gapmers, les morpholinos, les PMO (*phosphorodiamidate morpholino*) ou encore les MOE (*2'-O-methoxyethyl*). Ces différentes modifications permettent d'augmenter la spécificité de l'ASO, sa stabilité ou encore sa pénétration dans les cellules et permettent aussi d'activer plus ou moins la dégradation via la RNase H. Par exemple, la présence de gapmers augmente l'activité RNase H et les MOE favorisent la stabilité de l'ASO (Scoles et al. 2019). Ainsi, l'administration systémique dans les souris HSA^{LR} d'un MOE-gapmer-ASO ciblant l'ARNm *DMPK* a entraîné une dégradation efficace des transcrits CUG avec une correction pratiquement complète des défauts moléculaires et fonctionnels de ces souris (Wheeler et al. 2012). La société *Ionis Pharmaceuticals* a par la suite développé un ASO, IONIS-DMPK_{RX}, capable d'induire le clivage des répétitions CUG (Jauvin et al. 2017; Pandey et al. 2015). Malheureusement, l'essai clinique a dû être arrêté à la suite de la phase I-IIa en raison d'une concentration trop faible de l'ASO au niveau des muscles. Cette société est actuellement en train de développer une nouvelle

technologie, LICA (Ligand Conjugated Antisens) permettant d'améliorer la délivrance dans des tissus spécifiques comme le muscle squelettique. Par ailleurs, une étude récente a démontré que le couplage d'un ASO ciblant les répétitions CUG avec un petit peptide riche en arginine (Pip6a) améliorait considérablement la délivrance de l'ASO dans les muscles squelettiques des souris HSA^{LR} après administration systémique (Klein et al. 2019).

Parmi les différents ASO, certains ont été utilisés non pas pour induire la dégradation des ARNm mais pour bloquer de façon stérique les répétitions CUG. Une étude a révélé qu'un morpholino composé de 25 répétitions CAG était capable de se lier aux répétitions CUG et de déstabiliser les structures en épingle à cheveux. Cet encombrement stérique empêche alors les interactions avec les protéines MBNL1 et permet alors de corriger certains défauts d'épissage (Wheeler et al. 2009). Il existe également d'autres composés capables de bloquer les interactions entre les protéines de liaison à l'ARN et les triplets CUG. En effet, plusieurs petites molécules, tels que des inhibiteurs de CDK12, ont été identifiées par des criblages à haut-débit, notamment sur leur capacité à perturber la formation des *foci* (Ketley et al. 2020). Elles représentent également des candidats thérapeutiques prometteurs et ont l'avantage de pouvoir être administrées par voie orale en étant efficacement libérées dans les tissus. De plus, elles possèdent des coûts de fabrication généralement plus faibles (Reddy et al. 2019). Par exemple, il a été montré que l'érythromycine avait une grande affinité pour les répétitions CUG, induisant ainsi la libération des protéines MBNL et la correction de certains défauts d'épissage *in vitro*. Son administration par voie orale chez les souris HSA^{LR} a permis de confirmer ces résultats et une amélioration de la myotonie a également été observée (Nakamori et al. 2016). Le repositionnement de cet antibiotique sous le nom de MYD-0124 fait actuellement l'objet d'un essai clinique de phase II (jRCT2051190069) au Japon sur 30 patients DM1. Les chercheurs ont également montré qu'une autre molécule, la furamidine, était capable de corriger plusieurs défauts d'épissage dans des cellules musculaires issues de patients ainsi que dans le modèle murin HSA^{LR} (Jenquin et al. 2018). Ils ont pu démontrer par la suite que la combinaison de ces deux molécules provoquait ainsi un effet synergique (Jenquin et al. 2019).

Par ailleurs, un essai clinique de phase I a très récemment été conduit sur 12 patients DM1 aux États-Unis par la société *Expansion Therapeutics* pour tester l'effet d'une autre petite molécule, l'ERX-963, capable également de se lier aux répétitions CUG de l'ARNm *DMPK*. La société cherche à évaluer le potentiel effet de ce candidat-médicament sur l'hypersomnie et la somnolence diurne observées chez les patients. Les données sont pour le moment en cours d'analyse. L'ERX-963 et l'érythromycine font ainsi partie des 4 candidats-médicaments actuellement en cours d'évaluation thérapeutique pour la DM1 (**Figure 12**).

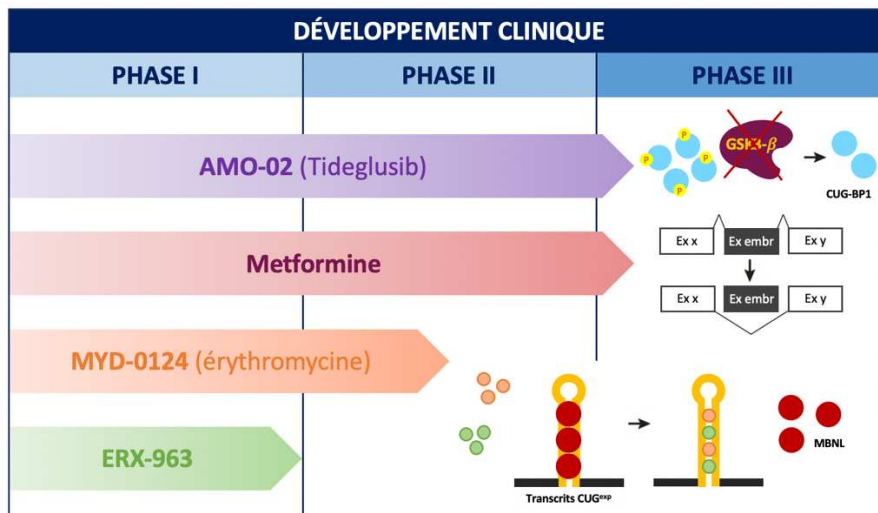


Figure 12 : Essais cliniques en cours de développement pour la DM1

Le Tideglusib inactive la protéine GSK3- β , permettant ainsi de normaliser l'expression et l'activité de CUG-BP1. La metformine permet de corriger plusieurs défauts d'épissage caractéristiques de la maladie mais son mécanisme d'action reste encore méconnu. L'érythromycine et l'ERX-963 sont capables de se lier aux répétitions CUG de l'ARNm *DMPK* entraînant la libération des protéines MBNL.

1.5.3. Cibler les protéines de liaison à l'ARN

Une autre approche thérapeutique consiste à cibler les protéines de liaison à l'ARN afin de restaurer leur niveau d'expression fonctionnel. Un essai clinique de phase II-III mis en place par la société *AMO-Pharma* est actuellement en cours dans plusieurs pays (Canada, États-Unis et Royaume-Uni) avec 56 patients atteints de la forme congénitale de la DM1 (**Figure 12**). Cet essai clinique porte sur le tideglusib (ou AMO-02), composé inhibant la voie GSK3- β et capable de réguler l'expression du CUG-BP1. Bien qu'on ne connaisse pas le mécanisme exact par lequel la voie GSK3- β agit sur l'expression de CUG-BP1, il a été démontré que le tideglusib augmentait significativement la survie des souris DMSXL et améliorait leur croissance et leur activité neuromotrice (Wang et al. 2019b). Les résultats de la phase II sont très encourageants avec une amélioration au niveau du SNC et des symptômes neuromusculaires des patients (Horrigan et al. 2020). Par ailleurs, une étude a révélé que la surexpression de la protéine MBNL1 chez les souris HSA^{LR} conduisait à la correction de défauts d'épissage et améliorait les atteintes musculaires de ces souris (Kanadia et al. 2006). Ainsi, plusieurs petites molécules ont été identifiées dans le but d'induire la surexpression des protéines MBNL1 comme par exemple le phénylbutazone (PBZ). Les souris HSA^{LR} traitées avec le PBZ ont montré des niveaux élevés de l'expression de Mbnl1 avec une correction partielle des défauts d'épissage (Chen et al. 2016). Plus récemment, un criblage par cytométrie en flux a permis d'identifier deux petites molécules inhibitrices des histones désacétylases (HDACi), l'ISOX et le Vorinostat, capables également d'entraîner la surexpression de MBNL1 avec la correction de certains défauts d'épissage dans

des fibroblastes dérivés de patients DM1 (Zhang et al. 2017a). Toutefois, ces résultats n'ayant pas été confirmés à ce jour, ils n'ont pas abouti au développement d'un essai clinique.

1.5.4. Cibler les défauts d'épissage

Une dernière stratégie thérapeutique consiste à agir plus en aval, au niveau des épissages dérégulés. Les ASO ont montré une utilité prometteuse avec les approches de saut d'exon. En se fixant et en masquant les séquences régulatrices de l'épissage (par exemple les sites accepteurs et les sites donneurs) les ASO vont ainsi perturber la machinerie de l'épissage alternatif et forcer artificiellement l'exclusion d'un exon spécifique. Cette stratégie de saut d'exon a été réalisée chez la souris HSA^{LR} pour forcer l'exclusion de l'exon 7a du Clcn1, rétablissant ainsi l'isoforme adulte du canal chlore et corrigeant la myotonie chez ces souris (Wheeler et al. 2007b). Cependant, la grande difficulté de cette approche réside dans le nombre important de défauts d'épissage présents chez les patients DM1. Ainsi, cibler tous ces défauts d'épissage semble compliqué à mettre en place. Par une approche basée sur le repositionnement de molécules thérapeutiques, l'équipe dans laquelle j'ai effectué cette thèse a montré que la metformine, un antidiabétique communément prescrit et bien toléré, présentait la capacité de pouvoir normaliser un certain nombre de défauts d'épissage associés à la DM1. Les mécanismes par lesquels la metformine régule ces épissages alternatifs semblent impliquer le facteur d'épissage RBM3 dont l'expression serait modulée via l'activation de la voie AMPK (Laustriat et al. 2015). Sur la base de ces résultats, un essai clinique de phase II a été conduit en collaboration avec l'hôpital Henri Mondor. Cet essai, nommé MYOMET, a été mené sur 40 patients adultes DM1 (20 placébos et 20 traités avec des doses croissantes de metformine) (**Figure 12**). L'efficacité du traitement a été évaluée notamment sur le test des '6 minutes de marche'. Cette étude a montré que les patients ayant reçu le traitement ont gagné une distance moyenne de marche de plus de 30 mètres (contre 3 mètres en moyenne pour les patients ayant reçu le placebo). Ce gain au niveau de la motricité est corrélé avec l'observation faite de l'effet de la metformine sur la posture globale des patients (Bassez et al. 2018). Un essai de phase III en Italie serait actuellement en préparation (données communiquées au cours du dernier congrès de l'IDMC [International Myotonic Dystrophy Consortium]).

L'ensemble de ces études illustre la difficulté de développer des pistes thérapeutiques prometteuses pour la DM1, en particulier en raison de la grande complexité des mécanismes physiopathologiques de la maladie. Il apparaît aussi, parmi les différentes cibles thérapeutiques, que la famille des protéines MBNL contribue de manière particulièrement importante aux atteintes des patients DM1. Cette famille de protéines a en effet fait l'objet de nombreuses études auxquelles nous allons nous intéresser dans la partie suivante.

2. La famille des protéines MBNL

Comme nous l'avons évoqué précédemment, cette famille de protéines a été identifiée pour être impliquée dans la DM1 au début des années 2000, en particulier pour sa capacité à se lier spécifiquement aux répétitions CUG sous leur forme double brin (Miller et al. 2000).

2.1. Les gènes MBNL

Le gène *muscleblind* fut initialement découvert quelques années auparavant chez la drosophile *D. melanogaster* où il a été montré qu'il jouait un rôle essentiel pour le développement des photorécepteurs et des muscles (Artero et al. 1998; Begemann et al. 1997). En effet, les gènes *MBNL* sont très conservés au cours de l'évolution. D'autres gènes orthologues ont été identifiés dans différents organismes comme *C. elegans* par exemple, chez lesquels on retrouve le gène unique *muscleblind*. En revanche, chez les mammifères, on dénombre trois gènes différents codant pour les protéines MBNL : les gènes *MBNL1*, *MBNL2* et *MBNL3* (Pascual et al. 2006). Dans la suite de ce chapitre, nous nous intéresserons tout particulièrement à ces trois formes présentes chez les mammifères.

2.1.1. Organisation génomique

Les trois gènes paralogues humains *MBNL* codent pour des protéines de liaison à l'ARN et plus particulièrement pour des protéines à doigts de zinc de type CCCH. Bien que chacun de ces gènes soit situé sur un chromosome différent (chromosome 3, 13 et X pour *MBNL1*, *2* et *3* respectivement), les trois membres de cette famille partagent plusieurs similarités au niveau de leur organisation structurale (**Figure 13**). En effet, ils sont constitués de 4 domaines en doigt de zinc (ZnF) qui fonctionnent en tandem. On retrouve ainsi les deux paires ZnF1/2 et ZnF3/4 séparées par une séquence de liaison également appelée "linker". Chaque domaine ZnF est constitué de motifs particuliers incluant trois cystéines (C) suivies d'une histidine (H), c'est la raison pour laquelle on parle de doigts de zinc de type CCCH (Konieczny et al. 2014). Il a été montré que les domaines ZnF1 et ZnF3 étaient très similaires, avec le même nombre d'acides aminés entre les cystéines et l'histidine : $CX_7CX_6CX_3H$, où X représente n'importe quel acide aminé. De la même manière les deux domaines ZnF2 et ZnF4 possèdent eux des motifs $CX_7CX_4CX_3H$ (Fardaei et al. 2002). Ces domaines sont très importants car ce sont eux qui permettent la liaison des protéines MBNL à l'ARN en reconnaissant un motif consensus 5'- YGCY -3' (Y étant une pyrimidine C ou T). La séquence "linker" codée par l'exon 3 qui relie ces deux paires de domaines ZnF est également très importante pour l'interaction protéine/ARN. Il a en effet été montré que sa délétion réduisait

fortement l'affinité de liaison de MBNL1 pour ses cibles ARNm (Tran et al. 2011). Les trois paralogues possèdent également un domaine dans la partie C-terminale leur permettant d'interagir entre eux. Cette séquence de liaison homotypique est codée par l'exon 7 et permet ainsi l'homodimérisation des protéines MBNL. Enfin, l'exon 5 et une partie de l'exon 6, composé d'une séquence très conservée des 5 acides aminés KRAEK, codent un signal de localisation nucléaire bipartite permettant la translocation des protéines MBNL au niveau du noyau de la cellule. Toutefois, la protéine MBNL3 ne possède pas ce signal (Konieczny et al. 2014; Tran et al. 2011). Ces différentes régions sont donc très importantes pour la fonction des protéines MBNL, autant au niveau de la localisation de la protéine, de la liaison avec l'ARN ainsi qu'au niveau de la régulation de l'épissage alternatif.

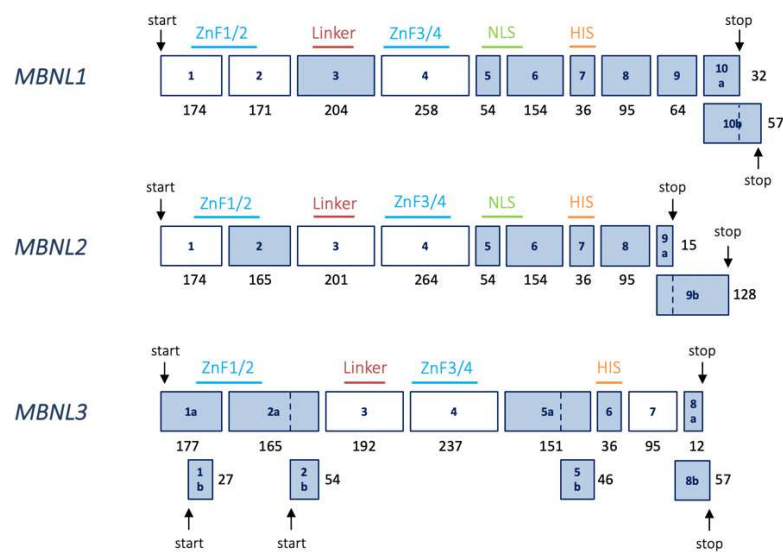


Figure 13 : Organisation des exons codants des gènes *MBNL*

Les exons alternatifs ou modifiés par la présence d'un codon START alternatif ou d'un décalage du cadre de lecture sont représentés en bleu. La taille (nombre de nucléotides) de chaque exon est indiquée. Les différents codons START et STOP sont également indiqués par des flèches. ZnF: domaines en doigt de zinc, Linker: séquence de liaison, NLS: séquence de localisation nucléaire, HIS: séquence d'interaction homotypique. *Adapté de (Mérien et al., in submission).*

Bien qu'ayant un rôle principal de facteur d'épissage, les ARNm *MBNL* sont eux aussi soumis au phénomène de l'épissage alternatif. Par exemple, le gène *MBNL1* humain est composé de 10 exons codants dont 6 sont sujets à l'épissage alternatif (7 si on compte l'exon 10 et ses deux tailles possibles en fonction du cadre de lecture modifié par l'exclusion de l'exon 6 ou de l'exon 8) (Fardaei et al. 2002) (**Figure 13**). L'épissage des ARNm *MBNL* va ainsi donner lieu à différentes isoformes allant d'environ 300 à 390 acides aminés pour MBNL1. De manière générale, les isoformes majeures de MBNL1 et MBNL2 ont été décrites en fonction de la présence ou non des exons 5, 7 et 8. Ainsi, au cours du développement, différentes isoformes des protéines MBNL seront exprimées, modulant alors leur localisation

et leurs fonctions (Tran et al. 2011; Pascual et al. 2006). Il a par exemple été montré que les isoformes de MBNL1 de 42 et 43 kDa contenant l'exon 5 étaient majoritairement nucléaires et qu'elles étaient préférentiellement exprimées dans les étapes précoces du développement (Botta et al. 2013; Tran et al. 2011).

Malgré les ressemblances au niveau de l'organisation structurale entre les trois gènes *MBNL*, ces derniers arborent en revanche des profils d'expression très différents.

2.1.2. Profils d'expression

De manière générale, très peu d'études se sont intéressées aux profils d'expression des protéines MBNL chez l'Homme. Quelques-unes ont toutefois montré que les protéines MBNL1 et MBNL2 étaient exprimées de façon ubiquitaire, contrairement à la protéine MBNL3 dont l'expression est beaucoup plus restreinte et très peu détectée dans les tissus adultes (Sznajder and Swanson 2019; Holt et al. 2009) (**Figure 14**). La protéine MBNL1 est le paralogue le plus exprimé des trois dans la majorité des tissus avec une prédominance encore plus importante au niveau des muscles squelettiques par exemple. Le cerveau représente cependant l'exception qui confirme la règle, avec une expression prédominante de MBNL2 (Konieczny et al. 2014).

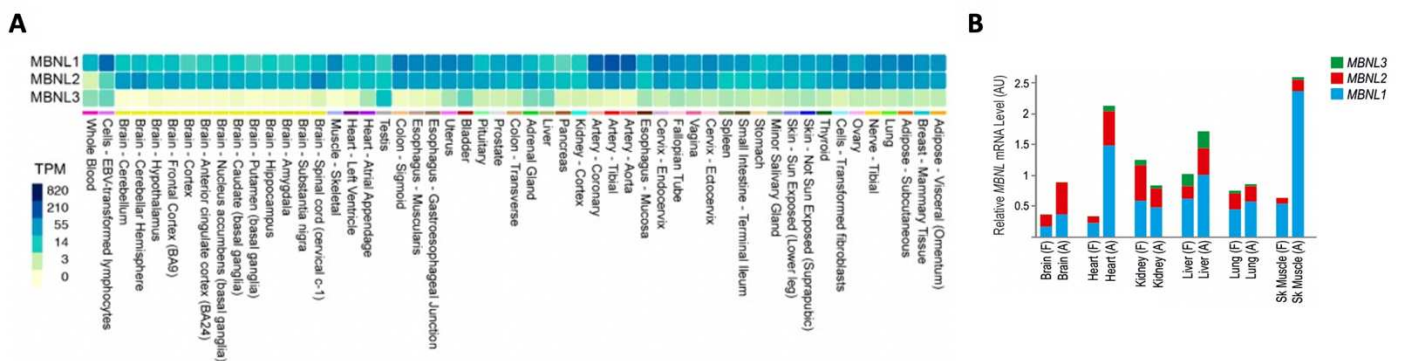


Figure 14 : Profils d'expression des gènes *MBNL* dans différents tissus humains

(A) Expression des trois paralogues dans 53 tissus adultes sains issus de plus de 700 individus. TPM: transcrits par million. **(B)** L'expression des gènes *MBNL* diffère également selon le stade développemental. (F): foetal, (A): adulte, AU: unités arbitraires.

Adapté de (Sznajder and Swanson 2019; Konieczny et al. 2014).

Le profil d'expression des protéines MBNL varie également au cours du développement, en particulier pour MBNL1 et MBNL2 dont l'expression augmente pendant la différenciation des tissus (**Figure 14-B**). Cela est d'autant plus flagrant dans certains tissus comme les muscles squelettiques, le cœur ou encore le cerveau où les niveaux d'expression des ARNm

MBNL1 et *MBNL2* sont nettement supérieurs dans les tissus adultes par rapport aux tissus foetaux (Konieczny et al. 2014).

En plus de l'augmentation de l'expression des protéines *MBNL1* et *MBNL2*, ce sont également différentes isoformes qui vont être exprimées au cours du développement. En effet, il a tout d'abord été montré que l'épissage de l'exon 7 de *Mbnl1* était modulé lors du développement post-natal du muscle squelettique chez la souris, jusqu'à être très peu inclus chez les souris adultes (Lin et al. 2006) (**Figure 15-A**). De la même manière, il a ensuite été observé au niveau du muscle cardiaque des souris que les isoformes ne contenant pas l'exon 5 étaient de plus en plus exprimées au cours du développement embryonnaire (Terenzi and Ladd 2010). Une étude plus récente s'est intéressée à ces différents profils d'épissage au cours du développement humain en comparant les pourcentages d'inclusion des exons 5, 7 et 8 des trois paralogues dans différents tissus foetaux et adultes à partir de 5 patients DM1 et 7 patients sains (Sznajder et al. 2016) (**Figure 15-B**).

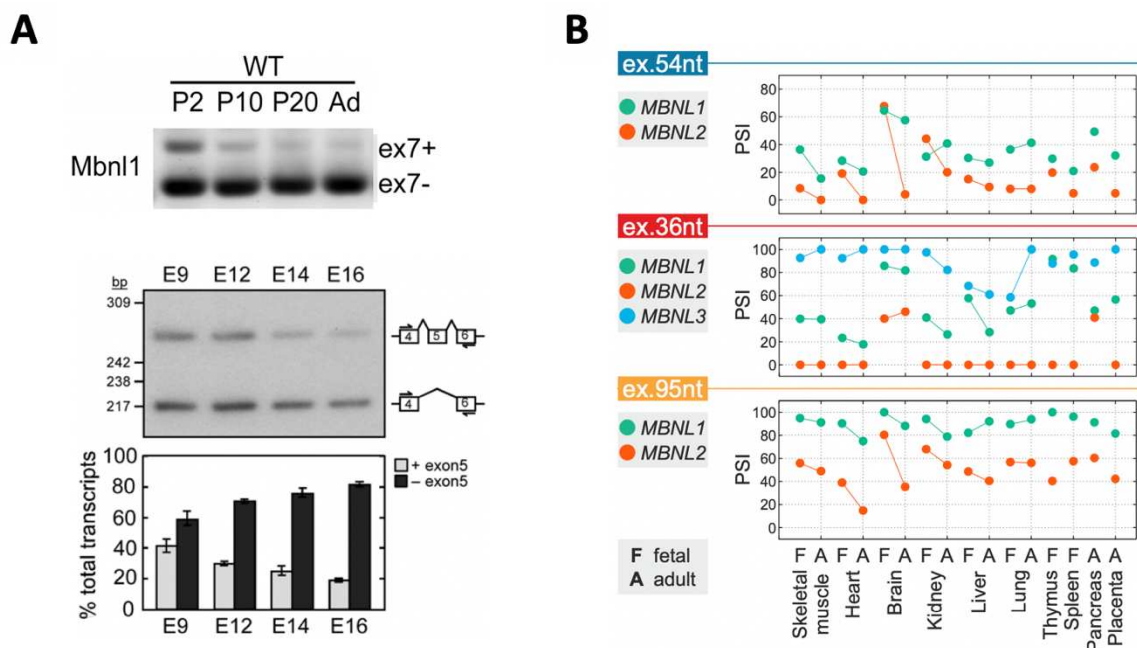


Figure 15 : Profils d'épissage des paralogues *MBNL* au cours du développement

(A) Épissage des exons 7 et 5 des transcrits *Mbnl1* chez la souris au cours du développement post-natal des muscles squelettiques et du développement embryonnaire du cœur respectivement. **(B)** Pourcentage d'inclusion des exons 5 (54nt), 7(36nt) et 8 (95nt) des trois paralogues *MBNL* dans différents tissus humains aux stades foetal (entre 20 et 36 semaines) et adulte.

Adapté de (Sznajder et al. 2016; Terenzi and Ladd 2010; Lin et al. 2006).

Cette étude a permis de confirmer certains résultats observés chez la souris comme par exemple l'importante diminution de l'inclusion de l'exon 5 de *MBNL1* au cours du

développement dans le cœur, avec environ 20% d'inclusion au stade adulte. Ils ont également observé un shift encore plus important de l'inclusion de cet exon pour *MBNL2* au niveau du cerveau. Concernant l'exon 7, la diminution de son inclusion n'a pas été observée dans le muscle squelettique bien que la majorité des transcrits *MBNL1* (60%) ne le contiennent pas. Celui-ci n'est d'ailleurs quasiment jamais inclus au sein des transcrits *MBNL2* dans les différents tissus, à l'exception du cerveau et du pancréas (Sznajder et al. 2016). Il est possible que certaines transitions au niveau des inclusions de ces exons n'aient pas été observées car celles-ci pourraient avoir eu lieu dans des stades encore plus précoces, au cours du développement embryonnaire par exemple.

De manière générale, ces différentes études ont ainsi pu mettre en évidence la transition foetale/adulte des profils d'épissage des transcrits *MBNL* au cours du développement. Il est intéressant de faire le parallèle entre cette régulation développementale de l'expression des protéines *MBNL* et leur rôle important dans la transition développementale de l'épissage alternatif.

2.2. Le rôle des protéines *MBNL* dans le développement

Un nombre croissant d'études sont apparues ces dernières années visant à étudier la contribution des protéines *MBNL* au niveau développemental. Chez la drosophile, il a été montré dès la fin des années 1990 que l'expression du gène *mbl* était requise pour la mise en place du système nerveux périphérique au cours du développement embryonnaire (Kania et al. 1995) et était également essentielle pour la différenciation terminale des bandes Z sarcomériques ainsi que pour la différenciation des cellules photoréceptrices (Artero et al. 1998; Begemann et al. 1997). Il semble ainsi que la protéine muscleblind joue un rôle important pour le développement embryonnaire de plusieurs tissus cellulaires distincts (Nikonova et al. 2019).

2.2.1. L'épissage alternatif

Ces protéines jouent des rôles importants au niveau de la régulation des ARN. En effet, elles sont impliquées dans différents processus comme la stabilité et la localisation des ARNm, la polyadénylation ou encore la biogénèse des miRNA. Les protéines *MBNL* ont été cependant principalement décrites dans la littérature pour leur rôle de régulateurs de l'épissage alternatif des ARN pré-messagers (Sznajder and Swanson 2019). L'élimination des introns des ARN pré-messagers est une étape d'épissage nécessaire pour l'expression de la grande majorité des gènes. En plus de ce processus d'épissage "classique", l'épissage alternatif des ARNm représente un mécanisme important pour la régulation de l'expression

d'isoformes protéiques à un instant t du développement. L'épissage alternatif apparaît être un mécanisme essentiel pour la diversité et la complexité des protéines codées par le génome humain. En effet, seulement 10% du génome humain est constitué d'exons codants, le reste étant majoritairement des introns et des régions UTR non codantes. En revanche, environ 95% des gènes multi-exoniques humains sont sujets à l'épissage alternatif, permettant ainsi la production de différentes isoformes de protéines, modifiant leurs fonctions ou encore leur localisation. Une mauvaise régulation de ce processus est associée à un large spectre de maladies chez l'Homme, en altérant ou en supprimant d'importantes fonctions physiologiques des protéines. Il est estimé qu'environ un tiers des mutations à l'origine de maladies génétiques impliquent des modifications de l'épissage alternatif (Montes et al. 2019). Ainsi, un équilibre précis de l'épissage alternatif est primordial (Kim et al. 2018).

Dans cette partie, nous n'aborderons que les grandes généralités de ce processus extrêmement complexe.

2.2.1.1. Mécanisme général de l'épissage alternatif

L'épissage est orchestré par un complexe macromoléculaire connu sous le nom de splicéosome. Il est composé de cinq petites ribonucléoprotéines (snRNP) (U1, U2, U4/U6 et U5) associées à plus de 150 autres protéines (Ule and Blencowe 2019). Ce complexe va se former via la reconnaissance de plusieurs séquences consensus situées au niveau des extrémités des exons et des introns (**Figure 16-A**). Parmi les différents composants du splicéosome, la snRNP U1 reconnaît une séquence au niveau de l'extrémité 5' de l'intron, à la jonction avec l'exon précédent, représentant la séquence d'épissage 5' (5'SS) aussi appelée site donneur. À l'inverse, le complexe U2AF reconnaît un site poly-pyrimidine intronique juxtaposé à la séquence d'épissage de l'extrémité 3' de l'intron (3'SS), également appelée site accepteur. Un autre site, appelé site de branchement, est quant à lui reconnu par la snRNP U2. La reconnaissance de ces séquences permet ainsi à la machinerie d'épissage de définir les exons et les introns (**Figure 16-B**). L'assemblage de ce complexe peut cependant être ciblé par différents facteurs régulateurs de l'épissage. On distingue de manière générale les facteurs d'épissage activateurs et répresseurs. Par exemple, la famille des protéines SR et d'autres protéines comme RBM5 ont majoritairement un rôle d'activateur de l'épissage contrairement à la famille des hnRNP qui sont le plus souvent des répresseurs de l'épissage (Ule and Blencowe 2019; Daguenet et al. 2015; Poulos et al. 2011).

Concernant la régulation de l'épissage alternatif, différents facteurs d'épissage peuvent interagir avec le splicéosome à travers différentes séquences aussi bien exoniques qu'introniques (**Figure 16-C**). Ceci va alors induire l'inclusion ou l'exclusion des exons

alternatifs et peut également donner lieu à des phénomènes de rétention d'intron. Ces séquences régulatrices sont multiples, certaines favorisant l'épissage (ESE et ISE) et d'autres au contraire l'inhibent (ESS et ISS). Ainsi, en fonction de l'emplacement des sites de liaisons des facteurs d'épissage vis à vis de ces séquences, ces derniers vont favoriser soit l'inclusion soit l'exclusion de l'exon en question. Il existe également d'autres types d'épissage alternatif plus complexes, comme par exemple le cas des exons mutuellement exclusifs ou encore la présence de sites donneurs ou accepteurs alternatifs (Ule and Blencowe 2019; Dagueuet et al. 2015).

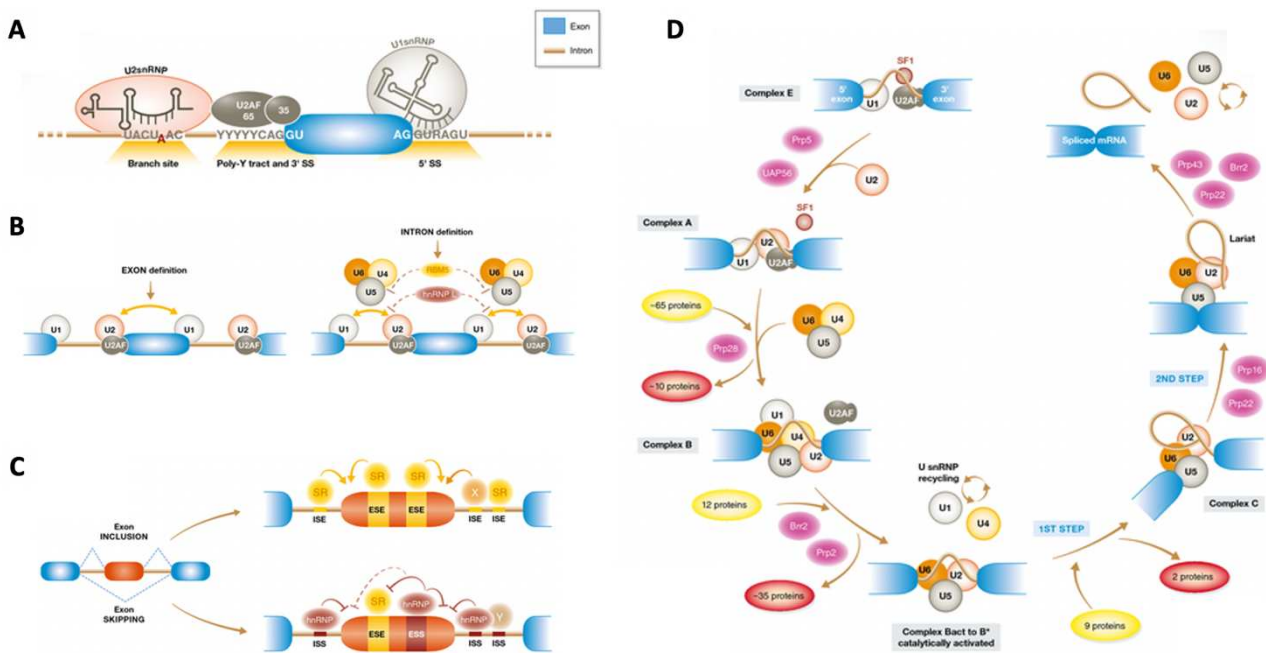


Figure 16 : Mécanisme général de l'épissage alternatif et assemblage du splicéosome
(A) L'initiation de la formation du splicéosome est régulée par des séquences consensus situées aux extrémités des exons et des introns. 5'SS: site donneur, Poly-Y tract (site poly-pyrimidine) et 3'SS: site accepteur, branch site: site de branchement. **(B)** La reconnaissance des sites donneurs et accepteurs par les snRNP U1 et U2 respectivement permettent le recrutement des trois autres snRNP aboutissant à l'épissage des introns. Ce processus peut être régulé par d'autres facteurs d'épissage. **(C)** L'inclusion ou l'exclusion de l'exon est régulée par des séquences exoniques ou introniques pouvant favoriser (ESE et ISE) ou masquer (ESS et ISS) la reconnaissance des sites d'épissage. Ainsi, en fonction de leurs emplacements, les facteurs d'épissage vont favoriser ou inhiber l'épissage de l'exon. **(D)** Représentation schématique de l'assemblage du splicéosome. ESE: Exon splice enhancer, ISE: intron splice enhancer, ESS: Exon splice silencer, ISS: intron splice silencer. *Adapté de (Dagueuet et al. 2015).*

Afin de procéder à l'épissage, plusieurs interactions vont s'établir au niveau des différents sites décrits ci-dessus. Dans un premier temps, la snRNP U1 va reconnaître le site donneur 5'SS alors que le site de branchement est reconnu par la protéine SF1 et le site donneur 3'SS par le complexe U2AF (**Figure 16-D**). Ces interactions vont permettre le recrutement de la

snRNP U2, prenant alors la place de la protéine SF1. Trois autres snRNP vont à leur tour être recrutées (U4/U6 et U5) ainsi qu'une multitude d'autres protéines, formant ainsi le spliceosome (complexe B) (**Figure 16-D**). Après plusieurs remodelages, ce complexe deviendra catalytiquement actif conduisant à la première étape de la réaction d'épissage formant un lasso intermédiaire (complexe C) (**Figure 16-D**). La deuxième étape consiste à l'excision de la région concernée qui sera alors dégradée. Parallèlement, les snRNP seront recyclés pour un autre cycle d'assemblage du spliceosome (Ule and Blencowe 2019; Daguene et al. 2015).

2.2.1.2. L'épissage alternatif au cours du développement

L'épissage alternatif est ainsi très finement régulé en fonction des tissus et du stade développemental. De nombreuses transitions d'isoformes résultant de l'épissage alternatif ont été observées au cours de la différenciation cellulaire et de la maturation des différents tissus (Baralle and Giudice 2017). Une étude princeps publiée en 2011 a même montré le rôle de l'épissage alternatif dans le maintien de la pluripotence des cellules souches embryonnaires. Cette étude a révélé qu'une isoforme spécifique aux cellules souches embryonnaires du facteur de transcription FOXP1 jouait un rôle majeur en stimulant l'expression de gènes nécessaires au maintien de la pluripotence tels que *OCT4* et *NANOG*. Ils ont observé que cette isoforme contenant l'exon 18b n'était plus exprimée dans les cellules différenciées. En effet, FOXP1 subit une transition développementale au niveau de l'épissage alternatif de l'exon 18/18b au cours des processus de différenciation, aboutissant à une isoforme contenant l'exon 18 et non plus l'exon 18b, et modifiant les préférences de liaison à l'ADN du facteur de transcription (Gabut et al. 2011). Depuis, d'autres études ont confirmé le rôle important de l'épissage alternatif dans la régulation de la pluripotence avec par exemple les isoformes d'*OCT4* ou encore *MDB2* (methyl-CpG binding domain protein 2) spécifiques de l'état de pluripotence (Kanitz et al. 2019; Liu et al. 2018).

Des analyses transcriptomiques ont permis de montrer que le cerveau et le muscle squelettique étaient parmi les tissus humains qui présentent le plus grand nombre d'exons alternatifs exprimés de manière différentielle (Pistoni et al. 2010; Wang et al. 2008). Une récente étude a d'ailleurs pu mettre en évidence une régulation temporelle précise de l'épissage alternatif au cours du développement neural chez la souris, en révélant deux transitions majeures, la première à la fin du développement foetal, au moment de la naissance et la seconde au cours des deux premières semaines post-natales. Les isoformes adultes générées par ces événements d'épissage alternatif sont spécifiques du cerveau adulte et essentiels pour l'établissement et le maintien de l'homéostasie des neurones matures (Weyn-Vanhentenryck et al. 2018). Cette étude a aussi montré l'importance de certains facteurs d'épissage spécifiques dans ces transitions développementales d'épissages

alternatifs. Plus précisément, les facteurs d'épissage tels que NOVA, RBFOX, PTBP ou encore les CELF semblent impliqués majoritairement lors des transitions précoces alors que les protéines MBNL ont plutôt été identifiées comme régulateurs des transitions d'épissages alternatifs plus tardifs, au cours du développement post-natal (Weyn-Vanhentenryck et al. 2018).

Ce même phénomène de transition a été observé dans les muscles squelettiques où l'expression de différentes isoformes de protéines varie également selon les muscles ou encore de manière plus précise, en fonction du type de fibres musculaires (Baralle and Giudice 2017). Plusieurs études ont ainsi observé ces transitions de l'épissage alternatif au cours du développement post-natal des muscles squelettiques et du cœur chez la souris, avec ici aussi une prédominance de ces événements dans les deux premières semaines après la naissance (Baralle and Giudice 2017; Brinegar et al. 2017; Kalsotra et al. 2008). Parmi les événements d'épissage ayant lieu au cours du développement des muscles squelettiques, plusieurs sont impliqués dans l'homéostasie calcique tels que l'exon 22 de SERCA1 et l'exon 29 de CACNA1S et une étude a montré que la grande majorité de ces derniers était régulée par la transition de l'activité des protéines antagonistes CELF et MBNL qui se produit pendant le développement post-natal (Brinegar et al. 2017; Lin et al. 2006).

2.2.2. Les protéines MBNL comme facteurs d'épissage

2.2.2.1. Mécanisme d'action des protéines MBNL dans l'épissage

Les protéines Mbnl se lient à leurs cibles via leurs domaines ZnF organisés en tandem et reconnaissant les motifs ARN 5'-YGCY-3' (Teplova and Patel 2008). De nombreuses études structurales se sont intéressées aux interactions entre ces facteurs d'épissage et leurs cibles, suggérant ainsi plusieurs modèles prédictifs (**Figure 17-A**). En effet, les protéines MBNL vont être capables de se lier à différentes conformations de l'ARN simple et double-brin selon plusieurs paramètres comme le nombre de motifs YGCY ou encore le nombre de nucléotides séparant ces motifs. Ces études ont notamment montré que la séquence "linker" reliant les domaines ZnF2 et 3 permettait une très grande flexibilité des deux paires de domaines ZnF, capables de lier deux motifs YGCY uniquement séparés par un seul nucléotide. Celle-ci permet ainsi aux protéines MBNL d'adopter de multiples conformations pour reconnaître une grande variété de cibles ARN (Konieczny et al. 2014; Cass et al. 2011). D'autres études ont également proposé des modèles dans lesquels les protéines MBNL s'intercaleraient au niveau des structures ARN sous une conformation de boucle et même d'épingle à cheveux, avec ses domaines ZnF se fixant entre les séquences complémentaires (Konieczny et al. 2014; Teplova and Patel 2008) (**Figure 17-A**). C'est notamment le cas de figure de l'ARN pré-messager de la Troponine T cardiaque (cTNT) pour lequel il a été montré que la liaison de

MBNL1 modifiait la conformation de sa structure tige-boucle (Park et al. 2017; Fu et al. 2012). Une étude récente a démontré que la structure des ARN cibles ainsi que le nombre et la disposition des motifs YGCY pouvaient moduler l'affinité de liaison des protéines MBNL ainsi que leur activité d'épissage. L'augmentation du nombre de motifs YGCY dans un contexte structural similaire corrélait ainsi avec une plus forte affinité et une activité d'épissage plus importante (Taylor et al. 2018).

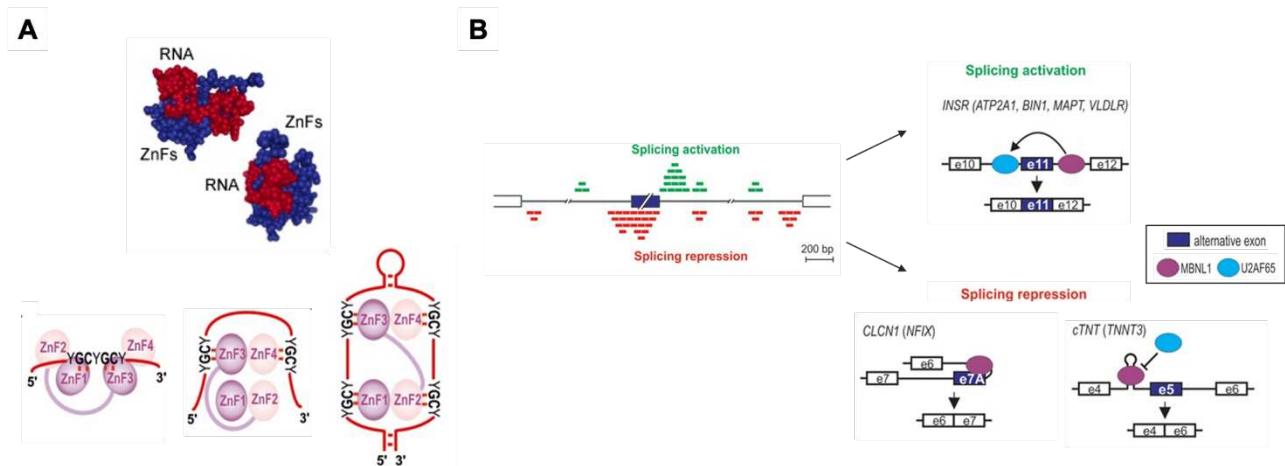


Figure 17 : Liaison des protéines MBNL à l'ARN et épissage alternatif

(A) Modèle structural des domaines ZnF de MBNL1 (en bleu) complexés avec le motif ARN YGCY (en rouge) et différents modes de liaisons proposés entre les tandems ZnF et différentes conformations d'ARN cibles simple- ou double-brin. (B) De manière générale, la liaison des protéines MBNL en amont (rouge) ou en aval (vert) de l'exon alternatif va entraîner préférentiellement l'exclusion ou l'inclusion de cet exon respectivement. *Adapté de* (Sznajder and Swanson 2019; Konieczny et al. 2014).

Les protéines MBNL peuvent réguler l'épissage des exons alternatifs de manière positive ou négative. La régulation positive, ou activation de l'épissage, entraîne l'inclusion de l'exon régulé alors que l'inhibition de l'épissage engendre à l'inverse l'exclusion de l'exon. Le positionnement des sites de liaison des protéines MBNL est également un facteur déterminant vis à vis de l'activité d'épissage de ces protéines favorisant alors l'inclusion ou l'exclusion de l'exon cible (Konieczny et al. 2014; Wang et al. 2012). En effet, il a été montré que la fixation des protéines MBNL1 et MBNL2 au niveau de l'exon alternatif ou au niveau des régions introniques en amont de celui-ci facilite généralement l'exclusion de l'exon, tandis que la fixation dans les régions introniques en aval favorise davantage l'inclusion (Konieczny et al. 2014) (**Figure 17-B**). Ce modèle de régulation suggère que les protéines MBNL induiraient l'inclusion des exons soit i) en entrant en compétition avec des inhibiteurs au niveau de sites introniques de répression de l'épissage (ISS), ou ii) via l'activation des séquences introniques activatrices (ISE), ou encore iii) en stimulant la reconnaissance du site donneur (5'SS) en aval de l'exon par le spliceosome. Cela indique ainsi à l'inverse que

l'inhibition de l'épissage résulte du blocage par les protéines MBNL des séquences activatrices (ESE et ISE), du site accepteur (3'SS) ou encore du site poly-pyrimidine et du site de branchement (cf. Figure 16) (Konieczny et al. 2014). En effet, il a par exemple été démontré dans le cadre de l'épissage alternatif de cTNT que MBNL1 interagissait avec une structure tige-boucle dans l'intron en amont de l'exon 5 et empêcherait le facteur U2AF65 de reconnaître le site poly-pyrimidine, inhibant ainsi le recrutement de la snRNP U2 et favorisant l'exclusion de l'exon 5 (Warf et al. 2009).

2.2.2.2. Rôle développemental des protéines MBNL

Comme nous l'avons vu précédemment, les protéines MBNL présentent des profils d'expression spécifiques que ce soit au cours du développement et en fonction des tissus suggérant leur rôle dans des processus de spécification tissulaire et développementale. Concordant avec le fait que l'expression de ces protéines augmente au cours du développement, une étude publiée en 2013 a montré l'importance des protéines MBNL1 et MBNL2 dans la régulation négative de la pluripotence. Il a tout d'abord été constaté de manière intéressante que parmi plus de 200 facteurs d'épissage analysés, MBNL1 et MBNL2 présentaient les niveaux d'expression les plus faibles dans les cellules humaines pluripotentes indifférenciées par rapport aux autres cellules et tissus différenciés (**Figure 18-A**) (Han et al. 2013).

Par des expériences de perte de fonction à l'aide de siRNA ciblant MBNL1 et MBNL2 dans différents types de cellules différenciées, les auteurs ont pu mettre en évidence de nombreux défauts d'épissage, dont plus d'une cinquantaine correspondent à des profils d'épissage spécifiques des cellules pluripotentes (**Figure 18-B**). Ils ont également démontré que le knock-down de MBNL1 et MBNL2 améliorait significativement l'efficacité de reprogrammation des cellules différenciées en cellules souches humaines induites à la pluripotence. À l'inverse, la surexpression de MBNL1 et MBNL2 dans des cellules indifférenciées entraîne la transition de nombreux profils d'épissage vers des profils d'épissage propres aux cellules différenciées. Enfin, cette étude a montré que l'épissage alternatif des exons 18/18b de FOXP1 faisait partie des événements régulés par les protéines MBNL. Ainsi, l'expression des protéines MBNL augmentant au cours de la différenciation cellulaire, celles-ci vont alors jouer un rôle majeur dans ces processus de différenciation, en donnant naissance à des profils d'épissage propres aux différents tissus et en réprimant les profils d'épissage caractéristiques des cellules indifférenciées, notamment en induisant la transition de l'épissage alternatif de FOXP1 (**Figure 18-C**). Les protéines MBNL sont ainsi considérées comme des régulateurs négatifs de la pluripotence (Han et al. 2013).

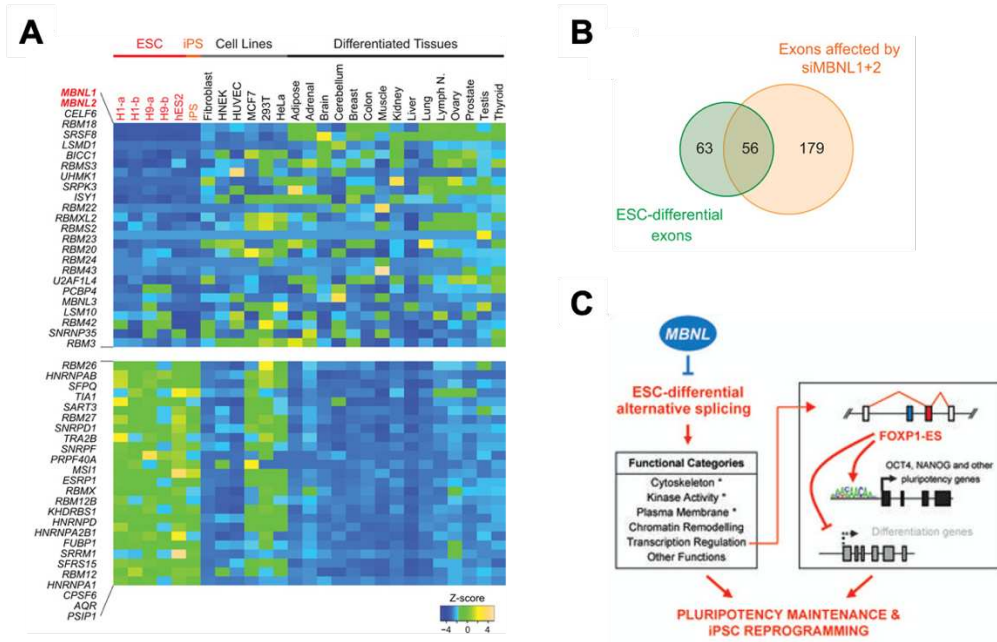


Figure 18 : MBNL1 et MBNL2 inhibent le maintien de la pluripotence

(A) Expression relative des ARNm des facteurs d'épissage dans des cellules souches indifférenciées par rapport à leur expression dans des cellules non pluripotentes et dans des tissus différenciés. **(B)** Diagramme de Venn montrant la proportion des exons alternatifs différemment épissés dans les cellules souches embryonnaires (vert) et retrouvés affectés par le knock-down de MBNL1 et MBNL2 par siRNA dans des cellules HeLa (orange). **(C)** Modèle proposé pour le rôle des protéines MBNL dans la régulation des événements d'épissage alternatif contrôlant le maintien de la pluripotence. *Adapté de (Han et al. 2013).*

Plusieurs études ont également démontré le rôle important des protéines MBNL dans la transition de profils d'épissages foetaux à des profils d'épissage adultes au cours du développement post-natal des tissus. A titre d'exemple, il a notamment été montré dans un modèle de souris, dans lequel le gène *Mbnl1* a été inactivé (*Mbnl1*^{-/-}), qu'en l'absence d'expression de Mbnl1 plusieurs transitions d'épissage alternatif post-natales et essentielles pour le développement des muscles squelettiques n'avaient pas lieu (Lin et al. 2006) (**Figure 19**). On observe chez les souris saines une transition des profils d'épissage au cours des deux premières semaines après la naissance (entre P2 et P20) avec une importante augmentation de l'isoforme de *Serca1* contenant l'exon 22 et à l'inverse une augmentation des isoformes de *Ldb3* (*Zasp*) et de la m-Titine dépourvues de l'exon 11 et de l'exon 5 respectivement. Ces transitions vont permettre à chacune de ces protéines d'acquérir une isoforme adulte, jouant des rôles essentiels dans l'homéostasie calcique (*Serca1*) ou encore dans l'architecture et le maintien des fibres musculaires avec les bandes Z (*Ldb3*) et M (m-Titine) par exemple. L'absence de la protéine Mbnl1 aboutit ainsi à la perturbation de ces profils d'épissage, démontrant l'importance de ce facteur d'épissage pour le bon développement des muscles squelettiques (Lin et al. 2006).

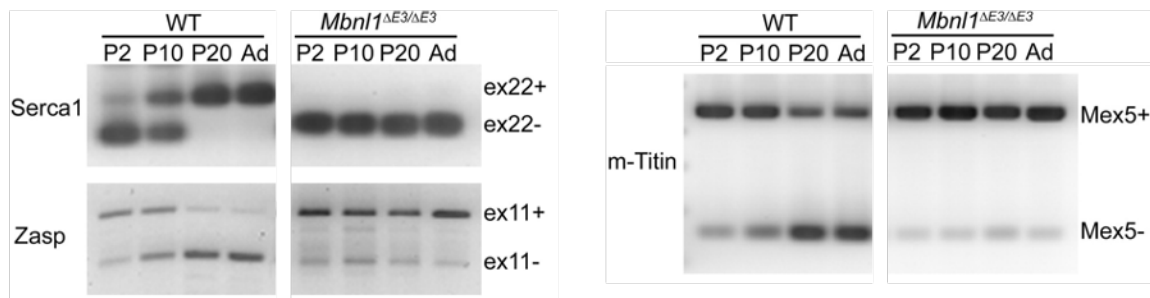


Figure 19 : Mbnl1 est responsable de plusieurs transitions d'épissage post-natales musculaires

Exemple des défauts d'épissage de l'exon 22 de SERCA1, l'exon 11 de LDB3 (ZASP) et l'exon 5 de la Titine dans les muscles squelettiques de souris knock-out pour le gène *Mbnl1* (Lin et al. 2006).

2.2.2.3. Différentes activités d'épissage selon les isoformes MBNL

L'activité d'épissage des protéines MBNL va aussi varier en fonction des tissus. Ainsi, MBNL1 régule l'épissage de nombreuses cibles au niveau des muscles squelettiques et du cœur alors que MBNL2 joue un rôle plus important dans l'épissage alternatif au niveau du système nerveux central. Concernant MBNL3, ce dernier a été très peu décrit dans la littérature, principalement en raison de ses faibles niveaux d'expression. Une étude publiée en 2016 a montré que, malgré le fait que les trois paralogues étaient capables de reconnaître les mêmes motifs ARN YGCY, leur affinité de liaison avec ces motifs et leur activité d'épissage étaient cependant différents. En effet, MBNL1 est globalement considéré comme le paralogue ayant l'activité d'épissage la plus élevée, suivi de près par MBNL2. Une interchangeabilité fonctionnelle entre ces deux paralogues a d'ailleurs été décrite dans plusieurs études (Sznajder et al. 2016; Wang et al. 2012) En revanche, MBNL3 possède une activité d'épissage beaucoup moins élevée. Nous avons vu précédemment que les ARNm *MBNL*, eux-aussi, étaient sujets à l'épissage alternatif, donnant ainsi naissance à plusieurs isoformes ce qui va influencer sur l'activité d'épissage. Il a tout d'abord été montré que les isoformes de MBNL1 dépourvues de l'exon 3, codant la séquence "linker" reliant les domaines ZnF, avaient une activité d'épissage réduite associée à une plus faible affinité pour les ARNm cibles (Tran et al. 2011). La présence de l'exon 5 dans MBNL1 et de MBNL2 semble jouer un rôle important. Dans la mesure où celui-ci code une séquence NLS, les isoformes MBNL1 et MBNL2 contenant l'exon 5 ont une localisation majoritairement nucléaire augmentant ainsi leur concentration dans le noyau des cellules et favorisant ainsi la régulation des épissages alternatifs (Sznajder et al. 2016; Tran et al. 2011). De manière intéressante il a été montré que les facteurs régulant l'épissage alternatif de l'exon 5 de MBNL1 et MBNL2 étaient les protéines MBNL1 et 2 elles-mêmes. Ces observations ont abouti à la proposition d'un modèle de boucle de rétrocontrôle dans lequel (Figure 20) ; lorsque le niveau d'expression des protéines MBNL est faible, celles-ci ne pourraient pas

inhiber l'inclusion de leur propre exon 5 induisant alors la production d'isoformes possédant un signal NLS complet. La présence de ce signal va ainsi permettre aux protéines MBNL1 et 2 d'être majoritairement localisées au niveau du noyau où elles vont pouvoir remplir leur fonction de régulateurs de l'épissage. A l'inverse, lorsque la concentration nucléaire des protéines MBNL est suffisante, celles-ci vont alors entraîner l'exclusion de leur exon 5 grâce à la présence de plusieurs motifs YGCY au niveau de l'extrémité 3' de l'intron en amont (Konieczny et al. 2018; Gates et al. 2011).

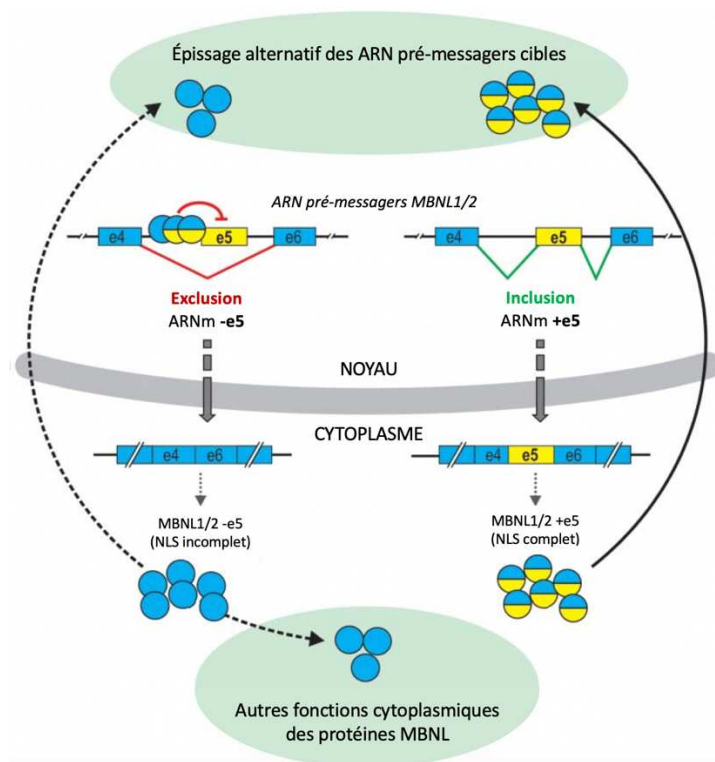


Figure 20 : Boucle de rétrocontrôle de l'épissage de l'exon 5 de MBNL1 et MBNL2

Les exons constitutifs sont représentés en bleu et l'exon 5 alternatif en jaune. *Adapté de (Konieczny et al. 2018).*

Ce modèle de rétrocontrôle est cohérent avec l'augmentation de l'expression des protéines MBNL au cours du développement et qui s'accompagne de la diminution des isoformes contenant l'exon 5 (Sznajder et al. 2016; Terenzi and Ladd 2010). Parmi les isoformes générées ne possédant alors qu'une séquence NLS partielle, seule une partie serait transportée vers le noyau permettant ainsi de maintenir une activité basale d'épissage, notamment pour garantir la génération de profils d'épissage "adultes" des ARNm cibles. L'ensemble de ces données indique également qu'une partie non-négligeable des protéines MBNL est localisée dans le cytoplasme où elles vont avoir des fonctions différentes toutes orientées autour de la biogénèse des ARN.

2.2.3. Autres fonctions des protéines MBNL

Des fonctions autres que celles visant la régulation de l'épissage alternatif ont été identifiées pour les protéines MBNL (**Figure 21**). Dans le noyau, elles vont également être impliquées dans le processus de la polyadénylation alternative (PAA). De manière générale, la partie 3' du dernier exon codant des ARN pré-messagers comporte des sites de polyadénylation qui vont être les cibles potentielles d'un clivage suivi de la synthèse d'une queue poly-A. Ce processus est essentiel pour le transport nucléo-cytoplasmique ainsi que pour la traduction des ARNm. Ainsi, plus il y aura de sites de PAA et plus il y aura de configurations de polyadénylation différentes. Ce mécanisme permet ainsi de générer des régions 3'UTR non traduites alternatives à la fin des ARNm cibles. Ces régions présentent des sites d'interaction avec différents facteurs, en particulier les protéines de liaison à l'ARN et les microARN, qui vont pouvoir moduler la localisation, la traduction et le turnover des ARN (Tian and Manley 2013). À l'aide de fibroblastes issus de souris dans lesquelles les gènes *Mbnl* ont été inactivés, une étude a révélé que l'absence des protéines Mbnl induisait une dérégulation massive des événements de PAA. Ils ont pu démontrer que les protéines Mbnl étaient capables d'interagir directement au niveau des sites de polyadénylation et qu'elles contribuaient ainsi à la transition de nombreux profils de PAA au cours du développement post-natal (Batra et al. 2014).

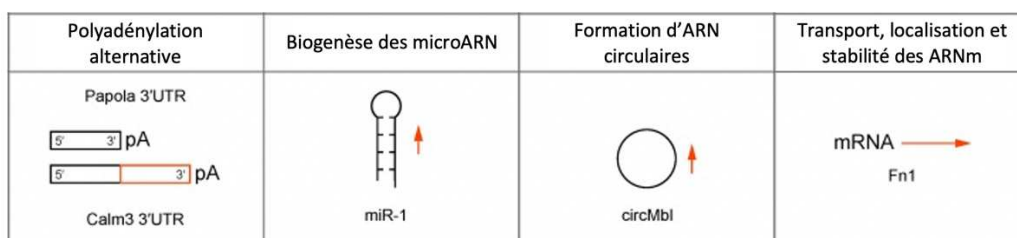


Figure 21 : Autres fonctions exercées par les protéines MBNL

Adapté de (Sznajder and Swanson 2019).

Les protéines MBNL pourraient également intervenir dans la biogenèse des micro-ARN. Il a ainsi été montré que MBNL1 interagissait au niveau du pré-micro-ARN miR-1 et favorisait ainsi sa biogenèse en miR-1 mature via le clivage par l'endonucléase Dicer (Rau et al. 2011). Toujours concernant les ARN non codants, il a été également montré chez la drosophile que la protéine Muscleblind était un facteur impliqué dans la synthèse des ARN circulaires et qu'elle permettait notamment de circulariser son propre ARN pré-messager au niveau de l'exon 2 (Ashwal-Fluss et al. 2014).

Parallèlement, plusieurs études se sont également intéressées aux fonctions cytoplasmiques des protéines MBNL en révélant un rôle important de celles-ci dans la régulation de la stabilité et de la localisation des ARNm (Wang et al. 2015; Masuda et al. 2012; Wang et al. 2012) (**Figure 22**). Eric Wang et ses collègues ont publié en 2012 une étude transcriptomique sur différents tissus murins, en particulier les muscles striés et le cerveau, dans laquelle ils ont montré que les protéines Mbnl se liaient aux régions 3'UTR de nombreux ARNm cibles codant principalement des protéines sécrétées et des protéines constituant des membranes, de la matrice extracellulaire ou encore des protéines synaptiques telles que l'Intégrine $\beta 1$ ou la protéine SNAP25. Ils ont observé que les protéines Mbnl participaient à leur localisation au niveau des membranes pour favoriser leur traduction (**Figure 22-A**). En effet, la déplétion des protéines Mbnl1 et Mbnl2 dans les cellules musculaires de souris entraîne la relocalisation de plusieurs ARNm comme la Fibronectine 1 (Fn1), loin des compartiments membranaires (Wang et al. 2012). Il avait aussi été démontré que Mbnl2 interagissait avec la partie 3'UTR de l'ARNm codant l'Intégrine $\alpha 3$ pour la transporter jusqu'à la membrane plasmique, lieu de sa traduction (Adereth et al. 2005).

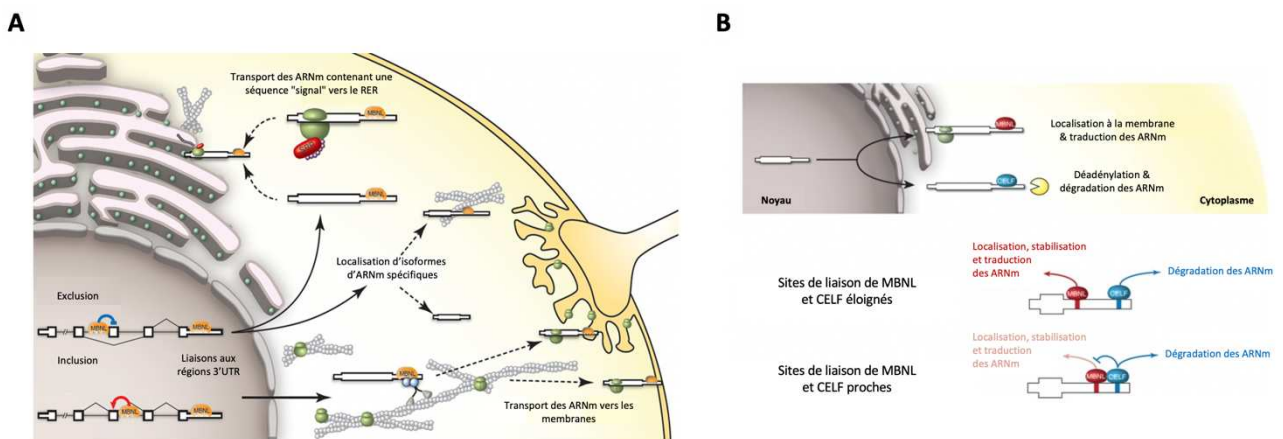


Figure 22 : Fonctions cytoplasmiques des protéines MBNL et antagonisme avec les protéines CELF

(A) Les protéines MBNL régulent le transport et la localisation d'ARNm cibles et favorisent ainsi leur traduction. RER: Réticulum Endoplasmique Rugueux. (B) Modèle de l'antagonisme fonctionnel entre les protéines MBNL et CELF au niveau du devenir des ARNm. *Adapté de* (Wang et al. 2015, 2012).

Contrairement aux protéines MBNL, l'implication des protéines CELF au niveau de la régulation de la stabilité des ARNm a été largement documentée. En effet, ces dernières jouent des rôles importants dans les processus de déadénylation et de dégradation des ARNm (Barreau et al. 2006). D'après les fonctions antagonistes entre les protéines MBNL et CELF observées dans la régulation de l'épissage alternatif au cours du développement, un modèle similaire a été proposé suggérant que les protéines MBNL réguleraient également

la stabilité des ARNm de façon compétitive avec les protéines CELF. Selon ce modèle, la proximité entre les sites de liaisons des protéines MBNL et CELF dans la région 3'UTR des ARNm cibles déterminerait le devenir de ces ARNm qui seraient majoritairement dégradés dans le cas où les sites de liaisons sont proches les uns des autres (Wang et al. 2015) (**Figure 22-B**).

2.3. Implication des protéines MBNL dans la DM1

Compte tenu des différents rôles des protéines MBNL, leur dérégulation chez les patients atteints de la DM1 va être à l'origine de multiples défauts d'ordre moléculaire contribuant à la mise en place de la pathologie.

2.3.1. Séquestration et rôle dans la dynamique de formation des *foci*

Comme nous l'avons vu dans le chapitre précédent, l'implication des protéines MBNL dans la DM1 a tout d'abord été mise en évidence par leur capacité à se lier spécifiquement aux transcrits DMPK mutants, de façon proportionnelle à la longueur des répétitions CUG sous forme double-brin caractéristiques de la pathologie ((Miller et al. 2000). Les protéines MBNL reconnaissant les motifs YGCY, les répétitions CUG constituent alors une multitude de sites de liaison pour les paralogues MBNL (**Figure 23-B**), ce qui a pour conséquence leur séquestration dans les *foci* intranucléaires formés par les expansions de répétitions CUG sous forme d'épingle à cheveux (Fardaei et al. 2002, 2001) (**Figure 23-A**). De manière intéressante, plusieurs études ont révélé que les interactions entre les protéines MBNL et les répétitions CUG étaient dynamiques avec notamment une agrégation stochastique des protéines MBNL au sein des *foci*. À l'aide de techniques d'imagerie sur cellules vivantes, il a été observé que les interactions directes des protéines MBNL avec les répétitions suivent une succession de phénomènes d'agrégaions/fragmentations de manière totalement aléatoire (Sznajder et al. 2016; Querido et al. 2011). À l'inverse, les *foci* peuvent aussi rester stables pendant plusieurs heures, en particulier dans le cas de figure où ils seraient saturés en protéines MBNL. En effet, il a été démontré qu'un faible niveau d'expression des ARN mutants ou des répétitions CUG de petites tailles étaient facilement saturés par les protéines MBNL, n'empêchant cependant pas une autre dynamique via un échange continu avec des protéines MBNL non liées au niveau du nucléoplasme (**Figure 23-C**). L'analyse *in silico* de la dynamique des interactions entre les protéines MBNL et les *foci* a révélé que les paralogues ne possédaient pas la même mobilité vis à vis des *foci*, MBNL1 étant de manière générale le paralogue le plus mobile et MBNL3 à l'inverse le moins mobile. Concernant MBNL2, il a été observé que l'inclusion de l'exon 8 diminuait significativement sa mobilité vis à vis des *foci*. Cependant aucune différence n'a été observée entre les isoformes de MBNL1 contenant ou non l'exon 8 (Sznajder et al. 2016).

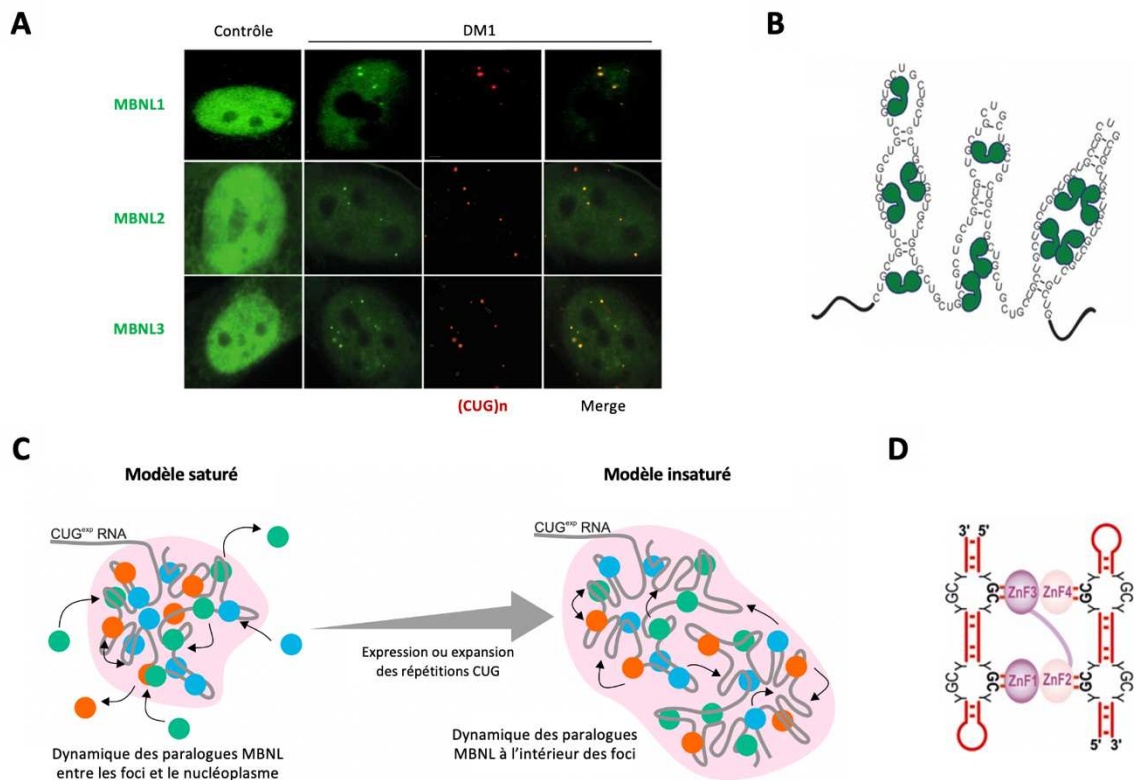


Figure 23 : Interactions entre les protéines MBNL et les répétitions CUG toxiques

(A) Co-localisation des trois paralogues MBNL (vert) avec les répétitions CUG (rouge) dans des fibroblastes DM1. (B) Les protéines MBNL reconnaissent une multitude de motifs YGCY définis par les répétitions CUG. (C) La dynamique des foci est définie selon deux modèles différant par le nombre de sites de liaison des protéines MBNL avec les répétitions CUG. Dans le modèle saturé (gauche) les protéines MBNL interagissent en compétition avec les répétitions CUG, créant une dynamique entre foci et nucléoplasme. A l'inverse, le modèle insaturé (droite) permet une séquestration plus efficace des protéines MBNL dont la dynamique s'effectue désormais au sein même des foci et contribuerait ainsi à leur instabilité. (D) Modèle structural dans lequel une protéine MBNL relierait deux épingles à cheveux suggérant un rôle des protéines MBNL dans la formation des foci. *Adapté de (Sznajder et al. 2016; deLorimier et al. 2014; Konieczny et al. 2014; Fardaei et al. 2002, 2001).*

Un autre modèle opposé a été décrit dans lequel une plus forte expression des transcrits toxiques ou une expansion plus importante du nombre de répétitions CUG n'entraîneraient pas une saturation des *foci* mais un épuisement drastique des paralogues MBNL, avec cette fois-ci des interactions dynamiques au sein même des *foci* (Figure 23-C) (Sznajder et al. 2016). Ainsi, l'assemblage des *foci* est probablement modulé non seulement par la longueur des répétitions mais aussi par le profil d'expression spatio-temporel de DMPK (Sznajder and Swanson 2019). Il a ainsi été observé que la taille des *foci* augmentait avec des concentrations plus élevées de protéines MBNL1 et qu'à l'inverse, la down-expression par shRNA des protéines MBNL entraînait une diminution de leur taille et de leur nombre,

suggérant ainsi un rôle dans l'assemblage ou la stabilisation des *foci* (Querido et al. 2011; Dansithong et al. 2005).

Par ailleurs, les protéines MBNL1 n'interagissent pas exactement de la même manière en fonction des différentes isoformes de la protéine. En effet, bien que toutes les isoformes de MBNL1 soient séquestrées par les *foci*, il est intéressant de noter que celles contenant l'exon 3 (séquence "linker") se lient aux répétitions CUG avec une plus grande affinité (Tran et al. 2011). L'exon 7 de MBNL1 pourrait lui-aussi jouer un rôle important dans la formation des *foci* puisque des dimérisations entre les isoformes contenant l'exon 7 ont été observées et favoriseraient la stabilisation des *foci* (Tran et al. 2011; Yuan et al. 2007). Enfin, un des modèles propose une liaison des protéines MBNL également à l'extérieur des structures en épingle à cheveux permettant ainsi aux paralogues MBNL de relier plusieurs épingles à cheveux entre elles, ce qui conduirait à un déploiement local des *foci* intranucléaires (**Figure 23-D**) (Sznajder and Swanson 2019; Konieczny et al. 2014). D'autres facteurs pourraient également jouer un rôle au niveau de la formation des *foci*. Par exemple, deux membres de la famille des DEAD-box, DDX5 et DDX6, ont été identifiés pour soit exacerber soit diminuer la toxicité sous-jacente des complexes *foci*-MBNL respectivement. En effet, DDX5 favoriserait le recrutement des protéines MBNL au niveau des *foci* alors que DDX6 entrerait en compétition avec les protéines MBNL pour se lier aux répétitions CUG et permettrait de déstabiliser les *foci* (Pettersson et al. 2015; Laurent et al. 2012).

2.3.2. Les modèles d'études des protéines MBNL

De part la séquestration des protéines MBNL et leur potentiel rôle dans la formation des *foci*, la majorité des études visant à mieux comprendre la contribution des protéines MBNL dans le processus pathologique s'est intéressée aux conséquences moléculaires et fonctionnelles de la perte de fonction de cette famille de protéines aboutissant à la génération de modèles murins knock-out soit par la délétion des gènes *Mbnl* seuls, soit par l'inactivation combinée de plusieurs d'entre eux voire même des trois gènes (**Tableau 3**). Les équipes du Pr. Swanson et du Dr. Thornton ont très largement contribué à la génération de ces modèles murins.

Tableau 3 : Les modèles murins pour l'étude de la contribution des protéines MBNL dans la DM1

Modèle murin	Mutation / construction	Description	Limitations	Référence(s)
<i>Mbn1</i> KO	Délétion de l'exon 3	Myotonie, cataracte, centralisation de noyaux des fibres musculaires, baisse de motivation et apathie. Nombreux défauts d'épissage dans les muscles et le cerveau. Surexpression de <i>Mbn2</i> dans le noyau des cellules.	Pas d'atrophie musculaire.	Kanadia (2003), Matynia (2010) et Suenaga (2012)
<i>Mbn2</i> KO	Insertion d'une mutation (gene trap) dans l'intron 2	Myotonie discrète et défauts histologiques musculaires (noyaux centraux, tailles des fibres hétérogènes) au bout de 6 mois.	Pas de défaut d'épissage musculaire. Pas d'étude des atteintes neuronales.	Hao (2008)
	Délétion de l'exon 2	Reproduction de nombreux défauts neuronaux (troubles du sommeil, somnolence diurne, défauts d'apprentissage et de mémorisation). Plusieurs défauts d'épissage, principalement dans le cerveau.	Pas d'atteintes musculaires.	Charizanis (2012)
<i>Mbn3</i> KO	Délétion de l'exon 2	Phénotypes liés à l'âge, atteinte légère et tardive des muscles squelettiques, défauts de régénération musculaire, intolérance au glucose, troubles cardiaques et cataracte.	Expression d'une isoforme courte de la protéine <i>Mbn3</i> . Peu de défauts moléculaires décrits.	Poulos (2013) et Choi (2016)
<i>Mbn1</i> , <i>Mbn2</i> DKO	Délétion de l'exon 3 de <i>Mbn1</i> et délétion constitutive ou conditionnelle (Myo-Cre ou Nestin-Cre) de l'exon 2 de <i>Mbn2</i>	Souris <i>Mbn1</i> ^{-/-} ; <i>Mbn2</i> ^{+/-} : durée de vie réduite (23 semaines maximum) et poids plus faible. Myotonie sévère, nombreux noyaux centraux, fibres musculaires fractionnées et hétérogènes, faiblesse et atrophie musculaires progressives, défauts de conduction cardiaque, défauts dans les jonctions neuromusculaires. Nombreux défauts d'épissage exacerbés dans les muscles striés et le cerveau. S'apparente à un phénotype congénital de la DM1.	La double délétion constitutive entraîne une létalité néonatale importante.	Lee (2013), et Goodwin (2015)
<i>Mbn1</i> , <i>Mbn3</i> DKO	Délétion de l'exon 3 de <i>Mbn1</i> et de l'exon 2 de <i>Mbn3</i>	Myotonie accrue, noyaux centraux et fibres musculaires atrophiées.	Peu de différences significatives avec les souris <i>Mbn1</i> KO.	Choi (2015)
<i>Mbn1</i> , <i>Mbn2</i> , <i>Mbn3</i> TKO	Délétion de l'exon 3 de <i>Mbn1</i> et délétions conditionnelles (Myo-Cre) de <i>Mbn2</i> et <i>Mbn3</i>	Très semblable aux souris <i>Mbn1,2</i> DKO. Détresse respiratoire, cyanose, faiblesse musculaire, cyphose, désorganisation des clusters de récepteurs à l'acétylcholine dans le diaphragme. Nombreux défauts d'épissage et d'expression génique.	Importante mortalité néonatale.	Thomas (2017)

2.3.2.1. Génération des lignées simple KO et spécialisation fonctionnelle

La première démonstration de l'implication des protéines MBNL dans la physiopathologie de la DM1 fut apportée par la génération des souris *Mbn1* KO chez lesquelles plusieurs atteintes principalement musculaires ont été observées (Kanadia et al. 2003). Ces souris présentent une myotonie expliquée en partie par un défaut d'épissage de l'exon 7 du *CLCN1* conduisant une importante perte d'expression de la protéine. Au niveau histologique, certaines fibres musculaires de ces souris sont fractionnées et possèdent des noyaux centralisés, un stigmate histologique fréquemment observé dans des biopsies de patients DM1 (André et al. 2018). D'autre part, ces souris développent également une cataracte sous-capsulaire typique de la pathologie (**Figure 24-A**). En revanche, ces souris ne présentent

aucune atrophie ni faiblesse musculaire et quasiment aucun défaut de conduction cardiaque bien que plusieurs défauts d'épissage aient été observés dans les muscles squelettiques et le cœur (Kanadia et al. 2003). Concernant le système nerveux central, ces souris ont montré des effets très modérés au niveau de la régulation de l'épissage alternatif (Suenaga et al. 2012) et ne présentent que très peu de déficits cognitifs. Seules une certaine apathie et une baisse de motivation ont été observées (Matynia et al. 2010). Afin de confirmer le rôle de *Mbnl1* dans les principales atteintes musculaires décrites, plusieurs expériences de surexpression ont été réalisées dans différents modèles murins de la DM1 et ont montré que la surexpression de *Mbnl1* était suffisante pour corriger la myotonie, la myopathie et plusieurs défauts d'épissage associés à l'expression des répétitions CUG (Chamberlain and Ranum 2012; Kanadia et al. 2006). Cependant, l'absence de nombreuses caractéristiques cliniques de la DM1 chez les souris *Mbnl1* KO indique que la perte de fonction de *Mbnl1* ne peut pas être responsable à elle-seule de l'étendue des atteintes multi-systémiques de la maladie. Les souris *Mbnl1* KO ne présentant pas de *foci* intranucléaires, les autres membres de la famille des protéines MBNL ne sont alors pas séquestrés et pourraient être d'autres intermédiaires clés de la DM1 (Braz et al. 2018).

L'inactivation du gène *Mbnl2* entraîne, contrairement à celle de *Mbnl1*, une pathologie musculaire très légère mais des atteintes cognitives beaucoup plus marquées. Un premier modèle de souris *Mbnl2* KO, développé en 2008 par insertion d'une mutation dans l'intron 2 a été le seul à montrer un défaut d'expression du CLCN1 associé à une myotonie légère au bout de 6 mois (Hao et al. 2008). Cependant le modèle murin *Mbnl2* KO le plus étudié encore aujourd'hui est celui généré par l'équipe du Pr. Swanson et qui repose sur une délétion de l'exon 2. Curieusement et sans explication véritable, il a tout d'abord été observé que les souris *Mbnl2* KO étaient plus petites pendant les 3 semaines qui suivent leur naissance mais qu'elles rattrapaient leur retard de croissance au bout d'un mois (Charizanis et al. 2012) (**Figure 24-B**). Bien que les souris DMSXL soient également plus petites que des souris saines, aucune étude n'a démontré la cause de ce phénotype commun aux deux modèles murins. Ce modèle de souris *Mbnl2* KO ne présente quasiment pas de symptômes musculaires en dehors de quelques défauts histologiques des muscles squelettiques au bout de 4 mois environ. En revanche, les souris *Mbnl2* KO développent plusieurs déficits cognitifs caractéristiques de la DM1 tels que des troubles du sommeil avec une augmentation des épisodes de sommeil paradoxal caractérisés par des mouvements oculaires rapides (MOR ou REM en anglais). Ces souris présentent également des déficits d'apprentissage et de mémoire spatiale causés notamment par une diminution de la transmission synaptique médiée par les récepteurs glutamatergiques NMDA au niveau de l'hippocampe (**Figure 24-B**). Concernant les défauts d'épissage, les auteurs n'en ont décrit aucun au niveau musculaire contrairement aux souris *Mbnl1* KO mais ont observé une importante splicéopathie au niveau cérébrale (**Figure 24-C**).

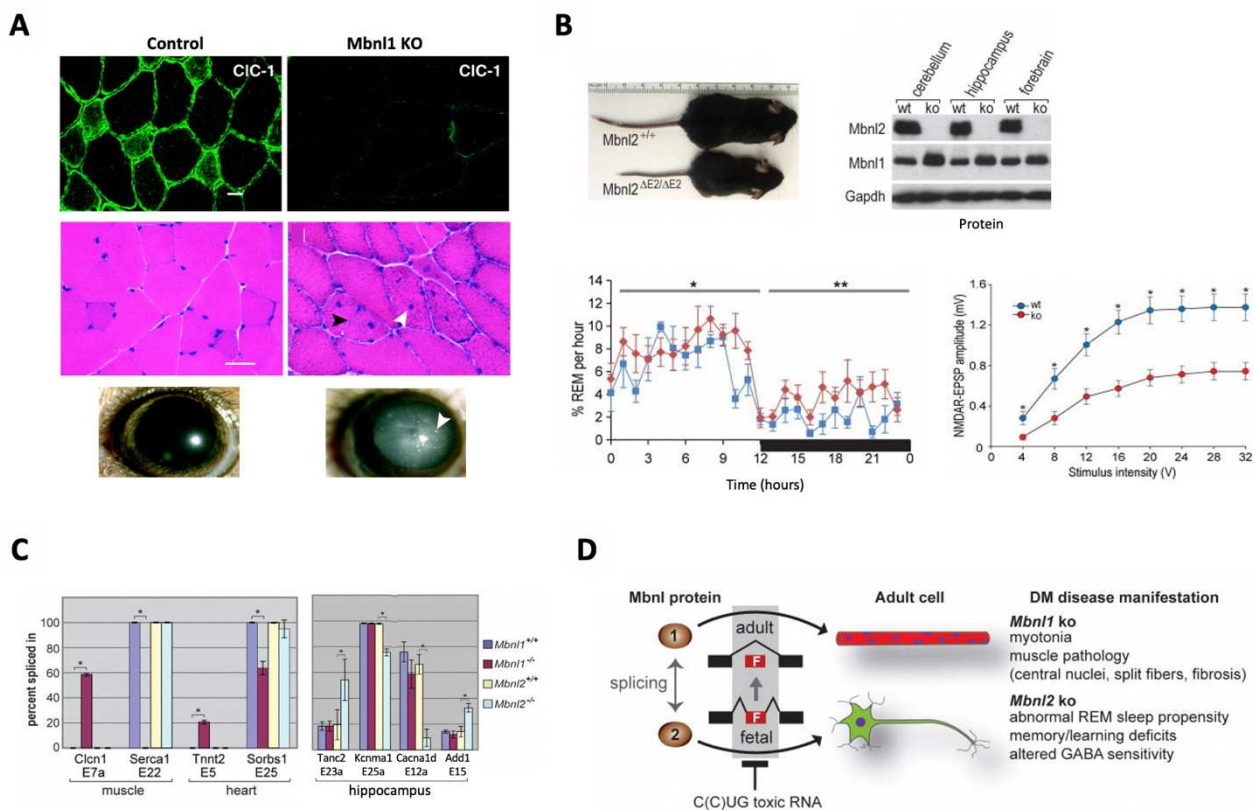


Figure 24 : Contribution tissu-spécifique des protéines Mbn1 et Mbnl2

(A) Parmi les phénotypes DM1 observés dans les souris *Mbnl1* KO on retrouve le défaut d'expression du CLCN1 (vert) associé à la myotonie, une myopathie et une cataracte.

(B) Les souris *Mbnl2* KO sont plus petites à la naissance et leur cerveau présente une surexpression des protéines Mbnl1. Elles développent des déficits cognitifs comme de trouble du sommeil et des défauts d'apprentissage et de mémorisation. **(C)** Les défauts d'épissage musculaires sont liés à la perte d'expression de Mbnl1 alors que les défauts d'épissage neuronaux sont liés à celle de Mbnl2.

(D) Modèle proposé pour la spécialisation tissulaire de Mbnl1 et Mbnl2. *Adapté de (Charizanis et al. 2012; Kanadia et al. 2003).*

Ensemble, ces résultats montrent une contribution tissu-spécifique des deux paralogues Mbnl1 et Mbnl2 suivant un modèle selon lequel la perte de fonction de Mbnl1 est à l'origine des atteintes musculaires alors que celle de Mbnl2 engendre les atteintes neuronales caractéristiques de la DM1 (Charizanis et al. 2012; Kanadia et al. 2003) **(Figure 24-D)**.

Concernant le troisième paragone, les souris *Mbnl3* KO ont montré des résultats intrigants car malgré la faible expression de celui-ci dans les tissus musculaires adultes, ces souris développent quelques atteintes légères liées à l'âge. En effet, un déclin de la régénération des muscles a par exemple été observé chez des souris âgées de plus de 32 semaines (Poulos et al. 2013). D'autres phénotypes liés à l'âge ont ensuite été décrits tels qu'une intolérance au glucose ou encore certains troubles cardiaques (Choi et al. 2016). Ainsi la perte d'expression de Mbnl3 pourrait contribuer à l'apparition de phénotypes liés à un potentiel vieillissement accéléré chez les patients DM1 (Mateos-Aierdi et al. 2015).

Cependant la contribution de MBNL3 dans la physiopathologie humaine de la DM1 reste plus incertaine car contrairement aux cellules murines, l'expression de la protéine MBNL3 n'est pas détectée dans les myoblastes humains (Holt et al. 2009). Par ailleurs, une étude a généré des souris *Mbn1,3* double KO (DKO) en réalisant une perte de fonction combinée de *Mbn1* et *Mbn3*. Malgré de faibles différences significatives, les auteurs ont tout de même observé une myotonie et une myopathie légèrement plus accrues chez ces souris par rapport aux souris *Mbn1* KO (Choi et al. 2015).

2.3.2.2. Déplétions combinées des protéines MBNL et compensation

De façon intéressante, l'étude de ces modèles de perte de fonction des protéines *Mbn1* a permis de mettre en avant l'existence d'un phénomène « compensatoire » entre les différents paralogues de cette famille. En effet, une surexpression de *Mbn1* est détectée dans le cerveau des souris *Mbn2* KO et une surexpression encore plus importante de *Mbn2* a été décrite dans les muscles squelettiques des souris *Mbn1* KO (Lee et al. 2013; Charizanis et al. 2012) (**Figure 24-B** et **Figure 25-A**). Ces phénomènes de compensation entre *Mbn1* et *Mbn2* pourraient expliquer pourquoi certains phénotypes musculaires DM1 comme l'atrophie des muscles squelettiques ne sont pas observés chez les souris *Mbn1* KO (Lee et al. 2013). Contrairement aux souris *Mbn1,3*DKO, les souris *Mbn1,2*DKO ont développé des atteintes beaucoup plus sévères. En effet, la double perte de fonction de *Mbn1* et *Mbn2* entraîne une létalité néonatale très importante, accompagnée d'une réduction du poids corporel des souris à la naissance. Les souris *Mbn1,2*DKO naissent cyanosées et souffrent de détresse respiratoire. Cette létalité néonatale a conduit à s'intéresser plutôt aux souris *Mbn1* KO et *Mbn2* hétérozygote (+/-) qui étaient viables mais dont la survie était tout de même significativement diminuée, avec une espérance de vie d'environ 3 semaines (Thomas et al. 2017; Lee et al. 2013) (**Figure 25-B**). Afin de pouvoir étudier l'effet de la double délétion au niveau musculaire ou neuronal, des pertes de fonction conditionnelles et tissu-spécifiques de *Mbn2* sur un fond génétique *Mbn1* KO ont été conduites à l'aide du système de la Cre-recombinase (Goodwin et al. 2015; Lee et al. 2013). Ainsi, il a été montré dans un contexte *Mbn1* KO que plus le niveau d'expression de *Mbn2* était faible et plus la sévérité de la myopathie des souris était importante. De même, l'inactivation d'une seule copie de *Mbn2* (souris *Mbn1*^{-/-}; *Mbn2*^{+/-}) est suffisante pour exacerber la myotonie (**Figure 25-C**). Ces souris ont même développé plusieurs phénotypes musculaires qui étaient absents de chez les souris *Mbn1* KO comme une atrophie et une faiblesse musculaire, des défauts de conduction cardiaque ou encore une maturation défectueuse des jonctions neuromusculaires (JNM), notamment très impactée au niveau du diaphragme (**Figure 25-D**).

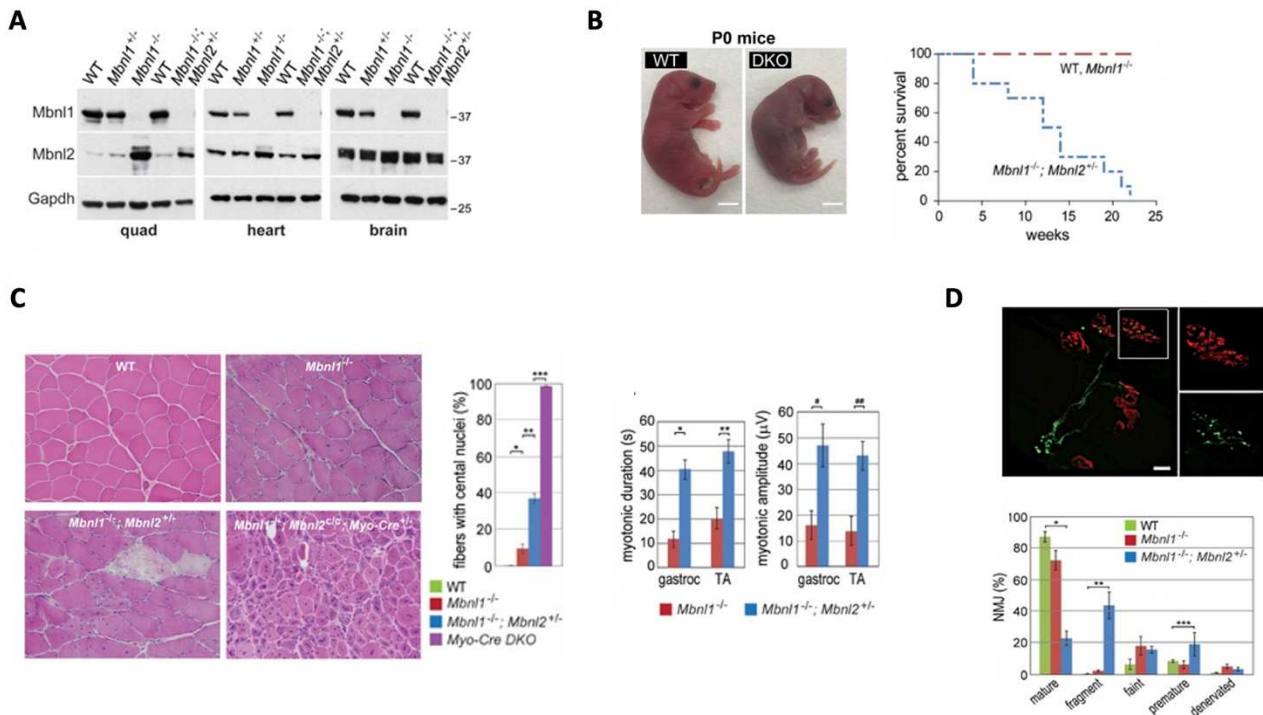


Figure 25 : Compensation fonctionnelle entre Mbnl1 et Mbnl2 et modélisation des formes sévères de la DM1

(A) La perte de fonction de Mbnl1 entraîne la surexpression de Mbnl2 dans les muscles squelettiques et plus légèrement dans le cœur. (B) La double déplétion entraîne une importante létalité néonatale avec des souris cyanosées et en détresse respiratoire. Les souris *Mbnl1*^(-/-); *Mbnl2*^(+/-) sont viables mais leur survie est significativement diminuée. (C) La double déplétion exacerbe la myopathie et la myotonie des souris de manière importante et engendre de nombreux défauts concernant la morphologie des jonctions neuromusculaires, en particulier au niveau du diaphragme (D). *Adapté de (Thomas et al. 2017; Lee et al. 2013).*

Les défauts au niveau des JNM ont également été observés dans le modèle de triple déplétion des protéines Mbnl (TKO) où il a été décrit une perturbation de l'organisation et de la morphologie des récepteurs à l'acétylcholine situés dans le compartiment post-synaptique, au niveau des cellules musculaires. De manière intéressante, une étude avait déjà suggéré par le passé des défauts au niveau des JNM avec la présence de *foci* dans ces structures (Wheeler et al. 2007a). Ensemble, ces observations au niveau du muscle responsable de la respiration pourraient ainsi expliquer les détresses respiratoires à l'origine de la forte mortalité néonatale des souris DKO et TKO (Thomas et al. 2017). De plus, l'insuffisance respiratoire représente une manifestation clinique majeure des formes congénitales de la DM1, suggérant ainsi l'utilité des souris Mbnl1,2 DKO et TKO pour modéliser les formes sévères de la pathologie (Thomas et al. 2017). Par ailleurs, une autre étude focalisée sur les défauts touchant le système nerveux central a montré que l'inactivation conditionnelle de Mbnl2 au niveau des tissus neuronaux dans un contexte *Mbnl1* KO réduisait également la survie des souris, de façon comparable aux souris *Mbnl1*^(-/-); *Mbnl2*^(+/-). Ces souris présentent des déficits moteurs importants et une

splicéopathie neuronale exacerbée comparativement aux souris *Mbnl2* KO (Goodwin et al. 2015). Plus récemment, des défauts dans la morphologie des épines dendritiques ainsi que dans la distribution des neurones corticaux ont également été rapportés chez les souris DKO (Lee et al. 2019).

L'ensemble de ces modèles murins a révélé une certaine spécialisation fonctionnelle entre les trois paralogues, en particulier pour MBNL1 et MBNL2, en mettant également en évidence des phénomènes de compensation fonctionnelle. Dans l'ensemble, ces études ont démontré que la perte de fonction des protéines MBNL était essentielle pour la mise en place de nombreuses atteintes cliniques observées chez les patients DM1.

2.3.3. Les défauts d'épissage associés aux protéines MBNL

Parmi les nombreux défauts phénotypiques et moléculaires reproduits dans les modèles murins de perte de fonction des protéines MBNL, les défauts d'épissage ont fait l'objet de nombreuses études. Celles-ci ont ainsi permis de mettre en évidence les défauts d'épissages de la pathologie directement causés par la séquestration des protéines MBNL dans les *foci* intranucléaires (**Tableau 4**). En analysant les défauts d'épissage observés dans un modèle murin musculaire de la DM1 (souris HSA^{LR}), il a été montré que plus de 80% de ceux-ci étaient retrouvés dérégulés dans les muscles des souris *Mbnl1* KO (Du et al. 2010). De plus, une étude a retrouvé, chez ces mêmes souris *Mbnl1* KO, une trentaine de défauts d'épissage observés chez les patients DM1 (Nakamori et al. 2013). En effet, *Mbnl1* régule de nombreux épissages alternatifs impliqués dans différentes atteintes musculaires caractéristiques de la DM1 comme par exemple l'exon 7a du *Clcn1* ou encore l'exon 29 de *Cacna1s* et l'exon 11 de *Bin1* associés respectivement à la myotonie ou à la myopathie générale des patients DM1. Ces défauts d'épissage observés dans les modèles murins DM1 et dans les souris *Mbnl1* KO sont directement liés au rôle de MBNL1 dans le contrôle des transitions d'épissage foetal-adulte au cours du développement (López-Martínez et al. 2020). L'identification de ces événements d'épissage alternatif régulés par les protéines MBNL nous aide non seulement à mieux comprendre l'implication de ces protéines dans la physiopathologie de la DM1 mais également à prédire le degré de sévérité de la maladie. En effet, il a été démontré que la sévérité de plusieurs défauts d'épissage dépendait directement de la concentration des protéines MBNL1, indiquant alors une régulation dose-dépendante (Wagner et al. 2016). Ainsi, utiliser le degré de sévérité d'une combinaison de défauts d'épissage comme biomarqueurs de la DM1 permettrait d'estimer la concentration des protéines MBNL1 fonctionnelles dans les muscles des patients, directement corrélée à la sévérité de leurs atteintes musculaires (Wang et al. 2019a; Wagner et al. 2016). Bien que *Mbnl1* soit le paraglogue le plus impliqué dans la splicéopathie musculaire de la DM1, les différents modèles murins de perte de fonction de ces protéines *Mbnl* ont permis de démontrer que

la perte d'expression de *Mbnl2* était en revanche responsable en grande partie de la splicéopathie neuronale observée dans la DM1, avec une très faible contribution de *Mbnl1* (Charizanis et al. 2012; Suenaga et al. 2012). Parmi les défauts d'épissage décrits dans les souris *Mbnl2*KO, le défaut d'exclusion de l'exon 5 de *Grin1* (ou *NmdaR1*) également retrouvé chez des patients DM1 pourrait notamment expliquer les diminutions de transmissions synaptiques observées au niveau de l'hippocampe de ces souris (Charizanis et al. 2012).

Tableau 4 : Principaux épissages alternatifs dérégulés dans les modèles murins knockout *Mbnl*

Gène	Exon/intron	Tissu	Référence(s)
<i>Abi2</i>	Exon 7	Muscle sq.	Lee (2013)
<i>Add1</i>	Exon 15	Cerveau, hippocampe	Charizanis (2012) et Goodwin (2015)
<i>Atp2a1 (Serca1)</i>	Exon 22	Muscle sq.	Lin (2006), Charizanis (2012) et Lee (2013)
<i>Bin1</i>	Exon 7	Muscle sq.	Thomas (2017)
	Exon 11	Muscle sq.	Lee (2013)
<i>Cacna1d</i>	Exon 12a	Cerveau, hippocampe	Charizanis (2012) et Goodwin (2015)
<i>Cacna1s</i>	Exon 29	Muscle sq., cœur	Lee (2013)
<i>Cacna2d1</i>	Exon 19	Muscle sq.	Thomas (2017)
<i>Camk2d</i>	Exons 14-16	Cerveau, hippocampe	Charizanis (2012), Suenaga (2012) et Goodwin (2015)
<i>Capn3</i>	Exon 16	Muscle sq.	Lin (2006)
<i>Clasp2</i>	Exon 16a/b	Cerveau	Goodwin (2015)
<i>Clcn1</i>	Exon 7a	Muscle sq.	Kanadia (2003), Charizanis (2012) et Lee (2013)
<i>Csda</i>	Exon 6	Muscle sq.	Lee (2013)
<i>Csnk1d</i>	Exon 9	Hippocampe	Charizanis (2012)
<i>Dgl2</i>	Exon 17b	Hippocampe	Charizanis (2012)
<i>Grin1 (NmdaR1)</i>	Exon 5	Cerveau, hippocampe	Charizanis (2012) et Goodwin (2015)
<i>Kcnma1</i>	Exon 25a	Cerveau, hippocampe	Charizanis (2012) et Goodwin (2015)
<i>Ldb3 (Zasp)</i>	Exon 11	Muscle sq.	Lin (2006)
<i>Lnp</i>	Exon 13	Muscle sq.	Lee (2013)
<i>Mapt</i>	Exon 2,3 et 10	Cerveau	Goodwin (2015)
<i>Mbnl1</i>	Exon 5	Muscle sq., cœur	Kalsotra (2008) et Du (2010)
	Exon 7	Muscle sq., cœur	Lin (2006) et Lee (2013)
<i>Mbnl2</i>	Exon 7	Muscle sq.	Lin (2006)
<i>Ndrg4</i>	Exon 14	Hippocampe	Charizanis (2012)
<i>Nfix</i>	Exon 9	Muscle sq.	Lee (2013)
<i>Nrap</i>	Exon 12	Muscle sq.	Lin (2006)
<i>Ryr1</i>	Exon 70 et 83	Muscle sq.	Lee (2013)
<i>Ryr2</i>	Exon 4 et 5	Cœur	Lee (2013)
<i>Sorbs1</i>	Exon 25	Muscle sq., cœur	Du (2010), Charizanis (2012) et Lee (2013)
<i>Spag9</i>	Exon 31	Cœur	Lee (2013)
<i>Spna2</i>	Exon 23	Hippocampe	Charizanis (2012)
<i>Tanc2</i>	Exon 23a	Hippocampe	Charizanis (2012)
<i>m-Titin</i>	Exon 5	Muscle sq.	Lin (2006)
<i>Tmem63b</i>	Exon 5	Cœur	Lee (2013)
<i>Tnnt2 (cTnt)</i>	Exon 5	Cœur	Kanadia (2003), Charizanis (2012) et Lee (2013)
<i>Tnnt3</i>	Exon F	Muscle sq.	Kanadia (2003), Lin (2006) et Lee (2013)

Comme nous l'avons vu précédemment, il existe une compensation d'expression entre les deux paralogues *Mbnl1* et *Mbnl2* qui se reflète au niveau fonctionnel. Il a notamment été montré que la perte d'expression de *Mbnl1* entraînait non seulement la surexpression de *Mbnl2* mais surtout une augmentation des interactions de celui-ci avec les ARNm cibles de *Mbnl1*, permettant ainsi d'atténuer partiellement la sévérité des défauts d'épissage dans les muscles des souris *Mbnl1* KO (Lee et al. 2013). Par ailleurs, nous avons vu précédemment que les protéines MBNL régulaient également leurs propres profils d'épissage et

particulièrement l'exon 5 qui joue un rôle important dans la localisation de celles-ci (Konieczny et al. 2018). Il a alors été proposé que la perte d'expression de MBNL1 dans les cellules musculaires (respectivement MBNL2 dans les cellules neuronales) favorisant l'inclusion de l'exon 5 de MBNL2 (respectivement MBNL1) entraînerait alors une proportion plus importante des isoformes nucléaires de ce paralogue, permettant ainsi la restauration partielle de certains défauts d'épissage (Lee et al. 2013; Charizanis et al. 2012). Les souris DKO ont par la suite pu démontrer cette compensation fonctionnelle au niveau de l'épissage alternatif en observant une exacerbation des défauts d'épissage au niveau des muscles squelettiques et du cœur (**Figure 26-A et B**). Il a d'ailleurs été montré que la simple inactivation d'un allèle de Mbnl2 était suffisante pour aggraver certains défauts d'épissage, en accord avec les expériences de régulation dose-dépendante mentionnées plus haut (Wagner et al. 2016; Lee et al. 2013).

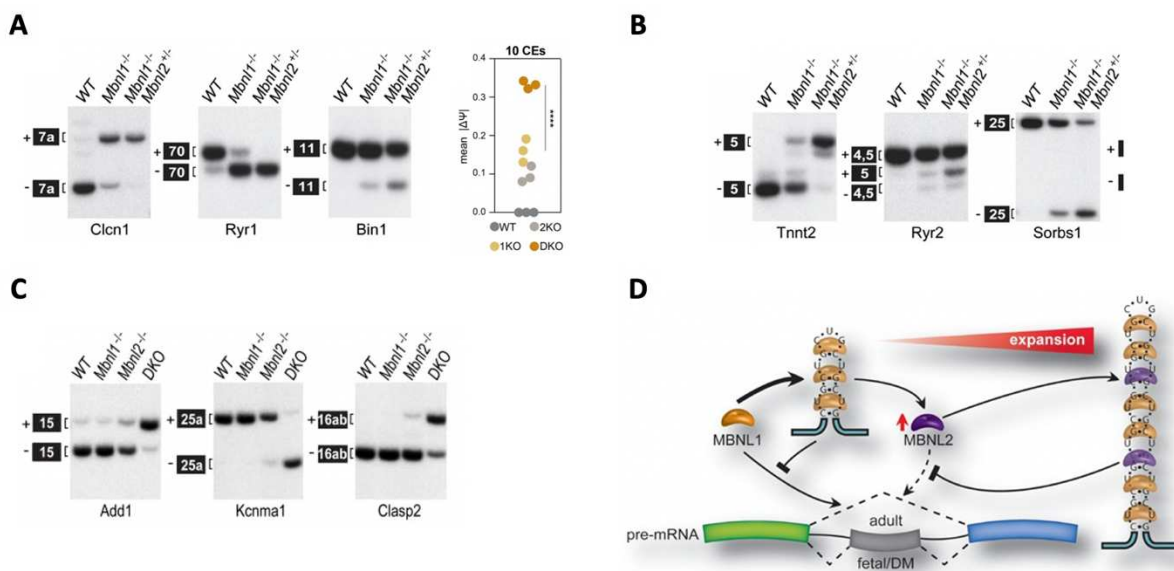


Figure 26 : Exacerbation des défauts d'épissage chez les souris Mbnl1,2 DKO

De manière générale, la perte de fonction de Mbnl1 associée à la déplétion totale ou partielle de Mbnl2 engendre une aggravation des défauts d'épissage caractéristiques de la DM1 et observés dans les muscles squelettiques (**A**) et cardiaque (**B**) des souris Mbnl1 KO ou encore dans le cerveau (cortex frontal et hippocampe) (**C**) des souris Mbnl2 KO. (**D**) Modèle proposé pour l'exacerbation des défauts d'épissage dans un contexte DM1. *Adapté de (Thomas et al. 2017; Goodwin et al. 2015; Lee et al. 2013).*

Ces phénomènes de compensation indiquent ainsi un modèle selon lequel les protéines MBNL2 prendraient le relais au niveau de la régulation des épissages alternatifs musculaires lorsque les protéines MBNL1 sont séquestrées par les *foci* intranucléaires. Cependant dans le cas d'expansions importantes de répétitions CUG chez les patients atteints de formes sévères de la maladie, les protéines MBNL2 finiraient elles-aussi par être séquestrées et ne pourraient alors plus atténuer les défauts d'épissages. Ce modèle est également applicable

aux tissus neuronaux en permutant les deux paralogues puisque la splicéopathie décrite dans le cerveau des souris *Mbnl2* KO est elle aussi aggravée dans le cas de la double déplétion de *Mbnl1* et *Mbnl2* (Goodwin et al. 2015; Lee et al. 2013) (**Figure 26-C et D**). En plus de l'exacerbation générale des défauts d'épissage, les souris DKO ont reproduit l'importante dérégulation de l'épissage des isoformes de la protéine *Mapt/Tau* observée dans la DM1. Cette tauopathie pourrait contribuer à l'altération de la plasticité synaptique chez les patients DM1 (Braz et al. 2018; Goodwin et al. 2015). Ainsi, plusieurs études suggèrent que les souris DKO présentent une splicéopathie comparable à celle des formes congénitales de la DM1. Dans ce contexte, la proportion anormalement élevée d'isoformes nucléaires des protéines MBNL ne vont pas pouvoir atténuer les défauts d'épissage comme dans les modèles simples KO mais vont au contraire être séquestrées à leur tour par les répétitions CUG dans le noyau des cellules, formant ainsi un véritable cercle vicieux (Konieczny et al. 2018; Thomas et al. 2017; Lee et al. 2013).

Concernant le troisième paralogue, les phénotypes décrits chez les souris *Mbnl3* KO n'étaient pas accompagnés de défauts d'épissage (Choi et al. 2016; Poulos et al. 2013). Contrairement au souris *Mbnl1,2* DKO, aucune compensation fonctionnelle n'a été observée concernant les défauts d'épissage puisque les souris *Mbnl1,3* DKO ne présentaient pas plus de dérégulation que les souris *Mbnl1* KO (Choi et al. 2015). De plus, bien que les souris TKO montrent une splicéopathie musculaire de type congénitale, celle-ci n'est que très légèrement augmentée par rapport aux souris *Mbnl1,2* DKO, ce qui démontre que cette splicéopathie est principalement causée par la perte de fonction de *Mbnl1* et *Mbnl2*. En effet, les protéines *Mbnl3* ont été principalement observées dans le cytoplasme, indiquant des fonctions autres que la régulation de l'épissage (Thomas et al. 2017).

2.3.4. Autres défauts attribués à la séquestration des protéines MBNL

Nous avons vu précédemment que les protéines MBNL jouaient également d'autres rôles que celui de facteurs d'épissage. D'importantes dérégulations au niveau de la polyadénylation ont par exemple été observées au niveau du cortex frontal et de l'hippocampe des souris *Mbnl1,2* DKO, dont plus de 500 évènements dérégulés communément entre les souris DKO et des biopsies post-mortem de patients DM1 (Goodwin et al. 2015). Plusieurs de ces défauts de PAA sont impliqués dans des voies associées aux fonctions neuronales mais aussi de manière intéressante dans la voie mTOR et dans la protéolyse médiée par l'ubiquitine, toutes les deux également identifiées via l'analyse des défauts de PAA dans les muscles squelettiques des patients (Goodwin et al. 2015; Batra et al. 2014).

Bien que Mbnl1 ne soit pas le paralogue majeur au niveau de la régulation de l'épissage dans le système nerveux central, plusieurs études ont montré que celui-ci était néanmoins impliqué dans d'autres rôles importants tels que la stabilité de plusieurs ARNm codant des protéines synaptiques ou encore la régulation de la longueur des dendrites et des axones neuronaux (Braz et al. 2018; Wang et al. 2017). Par exemple, la perte d'expression de Mbnl1 chez les souris reproduit la surexpression de la protéine RAB3A observée chez les patients DM1, indépendamment des défauts d'épissage. Cette surexpression perturberait la dynamique et la libération finement régulée des vésicules synaptiques contribuant à l'altération de la neurotransmission dans la DM1 (Hernández-Hernández et al. 2013).

Ces dernières années, plusieurs micro-ARN ont été décrits comme dérégulés dans différents tissus de patients DM1 comme les muscles squelettiques, le cœur ou encore dans le sang (Koehorst et al. 2020). La perte de fonction de MBNL1 entraîne des dérégulations dans la biogenèse des micro-ARN, en particulier celle du micro-ARN miR-1 associée aux défauts cardiaques des patients DM1. Il a été démontré que lorsque les protéines MBNL1 étaient séquestrées, LIN28 pouvait alors se lier au précurseur miR-1 et bloquer ainsi sa maturation. miR-1 ne pouvant plus réguler ses cibles, certaines sont surexprimées telles que le canal calcique spécifique du cœur CACNA1C. Il est très probable que la dérégulation de ce dernier soit impliquée dans les défauts cardiaques de la DM1 puisque des mutations gain-de-fonction du gène *CACNA1C* engendrent des arythmies et des morts soudaines (Rau et al. 2011).

De manière plus générale, plusieurs études ont révélé d'importantes modifications concernant l'expression génique associées à la perte de fonction des protéines MBNL, avec par exemple, près de 70% des défauts d'expression génique observés dans les muscles squelettiques des souris HSA^{LR} retrouvés dérégulés chez les souris *Mbnl1* KO (Du et al. 2010). De la même manière que pour les défauts d'épissage, les défauts d'expression génique pourraient également servir de biomarqueurs de la maladie. En effet, plusieurs de ces défauts ont été corrélés avec le degré de sévérité de certains symptômes comme la faiblesse musculaire, suggérant une signature moléculaire pour les changements d'expression génique dans la DM1 (Wang et al. 2019a). Enfin, contrairement à la faible augmentation de la splicéopathie, plus que 400 gènes ont été retrouvés différentiellement exprimés chez les souris TKO comparé à des souris saines, pour seulement 70 chez les souris *Mbnl1,2* DKO. De manière intéressante, ces défauts d'expression génique sont impliqués dans une cinquantaine de voies biologiques également dérégulées dans les muscles de patients atteints de la forme congénitale de la DM1, et dont plusieurs sont liées au stress comme l'hypoxie (Thomas et al. 2017).

Ensemble, tous ces résultats indiquent une contribution majeure des protéines MBNL dans la physiopathologie de la DM1 en raison d'importantes dérégulations de leurs ARN cibles et permettent d'établir de véritables signatures moléculaires de la maladie, autant au niveau des défauts d'épissage alternatifs qu'au niveau des défauts d'expression génique. Cependant, la majorité des études ayant été conduites par l'intermédiaire de modèles murins, on connaît finalement peu de choses sur ces protéines dans un contexte humain, ce qui est essentiellement lié à la difficulté d'avoir accès à des échantillons issus de patients. Une des solutions permettant de contourner cette limitation réside dans les cellules souches pluripotentes humaines. Isolées pour la première fois il y a plus de 20 ans, ces cellules sont devenues une opportunité unique pour à la fois mieux comprendre les étapes précoces du développement mais également les mécanismes impliqués dans le développement de maladies monogéniques.

3. Les cellules souches pluripotentes humaines

Les cellules souches pluripotentes (CSP) humaines offrent la possibilité d'étudier les processus physiopathologiques tout au long du développement et d'en décortiquer les mécanismes cellulaires et moléculaires sous-jacents. En effet, elles représentent tout d'abord une source inépuisable de matériel puisque leur amplification est en théorie illimitée. Il s'agit de leur capacité d'auto-renouvellement, leur permettant de se diviser à l'identique quasi-indéfiniment. D'autre part, elles sont également capables de se différencier et ainsi donner naissance à tous les types cellulaires qui formeront un organisme (endoderme, ectoderme et mésoderme). Cette dernière propriété est appelée pluripotence. Les CSP se distinguent d'une part des cellules souches totipotentes (œuf fécondé ou cellules issues de celui-ci jusqu'au 4^{ème} jour après fécondation) capables de donner naissance à un organisme complet ainsi qu'aux annexes embryonnaires ; et d'autre part des cellules souches uni ou multipotentes (déjà orientées vers un type cellulaire donné) qui peuvent être d'origine foetale ou adulte.

A ce jour, les CSP humaines peuvent être obtenues à partir de deux origines distinctes : les cellules souches embryonnaires (hESC) et les cellules souches reprogrammées à partir de cellules somatiques et dites induites à la pluripotence. La première source n'existe que de manière transitoire au cours du développement ; entre le 5^{ème} et le 7^{ème} jour après fécondation. Les hESC sont obtenues à partir d'embryons humains surnuméraires issus d'une fécondation *in vitro* (FIV) dans le contexte de programmes de procréation médicalement assistée (PMA) ou encore de diagnostics préimplantatoires (DPI). La deuxième source repose quant à elle sur les travaux de l'équipe du Pr. Shinya Yamanaka qui ont permis la conversion de cellules somatiques en cellules pluripotentes en surexprimant uniquement quatre gènes ; raison pour laquelle on nomme ces dernières « cellules souches induites à la pluripotence » (iPSC).

3.1. Obtention des cellules souches pluripotentes humaines

3.1.1. Les cellules souches embryonnaires humaines

Les hESC sont obtenues à partir d'une structure embryonnaire correspondant au stade de développement appelé « blastocyste » et compris entre le 5^{ème} et le 7^{ème} jour après fécondation *in vitro*. Au début de l'embryogénèse humaine, l'œuf fécondé ou zygote va effectuer des divisions successives pour atteindre dans un premier temps le stade morula à partir du 3^{ème} jour ; stade constitué de 16 à 64 cellules totipotentes. Lors de la division suivante, ces cellules donnent alors naissance à une nouvelle structure embryonnaire

appelée blastocyste. Le blastocyste est composé de trois sous-structures. Il est constitué d'une part par le trophoblaste formé par les cellules les plus externe du blastocyste. Ces cellules forment une membrane étanche et donneront naissance aux futurs tissus extra embryonnaires comme le placenta ou le cordon ombilical. Les cellules les plus internes du blastocyste forment quant à elles une structure compacte appelée masse cellulaire interne (MCI). Enfin le blastocyste est composé d'une troisième partie appelée blastocœle qui est une cavité centrale remplie de liquide. La MCI est elle-même composée de deux structures : l'hypoblaste et l'épiblaste. La première donnera naissance au sac vitellin ou à l'endoderme pariétal. L'épiblaste, plus en profondeur, est constitué de cellules pluripotentes qui donneront naissance à l'ensemble des cellules qui forment l'organisme mais ne peuvent pas, à elles seules, être à l'origine d'un embryon. C'est donc de la MCI que sont issues les hESC, et plus particulièrement des cellules pluripotentes épiblastiques (Shamblott et al. 1998; Thomson et al. 1998).

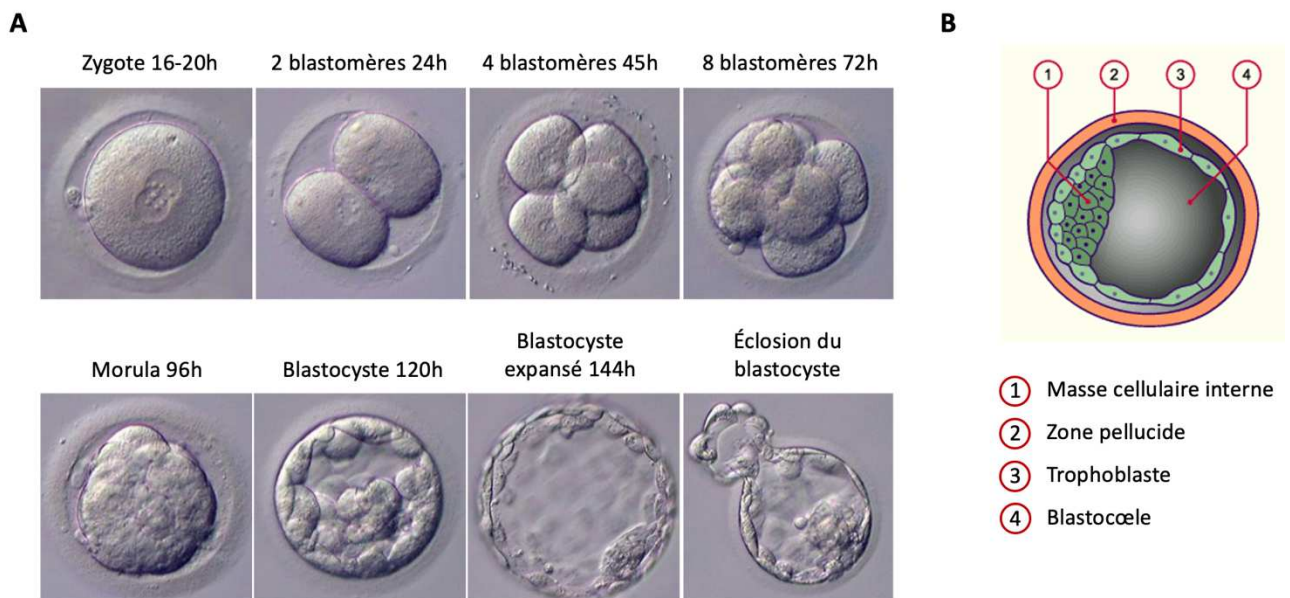


Figure 27 : L'embryogénèse, l'origine des cellules hES

(A) Cinétique de différenciation d'un embryon humain in vitro. A partir de l'œuf fécondé (zygote), les cellules vont suivre des divisions successives jusqu'à atteindre le stade blastocyste en passant par le stade morula. Survient alors la rupture de la zone pellucide permettant l'éclosion de l'embryon.

(B) Représentation schématique de la structure d'un blastocyste. (*Adapté du site internet <http://www.embryology.ch/francais/evorimplantation/furchung01.html>*)

Les lignées d'hESC peuvent être dérivées à partir d'embryons humains surnuméraires obtenus lors de diagnostics préimplantatoires (DPI). En effet, des stratégies de DPI ont été mises en place afin d'identifier des embryons porteurs de mutations causales de maladies monogéniques, souvent rares et incurables. Les embryons 'positifs' pour une telle maladie sont exclus du programme de PMA et peuvent alors être utilisés à des fins de recherche scientifique avec le consentement éclairé des parents, pour dériver des hESC porteuses elles-

aussi de la mutation causale de la maladie (Mateizel et al. 2006). Ainsi, environ 190 lignées d'hESC porteuses de mutations ont pu être générées à ce jour, permettant de modéliser plus de 50 maladies génétiques humaines (Ilic and Ogilvie 2017) (**Tableau 5**).

Disease	Line	Disease	Line	Disease	Line	Disease	Line
Adrenoleuko-dystrophy	SI-201			Ichthyosis	Lis 46_1chthytosis_2_N Lis 45_1chthytosis_1_N	Retinitis pigmentosa	UCLA7 GENEA085
	UM112-1 PGD UM112-2 PGD GENEA073	Factor VIII deficiency	GENEA050 GENEA096 HAD 3	Juvenile retinoschisis	GENEA072	Saethre-Chotzen syndrome	Lis04_Twist
Alpha thalassemia	Lis14_Alport_3	Familial adenomatous polyposis	STR-1-305-APC STR-1-355-APC STR-1-359-APC Lis34_FAP_3	Klinefelter's syndrome	Royan H4 KCL008 FC018 BG01V WIA16	Simpson Golabi Behmel syndrome	GENEA088
Amyotrophic lateral sclerosis; frontotemporal dementia	UM141-6 PGD			Leuko-encephalopathy	Lis 41_LTBL_N	Spinocerebellar ataxia	VUB10_SCA7
Androgen insensitivity	Lis07_AIS_1			Loeys-Dietz syndrome 2	GENEA083 GENEA084	Spinal muscular atrophy	STR-1-221-Scs2 UM134-1 PGD HUES PGD 1 HUES PGD 13 HUES PGD 14 KCL026
Aniridia (PAX6)	Lis08_AIS_2 UM29-2 PGD UM29-3 PGD SI-170	Fanconi's anemia	Lis34_FAP_2 Lis25_FAP_1 SI-128	Marfan syndrome	SI-154	Torsion dystonia	GENEA074 Lis09_DYS_1
Becker muscular dystrophy	SI-158	Fragile X syndrome	WCMC-37			Tuberous sclerosis 2	GENEA086 GENEA087
Beta Thalassemia	SI-164 Q2-8			Merosin deficiency 1A	STR-1-301-MFS VUB08_MFS MFS5 UM89-1 PGD UM89-4 PGD GENEA081	Turner syndrome	KCL041 GENEA069
	KCL035 GENEA058 GENEA059 VUB20_CMT1A			Multiple endocrine neoplasia-type 2A	STR-1-209-MEN2a STR-1-211-MEN2a UM57-1 PGD VUB03_DM1 VUB19_DM1 VUB24_DM1 HAD 1	Viteliform macular dystrophy	GENEA070 KCL015
Becker muscular dystrophy	UM89-3 PGD GENEA064 GENEA062 GENEA063 KCL003 KCL021 KLO42 KLO43 STR-1-203-CFTR STR-1-251-CFTR SI-257	Hemophilia B	Lis39_FXS8_N Lis38_FXS7_N Lis37_FXS10_N Lis29_FXS_7 Lis01_HEFX1 Lis02_FXS_2 Lis03_FXS_4 Lis24_FXS_5 Lis26_FXS_6 UM9-1PGD ES-11EM	Myotonic dystrophy	SI-148	Von Hippel-Lindau disease/syndrome	KCL016 KCL017 GENEA060 GENEA061 HUES PGD 2 KCL029
BRCA1	UM59-2 PGD UM59-4 PGD	Hereditary multiple exostoses	GENEA097 GENEA098 SI-186		SI-153 KCL018 GENEA066 GENEA067 Lis12_DM_1	Wiscott-Aldrich syndrome	KCL029 STR-1-229-MTXX
Charcot-Marie-tooth disease type 1	UM89-3 PGD GENEA064 GENEA062 GENEA063 KCL003 KCL021 KLO42 KLO43 STR-1-203-CFTR STR-1-251-CFTR SI-257	Huntington's disease	SI-187 SI-194 VUB05_HD VUB28_HD_MFS STR-1-155-HD KCL005 KCL012 KCL013 KCL027 KCL036 KCL078 HUES PGD 16 UM17-1 PGD GENEA017 GENEA018 GENEA046 GENEA090 GENEA091 GENEA089 HS799		LI-148 KCL018 GENEA066 GENEA067 Lis12_DM_1	X-kinked myopathy with excessive autophagy	STR-1-231-MTXX GENEA068
Cystic fibrosis	VUB04_CF VUB22_CF HAD 2			Nemaline myopathy 2	Lis19_DM_2 GENEA078 GENEA079 GENEA080	Wilm's tumor	GENEA068
	GENEA041 GENEA040 SI-180			Neurofibromatosis	SI-137 SI-138 SI-139 SI-140 SI-235 KCL024 KCL025 Lis 47_NF1_2_N		
Duchenne muscular dystrophy	HUES PGD 3 Lis48_DMD_6_N	Hydrocephaly	Lis49_Hydrocephaly_2_N Lis35_Hydrocephaly_1 UM15-4 PGD		Lis 42_NF1_1_N Lis43_connexin_3_N		
	Lis23_DMD_5 Lis10_DMD_1 Lis11_DMD_2 Lis20_DMD_3 Lis22_DMD_3 SI-245			Noonan syndrome	Lis17_Connexin_1 Lis18_Connexin_2 Lis21_Noonan_1		
Emery-Dreifuss muscular dystrophy	STR-1-171-GLA VUB09_FSHD	Hydroxysteroid dehydrogenase 4 deficiency	UM38-2 PGD	Osteogenesis imperfecta	VUB23_OI SA002		
Fabry disease	GENEA024 GENEA049	Hypertrophic cardiomyopathy (MYBPC3)	GENEA077 GENEA071 GENEA065	Patau syndrome (trisomy 13)	Mtz-hES13 FY-hES-5 SA002		
Factoscapulohumeral muscular dystrophy		Hypochondroplasia					
		Incontinentia pigmenti					
		Infantile neuroaxonal dystrophy					

Tableau 5 : Les lignées hESC porteuses de mutations génétiques dans le monde

Adapté de (Ilic and Ogilvie 2017).

Les premières lignées de cellules ES ont été dérivées à partir d'embryons de souris en 1981 (Evans and Kaufman 1981) permettant de maîtriser ces cellules ES murines en culture *in vitro*. Ces premières études ont permis de démontrer le caractère « immortel » de ces cellules sur un grand nombre de passages tout en conservant leur propriété de pluripotence ainsi qu'en évitant l'apparition d'anomalies caryotypiques (Evans and Kaufman 1981; Martin 1981). Pour ce faire, les ESC sont isolées par des techniques de microdissection des blastocystes, et sont ensuite ensemencées en co-culture sur des cellules nourricières (le plus souvent des fibroblastes embryonnaires murins) ayant pour but de sécréter des facteurs de maintien de la pluripotence. Au contact de ces cellules nourricières, les ESC forment des colonies qui sont par la suite amplifiées en les dissociant en petits fragments via des méthodes enzymatiques ou mécaniques, afin de générer à leur tour d'autres colonies.

C'est 17 ans plus tard que la première lignée d'hESC fut dérivée par une équipe américaine (Thomson et al. 1998) en utilisant les mêmes méthodes utilisées précédemment chez la souris. Alors que l'obtention d'hESC nécessitait la destruction des embryons *ex utero*, de nouvelles études ont permis par la suite de générer des lignées d'hESC à partir d'un unique blastomère sans interférer avec le bon développement de l'embryon (Klimanskaya et al. 2006; Geens et al. 2009). De plus, les méthodes de dérivation ont évolué, notamment avec l'utilisation de molécules synthétiques et de cytokines permettant une meilleure maîtrise des cultures cellulaires (Chen et al. 2014). D'autres stratégies ont également été mises en place dans le but de générer des lignées d'hESC de grade clinique en étant dérivées dans des conditions GMP (Good Manufacturing Practice) et dont les conditions de culture ne comportent aucun facteur xénogène (Tannenbaum et al. 2012; Ilic and Ogilvie 2017).

Cependant, l'utilisation des cellules souches embryonnaires humaines soulève un grand nombre de questions éthiques. La question la plus délicate est de savoir à partir de quel moment on considère l'embryon comme un être humain, soit le moment du début de la vie. En France, la question du statut juridique de l'embryon s'est posée dès 1975 avec la loi relative à l'interruption volontaire de grossesse. Il faut cependant attendre 1994 pour qu'une première loi de bioéthique soit instaurée, loi qui interdisait toute recherche sur l'embryon. Avec la dérivation des premières lignées de hESC et les premières démonstrations quant à l'intérêt de ces cellules pour des applications biomédicales, les législateurs ont été amenés à modifier la loi de bioéthique en 2004. Cette deuxième version de la loi de Bioéthique (Loi n°2004-800 du 6 août 2004 relative à la bioéthique) maintenait le régime d'interdiction de toute recherche sur l'embryon mais permettait dans le cadre de dérogations demandées à l'Agence de la Biomédecine, de développer des programmes de recherche sur les hESC. Cependant, cette loi était assortie d'un moratoire de 5 ans. La loi de Bioéthique Française a donc été révisée en 2011 puis en 2013 pour aboutir à une levée du système d'interdiction avec dérogation, désormais remplacé par un système d'autorisation encadré par l'Agence de la Biomédecine. Cette nouvelle législation, même si plus souple, nécessite que les programmes de recherche respectent quatre conditions bien spécifiques : la pertinence du projet scientifique, la finalité médicale de la recherche, l'absence de techniques alternatives et le respect des principes éthiques. En 2019 se sont tenus les Etats Généraux de Bioéthique qui visaient à reviser le texte de la loi pour intégrer en particulier les récents développements obtenus dans ce domaine. Ces états généraux ont abouti à une proposition de loi qui est actuellement en deuxième lecture au Sénat et qui devrait permettre de passer d'un système d'autorisation encadrée à un simple système déclaratif.

Afin de s'affranchir des questions éthiques soulevées par l'utilisation des hESC, plusieurs équipes dans le monde se sont intéressées à comprendre les mécanismes impliqués dans le

maintien de la pluripotence et ces travaux ont permis par la suite d'identifier des procédés permettant de générer des cellules ayant les mêmes caractéristiques que les cellules ES. C'est ainsi qu'à démarrer l'émergence des cellules souches induites à la pluripotence.

3.1.2. Les cellules souches induites à la pluripotence

En étudiant les mécanismes de la pluripotence, l'équipe japonaise du Pr. Yamanaka a identifié en 2006 la possibilité de convertir des cellules somatiques adultes de souris en cellules présentant les deux propriétés cardinales des cellules pluripotentes, l'auto-renouvellement et la capacité de différenciation vers les trois feuilletts embryonnaires. Ces cellules, appelées cellules souches induites à la pluripotence (iPSC), ont été générées par la surexpression de quatre facteurs de transcription introduits dans les cellules somatiques via des vecteurs rétroviraux (Takahashi and Yamanaka 2006). Cette étude a ainsi permis d'identifier quatre gènes (OKSM : Oct4, Klf4, Sox2 et c-Myc) nécessaires et suffisants pour reprogrammer ces cellules en iPSC (**Figure 28**).

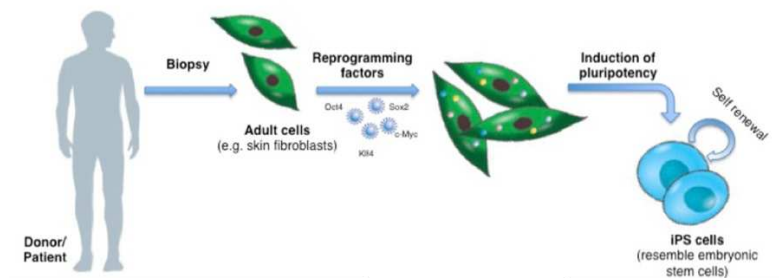


Figure 28 : Représentation schématique de l'obtention d'hiPSC
(Adapté de M. Roszbach, Factsheet EuroStemCell).

En 2007, la même équipe réussit à étendre ce protocole à des fibroblastes adultes humains permettant de générer la toute première lignée de cellules iPS humaines (hiPSC) (Takahashi et al. 2007). Des résultats similaires furent obtenus et décrits la même année par une équipe américaine qui publia à son tour un protocole de reprogrammation avec cependant un cocktail différent de facteurs de transcription (Oct4, Sox2, Nanog et Lin28) et avec l'aide cette fois-ci de lentivirus (Yu et al. 2007). La reprogrammation porte sur la réactivation du programme de pluripotence intrinsèque aux cellules. Il a été démontré que les gènes Oct4 et Sox2 étaient indispensables à cette réactivation alors que les autres gènes (Klf4, c-Myc, Nanog et Lin28) jouent un rôle au niveau de la structure de la chromatine ou encore sur le cycle cellulaire, facilitant ainsi la reprogrammation (Kim et al. 2009). Ces protocoles permettent la reprogrammation de cellules somatiques en hiPSC avec une

efficacité comparable, de l'ordre de 0,02% (Malik and Rao 2013). Comme les cellules ES, les hiPSC forment des colonies et peuvent également être isolées, amplifiées et différenciées.

Les travaux de recherche ayant permis l'obtention de cellules souches induites à la pluripotence ont valu, en 2012, le prix Nobel de médecine au Pr. Yamanaka ainsi qu'au Pr. Gurdon, pour ses travaux pionniers sur la reprogrammation par transfert nucléaire. En effet, cette découverte a ouvert le champ des possibles autant au niveau de la modélisation de maladies humaines qu'au niveau des domaines de la médecine régénérative et de la médecine personnalisée (Rowe and Daley 2019; Zakrzewski et al. 2019). De plus, la possibilité de pouvoir désormais passer d'un état différencié à un état pluripotent a complètement remis en question tous les concepts originels de différenciation cellulaire avec en particulier l'idée de hiérarchisation du destin des cellules, passant de manière unidirectionnelle d'un état immature (pluripotent) à mature (différencié). Ce concept fut illustré par Conrad Waddington en représentant une cellule descendant le sommet d'une montagne jusqu'au fond d'une vallée, avec l'idée qu'une fois au fond de la vallée, la cellule avait acquis son identité et ne pouvait plus en bouger (Waddington 1957). À la suite des travaux du Pr. Yamanaka, des études ont montré qu'il était possible, en plus du fait de pouvoir revenir à un état intermédiaire ou indifférencié, de passer directement d'un état différencié à un autre, en forçant l'expression de facteurs de transcription spécifiques de certains tissus. Par exemple, l'expression ectopique du facteur de transcription MYOD est suffisante pour convertir des fibroblastes directement en myoblastes (Takahashi and Yamanaka 2016). C'est ainsi qu'ont été revisités les notions de trans-différenciation et que sont apparus les concepts d'état intermédiaire de différenciation (**Figure 29**).

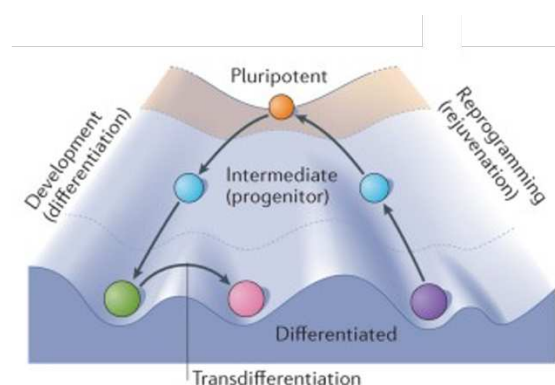


Figure 29 : Revisite du concept de destinée cellulaire et notion de trans-différenciation

Le concept de destinée cellulaire a totalement été remis en question avec les possibilités de reprogrammation et de trans-différenciation. En reprenant la représentation de Conrad Waddington, il est désormais possible de remonter la « montagne » ou encore de passer d'une « vallée » à une autre. *Adapté de (Takahashi and Yamanaka 2016).*

Les premières expériences de reprogrammation de lignées hiPSC utilisaient au départ des vecteurs intégratifs, comme des lentivirus, pouvant être à l'origine de mutations susceptibles d'altérer la fonction de certains gènes et donc de perturber la physiologie et l'intégrité des cellules. Il existe aujourd'hui plusieurs méthodes basées sur l'utilisation des ARNm, d'épisome ou encore du virus de Sendai pour pallier ce problème (**Figure 30**) (González et al. 2011; Beghini et al. 2020). D'autres études ont aussi permis d'identifier des molécules capables d'augmenter l'efficacité de la reprogrammation comme les inhibiteurs de HDAC ou les facteurs de remodelage de la chromatine (Huangfu et al. 2008). De même, alors que les premières descriptions de reprogrammation concernaient des fibroblastes cutanés, il est maintenant possible de reprogrammer des cellules sanguines issues d'une simple prise de sang, ou encore des cellules d'urine, ce qui est beaucoup moins invasif pour les patients (Staerk et al. 2010; Loh et al. 2010; Bento et al. 2020). Ces dernières avancées ont également permis d'optimiser l'efficacité de reprogrammation de 0,1 à 1% en fonction des types cellulaires (Malik and Rao 2013).

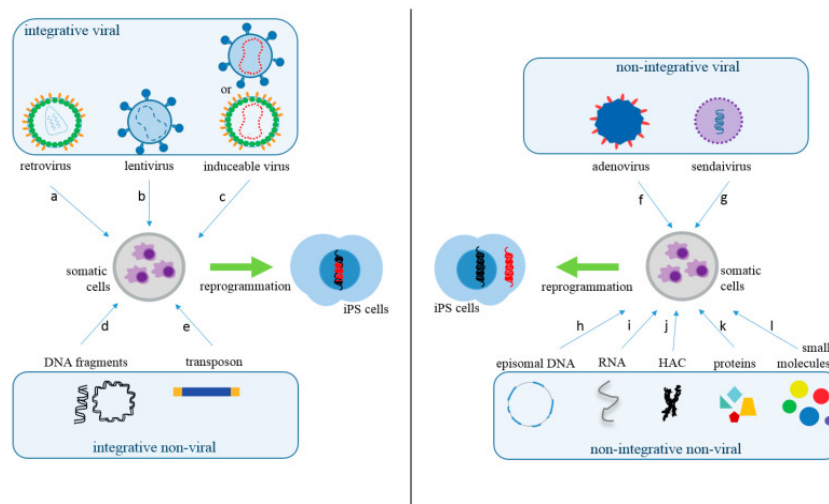


Figure 30 : Méthodes de reprogrammation des cellules somatiques pour la génération d'hiPSC

Elles peuvent être classées en quatre catégories selon leurs critères intégratif et viral. Les méthodes virales intégratives correspondent aux rétrovirus (a), aux lentivirus (b) ou encore aux rétrovirus ou lentivirus inducibles (c) ; et les méthodes non virales intégratives, aux fragments d'ADN linéaires ou circulaires (d) ou aux transposons (e). Concernant les méthodes non intégratives, elles peuvent aussi être virales avec les adénovirus (f) ou le virus Sendai (g) ; ou bien non virales avec l'ADN épisomal (h), les ARN (i), le chromosome artificiel humain (HAC) (j), les protéines (k) ou les petites molécules (l) (Beghini et al. 2020).

3.2. Caractéristiques des CSP humaines

3.2.1. Maintien de la pluripotence *in vitro*

Le maintien de hESC et hiPSC *in vitro* repose sur la conservation des deux propriétés majeures des CSP, à savoir l'auto-renouvellement et la pluripotence, tout en préservant leur intégrité caryotypique (Henry et al. 2018). Ceci nécessite la mise en place d'une balance finement régulée permettant de contrôler le passage d'un état à un autre. Il existe un grand nombre de méthodes pour cultiver ces cellules. Au départ, les premières méthodes utilisaient soit un sérum de remplacement (KSR pour Knock-out Serum Replacement) (Lee et al. 2006) ou encore des cellules nourricières, principalement des fibroblastes embryonnaires humains (MEF) dont la prolifération a été préalablement inactivée soit par traitement chimique à la Mitomycine C, soit par irradiation via rayons γ (Mallon et al. 2006). Désormais, les CSP sont majoritairement cultivées sur des coating comme la vitronectine ou le Matrigel, avec des milieux de culture contenant des petites molécules permettant d'activer ou d'inhiber certaines voies de signalisation impliquées dans la pluripotence et la différenciation cellulaire (Rivera et al. 2020). Ceci permet l'amplification des CSP sous forme de colonies qui sont scrupuleusement contrôlées (principalement visuellement sur des critères morphologiques). Ces colonies sont ensuite sélectionnées et dissociées mécaniquement (avec l'aide d'une aiguille par exemple) pour être à leur tour amplifiées. Il est également possible d'utiliser des méthodes enzymatiques pour dissocier les colonies. Bien que moins douces, celles-ci permettent une amplification plus importante des CSP en "single cell". De plus, l'utilisation d'un inhibiteur des Rho-kinases (ROCK) tel que le Y-27632, permet d'augmenter la survie des CSP une fois dissociées, améliorant la résistance de ces dernières notamment lors des étapes de décongélation (Watanabe et al. 2007).

Le maintien des CSP en culture est ainsi effectué à l'aide de molécules jouant un rôle d'une part sur l'activation des voies de signalisation impliquées dans la pluripotence et d'autre part sur l'inhibition des voies de différenciation. Plusieurs molécules ont été identifiées comme facteurs de régulation de la balance auto-renouvellement/différenciation. En effet, des molécules comme le FGF2 (*Fibroblast Growth Factor 2*), l'Activine A et le TGF β (*Transforming Growth Factor β*) permettent le maintien de la pluripotence en induisant l'expression des gènes de pluripotence *Oct4*, *Sox2* ou encore *Nanog* (Eiselleova et al. 2008). Ceux-ci sont capables de s'auto-réguler en formant une boucle de régulation de la pluripotence (**Figure 31**).

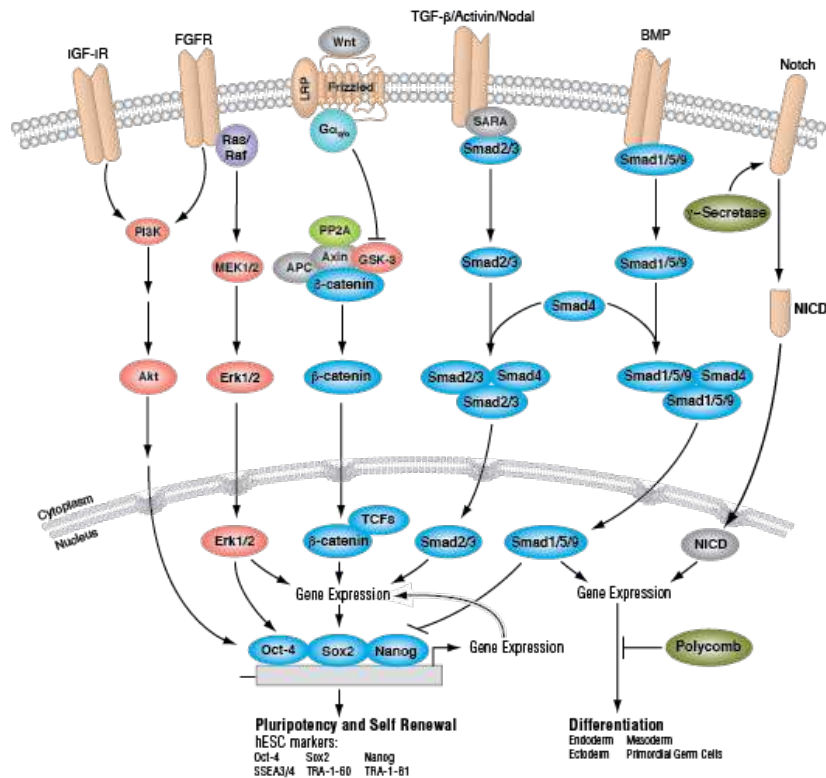


Figure 31 : Voies de signalisation impliquées dans la pluripotence et la différenciation des CSP

<http://www.cellsignal.com/contents/science-cst-pathways-developmental-biology/pluripotency-and-differentiation/pathways-pluripotency>

Le gène *Nanog* est considéré comme le facteur clé de la pluripotence car il a été montré par plusieurs études qu'il était essentiel pour bloquer la différenciation et induire l'état de pluripotence (Saunders et al. 2013). Le gène *Nanog* est activé par la voie de signalisation Activine/Nodale via les facteurs SMAD2/3 qui vont se lier et activer son promoteur (Vallier et al. 2009). Le FGF2, présent dans tous les milieux de culture des CSP, permet l'activation de la voie PI3K/AKT. Une fois activée, celle-ci réprime la voie RAF/MEK/ERK ainsi que la voie WNT permettant la fonction pro auto-renouvellement des SMAD2/3 (Singh et al. 2012) (**Figure 32-A**). A l'inverse, lorsque la voie PI3K/AKT n'est pas activée, les voies WNT et RAF/MEK/ERK ne sont plus inhibées, aboutissant à la séquestration de GSK3. Ceci a pour conséquence une modification de l'action des SMAD2/3 qui vont alors exercer une fonction pro-différenciative (**Figure 32-B**). De plus, la voie BMP va quant à elle jouer un rôle d'inhibiteur de pluripotence via l'action cette fois-ci des SMAD1/5/8-9 qui vont réprimer l'expression de *Nanog* (Suwinska and Ciernerych 2011).

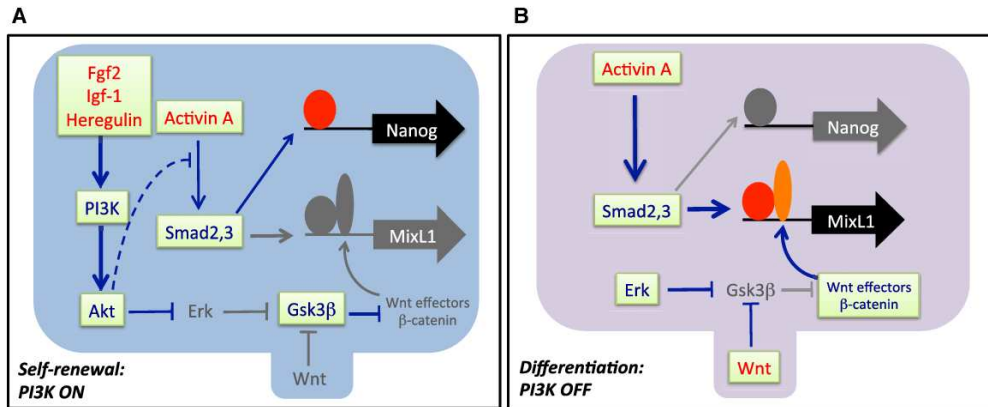


Figure 32 : Modèle des réseaux de signalisation régulant la balance pluripotence/différenciation des CSP humaines

(A) Lorsque la voie PI3K/AKT est activée, cela induit l'auto-renouvellement et le maintien de la pluripotence par inhibition des facteurs "pro-différenciation". (B) A l'inverse, lorsqu'elle est inactive toute la signalisation "pro-différenciative" n'est plus inhibée (Singh et al. 2012).

3.2.2. Caractéristiques moléculaires des CSP humaines et contrôle qualité

Afin de déterminer l'état de pluripotence des cellules et d'en assurer le contrôle, de nombreux marqueurs de surface ont été décrits comme associés à l'état indifférencié des CSP (Mateizel et al. 2006). Il s'agit de marqueurs de surface tels que SSEA3 (*Stage Specific Embryonic Antigen*), SSEA4, TRA-1-60 (*Tumour Rejection Antigen*), TRA-1-81, TRA-2-49, TRA-2-54 et GCTM-2 (*Germ Cell Tumour Marker*). De même l'expression des facteurs de transcription impliqués dans le maintien de la pluripotence, tels que Nanog ou Oct4, peut être suivie par des techniques de RT-qPCR ou de cytométrie en flux. Parallèlement à ces contrôles, il est également important de valider l'intégrité génomique des CSP humaines. En effet, même si ces cellules sont définies comme immortelles, plusieurs études ont montré ces dernières années la présence d'anomalies récurrentes avec le temps en culture. Parmi celles-ci, l'une des plus fréquentes est la duplication du chromosome 20q, et récemment plusieurs mutations spontanées du gène *TP53* codant la protéine suppresseur de tumeur P53 ont été décrites (Merkle et al. 2017; Avery et al. 2013; Lefort et al. 2009). Ceci est d'autant plus important pour les hiPSC car les premiers passages qui font suite à leur reprogrammation représentent une étape critique pour l'intégrité de ces cellules ainsi que pour l'établissement de leur nouvelle identité (Steichen et al., 2019).

La deuxième propriété majeure des CSP est leur potentiel de différenciation qui peut être évalué à la fois *in vitro* et *in vivo*. Afin de contrôler ce potentiel, une des méthodes les plus couramment utilisées est d'évaluer la capacité des CSP humaines à se différencier spontanément en culture en absence de coating et de facteurs extracellulaires. Celles-ci vont

alors former des corps embryoides (EB pour embryoid bodies) composés des trois feuillets embryonnaires : ectoderme, endoderme et mésoderme. La capacité de différenciation des CSP humaines peut également être évaluée *in vivo* en injectant sous cutanément ces cellules dans des souris immunodéficientes aboutissant à la formation de tumeurs embryonnaires, appelées tératomes, et composées de structures cellulaires issues des trois feuillets embryonnaires (Kim et al. 2009). Outre ces méthodes de différenciation spontanée, les CSP peuvent surtout être différenciées *in vitro* dans des conditions bien définies à l'aide de petites molécules et cytokines jouant sur les voies de signalisation qui régulent la différenciation cellulaire ou encore via des systèmes de surexpression de facteurs de transcription. Au cours des deux dernières décennies, une multitude de protocoles de différenciation permettant de guider les CSP humaines vers un type cellulaire précis ont été publiés. L'important intérêt de la recherche dans ce domaine a permis l'obtention de différents types cellulaires extrêmement variés à partir des CSP humaines (**Figure 33**) (Williams et al. 2012). Ces études ont également permis de démontrer que de manière générale les hiPSC répondent aux mêmes signaux extracellulaires que les hESC, avec toutefois quelques variations au niveau des cinétiques de différenciation ou encore au niveau de l'hétérogénéité des cultures différenciées (Cahan and Daley 2013).

Dans le cadre de cette thèse nous nous intéresserons tout particulièrement à la différenciation des CSP humaines en cellules musculaires squelettiques, cellules d'intérêt pour l'étude de la dystrophie myotonique de type 1.

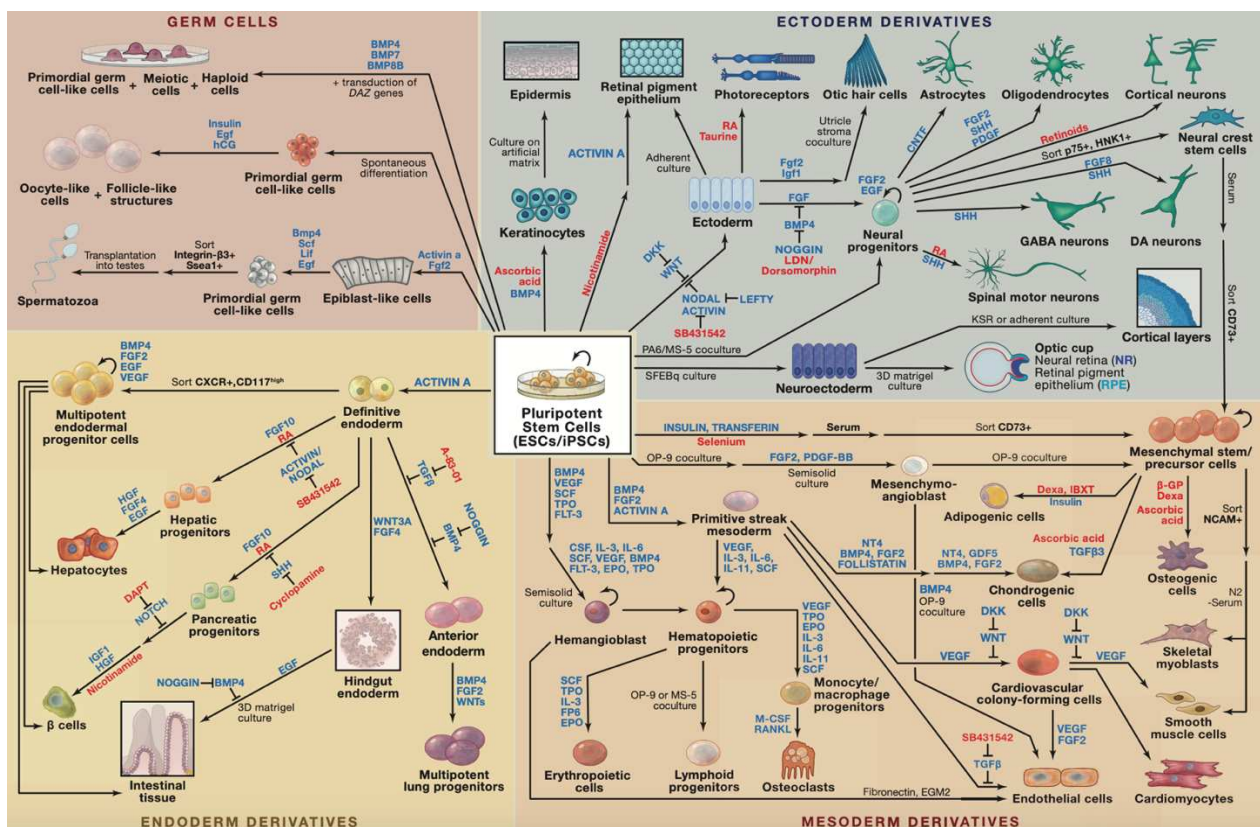


Figure 33 : Différenciation ciblée des CSP
Adapté de (Williams et al. 2012).

3.2.3. Différenciation ciblée des CSP humaines en cellules musculaires squelettiques

On peut distinguer deux méthodes majeures pour tenter de récapituler la myogenèse *in vitro* à partir des CSP (**Figure 34**). La première et la plus utilisée par la communauté correspond à la reprogrammation directe, en utilisant des stratégies de surexpression de facteurs de transcription myogéniques. Parmi ces stratégies, les plus courantes utilisent la surexpression du facteur MyoD (Maffioletti et al. 2015) ou encore celle des facteurs Pax3/7 (Filareto et al. 2015; Darabi and Perlingeiro 2016). Cela aboutit à une différenciation forcée des CSP vers le linéage du muscle squelettique. Ces stratégies de reprogrammation directe présentent néanmoins des inconvénients importants, à savoir un contournement des étapes précoces du développement ; critique dans le cadre de l'étude de maladies présentant des atteintes développementales, ainsi que le besoin la plupart du temps d'une étape de tri cellulaire (Chal and Pourquié 2017).

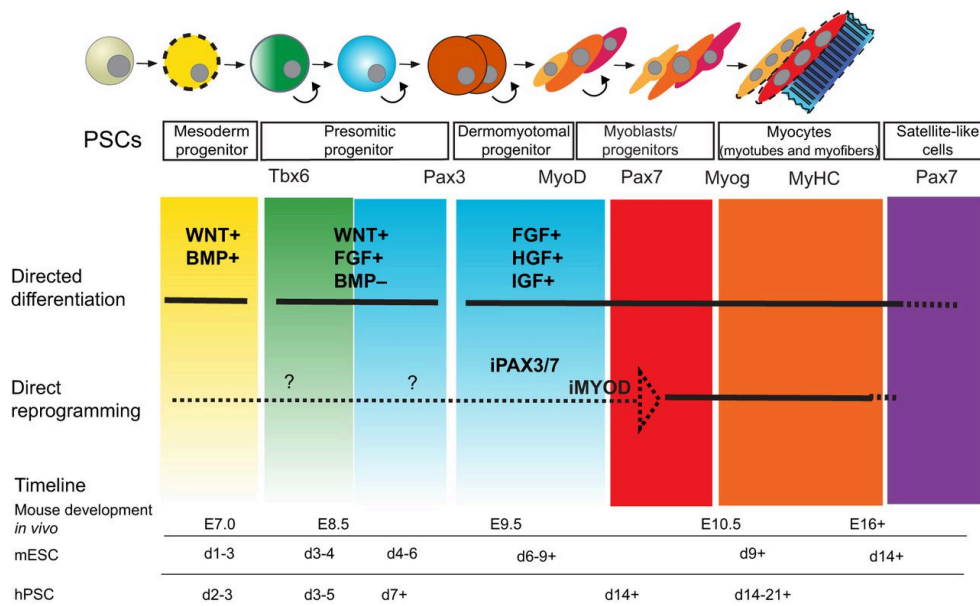


Figure 34 : Obtention de cellules musculaires squelettiques à partir des CSP

Les deux méthodes d'obtention de cellules myogéniques sont schématisées ici. La différenciation dirigée permet de suivre les étapes du développement *in vivo* alors que la reprogrammation directe contourne les étapes précoces de la myogenèse. L'expression de certains facteurs de transcription (Tbx6, Pax3, Pax7, MyoD ou MyoG) et de certaines protéines spécifiques du muscle (Desmine ou MyHC) permettent de suivre le déroulement de la différenciation myogénique (Chal and Pourquié 2017).

À l'opposé, la deuxième méthode permettant d'obtenir des cellules musculaires squelettiques à partir des CSP consiste en des stratégies de différenciations dirigées, dépourvues de transgènes et qui semblent plus appropriées d'un point de vue développemental. Il s'agit ici d'essayer de reproduire *in vitro* les étapes normales du développement *in vivo* à l'aide de petites molécules, cytokines, ou facteurs de croissance définis et permettant d'activer et/ou d'inhiber certaines voies de signalisation. Plusieurs études ont tenté d'établir un protocole de différenciation en suivant cette stratégie de différenciation dirigée. Parmi celles-ci, trois semblent plus prometteuses que les autres, notamment en raison de l'utilisation d'inhibiteurs GSK3 β , permettant d'induire le mésoderme dans les étapes les plus précoces de la différenciation (Kodaka et al. 2017) (**Tableau 6**). Les trois protocoles utilisent le CHIR99021 pour activer la voie Wnt et ayant pour autre effet d'inhiber la voie GSK3 β . Seule l'équipe du Pr. Pourquoié se distingue en utilisant également du LDN, un inhibiteur de la voie BMP, qui améliorerait alors l'obtention de progéniteurs présomitiques (Sakurai et al. 2012; Chal and Pourquoié 2017) (**Tableau 6**). Le protocole de l'équipe canadienne du Dr. Stanford (Shelton et al. 2014) diffère surtout au niveau du type de culture cellulaire puisque cette dernière propose un protocole d'induction myogénique relativement long et s'effectuant en corps embryoides (EB), agrégats tridimensionnels cultivés en suspension. Enfin le protocole initialement commercialisé par Genea Biocells (Caron et al. 2016) présente quant à lui un réel avantage puisqu'il permet l'obtention de myotubes après seulement 25 jours de culture cellulaire (moins de 4 semaines). Cependant ce dernier ne permet pas de générer de cellules progénitrices PAX7+, contrairement au protocole du Pr. Pourquoié (**Tableau 6**). L'obtention de ces cellules dites "satellite-like" est important d'un point de vue thérapeutique car elles ont la capacité de se diviser de manière asymétrique, ce qui permet d'une part l'auto-renouvellement et d'autre part la génération d'un myoblaste pouvant proliférer et contribuer à la régénération musculaire (Chal and Pourquoié 2017).

Type de culture	Type(s) cellulaire(s)	Protocole	Principaux facteurs utilisés	Référence
EB (corps embryoides)	ESC/iPSC (humaines et murines)		Inhibiteur GSK3 β , BMP, VEGF, Inhibine β , FGF2	Shelton et al., 2014
2D	ESC/iPSC (humaines)		Inhibiteur GSK3 β , inhibiteur BMP, FGF2, HGF, IGF1	Chal et al., 2016
2D	ESC (humaines)		Inhibiteur GSK3 β , acide ascorbique, inhibiteur Alk5, dexaméthasone, EGF, insuline	Caron et al., 2016

Tableau 6 : Exemple de protocoles d'induction myogénique par différenciation dirigée (C: CHIR99021, L: LDN, F: FGF2, K: KSR, H: HGF, I: IGF1 (Chal et al., 2016) / Insulin (Caron et al., 2016), HS: Horse serum, Ai: Alk5 inhibitor, E: EGF, P: PDGF, IG: IGF1, AA: ascorbic acid, O: oncostatin, N: necrosulfonamide, Dex: dexamethasone, SB: SB431542) *Adapté de (Chal et al. 2016; Caron et al. 2016; Shelton et al. 2014).*

Bien que l'on puisse observer la présence de sarcomères très organisés, le niveau de maturation des myotubes obtenus *in vitro* à partir des CSP reste toutefois limité. En effet, des analyses ultrastructurales ont montré par exemple que ces fibres musculaires ne parvenaient pas à former un réseau de tubules T mature, caractéristique des fibres musculaires adultes (Skoglund et al. 2014). En effet, sur la base de leurs caractéristiques morphologiques, ces myotubes dérivés des CSP présenteraient davantage un phénotype fœtal ou périnatal. Outre le caractère immature intrinsèque des fibres, cette maturation incomplète pourrait également s'expliquer par le manque d'innervation et de vascularisation dans ces cultures cellulaires (Chal and Pourquie 2017). Ainsi, des systèmes cellulaires plus complexes ont été développés, avec notamment la mise en place de systèmes tridimensionnels (3D) comme par exemple la génération d'un myocarde humain artificiel à partir de fibroblastes et de cardiomyocytes dérivés de CSP humaines contenus au sein d'un patch construit par impression 3D, permettant d'améliorer la maturation de ces cardiomyocytes (Tiburcy et al. 2017). Plus récemment, un tissu musculaire squelettique artificiel humain a été mis au point à partir de plusieurs types cellulaires dérivés d'hiPSC (cellules musculaires, motoneurones, péricytes et cellules endothéliales vasculaires) afin d'augmenter la complexité histologique. Cette étude décrit également l'utilisation d'hydrogels dans le but notamment d'aligner les cellules musculaires dérivées d'hiPSC et d'améliorer leur maturation (Maffioletti et al. 2018).

3.3. Les CSP humaines et la modélisation pathologique

L'utilisation des CSP humaines porteuses de la mutation causale d'une maladie génétique ouvre le champ des possibles pour modéliser ces pathologies afin de mieux comprendre, voire même de découvrir, les mécanismes moléculaires et cellulaires sous-jacents. En effet, leur capacité de se différencier dans tous les types cellulaires permet ainsi de reproduire certains phénotypes cellulaires et/ou moléculaires propres aux tissus et organes affectés par une maladie génétique.

3.3.1. Contribution des CSP humaines dans la DM1

Ces dernières années, un nombre croissant de recherches ont été publiées démontrant l'intérêt des CSP humaines pour l'étude de la DM1. Ces études ont par exemple permis l'étude de l'instabilité des répétitions CTG et de manière plus générale de modéliser la pathologie dans le but d'identifier de nouveaux mécanismes moléculaires et de nouvelles pistes thérapeutiques, notamment en conduisant des criblages moléculaires. Plusieurs lignées hESC (Mateizel et al. 2006; Dumevska et al. 2016) et hiPSC (Du et al. 2013; Xia et al. 2013; Ueki et al. 2017; Martineau et al. 2018; Mondragon-Gonzalez and Perlingeiro 2018; Wang et al. 2018b; Kim et al. 2019) porteuses de la mutation causale de la DM1 ont en effet été générées ces dernières années. L'obtention de CSP DM1 a notamment permis d'étudier l'instabilité des répétitions CTG avec le temps de culture. En effet, il a été démontré que le nombre de répétitions suivait une augmentation progressive corrélée au nombre de passage dans ces cellules indifférenciées (De Temmerman et al. 2008; Du et al. 2013; Ueki et al. 2017). Cependant aucun changement significatif de la taille des répétitions n'a été observé suite à la différenciation des CSP humaines DM1 (Ueki et al. 2017).

Différentes études ont tout d'abord démontré que les CSP humaines pouvaient reproduire certains stigmates moléculaires et cellulaires associés à la DM1 comme par exemple la présence d'agrégats intranucléaires d'ARN mutants ou encore en reproduisant certains défauts d'épissage propres à la pathologie comme le défaut d'inclusion de l'exon 11 de BIN1 et le défaut de l'exclusion de l'exon 5 du NMDA-R1 (Marteyn et al. 2011; Xia et al. 2013; Mondragon-Gonzalez and Perlingeiro 2018) (**Figure 35**).

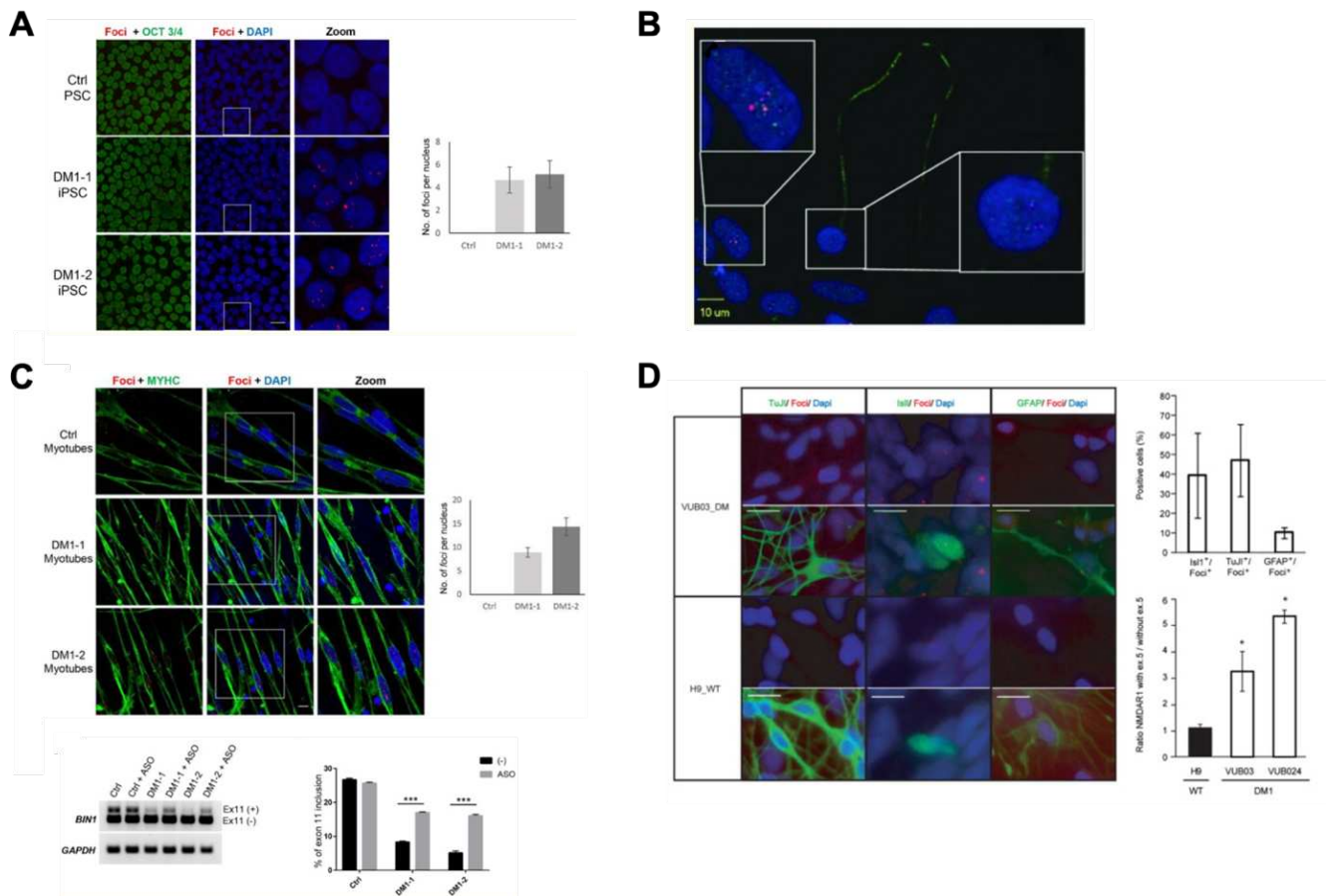


Figure 35 : Reproduction de phénotypes cellulaires et moléculaires associés à la DM1 à l'aide des CSP humaines et leurs dérivés

Présence de *foci* intranucléaires (en rouge) d'ARN DMPK mutants dans des hiPSC DM1 (**A**), des cellules neuronales (**B**), des myotubes dérivés d'hiPSC DM1 (**C**) et dans des neurones moteurs issus d'hESC DM1 (**D**). Reproduction du défaut d'inclusion de l'exon 11 de BIN1 dans des myotubes dérivés d'hiPSC DM1 (**C, panel du bas**) et du défaut d'exclusion de l'exon 5 de NMDAR1 dans des motoneurones issus d'hESC DM1 (**D, panel de droite**). Adapté de (Marteyn et al. 2011; Xia et al. 2013; Mondragon-Gonzalez and Perlingeiro 2018).

De par leur capacité de pluripotence, les CSP DM1 et leurs dérivés ont aussi contribué à la découverte de nouveaux mécanismes moléculaires. En effet, l'équipe dans laquelle j'ai effectué ma thèse a pu mettre en évidence au cours des dernières années différents défauts *in vitro* qui n'avaient encore jamais été décrits dans la pathologie (**Tableau 7**). En différenciant les CSP DM1 en précurseurs neuronaux, ils ont ainsi pu révéler une altération de la voie mTOR associée à des défauts de prolifération (Denis et al. 2013). Par ailleurs, ils ont également identifié des défauts d'expression génique avec par exemple une sous-expression du gène SLITRK4 engendrant une hyper-arborisation neuritiques dans des cultures motoneuronales (Marteyn et al. 2011).

Lignées cellulaires	Type de différenciation	Phénotype(s) étudié(s)	Référence
<u>hESCs</u> - VUB03 - VUB24	NSC (précurseurs neuronaux)	Altération de la voie mTOR → Augmentation de l'autophagie → Défaut de prolifération	Denis et al., 2013
	Motoneurons	Sous-expression du gène SLITRK4 → Hyper-arborisation neuritique → Altérations des jonctions neuromusculaires	Marteyn et al., 2011
	NSC & MSC (précurseurs neuronaux et mésenchymateux)	Sous-expression du gène ZNF37A → Défauts de réponse à l'IL13/IL4	Gauthier et al., 2013

Tableau 7 : Contribution des CSP humaines DM1 à l'identification de nouveaux mécanismes moléculaires

Parallèlement à cette recherche mécanistique, l'équipe a cherché à développer différentes approches pour évaluer la capacité d'utiliser ces CSP pour identifier des composés potentiellement thérapeutiques pour la DM1. Trois différentes stratégies ont ainsi été développées visant soit à diminuer le nombre de *foci* intranucléaires, soit à corriger certains défauts d'épissage caractéristiques de la DM1, ou encore de jouer sur la diminution de l'expression du gène *DMPK*. L'ensemble de ces stratégies ont été développées dans des progéniteurs mésenchymateux (MSC) issus de cellules hESC DM1 car ils reproduisent certains phénotypes moléculaires associés à la DM1 et qu'ils sont faciles à manipuler et amplifier.

La première approche a cherché à établir un criblage pharmacologique en imagerie automatisée pour identifier des composés capables de moduler le nombre de foci. Une chimiothèque composée de 12 264 molécules (issues de différentes banques chimiques : Prestwick, ChemiXinfinity et LOPAC), chacune utilisées à une seule et même concentration (5µM). Brièvement, l'algorithme développé au sein du laboratoire a permis de quantifier automatiquement le nombre de noyaux, le nombre de *foci* par noyaux ainsi que leur aire. Cette approche permet d'identifier une famille de composés, les glycosides cardiaques, qui présente la capacité de moduler le nombre et l'aire des *foci*. Parmi ces composés, l'équipe s'est particulièrement intéressée à l'un d'entre eux, la digoxine. Celle-ci est utilisée à ce jour dans le traitement de certaines insuffisances cardiaques. Au niveau cellulaire, elle inhibe les pompes Na⁺/K⁺, augmentant le taux d'ions sodium intracellulaires, et consécutivement, le taux calcique intracellulaire (Schoner and Scheiner-Bobis 2007). Cette augmentation améliore la contractilité des cellules cardiaques et entraîne un effet inotrope positif (augmentation de la force de contraction du muscle cardiaque). Parallèlement à son action sur les *foci* intranucléaires, la digoxine a aussi montré son efficacité sur la correction de certains défauts d'épissage associés à la DM1, toujours dans ces MSC dérivées d'hESC DM1 (**Figure 36-A**). L'équipe a ensuite montré par la suite que la digoxine pouvait également corriger ces défauts *in vivo*, en utilisant le modèle murin musculaire HSA^{LR} de la DM1. D'un

point de vue fonctionnel, la digoxine présente le potentiel de restaurer les défauts de différenciation myogénique observés sur des cellules myoblastiques primaires dérivées de patients DM1. Malgré ces résultats prometteurs, aucune étude clinique n'a été initiée avec la digoxine principalement en raison des effets secondaires qu'elle aurait pu engendrer chez les patients DM1 (Maury et al. 2019). En parallèle à cette approche à moyen débit, une deuxième approche cette fois-ci basée sur le repositionnement de molécules a été employée. L'équipe s'est tout particulièrement intéressée à la metformine, couramment prescrite dans le cadre de diabète de type II, et qui venait d'être montrée pour jouer sur certains épissages. L'équipe a ainsi montré que la metformine présentait la capacité de pouvoir normaliser un certain nombre de défauts d'épissage associés à la DM1 (**Figure 36-B**) (Laustriat et al. 2015). Comme évoqué précédemment dans la première partie de l'introduction, le repositionnement de la metformine a abouti à l'essai clinique MYOMET, dont la phase II a montré des résultats encourageants au niveau de la motricité des patients DM1 (Bassez et al. 2018).

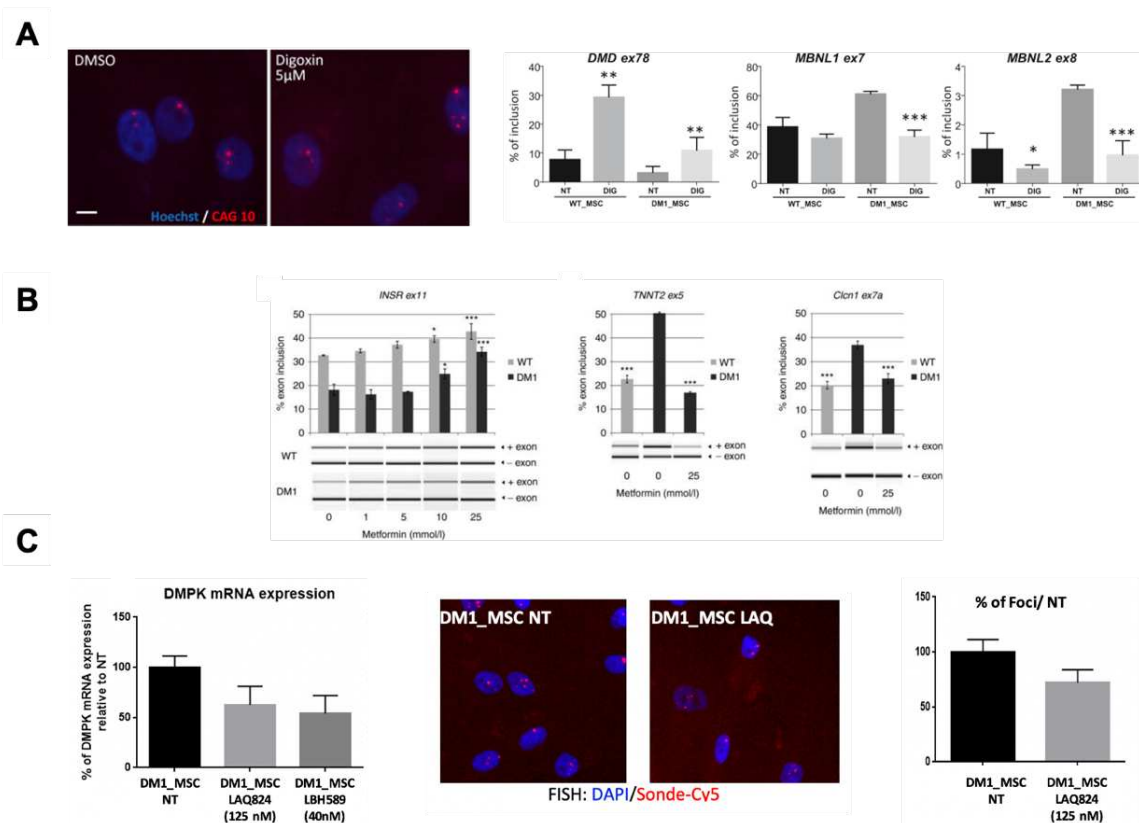


Figure 36 : Contribution des CSP humaines DM1 dans la recherche de composés potentiellement thérapeutiques

(A) La digoxine permet de diminuer le nombre et la taille des *foci* intranucléaires et corrige certains défauts d'épissage caractéristiques de la DM1. (B) La metformine permet elle aussi de normaliser de nombreux défauts d'épissage associés à la maladie. (C) Deux HDAC inhibiteurs, le LAQ824 et le LBH589 diminuent l'expression de l'ARNm de *DMPK* et sont aussi capables de diminuer le nombre de *foci*. Adapté de (Laustriat et al. 2015; Maury et al. 2019).

Plus récemment, la troisième approche, non encore publiée, a visé à identifier des composés capables de diminuer l'expression du gène *DMPK* dans les MSC dérivées de CSP DM1. La chimiothèque utilisée ici contenait 24 molécules connues pour modifier la structure de la chromatine. Ce criblage a permis d'identifier deux composés appartenant à la famille des inhibiteurs des histones désacétylases (HDACi) : le LAQ824 (ou Dacinostat) et le LBH589 (ou Panabinostat). Ces deux composés sont capables de diminuer significativement l'expression de l'ARNm *DMPK* et consécutivement de diminuer le nombre de *foci* par noyau et normaliser certains défauts d'épissage alternatif associés à la DM1 (**Figure 36-C**) (Lesueur et al., in preparation). Ces résultats sont en adéquation avec ceux obtenus dans une étude publiée en 2017 montrant l'effet bénéfique de deux autres HDACi, le Vorinostat et l'ISOX, sur certains défauts d'épissage dans des fibroblastes dérivés de patients DM1 en augmentant l'expression de *MBNL1* (Zhang et al. 2017a). Les mécanismes moléculaires par lesquels ces HDACi régulent l'expression du gène *DMPK* et/ou augmentent l'expression de la protéine *MBNL1* sont encore aujourd'hui peu connus.

Concernant les défauts myogéniques associés à la DM1, encore peu d'études utilisant les CSP humaines et leurs dérivés ont été réalisées. Ces études se sont principalement intéressées à la reproduction de défauts d'épissage ou encore des *foci* intranucléaires. De plus, ces dernières ont toutes utilisé des stratégies de surexpression pour la reprogrammation directe de CSP humaines en cellules myogéniques. En effet, une équipe japonaise a obtenu des cellules musculaires DM1 en forçant l'expression du facteur d'expression myogénique *MyoD* à partir d'hiPSC DM1 (Ueki et al. 2017). Plus récemment, une équipe américaine a elle-aussi utilisé la stratégie de surexpression de *MyoD* pour obtenir des cellules myogéniques DM1, non pas à partir d'hiPSC, mais à partir de cellules souches dérivées de l'urine directement (Kim et al. 2019). Enfin, des hiPSC DM1 ont également été différenciées en cellules musculaires en utilisant cette fois la surexpression du facteur *Pax7* (Mondragon-Gonzalez and Perlingeiro 2018). Comme mentionné précédemment, la surexpression de ces facteurs biaise néanmoins la chronobiologie de la myogenèse, pouvant ainsi affecter la modélisation pathologique. L'obtention de dérivés musculaires sans utilisation de transgène permettrait l'étude de la mise en place de défauts cellulaires et moléculaires tout au long de la myogenèse humaine, et aboutirait ainsi à une meilleure modélisation des atteintes myogéniques de la DM1 à partir de CSP humaines.

Depuis quelques années, un nouvel outil d'édition du génome a fait son apparition et est rapidement devenu incontournable dans les laboratoires du monde entier, notamment au niveau des maladies monogéniques, autant d'un point de vue thérapeutique que pour la modélisation pathologique de celles-ci : CRISPR/Cas9.

3.4. Les cellules souches pluripotentes et CRISPR/Cas9

3.4.1. CRISPR/Cas : origines et généralités

Avant cette récente découverte, il existait deux techniques majeures pour l'édition génomique : les ZFNs (Zinc Finger Nucleases) et les TALENs (Transcription activator-like effector nucleases). Ces deux techniques utilisent la nucléase Fok1 fusionnée à des domaines de liaison à l'ADN étant soit des doigts de zinc dans le cas des ZFNs, soit des effecteurs activateurs de la transcription (TAL) pour les TALENs. Dans les deux cas, ce sont ces domaines de liaisons à l'ADN qui vont jouer le rôle de localisateur génomique, pour ensuite induire une cassure double brin de l'ADN à un endroit précis via l'activité endonucléase de Fok1. C'est à la suite de cette cassure que la cellule va entreprendre la réparation de l'ADN sujette aux erreurs et pouvant alors entraîner des insertions ou des délétions (indels). Ces indels peuvent être responsables d'un décalage du cadre de lecture de la séquence codante aboutissant à éteindre le gène ciblé. Les TALENs sont capables de reconnaître un nucléotide précis alors que les ZFNs ne peuvent reconnaître que des trinuécléotides, ce qui complexifie de manière importante les interactions avec l'ADN et peut alors affecter les résultats escomptés d'édition du génome. Ainsi, les TALENs sont considérés comme plus efficaces et plus précis que les ZFNs (Yang and Huang 2019). Cependant, pour être suffisamment spécifique en reconnaissant une séquence composée au minimum de 18 nucléotides, il est nécessaire de concevoir 6 protéines dans le cas des ZFNs (Liu et al. 1997) et 18 effecteurs TAL pour les TALENs (Sanjana et al. 2012). Un tel encombrement affecte alors l'efficacité de ces méthodes d'édition génomique, en particulier dans les régions où la chromatine est inactive et peu accessible en raison de la méthylation de l'ADN et de la désacétylation des histones (Ben Jehuda et al. 2018).

C'est dans ce contexte que fut décrit en 2012 le système CRISPR/Cas9 et son potentiel dans l'édition du génome, par les équipes d'Emmanuelle Charpentier et de Jennifer Doudna (Jinek et al. 2012). Elles ont d'ailleurs reçu en 2020 le prix Nobel de Chimie pour récompenser la découverte de cet outil génétique révolutionnaire, plus communément appelé « ciseaux moléculaires ». Les systèmes CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) /Cas (CRISPR-associated) avaient été décrits, au préalable, comme systèmes de défense immunitaire adaptative des bactéries et des archées pour se défendre contre les bactériophages (Barrangou et al. 2007). Ils sont constitués de trois éléments :

- des petits ARN qui vont permettre la détection et l'élimination d'acides nucléiques étrangers (virus ou plasmides).
- des protéines endonucléases Cas codées par des gènes endogènes à la bactérie
- d'un locus CRISPR constitué de plusieurs séquences uniques capables de cibler spécifiquement le génome et séparées les unes des autres par des répétitions identiques.

Il a été démontré que ces systèmes immunitaires adaptatifs s'effectuent en trois étapes (**Figure 37**). Lors d'une primo-infection par un bactériophage, les bactéries et les archées vont intégrer de courts fragments de séquences étrangères (les protospacers) au niveau de leur locus CRISPR. Par la suite, ce locus CRISPR sera transcrit et donnera lieu à des ARN CRISPR précurseurs (pre-crRNA) qui seront ensuite clivés en crRNA, désormais capables de reconnaître les ADN cibles respectifs. La dernière étape consiste à la reconnaissance ainsi qu'à l'interférence des acides nucléiques étrangers. En effet, ces petits fragments de crRNA vont s'apparier avec les séquences protospacers complémentaires issues de l'ADN étranger et une fois la reconnaissance de la cible effectuée, les protéines Cas complexées avec les crRNA vont alors couper l'ADN cible, inhibant la réplication virale et procurant l'immunité aux cellules procaryotes. Les protéines Cas jouent ainsi le rôle d'effecteurs de l'immunité procaryotique (Jinek et al. 2012).

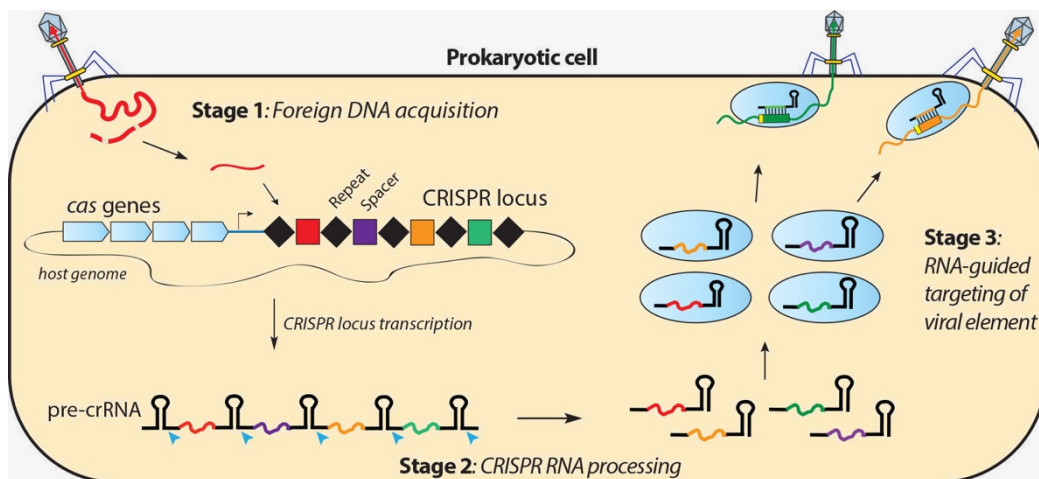


Figure 37 : Étapes de la réponse immunitaire adaptative médiée par les systèmes CRISPR/Cas chez les procaryotes

https://doudnalab.org/research_areas/crispr-systems/

Les systèmes CRISPR/Cas sont également définis par une grande diversité. En effet, ils sont classifiés selon deux grandes classes (classe 1 et classe 2), elles-mêmes regroupant plusieurs types. De manière générale, les types I, III et IV appartiennent à la classe 1 et la classe 2 est quant à elle composée des types II, V et VI (**Figure 38**) (Ishino et al. 2018). Les systèmes CRISPR/Cas composés de plusieurs sous-unités effectrices sont regroupés au sein de la classe 1 alors que les systèmes appartenant à la classe 2 ont pour effecteur une seule et grande protéine. La classification des différents types de systèmes CRISPR/Cas est très complexe et repose sur la conservation des gènes *cas* ainsi que sur l'organisation du locus. Jusqu'à présent, 13 gènes *cas* différents ont été décrits, illustrant bien la grande diversité de ces systèmes (Koonin and Makarova 2019). De plus, il a été montré qu'un même organisme

procaryote peut être composé de plusieurs de ces types, ce qui semble indiquer que ces systèmes seraient compatibles entre eux et partageraient des acteurs communs (Wiedenheft et al. 2012). Enfin, chaque type peut-être également composés de plusieurs sous-types (par exemple les sous-types II-A, II-B et II-C) selon des différences plus subtiles au niveau de l'organisation du locus et correspondant souvent à la présence de protéines Cas spécifiques (Koonin and Makarova 2019).

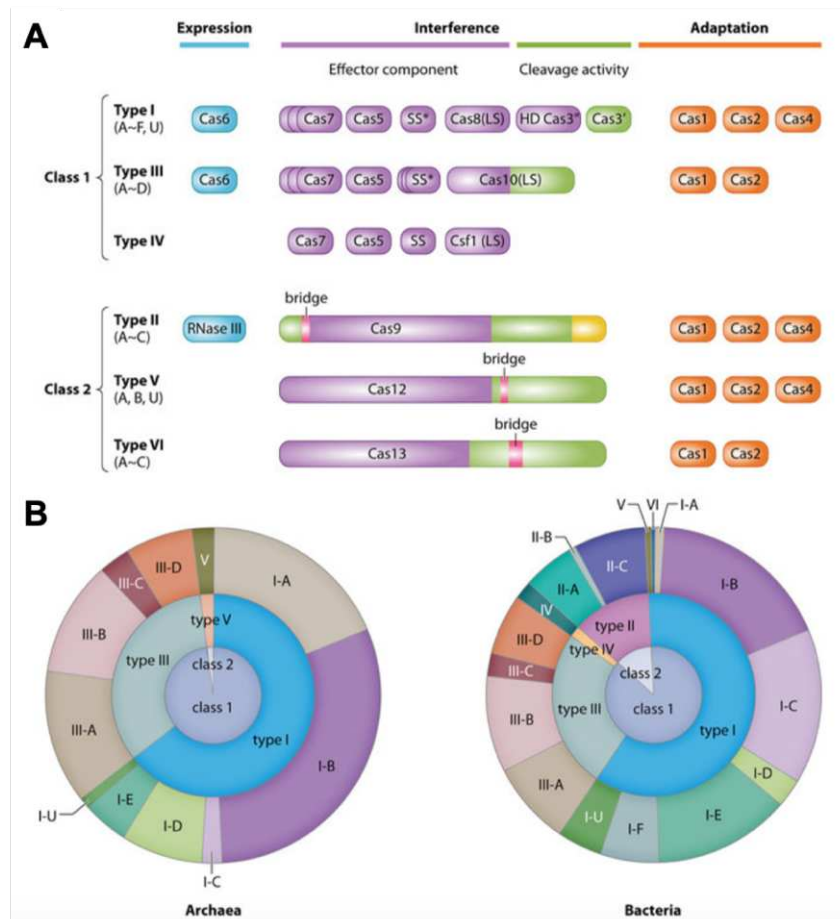


Figure 38 : Classification des systèmes CRISPR/Cas et répartition au sein des procaryotes

(A) Classification des systèmes CRISPR/Cas selon des analyses de séquences et selon l'organisation entre les différentes protéines Cas. Les protéines Cas peuvent être impliquées dans les trois grandes étapes de la réponse immunitaire adaptative. Les protéines Cas1 et Cas2 (et parfois Cas4) sont réquisitionnées lors de la phase d'adaptation pour processor les acides nucléiques étrangers en protospacers pour ensuite les intégrer dans le locus. (SS: small subunit, LS: large subunit). **(B)** Répartition des loci CRISPR/*cas* ayant été identifiés dans le génome des bactéries et des archées. Seules les loci encodant un système complet ont été répertoriés (Ishino et al. 2018).

3.4.2. CRISPR/Cas9 et édition du génome

En raison de l'organisation plus simple de leur locus CRISPR/*cas* et ne possédant qu'une seule grande protéine effectrice, les systèmes de la classe 2 se sont imposés comme les systèmes de choix pour développer une nouvelle génération de technologies d'édition génomique. Comme mentionné plus haut, ils sont composés des types II, V et VI ; le type II étant le grand représentant de la classe 2. Brièvement, les systèmes de type V sont principalement caractérisés par la Cas12 et sont constitués de deux sous-types majeurs : le sous-type V-A avec la Cas12a (également appelée Cpf1) et le sous-type V-B avec la Cas12b (ou C2c1). De la même manière, les types VI sont caractérisés par la Cas13, la seule protéine Cas ciblant l'ARN et non l'ADN, et comportent également deux sous-types principaux : VI-A avec la Cas13a (ou C2c2) et VI-B avec la Cas13b (ou C2c3) (Ishino et al. 2018). Bien que les systèmes de types V et VI présentent de belles perspectives, le type le plus communément étudié aujourd'hui reste le type II et sa protéine signature, la Cas9.

Plusieurs études ont décrit les systèmes de type II en trois sous-types selon l'organisation de l'opéron *cas*. Le sous-type II-A a un opéron composé de 4 gènes (*cas1*, *cas2*, *cas9* et *csn2*). Le rôle exact de la protéine Csn2 reste encore incertain mais elle serait impliquée dans l'intégration des protospacers au sein du locus CRISPR/*cas* (Barrangou et al. 2007). Le sous-type II-B possède lui aussi un opéron de 4 gènes (*cas1*, *cas2*, *cas9* et *cas4*) alors que l'opéron du sous-type II-C n'en possède que 3 (*cas1*, *cas2* et *cas9*). Le gène *cas9*, caractéristique des systèmes de type II, code une protéine à plusieurs domaines et capable d'effectuer à la fois le rôle d'effecteur et de clivage de l'ADN cible et intervient également lors du processus de maturation des crRNA. Ce dernier est catalysé par une enzyme bactérienne, la RNase III, aidée par un ARN antisens trans-activateur (le tracrRNA) encodé au sein du locus CRISPR/*cas*, et dont la séquence est complémentaire à celles des répétitions ARN séparant les futurs crRNAs (Wiedenheft et al. 2012). Le tracrRNA joue également un rôle important dans l'étape de reconnaissance de l'ADN cible, en formant un complexe avec le crRNA. C'est ce duo d'ARN tracrRNA:crRNA, aussi appelé ARN guide (gRNA), qui va alors guider la Cas9 pour qu'elle puisse exercer son activité endonucléase sur l'ADN cible (Jinek et al. 2012; Shmakov et al. 2017).

3.4.2.1. Mécanisme d'action

Les systèmes CRISPR/Cas9 ciblent des ADN double brin et requièrent un motif adjacent aux protospacers, appelé motif PAM. Celui-ci est situé immédiatement en aval de la séquence cible, sur le brin opposé à celui reconnu par le complexe tracrRNA:crRNA. Les motifs PAM sont composés de quelques paires de bases et leur séquence varie selon les systèmes CRISPR/Cas9 identifiés dans différentes espèces. Le système SpCas9 identifié chez

Streptococcus pyogenes fut le premier système CRISPR/Cas9 développé pour des expériences d'édition du génome et reste encore le plus largement utilisé (Cong et al. 2013; Mali et al. 2013). Ce système appartient au sous-type II-A et est constitué de la protéine SpCas9 d'une taille de 1368 acides aminés reconnaissant le motif PAM 5'-NGG-3' (Nishimasu et al. 2014). De manière générale, les protéines Cas9 sont constituées de deux lobes, un lobe de reconnaissance (REC) et un lobe à activité nucléase (NUC), reliés entre eux par un pont d'hélice alpha (Jiang and Doudna 2017; Ishino et al. 2018) (**Figure 39**). Il a été montré que lorsque la protéine Cas9 n'est pas liée au complexe tracrRNA:crRNA, elle adopte une conformation inactive. En revanche, lorsque le complexe ARN se lie à la Cas9, le domaine REC subit un changement conformationnel important permettant au complexe Cas9-gARN de processor l'ADN cible. La liaison avec l'ADN cible s'effectuera aussi grâce au domaine d'interaction avec le motif PAM (domaine PI), situé dans la région C-terminale du domaine NUC, qui va alors reconnaître cette séquence PAM et ainsi permettre l'hybridation du gRNA avec la séquence cible, immédiatement en amont du motif PAM. Cette hybridation va induire une séparation locale des deux brins d'ADN formant une boucle (appelée R-loop), entraînant un nouveau changement conformationnel de la Cas9 pour permettre le clivage des deux brins d'ADN environ 3 nucléotides en amont du motif PAM (Jiang and Doudna 2017) (**Figure 39**).

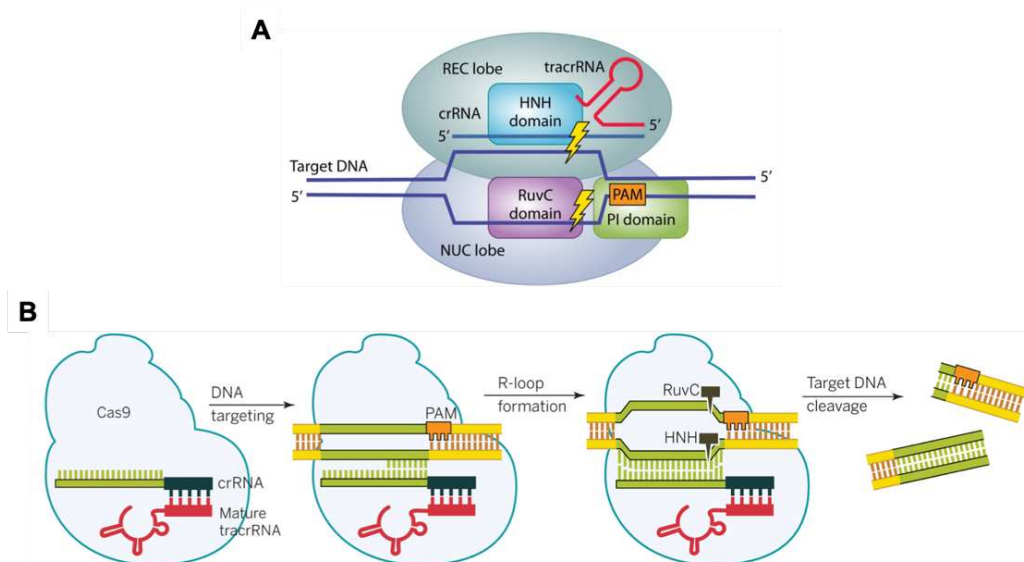


Figure 39 : Représentation schématique du système CRISPR/Cas9 de type II-A

(A) La protéine Cas9 est constituée de deux lobes REC et NUC. Le lobe REC est responsable de la reconnaissance de l'acide nucléique (target DNA) et le lobe NUC contient deux domaines, HNH et RuvC, à activité nucléase ainsi qu'un domaine d'interaction avec le motif PAM de l'ADN cible (domaine PI). (B) Le complexe tracrRNA:crRNA reconnaît l'ADN cible puis l'interaction du domaine PI avec la séquence PAM induit une séparation locale des deux brins de l'ADN en amont du PAM, qui seront ensuite clivés par les domaines à activité nucléase, entraînant ainsi une coupure double brin Adapté de (Ishino et al. 2018; Doudna and Charpentier 2014).

La coupure double brin est effectuée par deux domaines à activité nucléase situés dans le lobe NUC : le domaine HNH et le domaine RuvC. Le domaine HNH va couper le brin cible, c'est-à-dire celui hybridé avec le gRNA pendant que le domaine RuvC va quant à lui couper le brin opposé, sur lequel se trouve le motif PAM (Doudna and Charpentier 2014; Jiang and Doudna 2017).

3.4.2.2. Deux mécanismes de réparation de l'ADN double brin

Ainsi, pour un crRNA ciblant une séquence génomique donnée, le complexe Cas9-gRNA va être capable d'induire une coupure double brin à cet endroit précis, à la condition qu'un motif PAM soit adjacent à la séquence ciblée. Une fois cette coupure effectuée, la cellule va activer le processus de réparation de l'ADN. Plusieurs études ont montré que le processus de réparation de l'ADN double brin reposait sur deux mécanismes majeurs : la réparation par jonction d'extrémités non homologues (NHEJ) et la réparation par recombinaison homologue (HDR) (**Figure 40**).

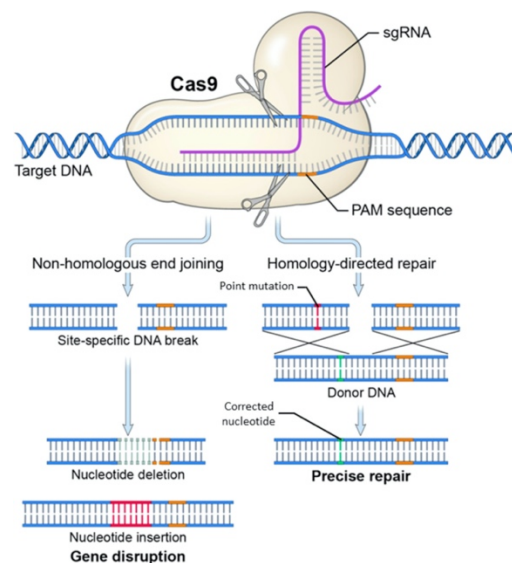


Figure 40 : Les mécanismes de réparation de l'ADN double brin

À la suite d'une cassure double brin médiée par CRISPR/Cas9, la cellule répare cette cassure par NHEJ ou HDR. Dans le cas de la NHEJ (à gauche) la réparation va être aléatoire et va induire des insertions ou des délétions pouvant entraîner la perte de fonction d'un gène. Dans le cas de la HDR (à droite), la réparation s'effectue en présence d'une séquence ADN homologue qui servira de matrice permettant une réparation précise et ciblée (Guitart et al. 2016).

La réparation par NHEJ implique une ligation de façon aléatoire des extrémités à bout franc de l'ADN au niveau de site de la cassure. Au cours de ce processus, des insertions ou

des délétions (indels) de nucléotides vont alors avoir lieu, pouvant entraîner plusieurs cas de figures. Soit les indels ne modifient pas le cadre de lecture et s'apparentent alors à des mutations de substitution. Celles-ci peuvent être silencieuses et n'auront pas d'effet. A l'inverse, elles peuvent modifier certains codons induisant soit un changement d'acides aminés (mutations faux-sens) soit l'apparition d'un codon STOP prématuré (mutation non-sens) et entraîner alors la perte de fonction du gène. Dans le cas où les indels ne correspondent pas à un multiple de trois, elles induisent un décalage du cadre de lecture aboutissant également dans la majorité des cas à l'apparition de codons STOP prématurés. Ce mécanisme est donc très intéressant pour supprimer l'expression d'un gène de façon définitive (Maruyama et al. 2015; Guitart et al. 2016).

Le deuxième mécanisme de réparation de l'ADN double brin s'effectue uniquement en présence d'une matrice, fournie par l'expérimentateur le plus souvent par transfection. Cette séquence d'ADN double brin synthétique comporte une forte homologie de séquence avec la région ciblée et va alors servir de modèle pour la réparation par recombinaison homologue. Cependant, en plus de n'être possible qu'en présence d'une matrice ADN, des études ont montré que la réparation par HDR possède une efficacité moins importante (entre 0,5 et 20% d'efficacité environ) que celle par NHEJ (de 20 à 60% environ). Ceci s'explique notamment en raison des protéines impliquées dans la recombinaison homologue qui sont exprimées uniquement lors des phases S et G2 du cycle cellulaire (Maruyama et al. 2015).

3.4.2.3. Limitations

L'une des limitations et préoccupations majeures de l'édition génomique médiée par CRISPR/Cas9 est le risque d'effets dits hors-cible ou "off-target". Bien que la spécificité de la Cas9 soit étroitement contrôlée par la séquence du gRNA de 20 nucléotides ainsi que par la présence d'un motif PAM adjacent à la séquence ciblée, une activité potentielle de clivage pourrait également se produire en dehors de cette séquence cible. En effet, contrairement aux nucléotides les plus proches du motif PAM, une hybridation parfaite entre le reste de la séquence du gRNA et la séquence cible n'est pas toujours nécessaire pour induire la coupure par la Cas9 (Zhang et al. 2015). La probabilité qu'une séquence ADN subisse un potentiel effet off-target va alors dépendre du nombre de mésappariements entre la séquence du gRNA et de l'ADN ainsi qu'en fonction de leur position relative au motif PAM. Ces coupures off-target peuvent alors avoir lieu bien qu'il y ait trois voire même cinq mésappariements entre l'ADN et la partie 5' (éloignée du motif PAM) du gRNA (Zhang et al. 2015). Dans le but de réduire ces effets off-target, plusieurs études ont réalisé une approche de mutagenèse dirigée pour obtenir un mutant de la protéine spCas9 plus spécifique. En ciblant des résidus jouant sur la structure de la protéine, l'équipe du Pr. Feng Zhang a pu générer l'eSpCas9

("enhanced specificity" SpCas9) et identifier plusieurs mutations capables de réduire les effets off-target tout en maintenant une activité ciblée robuste (Slaymaker et al. 2016). Parallèlement, une autre étude a identifié des mutations capables de réduire les contacts non spécifiques (principalement des liaisons Hydrogène) entre la protéine et l'ADN et a permis la génération de la SpCas9-HF1 (High-fidelity) dont les effets off-target sont considérablement atténués (Kleinstiver et al. 2016). Récemment, il a été démontré que ces deux mutants SpCas9 sont en fait bloqués dans un état inactif lorsqu'ils sont liés à des cibles mésappariées. En effet, cette étude montre le rôle d'un domaine non-catalytique, REC3, comme régulateur du domaine nucléase HNH de la Cas9. Cette découverte leur a ainsi permis de générer un nouveau variant, la HypaCas9 (Hyper-accurate), dont la spécificité est également considérablement augmentée par rapport à la SpCas9 (Chen et al. 2017).

Une autre limitation de la technologie CRISPR/Cas9 est la nécessité d'un motif PAM à proximité de la séquence ADN cible. La SpCas9 est l'une des Cas9 les plus utilisées car elle est capable de reconnaître un motif PAM relativement court : 5'-NGG-3' (où N est un nucléotide quelconque), contrairement à l'un de ses orthologues, la SaCas9 (*Staphylococcus aureus*) reconnaissant les séquences 5'-NNGRRT-3' ou 5'-NNGRR(N)-3' (où R est une purine quelconque) beaucoup plus contraignantes (Uddin et al. 2020). Pour pallier cette limitation et agrandir la fenêtre de ciblage des gènes, de nouveaux variants de la protéine SpCas9 ont été conçus par différentes équipes comme par exemple la SpCas9-NG capable de reconnaître le très court motif 5'-NG-3' (Nishimasu et al. 2018) ou encore la xCas9 capable quant à elle de reconnaître divers motifs PAM, y compris 5'-NG-3', 5'-GAA-3' et 5'-GAT-3' (Hu et al. 2018). Une récente étude comparative a montré que parmi ces deux variants, la SpCas9-NG aurait finalement une capacité de reconnaissance des motifs PAM plus importante que la xCas9, mais que cette dernière présenterait en revanche des effets off-target plus faibles (Kim et al. 2020).

3.4.2.4. Avancées et perspectives des systèmes CRISPR/Cas9

Combiné aux différents mécanismes de réparation de l'ADN, le système CRISPR/Cas9 peut être utilisé pour différentes stratégies d'édition du génome, lui conférant un large potentiel thérapeutique très intéressant. Associé à une réparation par NHEJ, CRISPR/Cas9 va être utilisé essentiellement pour des stratégies dites de knock-out (KO) conduisant à la perte d'expression d'un gène. À l'opposé, combiné CRISPR/Cas9 à une réparation par HDR va aboutir à des stratégies de knock-in (KI) qui vont permettre de corriger une mutation en la remplaçant par une séquence 'saine' ou au contraire d'induire une mutation précise pour étudier la maladie qui y est associée (**Figure 41-A**) (Wu et al. 2020).

En plus d'une utilisation plus simple comparée à celle des TALENs ou des ZFNs, CRISPR/Cas9 s'est rapidement imposé comme l'outil d'édition génomique par excellence aussi en raison de sa constante évolution, élargissant d'année en année ses champs d'application. En effet, de nombreuses études se sont intéressées aux améliorations possibles de ce système, notamment dans le but de réduire la fréquence des effets off-target. Tout d'abord, des mutations sur l'un des deux domaines à activité endonucléase ont permis de développer une nouvelle génération de Cas9 jouant désormais un rôle de nickases, n'induisant alors plus de cassure double brin mais uniquement une coupure simple brin (Jinek et al. 2012). La Cas9 D10A (substitution de l'aspartate en position 10 par une alanine au niveau du domaine RuvC) ne clive que le brin d'ADN cible (via le domaine HNH) alors que la Cas9 H840A (substitution d'une histidine par une alanine au niveau du domaine HNH) clive à l'inverse le brin opposé (via le domaine RuvC) (Cong et al. 2013; Jinek et al. 2012). Le principal intérêt de générer de telles Cas9 réside sur le principe qu'une simple coupure sur un brin n'activera pas la réparation par NHEJ et favorisera la réparation par HDR au détriment cependant de l'efficacité (Cong et al. 2013). En combinant deux Cas9 nickases avec deux gRNA appariés ciblant chacun un brin de l'ADN cible, il est alors possible d'induire une coupure double brin qui ne favorisera pas la réparation par HDR mais qui augmentera significativement la spécificité du système en réduisant les effets off-targets (**Figure 41-B**) (Doudna and Charpentier 2014; Ran et al. 2013).

Une Cas9 dépourvue de toute fonction catalytique, aussi appelée dCas9 (pour dead Cas9) a aussi pu être développée par l'introduction de mutations spécifiques au niveau des deux domaines à activité nucléase. Cette dCas9, fusionnée à différents effecteurs, peut servir de guide pour de multiples applications (**Figure 41-C**). Fusionnée avec des domaines activateurs ou répresseurs de la transcription, elle va pouvoir permettre d'activer ou de réprimer l'expression de gènes (Wu et al. 2020; Gilbert et al. 2014). Il est également possible d'agir au niveau de l'épigénome en fusionnant la dCas9 à des effecteurs permettant de moduler la chromatine, en ciblant par exemple la modification des histones ou l'état de méthylation de l'ADN (Pulecio et al. 2017). Enfin, la dCas9 a aussi montré un autre type d'application en étant fusionnée avec des protéines de fluorescence comme la GFP pour marquer des régions spécifiques de l'ADN. Ce système peut alors permettre de suivre la dynamique des changements conformationnels des chromosomes ou encore l'organisation nucléaire de manière plus générale (Chen et al. 2013).

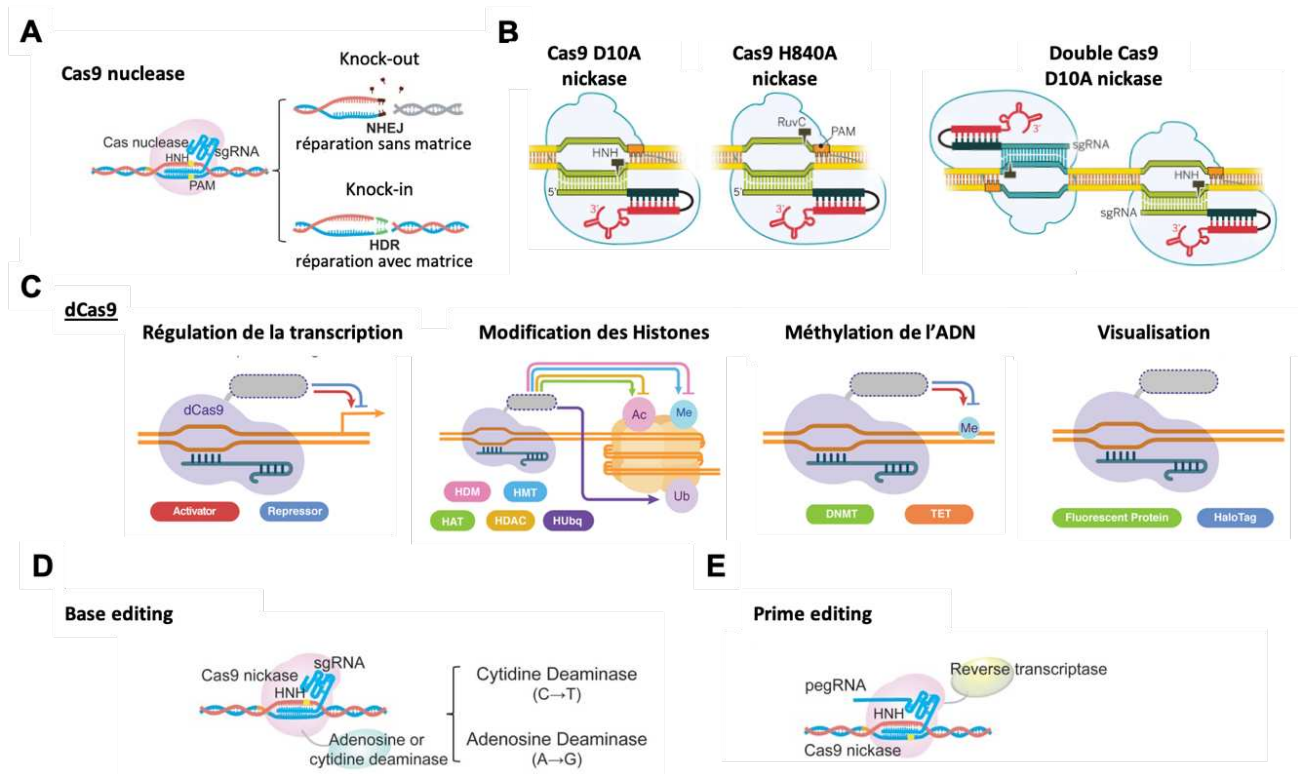


Figure 41 : Vue d'ensemble des avancées technologiques du système CRISPR/Cas9

(A) Le système CRISPR/Cas9 classique induisant une cassure double brin suivie par la réparation de l'ADN via la NHEJ ou l'HDR. (B) La Cas9 peut agir sous la forme d'une nickase lorsque l'un de ses domaines RuvC ou HNH est muté. Quand la Cas9 nickase est utilisée avec deux gRNA appariés, ciblant chacun un brin de l'ADN cible, elle va alors pouvoir induire une cassure double brin. (C) La dCas9 permet une manipulation fine de l'épigénome à différents niveaux selon qu'elle est fusionnée à des régulateurs transcriptionnels, des modificateurs de la chromatine ou encore des molécules fluorescentes. (D) Le base editing consiste en la fusion d'une Cas9 D10A nickase avec une cytidine ou une adenosine déaminase afin de substituer précisément un C par un T ou un A par un G respectivement. (E) Le prime editing est la fusion d'une Cas9 H840 nickase avec une reverse transcriptase qui va utiliser un guide ARN de prime editing (pegRNA) pour éditer directement le génome à l'endroit de la coupure. HDM: histone déméthylase, HMT: histone méthyltransférase, HAT: histone acétyltransférase, HDAC: histone désacétylase, HUBq: histone ubiquitine ligase, DNMT: DNA méthyltransférase, TET: ten-eleven translocation enzymes, Ac: acétylation, Me: méthylation, Ub: ubiquitination. *Adapté de (Wu et al. 2020; Pulecio et al. 2017; Doudna and Charpentier 2014).*

Plus récemment, une nouvelle génération de Cas9 a vu le jour : les éditeurs de base ou base editing (**Figure 41-D**). Il s'agit de la fusion d'une Cas9 D10A nickase avec une cytidine déaminase ou une adénosine déaminase, qui vont générer des mutations précises et ciblées, sans induire de cassure double brin. De manière générale, il a été montré que la coupure simple brin induite par l'activité nickase de la Cas9 D10A déclenche la réparation de ce brin en utilisant comme modèle le brin complémentaire qui aura été lui édité par la déaminase. Plusieurs types d'éditeurs de base ont été développés distinguant les éditeurs de base cytosine aux éditeurs de base adénine. Les premiers sont capables de convertir une paire de

basse C-G par la paire de bases T-A, alors que les deuxièmes convertissent une paire de bases A-T par une paire de bases G-C (Wu et al. 2020; Rees and Liu 2018).

Enfin, l'équipe du Pr. David Liu a développé en 2019 le prime editing, un nouveau système différent du base editing car ce dernier consiste en la fusion d'une Cas9 H840A nickase avec une reverse transcriptase (**Figure 41-E**). Cette nouvelle méthode d'édition du génome est plus polyvalente car le gRNA est ici remplacé par un guide ARN de prime editing (ou pegRNA) qui va, en plus de spécifier la séquence ADN cible, encoder l'édition souhaitée qui sera directement effectuée par la reverse transcriptase. Ainsi, il est théoriquement possible d'effectuer n'importe quelle conversion de base (12 possibles au total) aussi bien que des insertions ou des délétions, sans avoir aucunement recours à une cassure double brin ni à une matrice ADN. Ces deux dernières méthodes d'édition génomique sont encore très récentes et en constante amélioration (Wu et al. 2020; Anzalone et al. 2020).

Toutes ces différentes approches et avancées technologiques des systèmes CRISPR/Cas9 représentent ainsi des perspectives d'innovations très importantes pour l'étude et le traitement des maladies génétiques. En particulier, la combinaison des technologies CRISPR et hiPSC, toutes deux récompensées par un prix Nobel, a totalement changé notre vision à la fois au point de vue des stratégies thérapeutiques comme la thérapie génique et la thérapie cellulaire, mais également au niveau de la modélisation pathologique et du criblage de molécules, pour la découverte de nouveaux médicaments et dans le but de mieux comprendre ces pathologies (De Masi et al. 2020).

3.4.3. CRISPR/Cas9 combiné aux iPSC humaines comme outil thérapeutique pour les maladies génétiques

Comme illustré par le nombre exponentiel de publications combinant les approches CRISPR et hiPSC, la fusion de ces deux technologies a changé fondamentalement la recherche biomédicale ces dernières années (**Figure 42**). Dans le cadre des maladies monogéniques, la technologie CRISPR/Cas9 est de plus en plus utilisée pour générer des lignées isogéniques, en corrigeant ou en insérant la mutation causale de la maladie à partir d'iPSC humaines malades ou saines respectivement. L'obtention de ces lignées isogéniques est devenue primordiale pour les programmes de modélisation pathologique et est très fortement recommandée par les directives de l'ISSCR (International Society for Stem cell research) (**Figure 43**).

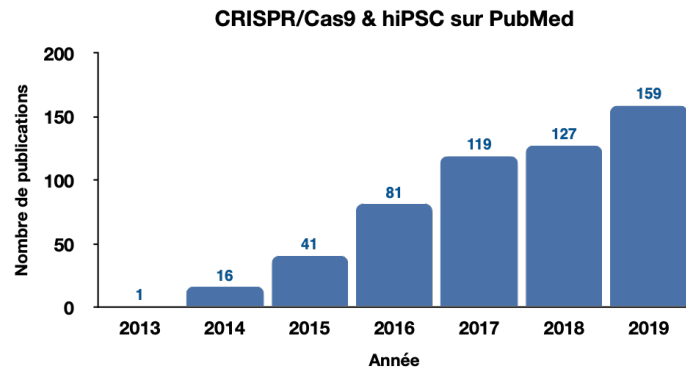


Figure 42 : Publications relatives à l'utilisation de CRISPR/Cas9 et des iPSC humaines
D'après PubMed, avec les mots-clés [CRISPR/Cas9] AND [hiPSC], en novembre 2020.

Les iPSC humaines et CRISPR/Cas9 sont deux technologies qui ont changé fondamentalement la recherche biomédicale ces dernières années, notamment au niveau de la médecine translationnelle. Supprimer l'expression d'un gène, ou knock-out, dans des lignées iPSC permet de pouvoir étudier la fonction de ce gène et de mieux comprendre son implication dans une pathologie donnée. A l'opposé, les stratégies de knock-in sont principalement utilisées soit pour corriger un gène ou au contraire pour induire une mutation précise dans ces lignées iPSC humaines. De plus, les bibliothèques de gRNA commencent à remplacer peu à peu les bibliothèques d'ARN interférence pour mettre en place des criblages génomiques à haut débit (Zhang et al. 2017b). En effet, l'utilisation des iPSC humaines est de plus en plus fréquemment associée à celle de CRISPR/Cas9. La technologie CRISPR/Cas9 peut également être utilisée pour mieux contrôler les étapes de spécification cellulaire, via les systèmes CRISPRa ou CRISPRi décrits plus haut, en activant ou inhibant certains facteurs de transcription régulant des voies de signalisation déterminantes pour l'acquisition d'une nouvelle identité cellulaire (Hsu et al. 2019).

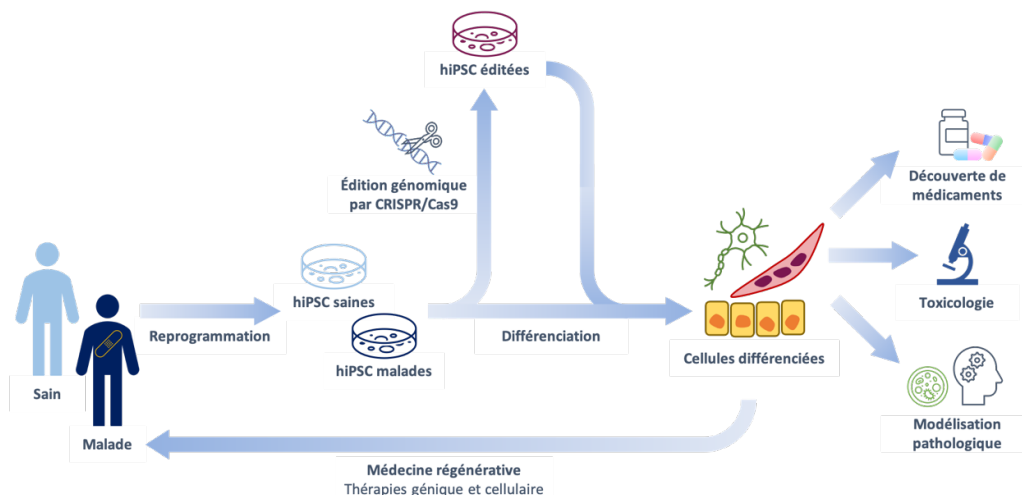


Figure 43 : Représentation du déroulement de la recherche impliquant les iPSC humaines et CRISPR/Cas9 pour de nouvelles alternatives thérapeutiques

Au final, la correction de mutations par la technologie CRISPR/Cas9 ouvre le champ des possibles dans les domaines de la thérapie génique et de la médecine régénérative (De Masi et al. 2020).

3.4.3.1. Thérapie génique et thérapie cellulaire

On distingue deux grandes sortes de thérapies géniques. La première est la thérapie génique ciblant la lignée germinale, c'est-à-dire l'édition du génome dans des embryons humains afin de corriger la mutation génétique directement à la source, avant le développement du fœtus et avant que la maladie ne puisse se mettre en place. De nombreux débats éthiques se tiennent aujourd'hui pour permettre ou non (et si oui dans quelles mesures) cette thérapie génique humaine. Certains y voient une utopie encore lointaine principalement pour des raisons techniques avec le trop grand risque d'effets indésirables, notamment concernant les potentiels effets off-targets générés par CRISPR/Cas9. Néanmoins, au vu des innovations et améliorations incessantes des systèmes CRISPR/Cas9, ce n'est qu'une question de temps pour que ceux-ci deviennent de plus en plus fiables, efficaces et spécifiques. Le plus important du débat réside davantage sur les futures réglementations de la thérapie génique, et il soulève d'importantes questions éthiques : voulons-nous modifier de manière irréversible la lignée germinale humaine ? C'est à ce genre de questions que sont confrontés actuellement les comités éthiques compétents dans le but de pouvoir autoriser peut-être un jour l'application clinique de cette technologie à des fins médicales (Hockemeyer and Jaenisch 2016).

L'autre grand axe de la thérapie génique consiste cette fois à cibler non pas la lignée germinale, mais les cellules somatiques. L'édition génétique de cellules somatiques, comme celle des hiPSC, est en constant progrès et commence à s'imposer comme une possibilité thérapeutique prometteuse (Hockemeyer and Jaenisch 2016). Dans ce cas de figure, deux méthodes se distinguent. La première consiste en ce qui est appelé la thérapie génique *ex vivo*, ou encore la thérapie génique pour la thérapie cellulaire. Elle appartient au grand domaine de la médecine régénérative (**Figure 43**). Cela consiste à modifier génétiquement les iPSC préalablement reprogrammées à partir de cellules du patient. Une fois la mutation causale de la pathologie corrigée par CRISPR/Cas9 *in vitro*, il faut différencier ces cellules éditées dans le type cellulaire d'intérêt pour pouvoir par la suite transplanter le patient avec ses propres cellules, désormais saines, qui vont alors prendre le relais sur les cellules malades. On parle de thérapie cellulaire autologue (Jacinto et al. 2020). Cette approche reste cependant encore compliquée à mettre en œuvre en particulier dans le cadre des maladies génétiques musculaires majoritairement en raison des problèmes de survie cellulaire et d'incompatibilité avec une délivrance systémique (Raaijmakers et al. 2019). De plus, l'obtention de cellules permettant la régénération musculaire comme les cellules satellites

ou les péricytes à partir d'iPSC humaines reste encore à être mieux définie (Dellavalle et al. 2011; André et al. 2018). La deuxième méthode consiste à injecter directement la Cas9 et son gRNA *in vivo*, afin qu'elle puisse corriger la mutation dans les cellules cibles. Par exemple, une étude publiée en 2018 a montré l'efficacité de l'édition génomique *in vivo* du gène *DMD* par CRISPR/Cas9 dans un modèle canin de la dystrophie musculaire de Duchenne. Ces chiens portent une mutation qui induit la perte de l'exon 50 de la protéine DMD aboutissant à l'apparition d'un codon STOP prématuré et à la perte de fonction de la protéine. Ils ont utilisé un virus adéno-associé (AAV) pour délivrer le système CRISPR/Cas9 soit par injection intramusculaire soit de manière systémique (deux chiens injectés dans chacune des conditions). L'objectif de cette édition du génome était dans un premier temps de restaurer l'expression de la protéine DMD dans les cellules musculaires en ciblant l'exon 51 afin de rétablir le cadre de lecture. Cette restauration fut très variable selon le type de muscle mais ils ont pu observer des résultats très encourageants à la suite des injections systémiques, notamment dans le muscle cardiaque, en restaurant l'expression de la protéine DMD à plus de 90% comparé à un chien non atteint. De plus, cette restauration fut accompagnée d'une amélioration musculaire au niveau histologique (Amoasii et al. 2018).

Concernant la DM1, les stratégies thérapeutiques sont multiples et certaines ont pu bénéficier de l'apport innovant des systèmes CRISPR/Cas9 (**Figure 44**). La première stratégie n'est autre que la correction de la mutation à proprement parler, c'est-à-dire l'élimination des répétitions CTG au niveau de la région 3'UTR du gène *DMPK*. Une telle excision est possible en induisant une cassure double brin de part et d'autre des répétitions à l'aide de deux gRNA (**Figure 44-A**). De nombreuses équipes ont suivi cette stratégie dans différents modèles de la pathologie : dans des myoblastes DM1 primaires et immortalisés (Dastidar et al. 2018; van Agtmaal et al. 2017), des fibroblastes DM1 immortalisés MyoD-inductibles (Provenzano et al. 2017), des hESC et hiPSC DM1 (Yanovsky-Dagan et al. 2019; Dastidar et al. 2018) ainsi que dans des cellules myogéniques ou neuronales (NSC) dérivées d'hiPSC DM1 (Dastidar et al. 2018; Wang et al. 2018a). La grande majorité de ces études a montré que l'excision des répétitions permettait de diminuer le nombre de *foci* intranucléaires, de corriger certains défauts d'épissage ainsi que d'améliorer la différenciation myogénique (Raaijmakers et al. 2019). Plus récemment, l'équipe du Dr Ana Buj-Bello à Généthon a également réalisé cette édition génomique des répétitions CTG dans des myoblastes DM1 immortalisés avant d'établir la première preuve de concept de la faisabilité de cette approche *in vivo*, dans un modèle murin DM1, les souris DMSXL, en injectant des vecteurs AAV exprimant le système CRISPR/Cas9 au niveau du muscle tibial antérieur. Ils ont ainsi pu observer une diminution des *foci* dans les muscles ayant été injectés (Scrudato et al. 2019).

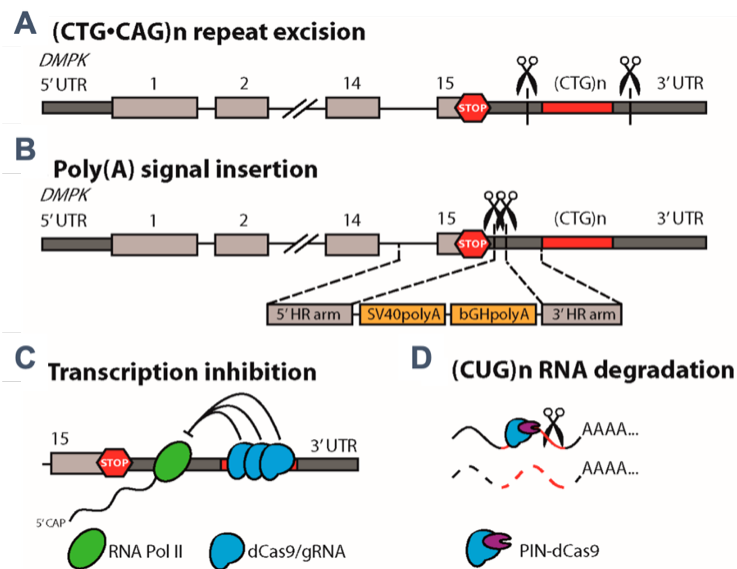


Figure 44 : Stratégies CRISPR/Cas9 ayant été utilisées dans des modèles DM1

(A) Les expansions anormales des répétitions des triplets CTG ont été excisées du gène *DMPK* par une double coupure double brin, en amont et en aval des répétitions. La transcription des répétitions CTG a été empêchée soit par (B) l'insertion d'un signal poly(A) prématuré entre le codon STOP du gène *DMPK* et les répétitions, soit par (C) le recrutement de dCas9 au niveau des répétitions en bloquant physiquement la progression de l'ARN polymérase II. (D) Les répétitions CUG des transcrits *DMPK* ont été dégradées via une dCas9 fusionnée au domaine PIN à activité ribonucléase (Raaijmakers et al., 2019).

La deuxième stratégie CRISPR/Cas9 utilisée est d'inhiber la transcription des répétitions. Il existe plusieurs moyens pour y parvenir. L'un d'eux est d'introduire, par une stratégie de knock-in, un signal de polyadénylation (polyA) entre le codon STOP du gène *DMPK* et les répétitions CTG (Figure 44-B). Pour augmenter la spécificité de l'insertion ciblée, l'équipe ayant réalisé cette stratégie de knock-in a utilisé le système Cas9 D10A nickase couplé à deux gRNA dans des hiPSC DM1 et des NSC dérivées de ces hiPSC DM1 (Wang et al. 2018a).

Une autre équipe a utilisé les systèmes dCas9 pour qu'ils ciblent les répétitions CTG afin de bloquer physiquement, par encombrement allostérique, la progression de l'ARN polymérase II, inhibant ainsi la transcription des répétitions (Figure 44-C). En plus d'avoir réalisé cette stratégie dans des myoblastes primaires DM1, cette équipe a également obtenu des résultats *in vivo* encourageants avec le modèle murin HSA^{LR} de la DM1, en observant notamment une diminution de la myotonie chez ces souris (Pinto et al. 2017). Enfin, les systèmes dCas9 ont également été utilisés pour cibler les répétitions au niveau de l'ARNm *DMPK* de myoblastes primaires DM1, dans le but de dégrader ces transcrits en fusionnant la dCas9 avec le domaine PIN à activité ribonucléase du gène *SMG6* (Figure 44-D) (Batra et al. 2017). Employer la stratégie d'excision des répétitions CTG médiée par CRISPR/Cas9 dans des ESC ou des iPSC DM1 humaines a ainsi permis d'obtenir les premières lignées isogéniques de cette pathologie (Dastidar et al. 2018; Yanovsky-Dagan et al. 2019).

3.4.3.2. Modélisation pathologique

L'utilisation de la technologie CRISPR pour générer des lignées isogéniques est devenue un élément important dans les approches de modélisation pathologique. L'utilisation de ces lignées permet en effet de pallier les variations du fond génétique ou encore d'autres variables comme le sexe et l'âge (De Masi et al. 2020). C'est pour cette raison que les systèmes CRISPR/Cas9 ne cessent de s'améliorer d'année en année au niveau de leur spécificité d'action, afin de réduire au maximum le risque d'effets off-target (Bassett 2017) et de permettre de corriger ou d'introduire une mutation au niveau d'un seul nucléotide à l'aide des systèmes de base et prime editing (Wu et al. 2020).

Dans le cadre de la DM1, l'obtention d'iPSC humaines déplétées des répétitions CTG a permis de générer un véritable contrôle pour la comparaison de phénotypes observés chez les cellules non-éditées malades (Dastidar et al. 2018). À l'inverse, la mise au point d'une stratégie CRISPR/Cas9 permettant cette fois d'insérer des répétitions CTG dans des iPSC humaines saines dans le but de modéliser la DM1 n'a pas encore été décrite à ce jour. Cela pourrait s'expliquer en raison de la grande instabilité de ces répétitions qui risquerait de compliquer cette approche (Tomé and Gourdon 2020). Seule une récente étude sur des souris a démontré la faisabilité de cette stratégie en insérant à l'aide de CRISPR/Cas9 des expansions de triplets CTG dans la partie 3'UTR du gène *Dmpk* de ces dernières (Nutter et al. 2019).

Outre la génération de lignées isogéniques, la technologie CRISPR combinée aux hiPSC peut également permettre de générer des nouveaux outils cellulaires permettant de mieux comprendre le rôle de certains facteurs « pathologiques » dans la spécification cellulaire. C'est dans ce cadre que se place l'étude que j'ai menée pendant mon doctorat visant à développer de nouveaux modèles pour mieux comprendre le rôle développemental des protéines MBNL dans un contexte humain.

OBJECTIFS ET DÉMARCHE EXPÉRIMENTALE

La DM1 est une maladie neuromusculaire héréditaire dont un des mécanismes physiopathologiques le plus reconnu consiste en la perte de fonction des protéines MBNL régulatrices de l'épissage alternatif. Le processus d'épissage alternatif est un mécanisme fondamental non seulement pour la diversification des isoformes des protéines mais aussi pour la régulation spatio-temporelle du développement. Par conséquent, une meilleure compréhension de la régulation de ce processus permettrait de déchiffrer les mécanismes pathologiques impliqués dans la DM1. Malgré les nombreuses études réalisées dans des modèles murins visant à démontrer la contribution des protéines MBNL dans la DM1, peu de données existent sur leur rôle chez l'Homme.

En s'appuyant sur la technologie d'édition génomique CRISPR/Cas9, l'objectif général de ma thèse a consisté à évaluer la possibilité d'utiliser les cellules souches pluripotentes humaines pour mieux comprendre la fonction des protéines MBNL en particulier dans la spécification myogénique, et ce dans le but de mieux comprendre leur implication dans la physiopathologie humaine de la DM1. En nous inspirant des études réalisées sur la génération des modèles murins *Mbnl* knock-out (KO) par les équipes du Pr. Maurice Swanson et du Dr. Charles Thornton (Kanadia et al. 2003; Charizanis et al. 2012; Lee et al. 2013; Thomas et al. 2017), notre démarche expérimentale fut de générer plusieurs lignées hiPSC différentiellement déplétées en protéines MBNL en ciblant les différents paralogues avec pour but principal d'obtenir une lignée hiPSC *MBNL1* KO, une lignée *MBNL2* KO ainsi qu'une lignée *MBNL1,2* DKO.

Cette thèse s'est également inscrite dans un contexte où les outils cellulaires musculaires humains pour l'étude de la DM1 présentent certaines limites. Ces derniers reposent en effet majoritairement sur l'utilisation i) de cellules musculaires primaires issues de biopsies de patients mais dont la prolifération *in vitro* reste limitée ; ii) de cellules musculaires primaires immortalisées et dont le processus d'immortalisation aboutit à la modification du cycle cellulaire de ces cellules (Arandel et al. 2017) ; iii) de cellules myogéniques différenciées à partir de cellules souches pluripotentes DM1 par la surexpression de facteurs de transcription myogéniques comme MyoD ou Pax7 (Ueki et al. 2017; Mondragon-Gonzalez and Perlingeiro 2018; Kim et al. 2019).

Ces approches de reprogrammation directe présentent l'inconvénient de contourner les premières étapes développementales de la myogenèse, étapes qui pourraient être primordiales dans la mise en place des mécanismes physiopathologiques de maladies génétiques neuromusculaires telles que la DM1 (Chal and Pourquoié 2017).

Ainsi, l'objectif de cette thèse a également été de mettre au point un protocole de différenciation myogénique au laboratoire pour l'étude de la DM1 en utilisant uniquement des petites molécules et des cytokines. L'utilisation d'un tel protocole combiné avec l'obtention des lignées hiPSC déficientes en protéines MBNL obtenues par CRISPR/Cas9, permet également de répondre à la question de la contribution des protéines MBNL dans la

spécification myogénique humaine en évaluant les conséquences moléculaires et cellulaires de cette perte de fonction sur la génération de cellules musculaires squelettiques. Pour cela, un des objectifs de ma thèse fut de réaliser des analyses transcriptomiques sur ces cellules musculaires squelettiques obtenues à partir des hiPSC DM1 ainsi que des hiPSC *MBNL* KO générées dans le but d'identifier de nouveaux mécanismes moléculaires physiopathologiques dépendants ou non des protéines MBNL. Cette dernière étude nous a ainsi permis d'identifier un nouveau défaut d'épissage alternatif, régulé par les protéines MBNL1 et MBNL2, présent dans des cellules de patients DM1 et qui semble affecter la spécification post-synaptique.

En parallèle de ces objectifs principaux, j'ai également pu collaborer avec une autre étudiante en thèse de l'équipe, Julie Tahraoui afin d'étudier notamment le rôle des protéines MBNL dans la spécification d'un autre type cellulaire : les motoneurones spinaux.

La DM1 a pendant longtemps été décrite comme une maladie purement musculaire mais plusieurs études ont décrit la présence de défauts de transmission synaptiques chez les patients (Bombelli et al. 2016) ou encore l'observation de foci au niveau des noyaux des motoneurones (Wheeler et al. 2007a). Ainsi, l'objectif de cette étude est de définir si la mutation DM1, et en particulier la perte de fonction des protéines MBNL, pourrait affecter la communication entre les motoneurones spinaux et leur cible musculaire.

TRAVAUX DE RECHERCHE

Article 1 : La déplétion des protéines MBNL dans les cellules souches pluripotentes révèle une transition d'épissage anormale de SORBS1 affectant le développement neuromusculaire dans la DM1

Loss of MBNL in pluripotent stem cells reveals abnormal splicing switch of SORBS1 compromising the neuromuscular development in Myotonic Dystrophy type 1

Antoine Mérien¹, Julie Tahraoui-Bories¹, Gilles Moulay⁴, Michel Cailleret¹, Jean-Baptiste Dupont¹, Céline Leteur², Jérôme Polentes², Alexandre Carteron², Hélène Polvèche², Jean-Paul Concordet³, Christian Pinset², Margot Jarrige², Stéphane Vassilopoulos⁴, Denis Furling⁴, Cécile Martinat^{1*}

¹ INSERM / UEPS UMR 861, Paris Saclay University, I-STEM, 91100 Corbeil-Essonnes, France.

² CECS/AFM, I-STEM, 91100 Corbeil-Essonnes, France.

³ INSERM U1154 / CNRS UMR7196, MNHN TACGENE, 75005, Paris, France.

⁴ Sorbonne Université, INSERM, Association Institut de Myologie, Centre de recherche en myologie, Paris, France.

* **Corresponding author** : C. Martinat, PhD, I-Stem, 28 rue Henri Desbruères 91100 Corbeil-Essonnes France, Tel: + 00 33 1 69 90 85 33, Fax: + 00 33 1 69 90 85 21, E-mail: cmartinat@istem.fr

L'objectif principal de cette étude a consisté à développer un nouveau modèle, basé sur l'utilisation des cellules souches humaines induites à la pluripotence (hiPSC), pour mieux comprendre à la fois la fonction des protéines MBNL dans la spécification myogénique humaine mais également leur contribution dans la physiopathologie de la DM1. Nous avons utilisé la technologie CRISPR/Cas9 dans le but de générer des hiPSC déplétées en protéines MBNL afin d'évaluer l'impact de la perte d'expression de ces protéines au niveau cellulaire et moléculaire au cours des différentes étapes développementales de la myogenèse. À l'aide d'une approche transcriptomique comparative, nous avons pu identifier le défaut d'inclusion de l'exon 25 de *SORBS1* dans des myotubes dérivés d'hiPSC DM1 et d'hiPSC *MBNL1,2* DKO ainsi que dans des échantillons musculaires de patients DM1. Enfin, par des expériences de saut d'exon médiées par des oligonucléotides anti-sens, nous avons démontré dans cette étude que ce défaut d'épissage engendrait des anomalies structurales du compartiment musculaire post-synaptique.

Génération des lignées hiPSCs MBNL KO(s) par CRISPR/Cas9

La première étape de ma thèse fut d'obtenir des lignées hiPSC *MBNL1* knock-out (KO) et hiPSC *MBNL2* KO à l'aide de la technologie CRISPR/Cas9.

Nous avons conçu des ARN guides (gRNA) ciblant spécifiquement le premier exon codant non-alternatif de chacun des gènes *MBNL1* et *MBNL2* afin d'induire l'apparition de codons STOP très tôt dans la séquence d'ADN et s'appliquant ainsi à toutes les isoformes protéiques de ces deux gènes. Afin d'identifier les clones hiPSC KO, nous avons effectué cette sélection

en deux temps : la première étape fut d'analyser par la méthode de RFLP (restriction fragment length polymorphism) la perte totale ou partielle d'un site de restriction enzymatique situé au niveau du PAM, nous indiquant ainsi si l'édition génomique a eu lieu ou non. Cette étape offre les avantages d'une sélection rapide et peu coûteuse en raison d'une simple réaction de digestion enzymatique de l'ADN génomique des clones obtenus. La deuxième étape de sélection a consisté à séquencer par la méthode de Sanger les clones sélectionnés dans le but d'identifier parmi ces clones édités ceux dont les Indels (insertions et délétions) générés par CRISPR/Cas9 ont induit l'apparition de codons STOP.

Par ailleurs, compte-tenu des résultats obtenus chez la souris et démontrant une importante compensation entre les deux paralogues *Mbnl1* et *Mbnl2*, nous avons également généré une lignée hiPSC *MBNL1(-/-) ; MBNL2(-/-)* (DKO) (Lee et al. 2013). La condition *MBNL1(-/-) ; MBNL2(+/-)* fut également générée en raison de la forte mortalité néonatale observée chez les souris DKO alors que les souris *Mbnl1(-/-) ; Mbnl2(+/-)* sont viables mais développent tout de même des phénotypes sévères (Goodwin et al. 2015; Lee et al. 2013). Le gRNA ciblant *MBNL2* ayant été moins efficace que celui-ci ciblant *MBNL1* (environ 9% d'efficacité d'édition contre 25% respectivement), nous avons donc choisi d'obtenir ces deux lignées hiPSC *MBNL1(-/-) ; MBNL2(-/-)* et *MBNL1(-/-) ; MBNL2(+/-)* en répétant l'expérience d'édition génomique de *MBNL1* à partir des clones *MBNL2(-/-)* et *MBNL2(+/-)* précédemment obtenus. Enfin, bien que le paralogue *MBNL3* ne semble que très peu exprimé dans des tissus adultes, et pour pallier tous éventuels phénomènes de compensation dans la lignée hiPSC DKO, nous avons généré une lignée *MBNL1(-/-) ; MBNL2(-/-) ; MBNL3(-/-)* (TKO). Les clones obtenus dans la suite de notre étude ont été validés par un contrôle qualité de pluripotence par cytométrie en flux. L'intégrité génomique fut également confirmée par des analyses caryotypiques et par le génotypage des SNP. Enfin, une analyse des sites off-target exoniques ainsi que des sites introniques ou intergéniques les plus susceptiblement reconnus par la Cas9, a été réalisé et a permis de valider l'absence d'effet off-target au niveau de ces séquences dans chacune des lignées utilisées dans cette étude.

La double déplétion des protéines MBNL1 et MBNL2 affecte les étapes tardives de la myogenèse et reproduit les caractéristiques principales de la DM1

En adaptant à nos cultures cellulaires un protocole de différenciation myogénique n'utilisant aucun transgène (Caron et al. 2016), nous avons pu évaluer la capacité des hiPSC déplétées en MBNL à se différencier en cellules musculaires squelettiques. Dans un premier temps, j'ai suivi la cinétique d'expression de ces protéines à partir d'hiPSC WT et observé l'augmentation graduelle des niveaux d'expression de *MBNL1* et *MBNL2* tout au long de la différenciation, celles-ci n'étant pas retrouvées exprimées au stade indifférencié. Concernant *MBNL3* nous n'avons observé qu'une très légère expression dans les premiers jours de

différenciation et aucune expression au niveau des myoblastes ou des myotubes, en accord avec ce qui avait été décrit dans des myoblastes primaires humains (Holt et al. 2009).

En différenciant par la suite nos différentes lignées *MBNL* KO, j'ai tout d'abord pu valider la perte d'expression des protéines MBNL selon les clones correspondants et observer une compensation entre les paralogues MBNL1 et MBNL2 avec une augmentation de l'expression de MBNL1 et de MBNL2 dans cellules musculaires dérivées respectivement d'hiPSC *MBNL2*(-/-) et *MBNL1*(-/-).

Concernant l'impact fonctionnel des protéines MBNL durant la différenciation myogénique, nous avons observé que la déplétion de ces protéines ne perturbait pas l'engagement myogénique, avec des taux similaires de cellules (autour de 80%) exprimant le marqueur des progéniteurs myogéniques PAX3 au 5^{ème} jour de différenciation par rapport aux cellules dérivées d'hiPSC DM1 ou WT. De la même manière, nous n'avons observé aucune différence significative concernant le pourcentage de cellules positive à la Desmine, un filament intermédiaire spécifique du muscle, au 16^{ème} jour de différenciation, indiquant que les précurseurs myogéniques sont capables de se différencier en myoblastes en l'absence d'expression des protéines MBNL. Malgré l'absence de phénotype cellulaire, certains défauts d'épissage alternatif caractéristiques de la DM1 sont cependant détectables dès ce stade myoblaste, tels que les défauts d'épissage de l'exon 78 de *DMD* et de l'exon 11 de *INSR*. Ces défauts d'épissage sont encore plus nombreux au niveau des myotubes avec une réelle aggravation dans les myotubes dérivés d'hiPSC DKO et TKO, notamment concernant l'exon 11 de *BIN1* ou encore de l'exon 29 de *CACNA1S*.

En parallèle des lignées déplétées en MBNL, nous avons également inclus dans notre étude deux lignées d'hiPSC générées à partir de deux patients différents ainsi qu'une lignée WT contrôle issue de la sœur d'un des deux patients et non atteinte par la maladie. L'utilisation de ces lignées nous ont permis de reproduire, comme cela a déjà été observé dans des cultures primaires et immortalisées issues de patients DM1 (Arandel et al. 2017), un défaut de fusion myogénique dans les myotubes dérivés d'hiPSC DM1 ainsi que plusieurs défauts d'épissage décrits chez les patients DM1.

De manière intéressante, on retrouve également ce phénotype en l'absence des protéines MBNL. Celui-ci est modéré dans les myotubes dérivés des hiPSC simple KO alors que nous le retrouvons aggravé dans les myotubes dérivés des hiPSC *MBNL2*(-/-) et *MBNL2*(+/-), DKO et TKO, avec aucune différence significative entre les DKO et les TKO. Plusieurs conclusions peuvent être dressées à ce stade. Tout d'abord, nos résultats démontrent la possibilité d'utiliser un protocole de différenciation myogénique efficace et n'utilisant pas de transgène. Appliqué à des lignées DM1, il permet de générer des cellules musculaires squelettiques qui reproduisent les stigmates moléculaires et cellulaires associées à la pathologie. Concernant les protéines MBNL, nos résultats montrent un rôle des protéines MBNL1 et MBNL2 dans les étapes tardives de la myogénèse *in vitro* et permettent de confirmer l'importante compensation fonctionnelle qui existe entre ces deux paralogues.

Analyse comparative des voies moléculaires dérégulées entre les myotubes dérivés d'hiPSC DM1, *MBNL1* KO et DKO

Dans le but d'étudier plus finement les conséquences de la déplétion des protéines MBNL sur la myogenèse humaine, nous avons réalisé une analyse transcriptomique comparative par séquençage AmpliSeq entre les myotubes dérivés d'hiPSC *MBNL1* KO, DKO et WT. En comparant les myotubes *MBNL1* KO et WT ainsi que les myotubes DKO et WT, nous avons identifié dans les deux cas plus de 120 gènes différentiellement exprimés (DEGs), avec seulement 12 gènes communément dérégulés dans ces deux listes. En revanche, 258 gènes sont retrouvés dérégulés entre myotubes DKO et *MBNL1* KO, suggérant un rôle compensatoire de MBNL2 qui pourrait expliquer le peu de gènes communément dérégulés entre myotubes *MBNL1* KO comparés aux myotubes WT ainsi que myotubes DKO et WT. De plus, l'analyse par "gene ontology" (GO) des 126 DEGs dans les myotubes *MBNL1* KO comparés aux WT ne révèle que peu de processus biologiques liés au muscle squelettique, avec la dérégulation de gènes majoritairement impliqués dans "l'organisation de la matrice extracellulaire", "la régulation de la migration cellulaire" ou encore "la régulation de la croissance cellulaire", alors que celles concernant les DEGs observés dans les myotubes DKO, qu'ils soient comparés aux myotubes *MBNL1* KO ou aux myotubes WT, montrent de nombreux procédés biologiques down-régulés intéressants comme "l'organisation de la structure d'actomyosine" ou encore "la contraction du muscle strié". Ces résultats soulignent encore une fois l'importante compensation de MBNL1 par MBNL2.

Nous avons poursuivi notre étude en réalisant un RNA-Seq sur ces mêmes cellules et en incluant les myotubes dérivés d'hiPSC DM1 pour analyser la contribution des protéines MBNL dans les défauts d'épissage observés dans DM1. Nous avons observé 696 exons différentiellement épissés (DSEs) entre myotubes DM1 et WT, 399 DSEs entre myotubes *MBNL1* KO et WT et 547 entre myotubes DKO et WT. Parmi les 696 DSEs observés dans les myotubes dérivés d'hiPSC DM1, 254 (36%) sont également dérégulés dans les myotubes DKO et 224 (32%) sont retrouvés dans les myotubes dérivés d'hiPSC *MBNL1* KO. Enfin, 126 DSEs ont été observés à la fois dans les myotubes dérivés d'hiPSC DM1, *MBNL1* KO et DKO. Parmi les 547 DSEs observés dans les myotubes DKO, 164 (30%) sont également retrouvés dans les myotubes *MBNL1* KO comparés aux myotubes WT, suggérant ici aussi une compensation fonctionnelle importante de MBNL2. De manière intéressante, les analyses par GO des DSEs observés dans les myotubes DM1 et de ceux observés dans les myotubes DKO ont révélé des processus biologiques dérégulés liés au muscle squelettique au niveau par exemple de la "contraction du muscle strié" et de "l'organisation des filaments d'actine", avec également la dérégulation de gènes impliqués dans "l'organisation sarcomérique" et le "développement de la jonction neuromusculaire" dans les myotubes DKO. Nous avons ainsi pu mettre en évidence une trentaine de DSEs impliqués dans la majorité de ces procédés biologiques et dont 21 présentent un défaut similaire dans les myotubes dérivés d'hiPSC DKO et DM1.

Identification d'un défaut d'épissage alternatif de la DM1, régulé par les protéines MBNL et responsable d'anomalies structurales du compartiment post-synaptique musculaire

Parmi les différents défauts d'épissage identifiés précédemment, nous avons choisi de nous intéresser plus particulièrement à celui qui touche le gène *SORBS1* pour plusieurs raisons. Tout d'abord, ce défaut d'épissage correspond à l'un des événements d'épissage retrouvés le plus fortement altérés dans les myotubes dérivés d'hiPSC DM1 et DKO. En nous référant à la littérature, nous avons également constaté que ce défaut d'épissage avait déjà été identifié à la fois dans le muscle cardiaque de souris *Mbnl1(-/-); Mbnl2(+/-)* (Lee et al. 2013) ainsi que dans des muscles squelettiques issus de biopsies de patients atteints des formes congénitales (Nakamori et al. 2017) et adultes (Wang et al. 2019a) de la DM1. Finalement, le dernier argument s'est reposé sur une étude récente démontrant l'implication de la protéine Sorbs1 dans l'aggrégation des récepteurs à l'acétylcholine (AChR) (Hallock et al. 2016). Même si ces résultats avaient été générés dans des cellules musculaires murines C2C12, ils suggéraient un rôle de Sorbs1 dans une des étapes essentielles dans la formation des jonctions neuromusculaires. De façon intéressante, différentes études avaient par le passé montré des défauts structurels et organisationnels des clusters des AChR à la fois dans un modèle murin sévère de la DM1 (Panaite et al. 2013) ainsi que dans les souris *Mbnl1*/DKO et TKO (Lee et al. 2013; Thomas et al. 2017).

Nous avons pu observer dans un premier temps un défaut au niveau de la transition développementale des isoformes de *SORBS1*, contenant ou non l'exon 25, dans les muscles squelettiques issus d'embryons atteints de formes congénitales de la DM1, comparés à des échantillons provenant d'embryons sains. Après avoir par la suite confirmé le défaut d'inclusion de l'exon 25 de *SORBS1* dans des cellules musculaires primaires issues de patient DM1 comparés à des myotubes primaires sains (16 et 96% d'inclusion respectivement), nous avons pu valider en analysant celui-ci dans nos différentes lignées CRISPR, que ce défaut d'épissage était bien régulé par les protéines MBNL avec une aggravation du défaut d'épissage avec la double déplétion de MBNL1 et MBNL2.

Afin d'étudier les conséquences fonctionnelles de ce défaut d'épissage, nous avons mis en place une stratégie de saut d'exon médiée par des oligonucléotides anti-sens (ASO). Nous avons dans un premier temps testé plusieurs séquences d'ASO ciblant différentes séquences régulatrices de l'épissage de cet exon en analysant leur effet sur son inclusion 48h post-transfection. J'ai ainsi pu observer que l'ASO ciblant la séquence ESE (exonic splicing enhancer) était le plus efficace car lui-seul permettait de diminuer le niveau d'inclusion de l'exon 25 de *SORBS1* dans des myotubes WT et ce, de façon similaire au niveau d'inclusion observé dans les myotubes DM1. Nous nous sommes par la suite assurés que ce saut d'exon ne modifiait pas l'expression de l'ARNm *SORBS1* dans ces myotubes sains transfectés.

Nous sommes interrogés quant à l'éventuel rôle du défaut d'épissage de l'exon 25 de *SORBS1* dans le clustering des AChR au niveau du compartiment musculaire post-synaptique. J'ai ainsi mis en place un protocole permettant de transfecter les cellules musculaires dérivées d'hiPSC WT avec l'ASO ciblant l'ESE, suivi par un traitement à l'Agrin dans le but d'induire la formation des clusters des AChR. Nous avons tout d'abord observé une diminution importante de l'immuno-marquage des AChR dans les myotubes traités par l'ASO comparés aux myotubes WT (mock). Afin d'aller plus loin dans notre analyse, nous avons évalué différents paramètres tels que la taille moyenne de ces clusters, l'aire totale par myotubes ou encore leur nombre en les catégorisant selon leur taille. En accord avec les défauts d'inclusion de l'exon 25 de *SORBS1*, nous avons dans un premier temps observé une diminution significative des clusters des AChR dans les myotubes dérivés d'hiPSC DM1 et DKO par rapport aux myotubes dérivés d'hiPSC WT. De manière intéressante, nous avons également observé une diminution de l'aire totale des AChR et de leur taille moyenne, ainsi qu'une diminution importante du nombre de gros clusters (supérieurs à $40\mu\text{m}^2$) dans les myotubes traités avec l'ASO. Nos résultats indiquent ainsi que l'exclusion de l'exon 25 de *SORBS1* médiée par l'ASO permet de reproduire à elle-seule les défauts de clustering des AChR observés dans les myotubes dérivés d'hiPSC DM1 et DKO.

Ayant fait par la suite l'observation de ces mêmes conséquences phénotypiques dans des myotubes primaires humains nous avons ainsi pu confirmer que l'exclusion de l'exon 25 de *SORBS1* était suffisante pour induire les défauts de clustering des AChR également dans ces cellules. Afin de mieux comprendre le rôle précis de l'exon 25 de *SORBS1*, nous nous sommes demandé si son exclusion avait un effet sur la stabilité de la protéine, sur sa localisation ou encore au niveau des interactions entre *SORBS1* et ses potentiels partenaires. Après avoir vérifié que l'ASO ne modifiait pas le niveau d'expression de la protéine, nous avons observé à l'aide d'analyses de microscopie confocale une différence de localisation de la protéine *SORBS1*, plus diffus après traitement et donc en absence d'exon 25. Ainsi, l'exon 25 de *SORBS1* semblerait avoir un rôle dans la localisation de la protéine qui serait essentielle pour la formation des clusters des AChR au niveau du compartiment post-synaptique. Ensemble, nos résultats pourraient expliquer d'un point de vue clinique certains défauts fonctionnels au niveau de la communication neuromusculaire chez les patients DM1.

Loss of MBNL in pluripotent stem cells reveals abnormal splicing switch of SORBS1 compromising the neuromuscular development in Myotonic Dystrophy type 1

Antoine Mérien¹, Julie Tahraoui-Bories¹, Gilles Moulay⁴, Michel Cailleret¹, Jean-Baptiste Dupont¹, Céline Leteur², Jérôme Polentes², Alexandre Carteron², Hélène Polvèche², Jean-Paul Concordet³, Christian Pinset², Margot Jarrige², Stéphane Vassilopoulos⁴, Denis Furling⁴, Cécile Martinat^{1*}

¹ INSERM / UEPS UMR 861, Paris Saclay University, I-STEM, 91100 Corbeil-Essonnes, France.

² CECS/AFM, I-STEM, 91100 Corbeil-Essonnes, France.

³ INSERM U1154 / CNRS UMR7196, MNHN TACGENE, 75005, Paris, France.

⁴ Sorbonne Université, INSERM, Association Institut de Myologie, Centre de recherche en myologie, UMR-S 974, Paris, France.

* **Corresponding author** : C. Martinat, PhD, I-Stem, 28 rue Henri Desbruères 91100 Corbeil-Essonnes France, Tel: + 00 33 1 69 90 85 33, Fax: + 00 33 1 69 90 85 21, E-mail: cmartinat@istem.fr

ABSTRACT

Alternative splicing has emerged as a fundamental mechanism not only for the diversification of protein isoforms but also for the spatiotemporal control of development. Therefore, a better understanding of how this mechanism is regulated has the potential to elucidate fundamental biological principles, and also to decipher pathological mechanisms involved in diseases where normal splicing networks are mis-regulated. Here, we took advantage of human pluripotent stem cells to decipher during human myogenesis the role of MBNL proteins, a family of tissue-specific splicing regulators whose loss of function is associated with Myotonic Dystrophy type 1 (DM1), an inherited neuromuscular disease. Thanks to the CRISPR/Cas9 technology, we generated human-induced pluripotent stem cells (hiPSCs) depleted in MBNL proteins and evaluated the consequences of this loss on the generation of skeletal muscle cells. Our results reveal the temporal requirement of MBNL proteins in human *in vitro* myogenesis and validate that the loss of function of these proteins recapitulate the main features of DM1. Combined to a transcriptomic analysis, we were able to highlight the molecular pathways regulated by these proteins during myogenesis that could be involved in the development of DM1. This led us to identify a new alternative splicing defect in DM1, regulated by MBNL proteins. This missplicing might affect the localization of SORBS1 that consequently leads to structural abnormalities of the muscular post-synaptic compartment. These findings may be clinically significant for the functional changes in neuromuscular connections associated with DM1.

Keywords: Human pluripotent stem cells, CRISPR/Cas9, Muscleblind-like, skeletal muscle cell differentiation, Myotonic dystrophy type 1.

INTRODUCTION

Alternative splicing has emerged in the post-genomic era as the main driver of proteome diversity. The identification that at least 95% of human genes produce multiple spliced RNA via alternative exon usage has revealed the prevalence of this additional layer of gene expression regulation (Wang et al. 2008; Pan et al. 2008). The regulation of alternative splicing varies with cell type, during development or upon cellular differentiation, thereby participating in the fine-tuning of a gene signature temporally and spatially (Merkin et al. 2012). To ensure this developmental and tissue-specific regulation, alternative splicing relies on complex mechanisms involving an intricate protein-RNA network. This leads to the precise and sequential recognition of cis-acting sequences on the pre-mRNA by trans-acting RNA-binding factors. It has become apparent that these coordinated splicing networks have important physiological functions in developmental processes and that disruption of these networks negatively impacts health and contributes to human diseases.

According to some estimates, up to 50% of all pathogenic mutations may affect splicing in some way (Cooper et al. 2009). Illustrating this, Myotonic Dystrophy Type 1 (DM1) represents one paradigmatic example of genetic disease in which loss of MBNL splicing factors is considered as the central pathological event. DM1 is an autosomal dominant neuromuscular disease caused by abnormal expansion of CTG repeats in the 3'untranslated region (UTR) of the *Dystrophia myotonia protein kinase (DMPK)* gene (Brook et al. 1992; Mahadevan et al. 1992). The pathogenesis of DM1 involves an RNA gain-of-function mechanism caused by the expression of mutant RNAs containing hundreds to thousands of CUG repeats that sequester Muscleblind-like RNA binding proteins (MBNL) involved in the regulation of alternative splicing (Miller et al. 2000; Fardaei et al. 2002). This functional loss of MBNL proteins results in several alternative splicing mis-regulation events that have been associated for some of them to clinical symptoms such as myotonia, muscle weakness and insulin resistance (Miller et al. 2000; Nakamori et al. 2013).

Three different genes encode human MBNL homologs: *MBNL1*, *MBNL2* and *MBNL3* which all share structural similarities (Fardaei et al. 2002; Kanadia et al. 2003b; Konieczny et al. 2014). *MBNL1* and *MBNL2* are ubiquitously expressed but while *MBNL1* is found in similar levels in all tissues, *MBNL2* is more abundant in the brain (Kanadia et al. 2003b; Konieczny et al. 2014). In a majority of tissues *MBNL1* and *MBNL2* expression rises during differentiation. In contrast, *MBNL3* is expressed at much lower levels, with peak expression in placenta, and its functions are poorly defined compared to its two other paralogs (Holt et al. 2009; Konieczny et al. 2014). According to this differential expression over time, *MBNL1* and *MBNL2* proteins have been shown to modulate the splicing profile of a large number of genes during fetal-to-adult transitions (Lin et al. 2006). At the functional level, knockout *Mbnl1* mice display skeletal muscle defects such as myotonia or central nuclei only at an adult stage (Kanadia et al. 2003a) whereas loss of *Mbnl2* mice leads predominantly to brain phenotypes with no major defects in skeletal muscles (Charizanis et al. 2012). However, the exploitation of these loss-of-function models was hampered by compensatory mechanisms which result in an increased expression of *Mbnl2* as observed in knock-out *Mbnl1* mice (Lee et al. 2013). In contrast, *Mbnl1*(-/-); *Mbnl2*(-/-) double knockouts are embryonic lethal while *Mbnl1*(-/-); *Mbnl2*(+/-) mice are viable but develop severe phenotypes such as reduced lifespan, heart conduction block, severe myotonia and progressive skeletal muscle weakness (Lee et al. 2013; Goodwin et al. 2015). Concerning *Mbnl3*(-/-) mice, results showed decreased skeletal muscle regeneration and reduced grip strength associated with age despite low expression of *Mbnl3* in adult muscles. Moreover, all the phenotypes described in *Mbnl3* (-/-) mice were not associated with significant changes in alternative splicing (Poulos et al. 2013; Choi et al. 2016). Altogether, these studies have highlighted the developmental function of *Mbnl* proteins in mice, but they also raise the question of

whether MBNL proteins are similarly involved in human development as well as the precise timing at which these proteins intervene.

In this study, we focused on the role of MBNL proteins during human skeletal myogenesis. Due to their capacity to differentiate into any cell types of an organism, we based our approach on the combined use of human induced pluripotent stem cells (hiPSCs) and CRISPR/Cas9 technology to generate hiPSCs depleted in MBNL homologs in order to evaluate the impact of MBNL proteins loss at the cellular and molecular level during the successive developmental steps of myogenesis. Thanks to a comparative transcriptomic approach, we identified the defective inclusion of *SORBS1* exon 25 in DM1 and *MBNL1,2* DKO hiPSC-derived myotubes and in skeletal muscle samples from DM1 patients. Through antisense oligonucleotide (ASO)-mediated exon skipping experiments, we demonstrated that this missplicing leads to structural abnormalities of the muscular post-synaptic compartment.

RESULTS

CRISPR/Cas9 strategy to generate MBNL KO(s) hiPSCs

To generate *MBNL* knockout hiPSC lines, sgRNAs were specifically designed to target one of the first coding exons as well as the presence of an enzymatic restriction site at the PAM sequence (Figure 1A and 1B) (Concordet and Haeussler 2018). Each sgRNA was complexed with SpCas9 protein to perform ribonucleoprotein transfection of dissociated hiPSCs (Ramakrishna et al. 2014). The selection of gene-edited clones was based on a two-step selection process. Thanks to the use of sgRNA targeting sequences within a restriction endonuclease site (Figure 1B; framed in purple) at the PAM level (Figure 1B; NGG in red), restriction fragment length polymorphism was first analyzed to rapidly detect CRISPR gene editing according to total or partial loss of restriction sites (Table S1). The second selection was based on Sanger DNA sequencing to select clones with INDELS that have generated STOP codons (Figure 1C. For *MBNL1*, we obtained an editing efficiency of 25% with 24 pre-selected clones out of 96 clones tested (10 with a homozygous profile and 14 heterozygous). After the sequencing of these 24 clones, 3 homozygous and 2 heterozygous clones were validated (Figure 1C and 1D and Supplementary Figure 1A). Concerning *MBNL2*, 9 clones were pre-selected out of 96 (editing efficiency of 9%) and 3 were finally validated by sequencing. We generated two heterozygous clones and obtained only one *MBNL2*(-/-) homozygous knockout clone (Figure 1C and Supplementary Figure 1A). Because of the compensatory phenomena observed between *Mbnl2* and *Mbnl1* in knockout mice (Lee et al. 2013), double knockout *MBNL1*(-/-); *MBNL2*(-/-) (DKO) were generated (Figure 1C and 1D and Supplementary Figure 1A). The condition *MBNL1*(-/-); *MBNL2*(+/-) was also generated as *Mbnl1*(-/-); *Mbnl2*(-/-) double mice knockouts are embryonic lethal whereas *Mbnl1*(-/-); *Mbnl2*(+/-) mice are viable but develop severe phenotypes (Lee et al. 2013; Goodwin et al. 2015). Finally, to decipher the contribution of *MBNL3*, we generated triple knockout of *MBNL1*(-/-); *MBNL2*(-/-); *MBNL3*(-/-). We obtained an editing efficiency of 22% and validated 5 different triple knockout (TKO) clones, two of which were considered for further analyses (Figure 1C and 1D and Supplementary Figure 1A). For each generated clone, pluripotency was validated by flow cytometry of the pluripotent markers SSEA3 and TRA1-81. Genomic integrity was also verified by a karyotype analysis and SNP genotyping (Supplementary Figure 1B and 1C). To determine off-target activity of our gRNAs, we analyzed the most likely off-target sites, and no mutation induced by genome editing was observed (Table S2).

No major impairment of the early myogenic commitment in the absence of MBNL proteins

The capacity of MBNL-depleted human pluripotent stem cells to differentiate into skeletal muscle cells was evaluated by using a transgene-free protocol which can be schematically divided into three steps (Caron et al. 2016) (Figure 2A). We first monitored the expression of the different isoforms of MBNL at the protein and mRNA level along this myogenic differentiation process. As previously reported (Venables et al. 2013), MBNL1 and MBNL2 were not found to be expressed in undifferentiated human pluripotent stem cells. Their level of expression gradually increased as differentiation progresses to reach a maximum expression after ~20 days of differentiation, corresponding to the beginning of the final myogenic maturation step (Figure 2B and Supplementary Figure 2A and 2D).

In contrast, MBNL3 expression slightly increased during the first days of differentiation followed by its decreased expression at later time points (Figure 2B and Supplementary Figure 2A-D). As previously described in human primary myoblasts (Holt et al. 2009), our results also confirmed that MBNL1 is the most abundant MBNL proteins in hiPSC-derived skeletal muscle cells although MBNL1 and MBNL2 proteins follow a similar kinetic of expression (Figure 2B).

The analysis at the protein level demonstrated a specific lack of expression of MBNL paralogs in the different gene-edited hiPSC lines differentiated into skeletal muscle cells (Figure 2C and 2D). As reported in mouse model of *Mbnl* loss of function (Lee et al. 2013), we also observed a compensatory increase in MBNL1 and MBNL2 expression in *MBNL2*(-/-) and *MBNL1*(-/-), respectively (Figure 2C and 2D and Supplementary Figure 2E). MBNL3 level was found not to be increased as a compensatory response to MBNL1 and MBNL2 loss (Supplementary Figure 2C and 2E).

To evaluate the functional impact of MBNL proteins loss during myogenic differentiation, the expression of stage-specific muscle markers was evaluated at different time points. As MBNL proteins are known to be important players in DM1 pathogenesis (Wang et al. 2012; Thomas et al. 2017), we used DM1 hiPSC as a reference (Supplementary Figure 3A and 3B). At 5 days of differentiation, approximately 80% of cells expressed the myogenic progenitor marker PAX3, with no significant difference between the different cell lines indicating that muscle lineage commitment is not affected by MBNL depletion (Figure 3A and Supplementary Figure 4A). Similar observation was made at 16 days of differentiation by the quantification of cells positive for DESMIN, a muscle-specific intermediate filament. Between 74 and 87% of cells were DESMIN⁺ with no significant difference between conditions showing that the myogenic precursors developed into skeletal myoblasts in the absence of MBNL proteins (Figure 3B and Supplementary Figure 4B).

Despite the absence of phenotypes at the cellular level, we sought to evaluate the molecular consequences of MBNL depletion. We chose to focus on several gene transcripts important for skeletal muscle function and previously shown to be mis-spliced in DM1 patients' biopsies (Nakamori et al. 2007, 2013). While these mis-splicing events were identified in adult tissues, RT-PCR analysis showed that the splicing profiles of *DMD* exon 78, *INSR* exon 11 and *MBNL1* exon 5 were significantly deregulated at the myoblast stage in DM1, *MBNL1*(-/-), DKO and TKO compared to non-DM1 control (WT) (Figure 3C and Supplementary Figure 4C). DKO showed more pronounced splicing changes than *MBNL1*(-/-) and DM1 conditions. A similar observation was made with the depletion of the three paralogs (Figure 3C). In contrast, the loss of MBNL2 alone did not significantly lead to mis-splicing events when compared to non-DM1 control (WT) myoblasts illustrating the compensation of MBNL2 by MBNL1 as previously reported (Wang et al. 2012) (Figure 3C). These observations suggested that MBNL proteins are dispensable for myogenic commitment even if MBNL loss of function led to early mis-splicing events that could interfere with

later stages of myogenesis.

MBNL1 and MBNL2 depletion impacts late myogenesis and reproduces the main features associated with DM1

We next evaluated the consequences of MBNL loss of function on the capacity of hiPSC-derived myoblasts to further differentiate and fuse into myotubes (Caron et al. 2016). As a point of reference, we first characterized this late differentiation by using DM1 hiPSC-derived myoblasts. Primary or immortalized muscle cell cultures derived from muscle biopsies of DM1 patients have been shown to present altered myogenic fusion as well as DM1-associated molecular features including mutated *DMPK* mRNAs aggregates that sequester MBNL1 and subsequently lead to alternative splicing defects (Arandel et al. 2017). Accordingly, we also observed that DM1 hiPSC-derived myoblasts under differentiation condition form smaller myotubes and present a significantly reduced fusion capacity when compared to WT hiPSC-derived myoblasts (Figure 4A and Supplementary Figure 5A). Moreover, the main DM1-associated molecular features including mutated *DMPK* mRNAs aggregates, nuclear sequestration of MBNL1 and MBNL2 and presence of several mis-spliced events were detected in these cells (Figure 4B-C, Supplementary Figure 3C-E and Supplementary Figure 5B). These data demonstrate that DM1 hiPSC-derived skeletal myotubes recapitulate phenotypic and molecular traits similar to those found in skeletal muscle cell cultures isolated from DM1 patients.

In comparison, loss of either MBNL1 or MBNL2 resulted in a moderate effect on myogenic fusion which is likely related to the compensatory phenomenon existing between the two paralogs (Figure 4A and Supplementary Figure 5A). Immunofluorescence analysis revealed that this upregulation was mainly localized in the nucleus (Figure 4B), in agreement with previous observations made in murine models of *Mbnl* loss of function (Lee et al. 2013). In contrast to single *MBNL* knock-out lines, *MBNL1*(-/-); *MBNL2*(+/-) and DKO hiPSC-derived myotubes showed an aggravated fusion defect (Figure 4A). Concordant with the absence of MBNL3 protein at this stage, depletion of MBNL3 in addition to the two paralogs did not exacerbate this phenotype in TKO lines when compared to DKO (Figure 4A).

At the molecular level, and concordant with the observations made at the myoblast stage, several DM1-associated mis-splicing events, known to contribute to muscle function and maintenance such as *BINI* exon 7 and *DMD* exon78, were found to be mis-spliced in DM1 hiPSC-derived myotubes (Figure 4C). For all targets examined, dramatic mis-splicing events were also observed in DKO myotubes (Figure 4C and Supplementary Figure 5B). These data strongly suggest that MBNL1 and MBNL2 proteins regulate RNA processing essential for normal late myogenesis.

Comparative analysis of the molecular pathways misregulated between DM1, MBNL1 KO and DKO hiPSC-derived myotubes

To better evaluate the consequence of MBNL1 and MBNL2 depletions during myogenesis, we sought to conduct a comparative transcriptomic analysis between DKO and WT hiPSC-derived myotubes. We also included *MBNL1*(-/-) hiPSC-derived myotubes to evaluate the effect of MBNL2 compensation. For this purpose, we used AmpliSeq targeted sequencing technology allowing analysis of over 20.000 distinct human RNA targets simultaneously using a highly multiplexed amplification method (Li et al. 2015). The principal component analysis (PCA) showed a distinct transcriptomic profile for each condition (Supplementary Figure 6A). The comparison between *MBNL1*(-/-) and WT hiPSC-derived myotubes led to the identification of 126 genes differentially expressed (false discovery rate (FDR) < 0.05; |Log2FoldChange| > 0.4), with 31 being up-regulated and 95 down-regulated in *MBNL1*(-/-) myotubes compared to WT (Supplementary Figure 6B). In the same order

of magnitude, 122 differentially expressed genes (DEGs) were found when comparing DKO to WT hiPSC-derived myotubes, 40 up-regulated and 82 down-regulated in DKO myotubes. A more pronounced difference was observed between DKO and *MBNLI*(-/-) hiPSC-derived myotubes with 258 differentially expressed genes (Supplementary Figure 6B). All the differentially expressed genes (DEGs) are listed in Table S3. Strikingly, only 12 genes were commonly deregulated in *MBNLI*(-/-) versus WT myotubes and in DKO versus WT myotubes (Figure 5A). Hierarchical clustering on the 122 DEGs detected in DKO compared to WT hiPSC-derived myotubes showed that *MBNLI*(-/-) hiPSC-derived myotubes were comparable to WT hiPSC-derived myotubes (Figure 5B). Accordingly, gene ontology (GO) analysis of DEGs in *MBNLI*(-/-) versus WT myotubes revealed enrichment in multiple biological processes, but few of these are related to skeletal muscle (Supplementary Figure 6C and Table S4). These observations could be explained, at least, by the functional compensation by MBNL2 in *MBNLI*(-/-) myotubes.

Concerning DEGs in DKO compared to WT hiPSC-derived myotubes, gene ontology analysis revealed enrichment in several biological processes related to skeletal muscle such as “striated muscle contraction” (adjusted pvalue: 0,00001), “actomyosin structure organization” (adjusted pvalue: 0,02341) or “muscle filament sliding” (adjusted pvalue: 0,02348) (Figure 5C). Strikingly, most of these muscle related-biological processes were found to be down-regulated in DKO hiPSC-derived myotubes (Figure 5C). In the same way, gene ontology analyses of DEGs in DKO compared to *MBNLI*(-/-) hiPSC-derived myotubes showed similar results with multiple down-regulated biological processes related to skeletal muscle (Supplementary Figure 6C-D). Among the list of down-regulated genes identified in DKO hiPSC-derived myotubes, we highlighted two interesting candidates for further explorations (Figure 5C): TCAP (also known as Telethonin) plays an essential role in the sarcomeric assembly by anchoring Titin in the Z-disk of the sarcomere (Bertz et al. 2009) and is known to be involved in Limb-Girdle Muscular Dystrophy 2G (LGMD2G) (Markert et al. 2010) ; MYOM1 is a constituent of the sarcomeric M band and plays an important role in sarcomere assembly and integrity (Prill et al. 2019). Interestingly, MYOM1 is found mis-spliced in muscular biopsies from DM1 patients (Nakamori et al. 2017) and this splicing defect could be caused by the down-regulation of MBNL proteins (Koebis et al. 2011).

Thus, we pursued our analysis by searching for cassette exons that demonstrate a significant difference in the level of inclusion between *MBNLI*(-/-), DKO and WT hiPSC-derived myotubes. A total of 547 cassette exons were found to be deregulated between DKO and WT hiPSC-derived myotubes, among which 296 and 251 exhibited a significant decrease and increase in inclusion, respectively (FDR < 0.05; Delta percent spliced-in (PSI) > 0.1) (Figure 5D and Table S5). Among the 547 differentially spliced-events (DSEs) observed in DKO hiPSC-derived myotubes, 164 (30%) were also found deregulated in *MBNLI*(-/-) hiPSC-derived myotubes compared to WT, suggesting here again an important functional compensation by MBNL2. We next compared these analyses with transcriptomic data generated from DM1 hiPSC-derived myotubes (Figure 5D). Of the 696 mis-splicing events observed in DM1 hiPSC-derived myotubes, 224 (32%) appeared in *MBNLI*(-/-) hiPSC-derived myotubes. In the same way, 254 (36%) appeared in DKO hiPSC-derived myotubes and among these, 126 were commonly mis-spliced in DM1, *MBNLI*(-/-) and DKO hiPSC-derived myotubes (Figure 5E). Interestingly, gene ontology analysis of these 126 DSEs showed deregulation of biological processes related to “vesicle-mediated transport”, “muscle contraction” or “cytoskeleton organization” (Table S6).

Functional annotation analysis of the DKO transcriptomic data revealed that the most relevant GO terms concerned skeletal muscle such as "sarcomere organization" (adjusted pvalue: 0,068) or “neuromuscular junction development” (adjusted pvalue: 0,06) (Figure 5F and Table S6). We highlighted 30 enriched DSEs involved in most of these muscle-related biological processes

deregulated in DKO myotubes (Supplementary Figure 6E), and found several already described alternative splicing defects associated with DM1 such as *DMD* exon 78, *BINI* exon 11 and *BINI* exon 7 (Figure 4C and Supplementary Figure 5B) (Fugier et al. 2011; Rau et al. 2015). Among these 30 DSEs, 21 were also mis-spliced in DM1 hiPSCs-derived myotubes in the same way they were in DKO myotubes (Supplementary Figure 6E). Only 3 DSEs (*ABI2* ex10, *MYL5* ex3 and *TPM1* ex9) mismatched in the sense of inclusion/exclusion. Furthermore, only half of these 30 DSEs were also found deregulated in *MBNLI*(-/-) myotubes, which could explain why the biological processes deregulated in these cells are less related to skeletal muscle. In general, when DSEs concern both *MBNLI*(-/-) and DKO myotubes, the level of deregulation was worsened in DKO myotubes (Supplementary Figure 6E). In addition to *DMD* exon 78, *BINI* exon 11 and *BINI* exon 7, we validated two other DSEs in our *MBNL* knockouts hiPSC-derived myotubes (data not shown), including the aberrant inclusion of *MYOMI* exon 18 in DM1 hiPSC-derived myotubes compared to WT (Supplementary Figure 6E and Table S5) which was also validated in DKO myotubes.

Altogether, these data revealed sets of disrupted biological pathways associated to the loss of MBNL1 and MBNL2 that may be particularly relevant in the pathological context of DM1.

Inclusion of *SORBS1* exon 25 is regulated by MBNL proteins and increases throughout human myogenesis

With the objective of identifying new molecular mechanisms that could be involved in the skeletal muscle pathology of DM1, we chose to focus on the abnormal inclusion of *SORBS1* exon 25 which appears to be one of the most severe mis-splicing events observed in DM1 and DKO hiPSC-derived myotubes (with a decrease in inclusion of approximately 70%) (Supplementary Figure 6E). This choice was also dictated by the fact that *SORBS1* mis-splicing has been previously described in cardiac muscle of *Mbnl1*(-/-); *Mbnl2*(+/-) mice (Lee et al. 2013) and in tibialis anterior biopsies from adult DM1 (Wang et al. 2019) and was one of 38 splicing events associated with muscle immaturity in patients with congenital muscular dystrophy (CMD) (Nakamori et al. 2017). Consistently, we observed abnormal *SORBS1* exon 25 skipping in skeletal muscle obtained from congenital DM1 embryos, as opposed to healthy controls undergoing exon 25 inclusion from early development stages (Figure 6A). We also validated *SORBS1* exon 25 inclusion defect in DM1 primary myotubes (Figure 6B). *SORBS1* exon 25 was almost totally included in WT primary myotubes ($93,3\% \pm 0,2$) while it was only included at $16,4\% \pm 2,8$ in DM1 primary myotubes. Regarding *MBNL* KOs hiPSC-derived myotubes, our results confirmed that *SORBS1* exon 25 inclusion is regulated by MBNL proteins driven by MBNL1 and MBNL2 gradient (Figure 6C). In addition, we observed that inclusion of *SORBS1* exon 25 increased throughout human myogenesis in WT hiPSC cells, reflecting a developmental transition not observed in DM1 and DKO cells (Supplementary Figure 7A). We further evaluated the potential cellular consequences of this mis-splicing event.

***SORBS1* exon 25 is involved in AChR clustering defects observed in DM1 and MBNL DKO myotubes**

Sorbin and SH3 domain containing 1 (*Sorbs1*) has been identified as potentially involved in the early steps of neuromuscular junction formation due to its role in the formation of acetylcholine receptor (AChR) aggregation in murine C2C12 cells (Hallock et al. 2016). Strikingly, previous studies showed that AChR organization and morphology were disrupted in a severe DM1 murine model (Panaite et al. 2013) as well as in *Mbnl* DKO mice (Lee et al. 2013; Thomas et al. 2017). We therefore sought to decipher whether *SORBS1* exon 25 missplicing could be associated with defective AChR clustering. For this purpose, agrin-induced AChR clustering was analyzed in WT, DM1 and DKO hiPSC-derived myotubes. Concordant with the defective inclusion of *SORBS1* exon 25, we

observed a significant decrease in AChR clustering in DM1 and DKO hiPSC-derived myotubes when compared to WT hiPSC-derived myotubes (Figure 6D-E). This difference was both found when considering the total area of AChR clusters and the number of AChR clusters of size greater than 20 μm^2 (Figure 6F-G).

To investigate the involvement of *SORBS1* exon 25 in AChR clustering, we used an antisense oligonucleotides (ASO)-based approach promoting *SORBS1* exon 25 skipping in WT-hiPSC-derived myotubes. Several ASOs targeting key sequences for exon 25 splicing were designed and tested for their efficiency to induce exon 25 exclusion 48h post-transfection in human primary myotubes (Supplementary Figure 7B). An ASO targeting a putative exonic splicing enhancer (ESE) sequence was found to induce an exclusion level comparable to that of DM1 and DKO hiPSC-derived myotubes without affecting *SORBS1* total mRNA expression (Figure 6E and Supplementary Figure 7C). As *SORBS1* human gene contains many alternative exons (Supplementary Figure 7D), we validated the specific effect of this ASO on *SORBS1* exon 25 splicing (Supplementary Figure 7D-E). At the functional level, transfection of the ASO in WT hiPSC-derived myotubes led to a similar decrease of AChR clustering as observed in DM1 and DKO hiPSC-derived myotubes (Figure 6D-G and Supplementary Figure 8A). We verified that the effect of ASO on AChR clusters was not related to any modulation of myogenic differentiation (Supplementary Figure 8A).

To validate these findings in another cellular model, WT control human primary skeletal muscle cells were transfected with the ASO targeting *SORBS1* exon 25 exclusion. Transfected cells displayed a dramatic decrease of AChR clustering correlating with the high rates of *SORBS1* exon 25 exclusion (Figure 7A-D and Supplementary Figure 8B).

Altogether, these findings demonstrated that *SORBS1* exon 25 exclusion was sufficient to induce the AChR clustering defects observed in DM1 and *MBNL* DKO hiPSC-derived myotubes.

To better decipher the functional role of *SORBS1* exon 25, we wondered whether its exclusion had an effect on the stability of the protein or on its localization. The ASO showed no effect on the total amount of *SORBS1* protein, suggesting that stability is not largely impacted (Supplementary Figure 8C). Confocal microscopic analyses revealed cytoplasmic localization of *SORBS1* under filamentous shape as well as a co-localization with AChR clusters in WT control human skeletal muscle cells (Figure 7E-F). In contrast, the localization of *SORBS1* protein displayed a more diffuse pattern in ASO-treated skeletal muscle cells (Figure 7F). Similar results were observed in DM1 and *MBNL* DKO hiPSC-derived myotubes as well as in ASO-treated WT hiPSC-derived myotubes (Figure 7G).

DISCUSSION

Cell-type-specific or developmental-stage-specific alternative splicing events are largely controlled by recruiting RNA-binding proteins (RBPs) that recognize specific regulatory sequences embedded in the pre-mRNA transcripts. In this study, we described the temporal requirement of the Muscleblind-like (MBNL) RBP family for defining human skeletal muscle maturation stages. Our results demonstrated that MBNL1 and MBNL2 were dispensable for the early myogenic commitment but played an important role for the late stages of myogenesis. Thanks to a comparative transcriptomic analysis, this study revealed a novel missplicing event associated to DM1 and regulated by MBNL proteins, that affects the developmentally regulated *SORBS1* exon 25 exclusion which ultimately results in disrupted post-synaptic muscle compartment. Given the technical limitations of studying

embryonic and fetal human tissues, our study establishes new human cellular models to explore the role of these proteins in tissue development. Furthermore, the different hiPSC-based models developed in this study also pave the way towards the development of new therapeutic strategies capable to cope with the loss of function of MBNL proteins.

While transcriptional regulation has been demonstrated to play an important role in myogenesis, there is a growing recognition of the contributions of alternative splicing to skeletal muscle development as well as the refinement of muscle function (Nikonova et al. 2020). Skeletal muscle was also one of the first tissues in which broad use of alternative splicing to increase cellular complexity was documented (Llorian and Smith 2011). A recent study has estimated the presence of hundreds of RBPs expressed in muscle of different species such as human, rodent, flies and worms, but only a small proportion (around 3%) have been studied (Nikonova et al. 2019). Specifically, the role of CELF (CUG-BP and ETR-3 like), Muscleblind-like (MBNL) and Fox families in controlling muscle-specific splice events has been largely documented in murine models. The Rbfox proteins are among the most heavily studied RBPs in the context of myogenesis. Gain- and loss-of-function experiments have demonstrated that Rbfox1 and Rbfox2 coordinately regulate splicing of the transcription factors Mef2a and Mef2d to maintain the expression and proper activity of muscle-specific isoforms (Nikonova et al. 2019, 2020). More recently, the generation of skeletal muscle-specific knockout of Rbfox1 and Rbfox2 in adult mice also revealed a role of these alternative splicing factors in the maintenance of skeletal muscle mass. This function could at least be attributable to the altered expression of calpain 3 splice isoforms that are not degraded rapidly leading to the accumulation of active calpain 3 and a defective proteostasis (Singh et al. 2018). As illustrated by this example, a major part of our knowledge has been gleaned from studies in animal models. Given the large number of RBPs expressed and with potential phenotypes in human muscle, this clearly illustrates the need for further studies examining the function of RBPs in human myogenesis.

In this context, human pluripotent stem cells represent an attractive cell source for producing myogenic cells and deciphering the function of different RBPs during this process. Until recently, the only efficient protocols allowing differentiation of human pluripotent stem cells into a reasonably mature muscle cells relied on the forced expression of transcription factors such as MyoD or Pax3/7 (Sato 2020). These direct reprogramming approaches aim to bypass early developmental stages by overexpressing myogenic regulators and the cellular events occurring during differentiation remain largely unclear (Chal and Pourquié 2017). More recently, stepwise induction protocols based on the use of small molecules and growth factors have been established and appeared as an alternative approach to better recapitulate the successive developmental stages involved in skeletal myogenesis (Chal et al. 2015; Caron et al. 2016). By using such approach, our study adds a new dimension to the role of MBNL proteins in defining human skeletal muscle maturation stages. The exact mechanisms by which MBNL proteins affect these processes are however unclear. In contrast to Rbfox1/2 for which depletion has been shown to inhibit myoblast fusion in vitro through the mis-splicing of Mef2D (Pedrotti et al. 2015; Runfola et al. 2015), our results indicate that depletion of MBNL proteins leads to the aberrant expression of hundreds of genes. Since several of these genes encode factors involved in sarcomere organization, actin dynamics and vesicular trafficking events required for membrane fusion (Sampath and Millay 2018), further experiments will be required to evaluate the exact MBNL-based mechanisms required for myogenic differentiation.

Among the three MBNL paralogs, our results confirmed that MBNL1 is the most abundant MBNL protein during in vitro myogenesis (Charizanis et al. 2012; Holt et al. 2009). We also observed a meticulous compensation of MBNL1 by MBNL2 since MBNL2 protein expression was increased in either *MBNL1*(-/-) and *MBNL1*(-/-); *MBNL2*(+/-) compared to WT and *MBNL2*(+/-) hiPSC-derived myotubes respectively. Moreover, transcriptomic analyses showed that several skeletal muscle-

related biological processes were down-regulated in DKO myotubes compared to *MBNL1(-/-)* myotubes. Taking together, our results strengthened this compensation phenomenon between MBNL1 and MBNL2 proteins in myogenesis as suggested in prior studies in mice (Goodwin et al. 2015; Lee et al. 2013). In contrast to MBNL1 and MBNL2, the expression of MBNL3 is found to be very low and limited in duration. The role of MBNL3 during myogenesis is not well understood. Whereas constitutive expression of *Mbnl3* in murine muscle precursors cells appears to inhibit terminal differentiation (Lee et al. 2008), it has been recently shown that *Mbnl3* depleted primary murine myoblasts displayed defective in vitro differentiation (Thomas et al. 2017). Our results showed that triple knockouts hiPSC-derived skeletal muscle cells recapitulated all the phenotypes identified in double knockouts suggesting that MBNL3 might not play such important role during human myogenesis.

Sequestration of MBNL proteins is considered a major player in DM1 pathology. Comparative transcriptomic analysis between *Mbnl1* knockout mice and transgenic mice expressing CUG expansion RNA have shown that the majority of changes induced by the abnormal expansions in skeletal muscle could be explained by reduced activity of *Mbnl1* (Osborne et al. 2009; Wang et al. 2012; Du et al. 2010). Moreover, DKO mice have been shown to exhibit a spliceopathy similar to the one observed in DM1 patients, particularly in patients with congenital form which is the most severe form of the disease associated with large numbers of CTG repeats (Lee et al. 2013; Thomas et al. 2017). In contrast, our study showed that the double loss of MBNL1 and MBNL2 proteins recapitulates approximately 36% of the splicing defects that we observed in skeletal muscle cells derived from DM1 hiPSC carrying more than 2500 CTG expansions. Altogether, these results strongly suggest that additional pathways may play non-negligible roles in the disease. Among the different hypotheses that can be formed, the stabilization and overexpression of CUG-BP1, due to its hyper-phosphorylation as a result of the alteration of the GSK3 β pathway in DM1, have been shown to lead to several splicing defects (Kalsotra et al. 2008; Wang et al. 2015). Furthermore, the splicing factor hnRNP-H found to be deregulated in DM1 may also contribute to the disease spliceopathy (Paul et al. 2006).

Nonetheless, we took advantage of the comparative transcriptomic analyses of MBNL depleted and DM1 hiPSC-derived myotubes to focus our attention on a strongly dysregulated alternative splicing event involving *SORBS1* exon 25 exclusion. Interestingly, this mis-splicing event was already described in databases from cardiac muscle of *Mbnl1(-/-)*; *Mbnl2(+/-)* mice (Lee et al. 2013) or in muscle biopsies from congenital (Nakamori et al. 2017) and adult DM1 patients (Wang et al. 2019). At the functional level, our results showed that missplicing of *SORBS1* exon 25 is sufficient to alter AChR clustering which consequently should result in post-synaptic defects. Interestingly, structural abnormalities in AChR clustering have been previously described in a severe murine model of DM1 (Panaite et al. 2013) as well as in *Mbnl* DKO and TKO mice (Lee et al. 2013; Thomas et al. 2017). Our study showed that the developmental transition of the splicing of *SORBS1* exon 25 occurs very early in human muscle development. Using severe cellular models of the disease either with DM1 hiPSC carrying large number of CTG repeats or with *MBNL1,2* DKO hiPSC-derived skeletal muscle cells, we were able to reproduce the abnormal splicing switch of *SORBS1* exon 25 observed in muscles obtained from congenital DM1 embryos.

However, the mechanisms by which the exclusion of *SORBS1* exon 25 result in altered post-synaptic formation are still not understood and will need further experiments. From a structural point of view, *SORBS1* exon 25 encodes the coiled-coil domain of the protein (Tomasovic et al. 2012) and numerous studies showed the involvement of these domains in oligomerization as well as being able to play the role of “molecular spacers” leading to the formation of large macromolecular complexes (Truebestein and Leonard 2016). Recently, a study has shown that murine *Sorbs1* could interact with

the MuSK/Dok-7/Crk/CrkL complex to enable the formation of AChR clusters (Hallock et al. 2016). In this study, targeted siRNA silencing of *Sorbs1* expression led to a defective AChR clustering in murine C2C12 muscle cells. In the current study, exclusion of *SORBS1* exon 25 did not seem to alter the level of expression of this protein but rather affected its cellular localization. Moreover, we showed that *SORBS1* exon 25 exclusion was sufficient to induce AChR clustering defects in both human primary and hiPSC-derived-myotubes. In conclusion, our results suggest a model in which sequestration of MBNL proteins by expanded CUG repeats in individuals with DM1 leads to *SORBS1* exon 25 skipping. The expression of this isoform ultimately affects the development of the post-synaptic compartment by mechanisms yet to be explored. These findings might be clinically significant for the functional changes in neuromuscular connections associated with DM1.

Finally, to this day, only a couple of studies have described the generation of DM1 hiPSC-derived skeletal muscle cells and the capacity of these cellular systems to reproduce the main molecular features associated with DM1, namely the nuclear aggregation of mutated mRNAs and some alternative splicing defects. Despite promising results, all these attempts to generate skeletal muscle cells from DM1 hiPSCs were based on the forced expression of myogenic transcription factors MyoD or Pax7 which, as discussed above, are often undefined (Mondragon-Gonzalez and Perlingeiro 2018; Ueki et al. 2017; Wang et al. 2018). The experimental system described here uses a transgene-free protocol to generate skeletal muscle cells from hiPSCs and opens new perspectives in term of disease modeling and drug screening. First, the different well-characterized steps of the protocol help to decipher the exact impact of the mutation on myogenic differentiation. Whereas most of the studies have focused on human primary skeletal muscle cells or immortalized skeletal muscle cells originating from DM1 patients' biopsies (Arandel et al. 2017), our results indicate that early myogenic commitment is not affected by the presence of the DM1 mutation. In addition, the transgene-free protocol described here includes a freezing step allowing the possibility to generate large cell banks. Consequently, this represent an unlimited source of material that could be used for drug screening.

In summary, our results provide important new experimental support for the MBNL loss-of-function model for DM1 and introduce new human cellular models to examine the tissue-specific consequences of MBNL loss.

MATERIALS AND METHODS

Cell culture

Human induced pluripotent stem cells were derived by Phenocell® from human primary fibroblasts reprogrammed using sendai virus vectors carrying OCT4, KLF4, SOX2 and c-Myc (Barrault et al. 2019). Two DM1 human induced pluripotent stem cell lines were generated from two independent patients' fibroblasts containing around 2500 CTG repeats. DM1 hiPSC carry a mutant DMPK gene containing approximatively 2700 CTG repeats. One control hiPSC line corresponding to the sibling matched control was also generated. Informed consents were obtained from all the patients included in this study, complying with the ethical guidelines of the institutions and with the legislation requirements of the country of origin. Experimental protocols were approved by the french minister of health (2019-A02599-48). All the hiPSC lines were used at passages around p10-p20 and were maintained in StemMACS iPS-Brew XF medium (Miltenyi Biotec®) in vitronectin (Gibco®) coated culture dishes. Culture medium was changed every two days and confluent cells were manually passaged every 5-7 days.

hiPSCs skeletal muscle differentiation experiments were performed using the commercially available Genea Biocells® skeletal muscle differentiation media following the protocol described previously (Caron et al. 2016). Briefly, hiPSCs were dissociated with StemPro Accutase Cell Dissociation Reagent (Gibco®) and plated at 3000 cells per cm² onto collagen I-coated plates (Gibco®) and maintained for 10 days in skeletal muscle induction medium (SKM01). At day 10, cells were dissociated with Trypsin-EDTA 0,05% (Gibco®), plated at 5000 cells per cm² onto collagen I-coated plates and maintained for 6 days in skeletal myoblast medium (SKM02). At day 16, myoblasts were dissociated and cryopreserved. For terminal differentiation, cells were thawed at 10.000 cells per cm² in SKM02 and after 2-3 days, when cells reached confluence, the medium was replaced by the myotube medium (SKM03). Cells were then maintained for 5 to 6 days in SKM03.

Human primary as well as immortalized WT and DM1 myoblasts were cultivated in a growth medium consisting of DMEM-F12 glutamax (Life Technologies) supplemented with 20% FBS (Sigma-Aldrich) and 0,1% penicillin/streptomycin (Life Technologies). Myogenic differentiation was induced by switching confluent cell cultures to DMEM supplemented with 5% horse serum (Life Technologies).

Control and congenital DM1 muscle samples were obtained from aborted fetuses showing, respectively, no sign of neuromuscular disease (control) and clinical symptoms of congenital DM1 form with large (CTG)_n > 1000 repeats (Rau et al. 2015).

sgRNAs design

sgRNAs and primers used in this study are reported in Table S1.

SpCas9 target sequences within the first coding exon of each *MBNL* paralog were determined by CRISPOR (<http://crispor.tefor.net>). sgRNAs with the highest predicted efficiency, the lowest number of potential off-targets and having an enzymatic restriction site at the PAM sequence were selected. crRNA and tracrRNA were synthesized by Integrated DNA Technologies®.

RNP transfection of hiPSCs

Control hiPS cells were dissociated with StemPro Accutase Cell Dissociation Reagent (Gibco®), plated in 24-well plates at 40.000 cells per cm² and transfected the next day with a mixed of 30nM of sgRNA (by duplexing RNA oligos: crRNA and tracrRNA), 30nM of SpCas9 purified protein (gift from Jean-Paul Concordet, MNHN-CNRS UMR 7196 / INSERM U1154) and Lipofectamine Stem

Transfection Reagent (ThermoFisher Scientific®), according to the manufacturer's protocol.

ASO(s) treatments

Proliferating human primary WT myoblasts or WT hiPSCs-derived myotubes were transfected with 50 nM of 2-OMe ASOs using RNAi Max transfection reagent (Life Technologies) according to the manufacturer's protocol. ASOs used in our study are reported in Table S1. Cells were harvested for analyses 48h after transfection and 6h after Agrin stimulation at 0.5 µg/ml (R&D systems).

Genomic DNA extraction and PCR

Genomic DNA was extracted from transfected hiPS cells either with QIAmp DNA Micro and Mini Kit (Qiagen®) or with QuickExtract DNA extraction solution (Lucigen®), according to the manufacturer's instructions. PCR were performed using Q5 High-Fidelity DNA Polymerase (NEB®) with 30ng gDNA, 38 cycles of 10'' at 98°C, 15'' at 60°C and 45'' at 72°C, with a final 10' extension and with primers framing the target area according to each guide used (see Table S1).

RFLP assays to identify edited clones

Restriction Fragment Length Polymorphism (RFLP) presents the advantage to easily and rapidly estimate the presence of gene editing in a mixed population or in single clones. Briefly, 2µL PCR products were digested by 5 units of restriction enzyme (NEB®): BsrI for MBNL1 and MBNL2 clones, and SmlI for MBNL3 clones. Cleaved fragments were then separated by gel electrophoresis. If CRISPR-mediated gene editing was successful, NHEJ-mediated INDELS abolished the restriction endonuclease recognition site. After restriction digest, clones that had lost their restriction site were selected and corresponding PCR products were sequenced by Sanger DNA sequencing (Eurofins®). TOPO-TA cloning (Invitrogen®) was used according to manufacturer's instructions to check each selected clone by Sanger DNA sequencing (Eurofins®), and the absence of subclones was assessed on at least 30 TA clones for each CRISPR cell line generated.

RT-PCR and Agilent DNA chips analysis for Alternative Splicing

Total RNA was extracted using the RNeasy Micro/Mini kit (Qiagen®) and reverse transcribed using random hexamers and Superscript III Reverse Transcriptase kit (Invitrogen®) according to the manufacturer's protocol. For splicing analysis, PCR amplification were carried out with recombinant Taq DNA polymerase (Invitrogen®) and primers listed in Table S1. The amplification was performed using a first step at 95°C for 3' followed by 35 cycles of 45'' at 95°C, 30'' at 60°C, 1' at 72°C, and a final 10' extension at 72°C. The PCR products were analyzed using Agilent® DNA chips and quantified with the BioAnalyzer 2100. The percent spliced-in (PSI) value correspond to the fraction of mRNAs that contains an exon and was calculated as the ratio of the density of the exon inclusion band to the sum of the densities of inclusion and exclusion bands, expressed as a percentage.

Gene expression analysis by quantitative PCR

Total RNA was extracted using the RNeasy Micro/Mini kit (Qiagen®) and was reverse transcribed using random hexamers and the Superscript III Reverse Transcriptase kit (Invitrogen®). Quantitative PCR reactions were carried out in 384-well plates using a QuantStudio 12K Flex Real-Time PCR System (Applied Biosystems®) with Power SYBR Green 2× Master Mix (Life Technologies®), 2.5 µl of cDNA and primers (Table S1) in a final volume of 10 µL. The $2^{-\Delta\Delta C_t}$ method was used to determine the relative expression level of each gene.

Fluorescent In situ Hybridization

Cell preparation was performed as previously described (Maury et al. 2019). Briefly, cells were fixed with PBS buffer solution containing 4% PFA for 8' at room temperature and incubated overnight at 4°C with 70% Ethanol (VWR®). After washes, cells were rehydrated with a solution of 5mM of MgCl₂ (Sigma-Aldrich®) for 10' at room temperature and then sequentially incubated with prehybridization buffer for 10' and hybridization buffer containing the probe overnight at 37°C under agitation. Washing steps consisted in pre-warmed washing buffer for 1 hour at 37°C. Finally, cells were washed twice in PBS buffer and incubated 5' in a Hoechst 33528 solution (Sigma-Aldrich®). Nuclear foci detection was analyzed using the Cellomics CX7 high content imaging system (ThermoFisher Scientific®) with the 20X objective and automatic focus on cells. Images were acquired in High resolution camera mode on two channels and quantification was performed as previously described with the software attached (Maury et al. 2019). Each image is a maximum projection of Z-stack acquisition.

Immunocytochemistry

After fixation with 4% PFA for 10' at room temperature, cells were incubated over night at 4°C with the following primary antibodies: DESMIN (R&D System®, 1:100), PAX3 (DSHB®, 1:30), α -ACTININ (Sigma®, 1:250), MF20 (DSHB®, 1:200), MBNL1 - MB1a(4A8) (DSHB®, 1:100), MBNL2 - MB2a(3B4) (DSHB®, 1:100), SORBS1 (Abcam®, 1:400), AChR (DSHB®, 1:100). Appropriated Alexa fluorescent secondary antibodies (Invitrogen®, 1:1000) (Table S8) and Hoechst (5 μ g/mL) were added for 1h. Every antibody used in our study is listed in Table S8. Image acquisitions were performed on the ImageXpress micro imaging system (Molecular Devices®) with the 20X objective and with three channels. Images were analyzed using MetaXpress software by quantifying the percentage of cells stained for PAX3 or DESMIN compared to the total number of cells.

Quantification of *in vitro* myoblasts fusion was performed counting the number of myotubes with at least 3 nuclei (manually counted) normalized to the total number of nuclei (at least 11.000 nuclei counted per experiment and per cell line).

AChR clusters morphometric analyses were performed using Fiji Software and using an exclusion threshold of 5 μ m².

SORBS1 cellular localization in human primary myotubes was performed by confocal microscopy using the Zeiss® LSM-880 microscope with the maximum intensity projection of acquired Z-stacks. Images of SORBS1 cellular localization in hiPSC-derived myotubes were acquired using Spinning Disk microscopy (Zeiss®) using the 20X objective and with the maximum intensity projection of acquired Z-stacks.

Protein extraction and Western blot analysis

Western blots analyses were performed as previously described (Maury et al. 2019). Briefly, cells were lysed in RIPA 1X buffer (Sigma®) containing protease inhibitors (Sigma®) and phosphatase inhibitors (Roche®). Proteins were quantified by Pierce BCA Protein Assay kit (Pierce®). Protein extracts (15 to 20 μ g) were loaded on a 4-12% SDS-PAGE gradient (NuPage Bis-Tris gels, Invitrogen®) and transferred onto Gel Transfer Stacks Nitrocellulose membranes (Invitrogen®) using the iBlot2 Dry Blotting System (Invitrogen®). Membranes were then incubated overnight at 4°C with the following primary antibodies: MBNL1 - MB1a(4A8) (DSHB®, 1:1000), MBNL2 - MB2a(3B4) (DSHB®, 1:1000), MBNL3 - 5A11(sc-136168) (Santa Cruz, 1:100). After hybridization of the peroxidase-conjugated secondary antibody (1:10000), immunoreactive bands were revealed by using Amersham ECL Select Western Blotting Detection Reagents (GE Healthcare®). Equal protein

loading was verified by the detection of β -Actin using the A3854 Monoclonal Anti- β -Actin–Peroxidase antibody (Sigma®, 1:10000).

Statistical analysis

All data were processed using Prism 8®. Values are represented as mean \pm SD. For comparisons of more than two groups, statistical analyses were performed with by ordinary one-way analysis of variance (ANOVA) using Tukey's multiple comparisons test. Values of $p < 0,05$ were considered significant (* $p < 0,05$; ** $p < 0,005$; *** $p < 0,0005$; **** $p < 0,0001$).

AUTHOR CONTRIBUTIONS

A.M., M.C. and C.M.: conception and design. A.M., J.T., G.M. and C.M.: investigation and data collection. A.M., J.T., J.B.D., H.P., M.J., J.P., C.L. and C.M.: formal analysis and interpretation. D.F. and C.M.: provision of different cellular resources. C.M.: financial support. A.M. and C.M.: manuscript writing.

DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST

The authors declare no conflicts of interest

ACKNOWLEDGEMENTS

I-Stem is part of the Biotherapies Institute for Rare Diseases (BIRD) supported by the Association Française contre les Myopathies (AFM-Téléthon). This project was supported by grants from Laboratoire d'Excellence Revive (Investissement d'Avenir; ANR-10-LABX-73) and from the doctoral school « Innovation Thérapeutique, du fondamental à l'appliqué » (ED 569) from Paris Saclay University. We thank Geneviève Gourdon's team for the southern blot analyses of the CTG repeats length. We gratefully acknowledge support from the PSMN (Pôle Scientifique de Modélisation Numérique) of the ENS of Lyon for computing resource. We thank Lina El-Kassar and Karine Giraud-Triboult for karyotyping the cell lines.

REFERENCES

- Arandel L, Polay Espinoza M, Matloka M, Bazinet A, De Dea Diniz D, Naouar N, Rau F, Jollet A, Edom-Vovard F, Mamchaoui K, et al. 2017. Immortalized human myotonic dystrophy muscle cell lines to assess therapeutic compounds. *Dis Model Mech* **10**: 487–497. <http://dmm.biologists.org/lookup/doi/10.1242/dmm.027367>.
- Barrault L, Gide J, Qing T, Lesueur L, Tost J, Denis JA, Cailleret M, Aubry L, Peschanski M, Martinat C, et al. 2019. Expression of miRNAs from the Imprinted DLK1/DIO3 Locus Signals the Osteogenic Potential of Human Pluripotent Stem Cells. *Cells* **8**: 1523. <https://www.mdpi.com/2073-4409/8/12/1523> (Accessed October 12, 2020).
- Bertz M, Wilmanns M, Rief M. 2009. The titin-telethonin complex is a directed, superstable molecular bond in the muscle Z-disk. *Proc Natl Acad Sci U S A* **106**: 13307–13310. <http://www.ncbi.nlm.nih.gov/pubmed/19622741> (Accessed October 12, 2020).
- Brook JD, Mccurrach ME, Harley HG, Buckler AJ, Church D, Aburatani H, Hunter K, Stanton VP, Thirion J-P, Hudson T, et al. 1992. Molecular Basis of Myotonic Dystrophy: Expansion of a Trinucleotide (CTG) Repeat at the 3' End of a Transcript Encoding a Protein Kinase Family Member. *Cell* **68**: 799–808. https://ac-els-cdn-com.gate2.inist.fr/0092867492901545/1-s2.0-0092867492901545-main.pdf?_tid=2610cd5e-6d4b-4b0f-9e07-37ef57f9e87d&acdnat=1524067745_836f90cc2a10d490db94052fee90f544 (Accessed April 18, 2018).
- Caron L, Kher D, Lee KL, McKernan R, Dumevska B, Hidalgo A, Li J, Yang H, Main H, Ferri G, et al. 2016. A Human Pluripotent Stem Cell Model of Facioscapulohumeral Muscular Dystrophy-Affected Skeletal Muscles. *Stem Cells Transl Med* **5**: 1145–1161. <http://doi.wiley.com/10.5966/sctm.2015-0224>.
- Chal J, Oginuma M, Al Tanoury Z, Gobert B, Sumara O, Hick A, Bousson F, Zidouni Y, Mursch C, Moncuquet P, et al. 2015. Differentiation of pluripotent stem cells to muscle fiber to model Duchenne muscular dystrophy. *Nat Biotechnol* **33**: 962–9. <http://www.ncbi.nlm.nih.gov/pubmed/26237517> (Accessed October 22, 2020).
- Chal J, Pourquié O. 2017. Making muscle: skeletal myogenesis in vivo and in vitro. *Development* **144**: 2104–2122. <http://www.ncbi.nlm.nih.gov/pubmed/28634270> (Accessed October 22, 2020).
- Charizanis K, Lee K-Y, Batra R, Goodwin M, Zhang C, Yuan Y, Shiue L, Cline M, Scotti MM, Xia G, et al. 2012. Muscleblind-like 2-Mediated Alternative Splicing in the Developing Brain and Dysregulation in Myotonic Dystrophy. *Neuron* **75**: 437–450. <http://dx.doi.org/10.1016/j.neuron.2012.05.029> (Accessed April 13, 2018).
- Choi J, Dixon DM, Dansithong W, Abdallah WF, Roos KP, Jordan MC, Trac B, Lee HS, Comai L, Reddy S. 2016. Muscleblind-like 3 deficit results in a spectrum of age-associated pathologies observed in myotonic dystrophy. *Sci Rep* **6**. <http://www.nature.com.gate2.inist.fr/articles/srep30999.pdf> (Accessed May 29, 2018).
- Concordet J-P, Haeussler M. 2018. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res* **46**: W242. <https://www.ncbi.nlm.nih.gov.proxy.insermbiblio.inist.fr/pmc/articles/PMC6030908/> (Accessed May 6, 2020).
- Cooper TA, Wan L, Dreyfuss G. 2009. RNA and disease. *Cell* **136**: 777–93. <http://www.ncbi.nlm.nih.gov/pubmed/19239895> (Accessed October 12, 2020).
- Du H, Cline MS, Osborne RJ, Tuttle DL, Clark TA, Donohue JP, Hall MP, Shiue L, Swanson MS, Thornton CA, et al. 2010. Aberrant alternative splicing and extracellular matrix gene expression in mouse models of myotonic dystrophy. *Nat Struct Mol Biol* **17**: 187–193. <http://www.nature.com/doifinder/10.1038/nsmb.1720>.
- Fardaei M, Rogers MT, Thorpe HM, Larkin K, Hamshere MG, Harper PS, Brook JD. 2002. Three proteins, MBNL, MBLL and MBXL, co-localize in vivo with nuclear foci of expanded-repeat transcripts in DM1 and DM2 cells. *Hum Mol Genet* **11**: 805–814. https://watermark.silverchair.com/ddf079.pdf?token=AQECAHi208BE49Oan9kkhW_Ercy7Dm3ZL_9Cf3qfKAc485ysgAAAaUwggGhBgkqhkiG9w0BBwagggGSMIIBjgIBADCCAYcGCSqGSIb3DQEĤATAeBglghkgBZQMEAS4wEQQM0KKYFTmScxhh4eL5AgEQgIIBWAOlh0dGgmlbQhKRTGVYs xhudPD-ho34CsXKuFATtw3wwiFm (Accessed April 18, 2018).
- Fugier C, Klein A, Hammer C, Vassilopoulos S, Ivarsson Y, Toussaint A, Tosch V, Vignaud A, Ferry A, Messaddeq N, et al. 2011. Misregulated alternative splicing of BIN1 is associated with T tubule alterations and muscle weakness in myotonic dystrophy. *Nat Med* **17**: 720–726. <http://www.nature.com.gate2.inist.fr/articles/nm.2374.pdf> (Accessed May 25, 2018).
- Goodwin M, Mohan A, Batra R, Lee KY, Charizanis K, Fernández Gómez FJ, Eddarkaoui S, Sergeant N, Buée L, Kimura T, et al. 2015. MBNL Sequestration by Toxic RNAs and RNA Misprocessing in the

- Myotonic Dystrophy Brain. *Cell Rep* **12**: 1159–1168.
- Hallock PT, Chin S, Blais S, Neubert TA, Glass DJ. 2016. Sorbs1 and -2 Interact with CrkL and Are Required for Acetylcholine Receptor Cluster Formation. *Mol Cell Biol* **36**: 262–70. <http://www.ncbi.nlm.nih.gov/pubmed/26527617> (Accessed February 12, 2021).
- Holt I, Jacquemin V, Fardaei M, Sewry CA, Butler-Browne GS, Furling D, Brook JD, Morris GE. 2009. Muscleblind-Like Proteins: Similarities and Differences in Normal and Myotonic Dystrophy Muscle. *Am J Pathol* **174**: 216–227. <https://www.ncbi.nlm.nih.gov.gate2.inist.fr/pmc/articles/PMC2631334/pdf/JPATH174000216.pdf> (Accessed April 23, 2018).
- Kalsotra A, Xiao X, Ward A, Castle J, Johnson J, Burge C, Cooper T. 2008. A Postnatal Switch of CELF and MBNL Proteins Reprograms Alternative Splicing in the Developing Heart. *Proc Natl Acad Sci U S A* **105**. <https://pubmed.ncbi.nlm.nih.gov/19075228/> (Accessed May 6, 2020).
- Kanadia RN, Johnstone KA, Mankodi A, Lungu C, Thornton CA, Esson D, Timmers AM, Hauswirth WW, Swanson MS. 2003a. A Muscleblind Knockout Model for Myotonic Dystrophy. *Science (80-)* **302**: 1978–1980.
- Kanadia RN, Urbinati CR, Crusselle VJ, Luo D, Lee Y-J, Harrison JK, Oh SP, Swanson MS. 2003b. Developmental expression of mouse muscleblind genes Mbnl1, Mbnl2 and Mbnl3. *Gene Expr Patterns* **3**: 459–462. https://ac-els-cdn-com.gate2.inist.fr/S1567133X03000644/1-s2.0-S1567133X03000644-main.pdf?_tid=a06156ea-9fc3-4c79-8f92-3047d6c51d88&acdnt=1524064722_be864982eae8cc343017e534614e76b6 (Accessed April 18, 2018).
- Koebis M, Ohsawa N, Kino Y, Sasagawa N, Nishino I, Ishiura S. 2011. Alternative splicing of myomesin 1 gene is aberrantly regulated in myotonic dystrophy type 1. *Genes to cells* **16**: 961–72. <http://www.ncbi.nlm.nih.gov/pubmed/21794030> (Accessed September 30, 2020).
- Konieczny P, Stepniak-Konieczna E, Sobczak K. 2014. MBNL proteins and their target RNAs, interaction and splicing regulation. *Nucleic Acids Res* **42**: 10873–10887. https://watermark.silverchair.com/gku767.pdf?token=AQECAHi208BE49Ooan9kkhW_Ercy7Dm3ZL_9Cf3qfKAc485ysgAAAagwggGkBgkqhkiG9w0BBwagggGVMiIBkQIBADCCAYoGCSqGSib3DQE_HATAeBglghkgBZQMEAS4wEQQMZEYyRkcyW2DSPJPTAgEQgIIBW6jN0x7jcc0aSKvbQMdVx_a2jK_3UjWJvfAqJaoMO9re4F4Q (Accessed April 18, 2018).
- Lee K-S, Smith K, Amieux PS, Wang EH. 2008. MBNL3/CHCR prevents myogenic differentiation by inhibiting MyoD- dependent gene transcription. *Differentiation* **76**: 299–309. https://ac-els-cdn-com.gate2.inist.fr/S0301468109600748/1-s2.0-S0301468109600748-main.pdf?_tid=f5202fe7-d9ae-4617-8501-2bef96febb94&acdnt=1527600092_2cb7dc27679a535fde293b8876cbea5c (Accessed May 29, 2018).
- Lee KY, Li M, Manchanda M, Batra R, Charizanis K, Mohan A, Warren SA, Chamberlain CM, Finn D, Hong H, et al. 2013. Compound loss of muscleblind-like function in myotonic dystrophy. *EMBO Mol Med* **5**: 1887–1900.
- Li W, Turner A, Aggarwal P, Matter A, Storvick E, Arnett DK, Broeckel U. 2015. Comprehensive evaluation of AmpliSeq transcriptome, a novel targeted whole transcriptome RNA sequencing methodology for global gene expression analysis. *BMC Genomics* **16**: 1069. <http://www.biomedcentral.com/1471-2164/16/1069> (Accessed October 12, 2020).
- Lin X, Miller JW, Mankodi A, Kanadia RN, Yuan Y, Moxley RT, Swanson MS, Thornton CA. 2006. Failure of MBNL1-dependent post-natal splicing transitions in myotonic dystrophy. *Hum Mol Genet* **15**: 2087–2097.
- Llorian M, Smith CWJ. 2011. Decoding muscle alternative splicing. *Curr Opin Genet Dev* **21**: 380–7. <http://www.ncbi.nlm.nih.gov/pubmed/21514141> (Accessed October 22, 2020).
- Mahadevan M, Tsilfidis C, Sabourin L, Shutler G, Amemiya C, Jansen G, Neville C, Narang M, Barceló J, O’Hoy K, et al. 1992. Myotonic dystrophy mutation: an unstable CTG repeat in the 3’ untranslated region of the gene. *Science (80-)* **255**: 1253–1255. <http://science.sciencemag.org.gate2.inist.fr/content/sci/255/5049/1253.full.pdf> (Accessed April 18, 2018).
- Markert CD, Meaney MP, Voelker KA, Grange RW, Dalley HW, Cann JK, Ahmed M, Bishwokarma B, Walker SJ, Yu SX, et al. 2010. Functional muscle analysis of the Tcap knockout mouse. *Hum Mol Genet* **19**: 2268–83. <http://www.ncbi.nlm.nih.gov/pubmed/20233748> (Accessed October 12, 2020).
- Maury Y, Poydenot P, Brinon B, Lesueur L, Gide J, Roquevière S, Côme J, Polvèche H, Auboeuf D, Alexandre Denis J, et al. 2019. Pluripotent Stem Cell-Based Drug Screening Reveals Cardiac Glycosides as Modulators of Myotonic Dystrophy Type 1. *iScience* **11**: 258–271. <http://www.ncbi.nlm.nih.gov/pubmed/30639849> (Accessed October 12, 2020).

- Merkin J, Russell C, Chen P, Burge CB. 2012. Evolutionary dynamics of gene and isoform regulation in Mammalian tissues. *Science* **338**: 1593–9. <http://www.ncbi.nlm.nih.gov/pubmed/23258891> (Accessed October 12, 2020).
- Miller JW, Urbinati CR, Thornton CA, Swanson MS. 2000. Recruitment of human muscleblind proteins to (CUG)_n expansions associated with myotonic dystrophy. *EMBO J* **19**: 4439–4448.
- Mondragon-Gonzalez R, Perlingeiro RCR. 2018. Recapitulating muscle disease phenotypes with myotonic dystrophy 1 induced pluripotent stem cells: a tool for disease modeling and drug discovery. *Dis Model Mech* **11**. <http://www.ncbi.nlm.nih.gov/pubmed/29898953> (Accessed October 22, 2020).
- Nakamori M, Hamanaka K, Thomas JD, Wang ET, Hayashi YK, Takahashi MP, Swanson MS, Nishino I, Mochizuki H. 2017. Aberrant Myokine Signaling in Congenital Myotonic Dystrophy. *Cell Rep* **21**: 1240–1252. <http://www.ncbi.nlm.nih.gov/pubmed/29091763> (Accessed September 26, 2019).
- Nakamori M, Kimura T, Fujimura H, Takahashi MP, Sakoda S. 2007. Altered mRNA splicing of dystrophin in type 1 myotonic dystrophy. *Muscle Nerve* **36**: 251–257. <http://doi.wiley.com/10.1002/mus.20809> (Accessed September 19, 2018).
- Nakamori M, Sobczak K, Puwanant A, Welle S, Eichinger K, Pandya S, Dekdebrun J, Heatwole CR, McDermott MP, Chen T, et al. 2013. Splicing biomarkers of disease severity in myotonic dystrophy. *Ann Neurol* **74**: 862–872. <http://doi.wiley.com/10.1002/ana.23992> (Accessed May 6, 2020).
- Nikonova E, Kao S-Y, Ravichandran K, Wittner A, Spletter ML. 2019. Conserved functions of RNA-binding proteins in muscle. *Int J Biochem Cell Biol* **110**: 29–49. <http://www.ncbi.nlm.nih.gov/pubmed/30818081> (Accessed October 22, 2020).
- Nikonova E, Kao S-Y, Spletter ML. 2020. Contributions of alternative splicing to muscle type development and function. *Semin Cell Dev Biol* **104**: 65–80. <https://www.sciencedirect.com/science/article/pii/S1084952119301260> (Accessed October 22, 2020).
- Osborne RJ, Lin X, Welle S, Sobczak K, O'Rourke JR, Swanson MS, Thornton CA. 2009. Transcriptional and post-transcriptional impact of toxic RNA in myotonic dystrophy. *Hum Mol Genet* **18**: 1471–1481.
- Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet* **40**: 1413–1415. <http://www.nature.com/articles/ng.259> (Accessed October 12, 2020).
- Panaite P-A, Kuntzer T, Gourdon G, Lobrinus JA, Barakat-Walter I. 2013. Functional and histopathological identification of the respiratory failure in a DMSXL transgenic mouse model of myotonic dystrophy. *Dis Model Mech* **6**: 622–31. <http://www.ncbi.nlm.nih.gov/pubmed/23180777> (Accessed February 11, 2021).
- Paul S, Dansithong W, Kim D, Rossi J, Webster NJG, Comai L, Reddy S. 2006. Interaction of muscleblind, CUG-BP1 and hnRNP H proteins in DM1-associated aberrant IR splicing. *EMBO J* **25**: 4271–83. <http://www.ncbi.nlm.nih.gov/pubmed/16946708> (Accessed February 17, 2021).
- Pedrotti S, Giudice J, Dagnino-Acosta A, Knoblauch M, Singh RK, Hanna A, Mo Q, Hicks J, Hamilton S, Cooper TA. 2015. The RNA-binding protein Rbfox1 regulates splicing required for skeletal muscle structure and function. *Hum Mol Genet* **24**: 2360–74. <http://www.ncbi.nlm.nih.gov/pubmed/25575511> (Accessed October 22, 2020).
- Poulos MG, Batra R, Li M, Yuan Y, Zhang C, Darnell RB, Swanson MS. 2013. Progressive impairment of muscle regeneration in muscleblind-like 3 isoform knockout mice. *Hum Mol Genet* **22**: 3547–3558.
- Prill K, Carlisle C, Stannard M, Windsor Reid PJ, Pilgrim DB. 2019. Myomesin is part of an integrity pathway that responds to sarcomere damage and disease ed. S.J. Du. *PLoS One* **14**: e0224206. <https://dx.plos.org/10.1371/journal.pone.0224206> (Accessed October 12, 2020).
- Ramakrishna S, Kwaku Dad A-B, Beloor J, Gopalappa R, Lee S-K, Kim H. 2014. Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. *Genome Res* **24**: 1020–7. <http://www.ncbi.nlm.nih.gov/pubmed/24696462> (Accessed May 6, 2020).
- Rau F, Lainé J, Ramanoudjame L, Ferry A, Arandel L, Delalande O, Jollet A, Dingli F, Lee KY, Peccate C, et al. 2015. Abnormal splicing switch of DMD's penultimate exon compromises muscle fibre maintenance in myotonic dystrophy. *Nat Commun* **6**.
- Runfola V, Sebastian S, Dilworth FJ, Gabellini D. 2015. Rbfox proteins regulate tissue-specific alternative splicing of Mef2D required for muscle differentiation. *J Cell Sci* **128**: 631–7. <http://www.ncbi.nlm.nih.gov/pubmed/25609712> (Accessed October 22, 2020).
- Sampath SC, Millay DP. 2018. Myoblast fusion confusion: the resolution begins. *Skelet Muscle* **8**: 3. <http://www.ncbi.nlm.nih.gov/pubmed/29386054> (Accessed October 22, 2020).
- Sato T. 2020. Induction of Skeletal Muscle Progenitors and Stem Cells from human induced Pluripotent Stem Cells. *J Neuromuscul Dis* **7**: 395–405. <https://www.medra.org/servlet/aliasResolver?alias=iospress&doi=10.3233/JND-200497> (Accessed

October 22, 2020).

- Singh RK, Kolonin AM, Fiorotto ML, Cooper TA. 2018. Rbfox-Splicing Factors Maintain Skeletal Muscle Mass by Regulating Calpain3 and Proteostasis. *Cell Rep* **24**: 197–208. <http://www.ncbi.nlm.nih.gov/pubmed/29972780> (Accessed October 22, 2020).
- Thomas JD, Sznajder ŁJ, Bardhi O, Aslam FN, Anastasiadis ZP, Scotti MM, Nishino I, Nakamori M, Wang ET, Swanson MS. 2017. Disrupted prenatal RNA processing and myogenesis in congenital myotonic dystrophy. *Genes Dev* **31**: 1122–1133. <http://genesdev.cshlp.org/content/31/11/1122.full.pdf> (Accessed May 29, 2018).
- Tomasovic A, Kurrle N, Banning A, Tikkanen R. 2012. *Role of Cbl-associated protein/ponsin in receptor tyrosine kinase signaling and cell adhesion.* <https://www.jmolbiochem.com/index.php/JmolBiochem/article/view/60/35> (Accessed January 21, 2021).
- Truebestein L, Leonard TA. 2016. Coiled-coils: The long and short of it. *Bioessays* **38**: 903–16. <http://www.ncbi.nlm.nih.gov/pubmed/27492088> (Accessed January 24, 2021).
- Ueki J, Nakamori M, Nakamura M, Nishikawa M, Yoshida Y, Tanaka A, Morizane A, Kamon M, Araki T, Takahashi MP, et al. 2017. Myotonic dystrophy type 1 patient-derived iPSCs for the investigation of CTG repeat instability. *Sci Rep* **7**: 1–7.
- Venables JP, Lapasset L, Gadea G, Fort P, Klinck R, Irimia M, Vignal E, Thibault P, Prinos P, Chabot B, et al. 2013. MBNL1 and RBFOX2 cooperate to establish a splicing programme involved in pluripotent stem cell differentiation. *Nat Commun* **4**: 2480. <http://www.nature.com/articles/ncomms3480> (Accessed October 12, 2020).
- Wang ET, Cody NAL, Jog S, Biancolella M, Wang TT, Treacy DJ, Luo S, Schroth GP, Housman DE, Reddy S, et al. 2012. Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. *Cell* **150**: 710–724.
- Wang ET, Sandberg R, Luo S, Khrebtkova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**: 470–476.
- Wang ET, Treacy D, Eichinger K, Struck A, Estabrook J, Olafson H, Wang TT, Bhatt K, Westbrook T, Sedehizadeh S, et al. 2019. Transcriptome alterations in myotonic dystrophy skeletal muscle and heart. *Hum Mol Genet* **28**: 1312–1321. <http://www.ncbi.nlm.nih.gov/pubmed/30561649> (Accessed October 22, 2020).
- Wang ET, Ward AJ, Cherone JM, Giudice J, Wang TT, Treacy DJ, Lambert NJ, Freese P, Saxena T, Cooper TA, et al. 2015. Antagonistic regulation of mRNA expression and splicing by CELF and MBNL proteins. *Genome Res* **25**: 858–871. <https://genome-cshlp-org.gate2.inist.fr/content/25/6/858.full.pdf> (Accessed April 18, 2018).
- Wang Y, Wang Z, Sun H, Shi C, Yang J, Liu Y, Liu H, Zhang S, Zhang L, Xu Y, et al. 2018. Generation of induced pluripotent stem cell line(ZZUi006-A)from a patient with myotonic dystrophy type 1. *Stem Cell Res* **32**: 61–64. <https://www.sciencedirect.com/science/article/pii/S1873506118302058> (Accessed October 26, 2020).

FIGURES

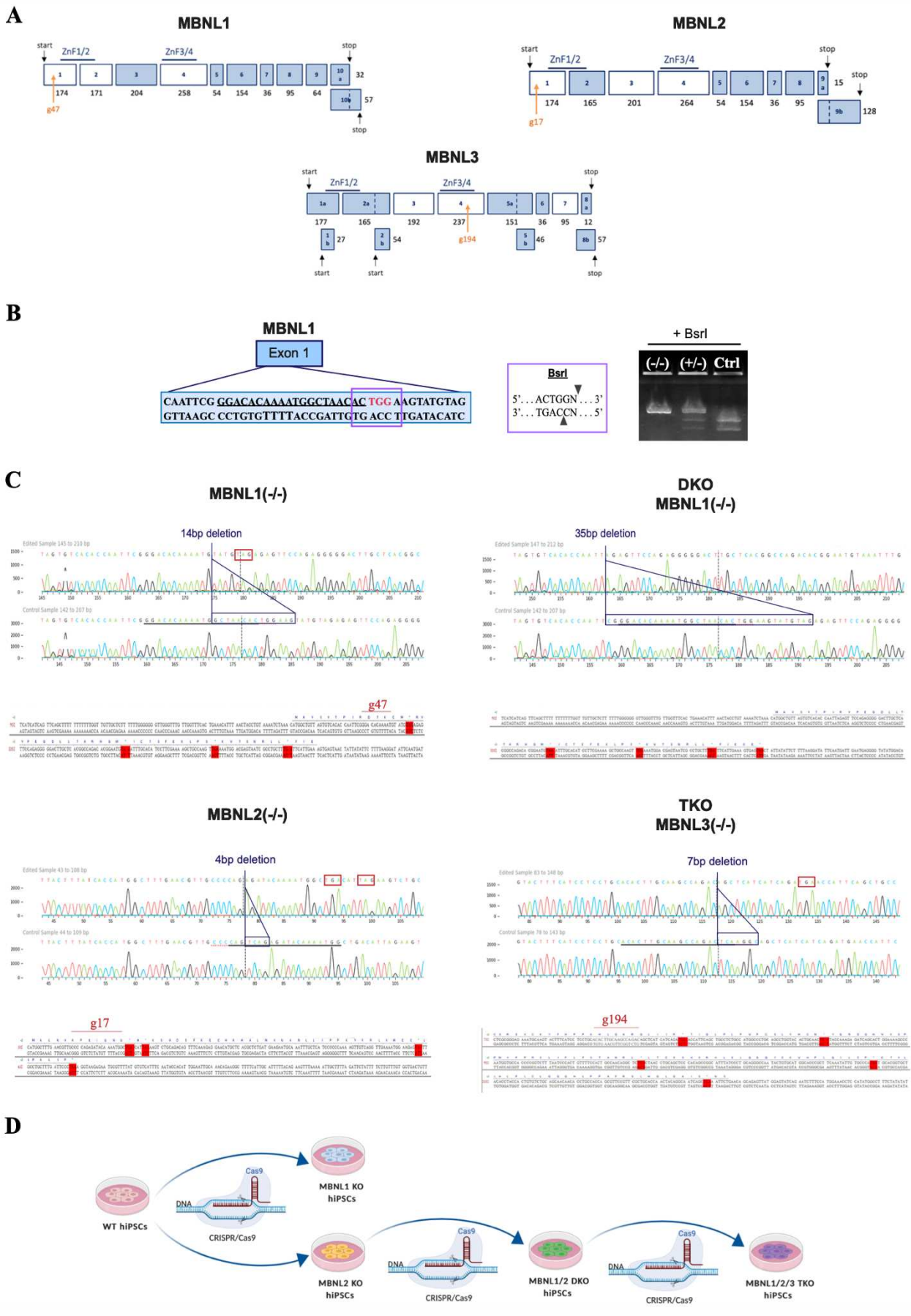


Figure 1

Figure 1: Generation of MBNL Knockout hiPSCs by CRISPR/Cas9

(A) Schematic representation of the different exons encoding Human MBNL1, MBNL2 and MBNL3. Alternative exons are indicated in blue and the targeted site of each guide-RNA is indicated by an orange arrow. Exon size in nucleotides is shown below each box. **(B)** Edited clones were first identified by RFLP (Restriction Fragment Length Polymorphism). The on-target site is underlined and the PAM sequence is in red. The restriction endonuclease recognition site (BsrI in this case) is framed in purple. After locus-specific PCR and restriction digest, cleaved fragments are separated by gel electrophoresis into readily distinguishable profiles, allowing an efficient and rapid screening of single clones to further analyze by sequencing. **(C)** Examples of sequencing of the targeting site in Human MBNL1, MBNL2 and MBNL3 genes after NHEJ-mediated knockout using the CRISPR/Cas9 system in WT hiPSCs. The MBNL1(-/-) clone harbors a deletion of 14bp; MBNL2(-/-) a deletion of 4bp; MBNL1(-/-); MBNL2(-/-) (DKO) a deletion of 35bp and MBNL1(-/-); MBNL2(-/-); MBNL3(-/-) (TKO) a deletion of 7bp. The DKO clone was generated from MBNL2(-/-) clone and the TKO clone from the DKO clone. STOP codons induced by INDELS are framed in red. **(D)** Schematic representation of the different gene editing experiments performed to obtain MBNL knockout hiPSCs clones. Created with Biorender.com

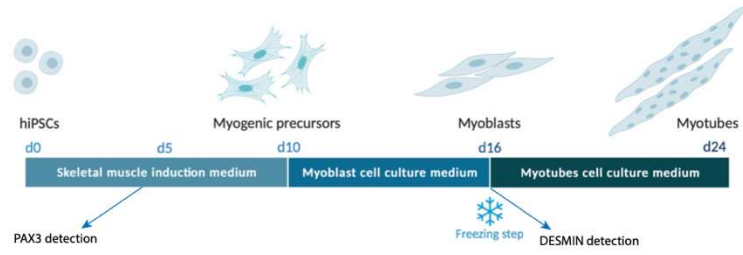
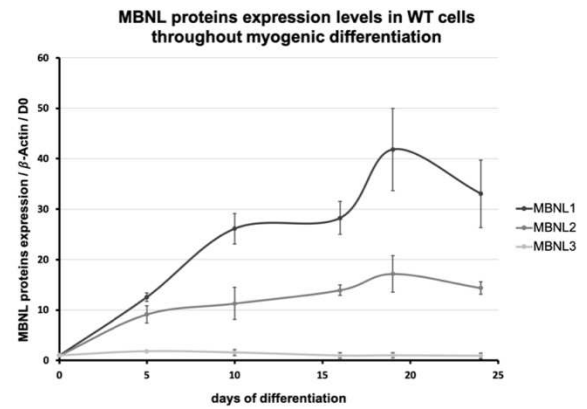
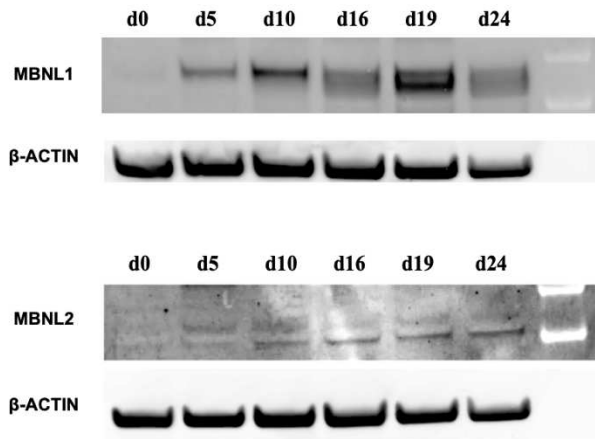
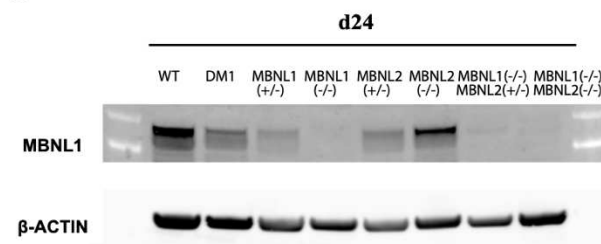
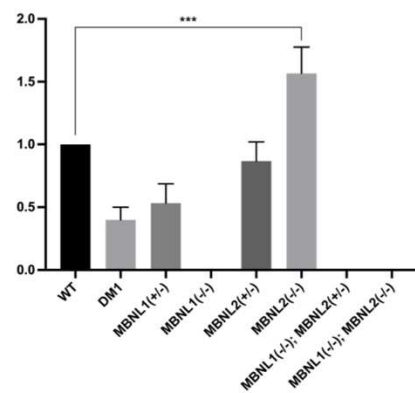
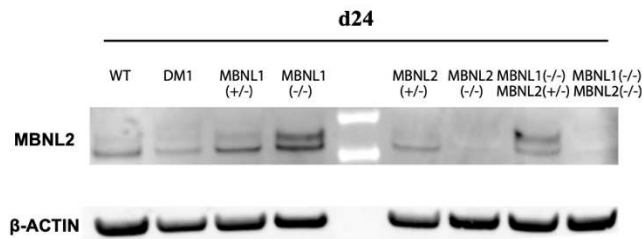
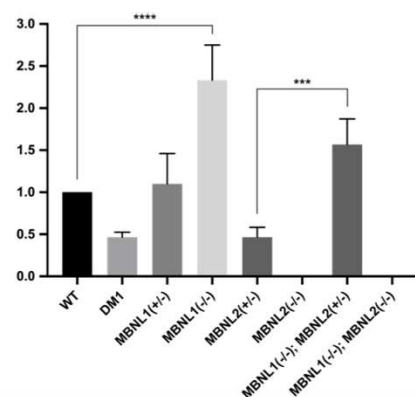
A**B****C****MBNL1 protein expression in myotubes (d24)****D****MBNL2 protein expression in myotubes (d24)****Figure 2**

Figure 2: Expression profile of the three MBNL paralogs during myogenic differentiation from hiPSCs

(A) Schematic representation of the differentiation protocol used to convert hiPSCs into skeletal muscle cells. Created with Biorender.com **(B)** Representative examples of western blot analysis for MBNL1 and MBNL2 at different time points of the myogenic differentiation. The graph represents the kinetics of expression of the three MBNL paralogs along the process of differentiation of WT hiPSCs. Western blot analysis of MBNL3 can be found in Supplementary Figure 2B. β -Actin was used as housekeeping protein for normalization. Data represent the mean \pm SD values from 2 independent experiments. **(C-D)** Example of a western blot analysis for MBNL1 **(C)** and MBNL2 **(D)** proteins after 24 days of differentiation (myotube stage). The graphs on the right panel represent the quantification of MBNL proteins expression in the different conditions compared to WT. β -Actin was used as housekeeping protein for normalization. Data represent means \pm SD values from 3 independent experiments (*, $P < 0.05$; **, $P < 0.005$; ***, $P < 0.0005$; ****, $P < 0.0001$ - Ordinary One-Way ANOVA, Tukey's multiple comparisons test).

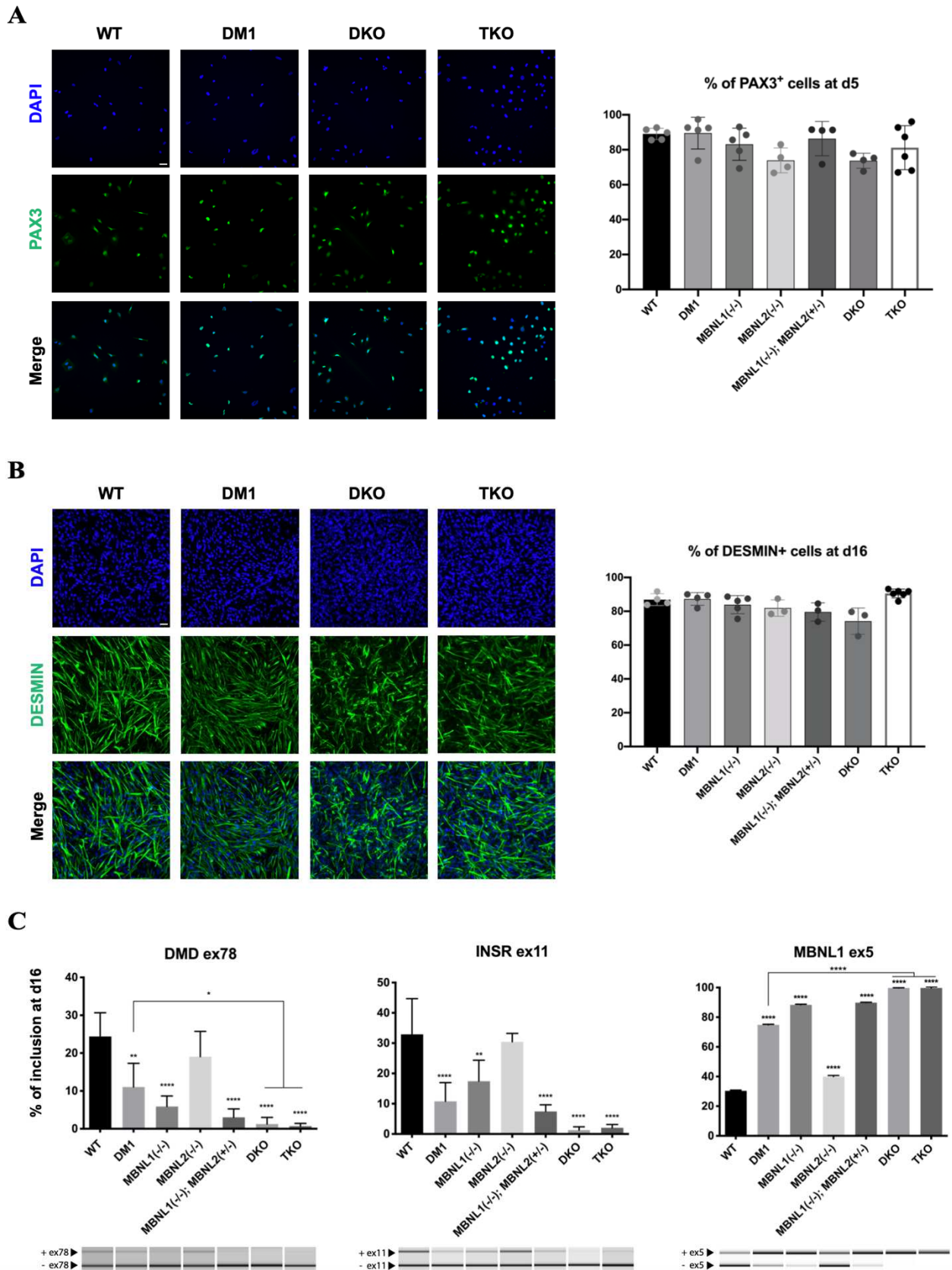


Figure 3

Figure 3: No impairment of the early myogenic commitment in the absence of the MBNL proteins

Representative images of immunolabeling for PAX3 (**A**) and DESMIN (**B**) on hiPSCs differentiated for 5 and 16 days respectively (PAX3 and DESMIN are in green, DAPI nuclear staining is in blue). Scale bar 10 μ M. The graphs on the left represent the percentage of PAX3⁺ or DESMIN⁺ cells, after automated quantification. Data represent the mean \pm SD values from at least 3 independent experiments in quadruplicate. (**C**) Alternative splicing analysis for *DMD* exon 78, *INSR* exon 11 and *MBNL1* exon 5 in hiPSCs-derived myoblasts (16 days of differentiation). Data represent the mean \pm SD values from 3 independent experiments in duplicate (*, P < 0.05; **, P < 0.005; ***, P < 0.0005; ****, P < 0.0001 - Ordinary One-Way ANOVA, Tukey's multiple comparisons test). Asterisks placed above individual bar correspond to the comparison with WT.

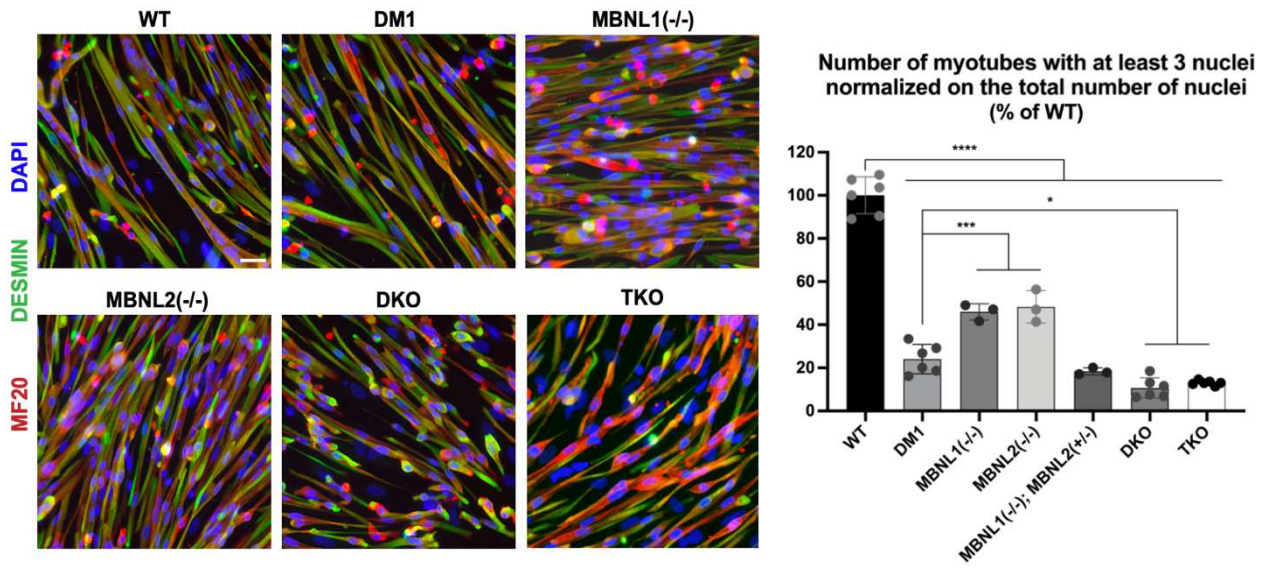
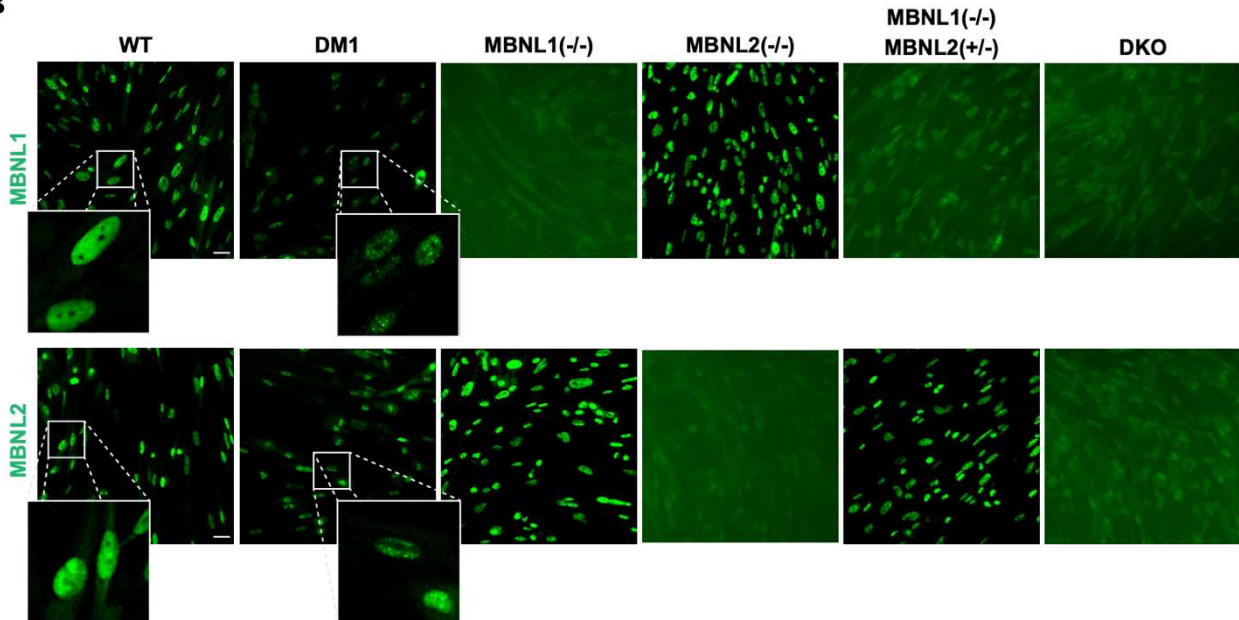
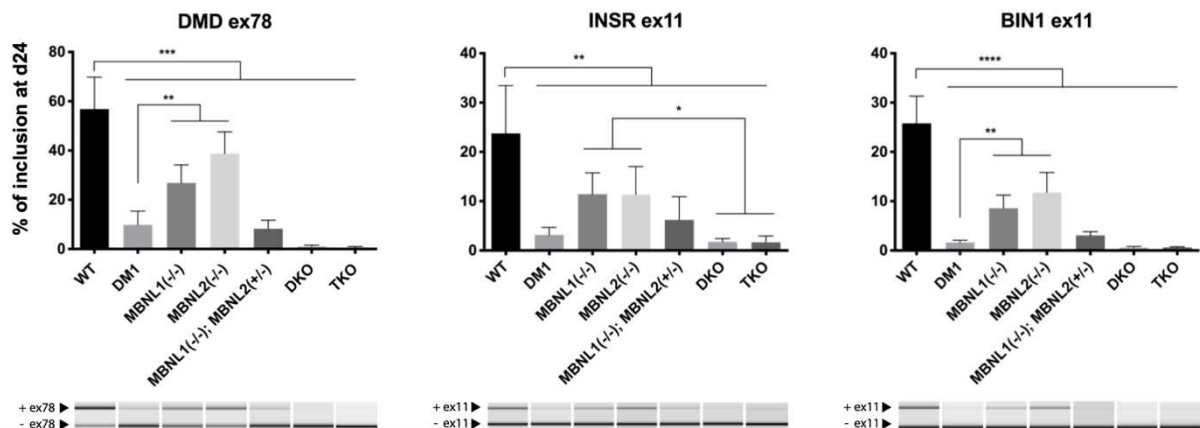
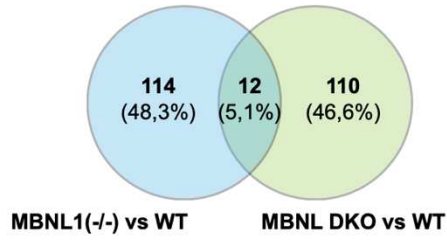
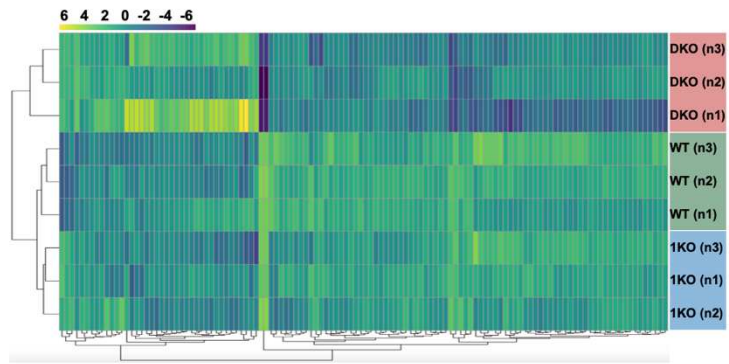
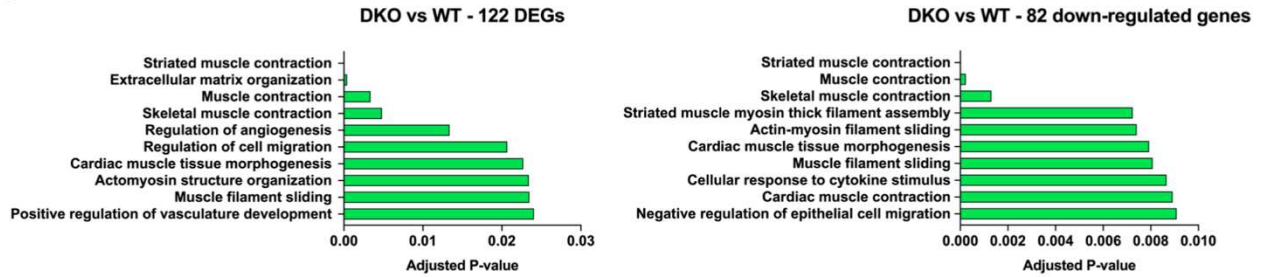
A**B****C****Figure 4**

Figure 4: Depletion of MBNL1 and MBNL2 proteins impairs myogenic fusion and reproduces DM1-associated phenotypes

(A) Representative images of immunolabeling for DESMIN (green), Myosin Heavy Chain (MF20) (red) and DAPI nuclear staining (blue) in skeletal muscle cells generated from hiPSCs after 24 days of differentiation. Scale bar 5 μ M. The graph on the right represents the number of myotubes containing at least 3 nuclei normalized to the total number of nuclei. A minimum of 11.000 nuclei has been quantified per experiment and per cell line. Data represent mean \pm SD values after normalization on the value obtained with WT hiPSC-derived skeletal muscle cells (n = at least 3 independent experiments). Data were analyzed with an ordinary One-Way ANOVA, Tukey's multiple comparisons test (*, P < 0.05; **, P < 0.005; ***, P < 0.0005; ****, P < 0.0001). (B) Representative images of immunolabeling for MBNL1 and MBNL2 on hiPSC-derived skeletal muscle cells (24 days of differentiation). Scale bar 10 μ M. Higher magnifications of the area in inset are provided to show the punctiform staining of MBNL1 and MBNL2 in DM1 hiPSC-derived skeletal muscle cells in comparison with WT hiPSC-derived skeletal muscle cells; Scale bar 5 μ M. (C) Alternative splicing analysis of *DMD* exon 78, *INSR* exon 11 and *BINI* exon 11 were performed by RT-PCR on hiPSC-derived skeletal muscle cells (24 days of differentiation). The mean \pm SD values from 3 independent experiments in duplicate are shown (*, P < 0.05; **, P < 0.005; ***, P < 0.0005; ****, P < 0.0001 - Ordinary One-Way ANOVA, Tukey's multiple comparisons test).

A**B****C**

Selected enriched genes set			
Gene_symbol	Log2 FC	Adjusted P-value	Role
RCS1	-2,23	0,007282878	Actin filament binding
TCAP	-2,16	0,00282123	Sarcomere assembly; structural constituent of muscle
MYL4	-2,00	0,001335221	Actin binding; structural constituent of muscle
MYOM1	-1,97	0,024514653	Striated muscle contraction; structural constituent of muscle
MYBPH	-1,77	0,015623657	Structural constituent of muscle
TNNT2	-1,75	0,044948796	Striated muscle contraction; actin filament binding
TNNC1	-1,63	0,034640984	Striated muscle contraction; actin filament binding

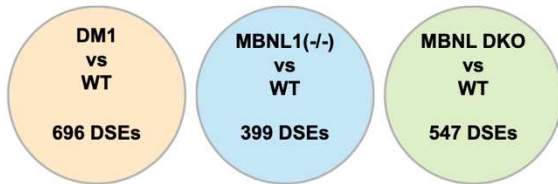
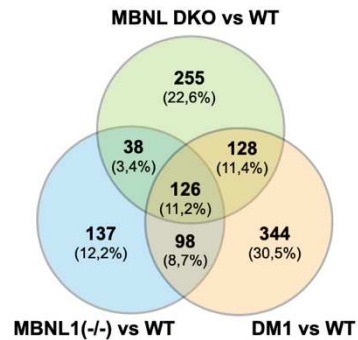
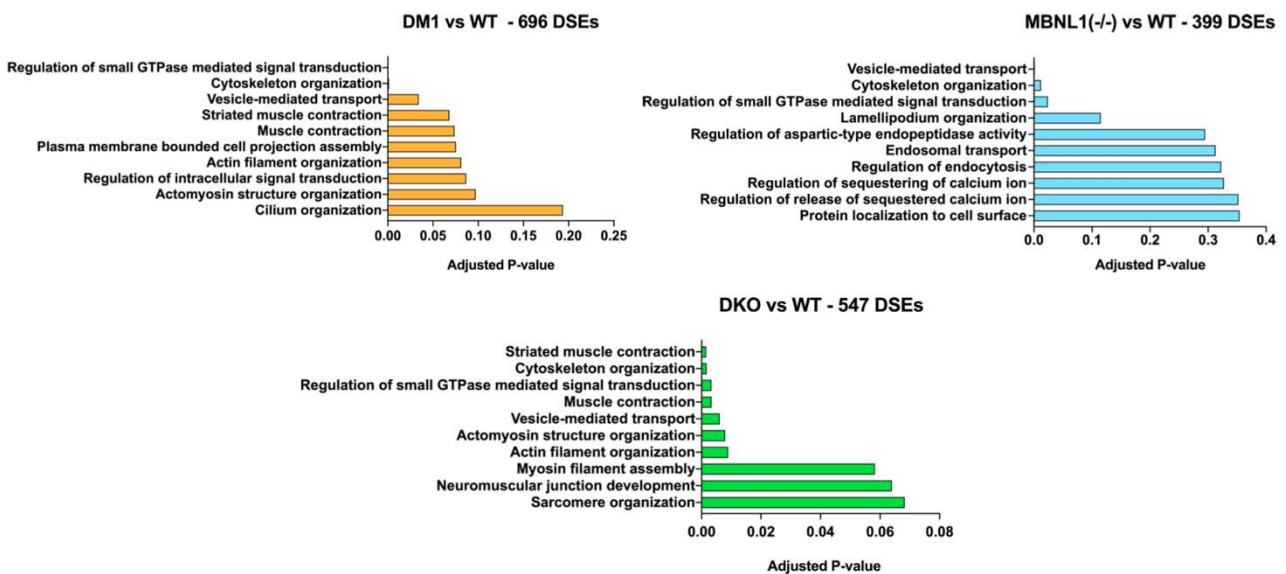
D**E****F****Figure 5**

Figure 5: Depletion of MBNL1 and MBNL2 affects the expression of gene sets and leads to a large number of alternative splicing defects related to skeletal muscle

(A) Venn diagram representation of the number of DEGs commonly deregulated between *MBNL1* KO and *MBNL1,2* DKO myotubes, both compared with WT hiPSC-derived myotubes. The proportion of each subset of DEGs relative to the total number of DEGs (236 in total) is indicated in brackets. **(B)** The hierarchical clustering of the 122 DEGs detected in DKO hiPSC-derived skeletal muscle compared to WT hiPSC-derived skeletal muscle cells. This analysis indicates that *MBNL1* KO are comparable to WT hiPSC-derived skeletal muscle cells regarding these DEGs. Gene expression is represented by a color code ranging from yellow for over-expressed genes to purple for under-expressed genes. **(C)** Gene set enrichment analysis using EnrichR on the 122 total DEGs (left) and 82 down-regulated genes (right) in DKO hiPSC-derived skeletal muscle cells when compared to WT hiPSC-derived skeletal muscle cells. The table on the right provides a list of specific genes down-expressed in DKO hiPSC-derived skeletal muscle cells (EnrichR analysis using GO Biological Process 2018 database). **(D)** Number of differentially spliced exons (DSEs) in DM1, *MBNL1* (-/-) and in DKO hiPSCs-derived skeletal muscle cells compared to WT myotubes. **(E)** Venn diagram highlighting the percentage of common and different DSEs between the three conditions listed in D. All the DSEs are provided in Table S5. **(F)** EnrichR based analysis of gene set enrichments using the lists of DSEs observed in DM1, *MBNL1*(-/-) and DKO hiPSC-derived skeletal muscle cells when compared to WT hiPSC-derived skeletal muscle cells. The top ten most deregulated biological processes are classified according to their adjusted p-value. All the data about the enriched biological processes are detailed in Table S6.

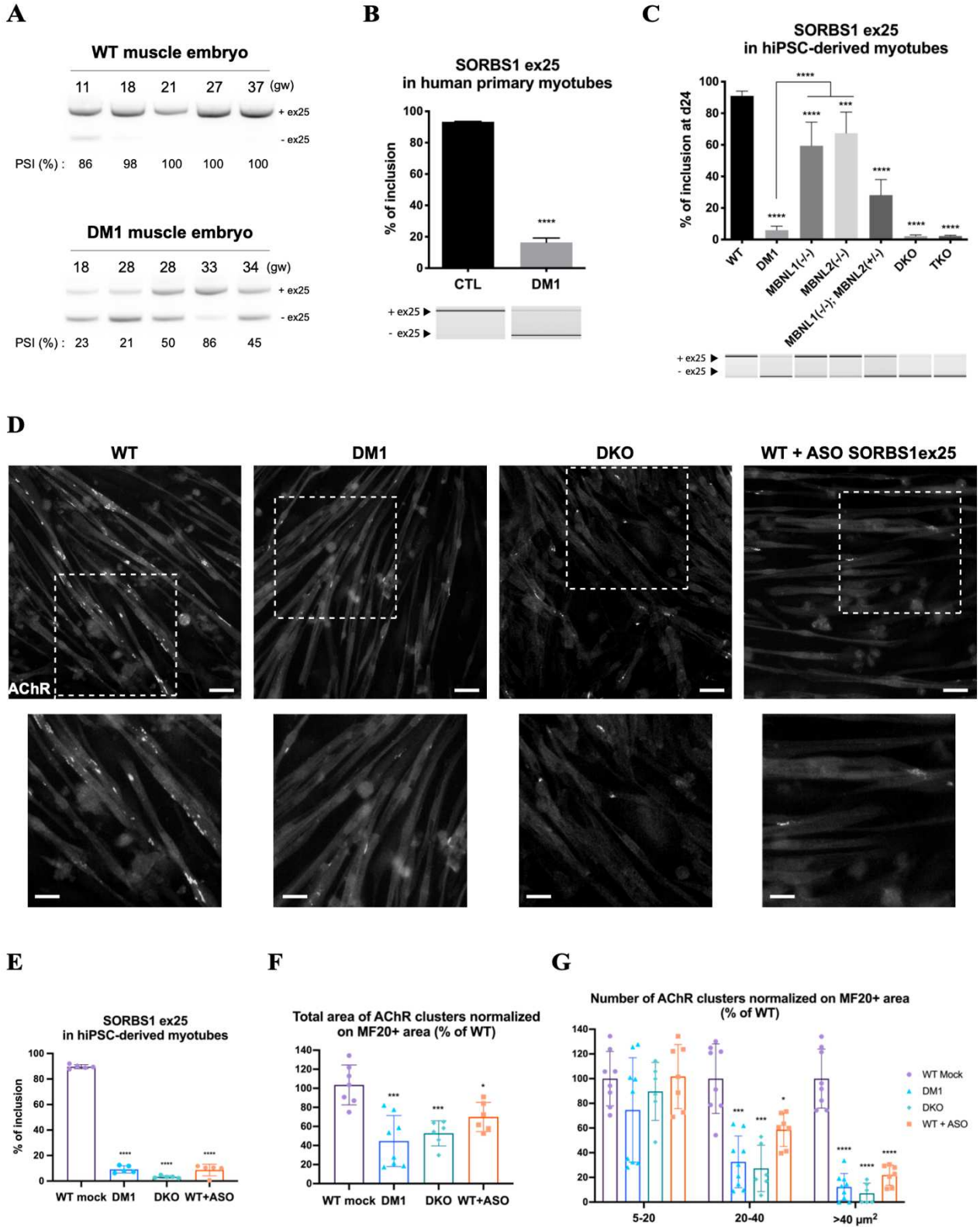


Figure 6

Figure 6: SORBS1 exon 25 splicing defect is regulated by MBNL proteins and leads to AChR clustering defects in hiPSC-derived myotubes

(A) RT-PCR analysis of *SORBS1* exon 25 inclusion in muscles obtained from WT and congenital DM1 human embryos of different ages (gw: gestational weeks). Percent spliced-in (PSI) scores are indicated. (B) RT-PCR analysis of *SORBS1* exon 25 inclusion in DM1 versus WT human primary myotubes. The mean \pm SD values from 3 independent experiments are shown (*, $P < 0.0332$; **, $P < 0.0021$; ***, $P < 0.0002$; ****, $P < 0.0001$; T test). (C) RT-PCR analysis of *SORBS1* exon 25 inclusion in DM1 and *MBNL* KOs myotubes compared to WT. The mean \pm SD values from 3 independent experiments are shown (*, $P < 0.0332$; **, $P < 0.0021$; ***, $P < 0.0002$; ****, $P < 0.0001$; One-Way ANOVA, Tukey's test). (D) Immunostaining of AChRs in Agrin-stimulated DM1, *MBNL* DKO and WT hiPSC-derived myotubes, transfected or not (WT mock) with ASO targeting the ESE of exon 25 (scale bar 10 μ m or 5 μ m for zooms). (E) *SORBS1* exon 25 inclusion was analyzed by RT-PCR in Agrin-stimulated DM1, *MBNL* DKO and WT-transfected myotubes compared to WT mock. The mean \pm SD values from 3 independent experiments are shown (*, $P < 0.0332$; **, $P < 0.0021$; ***, $P < 0.0002$; ****, $P < 0.0001$; One-Way ANOVA, Tukey's test). (F) Total area of AChR clusters and (G) number of AChR clusters according to their size. AChR clusters measuring between 5 and 20 μ m² are considered as small clusters, with medium ones between 20 and 40 μ m². Big clusters are larger than 40 μ m². Data were normalized on MF20+ area and compared to WT. (n=3 independent experiments with at least 2 wells per condition; each point representing the mean of 16 to 36 fields per well) (*, $P < 0.0332$; **, $P < 0.0021$; ***, $P < 0.0002$; ****, $P < 0.0001$; One-Way ANOVA, Tukey's test).

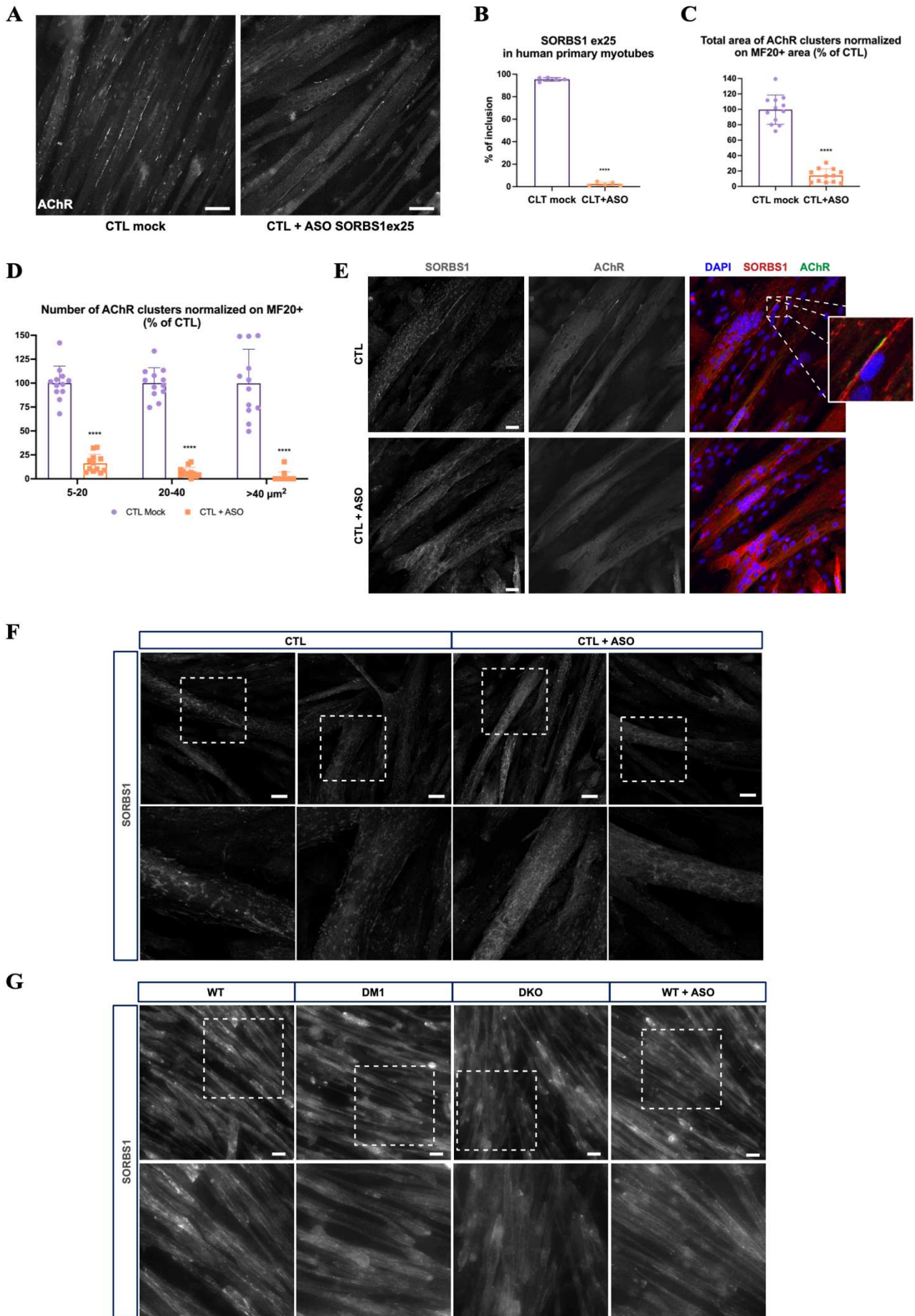


Figure 7

Figure 7: ASO-mediated exclusion of SORBS1 exon 25 induces AChR clustering defects in human primary myotubes and potential role of exon 25 on SORBS1 protein localization

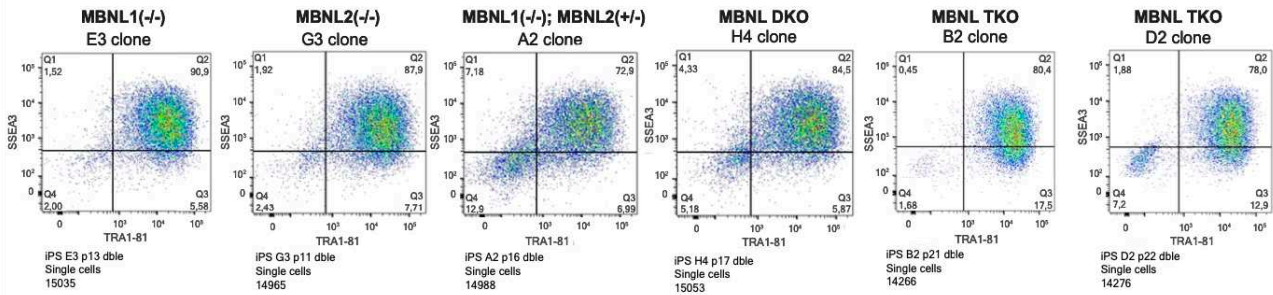
(A) Immunostaining of AChRs in Agrin-stimulated WT primary myotubes, transfected or not with ASO targeting the ESE of SORBS1 exon 25 (scale bars = 10 μ m). (B) RT-PCR analysis of SORBS1 exon 25 inclusion in WT human primary myotubes transfected or not with ASO. The mean \pm SD values from 3 independent experiments are shown (*, P < 0.0332; **, P < 0.0021; ***, P < 0.0002; ****, P < 0.0001; T test). (C) Total area of AChR clusters and (D) number of AChR clusters according to their size. Data were normalized on MF20+ area and compared to WT (n=3 independent experiments with at least 3 wells per condition; each point representing the mean of 16 fields per well) (*, P < 0.0332; **, P < 0.0021; ***, P < 0.0002; ****, P < 0.0001; t test). (E) Immunofluorescence images and merge of SORBS1 and AChR staining of Agrin-stimulated WT primary myotubes transfected or not with ASO. (AChR are in green, SORBS1 in red and nuclei in blue). Scale bars = 10 μ M. (F-G) Immunofluorescence images of SORBS1 staining of Agrin-stimulated WT primary myotubes transfected or not with ASO (F) and of Agrin-stimulated DM1 and *MBNLI,2* DKO hiPSC-derived myotubes, as well as WT hiPSC-derived myotubes treated or not with ASO (G). Scale bars = 10 μ M.

SUPPLEMENTARY FIGURES

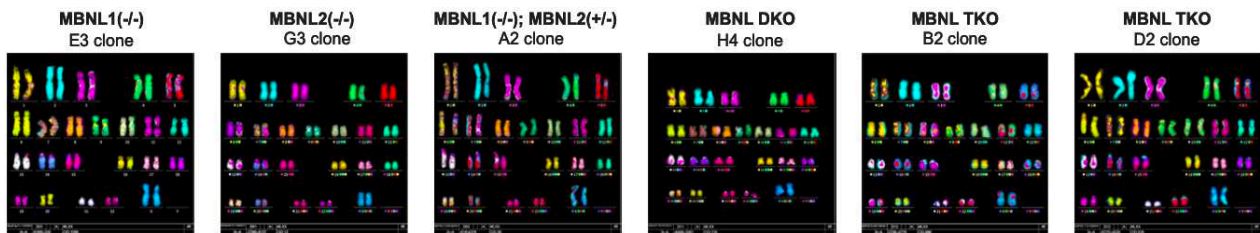
A

	Clones		Indel	STOP codons
MBNL1(-/-)	E3 clone	homozygous	- 14 bp	4
MBNL1(+/-)	C11 clone	heterozygous	WT + 1 bp	0 4
MBNL2(-/-)	G3 clone	homozygous	- 4 bp	4
MBNL2(+/-)	B3 clone	heterozygous	WT - 4 bp	0 4
MBNL1(-/-); MBNL2(-/-)	H4 clone	MBNL1 homozygous (G3 background)	- 35 bp	4
MBNL1(-/-); MBNL2(+/-)	A2 clone	MBNL1 homozygous (B3 background)	- 13 bp	4
MBNL1(-/-); MBNL2(-/-); MBNL3(-/-)	B2 clone	MBNL3 homozygous (H4 background)	- 16 bp	5
	D2 clone	MBNL3 homozygous (H4 background)	- 7 bp	5

B



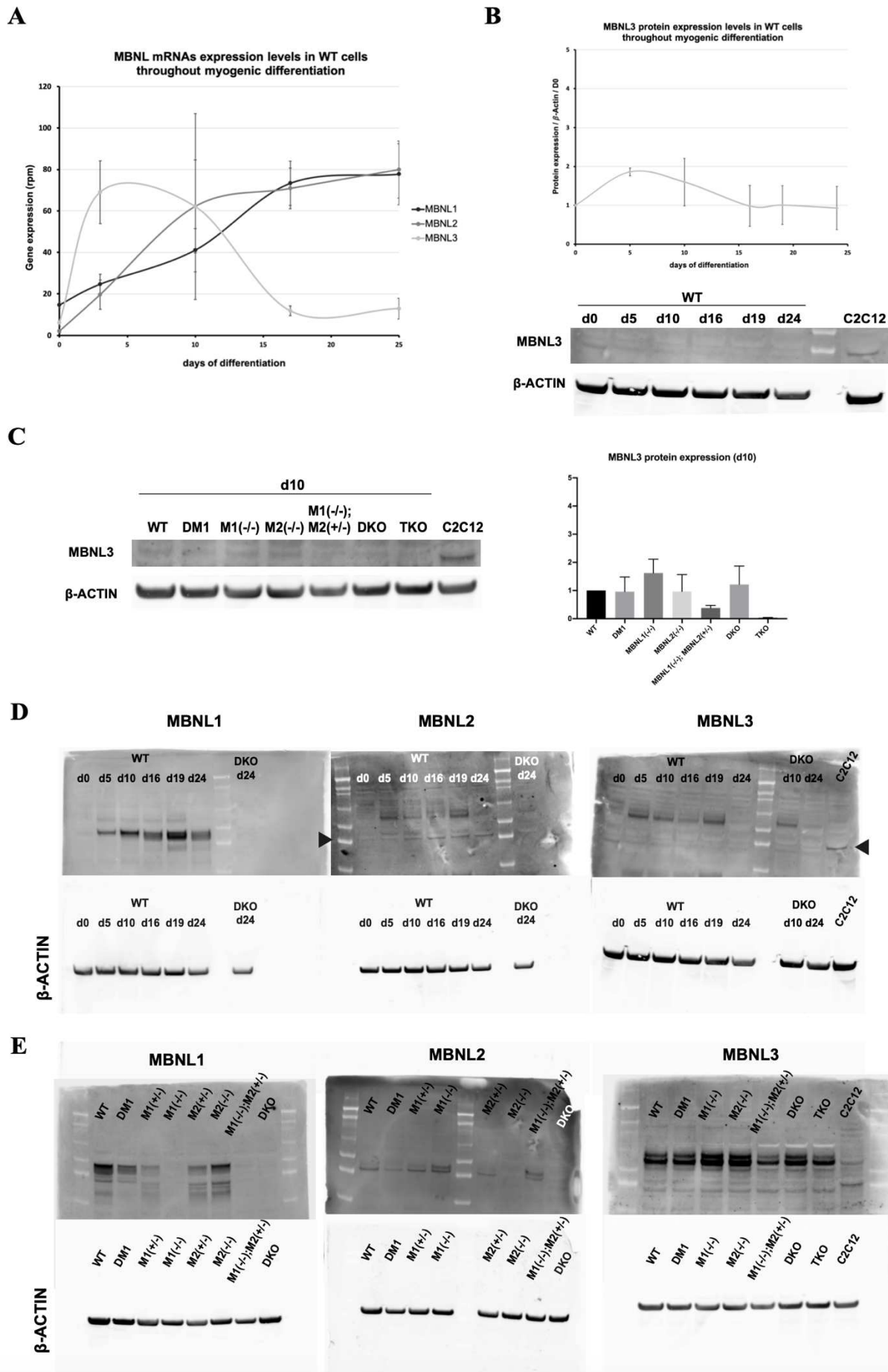
C



Supplementary Figure S1

Supplementary Figure S1: Description and characterization of the different clones obtained, related to Figure 1

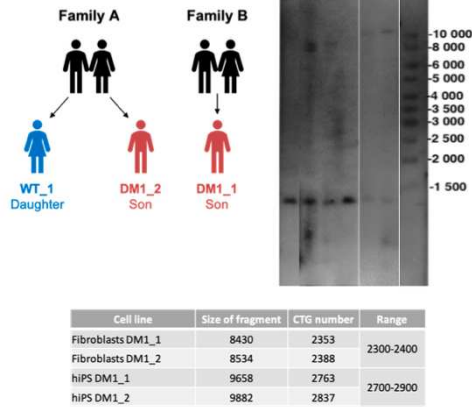
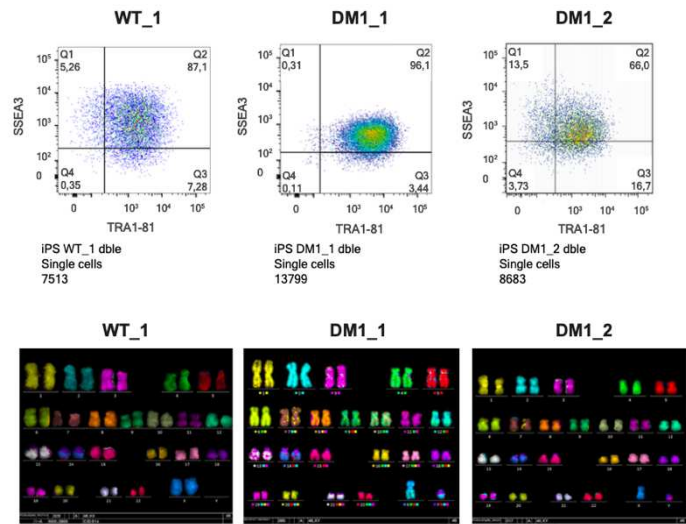
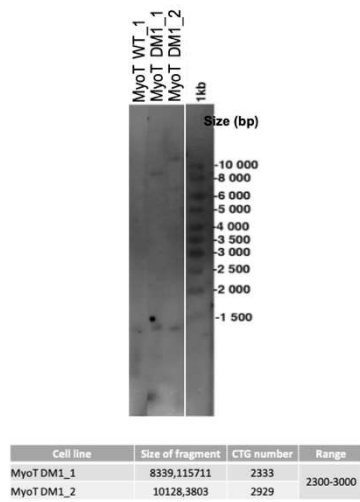
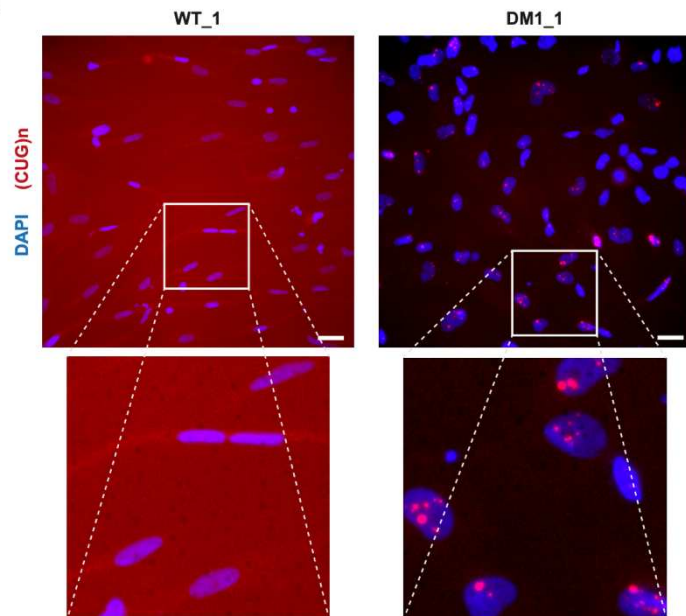
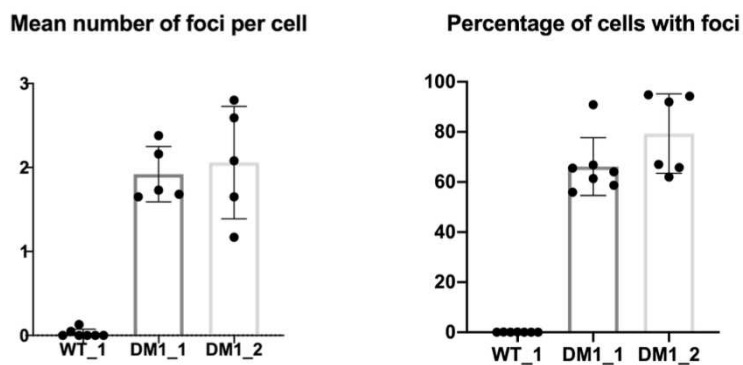
(A) Summary of the different *MBNL* knockout clones used in this study. All clones were analyzed by Sanger sequencing. Homozygous knockout clones were validated with 100% of allelic frequency. *MBNL1(-/-); MBNL2(+/-)* and *MBNL1(-/-); MBNL2(-/-)* clones were generated from *MBNL2(+/-)* and *MBNL2(-/-)* clones respectively. For further results, we focused on E3, G3, A2, H4, B2 and D2 clones. The 6 clones were validated in terms of pluripotency by flow cytometry for SSEA3 and TRA1-81 markers (B), and in terms of genomic integrity (C).



Supplementary Figure S2

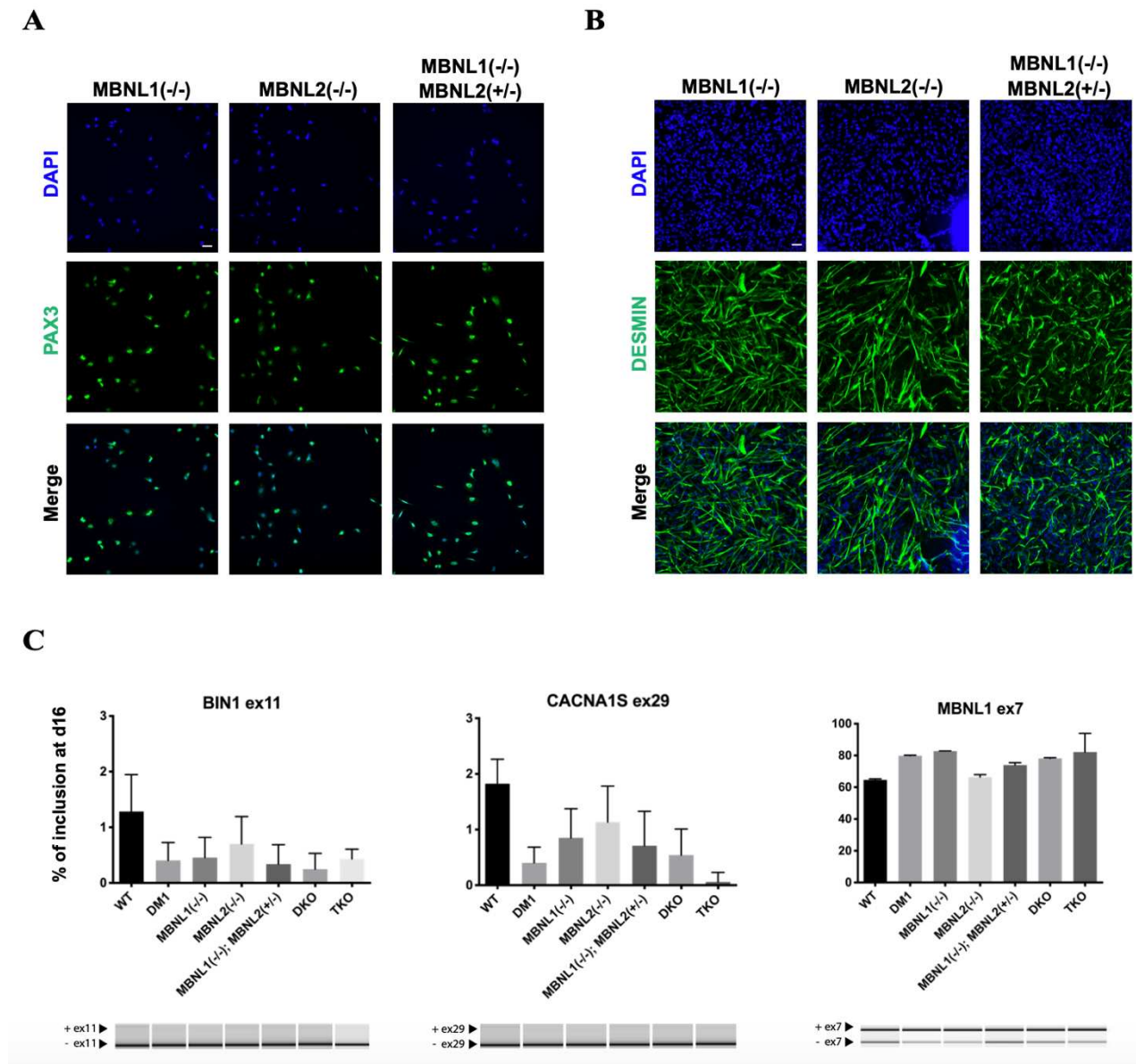
Supplementary Figure S2: Expression profile of MBNL3 at the RNA and protein level during hiPSC differentiation into myogenic cells, related to Figure 2

(A) Expression profile of the three *MBNL* paralog mRNAs in WT hiPSCs at different time points of the myogenic differentiation. Data represent the mean \pm SD values of expression of the transcripts expressed in read per million (rpm) from 3 independent experiments (ArrayExpress database <https://www.ebi.ac.uk/arrayexpress/experiments/E-METAB-8321/>). (B) Western blot and corresponding quantification showing the expression level of MBNL3 protein throughout myogenic differentiation of WT hiPSCs. C2C12 myoblasts were used as positive control. Data represent mean \pm SD values from 2 independent experiments. (C) Western blot analysis for MBNL3 protein in the different hiPSC lines used in this study and differentiated for 10 days. β -Actin was used as housekeeping protein for normalization. Data represent the mean \pm SD values from 3 independent experiments. (D) Original Western blots used in Fig2B and FigS2C. (E) Original Western blots used in Fig2C and FigS2D.

A**B****C****D****E****Supplementary Figure 3**

Supplementary Figure 3: Presence of CUG foci in skeletal muscle cells derived from DM1 hiPSCs, related to Figure 2

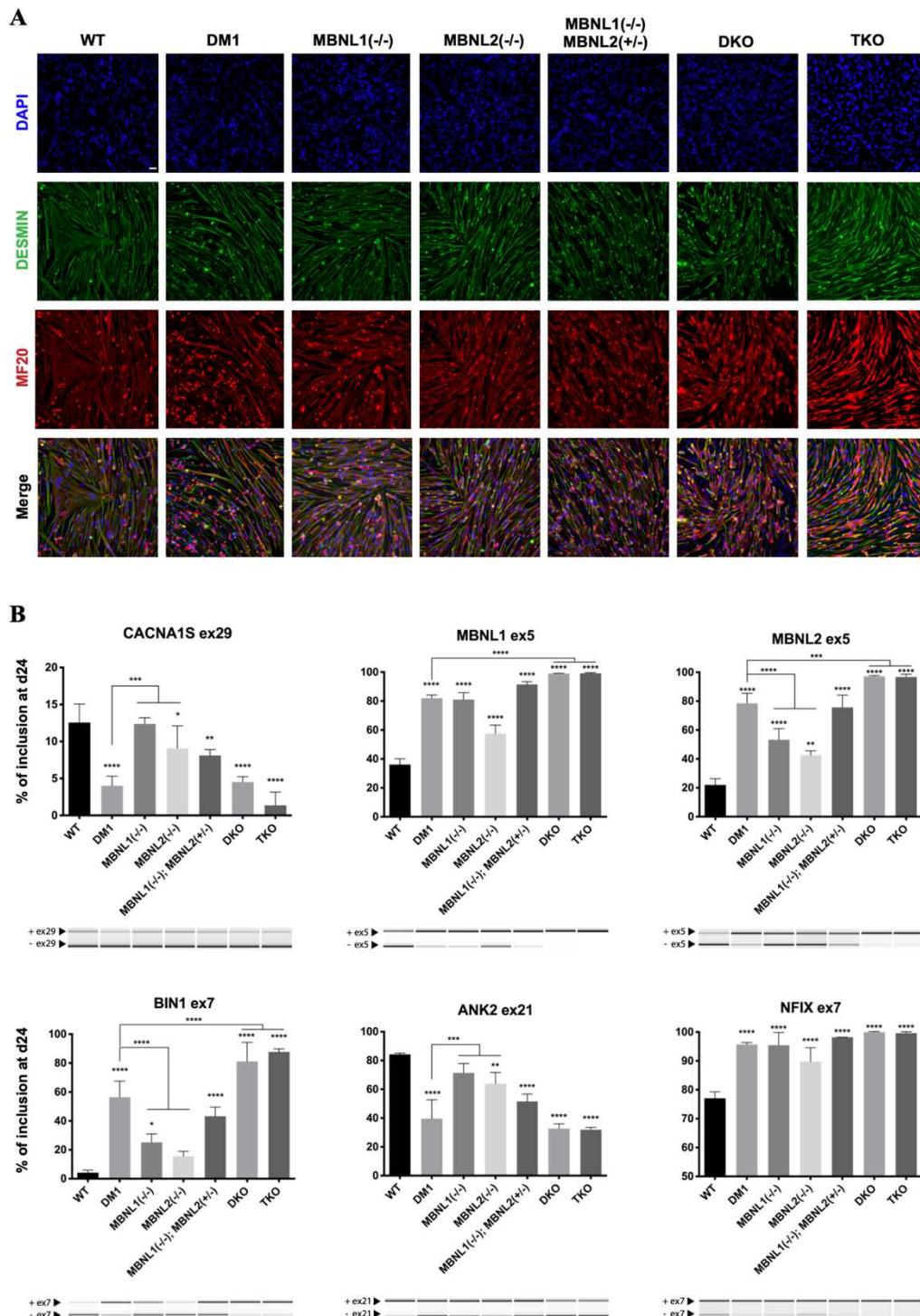
(A) Fibroblasts from two DM1 patients and one healthy donor subject (WT; sibling of one DM1 patient) were used for reprogramming into hiPSCs. Southern blot analysis of the CTG repeats length present in the different fibroblasts and hiPSCs. **(B)** Characterization of WT and DM1 hiPSCs by flow cytometry for two markers of pluripotency (SSEA3 and TRA1-81). Genomic integrity of these hiPSC lines was validated by karyotyping analysis. All the *MBNL* KO clones were generated from the WT_1 cell line. **(C)** Southern blot analysis of the CTG repeat length in hiPSC-derived skeletal muscle cells (24 days after differentiation). **(D)** Representative images of double immunolabeling for intranuclear CUG repeat foci (red) and DAPI nuclear staining (blue) of WT and DM1 hiPSC-derived skeletal muscle cells (24 days of differentiation). The images below are higher magnifications of the inset to better visualize the presence of CUG RNA foci in DM1 nuclei. Scale bar 50 μ M. **(E)** Quantification of the number of CUG repeat foci per cell as well as the percentage of cells with CUG RNA foci in skeletal muscle cells derived from WT and the two DM1 hiPSC lines (24 days of differentiation). The quantification of CUG repeat foci was performed as previously described (Maury et al., iScience 2019). Data represent the mean \pm SD values from at least 3 independent experiments in quadruplicate (*, $P < 0.05$; **, $P < 0.005$; ***, $P < 0.0005$; ****, $P < 0.0001$ - Ordinary One-Way ANOVA, Tukey's multiple comparisons test).



Supplementary Figure S4

Supplementary Figure S4: Characterization of myogenic precursors derived from *MBNL* knockouts hiPSCs, related to Figure 3

Representative images of immunolabeling for PAX3 (**A**) and DESMIN (**B**) on MBNL1 (-/-), MBNL2(-/-) and MBNL1(-/-); MBNL2(+/-) after 5 or 16 days of myogenic differentiation respectively. Scale bars 10 μ M. (**C**) Analysis of alternative splicing by RT-PCR on DESMIN+ cells derived from hiPSCs. Data represent the mean \pm SD values from 3 independent experiments (*, $P < 0.05$; **, $P < 0.005$; ***, $P < 0.0005$; ****, $P < 0.0001$ - Ordinary One-Way ANOVA, Tukey's multiple comparisons test).

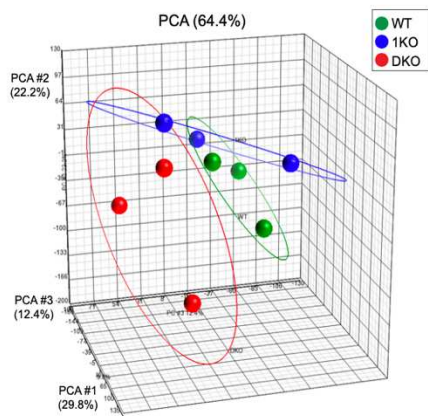


Supplementary Figure S5

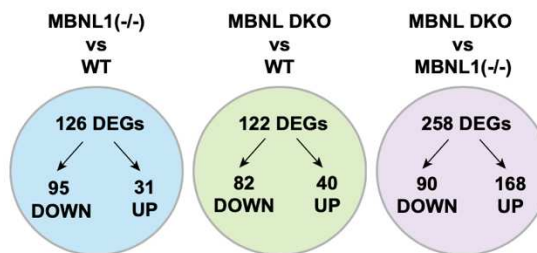
Supplementary Figure S5: Loss of MBNL1 and MBNL2 proteins reproduces DM1-associated splicing defects, related to Figure 4

(A) Representative images of immunolabeling for DESMIN (green), MF20 (red) and DAPI (blue) in skeletal muscle cells differentiated from hiPSCs for 24 days. Scale bar 10 μ M. (B) Analysis of alternative splicing of *CACNA1S* exon 25, *MBNL1* exon 5, *MBNL2* exon 5, *BIN1* exon 7, *ANK2* exon 21 and *NFIX* exon 7 on hiPSC-derived skeletal muscle cells (24 days of differentiation). Data represent the mean \pm SD values from 3 independent experiments (*, $P < 0.05$; **, $P < 0.005$; ***, $P < 0.0005$; ****, $P < 0.0001$ - Ordinary One-Way ANOVA, Tukey's multiple comparisons test). Asterisks placed above individual bar correspond to the comparison with WT.

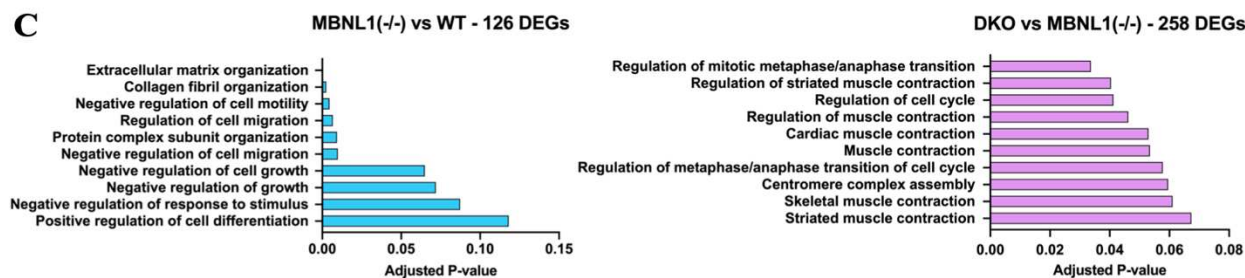
A



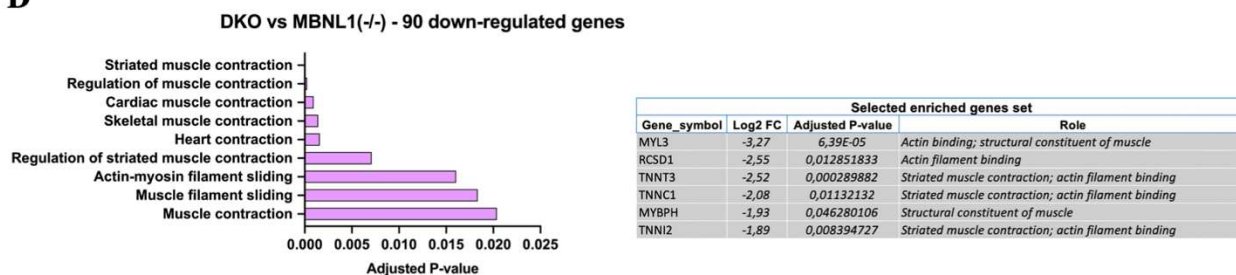
B



C



D



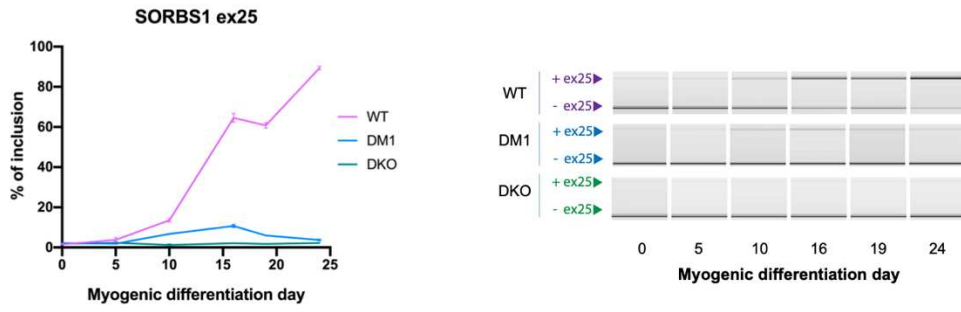
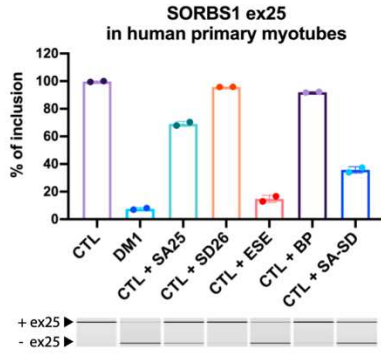
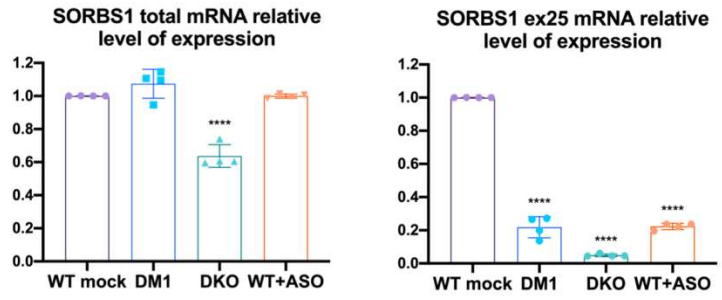
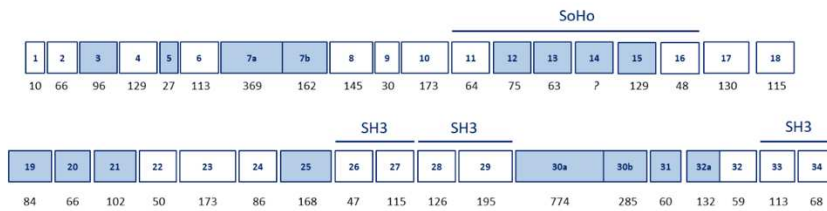
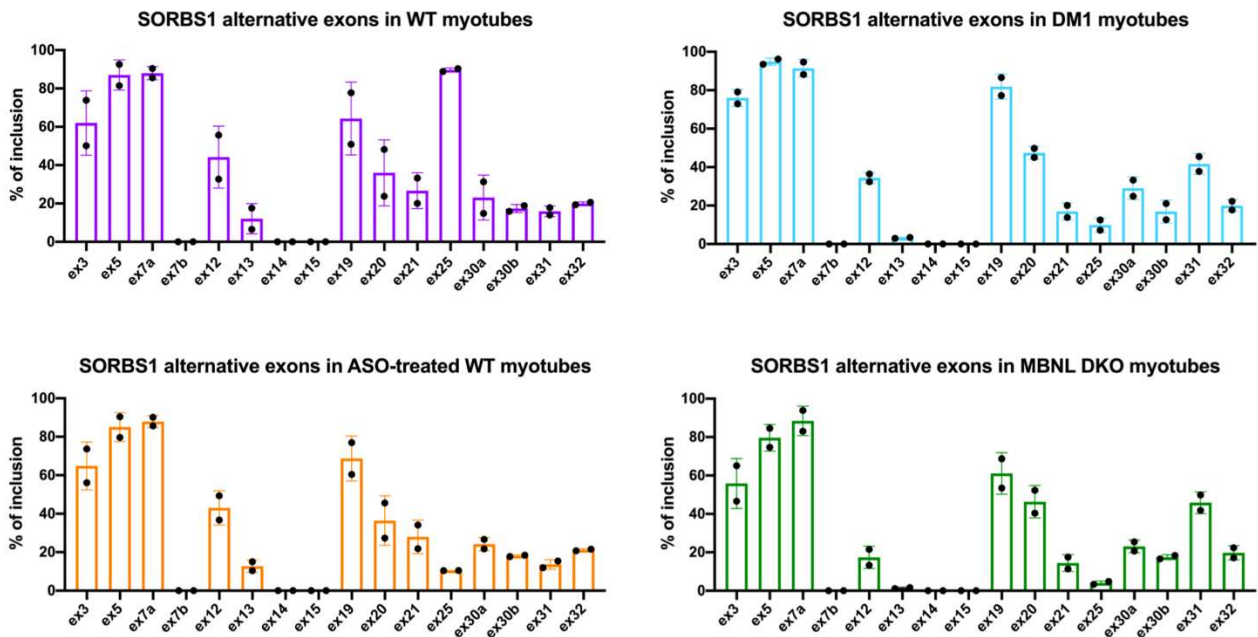
E

Gene_symbol	DKO vs WT		MBNL1 KO vs WT		DM1 vs WT		Role
	deltaPSI	pvalue_corrected_glm	deltaPSI	pvalue_corrected_glm	deltaPSI	pvalue_corrected_glm	
BIN1 ex8 (7)	0,78	0,00E+00	0,37	0,00E+00	0,55	0,00E+00	Formation of T-tubules; excitation-contraction coupling
SORBS1 ex30 (25)	-0,76	0,00E+00	-0,19	8,01E-11	-0,68	0,00E+00	Cytoskeletal protein binding; AChR clustering
KIF13A ex27 (32)	-0,75	0,00E+00	-0,23	0,00E+00	-0,62	0,00E+00	Microtubule motor activity
FLNB ex33	-0,71	0,00E+00	-0,33	0,00E+00	-0,45	0,00E+00	Actin binding
TRDN ex20	-0,57	0,00E+00	-0,32	0,00E+00	-0,43	0,00E+00	Muscle contraction
MYOM1 ex18	0,59	0,00E+00	0,31	0,00E+00	0,61	0,00E+00	Striated muscle contraction; structural constituent of muscle
TRDN ex32	-0,49	0,00E+00	-0,16	4,15E-14	-0,25	0,00E+00	Muscle contraction
DMD ex85 (78)	-0,47	0,00E+00	-0,20	0,00E+00	-0,36	0,00E+00	Structural constituent of cytoskeleton; calcium ion binding
SYNE2 ex117	-0,32	3,44E-06	-	-	-	-	Myonuclear anchorage
MYL6 ex7	-0,29	0,00E+00	-0,15	2,72E-06	-0,24	0,00E+00	Structural constituent of muscle
PICALM ex19	-0,29	0,00E+00	-0,13	0,00E+00	-0,12	3,83E-04	AP2-dependent clathrin-mediated endocytosis at the neuromuscular junction
ANK1 ex24	-0,26	1,62E-02	-	-	-0,23	4,77E-02	Structural constituent of cytoskeleton
SORBS1 ex36 (30a)	-0,25	3,02E-03	-	-	-0,28	7,75E-09	Cytoskeletal protein binding; AChR clustering
WASF3 ex7	0,24	2,84E-11	-	-	0,25	2,45E-08	Cytoskeletal organization
ANK3 ex41	0,24	0,00E+00	-	-	0,14	0,00E+00	Structural constituent of cytoskeleton
TPM3 ex8	-0,23	0,00E+00	-	-	-	-	Actin binding
SORBS1 ex26 (21)	-0,21	4,11E-05	-	-	-0,20	4,18E-06	Cytoskeletal protein binding; AChR clustering
TPM1 ex9	-0,20	1,50E-10	-	-	0,20	0,00E+00	Actin binding; cytoskeletal protein binding
BIN1 ex12 (11)	-0,20	0,00E+00	-0,13	0,00E+00	-0,17	0,00E+00	Formation of T-tubules; excitation-contraction coupling
MYL5 ex3	0,19	3,68E-05	-	-	-0,11	1,67E-03	Structural constituent of muscle
TNNT3 ex9	0,17	1,39E-09	0,10	6,85E-03	0,18	1,37E-10	Striated muscle contraction; actin filament binding
MICAL2 ex24	0,17	3,54E-07	-	-	0,30	0,00E+00	Actin binding
WASF3 ex8	-0,17	8,69E-11	-0,16	1,52E-05	-0,18	1,58E-11	Cytoskeletal organization
SYNE1 ex143	-0,15	0,00E+00	-0,11	2,65E-11	-0,15	0,00E+00	Myonuclear anchorage
MYBPHL ex6	-0,15	1,22E-02	-	-	-	-	Structural constituent of muscle
UNC13B ex38	0,15	0,00E+00	-	-	-	-	Synaptic vesicle cycle; neurotransmitter release cycle
ABIZ1 ex10	-0,14	8,18E-07	0,27	0,00E+00	0,37	0,00E+00	Regulation of actin cytoskeleton dynamics
MUSK ex9	-0,13	3,83E-06	-	-	-	-	Formation and maintenance of neuromuscular junctions
CHRNA1 ex4	-0,12	1,07E-03	-	-	-0,15	6,12E-05	Postsynaptic nicotinic acetylcholine receptors
MYBPC1 ex6	0,11	1,06E-03	-	-	-	-	Striated muscle contraction; actin binding

Supplementary Figure S6

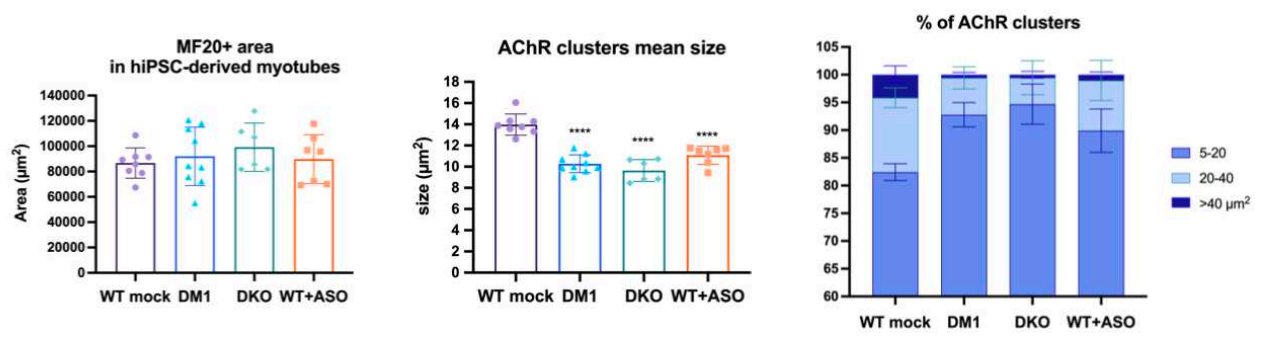
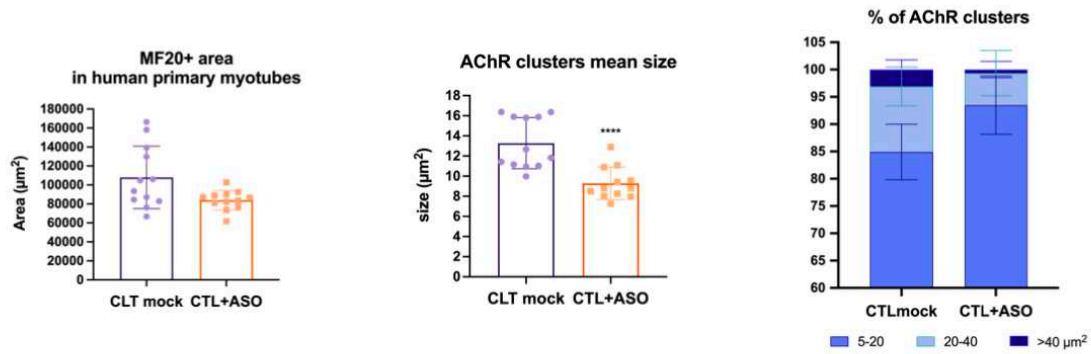
Supplementary Figure S6: A large number of skeletal muscle-related genes are down-regulated in DKO myotubes and identification of thirty muscle-related DSEs, related to Figure 5

(A) PCA (Principal Component Analysis) mapping indicating distinct transcriptomic profiles between WT, *MBNLI*(-/-) and *MBNLI*(-/-); *MBNL2*(-/-) hiPSC-derived skeletal muscle cells (24 days of differentiation). **(B)** Schematic representation of the number of differentially expressed genes (DEGs) in *MBNLI* KO hiPSCs-derived myotubes compared to WT (blue), DKO myotubes compared to WT (green) and DKO compared to *MBNLI* KO myotubes (purple). The detailed list DEGs for each condition is provided in Table S3. **(C)** EnrichR based gene set enrichment analysis of the DEGs identified between *MBNLI* KO hiPSCs and WT-derived skeletal muscle cells (left), and between DKO hiPSCs and *MBNLI* KO hiPSC-derived skeletal muscle cells (right). The most deregulated biological processes are classified according to their adjusted p-value. **(D)** Enriched biological processes after gene ontology analysis of the 90 down-regulated genes in DKO versus *MBNLI* KO hiPSCs-derived skeletal muscle cells. Some enriched DEGs, which are known to be involved in most of the deregulated biological processes related to skeletal muscle, are indicated. Biological interpretations of the list of DEGs were performed using the dedicated software EnrichR. The most deregulated biological processes are classified according to their adjusted p-value. Results of each GO analysis are detailed in Table S4. **(E)** List of thirty DSEs identified in DKO hiPSCs-derived skeletal muscle cells and related to skeletal muscle. Among these 30 differentially spliced exons (DSEs) selected, only 3 of them (*ABI2* exon 10, *MYL5* exon 3 and *TPM1* exon 9) show mismatches in the sense of inclusion/exclusion (red and blue respectively) between the DM1 and DKO conditions. For alternative splicing analysis, exons have been numbered according to FasterDB database. The numbers of the corresponding exons are indicated in brackets.

A**B****C****D****E****Supplementary Figure S7**

Supplementary Figure S7: Inclusion of *SORBS1* exon 25 increases throughout human myogenesis and specificity of ESE-targeting ASO, related to Figure 6

(A) RT-PCR analysis of *SORBS1* exon 25 inclusion throughout WT, DM1 and *MBNL* DKO hiPSCs differentiation into skeletal muscle cells. **(B)** RT-PCR analysis of *SORBS1* exon 25 splicing profiles in human primary myotubes 48h post transfection of ASOs targeting key sequences for exon 25 exclusion (n=2 independent experiments in duplicate). ISE: Intron Splice Enhancer, ISS: Intron Splice Suppressor, ESE: Exon Splice Enhancer, ESS: Exon Splice Suppressor, BP: Branching Point, SA: Acceptor Site, SD: Donor Site. The best skipping of exon 25 was obtained with the ASO targeting the ESE. Following experiments were then performed using this ASO. **(C)** RT-qPCR analysis of total *SORBS1* mRNA (left) and *SORBS1* exon 25 mRNA (right) expression levels in DM1, *MBNL* DKO and WT-transfected hiPSC-derived myotubes relative to WT. **(D)** Schematic representation of human *SORBS1* coding exons (alternative exons are colored in blue) and **(E)** RT-PCR analyses of splicing profiles of *SORBS1* alternative exons in DM1, *MBNL* DKO and WT hiPSC-derived myotubes transfected or not with ASO (n=2 independent experiments).

A**B****C****Supplementary Figure S8**

Supplementary Figure S8: *SORBS1* exon 25 ASO-mediated exclusion leads to AChR clustering defects, related to Figure 6

(A) MF20+ area and AChR clusters mean size in DM1, *MBNL* DKO and WT hiPSC-derived myotubes or in WT primary myotubes (B) transfected or not with ASO. Another representation of the number of clusters according to their size, expressed as a percentage, is shown, related to data in Figure 6G and Figure 7D, respectively. Concerning hiPSC-derived myotubes: n=3 independent experiments with at least 2 wells per condition; each point representing the mean of 16 to 36 fields per well (*, $P < 0.0332$; **, $P < 0.0021$; ***, $P < 0.0002$; ****, $P < 0.0001$; One-Way ANOVA, Tukey's test). And concerning human primary myotubes: n=3 independent experiments with at least 3 wells per condition; each point representing the mean of 16 fields per well (*, $P < 0.0332$; **, $P < 0.0021$; ***, $P < 0.0002$; ****, $P < 0.0001$; t test). (C) Western blots showing the expression levels of *SORBS1* protein in DM1, *MBNL* DKO and WT hiPSC-derived myotubes transfected or not with ASO. β -Actin was used as housekeeping protein to verify loading variations (* indicates bands that are not described as corresponding to any *SORBS1* protein isoform in terms of molecular weight).

SUPPLEMENTARY MATERIAL

Flow cytometry analysis

hiPS cells were dissociated with Trypsin-EDTA 0,05% (Gibco®). Immunostaining was performed on 100.000 cells per experiment at 4 °C for 30'. The antibodies used are described in Table S8. Cells were then incubated with 10 µg/ml DAPI solution. Analysis was performed on a MACSQuant Flow Cytometer (Miltenyi Biotec®) using FlowJo software (Tree Star, Inc.). About 15 000 events were recorded for each sample analyzed.

mFISH karyotyping analysis

Cells were conditioned with colchicine (Eurobio®) for 90', warmed with a hypotonic solution (5 mg/mL KCL) and fixed with Carnoy's fixative. mFISH 24Xcite probe (Metasystems®) and ProLong Gold Antifade Mountant with DAPI (Life Technologies®) were used for mFISH staining. Then, 10–50 metaphases were acquired with Metafer MetaSystems software coupled to an AxioImager Zeiss Z2 microscope equipped with a camera cool cube and 10X and 63X objectives. Images were analyzed with Isis software (MetaSystems®).

Off-target analysis

Off-target sites within exons as well as the most likely intronic or intergenic sites for each gRNA used (g47, g17 and g183) were determined with CRISPOR using the CFD off-target score (Doench et al. 2016). An AmpliSeq Custom panel was design covering the 23bp of the 47 off-target sites. gDNA extracted from each edited-hiPS cell line used in this study was used to generate AmpliSeq DNA library sequenced on a NextSeq 550 300bp run (2x150). The quality was assessed with FastQC (v0.11.9) (Andrews 2010) and the reads were trimmed with Cutadapt (v1.18) (Martin 2011) and PrinSeq (v0.20.4 _PhredScore < 20) (Schmieder and Edwards 2011) . Reads were aligned to the GRCH38 genome using BWA (v0.7.17-r1188) (Li and Durbin 2009) and multi-alignment were removed using Samtools (v0.1.18 r982:295_score < 10) (Li et al. 2009). SNP calling was done using BCFtools (v0.1.17-dev r973:277) (Li 2011) and filtered on the 23bp gRNA homologous region using bedtools intersect (v2.25.0) (Quinlan and Hall 2010). A true induced mutation was considered possible if the fraction of reads with a given variant was substantially higher than error rate of base calling. No off-target mutations were observed at any site (Table S2). Two already described SNP in chromosomes 8 and 22 were detected in WT hiPSCs and in every edited clones (Tables S2).

Ion AmpliSeq transcriptome human gene expression kit

For each sample, 50 ng of total RNA was reverse transcribed using the Ion AmpliSeq Transcriptome Human Gene Expression kit following the protocol of the manufacturer (ThermoFisher Scientific®). The cDNA libraries were amplified and barcoded using Ion AmpliSeq Transcriptome Human Gene Expression core panel and Ion Xpress Barcode Adapter (ThermoFisher Scientific®). The amplicons were quantified using Agilent High Sensitivity DNA kit before the samples were pooled in sets of eight. Emulsion PCR and Enrichment was performed on the Ion OT2 system Instrument using the Ion PI Hi-Q OT2 200 kit (ThermoFisher Scientific®). Samples were loaded on an Ion PI v3 Chip and sequenced on the Ion Proton System using Ion PI Hi-Q sequencing 200 kit chemistry (200 bp read length; ThermoFisher Scientific®).

The quality control of the sequencing data was evaluated using FastQC (v0.11.2) (Andrews 2010). The reads were trimmed using Prinseq (v0.20.4) (Schmieder and Edwards 2011) (--trim-right 20) and filtered by average quality score (--trim-qual 20). Reads were mapped to the Amplicons sequence

reference using bowtie2 local (v2.3.4.1) (Langmead and Salzberg 2012). Reads below a mapping score of 10 or multi-mapped were filtered using samtools (v0.1.19) (Li et al. 2009). The gene expression level in each sample was quantified with the specific GRCh38.92 annotation: Each amplicon sequence has been BLASTed (Camacho et al. 2009) against transcriptome and revealed that 21080 genes were correctly identified to be quantified (Table S7: correspondance_gene-RefSeq_Amplicons.xlsx).

The differential gene expression (DEG) between conditions was calculated with DESeq2 (Love et al. 2014) (v1.18.1 using R v3.4.1). Genes were considered differentially expressed when their adjusted pvalue was under 0.05 and their $|\log_2\text{FoldChange}|$ was upper than 0.4 (options: lfcThreshold=0.4, altHypothesis="greaterAbs"). Data are available on GEO database (GSE161029).

RNA-Seq library preparation and sequencing (TrueSeq polyA)

For each sample, 100ng of total RNA was used. mRNAs were purified using poly-T oligo from TruSeq Stranded kit (Illumina®) and were fragmented and reverse transcribed using the TruSeq Stranded kit following the protocol of the manufacturer. Final library was quantified using Agilent High Sensitivity DNA kit before the samples were pooled in sets of six for sequencing. 2nM of pooled libraries were denatured and the quantity of 1.8pM was used for cluster generation before paired-end sequencing on an Illumina Nextseq550 (Hight output 2*150pb run). Samples were sequenced with an average of 89 532 366 reads (paired-end) per sample.

The quality control of the sequencing data was evaluated using FastQC (v0.11.2) (Andrews 2010). The reads were trimmed using Prinseq (v0.20.4) (Schmieder and Edwards 2011) (--trim-right 20) and filtered by average quality score (--trim-qual 20) and cutadapt (v1.16) (Martin 2011).

For the alternative splicing analysis, the raw reads were annotated on GRCh37.75 Ensembl reference annotation by Tophat2 (Kim et al. 2013) (v2.0.13) and analyzed using only reads on exonic junctions with FaRLine pipeline (Benoit-Pilven et al. 2018) and FasterDB database (Mallinjou et al. 2014). Exons were considered differentially spliced when their adjusted pvalue was under 0.05 and their DeltaPSI (percent spliced-in) value was upper that 0.1. Data are available on GEO database (GSE161897).

REFERENCES

- Andrews S. 2010. FastQC A Quality Control tool for High Throughput Sequence Data. *Online*. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (Accessed October 12, 2020).
- Benoit-Pilven C, Marchet C, Chautard E, Lima L, Lambert M-P, Sacomoto G, Rey A, Cologne A, Terrone S, Dulaurier L, et al. 2018. Complementarity of assembly-first and mapping-first approaches for alternative splicing annotation and differential analysis from RNAseq data. *Sci Rep* **8**: 4307. <http://www.nature.com/articles/s41598-018-21770-7> (Accessed October 12, 2020).
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* **10**: 421. <http://www.ncbi.nlm.nih.gov/pubmed/20003500> (Accessed October 12, 2020).
- Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, Smith I, Tothova Z, Wilen C, Orchard R, et al. 2016. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* **34**: 184–191. <http://www.nature.com/articles/nbt.3437> (Accessed February 18, 2021).
- Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* **14**: R36. <http://www.ncbi.nlm.nih.gov/pubmed/23618408> (Accessed October 12, 2020).
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359. <http://www.nature.com/articles/nmeth.1923> (Accessed October 12, 2020).
- Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**: 2987–93. <http://www.ncbi.nlm.nih.gov/pubmed/21903627> (Accessed February 19, 2021).
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**: 1754–1760. <https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btp324> (Accessed February 19, 2021).
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Subgroup 1000 Genome Project Data Processing. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079. <https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btp352> (Accessed October 12, 2020).
- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550. <http://genomebiology.biomedcentral.com/articles/10.1186/s13059-014-0550-8> (Accessed October 12, 2020).
- Mallinoud P, Villemin J-P, Mortada H, Polay Espinoza M, Desmet F-O, Samaan S, Chautard E, Tranchevent L-C, Auboeuf D. 2014. Endothelial, epithelial, and fibroblast cells exhibit specific splicing programs independently of their tissue of origin. *Genome Res* **24**: 511–21. <http://www.ncbi.nlm.nih.gov/pubmed/24307554> (Accessed October 12, 2020).
- Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* **17**: 10. <http://journal.embnet.org/index.php/embnetjournal/article/view/200> (Accessed October 12, 2020).
- Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841–2. <http://www.ncbi.nlm.nih.gov/pubmed/20110278> (Accessed February 19, 2021).
- Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27**: 863–4. <http://www.ncbi.nlm.nih.gov/pubmed/21278185> (Accessed October 12, 2020).

SUPPLEMENTARY TABLES

RNA-guides used for CRISPR/Cas9-mediated genome editing

MBNL1	(g47fw)	GGACACAAAATGGCTAACAC TGG
MBNL2	(g17rev)	CCATTTTGTATCTCTGACTG GGG
MBNL3	(g194fw)	ACACTTGCAAGCCAGACTCA AGG

Primers used for RFLP assays

MBNL1	Fw	GGGGACATTTTCATCATCAGTTCA
	Rev	TTCCCCAGCCCAACATCATC
MBNL2	Fw	AACAGAAACAACAGCCCAAA
	Rev	AGGGAATCAAAGCAGGCAAT
MBNL3	Fw	CCGCTATGCTCACCTACTG
	Rev	GAAACAGGCCACAAGTGCAG

Restriction enzymes used for RFLP assays

MBNL1	Bsrl	NEB (R0527S)
MBNL2	Bsrl	NEB (R0597S)
MBNL3	SmlI	NEB (R0597S)

Primers used for Alternative Splicing analysis

DMDex78	Fw	TTAGAGGAGGTGATGGAGCA
	Rev	GATACTAAGGACTCCATCGC
INSRex11	Fw	GGATTACCTGCACAACGTGGTT
	Rev	ACGGCCACCGTCACATTC
BIN1ex11	Fw	AGAACCTCAATGATGTGTGG
	Rev	TCGTGGTTGACTCTGATCTCGG
CACNA1Sex29	Fw	GCTACTTTGGAGACCCCTGGAA
	Rev	AGGAGGGTTTCGCACTCCTTCTG
ANK2ex21	Fw	TACCTCCAGACCCCAACATC
	Rev	TCTTTACCACGGTGTGTCCATG
BIN1ex7	Fw	TCAAGTCACGCATTGCCAAGC
	Rev	GAAACCTACGCGGCTGTCCACAG
NFIEx7	Fw	GAGCCCTGTTGATGACGTGTCTA
	Rev	CTGCACAAACTCCTTCAGTGAGTC
MBNL1ex5	Fw	AGGGAGATGCTCTCGGGAAAAGTG
	Rev	GTTGGCTAGAGCCTGTTGGTATTGG
MBNL1ex7	Fw	CCAACAGGCTCTAGCCAACA
	Rev	CAGATGTTGTTGCTGCGGAC
MBNL2ex5	Fw	ACACTTGCAAGCCAAAAATC
	Rev	CTGCTGGTAGTGCAAGACG
KIF13Aex32	Fw	ACCTGTGCAGCATTACGGGACAC
	Rev	CTCGTCGTTTAATGAGTGCATCTG
MYOM1ex18	Fw	GGTCAGAGCAGTCAATGCAG
	Rev	ATCACAGGGTGGAGATGGTG
SORBS1ex25	Fw	ACTTGGAATCCACGGAAGAA
	Rev	TGTAGGTGCGTGGGAAGATT
SORBS1ex3-5	Fw	GCTGTGATGAATGGCTTGGC
	Rev	CACCTCTGTAGCCCTTGGCA
SORBS1ex7	Fw	TGCCAAGGGTACAGAAGTG
	Rev	TGGAGGGATAGGCGGATGAA
SORBS1ex12-15	Fw	ACACTGAGAGATCAAAAGACTGGT
	Rev	TGAGTATCTGGGAGAGTACAGAT
SORBS1ex19-21	Fw	CCTGGCAAGTCTCCGTCTCT
	Rev	TTTCCGGAGTAGGGCTGATG
SORBS1ex29-33	Fw	GCGACCACTGGTGA AAAACC
	Rev	CGATATCTCCATCGCGGAGT

Primers used for qPCR analysis

SORBS1 total (1)	Fw	AGAAGGTAGTCAAGAGGTCGG
	Rev	CGTAAATGCTCTTGGGCTCT
SORBS1 total (2)	Fw	CCCAAAAAGTTGACACCAAGTGC
	Rev	CCCTTCGTACCAGTTCTCATCT
SORBS1ex25 (1)	Fw	TCATTCCGACGCACCATCAG
	Rev	AATTGGCTCTGGCAGGTCT
SORBS1ex25 (2)	Fw	ACTTGGAATCCACGGAAGAAT
	Rev	AACTGATGGTGGTCCGGAAT

Antisense oligonucleotides (ASO) used for human *SORBS1* exon 25 skipping

2'-OMePS-ASO-SA25	AAGUUUCUGCCAUAAAAUUGCAUA
2'-OMePS-ASO-SD26	CAAAACAGACCUCUGACCCAGAUU
2'-OMePS-ASO-ESE	CCUCUUCUUGCUCGCGUUUAAGUC
2'-OMePS-ASO-BP	UGCAUAAAAUGGGCAGUGAAUCAU

2'-OMePS-ASO : 2'-O-methyl RNA antisens oligonucleotide with phosphorothioate links

Supplementary Table S1

gRNAs, primers, ASO and restriction enzymes used in our study.

guideId	guideSeq	offtargetSeq	mismatchPos	mismatchCount	miOfftargetScore	cdOfftargetScore	chrom	start	end	strand	locusDesc
47forw	GGACACAAAATGGCTAACACTGG	CTACGTAAATGGCTAACACGGG	***	4	0.789491758	0.303896104	chr3	161373172	161373194	-	intergenic:SPTSSB-RP11-3P17.4
47forw	GGACACAAAATGGCTAACACTGG	GGGCATAAACAGCTAACACTGG	***	4	0.263094174	0.166358025	chr14	88466738	88466760	+	exon:PTPN21
47forw	GGACACAAAATGGCTAACACTGG	GGAAACACAGTGGCAAAACAGG	***	4	0.35314247	0.116079673	chr9	12709850	12709872	-	exon:TYRP1
47forw	GGACACAAAATGGCTAACACTGG	GGCCACAGATGGCAAACTAGG	***	4	0.13178369	0.007975708	chr11	55572610	55572632	-	exon:CRAC16
47forw	GGACACAAAATGGCTAACACTGG	GGTCACAAATGGTTGACACAGG	***	4	0.037983778	0.073001508	chr4	65601747	65601769	+	exon:EPH45
47forw	GGACACAAAATGGCTAACACTGG	GAACCCAAATGGCTAACACAGG	***	4	0.372455417	0.04479638	chr1	186932766	186932788	-	exon:PLA2G4A
47forw	GGACACAAAATGGCTAACACTGG	GGACTAAAATGGCTAACACAGG	***	4	0.104467922	0.01919649	chr15	80375487	80375509	+	exon:RP11-210M15.2
47forw	GGACACAAAATGGCTAACACTGG	GGACACAAAATGGCAAGTCAAGG	***	4	0.011111171	0.007530931	chr6	31673304	31673326	-	exon:LY6G5B/CSNK2B-LY6G5B-562
47forw	GGACACAAAATGGCTAACACTGG	GGACACAAAATGGCTAACAGG	***	4	0.082720628	0.006459054	chr1	32365757	32365779	-	exon:ISDC1
47forw	GGACACAAAATGGCTAACACTGG	GGCCACAAATGGCTAACACTGA	***	4	0.289186667	0.004415228	chr19	19491912	19491934	+	intron:GATAD2A
47forw	GGACACAAAATGGCTAACACTGG	AGACACAAAAGGCTACACTGG	***	4	0.058493115	0	chr19	19712541	19712563	+	exon:ZNF14
47forw	GGACACAAAATGGCTAACACTGG	AGACAAAAGGCTACACTGG	***	4	0.091442725	0	chr19	12077094	12077116	+	exon:ZNF84
47forw	GGACACAAAATGGCTAACACTGG	GAATATAAAATGGCTAACACAGG	***	3	1.46807152	0.628571429	chr15	52843989	52844011	-	intergenic:RP11-209K10.2-RP11-209E8.1
47forw	GGACACAAAATGGCTAACACTGG	GGACACAAAATGACTAACAAAGG	***	2	0.22828097	0.032051282	chr1	61540246	61540268	+	intergenic:NFA1-AC099791.1
17rev	CCATTTGTATCTCTGACTGGGG	CATTTAATATCTGTGACTGTGG	***	4	0.919204167	0.449197861	chr2	37906107	37906129	-	intergenic:LINC00211-RMDN2
17rev	CCATTTGTATCTCTGACTGGGG	CTGTATTATCTCTGACTGTGG	***	4	1.34583333	0.170940171	chr5	96919211	96919233	+	exon:ERAP2
17rev	CCATTTGTATCTCTGACTGGGG	CCATTTGGTTCTGTCTGGGG	***	4	0.187530741	0.062055591	chr5	115755707	115755729	-	exon:CTC-5005.03
17rev	CCATTTGTATCTCTGACTGGGG	CCATTTGTATCAGGACTCTGG	***	4	0.008795944	0	chr17	81947124	81947146	+	exon:MYADML2/AC137723.5
17rev	CCATTTGTATCTCTGACTGGGG	CCATTTGTGTCTGAGCTGTGG	***	4	0.019326615	0	chr16	67562554	67562576	+	exon:CTCF
17rev	CCATTTGTATCTCTGACTGGGG	CCATTTGTCTCCCTCAAGAGG	***	4	0.027628997	0	chr17	35859068	35859090	-	exon:HEATR9
17rev	CCATTTGTATCTCTGACTGGGG	CACTTTTATCTCTGACTGAGG	***	3	0.508701473	0.117845118	chr5	66091905	66091927	+	intergenic:ERBB2IP-AC025442.3/SREK1
17rev	CCATTTGTATCTCTGACTGGGG	CCATTTGTAGTCTGATTGAGG	***	3	0.136127904	0.114285714	chr6	46881485	46881507	+	exon:GPR116
17rev	CCATTTGTATCTCTGACTGGGG	TCCTTTTGGATCTGACTGGAG	***	3	0.321982391	0.068873069	chr18	2259371	2259393	-	intergenic:RP11-16116.2-METTL4
17rev	CCATTTGTATCTCTGACTGGGG	GCATTTGATCTCTGACTGGGA	***	3	0.347443106	0.037297527	chr1	109679947	109679969	+	intergenic:GSTM2/GSTM4-GSTM2
17rev	CCATTTGTATCTCTGACTGGGG	TCATATTGTATCTGACTGTGAG	***	2	1.202531646	0.12962963	chr8	60072464	60072486	-	intron:RP11-27P7.1
17rev	CCATTTGTATCTCTGACTGGGG	CCAGTTTGTACTCTGACTGAG	***	2	0.786940299	0.108024691	chr3	155885019	155885041	-	intron:GMP5
17rev	CCATTTGTATCTCTGACTGGGG	CCATTTTGAACCTGACAGTAG	***	2	0.263625	0.055555556	chr2	53593130	53593152	-	intergenic:RNU6-997P-GPR75-ASB3
17rev	CCATTTGTATCTCTGACTGGGG	CCATTTTGTCTGACTGAGG	***	2	0.407958614	0.042420814	chr9	91992561	91992583	+	intergenic:RO2-SPTLC1
17rev	CCATTTGTATCTCTGACTGGGG	ACATTTGTATCTCTGACTGTGA	***	2	0.76633333	0.039682539	chr16	62886263	62886285	-	intergenic:AC010546.1-RP11-96H17.1
194forw	ACACTTGCAGCCAGACTCAAGG	TCACTTAGCAGCCAGACTCAAGG	***	4	0.589608261	0.367346939	chr10	121553004	121553026	+	intron:FGFR2
194forw	ACACTTGCAGCCAGACTCAAGG	GGACTTGGAGCCAGACTCAAGG	***	4	0.833979885	0.298136646	chr2	217475745	217475767	+	intergenic:DIRC3-AS1-DIRC3
194forw	ACACTTGCAGCCAGACTCAAGG	AAATTTGGAGCCAGACTCAAGG	***	4	0.60958333	0.183732057	chr17	78228553	78228575	+	exon:BRCC2
194forw	ACACTTGCAGCCAGACTCAAGG	AGAAATCTAGCCAGACTCAAGG	***	4	0.329465174	0.164983888	chr8	34130811	34130833	+	intergenic:RP11-317N12.1-RP11-431M3.1
194forw	ACACTTGCAGCCAGACTCAAGG	GCCCTTGCAGCCAGACTCAAGG	***	4	0.420174266	0.146938776	chr8	143810500	143810522	-	exon:SCRIB
194forw	ACACTTGCAGCCAGACTCAAGG	CCCGTGCAGCCAGACTCAAGG	***	4	1.286675824	0.117551021	chr17	48980652	48980674	-	intergenic:GIP-RP11-501C14.5
194forw	ACACTTGCAGCCAGACTCAAGG	AGATTTGCAAGCCAGACTCAAGG	***	4	0.026676601	0.067571644	chr3	152445455	152445477	+	exon:MBNL1
194forw	ACACTTGCAGCCAGACTCAAGG	AGCATGCAAGCCAGACTCAAGG	***	4	0.257335165	0.032545131	chr4	39964413	39964435	+	intergenic:RNASSP159-AC098591.1
194forw	ACACTTGCAGCCAGACTCAAGG	GGAAATGAAAGCCAGACTCAAGG	***	4	0.272988506	0.028442919	chr20	37615767	37615789	+	intergenic:RP11-425M5.8-LINC00489
194forw	ACACTTGCAGCCAGACTCAAGG	ACACTTGCAGCCAGACTCAAGG	***	4	0.01256604	0.02641026	chr4	75049743	75049765	-	exon:PARG1
194forw	ACACTTGCAGCCAGACTCAAGG	ACACATGCAAGCCAGACTCAAGG	***	4	0.030398918	0.022222222	chr22	20064998	20065020	+	intergenic:TPM2/AC006547.15/AC006547.13
194forw	ACACTTGCAGCCAGACTCAAGG	ACACCAGCCAGCCAGACTCAAGG	***	3	0.896991699	0.495238095	chr22	38267351	38267373	+	intergenic:RP1-506.6-TMEM184B
194forw	ACACTTGCAGCCAGACTCAAGG	AGCCTAGCAAGCCAGACTCAAGG	***	3	0.289903704	0.06988728	chr9	154403463	154403485	+	intergenic:SAP30L-AS1-GALNT10
194forw	ACACTTGCAGCCAGACTCAAGG	CCATTTAAAGCCAGACTCAAGG	***	3	0.347443106	0.038690476	chr10	25565643	25565665	+	intron:APBB1P
194forw	ACACTTGCAGCCAGACTCAAGG	TCATTTGAAAGCCAGACTCAAGG	***	3	0.508701473	0.036111111	chr19	6574558	6574580	+	intergenic:Y_RNA-CD70
194forw	ACACTTGCAGCCAGACTCAAGG	ACAAATTCAGCCAGACTCAAGG	***	3	0.316898657	0.016708438	chr9	99712089	99712111	-	intron:RP11-554F20.1
194forw	ACACTTGCAGCCAGACTCAAGG	CCATTTGAAAGCCAGACTCAAGG	***	2	1.417910448	0.038690476	chr8	101610816	101610838	+	intron:GPHL2

Supplementary Table S2

Off-target sites analyzed for each gRNA used in our study.

Supplementary Table S3 and S4, related to Figure 5 and Supplementary Figure 6 GSE161029 (GEO database)

Ion Ampli-Seq data (DEGs of each list and the 12 DEGs common in *MBNL1* KO vs WT and DKO vs WT) (Table S3) and enriched biological processes of DEGs (EnrichR GO analyses) (Table S4).

Supplementary Tables S5 and S6, related to Figure 5 and Supplementary Figure 6 GSE 161897 (GEO database)

RNA-Seq data (DSEs of each list, the 126 DSEs common to the three lists and the 344 DSEs only found in DM1 hiPSC-derived myotubes vs WT) (Table S5) and enriched biological processes of the DSEs of each list (EnrichR GO analyses) (Table S6).

Supplementary Table S7, related to Figure 5 and Supplementary Figure 6

Correspondance_gene_RefSeq_amplicons.

Antibodies used for Immunocytochemistry

Mouse anti-Pax3	DSHB® (#PAX3; 1:30)
Goat anti-Desmin	R&D Systems® (#AF3844; 1:100)
Mouse anti- α -Actinin	Sigma® (#A7811; 1:250)
Mouse anti-MF20	DSHB® (#MF20; 1:200)
Mouse anti-MBNL1	DSHB® (#MB1a(4A8); 1:100)
Mouse anti-MBNL2	DSHB® (#MB2a(3B4); 1:100)
Rat anti-AChR	DSHB® (#mAb35; 1:200)
Rabbit anti-SORBS1	Abcam® (#ab224129; 1:400)
Donkey anti-Mouse IgG (H+L) Alexa Fluor Plus 488	Invitrogen® (#A32766; 1:1000)
Donkey anti-Goat IgG (H+L) Alexa Fluor Plus 488	Invitrogen® (#A32814; 1:1000)
Donkey anti-Mouse IgG (H+L) Alexa Fluor Plus 555	Invitrogen® (#A32773; 1:1000)
Donkey anti-Rat IgG (H+L) Alexa Fluor Plus 488	Invitrogen® (#A21208; 1:1000)
Donkey anti-Rabbit IgG (H+L) Alexa Fluor Plus 647	Invitrogen® (#A32795; 1:1000)

Antibodies used for pluripotency characterization by Flow Cytometry

Alexa Fluor 647 Mouse anti-human TRA-1-81	Clone TRA-1-81	Biolegend® (#330706; 1:50)
Alexa Fluor 647 Mouse IgM, kappa Isotype Control	Clone MM-30	Biolegend® (#401618; 1:50)
PE Rat anti-human SSEA-3	Clone MC-631	Biolegend® (#330312; 1:10)
PE anti-human SSEA-4	Clone REA101	Miltenyi® (#130-122-914; 1:50)
PE Rat IgM, kappa Isotype Control	Clone RTK2118	Biolegend® (#400808; 1:50)
PE REA, Isotype Control	Clone REA293	Miltenyi® (#130-113-462; 1:50)

Supplementary Table S8

Antibodies used in our study.

Article 2 : Altération de la connectivité développementale dépendante des protéines MBNL au niveau de la communication neuromusculaire dans la DM1

MBNL dependent-impaired developmental connectivity within neuromuscular circuits in Myotonic Dystrophy type 1.

Julie Tahraoui-Bories¹, Antoine Mérien¹, Anchel Gonzalez-Barriga³, Jeanne Lainé³, Céline Leteur², Hélène Polvèche², Alexandre Carteron², Juliette Duchesne De Lamotte⁴, Jérôme Polentes², Margot Jarrige², Mario Gomes-Pereira³, Denis Furling³, Laurent Schaeffer⁶, Claire Legay⁵ and Cécile Martinat¹.

¹ INSERM/ UEVE UMR 861, Université Paris Saclay, I-STEM, 91100 Corbeil-Essonnes, France.

² CECS/AFM, I-STEM, 91100 Corbeil-Essonnes, France.

³ Sorbonne Université UPMC Université Paris 06, INSERM, CNRS, Centre de Recherche en Myologie, Institut de Myologie, Paris, France.

⁴ Ipsen Innovation, Les Ulis, France.

⁵ Université de Paris, CNRS, SPINN – Saint-Pères Paris Institute for the Neurosciences, Paris, France.

⁶ INMG, Inserm U1217, CNRS UMR5310, Université Lyon 1, Université de Lyon, Lyon, France.

* **Corresponding author:** C. Martinat, PhD, I-Stem, 28 rue Henri Desbruères 91100 Corbeil-Essonnes France, Tel: + 00 33 1 69 90 85 33, Fax: + 00 33 1 69 90 85 21, E-mail: cmartinat@istem.fr

Afin d'explorer plus largement la contribution des protéines MBNL dans la spécification cellulaire, j'ai participé, en collaboration avec une autre étudiante en thèse au sein de l'équipe, à une étude visant à mieux comprendre les mécanismes physiopathologiques impliqués dans les atteintes neuromusculaires de la DM1.

En effet, plusieurs études ont montré des atteintes du système neuromusculaire dans la DM1 avec par exemple l'observation de foci au sein de noyaux de motoneurons ainsi que dans des noyaux sous-synaptiques au niveau de jonctions neuromusculaires issues de biopsies de patients post-mortem (Wheeler et al. 2007a). Plus récemment, des défauts de transmission synaptique ont également été observés chez les patients DM1 (Bombelli et al. 2016) ainsi que des défauts neuritiques *in vitro*, identifiés précédemment par l'équipe dans des motoneurons spinaux dérivés d'hiPSC DM1 (Marteyn et al. 2011) mais qui étaient basés sur un protocole de différenciation long et peu efficace.

En se basant sur le développement récent au sein de l'équipe d'un protocole de différenciation optimisé qui permet la conversion de cellules souches pluripotentes humaines en une population homogène de neurones moteurs spinaux (MNs) (Maury et al. 2015), l'objectif de cette étude a visé à mieux comprendre les mécanismes par lesquels la mutation causale de la DM1 pouvait modifier la fonctionnalité de ces cellules.

Nous avons pu confirmer dans un premier temps le défaut d'arborisation neuritique précédemment décrit par l'équipe et nous avons également retrouvé ce défaut dans les MNs dérivés des hiPSC *MBNL* KOs, suggérant ainsi un rôle important de ces protéines au niveau des mécanismes dérégulés dans les MNs DM1.

Afin d'évaluer les conséquences fonctionnelles de ces anomalies, un modèle de co-culture a été développé entre MNs dérivés d'hiPSC et cellules musculaires primaires saines, notamment dans le but de définir si la mutation causale de la pathologie, et plus particulièrement la perte de fonction des protéines MBNL, pouvaient affecter la communication entre les MNs et leur cible musculaire. Ainsi, cette étude a pu montrer, par des analyses fonctionnelles, des altérations de la connectivité entre les deux types cellulaires, en observant notamment des défauts de clusterisation des AChR ainsi qu'une dérégulation du flux calcique dans les cellules musculaires lorsque celles-ci sont en présence de MNs dérivés d'hiPSC DM1 ou dérivés d'hiPSC *MBNL* KOs, avec généralement un phénotype plus sévère dans le cas de MNs DKO.

Enfin, une analyse transcriptomique comparative a permis d'identifier plus de 320 défauts d'expression génique ainsi que 236 défauts d'épissage alternatif dans les MNs DM1 comparés au MNs WT. De manière intéressante, on retrouve environ 40% de ces gènes également dérégulés dans les MNs DKO et environ 20% des défauts d'épissage, soulignant d'une part le rôle important des protéines MBNL et suggérant d'autre part l'implication d'autres mécanismes pathologiques. Les analyses par Gene Ontology ont révélé la dérégulation de certains procédés biologiques tels que la "signalisation trans-synaptique antérograde" ou encore le "développement du système nerveux".

Ensemble, ces résultats montrent des défauts pré-synaptiques dans les MNs DM1, qui semblent régulés par les protéines MBNL et qui altèrent la communication entre le motoneurone et la cellule musculaire. Ceci suggère l'importance d'une part du compartiment pré-synaptique dans la physiopathologie de la DM1, et d'autre part de la contribution des protéine MBNL dans les mécanismes impliqués.

Concernant ma contribution à cette étude, et outre l'utilisation des lignées hiPSC *MBNL* KOs et des contrôles qualité nécessaires que j'avais pu générer, j'ai participé à la démarche scientifique générale et aider à l'analyse ainsi qu'à la discussion des résultats.

MBNL dependent-impaired developmental connectivity within neuromuscular circuits in Myotonic Dystrophy type 1.

Julie Tahraoui-Bories¹, Antoine Mérien¹, Anchel Gonzalez-Barriga³, Jeanne Lainé³, Céline Leteur², Hélène Polvêche², Alexandre Carteron², Juliette Duchesne De Lamotte⁴, Jérôme Polentes², Margot Jarrige², Mario Gomes-Pereira³, Denis Furling³, Laurent Schaeffer⁶, Claire Legay⁵ and Cécile Martinat¹.

¹ INSERM/ UEVE UMR 861, Université Paris Saclay, I-STEM, 91100 Corbeil-Essonnes, France.

² CECS/AFM, I-STEM, 91100 Corbeil-Essonnes, France.

³ Sorbonne Université UPMC Université Paris 06, INSERM, CNRS, Centre de Recherche en Myologie, Institut de Myologie, Paris, France.

⁴ Ipsen Innovation, Les Ulis, France.

⁵ Université de Paris, CNRS, SPINN – Saint-Pères Paris Institute for the Neurosciences, Paris, France.

⁶ INMG, Inserm U1217, CNRS UMR5310, Université Lyon 1, Université de Lyon, Lyon, France.

* **Corresponding author:** C. Martinat, PhD, I-Stem, 28 rue Henri Desbruères 91100 Corbeil-Essonnes France, Tel: + 00 33 1 69 90 85 33, Fax: + 00 33 1 69 90 85 21, E-mail: cmartinat@istem.fr

ABSTRACT

Myotonic dystrophy type I (DM1) is one of the most frequent muscular dystrophies in adults. Although DM1 has long been considered mainly as a muscle disorder, growing evidence suggests the involvement in peripheral nerves in the pathogenicity of DM1 raising the question whether motoneurons actively contribute to neuromuscular defects in DM1. By using a micropatterned 96-well plate as a co-culture platform, we generated a functional humanized cellular model combining DM1 hiPSC-derived MNs and healthy skeletal muscle cells. Such approaches led to the identification of pre-synaptic defects which affect development or stability of the neuromuscular junction at an early developmental stage. These neuropathological defects could be reproduced by the loss of RNA-binding MBNL proteins, whose loss of function is associated with muscular defects associated with DM1. These experiments suggested that the functional defects associated to MNs can be directly attributed to the MBNL family proteins. Altogether, these findings hold several new implications for DM1 pathogenesis.

Keywords: Myotonic Dystrophy type 1, Motoneurons, Induced pluripotent stem cells, *Muscleblind* proteins, neuromuscular junction.

INTRODUCTION

Myotonic Dystrophy type 1 (DM1) is a dominant hereditary muscular dystrophy which results from the expansion of an unstable CTG repeat in the 3' untranslated region of *DM protein kinase* (*DMPK*) (Brook et al. 1992; Mahadevan et al. 1992). DM1 is considered as the most common adult-onset muscular dystrophy with considerable clinical heterogeneity, mainly characterized by muscle weakness and myotonia but also systemic manifestations involving cognitive impairments, cataract, cardiac conduction abnormalities. A certain correlation has been observed between the severity of the symptoms associated with DM1 and the number of CTG repeats with longer CTG repeat expansions correlating with an earlier age of onset and more severe disease.

At the level of skeletal muscles, the main DM1 pathogenic process arises from a RNA-mediated mechanism involving the dysfunction of alternative splicing regulator proteins by mutant DMPK transcripts which results in inappropriate expression of developmental splice isoforms in adult tissues (Ho et al. 2004). One of the most well described mechanism at the origin of this mis-regulation of alternative splicing resides in the direct sequestration of proteins from the MBNL family within nuclear aggregates formed by mutant RNAs containing expanded CUG repeats (Fardaei, 2002; Miller et al. 2000). Dozens of splicing events are thus found to be mis-regulated in skeletal muscle biopsies from DM1 patients (Nakamori et al. 2013). A couple of these events have been correlated directly to DM1 symptoms such as, the myotonia that characterized DM1 has been associated to the abnormal splicing of the skeletal muscle-specific chloride channel 1 (ClC-1) or muscle weakness more recently associated with the defective splicing of DMD exon 78 (Charlet-B. et al. 2002; Mankodi et al. 2002; Kanadia et al. 2006).

Although DM1 has long been considered mainly as a muscle disorder, growing evidence suggests the involvement of peripheral nerves in the pathogenicity of DM1. Several abnormalities, including multiple endplates on the same muscle fiber, abnormally endplate size and expanded terminal arborizations have been observed in patients' biopsies (Allen et al. 1969; Coërs et al. 1973). More recently, neurophysiological studies in two different cohorts of DM1 patients have reported neuromuscular transmission abnormalities suggesting defective axonal transmission (Bombelli et al. 2016; Krishnan et al. 2006). These observations were at a certain extent confirmed in a severe murine model of DM1 in which abnormalities in the size and the complexity of endplates were observed in the diaphragm of these animals (Panaite et al. 2013). However, the mechanisms by which the presence of the mutation might affect the neuro-muscular system are not clearly understood.

The detection of nuclear mutant RNA aggregates in association with MBNL proteins in motor neurons from post-mortem biopsies of DM1 patients (Wheeler et al. 2007) raises the possibility, as described for skeletal muscle cells, of a toxic RNA mediated mechanism within the motoneuronal compartment which may affect development, function or stability of the neuromuscular junction. Especially, the potential contribution of the MBNL proteins within these defects remains an open question.

To test this proposal, we took advantage of the recent progress made to date towards the combination of human pluripotent stem cell technology with advanced tissue culture system to recapitulate human neuromuscular structure and function (Barbeau et al. 2020). Thus, we based our approach both on our recent optimized protocol to rapidly and efficiently convert hPSCs into homogenous population of spinal motoneurons (hPSCs-spinal MNs) (Maury et al. 2015) and on the development of co-cultures between hPSC-spinal MNs and human skeletal muscle cells using micropatterned 96 well plates (Young et al. 2018). Combining them with human induced pluripotent

stem cells (hiPSC) derived from DM1 patients or CRISPR/Cas9-modified, we investigated the effect of the DM1 mutation as well as the depletion of MBNL proteins on neuromuscular connectivity.

RESULTS

Depletion of MBNL1 and MBNL2 proteins affects neuritogenesis as observed in DM1 hiPSC-derived spinal MNs

Our group has previously demonstrated the potential of using human embryonic stem cells carrying the causal mutation of DM1 (DM1_hESC) to confirm the presence of nuclear foci and neuritic outgrowth abnormalities within DM1_hESC derived motoneuronal cultures (Marteyn et al. 2011). Nonetheless, this study was based on the use of a long and ineffective protocol of differentiation giving rise to only 30% of spinal motoneurons in 25 days with a high proportion of contaminating non-relevant cells that clearly hampered efforts to pathological modeling.

To circumvent this limitation, we based our approach on our recent protocol allowing the efficient specification of hPSCs into homogenous population of spinal motoneurons (Supplementary Figure 1A) (Maury et al. 2015). We applied this protocol to differentiate hiPSC lines recently developed (Mérien et al. 2021) which include: one hiPSC line derived from a non-affected individual and three independent DM1 hiPSC lines derived from different patients (named DM1_1, DM1_2 and DM1_3) (Mérien et al. 2021). In addition, we generated one isogenic DM1 hiPSC line in which the 2700 CTG expansion has been completely excised by CRISPR/ Cas 9 technology (DM1_3_ΔCTG) (Supplementary Figure 1B). According to the protocol of differentiation, we generated around 70% of motoneurons regardless of the cell line demonstrating that the specification into MNs is not affected by the presence of DM1 mutation (Figure 1A and 1B). Mutant RNAs aggregates (Figure 1C) were detected in approximately 60% of DM1 hiPSC-derived MNs, while none were observed in isogenic DM1_3_ΔCTG hiPSC-derived MNs (Figure 1D). Validating our previous study, an increased neuritic outgrowth was observed specifically in DM1 hiPSC-derived MNs whereas no defect was observed in DM1_3_ΔCTG hiPSC-derived MNs (Figure 1E).

Due to their pivotal role in the muscular pathogenicity of DM1, we sought to investigate the contribution of MBNL1 and MBNL2 proteins in this neuritic phenotype by taking advantage of different hiPSC lines in which the genes coding these proteins have been depleted using CRISPR/Cas9 technology (Mérien et al. 2021). We first monitored the expression of these different paralogues at the mRNA and protein levels at different times along the differentiation of WT-hPSCs into spinal MNs. Their level of expression gradually increased as differentiation progresses to reach a maximum around ~30 days of differentiation, corresponding to mature MNs (Supplementary Figure 2A-B). The analysis at the protein level demonstrated a specific lack of expression of the MBNL paralogues in the different gene-edited hiPSC lines differentiated into MNs (Supplementary Figure 2C). As reported in mouse model of *Mbnl* loss of function (Lee et al. 2013), we also observed a compensatory increase in MBNL1 and MBNL2 expression in *MBNL2*(-/-) and *MBNL1*(-/-), respectively (Supplementary Figure 2D). To avoid any functional compensatory phenomenon between the two paralogues, we also investigated the effects of combined MBNL1 and MBNL2 deficiency by using a double knock-out hiPSC line (DKO) (Supplementary Figure 2B-D). Whereas depletion of MBNL1 and MBNL2 did not affect the generation of MNs (Figure 1B), quantification of the neuritic length revealed a significant increased neuritic arborization in *MBNL1*(-/-), *MBNL2*(-/-) and DKO hiPSC-derived MNs as observed in DM1 condition (Figure 1E).

A 96-well plate co-culture to study the communication between human skeletal muscle cells and hiPSC-derived motoneurons.

Our findings suggested that sequestration and functional inhibition of MBNL1 and MBNL2 paralogues by mutant mRNAs may affect the pre-synaptic development of MNs. Along those lines, it has been recently shown that the down-expression of *Mbnl1* in murine hippocampal neurons *in vitro* led to neurite outgrowth defects (P. Y. Wang et al. 2018) that the deprivation of both *Mbnl1* and *Mbnl2* also resulted in defective synaptic densities in murine brain (K. Y. Lee et al. 2019). Altogether, these findings prompted us to investigate whether this modified pre-synaptic development of MNs may affect their communication with their skeletal muscle targets.

We thus sought to develop an *in vitro* humanized co-culture between hiPSC-derived MNs and human skeletal muscle cells compatible with cellular and functional analyses. For this purpose, we used micropatterned 96 well plates in which each well contains around 200 micropatterns that were shown to favor the formation of aligned and striated mature myofibers from human primary skeletal muscle cells (Young et al. 2018). The use of a 96 well plate format also offers the possibility to analyze in parallel a large number of conditions (Figure 2A).

To evaluate the maturity and the functionality of this co-culture system, we first characterized co-cultures generated with healthy cells. Thus, motoneurons derived from WT hiPSC were seeded on top of human primary skeletal muscle cells derived from a healthy control (Figure 2B). The morphology of the co-cultures was followed by a timelapse phase contrast microscopy showing the development of neurite arborization with the fusion of skeletal muscle cells (Supplementary Figure 3A and 3B). The characterization for different markers confirmed appropriate maturation of hiPSC-derived motoneurons (*ISL1*) and myogenic differentiation in mature myotubes (Myosin Heavy Chain, MHC) (Figure 2C-D). Whereas the addition of hiPSC-derived MNs did not affect the level of myotube area (Figure 2C), a significant increase of AChR clustering at the post-synaptic level was detected in the presence of hiPSC-derived MNs when compared to myotubes alone (Figure 2E and Supplementary Figure 3C). The detection and localization of different pre-synaptic markers such as synaptophysin, the vesicular acetylcholine transporter and bassoon, in close proximity of AChR clusters confirmed the presence of endplate maturation (Figure 3A). These observations were validated by transmission electron microscopy revealing close contact between skeletal muscle cells and hiPSC-derived motoneurons (Figure 3B, left), the presence of electron-dense active zones in association with synaptic vesicles suggesting the AChR clustering at junctional zones (Figure 3B, right).

In term of functionality, muscular contractions were observed 4 days after the addition of hiPSC-derived MNs (Supplementary Figure 4A). Quantification of these twitch activities revealed a significant effect of hiPSC-derived MNs on the capacity of human primary skeletal muscle cells to contract (Figure 4A). Another important functional parameter consisted in the measurement of the capacity of hiPSC-derived MNs to induce Ca^{2+} transients in skeletal muscle cells. Extensive characterization of this parameter has been performed under different conditions. We first demonstrated that addition of hiPSC-derived MNs led to a significant increase in Ca^{2+} transients when compared to human primary skeletal muscle cells alone (Figure 4B and Supplementary Figure 4B). The specificity of these Ca^{2+} transients was therefore evaluated using pharmacological approaches. Thus, bath application of the neurotransmitter glutamate triggered a significant increase of the Ca^{2+} signal peaks (Figure 4C and Supplementary Figures 5A-B). In contrast, the excitation of skeletal muscle cells was totally abolished by d-tubocurarine, a competitive inhibitor of AChR, suggesting that the Ca^{2+} oscillations were mediated by AChR signaling (Figure 4D and Supplementary Figure 5C). Finally, treatment with Botulinium neurotoxin (BoNT/A), known to induce a prolonged blockage of neuroexocytosis, led to a significant reduction of Ca^{2+} activity validating that Ca^{2+}

oscillations observed in skeletal muscle cells are dependent of hiPSC-derived motoneurons (Figure 4E and Supplementary Figure 5D). These functional analyses indicate that these micropatterned co-cultures at this phase (day 7) recapitulate key features of early neuromuscular junction development and therefore offer new perspectives to study impairment of this inter-cellular structure in diseases.

Impaired communication of DM1 and MBNL depleted hiPSC-derived spinal MNs with skeletal muscle targets.

Based on these micropatterned co-cultures, we next evaluated the capacity of DM1 and MBNL depleted hiPSC-derived MNs to properly and functionally interact with skeletal muscle cells. With this aim, we designed different hybrid systems in which wild-type control primary skeletal muscle cells were co-cultured with DM1 or MBNL depleted hiPSC-derived MNs (Figure 5A). As previously observed in the cultures of hiPSC-derived MNs alone, a significant increased neurite outgrowth was quantified in DM1 and MBNL depleted conditions when compared to WT or CTG depleted DM1 hiPSC lines, phenotype that could not be attributed to a difference in the number of hiPSC-derived MNs or a difference of myogenic area (Figure 5B-C and Supplementary Figure 6A). In parallel to the neuritic phenotypes, quantitative analysis of the mean size of AChR clusters revealed impaired endplate maturation in DM1 and MBNL depleted hiPSC-derived MNs as compared to control or isogenic control DM1 (Figure 5D). Likewise, size distribution of AChR clusters revealed that even if the total area of AChR is not affected, AChR clusters characterized by a bigger size were more present in controls (WT and DM1_3_ΔCTG) whereas DM1 and MBNL depleted hiPSC-derived MNs resulted to a greater proportion of small AChR clusters (Figure 5D and Supplementary Figure 6C).

To evaluate the functional consequences of these phenotypes, Ca²⁺ oscillations induced hiPSC-derived MNs were measured at the level of skeletal muscle cells. A significant increase in the frequency of calcium transient was observed in the presence of DM1 and MBNL depleted hiPSC-derived MNs (Figure 6A-B and Supplementary Figure 7). In contrast, no difference was observed between WT and DM1-CTG depleted co-culture (Figure 6A-B). Because the muscular component consisted in healthy control skeletal muscle cells, we concluded that abnormal behavior of the myofibers could be attributable to the presence of DM1 and MBNL depleted hiPSC-derived spinal MNs. Our data strongly suggest that, under pathological conditions, the abnormal post-synaptic phenotypes depend on the exchange of a non-extracellular signal or signals between muscle and nerve cells.

In our previous study in which we observed a neuritic outgrowth in DM1 hESC-derived motoneuronal cultures, we identified that a defective expression of SLITRK2 gene can be correlated to this phenotype (Marteyn et al. 2011). Consistent with our previous results, we confirmed this misexpression in DM1 hiPSC-derived motoneurons as well as the normalization of this expression in CTG depleted DM1 hiPSC-derived motoneurons (Supplementary Figure 6D). Strikingly, the misexpression of SLITRK2 was not observed in the absence of MBNL proteins suggesting the involvement of additional molecular mechanisms in the synaptic defects observed.

Common misregulation of synaptic pathways in DM1 and MBNL depleted hiPSC-derived motoneurons.

To go further in understanding the molecular mechanisms that could explain the different phenotypes revealed in the miscommunication between motoneurons and skeletal muscle cells, we sought to conduct comparative transcriptomic analyses between DM1, MBNL depleted and WT hiPSC-derived motoneurons. A distinct transcriptomic profile between these three conditions was

validated by the principal component analysis (PCA) and heatmap representation (Figure 7A and Supplementary Figure 8A). The comparison between DM1 to wild-type hiPSC-derived motoneurons led to the identification of 328 genes differentially expressed, with 199 being up-regulated and 129 down-regulated in DM1 motoneurons compared to wild-type (Figure 7B). A more pronounced difference was observed between DKO and WT hiPSC-derived motoneurons with a total of 723 differentially expressed genes (DEG), 368 up-regulated and 355 down-regulated in DKO motoneurons (Figure 7B). All the differentially expressed genes (DEGs) are listed in Supplementary Table 4. Gene ontology analysis of DEGs in DM1 versus WT hiPSC-derived motoneurons revealed enrichment in pathways related to synaptic functions: “Nervous system development”, “Anterograde trans-synaptic signaling” and “Chemical synaptic transmission” (Supplementary Figure 8B). Strikingly, most of these synaptic related-biological processes were found to be down-regulated in DM1 hiPSC-derived motoneurons (Figure 7C). Similar observations were found in DKO hiPSC-derived motoneurons compared to WT (Figure 7D and Supplementary Figure 8C). Among the 328 DEGs observed in DM1 hiPSC-derived motoneurons, 137 (42%) were also found deregulated in DKO hiPSC-derived motoneurons (Figure 7E) suggesting the important role played by MBNL depletion in these deregulations. Interestingly, functional annotation of these common DEGs also revealed that the most relevant GO terms concerned “Nervous system development”, “Anterograde trans-synaptic signaling” and “Chemical synaptic transmission” (Figure 7E). Among the down-regulated genes identified in the common list, several of them caught our attention (Figure 7F and Supplementary Figure 8D). Thus, DLK1 gene has been shown to regulate several aspects of neuronal development and maintenance, including synaptic development (Nakata et al. 2005; Kurup et al. 2015). Another example is CXCR4 for which a murine knockout model revealed aberrant axonal projections at the NMJ level (Lieberam et al. 2005). Finally, Wnt16 could be an interesting candidate to explain the post-synaptic defects observed in the presence of DM1 and DKO hiPSC-derived motoneurons, as the secretion of Wnt16 by motoneurons has been shown to modulate AChR clustering independently of agrin signaling pathway (Zhang et al. 2012). Interestingly, the misexpression of DLK1 and Wnt16 was also observed in MBNL2^(-/-) hiPSC-derived motoneurons (Supplementary Figure 9A).

Due to the important role played by MBNL in the regulation of alternative splicing, we pursued our analysis by evaluating the cassette exons demonstrating a significant difference in the level of inclusion between DM1, DKO and WT hiPSC-derived motoneurons. Functional analyses of the DSEs identified in these different conditions also revealed different terms related to synapse structure and function (Figure 8A-B). Interestingly, the analysis of the 232 DSEs identified in DKO versus WT hiPSC-derived motoneurons revealed terms related to axonal transport and synaptic plasticity such as “microtubule motor activity”, “actin binding” or “ATP-dependent microtubule motor activity” (Figure 8B) suggesting a role of MBNL proteins in these cellular processes. Among the 236 DSE observed in DM1 hiPSC-derived motoneurons, only 48 DSE (20%) were also observed in DKO condition (Figure 8C-D) suggesting that other mechanisms might be involved in the alternative splicing defects observed in DM1 condition. Functional annotation analysis of the common transcriptomic data revealed that the most relevant GO terms concerned neuronal outgrowth such as “axogenesis” or “axon guidance” (Figure 8C). Interestingly, among these DSE, CAST and COL13A1 were validated by RT-PCR and previously shown to be respectively involved in the organization of pre-synaptic active zones (Ohtsuka, 2013) and in the motor endplate maintenance (Härönen et al. 2019) (Supplementary Figure 8E).

As Mbnl2 has been shown to be an important splicing factor involved in the developing brain (Charizanis et al. 2012), we evaluated the consequences of its loss on the gene expression defects. Interestingly, among the 200 DEG in MBNL2^(-/-) hiPSC-derived motoneurons (Supplementary Figure

9A), 123 genes (61,5%) were common to DKO hiPSC-derived motoneurons (Supplementary Figure 9B). As observed in DKO and DM1 conditions, the enriched pathways deregulated in the absence of MBNL2 concerned the nervous system and synaptic functions (Supplementary Figure 9A). The gene ontology analysis of the 236 mis-spliced events observed in the MBNL2^(-/-) hiPSC-derived MNs (Supplementary Figure 9C) showed that pathways involved in protein localization and actin cytoskeleton were enriched. Less redundancy was however observed concerning the mis-splicing events with only 30% of common events in MBNL2^(-/-) and DKO conditions that might suggest a significant role of MBNL1 in these defects (Supplementary Figure 9D).

DISCUSSION

The main findings of the present study were the identification of functional neuropathological defects in DM1 that impaired the interaction between hiPSC-derived motoneurons and their muscular targets at an early developmental stage. These neuropathological defects could be reproduced by the loss of RNA-binding MBNL proteins, mimicking the neuro-muscular defects associated to the expression of the DM1 mutation.

Although DM1 is traditionally considered to be a disease that affects the muscles, DM1 present many other manifestations including multiple debilitating neurological manifestations (Gourdon and Meola 2017). Potential defects in synaptogenesis have been described in a severe murine model of DM1 as well as in the frontal cortex of DM1 patients (Hernández-Hernández et al. 2013; Sicot et al. 2017). At the neuromuscular level, structural and electrophysiological studies in DM1 patients have also suggested the involvement of the neuronal compartment (Bombelli et al. 2016; Krishnan and Kiernan 2006). Thus, irregular and expanded terminal arborizations and profuse axonal proliferation, multiple endplates on the same muscle fiber and abnormally large endplate size have been observed in patients (Macdermot, 1961; Allen et al. 1969; Coërs et al. 1973; Fardeau and Tome, 1980). Furthermore, the DMPK protein was proposed to have a role in the formation of the neuromuscular junction and signal transmission (van Der Ven et al. 1993; Whiting et al. 1995). Finally, one of the most striking observation was the detection of mutant mRNA aggregates in the nuclei of spinal motoneurons in DM1 patients (Wheeler et al. 2007) raising the question whether motoneurons actively contribute to neuromuscular defects in DM1. Along those lines, our group has previously demonstrated the potential of using human pluripotent stem cells carrying the causal mutation of DM1(DM1_hPSC) to confirm the presence of nuclear foci and neuritic outgrowth abnormalities within DM1_hPSC derived motoneuronal cultures (Marteyn et al. 2011). Nonetheless, this study was based on the use of a long and ineffective protocol of differentiation giving rise only 30% of spinal motoneurons in 25 days with a high proportion of contaminating non-relevant cells that clearly hampered efforts to pathological modeling. To more broadly analyze the molecular and functional effects of the DM1 mutation in spinal motoneurons, we have taken advantage of our progress made in differentiation protocols to rapidly and efficiently convert hPSCs into homogenous and synchronous population of spinal motoneurons (Maury et al. 2015). Thanks to this development, we confirmed that the presence of DM1 mutation affects not only neuritogenesis but also the functional interaction between motoneurons and their muscular targets at an early developmental stage. Using a co-culture between these two cell types, our results indeed suggested that neuromuscular defects in DM1 may occur in an anterograde fashion with DM1 hiPSC-derived MNs that “fine-tune” muscle post-synaptic development. As the muscular component consisted in healthy control skeletal muscle cells, our results demonstrated that abnormal behavior of the myofibers could

be attributable to the presence of DM1 hiPSC-derived spinal MNs. Interestingly, the depletion of the long stretch of CTG repeats by CRISPR/ Cas9 rescued these pre/post synaptic phenotypes validating their direct association with the presence of the DM1 mutation.

At the mechanistic level, our data highlight a significant role of MBNL proteins in these neuromuscular phenotypes. Motoneurons generated from hiPSC lines depleted in the different paralogs MBNL1 and MBNL2 reproduced the early development defects observed in DM1 condition. Consistent with our observations, immature NMJs with fragmented AChR clusters have been previously described in *Mbnl1*^(-/-); *Mbnl2*^(+/-) mice (Lee et al. 2013). In the same line, *Caenorhabditis Elegans* Muscleblind homolog *mbl-1* was shown to be required for proper formation of neuromuscular junction synapses (Spilker et al. 2012). Surprisingly, expressing a *mbl-1* transgene in the presynaptic neuron was sufficient to rescue the synaptic defect, while muscle expression had no effect (Spilker et al. 2012). Whereas the major pathological features of the DM brain have been previously shown to be associated with the disruption of the MBNL2-mediated developmental splicing program (Charizanis et al. 2012), deletion of MBNL1 or MBNL2 led to the same functional impact on the synaptic defect at the neuromuscular level. We can therefore hypothesize that both paralogues might account for neuromuscular defects. These results raised the question of how and by which mechanisms MBNL proteins might be involved in the development of neuromuscular interactions.

Although purely speculative at this stage, the comparative transcriptomic analyses between DM1 and MBNL depleted hiPSC-derived MNs revealed the deregulation of several pathways related to nervous system development, anterograde trans-synaptic signaling and cytoskeletal elements of axonal transport. Interestingly, these results are concordant with previous transcriptomic analyses performed in the hippocampi of *Mbnl2*^(-/-) mice (Goodwin et al. 2015). Due to the large number of deregulated genes we identified, the hypothesis of a unique gene underlying the different phenotypes may be unlikely and combined misexpression of several actors involved in synaptic signaling pathways might be considered. MBNL proteins also have significant cytoplasmic expression and have been proposed to contribute to regulation of mRNAs stability (Du et al. 2010; Osborne et al. 2009; Wang et al. 2015; 2012; Masuda et al. 2012) or localization (Adereth et al. 2005; Wang et al. 2012). MBNL proteins could therefore determine the localization of mRNA by the splicing mechanism, long 3'UTR transcripts being trafficked to dendrites while short 3'UTR were retained in the soma (An et al. 2008; Shigeoka et al. 2016). Moreover, Wang and colleagues showed that MBNL1 could bind 3'UTR mRNA of SNAP25 and VAMP1, two synaptic proteins, to regulate their neuronal localization (Wang et al. 2012). MBNL proteins associated to mRNA form a messenger ribonucleoprotein (mRNP) complex itself integrated in granules allowing the transport of mRNA along axons thanks to a cytoskeletal machinery (Van Driesche and Martin 2018). These data led to an emerging hypothesis that RNA localization functions of MBNL are important for proper synapse function, and that mis-localized RNAs might account for some neurological features associated to DM1. Interestingly, the transcriptomic analysis of hiPSC-derived MNs revealed several enriched deregulated pathways involved in the cytoskeleton machinery, molecular motor, protein localization and organelle transport along microtubules in DM1 and MBNL depleted conditions. It could be therefore interesting to evaluate the synaptic vesicles trafficking along microtubules and the quantification of synaptic vesicles at the pre-synaptic active zone under the different hybrid conditions to assess the potential role of MBNL proteins.

Altogether our results suggested that DM1 might not be considered a muscle disease but rather a disease of the interaction between motor nerve and its muscle. We therefore suggest that MNs could also be considered as a target of future therapies to cope with the genetic defect leading to this fatal

neuromuscular disease. Consequently, the role of MNs in the establishment and the defective function of neuromuscular junctions in DM1 might deserve further attention and investigation.

MATERIAL & METHOD

hiPSC and differentiated cells

Human pluripotent stem cells were generated from fibroblasts derived from non-affected and DM1 affected patients as previously described (Mérien et al. 2021). Informed consents were obtained from all the patients included in this study, complying with the ethical guidelines of the institutions and with the legislation requirements. Experimental protocols were approved by the french minister of health (2019-A02599-48). Three DM1 human induced pluripotent stem cell lines (named DM1_1, DM1_2, DM1_3) were generated from three different patients' fibroblasts. The length of the repeat expansion was previously estimated to approximatively 2.700 CTG repeats in each DM1 hiPSC line generated (Mérien et al. 2021). Briefly, hiPS were grown on culture dishes-coated vitronectine (Gibco®) and maintained in iPS-Brew XF medium (Miltenyi Biotec®). Cell passaging was performed manually every 5 days and culture medium was changed every two days. Knock-out MBNL hiPS (MBNL1^(-/-), MBNL2^(-/-) and MBNL1^(-/-); MBNL2^(-/-) – named DKO) cells were generated from a non-affected hiPSC line by using CRISPR Cas9 technology as previously described (Mérien et al. 2021). CTG-depleted DM1 cell lines were developed from a DM1 cell line (DM1_3) by using CRISPR/ Cas 9 technology (Lo Scudato et al. 2019). The conversion of hiPSC into spinal motoneurons was performed as previously described (Maury et al. 2015). Briefly, hiPS were dissociated enzymatically using Stem Pro Accutase (ThermoFisher®) during 5min at 37°C, 5% CO₂ and plated in 25cm² flasks (Dutscher®) at 1,5million per flask in a induced-motoneuronal medium supplemented with cytokines every two days (Maury et al. 2015). After 14 days of differentiation, embryoid bodies were dissociated and motoneuron phenotype was assessed by immunolabeling for Islet1. Human primary control myoblasts (gift of D. Furling, Institute of Myology, Paris) were cultured in a growth medium and cryopreserved. After several cell passaging, primary myoblasts were purified on CD56 microbeads column (Miltenyi Biotec®) followed by immunolabelling for Desmin to estimate the percentage of myoblast cells.

Co-culture protocol

96 micropatterned-well plates were designed (Young et al. 2018) and provided by CYTOO SA®. At D0, myoblasts were seeded at 20 000 cells per well in DMEM-F12 glutamax (ThermoFisher®) supplemented by 20% fetal bovine serum (Sigma-Aldrich®) and 0,1% penicillin/streptomycine (ThermoFisher®) (see Table S1). At D2, after two washes with basal medium, the growth medium was replaced by the differentiation medium composed by DMEM-F12 glutamax and 2% horse serum (ThermoFisher®). At D4, hiPSC-derived MNs were seeded at 40 000 cells per well in DMEMF12-glutamax, neurobasal medium (Vol:Vol) supplemented with BDNF (Peprotech®), GDNF (Peprotech®), DAPT (Tocris®) and Y-27632 (Stem Cell®) (see Table S1). Co-cultures were maintained for at least 7 days.

Immunostaining

Motoneurons derived from hiPS were fixed with a solution containing 4% paraformaldehyde (PFA-Euromedex®) for 8min at room temperature. Co-cultures were fixed vol:vol with a solution containing 8% paraformaldehyde (PFA-Euromedex®) for 10min at room temperature. Primary antibodies were

diluted in PBS/ 1% BSA/0,2% Triton X-100 and incubated overnight at 4°C. Appropriated secondary antibodies were added for 1h at RT in the presence of Hoechst 33528 (ThermoFisher[®], 1mg/ml). Primary and secondary antibodies are listed in Table S2.

Fluorescent In Situ Hybridization (FISH)

FISH experiments were performed as previously described (Maury et al. 2019). Briefly, cells were fixed with PBS buffer containing 4% PFA (Electron Microscopy Science, 20µl/well) for 20 minutes at room temperature and washed with 200µL of PBS. Cells were then incubated overnight at 4 °C with 70% Ethanol (60 µl/well). After PBS wash, cells were rehydrated with a solution of PBS containing 5mM MgCl₂ for 15-30 min and then sequentially incubated with prehybridization buffer (50mM Phosphate buffer, 40% formamide, 2X SSC, 50µl/ well) for 15-30 min and hybridization buffer (Prehybridization buffer with 0,2% BSA and 7% Dextran, 50µl/well) containing 300ng/mL of the (CAG)₁₀-Cy5 probe (5'Alexa 647- TTCTTATTCTTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG3') (Operon) overnight at 37°C. Washing steps consisted in pre-warmed Washing buffer (Prehybridization buffer + 0.2% BSA, 60µL) for 15 min at RT followed by addition of washing buffer (80µL) and incubation for 30 min at 37°C. Cells were then washed twice in PBS before nuclei counterstaining Hoechst 33258 (InvitroGen[®], 5µg/mL). Plates were stored at 4°C until analysis. Nuclear foci detection was performed using the Cellomics Array Scan VTI high content imaging system (Thermofisher[®]) with the 20X objective to automatically focus cell preparation. Nuclear foci detection was analyzed using the Cellomics CX7 high content imaging system (Thermofisher[®]) with the 20X objective and an automatically focus cell preparation. Images were acquired in High resolution camera mode on two channels and quantification was performed as previously described (Maury et al. 2019).

Image acquisition and analysis

Images for hiPSC-derived motoneurons were acquired using ImageXpress micro imaging system (Molecular Devices[®]) and were analyzed using MetaXpress software for automated quantification of the percentage of cells stained for ISL1 compared to the total number of cells and using the bioapplication “find fibers” to measure the neuritic arborization. Images of co-cultures were acquired using Spinning Disk microscopy (Zeiss[®]) with the 20X and 40X objectives. Each image was taken with z-stack and an orthogonal view was performed on Fiji Software. For quantification of AChR clustering, images were analyzed by using Fiji Software to manually quantify total area and the mean size of AChR clusters (Algorithm in Annexe 1). By application of background correction and intensity threshold, the neuritic arborization was quantified by the determination of the neuritic length by micropattern. Islet1 quantification was determined by automatically counting the number of Islet1-positive cells by micropattern. Myotube area was quantified by delimitating the cells stained by MF20 antibody and application of intensity threshold.

From day 4 in co-culture, visual contractions were recorded in phase contrast using Spinning Disk (Zeiss[®]) under controlled atmosphere (37°C, 5% CO₂). Movies acquisition lasted 2 min and were analyzed using the detection of pixels variations on Fiji Software and then, visual contractions were analyzed by the kymograph tool.

Calcium Flux assay

Living cells were incubated in a HBSS^{+/+} solution (see Table S1) containing a Cal520 probe (2uM-Abcam[®]) at 37°C, 5% CO₂. After 25min, cells were washed twice with a HBSS^{-/-} solution and kepted at RT for 15min. Recording was performed using Spinning Disk Microscopy (Zeiss[®]) in a controlled

atmosphere (37°C, 5% CO₂). Basal acquisitions were set at 200ms intervals and two different fields were acquired per well. Images were processed and analyzed on Fiji Software using ROI manager to quantify fluorescence intensity inside each myotube (= 1 ROI) and corrected for background. The inhibition of synaptic activity was evaluated after different drug treatments. After a basal acquisition of each field, (1) 100uM of tubocurarine chloride (Sigma-Aldrich®) was applied to each well and a stream acquisition was performed during 1min; (2) the neurotoxin botulinium A (Ipsen®) was added at 5nM and acquisitions were performed after 4 hours and 8 hours of treatment. In contrast, the activation of synaptic activity was evaluated after treatment with L-Glutamic acid (100uM- Sigma-Aldrich®). Placebo experiments were conducted by adding drugs on myotubes alone without hiPSC-derived MNs or by treating co-cultures with a buffer solution that did not contain any of the chemical reagents. The curve of calcium response is represented as time as a function of the ratio $\Delta F/F_0$ in which F_0 represents the mean fluorescence intensity measured before drug treatment whereas ΔF represents the mean fluorescence (Young et al. 2018).

RNA isolation and RT-qPCR analysis

Total RNAs were isolated using the RNAeasy Mini extraction kit (Qiagen®) according to the manufacture's protocol. RNA levels and quality were checked using the NanoDrop® technology. cDNA was synthesized using SuperScript III (ThermoFisher®). Quantitative polymerase chain reaction (qPCR) analysis was performed using a QuantStudio 12K Flex real-time PCR system (Applied Biosystem®). Primers sequences are detailed in Table S3. All analyses were performed with at least three technical replicates per plate. The $2^{-\Delta\Delta}$ Ct method was used to determine the relative expression level of each gene. All expression data were normalized to three housekeeping genes (18S, cyclophilinA and GAPDH).

Protein extraction and Western blot analysis

Western blots analyses were performed as previously described (Maury et al., 2019). Briefly, cells were lysed in RIPA 1X buffer (Sigma®) containing protease inhibitors (Sigma®) and phosphatase inhibitors (Roche®). Proteins were quantified by Pierce BCA Protein Assay kit (Pierce®) using a multiplate colorimetric reader, CLARIOstar (BMG Labtech®). Protein extracts (20 to 30µg) were loaded on a 4-12% SDS-PAGE gradient (NuPage Bis-Tris gels, Invitrogen®) and transferred onto Gel Transfer Stacks Nitrocellulose membranes (Invitrogen®) using the iBlot2 Dry Blotting System (Invitrogen®). Membranes were then incubated overnight at 4°C with the following primary antibodies: MBNL1 - MB1a(4A8) (DSHB®, 1:1000), MBNL2 - MB2a(3B4) (DSHB®, 1:1000), SYP (ThermoFisher®, 1:500) and VACHT (Synaptic System®, 1:500). After hybridization of the peroxidase-conjugated secondary antibody (from 1:50000 to 1:10000), immunoreactive bands were revealed by using Amersham ECL Select Western Blotting Detection Reagents (GE Healthcare®). Equal protein loading was verified by the detection of β -Actin using the A3854 Monoclonal Anti- β -Actin-Peroxidase antibody (Sigma®, 1:10000).

Electronic Microscopy

Co-cultures were processed in situ as previously described (Lainé et al. 2018), except for a few modifications related to the use of glass instead of plastic coverslips. Following resin polymerisation, glass coverslips were dissolved with hydrofluoric acid, and plastic-embedded samples were glued on an epoxy block and cut into 70 nm ultrathin sections.

RNA-Seq library preparation and sequencing (TrueSeq polyA)

Total RNAs were isolated using the RNeasy Mini extraction kit (Qiagen®) according to the manufacturer's protocol. RNA levels were assessed using the NanoDrop® technology and quality control was evaluated by determining RNA Integrity Number (RIN) using the RNA screen tape technology (Agilent®). For each sample, 100ng of total RNA was used. mRNAs were purified using poly-T oligo from TruSeq Stranded kit (Illumina®) and were fragmented and retro-transcribed using the TruSeq Stranded kit following the protocol of the manufacturer (Illumina®). Final library was quantified using Agilent High Sensitivity DNA kit before the samples were pooled in sets of six for sequencing. 2nM of pooled libraries were denatured and the quantity of 1.8pM was used for cluster generation before paired-end sequencing on an Illumina Nextseq550 (High output 2*150pb run). Samples were sequenced with an average of 89 532 366 reads (paired-end) per sample. The quality control of the sequencing data was evaluated using FastQC (v0.11.2) ("Babraham Bioinformatics - FastQC A Quality Control Tool for High Throughput Sequence Data"). The reads were trimmed using Prinseq (v0.20.4) (Schmieder and Edwards 2011) (--trim-right 20) and filtered by average quality score (--trim-qual 20) and cutadapt (v1.16) (Martin, 2011). (Martin 2011) For the Alternative splicing the raw reads was annotated on GRCh37.75 Ensembl reference annotation by Tophat2 (Kim et al. 2013) (v2.0.13) and analyzed using only reads on exonic junctions with FaRLine pipeline (Benoit-Pilven et al. 2018) and FasterDB database (Mallinjou et al. 2014).

Statistical Evaluation

All data were processed using Prism 8®. Values are represented as mean ± SD. Student's t-test was used to analyze two groups of data (*p < 0,05; **p < 0,005; ***p < 0,0005; ****p < 0,0001). For comparisons of more than two groups, statistical analyses were performed with by ordinary one-way analysis of variance (ANOVA) using Tukey's multiple comparisons test. Values of p < 0,05 were considered significant (*p < 0,05; **p < 0,005; ***p < 0,0005; ****p < 0,0001).

AUTHOR CONTRIBUTIONS

J.T.B., C.L. and C.M.: conception and design. J.T.B. and C.M.: investigation and data collection. J.T.B., C.L., J.P., and C.M.: formal analysis and interpretation. J.L.: electronic microscopy. J.D.D.L.: Neurotoxin experiment. A.M., A.G., M.G.-P.: Engineered cell lines. D.F. and C.M.: provision of different cellular resources. C.M.: financial support. J.T.B. and C.M.: manuscript writing.

DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST: The authors declare no conflicts of interest.

ACKNOWLEDGEMENTS

I-Stem is part of the Biotherapies Institute for Rare Diseases (BIRD) supported by the Association Française contre les Myopathies (AFM-Téléthon). This project was supported by grants from Laboratoire d'Excellence Revive (Investissement d'Avenir; ANR-10-LABX-73), the Region Ile-de-France via the doctoral school « Innovation Thérapeutique, du fondamental à l'appliqué » (ED 569) from Paris Saclay University and Agence Nationale de la Recherche (ANR). We thank Lina El-Kassar and Karine Giraud-Triboult for karyotyping the cell lines and the bioproduction platform for the large-scale production of control myoblasts.

REFERENCES

- Adereth, Yair, Vincent Dammai, Nurgun Kose, Runzhao Li, and Tien Hsu. 2005. "RNA-Dependent Integrin A3 Protein Localization Regulated by the Muscleblind-like Protein MLP1." *Nature Cell Biology* 7 (12). <https://doi.org/10.1038/ncb1335>.
- Allen, D E, A G Johnson, and A L Woolft. 1969. "The Intramuscular Nerve Endings in Dystrophia Myotonica—a Biopsy Study by Vital Staining and Electron Microscopy." *J. Anat.* Vol. 105.
- An, Juan Ji, Kusumika Gharami, Guey-Ying Liao, Newton H. Woo, Anthony G. Lau, Filip Vanevski, Enrique R. Torre, et al. 2008. "Distinct Role of Long 3' UTR BDNF mRNA in Spine Morphology and Synaptic Plasticity in Hippocampal Neurons." *Cell* 134 (1). <https://doi.org/10.1016/j.cell.2008.05.045>.
- "Babraham Bioinformatics - FastQC A Quality Control Tool for High Throughput Sequence Data." n.d. Accessed February 23, 2021. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- Barbeau, Susie, Julie Tahraoui-Bories, Claire Legay, and Cécile Martinat. 2020. "Building Neuromuscular Junctions *in Vitro*." *Development* 147 (22): dev193920. <https://doi.org/10.1242/dev.193920>.
- Benoit-Pilven, Clara, Camille Marchet, Emilie Chautard, Leandro Lima, Marie Pierre Lambert, Gustavo Sacomoto, Amandine Rey, et al. 2018. "Complementarity of Assembly-First and Mapping-First Approaches for Alternative Splicing Annotation and Differential Analysis from RNAseq Data." *Scientific Reports* 8 (1): 4307. <https://doi.org/10.1038/s41598-018-21770-7>.
- Bombelli, Francesco, Ludovico Lispi, Sandro Costanzi Porrini, Manlio Giacanelli, Chiara Terracciano, Roberto Massa, and Antonio Petrucci. 2016. "Neuromuscular Transmission Abnormalities in Myotonic Dystrophy Type 1: A Neurophysiological Study." *Clinical Neurology and Neurosurgery* 150 (November): 84–88. <https://doi.org/10.1016/j.clineuro.2016.08.020>.
- Brook, J.David, Mila E. McCurrach, Helen G. Harley, Alan J. Buckler, Deanna Church, Hiroyuki Aburatani, Kent Hunter, et al. 1992. "Molecular Basis of Myotonic Dystrophy: Expansion of a Trinucleotide (CTG) Repeat at the 3' End of a Transcript Encoding a Protein Kinase Family Member." *Cell* 68 (4). [https://doi.org/10.1016/0092-8674\(92\)90154-5](https://doi.org/10.1016/0092-8674(92)90154-5).
- Charizanis, Konstantinos, Kuang-Yung Lee, Ranjan Batra, Marianne Goodwin, Chaolin Zhang, Yuan Yuan, Lily Shiue, et al. 2012. "Article Muscleblind-like 2-Mediated Alternative Splicing in the Developing Brain and Dysregulation in Myotonic Dystrophy." *Neuron* 75: 437–50. <https://doi.org/10.1016/j.neuron.2012.05.029>.
- Charlet-B., Nicolas, Rajesh S. Savkur, Gopal Singh, Anne V. Philips, Elizabeth A. Grice, and Thomas A. Cooper. 2002. "Loss of the Muscle-Specific Chloride Channel in Type 1 Myotonic Dystrophy Due to Misregulated Alternative Splicing." *Molecular Cell* 10 (1). [https://doi.org/10.1016/S1097-2765\(02\)00572-5](https://doi.org/10.1016/S1097-2765(02)00572-5).
- Coërs, Christian, Nicole Telerman Toppet, and Jean Marie Gerard. 1973. "Terminal Innervation Ratio in Neuromuscular Disease: II. Disorders of Lower Motor Neuron, Peripheral Nerve, and Muscle." *Archives of Neurology* 29 (4): 215–22. <https://doi.org/10.1001/archneur.1973.00490280027003>.
- Driesche, Sarah J. Van, and Kelsey C. Martin. 2018. "New Frontiers in RNA Transport and Local Translation in Neurons." *Developmental Neurobiology* 78 (3): 331–39. <https://doi.org/10.1002/dneu.22574>.
- Du, Hongqing, Melissa S Cline, Robert J Osborne, Daniel L Tuttle, Tyson A Clark, John Paul Donohue, Megan P Hall, et al. 2010. "Aberrant Alternative Splicing and Extracellular Matrix Gene Expression in Mouse Models of Myotonic Dystrophy." *Nature Structural & Molecular Biology* 17 (2). <https://doi.org/10.1038/nsmb.1720>.
- Fardaei, M. 2002. "Three Proteins, MBNL, MBLL and MBXL, Co-Localize *In Vivo* with Nuclear Foci of Expanded-Repeat Transcripts in DM1 and DM2 Cells." *Human Molecular Genetics* 11 (7). <https://doi.org/10.1093/hmg/11.7.805>.
- Fardeau M, Tome FMS. 1980. "Light and Electron Microscopic Study of Motor Endplates in the Adult and Neo-Natal Forms of Dystrophia Myotonica." In , edited by Taxi J, 287–98.
- Goodwin, Marianne, Apoorva Mohan, Ranjan Batra, Kuang Yung Lee, Konstantinos Charizanis, Francisco José Fernández Gómez, Sabiha Eddarkaoui, et al. 2015. "MBNL Sequestration by Toxic RNAs and RNA Misprocessing in the Myotonic Dystrophy Brain." *Cell Reports* 12 (7): 1159–68. <https://doi.org/10.1016/j.celrep.2015.07.029>.
- Gourdon, Genevieve, and Giovanni Meola. 2017. "Myotonic Dystrophies: State of the Art of New Therapeutic Developments for the CNS." *Frontiers in Cellular Neuroscience* 11 (April). <https://doi.org/10.3389/fncel.2017.00101>.
- Härönen, Heli, Zarin Zainul, Nikolay Naumenko, Raija Sormunen, Ilkka Miinalainen, Anastasia Shakirzyanova, Sabrina Santoleri, et al. 2019. "Correct Expression and Localization of Collagen XIII Are Crucial for the Normal Formation and Function of the Neuromuscular System." *European Journal*

- of Neuroscience* 49 (11): 1491–1511. <https://doi.org/10.1111/ejn.14346>.
- Hernández-Hernández, Oscar, Céline Guiraud-Dogan, Géraldine Sicot, Aline Huguet, Sabrina Luilier, Esther Steidl, Stefanie Saenger, et al. 2013. “Myotonic Dystrophy CTG Expansion Affects Synaptic Vesicle Proteins, Neurotransmission and Mouse Behaviour.” *Brain* 136 (3): 957–70. <https://doi.org/10.1093/brain/aws367>.
- Ho, Thai H., Nicolas Charlet-B, Michael G. Poulos, Gopal Singh, Maurice S. Swanson, and Thomas A. Cooper. 2004. “Muscleblind Proteins Regulate Alternative Splicing.” *EMBO Journal* 23 (15): 3103–12. <https://doi.org/10.1038/sj.emboj.7600300>.
- Kanadia, R. N., J. Shin, Y. Yuan, S. G. Beattie, T. M. Wheeler, C. A. Thornton, and M. S. Swanson. 2006. “Reversal of RNA Missplicing and Myotonia after Muscleblind Overexpression in a Mouse Poly(CUG) Model for Myotonic Dystrophy.” *Proceedings of the National Academy of Sciences* 103 (31): 11748–53. <https://doi.org/10.1073/pnas.0604970103>.
- Kim, Daehwan, Geo Pertea, Cole Trapnell, Harold Pimentel, Ryan Kelley, and Steven L. Salzberg. 2013. “TopHat2: Accurate Alignment of Transcriptomes in the Presence of Insertions, Deletions and Gene Fusions.” *Genome Biology* 14 (4): R36. <https://doi.org/10.1186/gb-2013-14-4-r36>.
- Krishnan, Arun V., and Matthew C. Kiernan. 2006. “Axonal Function and Activity-Dependent Excitability Changes in Myotonic Dystrophy.” *Muscle and Nerve* 33 (5): 627–36. <https://doi.org/10.1002/mus.20516>.
- Kurup, Naina, Dong Yan, Alexandr Goncharov, and Yishi Jin. 2015. “Dynamic Microtubules Drive Circuit Rewiring in the Absence of Neurite Remodeling.” *Current Biology* 25 (12): 1594–1605. <https://doi.org/10.1016/j.cub.2015.04.061>.
- Lainé, Jeanne, Gunnar Skoglund, Emmanuel Fournier, and Nacira Tabti. 2018. “Development of the Excitation-Contraction Coupling Machinery and Its Relation to Myofibrillogenesis in Human iPSC-Derived Skeletal Myocytes.” *Skeletal Muscle* 8 (1): 1. <https://doi.org/10.1186/s13395-017-0147-5>.
- Lee, Kuang-Yung, Moyi Li, Mini Manchanda, Ranjan Batra, Konstantinos Charizanis, Apoorva Mohan, Sonisha A. Warren, et al. 2013. “Compound Loss of Muscleblind-like Function in Myotonic Dystrophy.” *EMBO Molecular Medicine* 5 (12): 1887–1900. <https://doi.org/10.1002/emmm.201303275>.
- Lee, Kuang Yung, Ho Ching Chang, Carol Seah, and Li Jen Lee. 2019. “Deprivation of Muscleblind-Like Proteins Causes Deficits in Cortical Neuron Distribution and Morphological Changes in Dendritic Spines and Postsynaptic Densities.” *Frontiers in Neuroanatomy* 13 (July): 75. <https://doi.org/10.3389/fnana.2019.00075>.
- Lee, Kuang Yung, Moyi Li, Mini Manchanda, Ranjan Batra, Konstantinos Charizanis, Apoorva Mohan, Sonisha A. Warren, et al. 2013. “Compound Loss of Muscleblind-like Function in Myotonic Dystrophy.” *EMBO Molecular Medicine* 5 (12): 1887–1900. <https://doi.org/10.1002/emmm.201303275>.
- Lieberam, Ivo, Dritan Agalliu, Takashi Nagasawa, Johan Ericson, and Thomas M. Jessell. 2005. “A Cxcl12-Cxcr4 Chemokine Signaling Pathway Defines the Initial Trajectory of Mammalian Motor Axons.” *Neuron* 47 (5): 667–79. <https://doi.org/10.1016/j.neuron.2005.08.011>.
- Macdermot, Violet. 1961. “THE HISTOLOGY OF THE NEUROMUSCULAR JUNCTION IN DYSTROPHIA MYOTONICA.” <http://brain.oxfordjournals.org/>.
- Mahadevan, M, C Tsilfidis, L Sabourin, G Shutler, C Amemiya, G Jansen, C Neville, et al. 1992. “Myotonic Dystrophy Mutation: An Unstable CTG Repeat in the 3’ Untranslated Region of the Gene.” *Science* 255 (5049). <https://doi.org/10.1126/science.1546325>.
- Mallinjou, Pierre, Jean Philippe Villemin, Hussein Mortada, Micaela Polay Espinoza, Francois Olivier Desmet, Samaan Samaan, Emilie Chautard, Léon Charles Tranchevent, and Didier Auboeuf. 2014. “Endothelial, Epithelial, and Fibroblast Cells Exhibit Specific Splicing Programs Independently of Their Tissue of Origin.” *Genome Research* 24 (3): 511–21. <https://doi.org/10.1101/gr.162933.113>.
- Mankodi, Ami, Masanori P. Takahashi, Hong Jiang, Carol L. Beck, William J. Bowers, Richard T. Moxley, Stephen C. Cannon, and Charles A. Thornton. 2002. “Expanded CUG Repeats Trigger Aberrant Splicing of ClC-1 Chloride Channel Pre-mRNA and Hyperexcitability of Skeletal Muscle in Myotonic Dystrophy.” *Molecular Cell* 10 (1). [https://doi.org/10.1016/S1097-2765\(02\)00563-4](https://doi.org/10.1016/S1097-2765(02)00563-4).
- Marteyn, Antoine, Yves Maury, Morgane M. Gauthier, Camille Lecuyer, Remi Vernet, Jérôme A. Denis, Geneviève Pietu, Marc Peschanski, and Cécile Martinat. 2011. “Mutant Human Embryonic Stem Cells Reveal Neurite and Synapse Formation Defects in Type 1 Myotonic Dystrophy.” *Cell Stem Cell* 8 (4): 434–44. <https://doi.org/10.1016/j.stem.2011.02.004>.
- Martin, Marcel. 2011. “Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads.” *EMBNET Journal* 17 (1): 10. <https://doi.org/10.14806/ej.17.1.200>.
- Masuda, Akio, Henriette Skovgaard Andersen, Thomas Koed Doktor, Takaaki Okamoto, Mikako Ito, Brage Storstein Andresen, and Kinji Ohno. 2012. “CUGBP1 and MBNL1 Preferentially Bind to 3’ UTRs and Facilitate mRNA Decay.” *Scientific Reports* 2 (1): 1–10. <https://doi.org/10.1038/srep00209>.

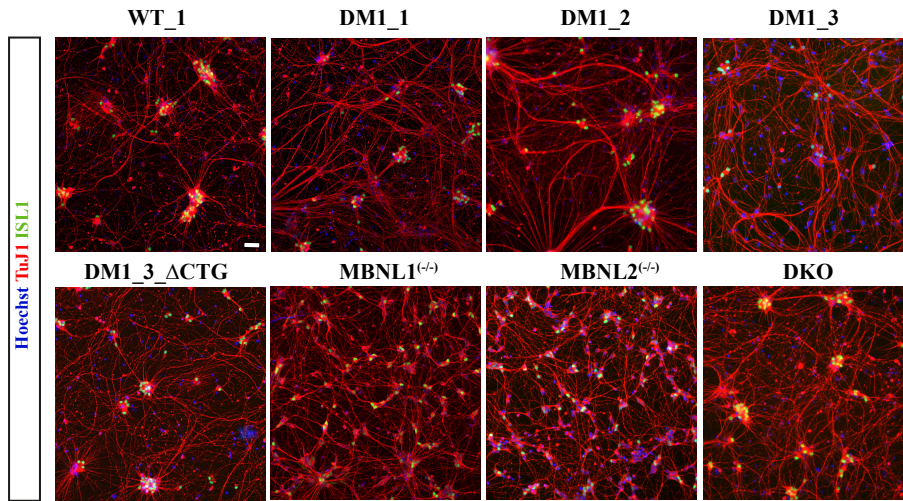
- Maury, Yves, Julien Côme, Rebecca A. Piskorowski, Nouzha Salah-Mohellibi, Vivien Chevaleyre, Marc Peschanski, Cécile Martinat, and Stéphane Nedelec. 2015. "Combinatorial Analysis of Developmental Cues Efficiently Converts Human Pluripotent Stem Cells into Multiple Neuronal Subtypes." *Nature Biotechnology* 33 (1): 89–96. <https://doi.org/10.1038/nbt.3049>.
- Maury, Yves, Pauline Poydenot, Benjamin Brinon, Lea Lesueur, Jacqueline Gide, Sylvain Roquevière, Julien Côme, et al. 2019. "Pluripotent Stem Cell-Based Drug Screening Reveals Cardiac Glycosides as Modulators of Myotonic Dystrophy Type 1." *IScience* 11 (January): 258–71. <https://doi.org/10.1016/j.isci.2018.12.019>.
- Miller, Jill W., Carl R. Urbinati, Patana Teng-Ummuay, Myrna G. Stenberg, Barry J. Byrne, Charles A. Thornton, and Maurice S. Swanson. 2000. "Recruitment of Human Muscleblind Proteins to (CUG)(n) Expansions Associated with Myotonic Dystrophy." *EMBO Journal* 19 (17): 4439–48. <https://doi.org/10.1093/emboj/19.17.4439>.
- Nakamori, Masayuki, Krzysztof Sobczak, Araya Puwanant, Steve Welle, Katy Eichinger, Shree Pandya, Jeanne Dekdebrun, et al. 2013. "Splicing Biomarkers of Disease Severity in Myotonic Dystrophy." *Annals of Neurology* 74 (6): 862–72. <https://doi.org/10.1002/ana.23992>.
- Nakata, Katsunori, Benjamin Abrams, Brock Grill, Alexandr Goncharov, Xun Huang, Andrew D. Chisholm, and Yishi Jin. 2005. "Regulation of a DLK-1 and P38 MAP Kinase Pathway by the Ubiquitin Ligase RPM-1 Is Required for Presynaptic Development." *Cell* 120 (3): 407–20. <https://doi.org/10.1016/j.cell.2004.12.017>.
- Ohtsuka, Toshihisa. 2013. "CAST: Functional Scaffold for the Integrity of the Presynaptic Active Zone." *Neuroscience Research* 76 (1–2): 10–15. <https://doi.org/10.1016/j.neures.2013.03.003>.
- Osborne, Robert J., Xiaoyan Lin, Stephen Welle, Krzysztof Sobczak, Jason R. O'Rourke, Maurice S. Swanson, and Charles A. Thornton. 2009. "Transcriptional and Post-Transcriptional Impact of Toxic RNA in Myotonic Dystrophy." *Human Molecular Genetics* 18 (8): 1471–81. <https://doi.org/10.1093/hmg/ddp058>.
- Panaite, Petrica Adrian, Thierry Kuntzer, Geneviève Gourdon, and Ibtissam Barakat-Walter. 2013. "Respiratory Failure in a Mouse Model of Myotonic Dystrophy Does Not Correlate with the CTG Repeat Length." *Respiratory Physiology and Neurobiology* 189 (1): 22–26. <https://doi.org/10.1016/j.resp.2013.06.014>.
- Schmieder, Robert, and Robert Edwards. 2011. "Quality Control and Preprocessing of Metagenomic Datasets." *Bioinformatics* 27 (6): 863–64. <https://doi.org/10.1093/bioinformatics/btr026>.
- Scrudato, Mirella Lo, Karine Poulard, Célia Sourd, Stéphanie Tomé, Arnaud F. Klein, Guillaume Corre, A. Huguet, D. Furling, Geneviève Gourdon, and Ana Buj-Bello. 2019. "Genome Editing of Expanded CTG Repeats within the Human DMPK Gene Reduces Nuclear RNA Foci in the Muscle of DM1 Mice." *Molecular Therapy* 27 (8): 1372–88. <https://doi.org/10.1016/j.ymthe.2019.05.021>.
- Shigeoka, Toshiaki, Hosung Jung, Jane Jung, Benita Turner-Bridger, Jiyeon Ohk, Julie Qiaojin Lin, Paul S Amieux, and Christine E Holt. 2016. "Dynamic Axonal Translation in Developing and Mature Visual Circuits." *Cell* 166 (1): 181–92. <https://doi.org/10.1016/j.cell.2016.05.029>.
- Sicot, Géraldine, Laurent Servais, Diana M. Dinca, Axelle Leroy, Cynthia Prigogine, Fadia Medja, Sandra O. Braz, et al. 2017. "Downregulation of the Glial GLT1 Glutamate Transporter and Purkinje Cell Dysfunction in a Mouse Model of Myotonic Dystrophy." *Cell Reports* 19 (13): 2718–29. <https://doi.org/10.1016/j.celrep.2017.06.006>.
- Spilker, Kerri A, George J Wang, Madina S Tugizova, and Kang Shen. 2012. "Caenorhabditis Elegans Muscleblind Homolog Mbl-1 Functions in Neurons to Regulate Synapse Formation." *Neural Development*. Vol. 7. <https://doi.org/10.1186/1749-8104-7-7>.
- Ven, Peter F.M. van Der, Gert Jansen, Toln H.M.S.M. van Kuppevelt, M. Benjamin Perryman, Mark Lupa, Patrick W. Dunne, Henk J.ter Laak, et al. 1993. "Myotonic Dystrophy Kinase Is a Component of Neuromuscular Junctions." *Human Molecular Genetics* 2 (11): 1889–94. <https://doi.org/10.1093/hmg/2.11.1889>.
- Wang, Eric T., Amanda J. Ward, Jennifer M. Cherone, Jimena Giudice, Thomas T. Wang, Daniel J. Treacy, Nicole J. Lambert, et al. 2015. "Antagonistic Regulation of mRNA Expression and Splicing by CELF and MBNL Proteins." *Genome Research* 25 (6): 858–71. <https://doi.org/10.1101/gr.184390.114>.
- Wang, Eric T, Neal A L Cody, Sonali Jog, Michela Biancolella, Thomas T Wang, Daniel J Treacy, Shujun Luo, et al. 2012. "Transcriptome-Wide Regulation of Pre-mRNA Splicing and mRNA Localization by Muscleblind Proteins." *CELL* 150: 710–24. <https://doi.org/10.1016/j.cell.2012.06.041>.
- Wang, Pei Ying, Kuei Ting Chang, Yu Mei Lin, Ting Yu Kuo, and Guey Shin Wang. 2018. "Ubiquitination of MBNL1 Is Required for Its Cytoplasmic Localization and Function in Promoting Neurite Outgrowth." *Cell Reports* 22 (9): 2294–2306. <https://doi.org/10.1016/j.celrep.2018.02.025>.

- Wheeler, T. M., M. C. Krym, and C. A. Thornton. 2007a. "Ribonuclear Foci at the Neuromuscular Junction in Myotonic Dystrophy Type 1." *Neuromuscular Disorders* 17 (3): 242–47. <https://doi.org/10.1016/j.nmd.2006.12.015>.
- Whiting, Elisabeth J., James D. Waring, Katsuyuki Tamai, Martin J. Somerville, Maxwell Hincke, William A. Staines, Joh-E Ikeda, and Robert G. Korneluk. 1995. "Characterization of Myotonic Dystrophy Kinase (DMK) Protein in Human and Rodent Muscle and Central Nervous Tissue." *Human Molecular Genetics* 4 (6). <https://doi.org/10.1093/hmg/4.6.1063>.
- Young, Joanne, Yoran Margaron, Mathieu Fernandes, Eve Duchemin-Pelletier, Joris Michaud, Mélanie Flaender, Oana Lorintiu, Sébastien Degot, and Pauline Poydenot. 2018. "MyoScreen, a High-Throughput Phenotypic Screening Platform Enabling Muscle Drug Discovery." *SLAS Discovery* 23 (8): 790–806. <https://doi.org/10.1177/2472555218761102>.
- Zhang, Bin, Chuan Liang, Ryan Bates, Yiming Yin, Wen-Cheng Xiong, and Lin Mei. 2012. "Wnt Proteins Regulate Acetylcholine Receptor Clustering in Muscle Cells." *Molecular Brain* 5 (1): 7. <https://doi.org/10.1186/1756-6606-5-7>.

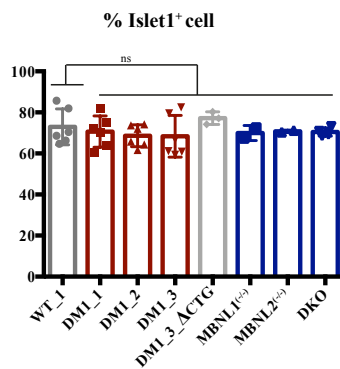
FIGURES

Figure 1

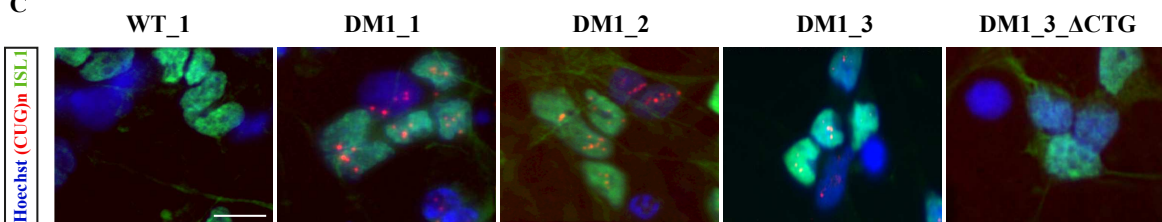
A



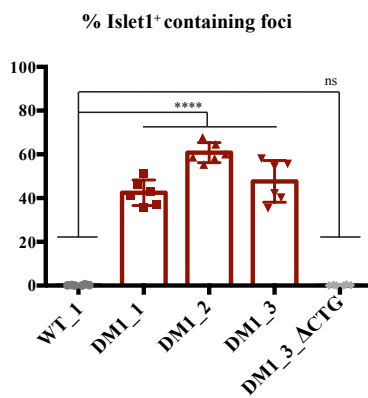
B



C



D



E

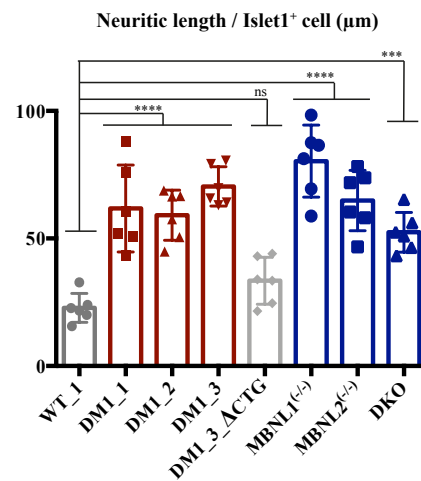


Figure 1: DM1 hiPSC-derived MN reproduced the mean features associated to DM1.

(A) Representative images of labeling for motoneuronal markers in hiPSC-derived MN after 30 days of differentiation. MN were identified by ISL1 staining (Green) and TuJI (Red) immunostaining. Nuclei were revealed by Hoechst staining. Scale bar 10 μ m. (B) Quantification of the percentage of hiPSC-derived MN containing foci. Data were analyzed with an ordinary One-Way anova, Tukey's multiple comparisons test compared to WT (****: $P \leq 0,0001$, ns: not significant). The mean \pm SD values from at least 5 independent experiments. (C) Representative images of mutant DMPK mRNA foci (red) detected by mRNA fluorescence in situ hybridization (FISH) combined with ISL1 immunolabeling (in green) in DM1, WT and DM1 CTG-deleted hiPSC-derived MN after 30 days of differentiation. Scale bar 10 μ m. (D) Quantification of the percentage of hiPSC-derived MN containing foci per ISL1⁺ve cells. Data were analyzed with an ordinary One-Way anova, Tukey's multiple comparisons test compared to WT (****: $P \leq 0,0001$, ns: not significant). The mean \pm SD values from at least 5 independent experiments. (E) Neuritic length per hiPSC-derived MN was determined after 30 days of differentiation by staining the motoneuronal cultures with TuJI and ISL1. Data were analyzed with an ordinary One-Way anova, Tukey's multiple comparisons test compared to WT (***: $P \leq 0,001$, ****: $P \leq 0,0001$, ns: not significant). The mean \pm SD values from at least 5 independent experiments.

Figure 2

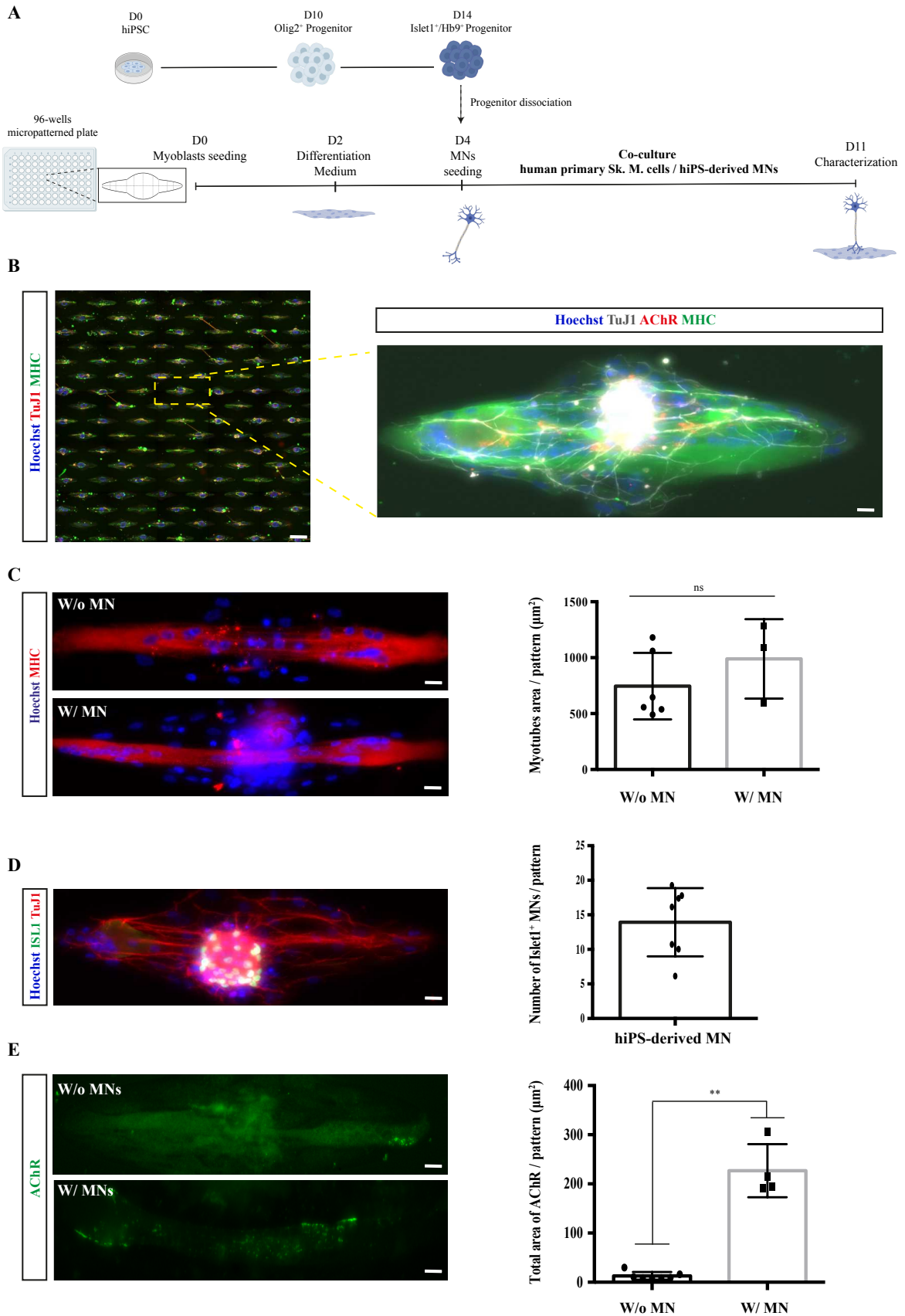


Figure 2: hiPSC-derived MN enhanced AChR clustering after 7 days of coculture.

(A) Schematic representation of the co-culture protocol. MN were differentiated from hiPSC for 14 days and plated on primary skeletal muscle cells differentiated for 2 days. Co-cultures were kept for a minimal of 7 days. (B) Representative image of immunostaining (Nuclei in blue, Myosin Heavy Chain MF20 in green and TuJ1 for neurites in red) on an entire well from a 96 well plate. Scale bar: 100µm. Each well contains about 200 micropatterns. The image on the right is a higher magnification of the insert to better visualize the staining for neurons by TuJ1 (in grey), acetylcholine receptors (AChR, in red) and human primary skeletal muscle cells stained for MF20 (in green). Nuclei were labeled with Hoechst (Blue). Scale Bar 10 µm. (C) Representative images of skeletal muscle cells immunolabeling using myosin heavy chain marker in the presence and absence of hiPSC-derived MN. On the right, the histogram indicated the mean of myotube area per pattern in the absence and the presence of hiPSC-derived MN. Data were analyzed with Student's t-test (ns: not significant). The mean ± SD values from at least 3 independent experiments in triplicate (at least 100 micropatterns). (D) Representative image of staining for MN by using ISL1 (Green), TuJ1 (Red) and Hoechst for nuclei (blue). Scale bar 10 µm. On the right, the histogram indicated the mean number of hiPSC-derived MN per pattern. Data were analyzed with Student's t-test (*: $p < 0,05$; **: $p < 0,005$; ***: $p < 0,0005$; ****: $p < 0,0001$, ns: no significant). The mean ± SD values from 7 independent experiments in triplicate (at least 100 micropatterns). (E) Representative images of immunolabelling for AChR (green). Quantification of the total area of AChR per pattern is indicated on the histogram. Scale Bar 10 µm. Data were analyzed with Student's t-test (**: $p < 0,005$). The mean ± SD values from at least 4 independent experiments in triplicate (at least 100 micropatterns).

Figure 3

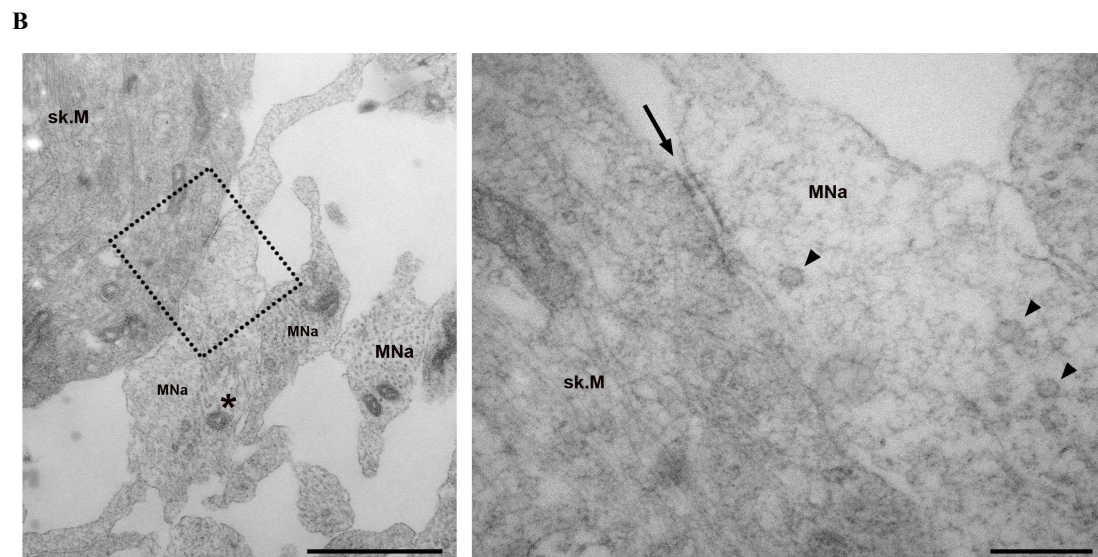
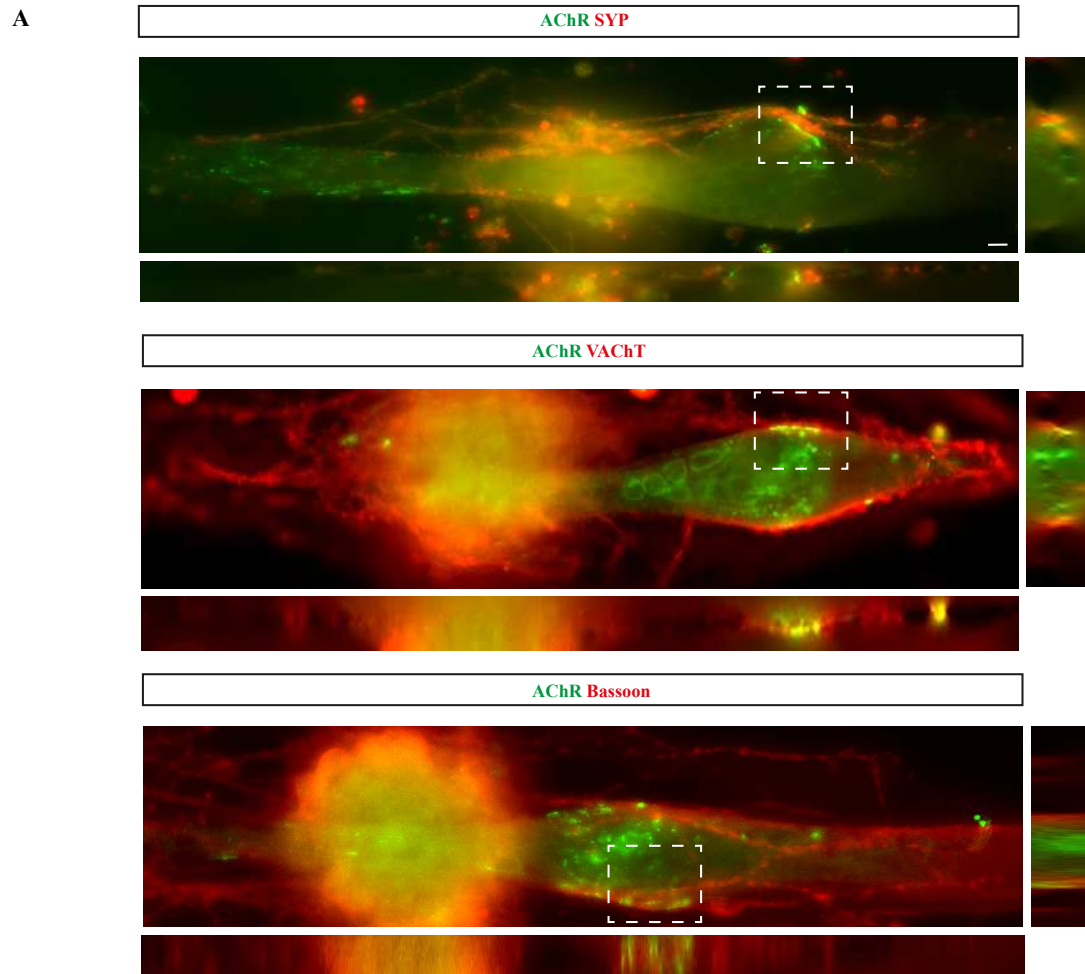


Figure 3: Immunofluorescence and ultrastructural study revealed contacts between pre- and post-synaptic compartments.

(A) Co-immunostaining of AChR and synaptic vesicles markers (from top to bottom: Synaptophysin (SYP), Vesicular acetylcholine transporter (VACHT) and Bassoon). Orthogonal views from different planes (x/z , y/z) of the white insert are represented. Scale bar: $10\mu\text{m}$. (B) Transmission electron micrographs of a contact between a hiPSC-derived MN axon (obliquely cut microtubules indicated by asterisk) and a skeletal muscle cell. Scale bar = $1\mu\text{m}$. The dotted square is enlarged in the right micrograph, where black arrowheads indicate synaptic vesicles in the pre-synaptic zone and the arrow points to plasmalemmal electron densities on both sides of the contact. Scale bar = 200nm .

Figure 4

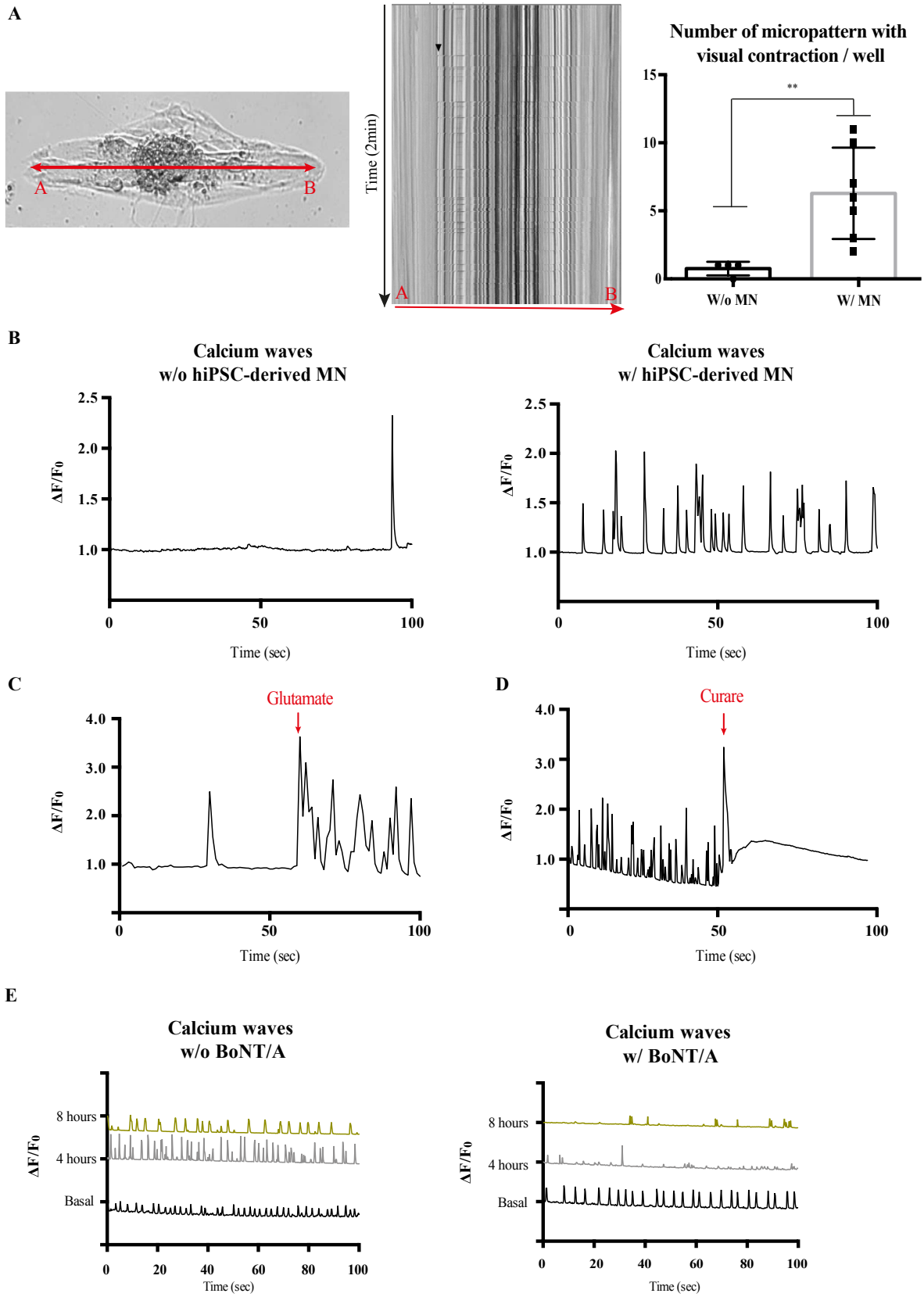


Figure 4: hiPSC-derived MN induced muscle calcium waves and contractions.

(A) Observation of muscle contractions by time lapse imaging during 2min in phase contrast microscopy using Spinning Disk (Zeiss®) with controlled atmosphere. The kymograph representation indicated the detection of pixel movements from A to B points, corresponding to visual muscle contractions. The quantification of the number of micropatterns presenting a visual contraction is indicated on the histogram on the right. Data were analyzed with Student's t-test (**: $p < 0,005$). The mean \pm SD values from at least 3 independent experiments in quadruplicate with two fields (40 micropatterns/field) per well. (B) Representative diagram of calcium waves recorded during 100s in the presence or absence of hiPSC-derived MN. $\Delta F/F_0$ representation after background correction. (C-D) Representative diagram of calcium waves recorded in skeletal muscle cells after (C) addition of a solution of 100 μ M L-glutamate (stimulation of NMDA receptors) and (D) inhibition with a solution of 100 μ M curare (competitive inhibition of AChR). $\Delta F/F_0$ representation after background correction. (E) Graphical representation of calcium waves recorded in basal condition and 4 and 8 hours after addition of a solution of 5nM Botulinum neurotoxin (BoNT/A). $\Delta F/F_0$ representation after background correction.

Figure 5

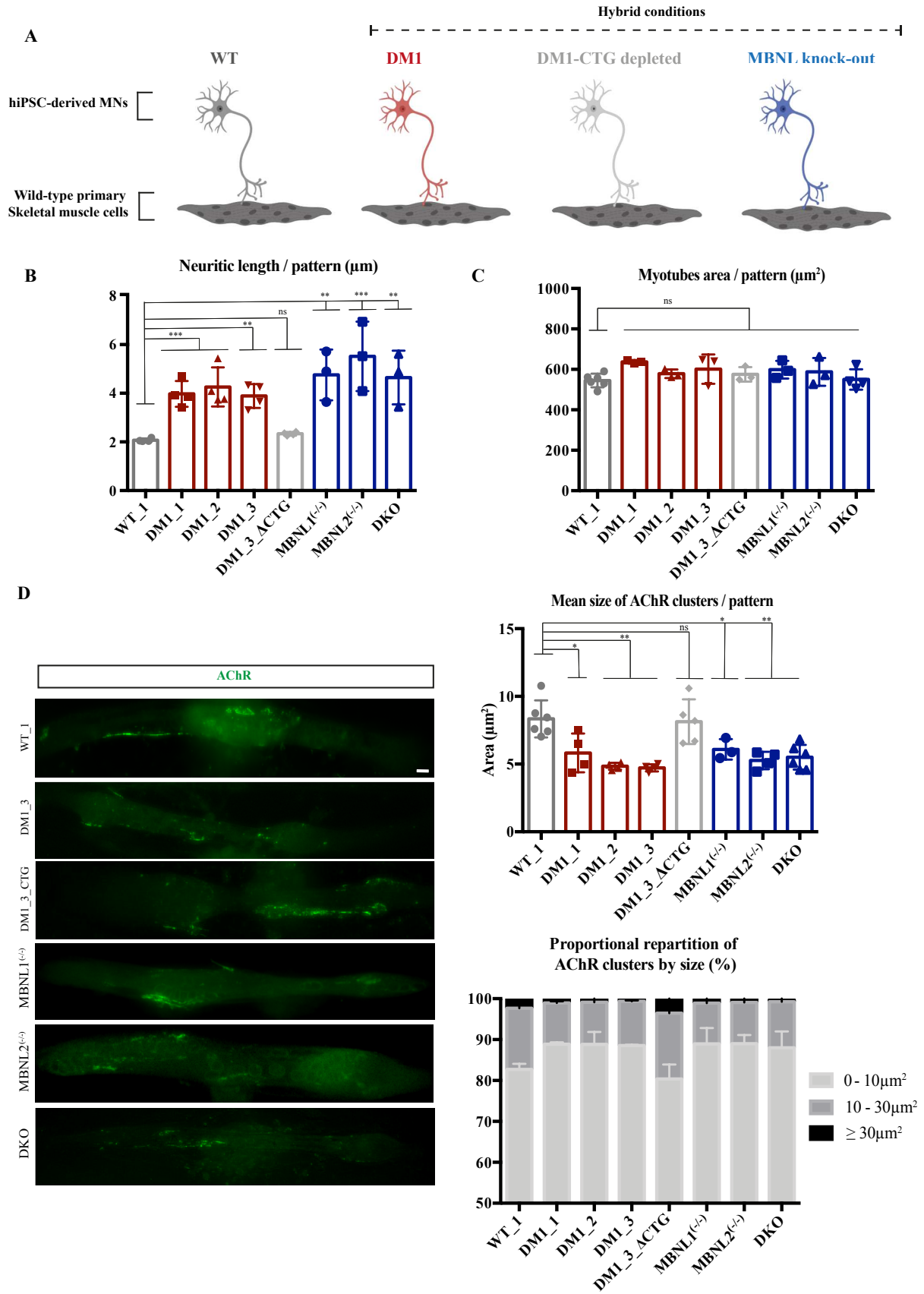
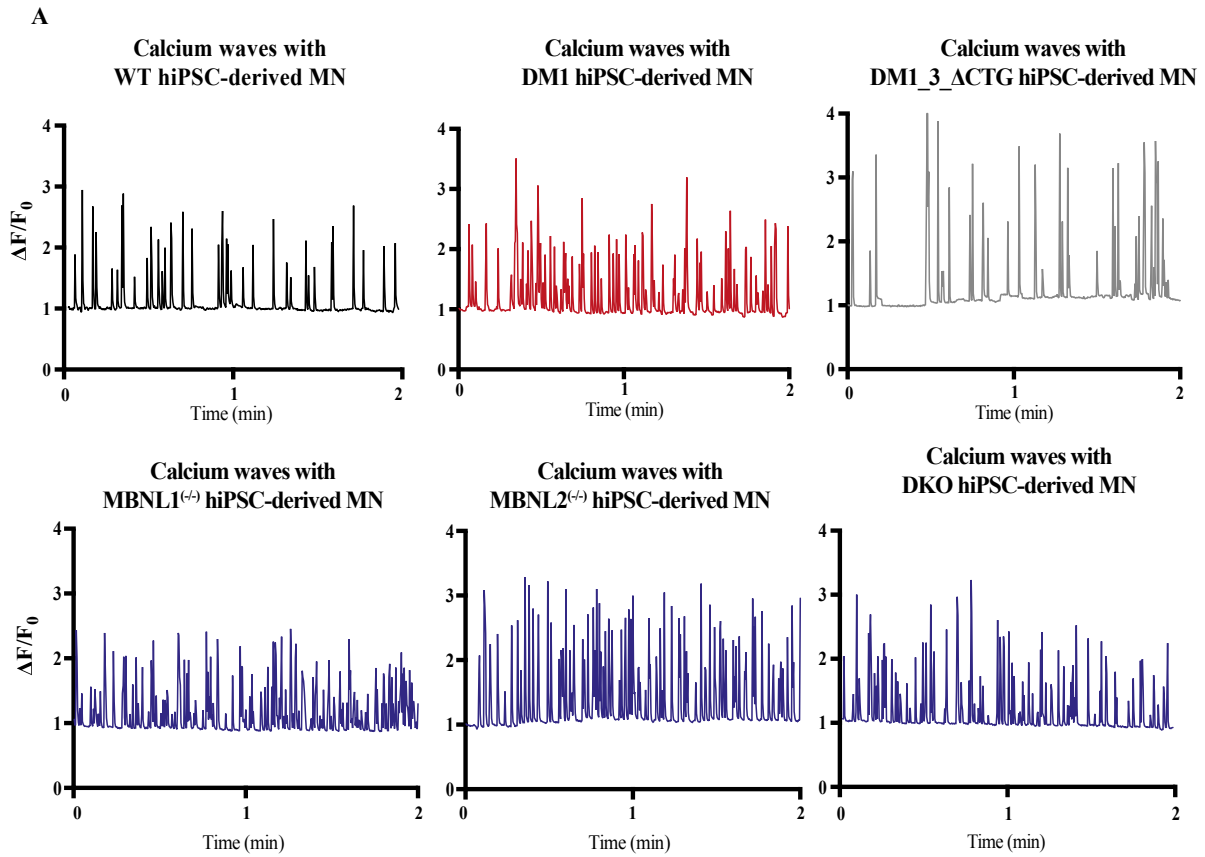


Figure 5: Synaptic defects observed at pre- and post- levels in DM1 and MBNL depleted hiPSC-derived MN.

(A) Schematic representation of hybrid co-culture systems developed in the study. (B) Quantification of the neuritic length (μm) per pattern after 11 days of co-culture. Data were analyzed with an ordinary One-Way anova, Tukey's multiple comparisons test compared to WT (**: $P \leq 0,01$, ***: $P \leq 0,001$, ****: $P \leq 0,0001$, ns: not significant). The mean \pm SD values from at least 3 independent experiments in triplicate (at least 150 micropatterns per experiment). (C) Quantification of myotube area per pattern (μm^2) after 11 days of co-culture. Data were analyzed with an ordinary One-Way anova, Tukey's multiple comparisons test compared to WT (ns: not significant). The mean \pm SD values from at least 3 independent experiments in triplicate (at least 150 micropatterns per experiment). (D) Representative images of AChR immunolabelling (green) in the co-cultures kept for 11 days with MN derived from the different hiPSC lines. Scale bar: $10\mu\text{m}$. On the top left, quantification of the mean size of AChR clusters per pattern. Data were analyzed with an ordinary One-Way anova, Tukey's multiple comparisons test compared to WT (*: $P \leq 0,05$, **: $P \leq 0,01$, ns: not significant). The mean \pm SD values from at least 3 independent experiments in triplicate (at least 50 micropatterns per experiment). On the bottom left, the proportional repartition (%) of AChR clusters by size ($0-10\mu\text{m}^2$, $10-30\mu\text{m}^2$, $30\mu\text{m}^2$ and more) in different hybrid co-culture systems.

Figure 6



B

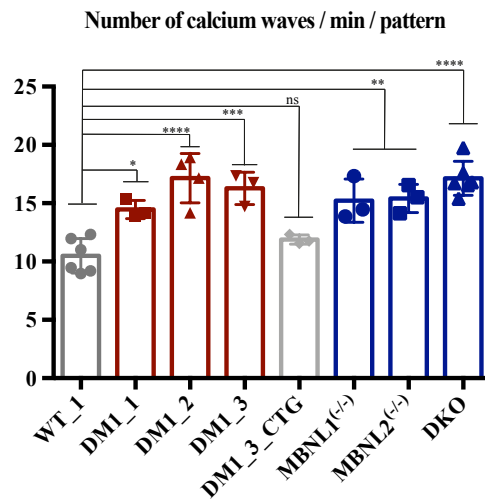


Figure 6: Communication defect between hiPSC-derived MN and skeletal muscle cells in DM1 and MBNL depletion conditions.

(A) Representative diagram of calcium waves (time=2min) recorded in human primary skeletal muscle cells in different hybrid conditions (WT in dark grey, DM1 in burgundy, DM1_3_ΔCTG in grey and MBNL knock-out conditions in blue). $\Delta F/F_0$ representation after background correction. (B) Quantification of the number of calcium waves recorded in human primary skeletal muscle cells per pattern and per min after 11 days of co-culture. Data were analyzed with an ordinary One-Way anova, Tukey's multiple comparisons test compared to WT (*: $P \leq 0,05$, **: $P \leq 0,01$, ***: $P \leq 0,001$, ****: $P \leq 0,0001$, ns: not significant). The mean \pm SD values from at least 3 independent experiments in triplicate (at least 100 micropatterns per experiment).

Figure 7

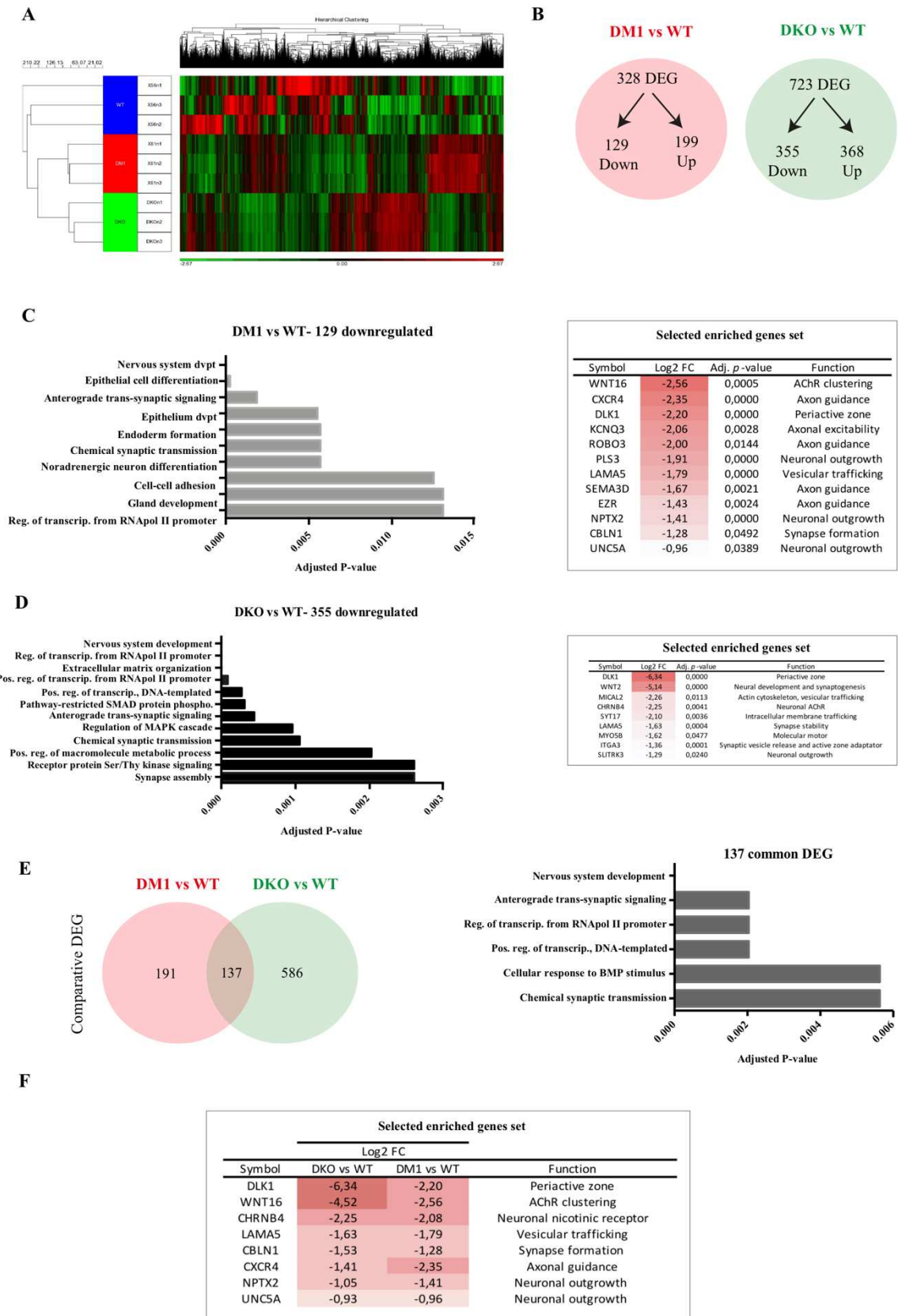


Figure 7: DM1 mutation and MBNL depletion affect the expression of gene sets related to synaptic function.

(A) The hierarchical clustering of differentially expressed genes (DEG) detected in DM1 and DKO hiPSC-derived MN compared to WT hiPSC-derived MN. Gene expression is represented by a color code ranging from green for under-expressed genes to red for over-expressed genes. **(B)** Schematic representation of the number of DEG in DM1 (red) and DKO (green) hiPSC-derived MN. **(C)** Genes set enrichment analysis using EnrichR on the 129 downregulated genes in DM1 hiPSC-derived MN when compared to WT hiPSC-derived MN (EnrichR analysis using GO Biological process 2018 database). The table on the right provides a list of specific genes differentially expressed in DM1 hiPSC-derived MN. **(D)** Genes set enrichment analysis using EnrichR on the 355 downregulated genes in DKO hiPSC-derived MN when compared to WT hiPSC-derived MN (EnrichR analysis using GO Biological process 2018 database). The table on the right provides a list of specific genes differentially expressed in DKO hiPSC-derived MN. **(E)** Venn diagram representation of the number of DEG deregulated between both different conditions analyzed. On the right, genes set enrichment analysis using EnrichR on the 137 common deregulated genes between DM1 versus WT hiPSC-derived MN and DKO versus WT hiPSC-derived MN conditions (EnrichR analysis using GO Biological process 2018 database). **(F)** List of specific deregulated genes in the panel of 137 genes commonly identified between DM1 versus WT hiPSC-derived MN and DKO versus WT hiPSC-derived MN conditions.

Figure 8

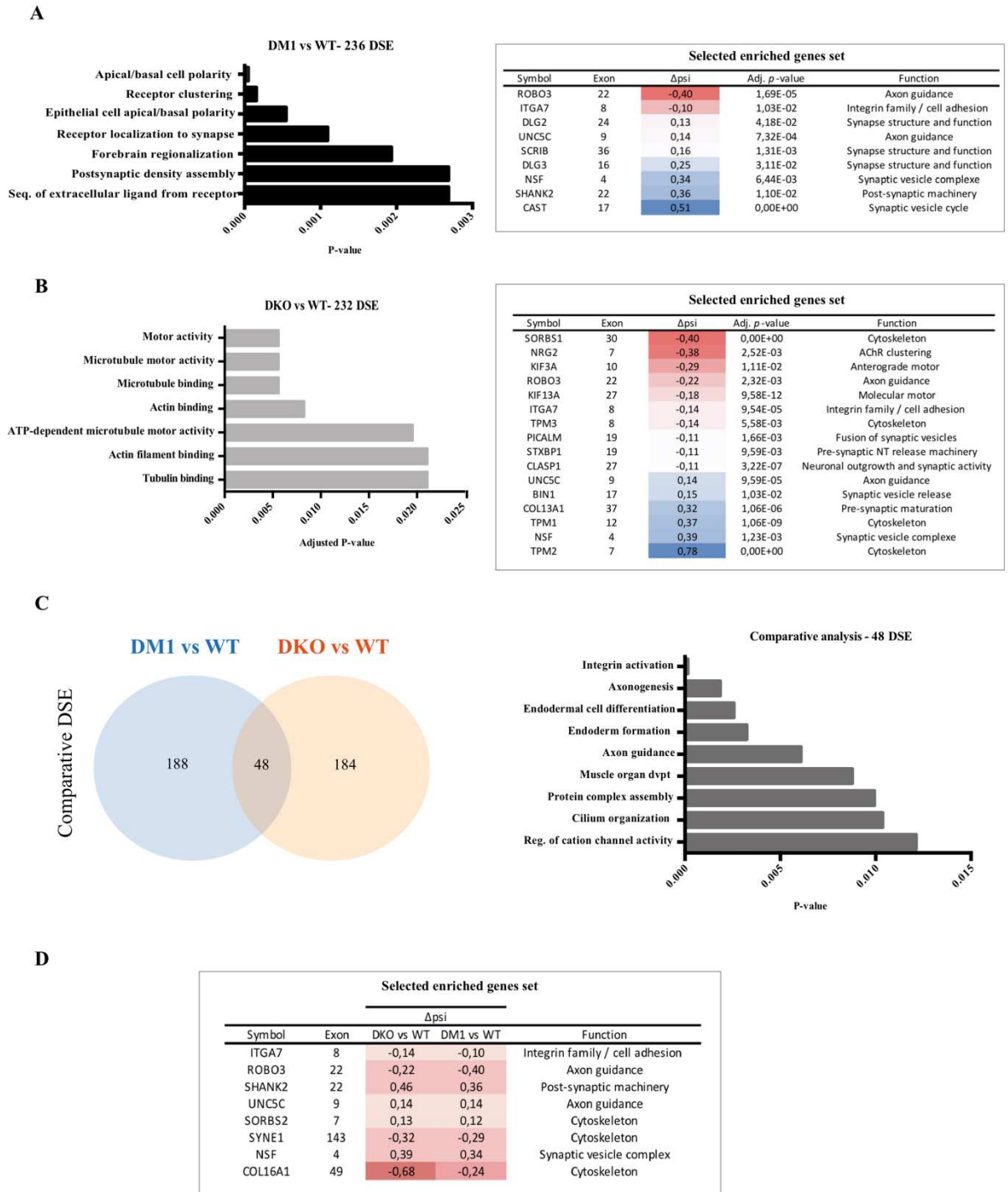


Figure 8: DM1 mutation and loss of MBNL1 and MBNL2 proteins lead to a large number of alternative splicing defects related to molecular motor activity and synaptic function.

(A) Genes set enrichment analysis using EnrichR on the 236 differentially spliced exons (DSE) in DM1 hiPSC-derived MN when compared to WT hiPSC-derived MN (EnrichR analysis using GO Biological process 2018 database). The table on the right provides a list of specific genes differentially spliced in DM1 hiPSC-derived MN. **(B)** Genes set enrichment analysis using EnrichR on the 232 differentially spliced exons (DSE) in DKO hiPSC-derived MN when compared to WT hiPSC-derived MN (EnrichR analysis using GO Molecular Function 2018 database). The table on the right provides a list of specific genes differentially spliced in DKO hiPSC-derived MN. **(C)** Venn diagram representation of the number of DSE deregulated between both different conditions analyzed. On the right, genes set enrichment analysis using EnrichR on the 48 common genes between DM1 versus WT hiPSC-derived MN and DKO versus WT hiPSC-derived MN conditions (EnrichR analysis using GO Biological process 2018 database). **(D)** List of specific mis-spliced genes in the panel of 48 genes commonly identified between DM1 versus WT hiPSC-derived MN and DKO versus WT hiPSC-derived MN conditions.

SUPPLEMENTARY FIGURES

Supplementary Figure 1

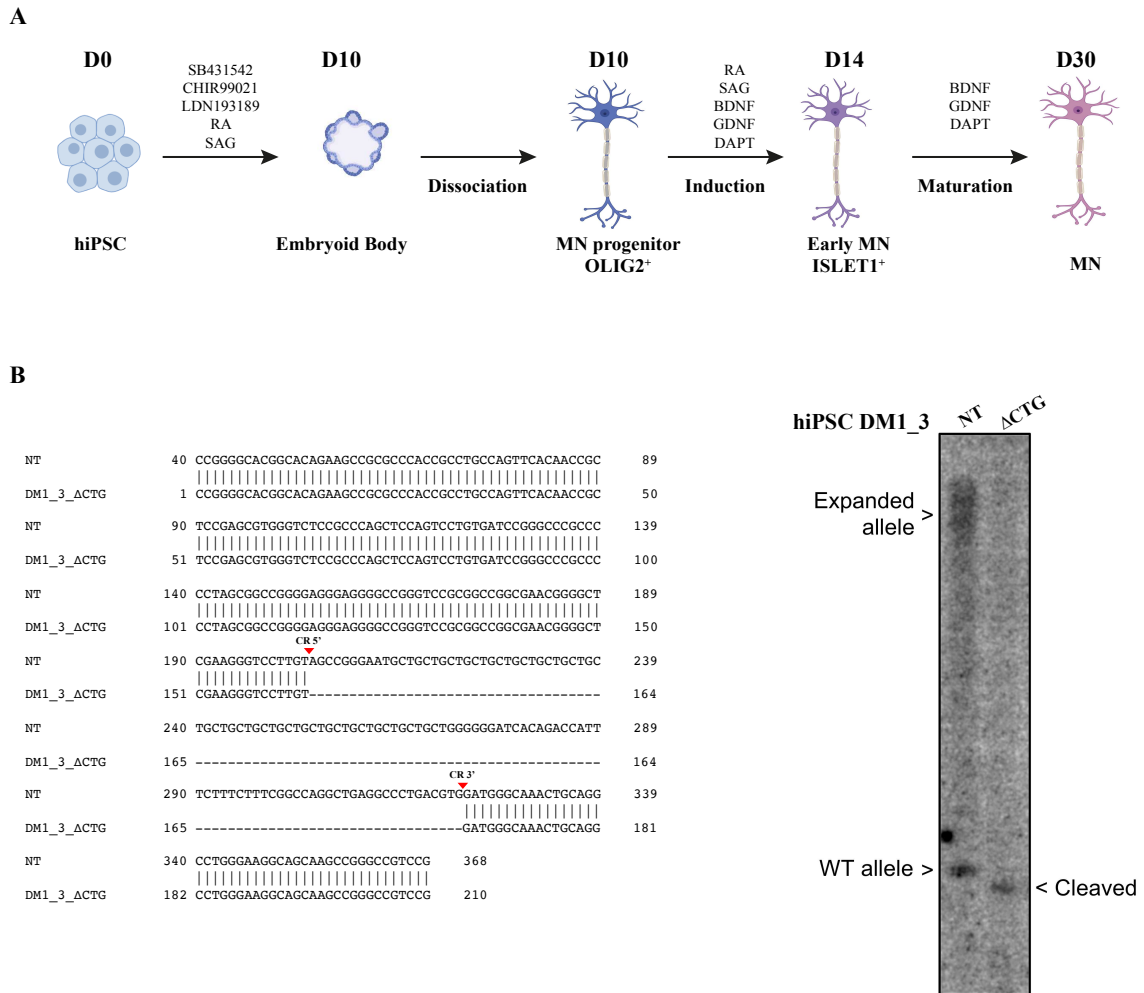


Figure S1: *In vitro* generation of human motoneurons.

(A) Schematic representation of the differentiation protocol from hiPSC to motoneurons. **(B)** Generation of the DM1-CTG depleted hiPSC cell line by CRISPR/Cas9 technology. At the left, sequencing analysis of not treated DM1_3 versus treated DM1_3 (DM1_3_ΔCTG), the guides are mentioned with red arrows. At the right, the southern blot analysis of DM1_3 and DM1_3_ΔCTG showing the complete deletion of CTG expansion.

Supplementary Figure 2

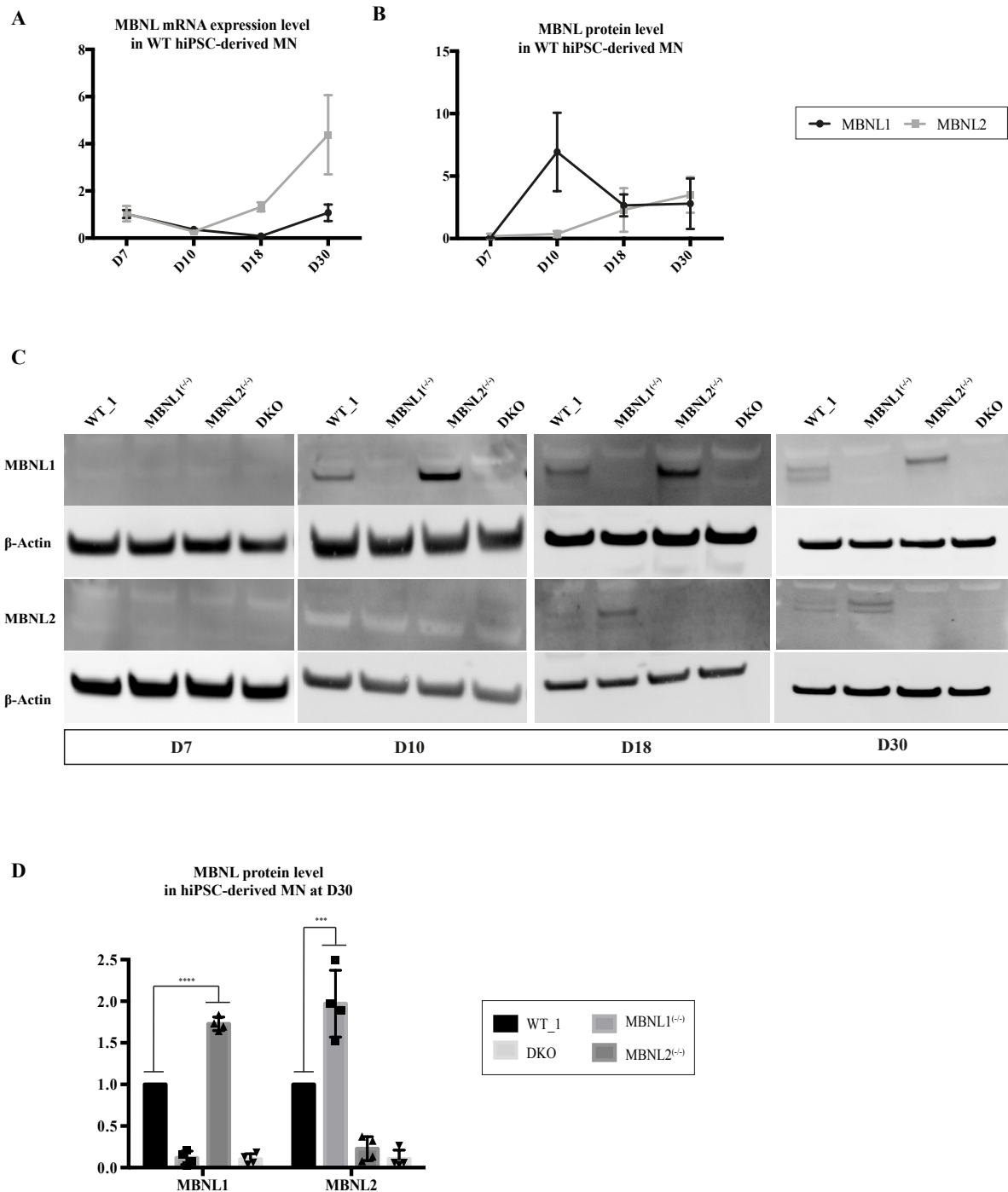


Figure S2: Kinetic of expression of MBNL proteins during differentiation of hiPSC into MN

(A-B) Kinetic expression of MBNL1 and MBNL2 (A) mRNA and (B) protein expression level in WT hiPSC-derived MN at different time points (day 7, 10, 18 and 30) of the differentiation process into MN. Results were normalized at D7 of differentiation. Data were obtained from 3 independent experiments. (C) Representative examples of western blot analysis for MBNL1 and MBNL2 at day 7, 10, 18 and 30 of hiPSC-derived MN differentiation. β -actin was used as housekeeping genes. (D) The graph represents the quantification of MBNL protein expression in MBNL1^{-/-}, MBNL2^{-/-} and DKO hiPSC-derived MN at D30 of differentiation. Results were normalized on the value of WT hiPSC-derived MN. Data were analyzed with an ordinary One-Way anova, Tukey's multiple comparisons test compared to WT, (***: $P \leq 0,001$, ****: $P \leq 0,0001$, ns: not significant). The mean \pm SD values from at least 3 independent experiments.

Supplementary Figure 3

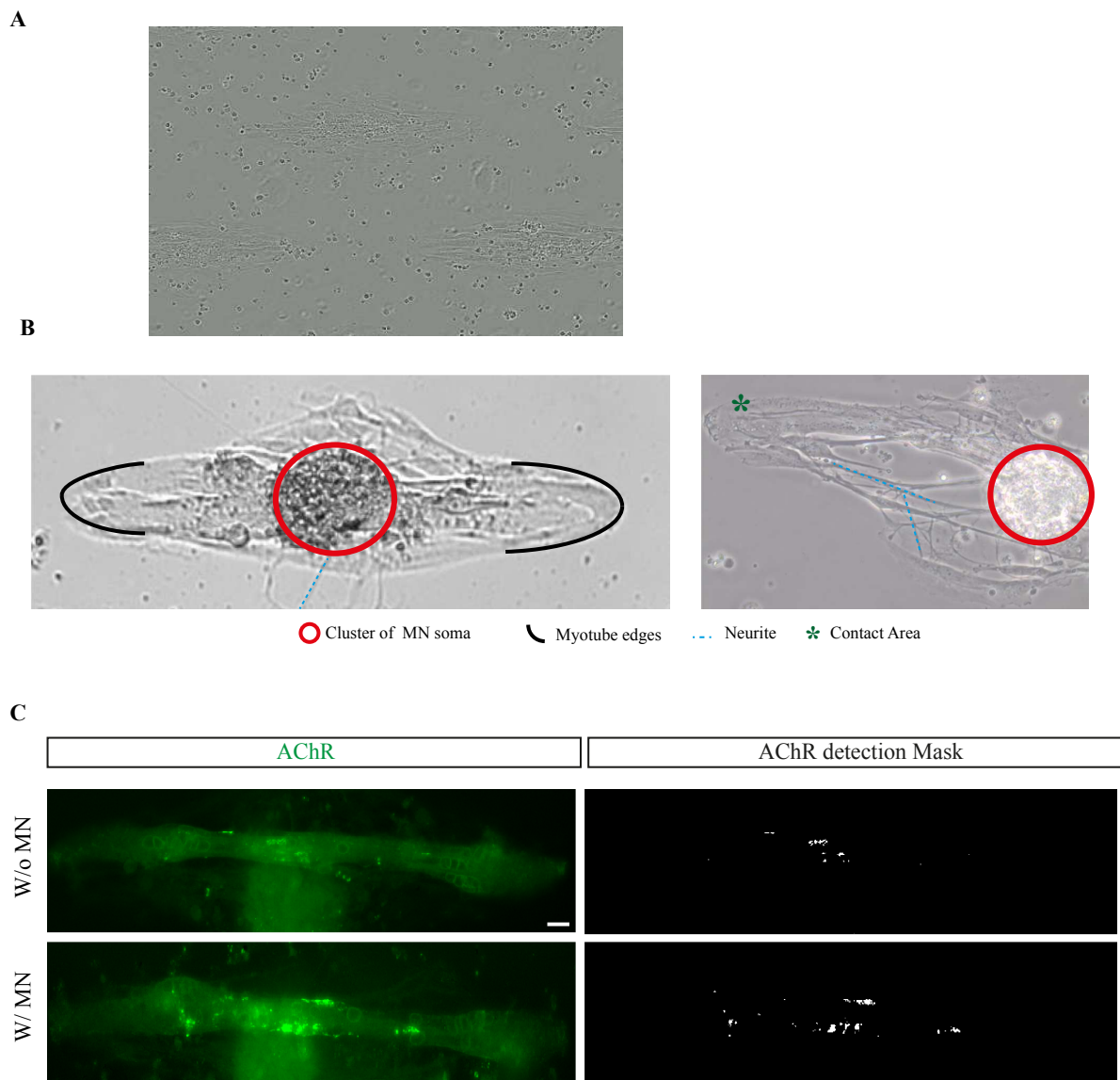
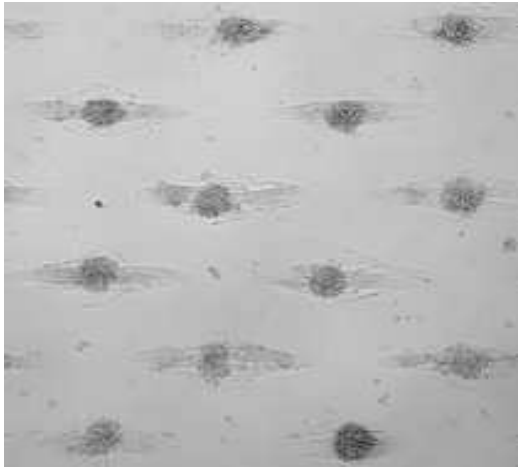


Figure S3 : Characterization of skeletal muscle cells and hiPSC-derived MN and morphology of the co-culture protocol.

(A) Movie of successive phase contrast images taken every 4 hours using Incucyte® to evaluate the set up and morphology of the co-culture system. (B) Morphology of the co-culture system observed by phase contrast microscopy. At the middle of the micropattern, the clustering of hiPSC-derived MN soma surrounded in red, the micropattern ends in black, neurites highlighted in blue and contact area marked by a green asterisk. (C) Representative examples of filter masks generated by an algorithm on Fiji Software for AChR detection based on representative images of AChR immunolabeling (in green). Scale bar 10µm.

Supplementary Figure 4

A



B

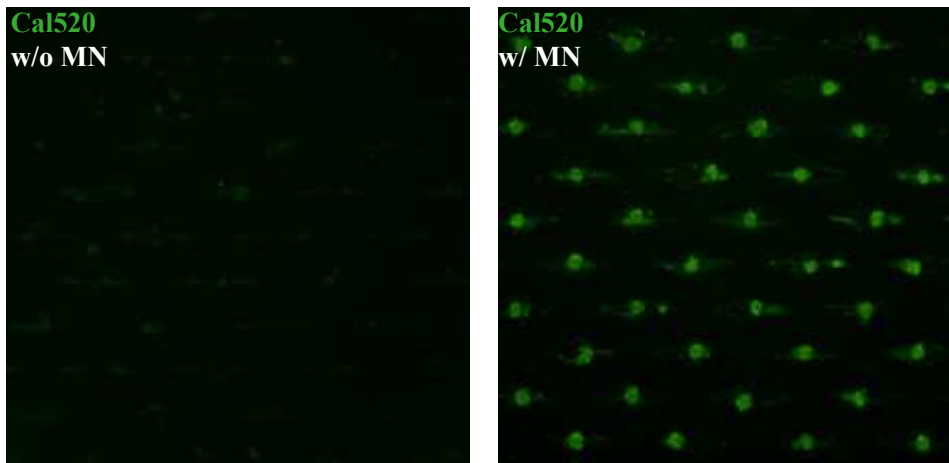


Figure S4 : Induction of muscle contraction and calcium flux by addition of hiPSC-derived MN.

(A) Live time lapse movie of visual muscle contraction 4 days after the addition of WT hiPSC-derived MN. Recording performed in phase contrast microscopy and under controlled atmosphere. **(B)** Timelapse movies of calcium flux in absence and presence of WT hiPSC-derived MN. Cells were incubated with the Cal520 calcium indicator. Recording performed under controlled atmosphere.

Supplementary Figure 5

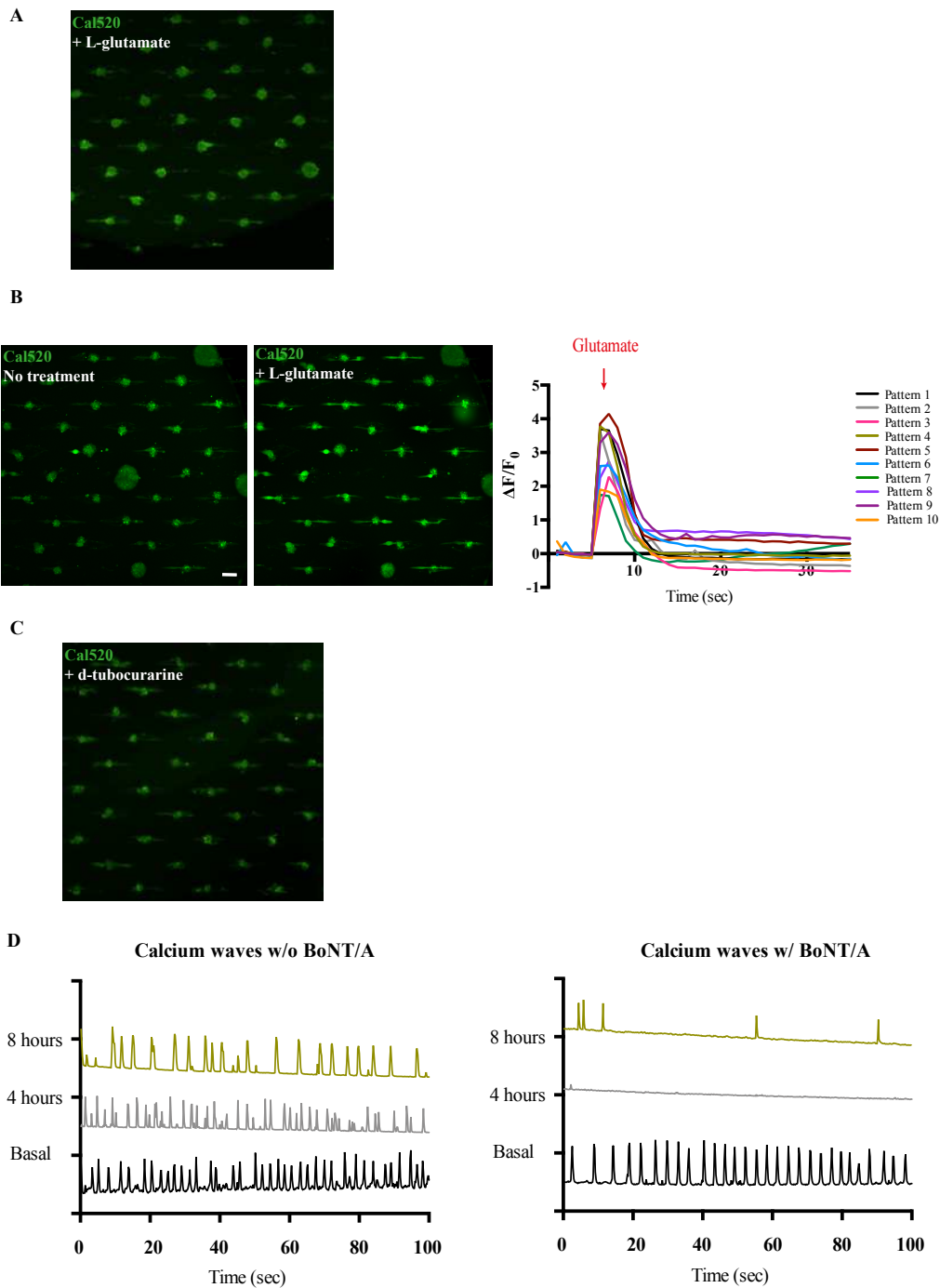


Figure S5: Pharmacological drugs control the communication between hiPSC-derived MN and skeletal muscle cells.

(A) Timelapse movie of calcium flux imaging in co-culture after the treatment of L-glutamate. Cells were incubated with the Cal520 calcium indicator. Recording performed under controlled atmosphere.

(B) Screenshots of timelapse movies of Ca^{2+} flux imaging (Green= Cal520 probe) in basal condition and after L-glutamate treatment ($100\mu\text{M}$). The diagram shows the micropatterns synchronization in the same well after L-glutamate treatment. $\Delta F/F_0$ representation after background correction. (C) Timelapse movie of calcium flux imaging in co-culture after the treatment of d-tubocurarine. Cells were incubated with the Cal520 calcium indicator. Recording performed under controlled atmosphere. (D) Representative diagram of Ca^{2+} transients in basal condition and 4 and 8 hours with or without BoNT/A treatment. $\Delta F/F_0$ representation after background correction.

Supplementary Figure 6

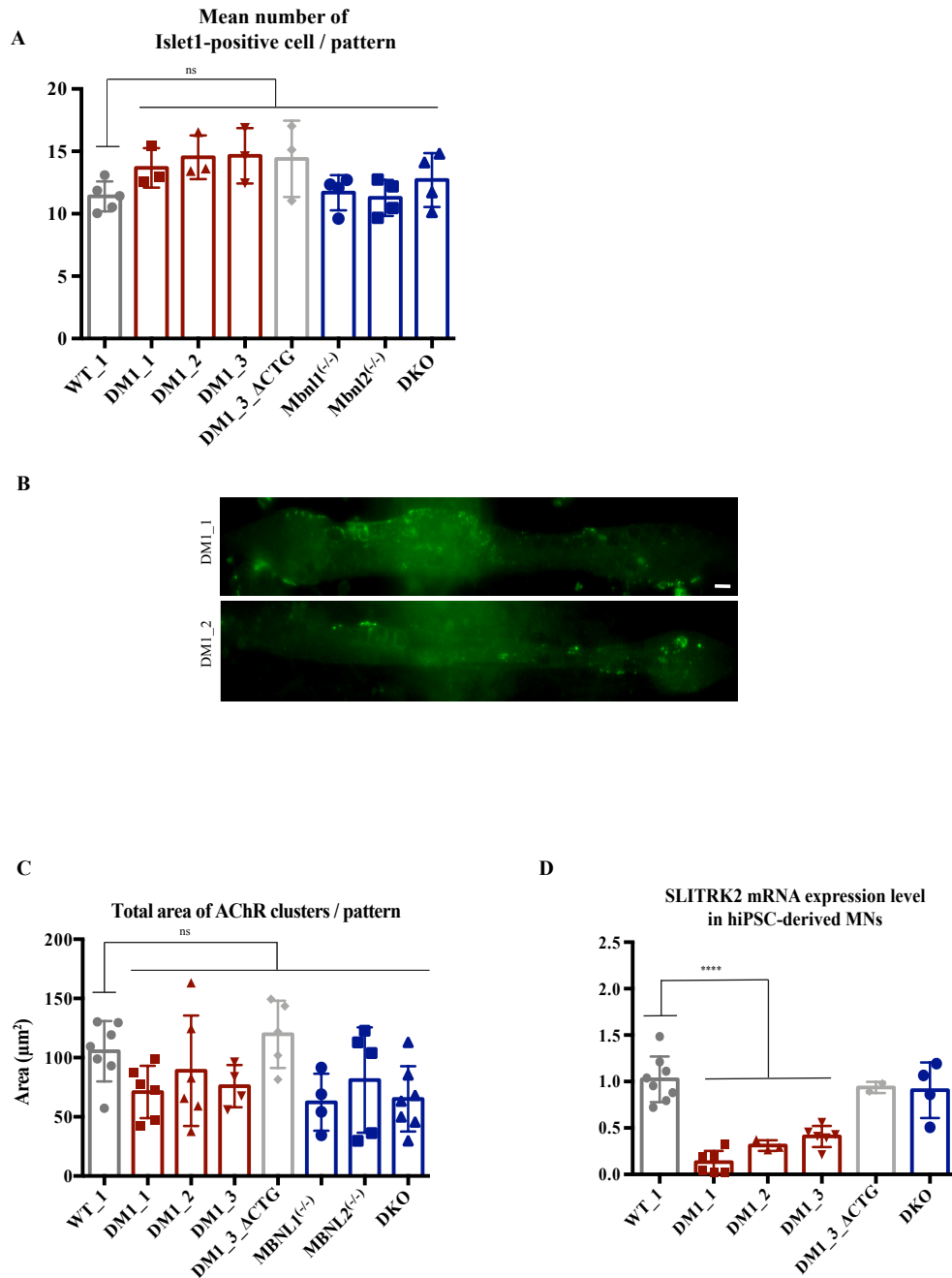


Figure S6: Negative effect of pathological hiPSC-derived MN on AChR clustering and gene expression.

(A) Quantification of Islet-1 positive cells per pattern. Data were analyzed with an ordinary One-Way anova, Tukey's multiple comparisons test compared to WT (ns: not significant). The mean \pm SD values from 3 independent experiments in triplicate and at least 150 micropatterns per quantification. (B) Representative images of AChR immunolabeling (in green) in DM1 hybrid co-culture conditions. Scale bar 10 μ m. (C) Quantification of the total area (μ m²) of AChR clusters per pattern. Data were analyzed with an ordinary One-Way anova, Tukey's multiple comparisons test compared to WT (ns: not significant). The mean \pm SD values from 3 independent experiments in triplicate and at least 150 micropatterns per quantification. (D) SLITRK2 mRNA expression level in hiPSC-derived MN. (n= at least 3 independent experiments in technical triplicate). Data were analyzed with an ordinary One-Way anova, Tukey's multiple comparisons test compared to WT (****: $P \leq 0,0001$). The mean \pm SD values from at least 3 independent experiments in triplicate.

Supplementary Figure 7

A

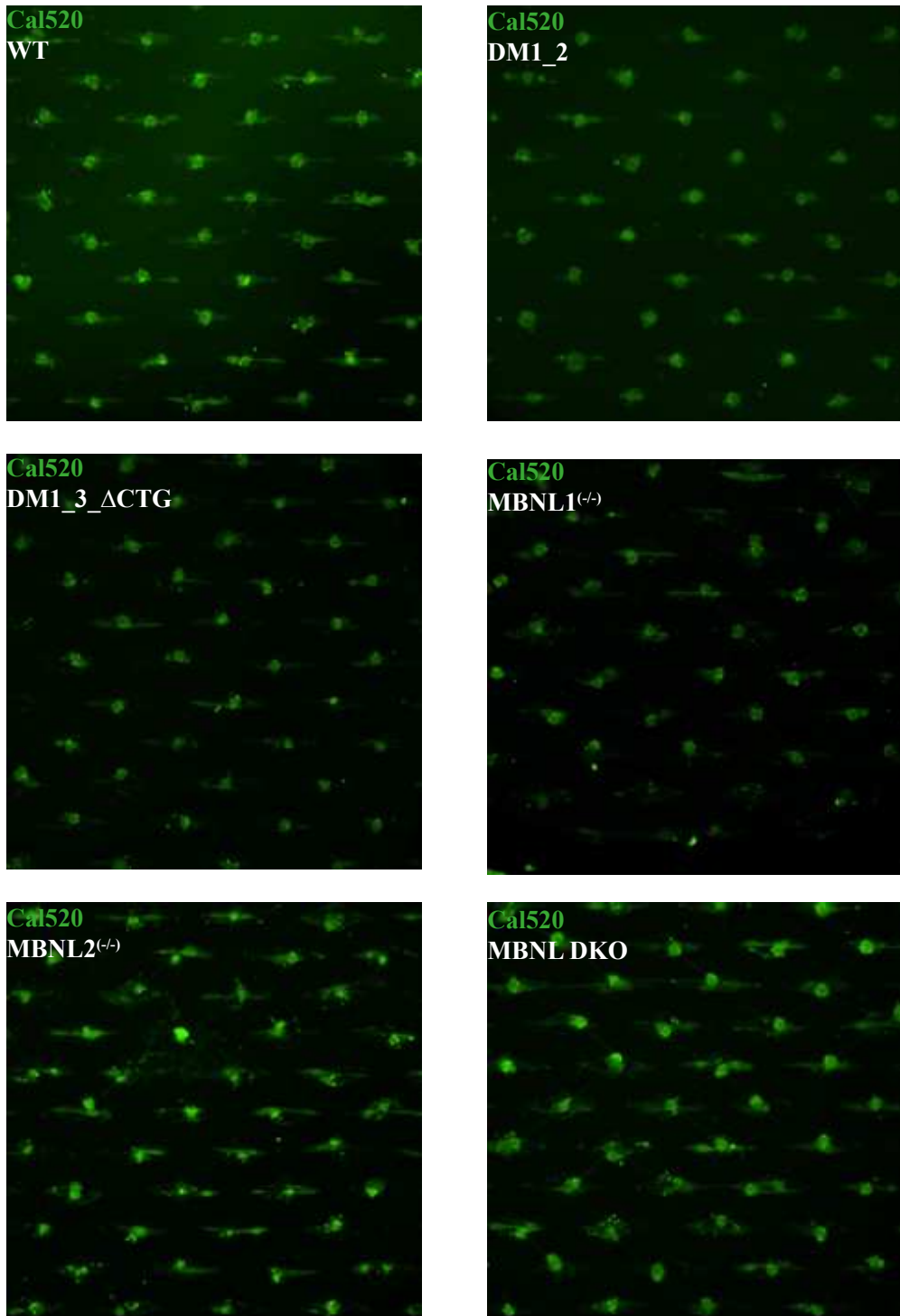
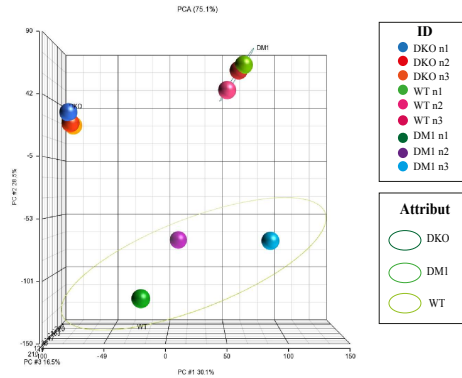


Figure S7: Communications defect between hiPSC-derived MN and skeletal muscle cells in hybrid systems.

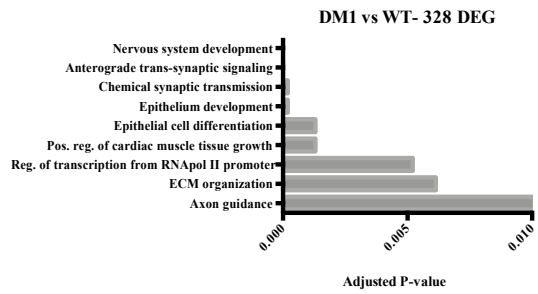
Timelapse movies of calcium flux imaging analysis (Green= Cal520, calcium indicator) in control and pathological hybrid conditions. Recording performed under controlled atmosphere.

Supplementary Figure 8

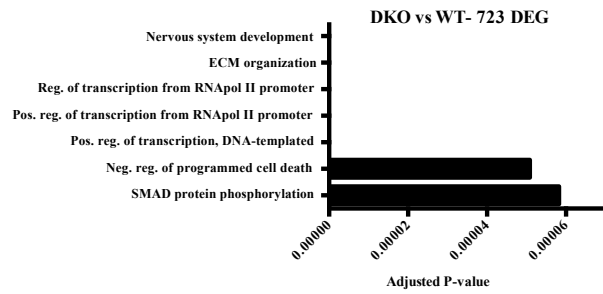
A



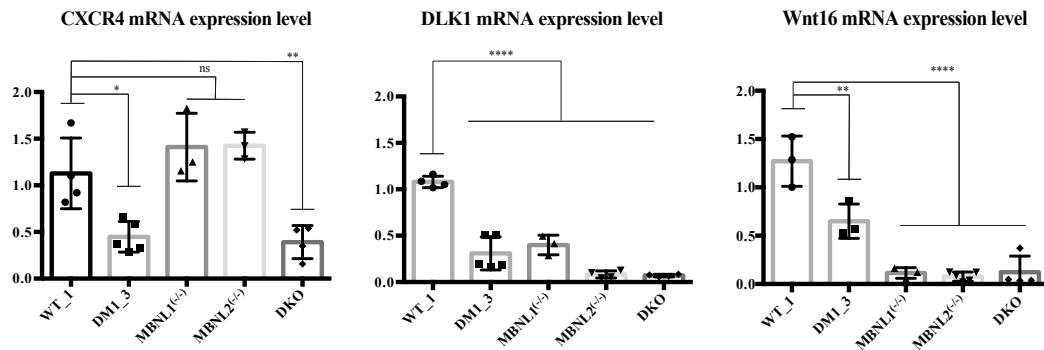
B



C



D



E

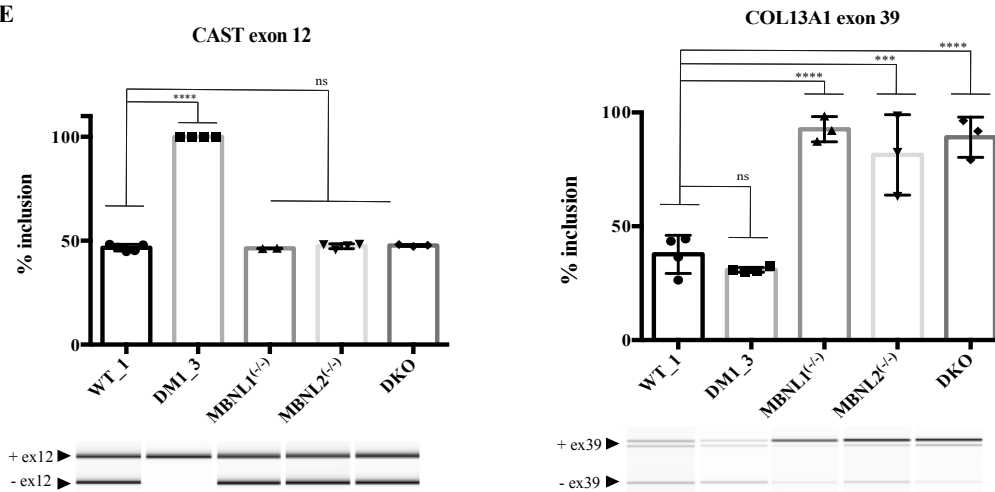
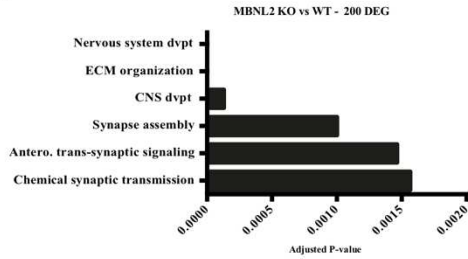


Figure S8: Identification of transcriptomic defects in DM1 and DKO hiPSC-derived MN.

(A) PCA (Principal Component Analysis) mapping indicating distinct transcriptomic profiles between WT, DM1 and DKO hiPSC-derived MN. **(B)** Genes set enrichment analysis using EnrichR on the 328 DEG in DM1 hiPSC-derived MN when compared to WT hiPSC-derived MN. (EnrichR analysis using GO Molecular Function 2018 database). **(C)** Genes set enrichment analysis using EnrichR on the 723 DEG in DKO hiPSC-derived MN when compared to WT hiPSC-derived MN. (EnrichR analysis using GO Molecular Function 2018 database). **(D)** CXCR4, DLK1 and Wnt16 mRNA expression level defect in hiPSC-derived MN. Data were analyzed with an ordinary One-Way anova, Tukey's multiple comparisons test compared to WT (*: $P \leq 0,05$, **: $P \leq 0,01$, ***: $P \leq 0,001$, ****: $P \leq 0,0001$, ns: not significant). The mean \pm SD values from 3 independent experiments in triplicate. **(E)** Alternative splicing analysis for CAST exon 12 and COL13A1 exon 39 in hiPSC-derived motoneuron. Data were analyzed with an ordinary One-Way anova, Tukey's multiple comparisons test compared to WT (***: $P \leq 0,001$, ****: $P \leq 0,0001$, ns: not significant). The mean \pm SD values from 3 independent experiments in triplicate.

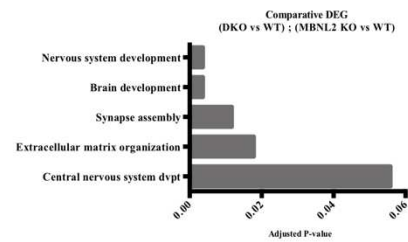
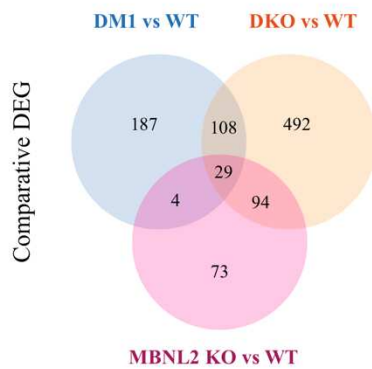
Supplementary Figure 9

A



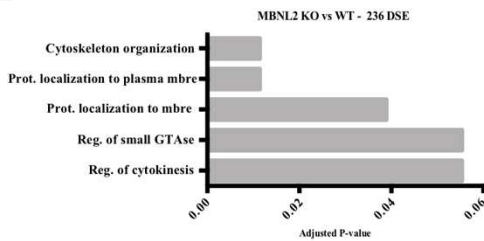
Selected enriched genes set			
Symbol	Log2 FC	Adj. p-value	Function
DLK1	-3,08	0,0000	Periactive zone
WNT16	-2,84	0,0013	Neural development and synaptogenesis
KIRREL3	-1,53	0,0141	Synapse formation
PLS3	1,54	0,0023	Neuronal outgrowth
LIN28A	2,59	0,0000	Axon regeneration

B



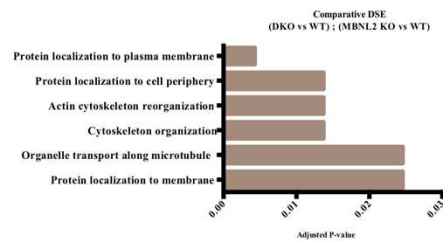
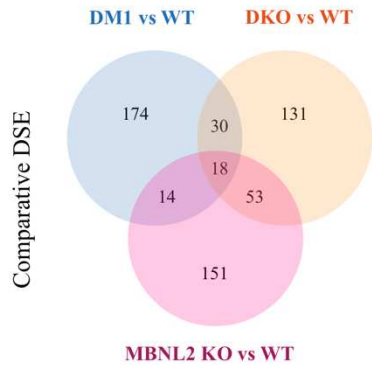
Symbol	Log2 FC			Function
	DKO vs WT	DM1 vs WT	M2 KO vs WT	
DLK1	-6,34	-2,20	-3,08	Periactive zone
WNT16	-4,52	-2,56	-2,84	AChR clustering
PLS3	1,05	-1,91	1,54	Neuronal outgrowth
CALB2	0,96	0,97	-1,38	Neuronal excitability

C



Selected enriched genes set				
Symbol	Exon	Δ psi	Adj. p-value	Function
SCN8A	8	-0,28	7,94E-03	Neuronal voltage-gated sodium channel
FLNA	30	-0,27	9,06E-03	Actin cytoskeleton
SNAP25	8	-0,24	9,05E-05	Fusion of synaptic vesicles
SORBS1	30	-0,22	1,83E-02	Cytoskeleton
PICALM	19	-0,19	8,53E-08	Fusion of synaptic vesicles
KIF13A	27	-0,14	1,55E-04	Molecular motor
CLASP1	27	-0,13	1,70E-04	Neuronal outgrowth and synaptic activity
UNC5C	9	0,32	1,88E-09	Axon guidance
DLG2	22	0,33	1,27E-03	Synapse structure and function

D



Selected enriched genes set				
Symbol	Exon	Δ psi		Function
		DKO vs WT	DM1 vs WT	
CAMTA1	24	0,13	0,13	Cell function and survival
UNC5C	9	0,14	0,14	Axon guidance

Figure S9: Identification of transcriptomic defects in MBNL2 knock-out hiPSC-derived MN.

(A) Genes set enrichment analysis using EnrichR on the 200 DEG genes in MBNL2^(-/-) hiPSC-derived MN when compared to WT hiPSC-derived MN (EnrichR analysis using GO Biological process 2018 database). The table on the right provides a list of specific genes differentially expressed in MBNL2^(-/-) hiPSC-derived MN. (B) Venn diagram representation of the number of DEG deregulated between the three (DM1, MBNL2^(-/-) and MBNL double knock-out) different conditions analyzed. On the top right, genes set enrichment analysis using EnrichR on the 123 DEG genes between DKO versus WT hiPSC-derived MN and MBNL2^(-/-) versus WT hiPSC-derived MN conditions (EnrichR analysis using GO Biological process 2018 database). On the bottom right, the table on the right provides a list of commonly DEG in the three conditions (DM1, MBNL2^(-/-), MBNL double knock-out) compared to WT hiPSC-derived MN. (C) Genes set enrichment analysis using EnrichR on the 236 DSE genes in MBNL2^(-/-) hiPSC-derived MN when compared to WT hiPSC-derived MN (EnrichR analysis using GO Biological process 2018 database). The table on the right provides a list of specific genes differentially spliced in MBNL2^(-/-) hiPSC-derived MN. (D) Venn diagram representation of the number of DSE deregulated between the three (DM1, MBNL2^(-/-) and MBNL double knock-out) different conditions analyzed. On the top right, genes set enrichment analysis using EnrichR on the 71 DSE genes between DKO versus WT hiPSC-derived MN and MBNL2^(-/-) versus WT hiPSC-derived MN conditions (EnrichR analysis using GO Biological process 2018 database). On the bottom right, the table on the right provides a list of commonly DSE in the three conditions (DM1, MBNL2^(-/-), MBNL double knock-out) compared to WT hiPSC-derived MN.

SUPPLEMENTARY MATERIAL & METHOD

Annexe 1: Detection of AChR on Fiji software

```
run("Brightness/Contrast...");
waitForUser("Regler Luminosité Contraste et cliquer sur OK");
run("8-bit");
run("Threshold...");
waitForUser("le threshold, cliquer sur Apply puis OK");
setOption("BlackBackground", true);
run("Convert to Mask");
run("Analyze Particles...", " show=Masks display summarize");
```

Supplementary Table 1: Medium and solutions composition

Growth Myoblasts Medium	Provider	Reference	C_{finale}
DMEM/F-12 Glutamax	Invitrogen	31331028	
Fœtal Bovine serum (FBS)	Sigma	F7524	20%
Penicilline/Streptomycine	Invitrogen	15140122	0,10%

Differentiation Myoblasts Medium	Provider	Reference	C_{finale}
DMEM/F-12 Glutamax	Invitrogen	31331028	
Horse serum	Sigma	H1270	2%

Motoneurons Medium	Provider	Reference	C_{finale}
DMEM/F-12 Glutamax	Invitrogen	31331028	1/2
Neurobasal medium	Invitrogen	21103049	1/2
Y-27632	StemCell	72302	10uM
B27 supplement, sans Vit A - 50X	Invitrogen	12587010	1/50
N2 supplement, 100X	Invitrogen	17502048	1/100
β-Mercapto Ethanol	Invitrogen	31350010	0,10%
BDNF	Miltenyi biotec	130-096-286	10 ng/ml
GDNF	Miltenyi biotec	130-096-449	10 ng/ml
DAPT	Sigma	D5942	10 μM

HBSS⁺⁺ solution	Provider	Reference	
10x HBSS	Invitrogen	14065-049	10%
Hepes Buffer 1M	Invitrogen	15630-056	1/50
NaOH 1N	Sigma	655104	1/400
Distilled water	Invitrogen	10977035	qsp

HBSS⁺ solution	Provider	Reference	
10x HBSS	Invitrogen	14175-053	10%
Hepes Buffer 1M	Invitrogen	15630-056	1/50
NaOH 1N	Sigma	655104	1/350
CaCl ₂ 1M	Sigma	C3306	1/500
Distilled water	Invitrogen	10977035	qsp

Supplementary Table 2: Antibodies

	Designation	Provider	Species	Dilution
Primary antibodies	Myosin Heavy Chain (MF20)	DSHB	Mouse	1/200
	Islet1	Neuromics	goat	1/1000
	Tubulin β III	BioLegend	Rabbit	1/1000
	Tubulin β III	BioLegend	Mouse	1/1000
	Acetylcholine receptor	DSHB	Rat	1/250
	Synaptophysin	Thermoscientific	Rabbit	1/250
	Bassoon	Synaptic System	Mouse	1/400
	VACHT	Synaptic System	Rabbit	1/500
	MBNL1	DSHB	Mouse	1/1000
	MBNL2	DSHB	Mouse	1/1000
	Anti- β -Actin-Peroxydase, monoclonal	Sigma	Mouse	1/10000
Secondary antibodies	Donkey anti-mouse IgG(H+L) alexa fluor 568	ThermoFisher		1/1000
	Donkey anti-mouse IgG(H+L) alexa fluor 647	ThermoFisher		1/1000
	Donkey anti-mouse IgG(H+L) alexa fluor 488	ThermoFisher		1/1000
	Donkey anti-goat IgG(H+L) alexa fluor 488	ThermoFisher		1/1000
	Donkey anti-goat IgG(H+L) alexa fluor 647	ThermoFisher		1/1000
	Donkey anti-rabbit IgG(H+L) alexa fluor 647	ThermoFisher		1/1000
	Donkey anti-rat IgG(H+L) alexa fluor 488	ThermoFisher		1/1000
	Peroxidase-conjugated secondary antibody	ThermoFisher		1/10000
	Hoechst (1 mg/ml)	Invitrogen		1/1000

Supplementary Table 3: Primer sequences

Gene	Forward	Reverse
GAPDH	GTCTCCTCTGACTTCAACAGCG	ACCACCCTGTTGCTGTAGCCAA
SLITRK2	TGCAGTCATTCAGGAAGGTG	GCTCTGCAGTCCATCAAACA
MBNL1	TTGCAAATTCAGCTGTGAGG	CGAATTTCCAAGCTGCTTTC
MBNL2	GCCCATGCCAACAGTCTAAT	GGCAACTGGATGGTGAGTTT
18S	GAGGATGAGGTGGAACGTGT	TCTTCAGTCGCTCCAGGTCT
Cyclophilin A	CCCACCGTGTCTTCGACAT	CCAGTGCTCAGAGCACGAAA

DISCUSSION GÉNÉRALE

Comme nous l'avons vu, les iPSC humaines associées au système CRISPR/Cas9 ont été les outils cellulaires et moléculaires centraux de cette thèse. Ces deux technologies ainsi que les récentes avancées concernant les protocoles de différenciation cellulaire m'ont permis de développer un nouveau modèle cellulaire pour l'étude du rôle des protéines MBNL, et particulièrement de leur contribution dans la physiopathologie musculaire de la DM1. Dans cette étude nous avons ainsi pu développer et mettre en évidence différents points :

1) Mes travaux de recherche ont permis le développement de nouveaux modèles d'étude de la DM1 caractérisé par la différenciation ciblée d'hiPSC porteuses de la mutation causale de la pathologie en cellules musculaires squelettiques et ce en l'absence de transgène.

2) Nous avons pu dans un second temps démontrer la temporalité d'action des protéines MBNL au long de la myogenèse humaine *in vitro* ainsi que leur contribution dans les principaux défauts moléculaires et cellulaires associés à la DM1.

3) La génération de ces lignées *MBNL* KO a également contribué à mettre en évidence les phénomènes de compensation entre les deux paralogues MBNL1 et MBNL2, notamment par des analyses transcriptomiques aboutissant à l'identification de nombreuses cibles dérégulées par la déplétion des protéines MBNL, au niveau de la régulation de l'expression génique ainsi que de l'épissage alternatif.

4) Ces analyses transcriptomiques et l'utilisation des techniques de saut d'exon médiées par les oligonucléotides anti-sens nous ont finalement permis d'associer le défaut d'épissage de l'exon 25 de *SORBS1* à un défaut de formation des clusters des récepteurs à l'acétylcholine (AChR) ; phénotype retrouvé dans les myotubes dérivés d'hiPSC DM1 et DKO

Dans le cadre du projet de thèse d'une autre doctorante de l'équipe, la génération des lignées hiPSC KO pour les différents gènes MBNL a parallèlement contribué à l'étude du rôle de ces protéines dans un autre type cellulaire, les motoneurones, dans le but de mieux comprendre la contribution de la composante pré-synaptique dans la DM1. Nous discuterons des résultats de ces recherches à la fin de cette partie.

A. Stratégie de modélisation pathologique pour mieux comprendre les désordres musculaires associés à la DM1

À ce jour, seules quelques études ont décrit la génération de cellules musculaires squelettiques dérivées d'hiPSC DM1. Celles-ci ont permis de reproduire certaines caractéristiques moléculaires de la pathologie, à savoir l'agrégation nucléaire des ARNm mutants et certains défauts d'épissage. Malgré ces résultats prometteurs, toutes ces études

ont utilisé des protocoles de différenciation basés sur l'expression forcée de facteurs myogéniques, le plus souvent MyoD ou Pax7 (Mondragon-Gonzalez and Perlingeiro 2018; Wang et al. 2018a; Ueki et al. 2017). La surexpression de ces facteurs de transcription a ainsi montré ses preuves pour la modélisation pathologique de maladies musculaires ainsi que pour le criblage de médicaments. Toutefois, ces approches de reprogrammation directe ont l'inconvénient de contourner les premières étapes développementales de la myogenèse, étapes qui pourraient être primordiales dans la mise en place de pathologies, en particulier dans le cadre de maladies génétiques (Chal and Pourquié 2017).

Récemment, plusieurs protocoles d'induction myogénique basés sur l'utilisation de petites molécules et de facteurs de croissance ont été développés et représentent une alternative prometteuse pour mieux récapituler de étapes développementales successives impliquées dans la myogenèse (Caron et al. 2016; Chal et al. 2016; Shelton et al. 2014). Le protocole expérimental décrit dans notre étude utilise un système sans transgène et ouvre ainsi de nouvelles perspectives quant à la modélisation de la DM1. Ce protocole de différenciation myogénique présente cependant certaines limitations comme par exemple i) l'absence de cellules PAX7⁺, dites "satellite-like", et importantes dans les processus de régénération musculaire, ou encore ii) une maturation des myotubes limitée, comparable à un phénotype fétal (Chal and Pourquié 2017). Toutefois, les différentes étapes bien caractérisées de ce protocole permettent de décrypter l'impact exact de la mutation tout au long de la différenciation myogénique. Alors que la plupart des études ont porté sur des cellules musculaires squelettiques primaires humaines ou des cellules musculaires immortalisées issues de biopsies de patients DM1 (Arandel et al. 2017), nos résultats indiquent que l'engagement myogénique précoce n'est pas affecté par la mutation causale de la pathologie.

En outre, le protocole décrit ici est constitué d'une étape de congélation nous permettant de générer de grandes banques cellulaires représentant une source illimitée de matériel qui pourrait être utilisée par la suite pour des criblages thérapeutiques par exemple. Il est également envisageable d'utiliser ces cellules pour aller vers des essais plus fonctionnels, en les cultivant par exemple dans des systèmes tridimensionnels plus complexes, avec d'autres types cellulaires comme des motoneurons ou encore des cellules endothéliales vasculaires (Maffioletti et al. 2018). Le protocole de différenciation en trois étapes que nous décrivons offre aussi l'opportunité d'étudier le rôle d'une protéine à différents stades développementaux. Il permet notamment d'envisager l'utilisation de systèmes inductibles pouvant inhiber ou activer l'expression d'un gène à un temps bien spécifique de la différenciation myogénique. En parallèle de cette étude, j'ai ainsi pu contribuer au sein du laboratoire à la génération d'une lignée hiPSC Cas9 inductible (iCas9) via la doxycycline aboutissant par exemple à l'obtention d'une banque de myoblastes Cas9 inductibles. Par ailleurs, ces hiPSC Cas9 inductibles m'ont été très utiles afin de tester l'efficacité des guides ARN ciblant les gènes *MBNL* (**Figure 45**) avant de procéder aux expériences de transfections

avec la Cas9 sous forme de complexes RNP et d'obtenir ainsi les différents clones. À l'avenir, on pourrait imaginer utiliser ces hiPSC Cas9 inductibles dans n'importe quel autre protocole de différenciation cellulaire et ainsi induire une extinction génique à des moments précis du développement.

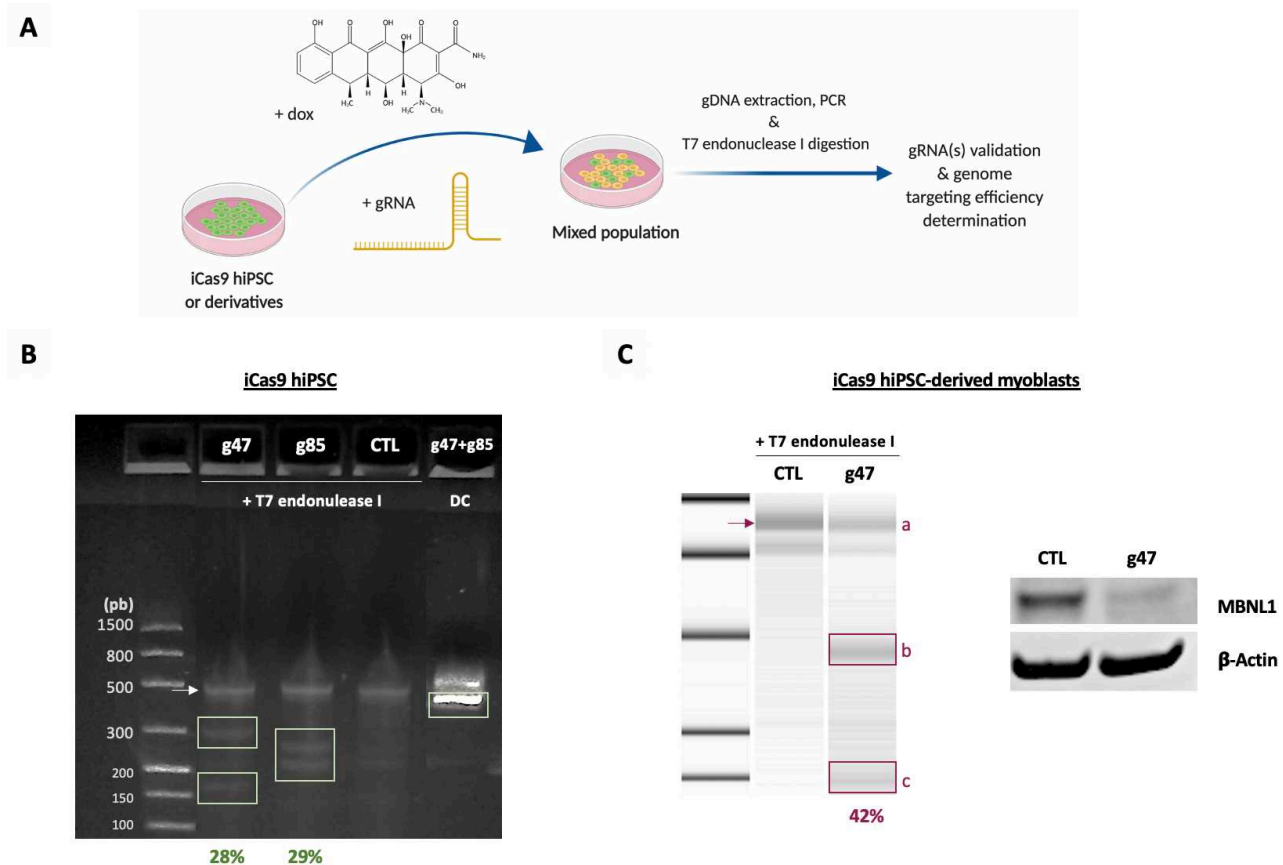


Figure 45 : Utilisation des hiPSC iCas9 et de leurs dérivés

(A) Représentation schématique de la stratégie générale utilisée pour valider les guides ARN (gRNA) et afin de déterminer leur efficacité d'édition génomique. Dox : doxycycline. (Crée à partir de *Biorender.com*). **(B-C)** Exemples de validation de guides ARN (g47 et g85) ciblant le premier exon codant du gène *MBNL1* dans des hiPSC iCas9 **(B)** ou dans des myoblastes dérivés d'hiPSC iCas9 **(C)**. La double coupure (DC) observée en **B** et induite par l'action des deux guides, permet de valider rapidement que les deux guides fonctionnent. L'efficacité d'édition de chaque guide est ensuite calculée après digestion des produits de PCR par la T7 endonucléase I et est représentée par le $\% \text{Indel} = 100 \times (1 - \sqrt{1 - F_{\text{cut}}})$ où $F_{\text{cut}} = (b+c) / (a+b+c)$, avec a correspondant à l'intégrale de l'intensité du produit de PCR non digéré (indiqué par une flèche), et b et c celle de chaque produit de clivage. L'analyse par Western blot des myoblastes iCas9 transfectés ou non avec le guide ARN ciblant le gène *MBNL1* **(C - à droite)** confirme l'édition génique partielle de la population cellulaire totale, avec une diminution de l'expression protéique de MBNL1.

B. Rôle des protéines MBNL dans la myogénèse humaine et contribution dans la physiopathologie musculaire de la DM1

- Fonctions des protéines MBNL1 et MBNL2 dans la fusion myogénique ?

Par la génération via CRISPR/Cas9 de lignées hiPSC KO pour les différents gènes *MBNL*, nous avons pu ainsi montrer que les protéines MBNL n'étaient pas essentielles pour l'engagement myogénique précoce mais qu'elles étaient en revanche nécessaires pour les stades de différenciation plus tardifs, indiquant une réelle réquisition temporelle de ces protéines.

Ces résultats sont en accord avec une étude récente montrant que certains facteurs d'épissage tels que Nova ou Rbfox interviennent majoritairement lors de transitions développementales précoces d'épissage alternatif alors que les protéines MBNL seraient plutôt identifiées comme régulateurs de transitions plus tardives au cours du développement neural post-natal (Weyn-Vanhentenryck et al. 2018). De plus, les dérivés myogéniques de ces hiPSC KO *MBNL* récapitulent les défauts moléculaires et cellulaires caractéristiques de la DM1 avec une mise en place de certains défauts d'épissage dès le stade myoblaste. Ces défauts sont d'autant plus importants au stade myotube avec également l'observation d'un défaut de fusion myogénique à la fois dans les myotubes dérivés d'hiPSC DM1 et d'hiPSC *MBNL* DKO. En utilisant une telle approche, notre étude ajoute une nouvelle dimension au rôle des protéines MBNL1 et MBNL2 dans la définition des étapes de maturation des muscles squelettiques humains.

Les mécanismes exacts par lesquels les protéines MBNL affectent ce processus de fusion myogénique restent cependant indéterminés. Contrairement à Rbfox1/2 pour qui il a été démontré que la déplétion inhibait la fusion des myoblastes *in vitro* via un défaut d'épissage de Mef2D (Pedrotti et al. 2015; Runfola et al. 2015), nos résultats indiquent que la déplétion des protéines MBNL conduit à l'expression aberrante de plus d'une centaine de gènes. Parmi ces gènes dérégulés, plusieurs codent des facteurs impliqués dans l'organisation sarcomérique, dans la dynamique du réseau d'actine ou encore dans les événements de trafic vésiculaire nécessaires à la fusion membranaire (Sampath and Millay 2018). Toutefois, nous n'avons pas observé de dérégulation concernant l'expression de Myomaker ou de Myomerger, deux protéines membranaires impliquées dans la fusion des myoblastes (Leikina et al. 2018; Millay et al. 2013), suggérant l'implication d'autres mécanismes indépendants de ces deux facteurs. D'autres expériences seront nécessaires pour évaluer les mécanismes dépendants des protéines MBNL et nécessaires à la différenciation myogénique. Dans un contexte moins mécanistique, notre étude fournit une liste de gènes dérégulés par l'absence des protéines MBNL dont il pourrait être intéressant d'évaluer leur intérêt pour de futures applications thérapeutiques.

- Importance des phénomènes compensatoires entre les paralogues MBNL1 et MBNL2

La différenciation myogénique des lignées hiPSC différentiellement déplétées en protéines MBNL nous a permis de confirmer que MBNL1 est la protéine MBNL la plus abondante au cours de la différenciation myogénique *in vitro* (Charizanis et al. 2012; Holt et al. 2009) et qu'un phénomène de compensation existe entre ces deux paralogues MBNL1 et MBNL2, notamment illustrée par la surexpression de la protéine MBNL2 à la fois dans les myotubes dérivés des hiPSC *MBNL1(-/-)* et des hiPSC *MBNL1(-/-); MBNL2(+/-)* par rapport aux myotubes dérivés d'hiPSC WT et d'hiPSC *MBNL2(+/-)* respectivement. De manière générale, les myotubes dérivés des hiPSC DKO montrent des défauts moléculaires et cellulaires aggravés par rapport aux myotubes dérivés des hiPSC simples KO *MBNL1* et *MBNL2*. Nous avons pu montrer par des analyses transcriptomiques que plusieurs processus biologiques étroitement liés au développement et fonctionnement du muscle squelettique étaient down-régulés dans les myotubes dérivés d'hiPSC DKO par rapport aux myotubes dérivés d'hiPSC *MBNL1(-/-)*. Comme cela avait été observé chez la souris dans des études antérieures (Goodwin et al. 2015; Lee et al. 2013), nos résultats renforcent l'idée qu'un processus de compensation entre les deux paralogues MBNL1 et MBNL2 se met naturellement en place. Cette capacité de pouvoir compenser la perte d'expression d'un gène, ayant un rôle important au niveau développemental par exemple, contribue au principe de robustesse génétique qui représente la faculté d'un organisme vivant à maintenir sa viabilité et ses fonctions malgré les perturbations environnementales ou génétiques (El-Brolosy and Stainier 2017; Diss et al. 2014). Cependant les mécanismes moléculaires de compensation fonctionnelle entre gènes paralogues sont encore aujourd'hui insuffisamment compris. Une première hypothèse pourrait être l'existence d'un signal de rétro-contrôle négatif en temps normal, et qui ne serait plus mis en place lors de la perte de fonction d'un des deux gènes paralogues. De manière intéressante, l'absence d'une telle réponse compensatoire dans des modèles de knock-down (par siRNA par exemple), par rapport aux mutants knock-out correspondants, indique qu'un mécanisme en amont de la fonction protéique est en jeu. Ce mécanisme pourrait se mettre en place au niveau des ARNm mutants voire même au niveau de la mutation génomique directement (El-Brolosy and Stainier 2017).

Contrairement à ces deux paralogues, l'expression de MBNL3 est beaucoup plus faible et limitée. Le rôle du troisième paragone au cours de la myogénèse reste encore aujourd'hui difficilement perçu. Alors que l'expression constitutive de *Mbnl3* dans des cellules musculaires murines semble inhiber la différenciation terminale (Lee et al. 2008a), il a été récemment démontré que les myoblastes primaires murins déplétés en *Mbnl3* présentaient une différenciation *in vitro* défectueuse (Thomas et al. 2017). Cependant, ne détectant pas l'expression de la protéine MBNL3 dans les myoblastes ou myotubes dérivés d'hiPSC, nos résultats sont en accord avec les observations d'une étude antérieure montrant que son expression était détectée dans les myoblastes primaires murins mais pas dans les myoblastes primaires humains (Holt et al. 2009). En effet, nous n'avons observé aucune différence

significative au niveau des phénotypes cellulaires et moléculaires entre myotubes dérivés d'hiPSC DKO et TKO, suggérant un rôle beaucoup moins important de la protéine MBNL3 au cours de la myogénèse humaine. Toutefois, la génération d'une lignée hiPSC simple KO *MBNL3(-/-)* ainsi que des analyses transcriptomiques plus poussées sur les cellules musculaires dérivées d'hiPSC TKO seraient nécessaires pour confirmer ces interprétations.

- Contribution des protéines MBNL1 et MBNL2 dans la DM1

La séquestration des protéines MBNL est considérée comme l'un des facteurs majeurs dans la mise en place de la DM1. Des analyses transcriptomiques comparatives entre des souris *Mbnl1(-/-)* et des souris transgéniques exprimant des répétitions CUG ont montré qu'environ 70% des changements induits par les expansions anormales dans les muscles squelettiques pourraient s'expliquer par une perte de l'activité de Mbnl1. De la même manière, environ 80% des défauts d'épissage musculaires identifiés dans les souris HSA^{LR} ont été retrouvés dérégulé dans les souris *Mbnl1(-/-)* (Du et al. 2010; Osborne et al. 2009; Wang et al. 2012). De plus, il a été démontré que les souris DKO présentaient une splicéopathie comparable à celle observée chez les patients DM1, et plus particulièrement chez les patients atteints de la forme congénitale, forme la plus sévère de la maladie et associée à un grand nombre de répétitions CTG (Lee et al. 2013; Thomas et al. 2017). En revanche, nous avons montré dans notre étude que la double déplétion des protéines MBNL1 et MBNL2 récapitulait environ 36% des défauts d'épissage que nous avons pu observer dans les cellules musculaires squelettiques dérivées des hiPSC DM1 portant plus de 2500 répétitions CTG. Dans l'ensemble, nos résultats suggèrent ainsi l'implication de mécanismes supplémentaires et indépendants des protéines MBNL dans la physiopathologie humaine de la maladie. En effet, bien que la déplétion des protéines MBNL soit un événement pathologique majeur de la DM1, de nombreuses autres voies peuvent jouer un rôle important dans la splicéopathie associée à la maladie, en commençant par la dérégulation d'une autre famille de facteurs d'épissage, les protéines CELFs. En effet, la stabilisation et la surexpression de CUG-BP1, dues à son hyper-phosphorylation en raison de l'altération de la voie GSK3 β dans la DM1, est également à l'origine de nombreux défauts d'épissage (Kalsotra et al. 2008; Wang et al. 2015). Par ailleurs, la hnRNP-H est un autre facteur d'épissage également retrouvé dérégulé dans la DM1 et pourrait aussi contribuer à la splicéopathie caractéristique de la maladie (Paul et al. 2006).

Néanmoins, nos résultats mettent en évidence de nombreux processus biologiques liés à l'homéostasie et à l'organisation du muscle squelettique dérégulés à la fois dans les myotubes dérivés d'hiPSC DM1 et d'hiPSC DKO. Ces analyses nous ont permis d'identifier ainsi une trentaine de défauts d'épissage couvrant la plupart de ces processus biologiques dérégulés liés au muscle. Certains des défauts d'épissage observés ont déjà été décrits

comme spécifiquement associés aux signes cliniques de la pathologie DM1. Par exemple, des études ont démontré que le défaut d'inclusion de l'exon 78 du gène *DMD* était suffisant pour provoquer une désorganisation des fibres musculaires (Rau et al. 2015) et que l'exclusion de l'exon 11 du gène *BIN1* était également suffisante pour induire une faiblesse musculaire (Fugier et al. 2011).

C. Mise en évidence du défaut d'épissage de l'exon 25 de *SORBS1* nécessaire à la formation des clusters des AChR

- Pourquoi le défaut d'épissage de l'exon 25 de *SORBS1* ?

Avec pour objectif d'identifier de nouveaux mécanismes moléculaires qui pourraient être impliqués dans la pathologie musculaire de la DM1, nous nous sommes intéressés à l'étude d'autres défauts d'épissage pour lesquels aucun phénotype cellulaire n'avait été rapporté jusqu'à présent. Parmi les candidats potentiels, *SORBS1* nous a particulièrement intéressés pour plusieurs raisons. 1) Le défaut d'inclusion de l'exon 25 de *SORBS1* est l'un des défauts d'épissage les plus sévères observés communément dans les myotubes dérivés des hiPSC DM1 et des hiPSC DKO ; 2) cette anomalie d'épissage a également été recensée dans plusieurs études transcriptomiques réalisées sur des biopsies musculaires de patients adultes DM1 (Wang et al. 2019a) et a été identifiée comme l'un des 38 défauts d'épissage corrélés avec l'immaturité musculaire chez les patients atteints de la forme congénitale de la DM1 (Nakamori et al. 2017). Ce défaut a également été retrouvé dans le muscle cardiaque de souris *Mbnl1(-/-)* ; *Mbnl2(+/-)* (Lee et al. 2013). 3) L'homologue murin de *SORBS1* a été identifié pour interagir avec le complexe MuSK/Dok-7/Crk et la diminution de l'expression de *Sorbs1* aboutit à des défauts d'aggrégation des récepteurs à l'acétylcholine (AChR) dans des cellules murines C2C12, et donc à un défaut de mise en place du compartiment post-synaptique (Hallock et al. 2016). De manière intéressante, des études ont pu observer chez les souris *Mbnl* DKO et TKO de nombreux défauts morphologiques des JNM du diaphragme, celles-ci étant moins matures, moins complexifiées et beaucoup plus fragmentées (Thomas et al. 2017; Lee et al. 2013).

- Quelles peuvent être les conséquences fonctionnelles du défaut d'épissage de *SORBS1* ?

Concordant avec les résultats obtenus sur l'homologue murin, notre étude indique que le défaut d'inclusion de l'exon 25 de *SORBS1* est responsable d'un défaut de mise en place ou de maintien des clusters des AChR, également observé dans les myotubes dérivés d'hiPSC DM1 et DKO. Afin de compléter les expériences décrites dans notre étude et qui visaient

à induire l'exclusion de l'exon 25 de *SORBS1* dans des cellules musculaires contrôles, j'ai également initié des expériences inverses qui visent à corriger ce défaut d'épissage dans des myotubes dérivés d'hiPSC DM1, sous-entendant forcer l'inclusion de l'exon 25 dans les conditions DM1. En me basant également sur la technologie CRISPR/Cas9, une stratégie a été établie afin d'intégrer au niveau génomique par knock-in dans des hiPSC DM1, la séquence d'ADNc du transcrite majeur de *SORBS1*, au détail près d'y avoir inséré l'exon 25 (**Figure 46**).

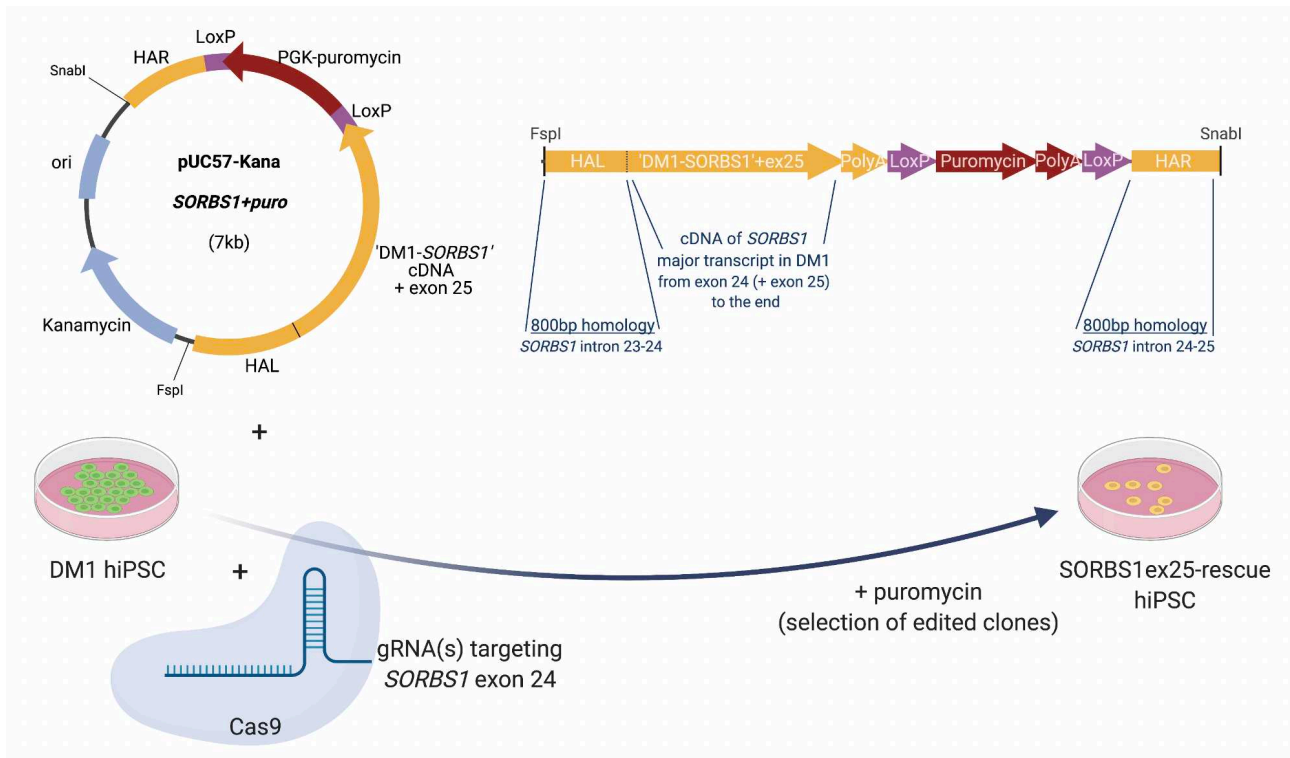


Figure 46 : Stratégie de knock-in pour corriger le défaut d'épissage de l'exon 25 de *SORBS1* dans des hiPSC DM1

La protéine Cas9 complexée au guide ARN (gRNA) induit une coupure double brin au niveau de l'exon 24 du gène *SORBS1*, permettant l'introduction de la construction exogène par recombinaison homologue. Celle-ci contient en effet deux bras d'homologie gauche ou droit (HAL et HAR) d'environ 800 paires de bases chacun et complémentaires des régions introniques qui encadrent l'exon 24. Elle contient également un gène de résistance à la puromycine afin de faciliter par la suite la sélection des clones édités. (Crée à partir de Biorender.com).

Ainsi, après différenciation, les transcrits *SORBS1* exprimés par les myotubes dérivés de ces hiPSC DM1 édités différeront uniquement par la présence de l'exon 25. J'ai pu obtenir plusieurs clones actuellement en cours d'analyse. Nous évaluerons par la suite si l'inclusion de l'exon 25 permet à elle-seule de corriger le défaut de clustering des AChR dans ces myotubes DM1 *SORBS1*ex25-rescue.

Nous n'avons malheureusement pas eu l'opportunité de valider ce phénotype sur des biopsies au niveau des jonctions neuromusculaires (JNM) de patients en raison de la

difficulté de leur obtention. Afin de confirmer nos résultats, il sera nécessaire par la suite de valider nos observations sur des modèles *in vivo*.

Dans ce but, des expériences reprenant la stratégie des ASO pour le saut d'exon de l'exon 25 de *SORBS1* ont été récemment effectuées sur des souris en collaboration avec l'équipe de Denis Furling à l'Institut de Myologie. Brièvement, la stratégie de saut d'exon de l'exon 25 de *SORBS1* utilisée dans notre étude a été adaptée en injectant, dans le muscle tibial antérieur de souris, un vecteur AAV exprimant un snRNA U7 modifié, contenant la séquence de l'ASO ciblant la séquence ESE murine de l'exon 25 de *Sorbs1*. Le but de ces expériences est d'analyser, deux mois post-injection, les effets de l'exclusion de cet exon au niveau de la morphologie des jonctions neuromusculaires (JNM) et de l'histologie plus globale des fibres musculaires, ainsi que d'un point de vue fonctionnel, en évaluant la force musculaire de ces souris.

Les résultats préliminaires suggèrent que l'exclusion de l'exon 25 de *Sorbs1* n'entraîne pas d'atteintes fonctionnelles significatives. En revanche, une première étude morphologique des JNM semble indiquer un potentiel effet de ce défaut d'épissage, avec l'observation plus fréquente de structures rétractées (**Figure 47-A,B**). L'importante hétérogénéité morphologique des JNM au sein d'un même muscle, injecté ou non, complique cependant les analyses et ne nous permet pas pour l'instant de confirmer ces observations. De façon intéressante, une analyse par microscopie électronique semble suggérer des anomalies au niveau des JNM avec la présence de structures dégénératives dans l'espace péri-synaptique (**Figure 47-C,D**). Des analyses supplémentaires sur d'autres souris injectées seront nécessaires pour confirmer ces premiers résultats.

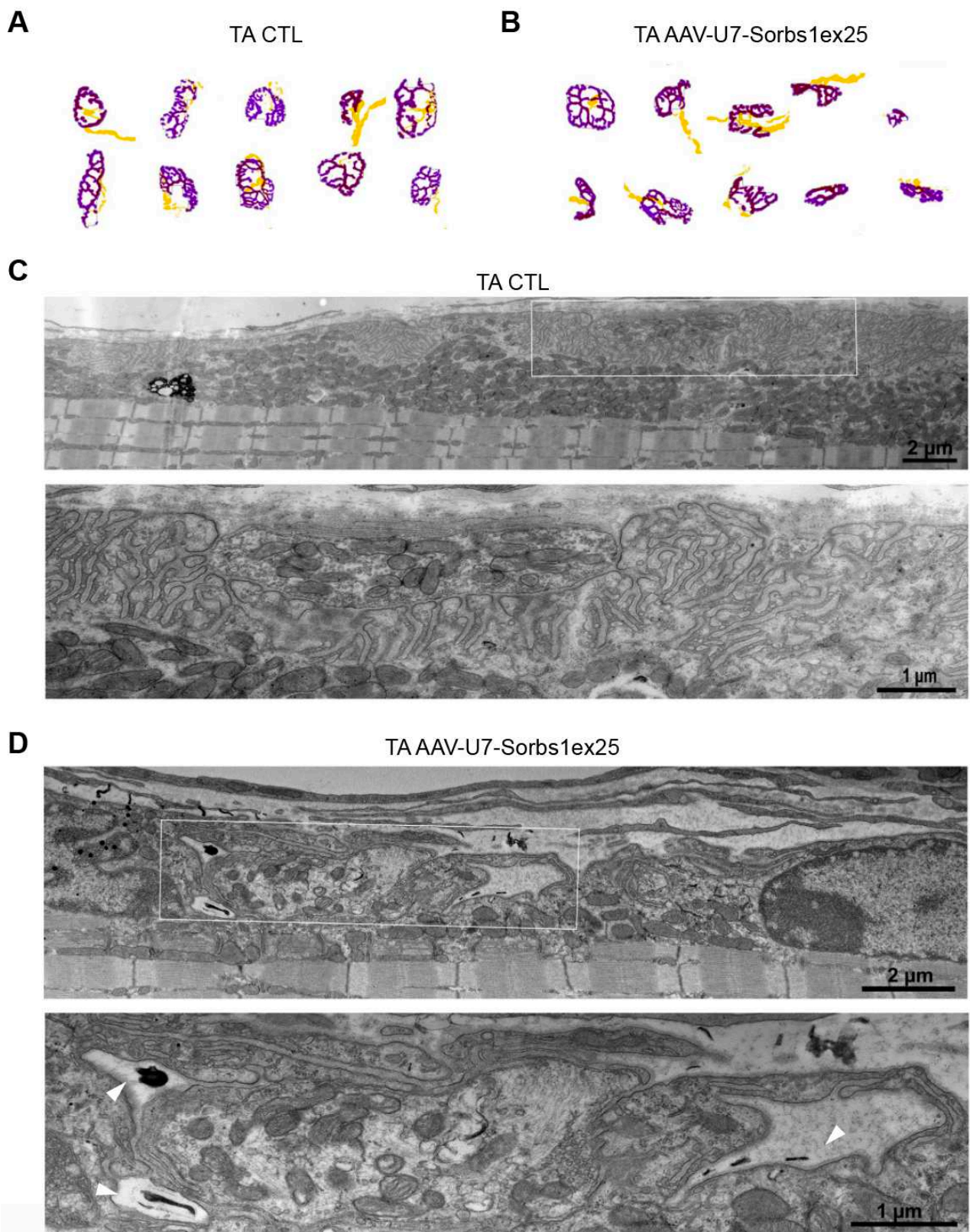


Figure 47 : Potentiels effets de l'exclusion de l'exon 25 de Sorbs1 au niveau des jonctions neuromusculaires chez la souris

(A-B) Exemple de patterns obtenus après immunofluorescence des compartiments pré et post-synaptiques (en jaune et violet respectivement) de JNM sur fibres isolées à partir du muscle tibial antérieur (TA) contrôle (A) ou injecté avec le vecteur AAV-U7-Sorbs1ex25 (B). (C-D) Images de microscopie électronique à transmission de JNM du muscle TA contrôle (C) ou injecté avec le vecteur (D). Les flèches blanches indiquent la présence de structures dégénératives au niveau de l'espace péri-synaptique.

- Quelles peuvent être les hypothèses mécanistiques reliant le défaut d'épissage de *SORBS1* au défaut d'aggrégation des récepteurs à l'acétylcholine ?

Les mécanismes par lesquels l'inclusion de l'exon 25 au niveau de la protéine SORBS1 influe sur la mise en place ou le maintien des clusters des AChR restent à ce jour uniquement hypothétiques. Dans leur étude sur les cellules murines C2C12, Hallock et al ont montré que le défaut de clustering des AChR était causé par la diminution d'expression de la protéine Sorbs1 (Hallock et al. 2016). Cependant, nous n'observons pas d'effet de l'exclusion de l'exon 25 sur la stabilité de la protéine SORBS1. La difficulté de trouver un anticorps spécifique ne nous permet pas en revanche d'identifier distinctement les différentes isoformes de la protéine et limite ainsi nos interprétations.

La protéine SORBS1, aussi connue sous le nom de CAP (c-Cbl associated protein) ou Ponsin, a été identifiée comme impliquée au niveau des sites d'adhésion focale ainsi que dans la régulation de l'organisation du cytosquelette en interagissant avec des protéines telles que la vinculine, la paxilline, l'afadine, la dynamine ou encore l'actine (Tosoni and Cestra 2009; Zhang et al. 2006; Kioka et al. 2002; Ribon et al. 1998). D'un point de vue structurale, la protéine SORBS1 est constituée de plusieurs domaines dont un domaine SoHo (Sorbin homology), un domaine coiled-coil (ou superhélice) et trois domaines SH3 (Src homology 3) et de manière intéressante, il a été montré que l'exon 25 de SORBS1 code le domaine coiled-coil de la protéine (UniProtKB – E9QNA7) (Tomasovic et al. 2012; Zhang et al. 2003). De nombreuses études ont montré l'implication des domaines coiled-coil dans l'oligomérisation ainsi que pouvant jouer le rôle "d'espaceurs moléculaires" permettant soit de séparer des domaines fonctionnels ou encore permettant la formation de grands complexes macromoléculaires. En effet, leurs propriétés physiques, telles que leur flexibilité, propre aux hélices α , ont des conséquences structurales et fonctionnelles importantes (Truebestein and Leonard 2016). Récemment, une étude a démontré que SORBS1 s'associait avec le complexe MuSK/Dok-7/Crk/CrkL pour permettre la formation des clusters des AChR (Hallock et al. 2016). La mise en place de l'organisation post-synaptique fait intervenir en aval du signal de transduction de la voie Agrin-LRP4-MuSK la Rapsyne, une protéine permettant l'ancrage des AChR et ainsi la formation de la JNM. Cependant, les mécanismes par lesquels le signal est transmis jusqu'à la Rapsyne sont encore méconnus (Xing et al. 2019). Il a par ailleurs été montré que la Rapsyne interagissait avec la queue cytoplasmique des AChR via son domaine coiled-coil (Xing et al. 2020; Lee et al. 2008b). Ainsi, plusieurs hypothèses sont envisageables quant au rôle de l'exon 25 de *SORBS1* dans la formation des clusters d'AChR.

On pourrait émettre l'hypothèse selon laquelle l'exon 25 serait nécessaire pour la localisation de la protéine SORBS1 ou encore pour son interaction avec des partenaires. Dans cet esprit, nous avons pu observer une certaine modification du marquage de la protéine

SORBS1, plus diffuse, dans les myotubes transfectés avec l'ASO ainsi que dans les myotubes dérivés d'hiPSC DM1 et DKO.

Il est envisageable que la présence ou l'absence de l'exon 25 puisse moduler la structure tridimensionnelle de la protéine, en particulier au niveau des structures adjacentes. Ainsi, en raison d'une part de la proximité du domaine coiled-coil, codé par l'exon 25, avec les domaines SH3 (**Figure 48-A**), et d'autre part de l'importance de ces domaines dans les interactions entre SORBS1 et ses partenaires, l'absence de l'exon 25 pourrait alors entraîner un changement conformationnel de la protéine qui modulerait l'activité d'interaction de SORBS1 avec des partenaires impliqués dans la voie de transduction conduisant au clustering des AChR.

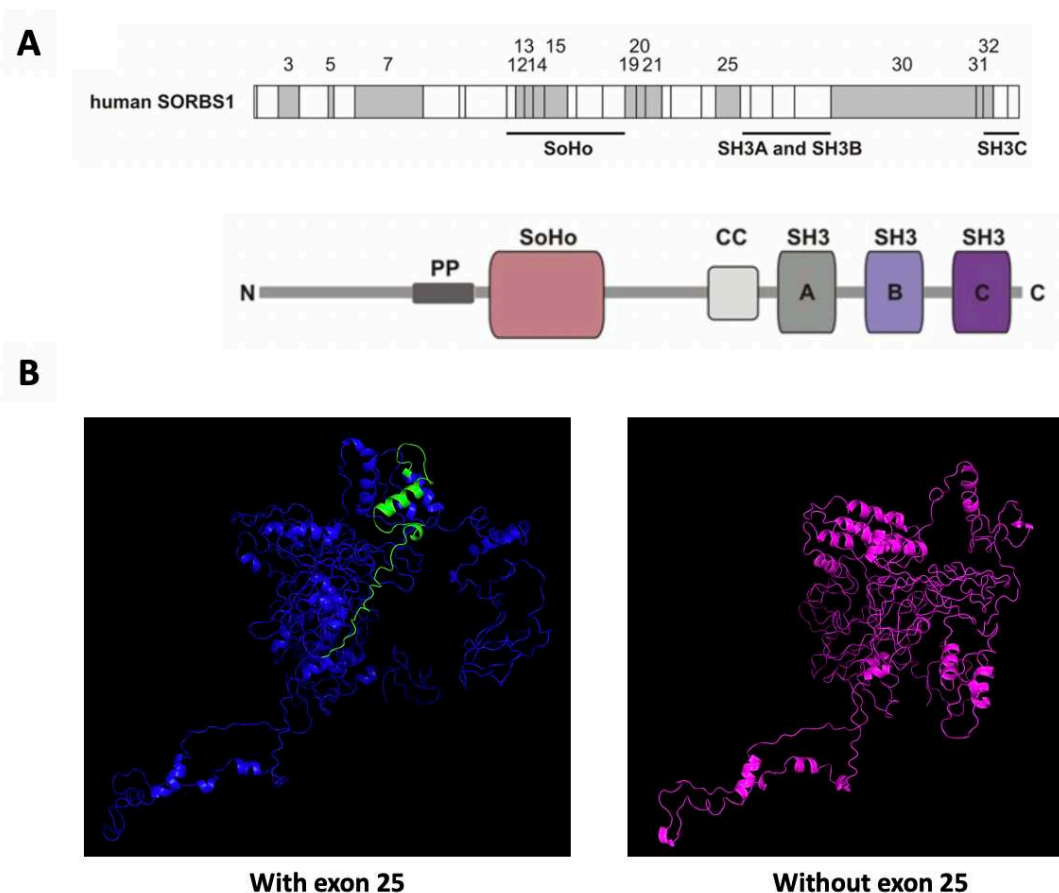


Figure 48 : Représentation schématique de SORBS1 et prédictions tridimensionnelles
(A) Représentation schématique du gène humain et de la protéine SORBS1. Les exons alternatifs sont indiqués en gris et les trois domaines SH3 sont répertoriés par ordre alphabétique (A, B et C). PP : région riche en Proline ; CC : domaine coiled-coil. *Adapté de (Tomasovic et al. 2012).*
(B) Modèles tridimensionnels prédictifs de la protéine SORBS1 avec (à gauche) ou sans (à droite) l'exon 25 (en vert). Ces prédictions structurales ont été obtenues avec le serveur I-TASSER à l'aide de la séquence protéique des isoformes de SORBS1 majoritairement retrouvées dans les myotubes dérivés d'hiPSC WT (référence UniProt Q9BX66-5) ou DM1 (Q9BX66-10), contenant ou non l'exon 25 respectivement.

Pour tester cette hypothèse, nous avons utilisé un outil de prédiction de structures protéiques tridimensionnelles nommé I-TASSER (Iterative Threading ASSEMBLY Refinement) développé par le Pr. Yang Zhang (Zhang 2008). À l'aide des analyses transcriptomiques effectuées dans notre étude, nous avons pu identifier le transcrit *SORBS1* majoritaire à la fois dans les myotubes dérivés d'hiPSC WT et d'hiPSC DM1, et ainsi déterminer l'isoforme protéique correspondante à chacun de ces deux transcrits. Selon les modèles obtenus, il semblerait que la présence ou l'absence de l'exon 25 induise un changement conformationnel de la partie C-terminale de la protéine (**Figure 48-B**).

Le premier domaine SH3 de la protéine étant situé juste en aval de l'exon 25 et ayant été montré comme interagissant avec des protéines du cytosquelette comme la vinculine (Mandai et al. 1999), l'absence de l'exon 25 pourrait alors modifier cette interaction aboutissant à un défaut de clusterisation des AChR. En effet, la vinculine fait partie, avec l'alpha-actinine ou encore la filamine, des protéines du cytosquelette associées aux clusters des AChR et qui pourraient contribuer à leur stabilisation ou à leur maintenance (Shadiack and Nitkin 1991; Daniels et al. 1990). De plus, il a également été montré que *SORBS1* jouait un rôle dans la signalisation des événements associés à l'adhésion cellulaire médiée par les intégrines puisque sa surexpression dans des cellules murines induit une forte augmentation du nombre des fibres de stress d'actine ainsi que du nombre de sites d'adhésion focale (Ribon et al. 1998). De manière intéressante, *SORBS1* a également été décrit comme interagissant avec la Filamine C (spécifique des cellules musculaires) au niveau des costamères (Zhang et al. 2007), points d'ancrage des disques Z avec le sarcolemme via le complexe dystroglycan/sarcoglycan et la matrice extra-cellulaire via les intégrines (Gontier et al. 2005). Ainsi, l'absence de l'exon 25 pourrait moduler les interactions entre *SORBS1* et des protéines associées à la matrice extra-cellulaire jouant un rôle important le développement ou le maintien des JNM (Singhal and Martin 2011).

Très récemment, une étude a démontré l'importance des domaines coiled-coil de la protéine SHTN1 (Shootin1) dans la modulation de son activité de liaison avec l'actine (Ergin and Zheng 2020). Bien que la protéine *SORBS1* se lie avec les filaments d'actine au niveau de sa partie N-terminale (Zhang et al. 2006), l'absence de l'exon 25 codant le domaine coiled-coil pourrait également modifier cette interaction, perturbant alors le recrutement, au niveau des AChR, des protéines du cytosquelette impliquées dans leur clusterisation (Xing et al. 2020). Pour tester cette dernière hypothèse, nous avons effectué un immunomarquage des filaments d'actine avec la phalloïdine dans des cellules musculaires humaines primaires transfectées ou non avec l'ASO, mais nous n'avons observé aucune anomalie au niveau de l'organisation du réseau d'actine (**Figure 49**).

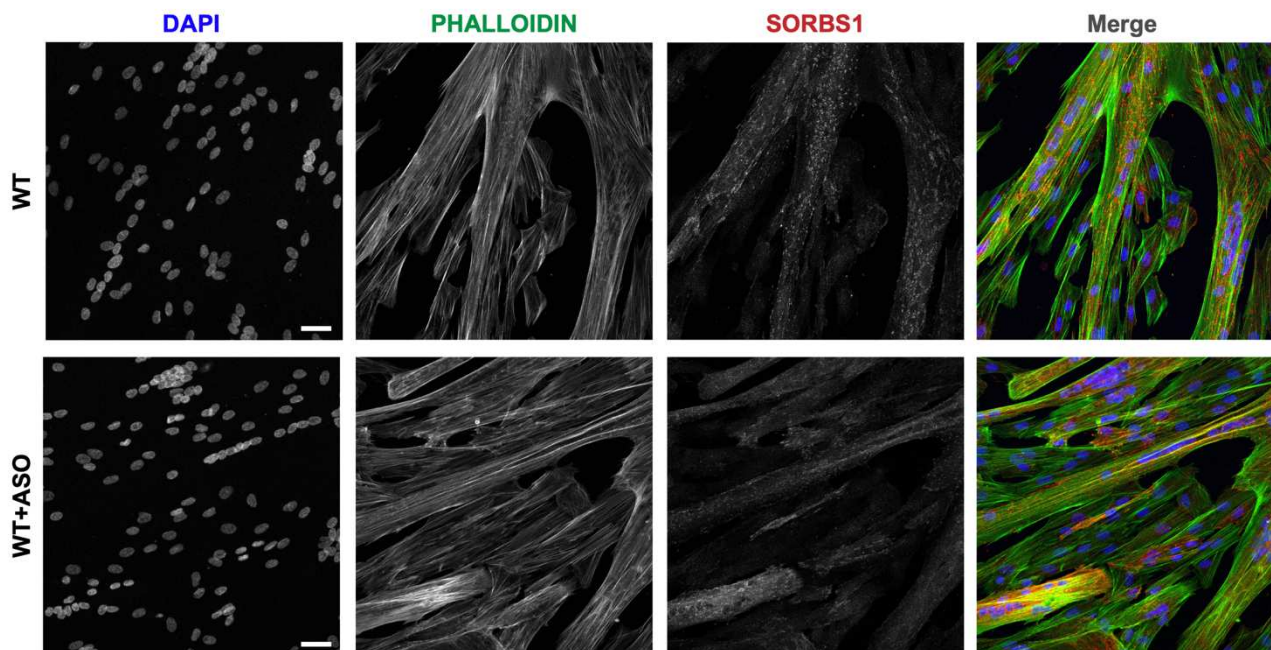


Figure 49 : L'exclusion de l'exon 25 de SORBS1 ne semble pas affecter l'organisation du réseau d'actine

Images d'immonofluorescence de SORBS1 et des filaments d'actine (Phalloïdine) dans des myotubes primaires stimulés à l'Aggrin et transfectés ou non avec l'ASO induisant l'exclusion de l'exon 25. Les noyaux sont représentés en bleu, SORBS1 en rouge et les filaments d'actine en vert. Barres d'échelle = 10 μ m.

Comme précisé en introduction de ce sous-chapitre, il ne s'agit ici que d'hypothèses mécanistiques. Des études complémentaires sont nécessaires pour conforter ou non ces hypothèses. En particulier, une étude précise, par cristallographie aux rayons X ou par résonance magnétique nucléaire, de la structure tridimensionnelle de ces deux isoformes de la protéine serait nécessaire pour valider ces premières observations.

D. Étude du rôle des protéines MBNL dans un autre type cellulaire : les motoneurones

En parallèle, j'ai participé, dans le cadre de travaux de recherche menés par une autre doctorante de l'équipe, à l'étude du rôle des protéines MBNL dans un autre type cellulaire : les motoneurones. En plus de confirmer la prédominance du paralogue MBNL2 dans les neurones (Charizanis et al. 2012) cette étude montre toutefois une compensation entre MBNL1 et MBNL2, suggérant ainsi un rôle non négligeable de MBNL1 dans les populations neuronales. En effet, les motoneurones dérivés d'hiPSC *MBNL1* KO, *MBNL2* KO ou encore DKO reproduisent le défaut d'arborisation neuritique décrit par le passé dans des motoneurones dérivés d'ESC DM1 (Marteyn et al. 2011). Grâce à la mise en place d'un système de co-culture entre des motoneurones dérivés d'hiPSC et des cellules musculaires primaires saines, un défaut de formation des clusters des AChR a pu être observé dans ces myotubes primaires en présence de motoneurones dérivés d'hiPSC DM1 et *MBNL* KO,

démontrant d'une part l'implication de la composante pré-synaptique dans la pathologie et d'autre part, le caractère MBNL-dépendant de celle-ci. Par des analyses de flux calcique, ces mêmes co-cultures ont également pu mettre en évidence une perturbation de l'homéostasie calcique musculaire liée à la déplétion des protéines MBNL dans les motoneurones.

Ces résultats sont en accord avec de nombreuses observations suggérant la contribution du compartiment pré-synaptique dans la physiopathologie de la DM1. En effet, plusieurs études ont montré une atteinte de la JNM chez les patients DM1 avec notamment l'observation de foci au niveau des noyaux des motoneurones et des noyaux sous-synaptiques ainsi qu'une co-localisation des transcrits DMPK et des AChR au niveau des JNM issues de biopsies de patients post-mortem (Wheeler et al. 2007a). Par le passé, bien avant l'identification de la DM1, des études post-mortem de patients diagnostiqués comme dystrophiques ont décrit des anomalies au niveau des plaques motrices (Macdermot 1961), lieu de la formation des clusters des AChR, ainsi qu'une réduction de la densité des vésicules présynaptiques (Allen et al. 1969). Récemment, des défauts de transmission synaptique ont été observés chez les patients DM1, avec en particulier des électromyogrammes sur fibres musculaires isolées anormaux (Bombelli et al. 2016). Toutefois, nos résultats ayant été obtenus essentiellement dans un nouveau modèle de JNM *in vitro*, une étude fonctionnelle plus approfondie serait nécessaire. En effet, l'utilisation de modèles cellulaires tridimensionnels plus complexes, tels que les tissus musculaires squelettiques artificiels (Maffioletti et al. 2018) ou encore les assemblöides cortico-moteurs (Andersen et al. 2020) développés par les équipes du Pr. Francesco Tedesco et du Dr. Sergiu Paşca respectivement, permettrait de valider nos résultats et de réaliser des analyses fonctionnelles plus poussées.

CONCLUSIONS ET PERSPECTIVES

Grâce aux cellules pluripotentes humaines ainsi qu'à la technologie CRISPR/Cas9, nous avons pu étudier les conséquences moléculaires et cellulaires de la perte de fonction des protéines MBNL au niveau de la génération de cellules musculaires squelettiques et révéler la temporalité d'action de ces protéines dans la myogenèse humaine. Notre étude nous a également permis de mieux comprendre la contribution des protéines MBNL dans la physiopathologie musculaire de la DM1, de souligner l'importance des phénomènes de compensation entre les deux paralogues MBNL1 et MBNL2 et de mettre en évidence, par une analyse transcriptomique, les voies moléculaires dérégulées dans la pathologie et dépendantes de ces protéines. Enfin, nous avons pu identifier le défaut d'épissage de l'exon 25 de *SORBS1* dans la DM1, régulé par les protéines MBNL et aboutissant à des anomalies structurales du compartiment post-synaptique musculaire (**Figure 50**).

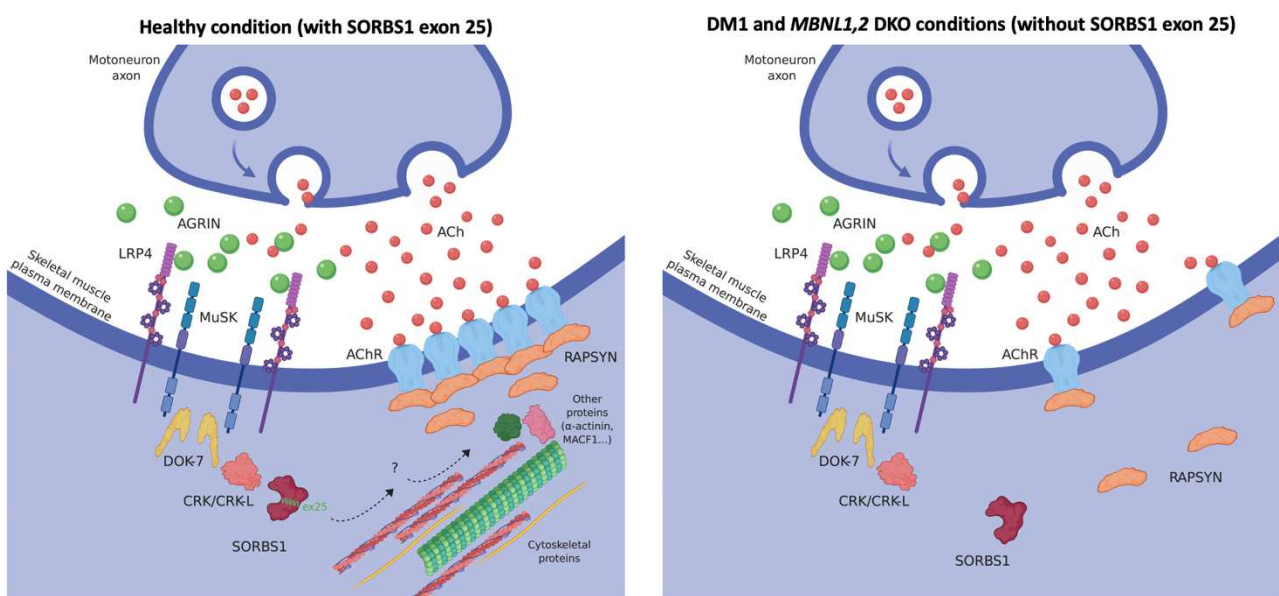


Figure 50 : Hypothèse du rôle de l'exon 25 de SORBS1 dans la formation des clusters des AChR

(Créé à partir de Biorender.com).

Par la suite, une étude plus précise des mécanismes moléculaires et des conséquences fonctionnelles du défaut d'épissage de l'exon 25 de *SORBS1* permettra de mieux comprendre sa contribution dans la physiopathologie de la DM1. En effet, une étude mécanistique sera nécessaire afin de mieux définir le rôle de SORBS1 et de son domaine coiled-coil dans la transduction du signal conduisant à la formation des clusters des AChR. Ces résultats s'ajouteraient ainsi aux nombreuses études tentant de déchiffrer les mécanismes moléculaires qui régulent le développement de la jonction neuromusculaire (Tintignac et al. 2015).

La génération des lignées hiPSC *MBNL* KO nous a ainsi permis d'étudier la contribution des protéines MBNL dans la physiopathologie musculaire humaine de la DM1, en nous abstrayant de tout autre mécanisme pathologique impliqué dans la maladie. Toutefois, en

modélisant la DM1 avec cette stratégie de knock-out, c'est-à-dire en déplaçant totalement les protéines MBNL, nous exacerbons d'une certaine manière la pathologie, dans laquelle ces protéines sont séquestrées au sein des *foci* intranucléaires. Si l'on se réfère à l'hypothèse mécanistique selon laquelle plus la taille des répétitions CTG est importante et plus l'haploinsuffisance en protéines MBNL le serait également ; le fait que les hiPSC DM1 utilisées dans notre étude proviennent de patients fortement atteints et contiennent un très grand nombre de répétitions CTG, pourrait expliquer pourquoi les myotubes dérivés d'hiPSC *MBNL* DKO récapitulent à ce point les défauts moléculaires et cellulaires observés dans les myotubes dérivés de ces hiPSC DM1.

Par ailleurs, une étude du rôle des protéines MBNL au niveau des motoneurones spinaux a également pu être réalisée, afin notamment de mieux comprendre la contribution du compartiment pré-synaptique dans la DM1. De la même manière, les hiPSC *MBNL* KO vont permettre par la suite d'étudier l'implication de ces protéines dans la spécification d'autres types cellulaires également touchés chez les patients DM1, en les différenciant par exemple en neurones corticaux, en astrocytes ou encore en cardiomyocytes. Cela permettrait d'identifier des mécanismes pathologiques communs, qui seraient régulées par les protéines MBNL indépendamment du tissu, ou au contraire d'étudier les voies moléculaires dérégulées spécifiquement dans chaque type cellulaire.

De manière plus générale, les lignées hiPSC *MBNL* KO que j'ai pu générer au cours de cette thèse constituent des outils cellulaires qui devraient permettre, à plus long terme, de faciliter l'identification de nouvelles stratégies thérapeutiques pour la DM1. Parce qu'il a été montré qu'une surexpression de *Mbnl1* était suffisante pour améliorer les atteintes musculaires et pour corriger certains défauts d'épissage chez les souris HSA^{LR} (Kanadia et al. 2006), de nombreuses études ont cherché jusqu'ici à restaurer un niveau d'expression fonctionnel des protéines MBNL, en réalisant notamment des criblages moléculaires afin d'identifier des petites molécules capables soit de libérer les protéines MBNL des répétitions CUG (Ketley et al. 2014), soit de réguler directement l'expression de ces protéines, en particulier MBNL1 (Zhang et al. 2017a; Chen et al. 2016). Dans ce contexte, les hiPSC *MBNL* KO représentent une alternative prometteuse et ouvrent la voie à la découverte de nouvelles pistes thérapeutiques, en mettant par exemple en place des criblages visant à identifier des composés cette fois-ci capables de contourner la perte de fonction des protéines MBNL.

* * *

Je vous remercie pour le temps consacré à la lecture de ce manuscrit !

REFERENCES BIBLIOGRAPHIQUES

- Adereth Y, Dammai V, Kose N, Li R, Hsu T. 2005. RNA-dependent integrin alpha3 protein localization regulated by the Muscleblind-like protein MLP1. *Nat Cell Biol* **7**: 1240–7. <http://www.ncbi.nlm.nih.gov/pubmed/16273094> (Accessed December 9, 2020).
- Allen DE, Johnson AG, Woolf AL. 1969. The intramuscular nerve endings in dystrophia myotonica--a biopsy study by vital staining and electron microscopy. *J Anat* **105**: 1–26. <http://www.ncbi.nlm.nih.gov/pubmed/4184783> (Accessed February 2, 2021).
- Alwazzan M, Newman E, Hamshere MG, Brook JD. 1999. Myotonic Dystrophy Is Associated with a Reduced Level of RNA from the DMWD Allele Adjacent to the Expanded Repeat. *Hum Mol Genet* **8**: 1491–1497. <https://academic.oup.com/hmg/article-lookup/doi/10.1093/hmg/8.8.1491> (Accessed November 26, 2020).
- Amoasii L, Hildyard JCW, Li H, Sanchez-Ortiz E, Mireault A, Caballero D, Harron R, Stathopoulou T-R, Massey C, Shelton JM, et al. 2018. Gene editing restores dystrophin expression in a canine model of Duchenne muscular dystrophy. *Science* eaau1549. <http://www.ncbi.nlm.nih.gov/pubmed/30166439> (Accessed September 11, 2018).
- Andersen G, Hedermann G, Witting N, Duno M, Andersen H, Vissing J. 2017. The antimyotonic effect of lamotrigine in non-dystrophic myotonias: a double-blind randomized study. *Brain* **140**: 2295–2305. <http://www.ncbi.nlm.nih.gov/pubmed/29050397> (Accessed December 3, 2020).
- Andersen J, Revah O, Miura Y, Thom N, Amin ND, Kelley KW, Singh M, Chen X, Thete MV, Walczak EM, et al. 2020. Generation of Functional Human 3D Cortico-Motor Assembloids. *Cell* **183**: 1913–1929.e26. <http://www.ncbi.nlm.nih.gov/pubmed/33333020> (Accessed February 17, 2021).
- André LM, Ausems CRM, Wansink DG, Wieringa B. 2018. Abnormalities in Skeletal Muscle Myogenesis, Growth, and Regeneration in Myotonic Dystrophy. *Front Neurol* **9**. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5985300/> (Accessed November 23, 2020).
- Anzalone A V., Koblan LW, Liu DR. 2020. Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. *Nat Biotechnol* **38**: 824–844. <http://www.nature.com/articles/s41587-020-0561-9> (Accessed November 23, 2020).
- Arandel L, Polay Espinoza M, Matloka M, Bazinet A, De Dea Diniz D, Naouar N, Rau F, Jollet A, Edom-Vovard F, Mamchaoui K, et al. 2017. Immortalized human myotonic dystrophy muscle cell lines to assess therapeutic compounds. *Dis Model Mech* **10**: 487–497. <http://dmm.biologists.org/lookup/doi/10.1242/dmm.027367>.
- Artero R, Prokop A, Paricio N, Begemann G, Pueyo I, Mlodzik M, Perez-Alonso M, Baylies MK. 1998. The muscleblind gene participates in the organization of Z-bands and epidermal attachments of Drosophila muscles and is regulated by Dmef2. *Dev Biol* **195**: 131–43. <http://www.ncbi.nlm.nih.gov/pubmed/9520330> (Accessed December 4, 2020).
- Ashwal-Fluss R, Meyer M, Pamudurti NR, Ivanov A, Bartok O, Hanan M, Evantal N, Memczak S, Rajewsky N, Kadener S. 2014. circRNA biogenesis competes with pre-mRNA splicing. *Mol Cell* **56**: 55–66. <http://www.ncbi.nlm.nih.gov/pubmed/25242144> (Accessed December 9, 2020).
- Avery S, Hirst AJ, Baker D, Lim CY, Alagaratnam S, Skotheim RI, Lothe RA, Pera MF, Colman A, Robson P, et al. 2013. BCL-XL Mediates the Strong Selective Advantage of a 20q11.21 Amplification Commonly Found in Human Embryonic Stem Cell Cultures. *Stem Cell Reports* **1**: 379. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3841249/> (Accessed December 16, 2020).
- Baralle FE, Giudice J. 2017. Alternative splicing as a regulator of development and tissue identity. *Nat Rev Mol Cell Biol* **18**: 437–451. <http://www.nature.com/articles/nrm.2017.27> (Accessed March 29, 2020).
- Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science (80-)* **315**: 1709–12. <http://www.ncbi.nlm.nih.gov/pubmed/17379808> (Accessed November 23, 2020).
- Barreau C, Paillard L, Méreau A, Osborne H. 2006. Mammalian CELF/Bruno-like RNA-binding proteins: molecular characteristics and biological functions. *Biochimie* **88**.

<https://pubmed.ncbi.nlm.nih.gov/16480813/> (Accessed December 9, 2020).

- Bassett AR. 2017. Editing the genome of hiPSC with CRISPR/Cas9: disease models. *Mamm genome* **28**: 348–364. <http://www.ncbi.nlm.nih.gov/pubmed/28303292> (Accessed November 23, 2020).
- Bassez G, Audureau E, Hogrel J-Y, Arrouasse R, Baghdoyan S, Bhugaloo H, Gourlay-Chu M-L, Le Corvoisier P, Peschanski M. 2018. Improved mobility with metformin in patients with myotonic dystrophy type 1: a randomized controlled trial. *Brain*. <https://academic.oup.com/brain/advance-article/doi/10.1093/brain/awy231/5086409> (Accessed September 3, 2018).
- Batra R, Charizanis K, Manchanda M, Mohan A, Li M, Finn DJ, Goodwin M, Zhang C, Sobczak K, Thornton CA, et al. 2014. Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-mediated disease. *Mol Cell* **56**: 311–322. <http://dx.doi.org/10.1016/j.molcel.2014.08.027>.
- Batra R, Nelles DA, Pirie E, Blue SM, Marina RJ, Wang H, Chaim IA, Thomas JD, Zhang N, Nguyen V, et al. 2017. Elimination of Toxic Microsatellite Repeat Expansion RNA by RNA-Targeting Cas9. *Cell* **170**: 899–906.e10. <http://dx.doi.org/10.1016/j.cell.2017.07.010> (Accessed July 6, 2018).
- Begemann G, Paricio N, Artero R, Kiss I, Perez-Alonso M, Mlodzik M. 1997. muscleblind, a gene required for photoreceptor differentiation in Drosophila, encodes novel nuclear Cys3His-type zinc-finger-containing proteins. *Development* **124**: 4321–4331. <https://pubmed.ncbi.nlm.nih.gov/9334280/> (Accessed December 4, 2020).
- Beghini DG, Iwao Horita S, Cascabulho CM, Anastácio Alves L, Henriques-Pons A. 2020. Induced Pluripotent Stem Cells: Hope in the Treatment of Diseases, including Muscular Dystrophies. *Int J Mol Sci* **21**. <http://www.ncbi.nlm.nih.gov/pubmed/32751747> (Accessed November 22, 2020).
- Ben Jehuda R, Shemer Y, Binah O. 2018. Genome Editing in Induced Pluripotent Stem Cells using CRISPR/Cas9. *Stem Cell Rev Reports* **14**: 323–336. <http://link.springer.com/10.1007/s12015-018-9811-3> (Accessed November 23, 2020).
- Bento G, Shafigullina AK, Rizvanov AA, Sardão VA, Macedo MP, Oliveira PJ. 2020. Urine-Derived Stem Cells: Applications in Regenerative and Predictive Medicine. *Cells* **9**. <http://www.ncbi.nlm.nih.gov/pubmed/32121221> (Accessed November 22, 2020).
- Bird TD. 2020. *Myotonic Dystrophy Type 1*. 1999th–2020th ed. ed. Et al. Adam MP, Ardinger HH, Pagon RA. University of Washington, Seattle, Seattle (WA) <http://www.ncbi.nlm.nih.gov/pubmed/20301344> (Accessed November 23, 2020).
- Bombelli F, Lispi L, Porrini SC, Giacanelli M, Terracciano C, Massa R, Petrucci A. 2016. Neuromuscular transmission abnormalities in myotonic dystrophy type 1: A neurophysiological study. *Clin Neurol Neurosurg* **150**: 84–88. <http://www.ncbi.nlm.nih.gov/pubmed/27611986> (Accessed January 21, 2021).
- Botta A, Malena A, Tibaldi E, Rocchi L, Loro E, Pena E, Cenci L, Ambrosi E, Bellocchi MC, Pagano MA, et al. 2013. MBNL142 and MBNL143 gene isoforms, overexpressed in DM1-patient muscle, encode for nuclear proteins interacting with Src family kinases. *Cell Death Dis* **4**: e770. <http://www.ncbi.nlm.nih.gov/pubmed/23949219> (Accessed December 10, 2020).
- Braz SO, Acquire J, Gourdon G, Gomes-Pereira M. 2018. Of Mice and Men: Advances in the Understanding of Neuromuscular Aspects of Myotonic Dystrophy. *Front Neurol* **9**: 519. <https://www.frontiersin.org/article/10.3389/fneur.2018.00519/full> (Accessed July 25, 2018).
- Brinegar AE, Xia Z, Loehr JA, Li W, Rodney GG, Cooper TA. 2017. Extensive alternative splicing transitions during postnatal skeletal muscle development are required for calcium handling functions. *Elife* **6**. <http://www.ncbi.nlm.nih.gov/pubmed/28826478> (Accessed March 29, 2020).
- Brook JD, Mccurrach ME, Harley HG, Buckler AJ, Church D, Aburatani H, Hunter K, Stanton VP, Thirion J-P, Hudson T, et al. 1992. Molecular Basis of Myotonic Dystrophy: Expansion of a Trinucleotide (CTG) Repeat at the 3' End of a Transcript Encoding a Protein Kinase Family Member. *Cell* **68**: 799–808. <https://ac-els-cdn-com.gate2.inist.fr/0092867492901545/1-s2.0-0092867492901545->

main.pdf?_tid=2610cd5e-6d4b-4b0f-9e07-

37ef57f9e87d&acdnat=1524067745_836f90cc2a10d490db94052fee90f544 (Accessed April 18, 2018).

- Buxton J, Shelbourne P, Davies J, Jones C, Tongeren T Van, Aslanidis C, de Jong P, Jansen G, Anvret M, Riley B, et al. 1992. Detection of an unstable fragment of DNA specific to individuals with myotonic dystrophy. *Nature* **355**: 547–548. <http://www.nature.com/articles/355547a0> (Accessed November 23, 2020).
- Cahan P, Daley GQ. 2013. Origins and implications of pluripotent stem cell variability and heterogeneity. *Nat Rev Mol Cell Biol* **14**: 357–68. <http://www.ncbi.nlm.nih.gov/pubmed/23673969> (Accessed January 13, 2021).
- Caron L, Kher D, Lee KL, McKernan R, Dumevska B, Hidalgo A, Li J, Yang H, Main H, Ferri G, et al. 2016. A Human Pluripotent Stem Cell Model of Facioscapulohumeral Muscular Dystrophy-Affected Skeletal Muscles. *Stem Cells Transl Med* **5**: 1145–1161. <http://doi.wiley.com/10.5966/sctm.2015-0224>.
- Cass D, Hotchko R, Barber P, Jones K, Gates DP, Berglund JA. 2011. The four Zn fingers of MBNL1 provide a flexible platform for recognition of its RNA binding elements. *BMC Mol Biol* **12**: 20. <http://www.ncbi.nlm.nih.gov/pubmed/21548961> (Accessed December 8, 2020).
- Chal J, Al Tanoury Z, Hestin M, Gobert B, Aivio S, Hick A, Cherrier T, Nesmith AP, Parker KK, Pourquié O. 2016. Generation of human muscle fibers and satellite-like cells from human pluripotent stem cells in vitro. *Nat Protoc* **11**: 1833–1850. <http://www.ncbi.nlm.nih.gov/pubmed/27583644> (Accessed November 23, 2020).
- Chal J, Pourquié O. 2017. Making muscle: skeletal myogenesis in vivo and in vitro. *Development* **144**: 2104–2122. <http://www.ncbi.nlm.nih.gov/pubmed/28634270> (Accessed October 22, 2020).
- Chamberlain CM, Ranum LPW. 2012. Mouse model of muscleblind-like 1 overexpression: Skeletal muscle effects and therapeutic promise. *Hum Mol Genet* **21**: 4645–4654.
- Charizanis K, Lee K-Y, Batra R, Goodwin M, Zhang C, Yuan Y, Shiue L, Cline M, Scotti MM, Xia G, et al. 2012. Muscleblind-like 2-Mediated Alternative Splicing in the Developing Brain and Dysregulation in Myotonic Dystrophy. *Neuron* **75**: 437–450. <http://dx.doi.org/10.1016/j.neuron.2012.05.029> (Accessed April 13, 2018).
- Chau A, Kalsotra A. 2015. Developmental insights into the pathology of and therapeutic strategies for DM1: Back to the basics. *Dev Dyn* **244**: 377–390. <http://doi.wiley.com/10.1002/dvdy.24240> (Accessed November 25, 2020).
- Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li G-W, Park J, Blackburn EH, Weissman JS, Qi LS, et al. 2013. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* **155**: 1479–91. <http://www.ncbi.nlm.nih.gov/pubmed/24360272> (Accessed November 23, 2020).
- Chen G, Masuda A, Konishi H, Ohkawara B, Ito M, Kinoshita M, Kiyama H, Matsuura T, Ohno K. 2016. Phenylbutazone induces expression of MBNL1 and suppresses formation of MBNL1-CUG RNA foci in a mouse model of myotonic dystrophy. *Sci Rep* **6**: 25317. <http://www.ncbi.nlm.nih.gov/pubmed/27126921> (Accessed December 3, 2020).
- Chen JS, Dagdas YS, Kleinstiver BP, Welch MM, Sousa AA, Harrington LB, Sternberg SH, Joung JK, Yildiz A, Doudna JA. 2017. Enhanced proofreading governs CRISPR–Cas9 targeting accuracy. *Nature* **550**: 407–410. <http://www.nature.com/articles/nature24268> (Accessed January 28, 2021).
- Chen KG, Mallon BS, McKay RDG, Robey PG. 2014. Human pluripotent stem cell culture: considerations for maintenance, expansion, and therapeutics. *Cell Stem Cell* **14**: 13–26. <http://www.ncbi.nlm.nih.gov/pubmed/24388173> (Accessed November 22, 2020).
- Choi J, Dixon DM, Dansithong W, Abdallah WF, Roos KP, Jordan MC, Trac B, Lee HS, Comai L, Reddy S. 2016. Muscleblind-like 3 deficit results in a spectrum of age-associated pathologies observed

- in myotonic dystrophy. *Sci Rep* **6**. <http://www.nature.com.gate2.inist.fr/articles/srep30999.pdf> (Accessed May 29, 2018).
- Choi J, Personius KE, Difranco M, Dansithong W, Yu C, Srivastava S, Dixon DM, Bhatt DB, Comai L, Vergara JL, et al. 2015. Muscleblind-Like 1 and Muscleblind-Like 3 Depletion Synergistically Enhances Myotonia by Altering Clc-1 RNA Translation. *EBioMedicine* **2**: 1034–1047. [https://www.ebiomedicine.com/article/S2352-3964\(15\)30080-3/pdf](https://www.ebiomedicine.com/article/S2352-3964(15)30080-3/pdf) (Accessed May 29, 2018).
- Chuah M, Chai YC, Dastidar S, VandenDriessche T. 2019. Gene Therapy and Gene Editing for Myotonic Dystrophy. In *Muscle Gene Therapy*, pp. 525–548, Springer International Publishing, Cham http://link.springer.com/10.1007/978-3-030-03095-7_30 (Accessed November 25, 2020).
- Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, et al. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science (80-)* **339**: 819–23. <http://www.ncbi.nlm.nih.gov/pubmed/23287718> (Accessed November 23, 2020).
- Cox DC, Guan X, Xia Z, Cooper TA. 2020. Increased nuclear but not cytoplasmic activities of CELF1 protein leads to muscle wasting. *Hum Mol Genet* **29**: 1729–1744. <http://www.ncbi.nlm.nih.gov/pubmed/32412585> (Accessed November 29, 2020).
- Daguenet E, Dujardin G, Valcárcel J. 2015. The pathogenicity of splicing defects: mechanistic insights into pre- mRNA processing inform novel therapeutic approaches. *EMBO Rep* **16**: 1640–1655. <https://onlinelibrary.wiley.com/doi/abs/10.15252/embr.201541116> (Accessed November 27, 2020).
- Daniels MP, Krikorian JG, Olek AJ, Bloch RJ. 1990. Association of cytoskeletal proteins with newly formed acetylcholine receptor aggregates induced by embryonic brain extract. *Exp Cell Res* **186**: 99–108. <http://www.ncbi.nlm.nih.gov/pubmed/2105221> (Accessed February 3, 2021).
- Dansithong W, Jog SP, Paul S, Mohammadzadeh R, Tring S, Kwok Y, Fry RC, Marjoram P, Comai L, Reddy S. 2011. RNA steady-state defects in myotonic dystrophy are linked to nuclear exclusion of SHARP. *EMBO Rep* **12**: 735–42. <http://www.ncbi.nlm.nih.gov/pubmed/21637295> (Accessed November 28, 2020).
- Dansithong W, Paul S, Comai L, Reddy S. 2005. MBNL1 Is the Primary Determinant of Focus Formation and Aberrant Insulin Receptor Splicing in DM1. *J Biol Chem* **280**: 5773–5780. <https://www-jbc-org.proxy.insermbiblio.inist.fr/content/280/7/5773> (Accessed May 6, 2020).
- Darabi R, Perlingeiro RCR. 2016. Derivation of skeletal myogenic precursors from human pluripotent stem cells using conditional expression of PAX7. In *Methods in Molecular Biology*, Vol. 1357 of, pp. 423–439, Humana Press Inc.
- Dasgupta T, Ladd AN. 2012. The importance of CELF control: molecular and biological roles of the CUG-BP, Elav-like family of RNA-binding proteins. *Wiley Interdiscip Rev RNA* **3**: 104–21. <http://www.ncbi.nlm.nih.gov/pubmed/22180311> (Accessed November 28, 2020).
- Dastidar S, Ardui S, Singh K, Majumdar D, Nair N, Fu Y, Reyon D, Samara E, Gerli MFM, Klein AF, et al. 2018. Efficient CRISPR/Cas9-mediated editing of trinucleotide repeat expansion in myotonic dystrophy patient-derived iPS and myogenic cells. *Nucleic Acids Res* **46**: 8275–8298. <http://www.ncbi.nlm.nih.gov/pubmed/29947794> (Accessed November 23, 2020).
- Davis BM, McCurrach ME, Taneja KL, Singer RH, Housman DE. 1997. Expansion of a CUG trinucleotide repeat in the 3' untranslated region of myotonic dystrophy protein kinase transcripts results in nuclear retention of transcripts. *Proc Natl Acad Sci U S A* **94**: 7388–93. <http://www.ncbi.nlm.nih.gov/pubmed/9207101> (Accessed November 27, 2020).
- De Antonio M, Dogan C, Hamroun D, Mati M, Zerrouki S, Eymard B, Katsahian S, Bassez G. 2016. Unravelling the myotonic dystrophy type 1 clinical spectrum: A systematic registry-based study with implications for disease classification. *Rev Neurol (Paris)* **172**: 572–580. <https://linkinghub.elsevier.com/retrieve/pii/S0035378716302053> (Accessed November 24, 2020).
- De Masi C, Spitalieri P, Murdocca M, Novelli G, Sangiuolo F. 2020. Application of CRISPR/Cas9 to

- human-induced pluripotent stem cells: from gene editing to drug discovery. *Hum Genomics* **14**: 25. <http://www.ncbi.nlm.nih.gov/pubmed/32591003> (Accessed November 23, 2020).
- De Temmerman N, Seneca S, Van Steirteghem A, Haentjens P, Van der Elst J, Liebaers I, Sermon KD. 2008. CTG repeat instability in a human embryonic stem cell line carrying the myotonic dystrophy type 1 mutation. *Mol Hum Reprod* **14**: 405–412. <https://academic.oup.com/molehr/article-lookup/doi/10.1093/molehr/gan034> (Accessed November 23, 2020).
- De Temmerman N, Sermon K, Seneca S, De Rycke M, Hilven P, Lissens W, Van Steirteghem A, Liebaers I. 2004. Intergenerational instability of the expanded CTG repeat in the DMPK gene: studies in human gametes and preimplantation embryos. *Am J Hum Genet* **75**: 325–9. <http://www.ncbi.nlm.nih.gov/pubmed/15185171> (Accessed November 25, 2020).
- Dean NL, Tan SL, Ao A. 2006. Instability in the transmission of the myotonic dystrophy CTG repeat in human oocytes and preimplantation embryos. *Fertil Steril* **86**: 98–105. <http://www.ncbi.nlm.nih.gov/pubmed/16716318> (Accessed November 25, 2020).
- Dellavalle A, Maroli G, Covarello D, Azzoni E, Innocenzi A, Perani L, Antonini S, Sambasivan R, Brunelli S, Tajbakhsh S, et al. 2011. Pericytes resident in postnatal skeletal muscle differentiate into muscle fibres and generate satellite cells. *Nat Commun* **2**: 499. <http://www.nature.com/articles/ncomms1508> (Accessed November 23, 2020).
- deLorimier E, Coonrod LA, Copperman J, Taber A, Reister EE, Sharma K, Todd PK, Guenza MG, Berglund JA. 2014. Modifications to toxic CUG RNAs induce structural stability, rescue mis-splicing in a myotonic dystrophy cell model and reduce toxicity in a myotonic dystrophy zebrafish model. *Nucleic Acids Res* **42**: 12768–78. <http://www.ncbi.nlm.nih.gov/pubmed/25303993> (Accessed December 9, 2020).
- Denis JA, Gauthier M, Rachdi L, Aubert S, Giraud-Triboulet K, Poydenot P, Benchoua A, Champon B, Maury Y, Baldeschi C, et al. 2013. mTOR-dependent proliferation defect in human ES-derived neural stem cells affected by myotonic dystrophy type 1. *J Cell Sci* **126**: 1763–1772. <http://jcs.biologists.org.gate2.inist.fr/content/joces/126/8/1763.full.pdf?with-ds=yes> (Accessed June 12, 2018).
- Diss G, Ascencio D, DeLuna A, Landry CR. 2014. Molecular mechanisms of paralogous compensation and the robustness of cellular networks. *J Exp Zool Part B Mol Dev Evol* **322**: 488–499. <http://doi.wiley.com/10.1002/jez.b.22555> (Accessed February 17, 2021).
- Dogan C, De Antonio M, Hamroun D, Varet H, Fabbro M, Rougier F, Amarof K, Arne Bes M-C, Bedat-Millet A-L, Behin A, et al. 2016. Gender as a Modifying Factor Influencing Myotonic Dystrophy Type 1 Phenotype Severity and Mortality: A Nationwide Multiple Databases Cross-Sectional Observational Study. *PLoS One* **11**: e0148264. <http://www.ncbi.nlm.nih.gov/pubmed/26849574> (Accessed November 24, 2020).
- Doudna JA, Charpentier E. 2014. The new frontier of genome engineering with CRISPR-Cas9. *Science (80-)* **346**: 1258096. <https://www.sciencemag.org/lookup/doi/10.1126/science.1258096> (Accessed November 23, 2020).
- Du H, Cline MS, Osborne RJ, Tuttle DL, Clark TA, Donohue JP, Hall MP, Shiue L, Swanson MS, Thornton CA, et al. 2010. Aberrant alternative splicing and extracellular matrix gene expression in mouse models of myotonic dystrophy. *Nat Struct Mol Biol* **17**: 187–193. <http://www.nature.com/doi/10.1038/nsmb.1720>.
- Du J, Campau E, Soragni E, Jespersen C, Gottesfeld JM. 2013. Length-dependent CTG{middle dot}CAG triplet-repeat expansion in myotonic dystrophy patient-derived induced pluripotent stem cells. *Hum Mol Genet* **22**: 5276–5287. <https://academic.oup.com/hmg/article-lookup/doi/10.1093/hmg/ddt386> (Accessed November 23, 2020).
- Dumevska B, Main H, McKernan R, Goel D, Schmidt U. 2016. Derivation of DM1 affected human embryonic stem cell line Genea067. *Stem Cell Res* **16**: 437–439.

- <https://www.sciencedirect.com/science/article/pii/S1873506116000568> (Accessed November 22, 2020).
- Eiselleova L, Peterkova I, Neradil J, Slaninova I, Hampl A, Dvorak P. 2008. Comparative study of mouse and human feeder cells for human embryonic stem cells. *Int J Dev Biol* **52**: 353–363. <http://www.ncbi.nlm.nih.gov/pubmed/18415935> (Accessed November 22, 2020).
- El-Brolosy MA, Stainier DYR. 2017. Genetic compensation: A phenomenon in search of mechanisms. *PLoS Genet* **13**: e1006780. <http://www.ncbi.nlm.nih.gov/pubmed/28704371> (Accessed February 16, 2021).
- Elkins JM, Amos A, Niesen FH, Pike ACW, Fedorov O, Knapp S. 2009. Structure of dystrophin myotonia protein kinase. *Protein Sci* **18**: 782–91. <http://www.ncbi.nlm.nih.gov/pubmed/19309729> (Accessed November 25, 2020).
- Ergin V, Zheng S. 2020. Putative Coiled-Coil Domain-Dependent Autoinhibition and Alternative Splicing Determine SHTN1's Actin-Binding Activity. *J Mol Biol* **432**: 4154–4166. <http://www.ncbi.nlm.nih.gov/pubmed/32371045> (Accessed February 12, 2021).
- Evans MJ, Kaufman MH. 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**: 154–156. <http://www.nature.com.gate2.inist.fr/articles/292154a0.pdf> (Accessed June 12, 2018).
- Fardaei M, Larkin K, Brook JD, Hamshere MG. 2001. In vivo co-localisation of MBNL protein with DMPK expanded-repeat transcripts. *Nucleic Acids Res* **29**: 2766–2771. <https://www.ncbi.nlm.nih.gov.gate2.inist.fr/pmc/articles/PMC55763/pdf/gke393.pdf> (Accessed April 18, 2018).
- Fardaei M, Rogers MT, Thorpe HM, Larkin K, Hamshere MG, Harper PS, Brook JD. 2002. Three proteins, MBNL, MBLL and MBXL, co-localize in vivo with nuclear foci of expanded-repeat transcripts in DM1 and DM2 cells. *Hum Mol Genet* **11**: 805–814. https://watermark.silverchair.com/ddf079.pdf?token=AQECAHi208BE49Ooan9kkhW_Ercy7Dm3ZL_9Cf3qfKAc485ysgAAAaUwggGhBgkqhkiG9w0BBwagggGSMIIBjgIBADCCAYcGCSqGSIb3DQEHATAeBglghkgBZQMEAS4wEQQM0KKYFTmScxhh4eL5AgEQgIIBWAOlh0dGgmIbQhKRTGVYsxhudPD-ho34CsXKuFATtw3wwiFm (Accessed April 18, 2018).
- Filareto A, Rinaldi F, Arpke RW, Darabi R, Belanto JJ, Toso EA, Miller AZ, Ervasti JM, McIvor RS, Kyba M, et al. 2015. Pax3-induced expansion enables the genetic correction of dystrophic satellite cells. *Skelet Muscle* **5**: 36. <http://www.ncbi.nlm.nih.gov/pubmed/26504514> (Accessed November 23, 2020).
- Flower M, Lomeikaite V, Ciosi M, Cumming S, Morales F, Lo K, Hensman Moss D, Jones L, Holmans P, Monckton DG, et al. 2019. MSH3 modifies somatic instability and disease severity in Huntington's and myotonic dystrophy type 1. *Brain* **142**: 1876–1886. <https://academic.oup.com/brain/article/142/7/1876/5520687> (Accessed November 25, 2020).
- Foisy L, Dong L, Savouret C, Hubert L, te Riele H, Junien C, Gourdon G. 2006. Msh3 is a limiting factor in the formation of intergenerational CTG expansions in DM1 transgenic mice. *Hum Genet* **119**: 520–6. <http://www.ncbi.nlm.nih.gov/pubmed/16552576> (Accessed November 25, 2020).
- Fougerousse F, Durand M, Lopez S, Suel L, Demignon J, Thornton C, Ozaki H, Kawakami K, Barbet P, Beckmann JS, et al. 2002. Six and Eya expression during human somitogenesis and MyoD gene family activation. *J Muscle Res Cell Motil* **23**: 255–64. <http://www.ncbi.nlm.nih.gov/pubmed/12500905> (Accessed January 12, 2021).
- Franck S, Barbé L, Ardui S, De Vlaeminck Y, Allemeersch J, Dziedzicka D, Spits C, Vanroye F, Hilven P, Duqué G, et al. 2020. MSH2 knock-down shows CTG repeat stability and concomitant upstream demethylation at the DMPK locus in myotonic dystrophy type 1 human embryonic stem cells. *Hum Mol Genet*. <http://www.ncbi.nlm.nih.gov/pubmed/33242073> (Accessed December 1, 2020).
- François V, Klein AF, Beley C, Jollet A, Lemercier C, Garcia L, Furling D. 2011. Selective silencing of mutated mRNAs in DM1 by using modified hU7-snrRNAs. *Nat Struct Mol Biol* **18**: 85–7.

- <http://www.ncbi.nlm.nih.gov/pubmed/21186365> (Accessed December 3, 2020).
- Frisch R, Singleton KR, Moses PA, Gonzalez IL, Carango P, Marks HG, Funanage VL. 2001. Effect of triplet repeat expansion on chromatin structure and expression of DMPK and neighboring genes, SIX5 and DMWD, in myotonic dystrophy. *Mol Genet Metab* **74**: 281–91. <http://www.ncbi.nlm.nih.gov/pubmed/11592825> (Accessed November 26, 2020).
- Fu Y, Ramisetty SR, Hussain N, Baranger AM. 2012. MBNL1-RNA recognition: contributions of MBNL1 sequence and RNA conformation. *ChemBiochem* **13**: 112–9. <http://www.ncbi.nlm.nih.gov/pubmed/22106026> (Accessed December 8, 2020).
- Fu YH, Pizzuti A, Fenwick RG, King J, Rajnarayan S, Dunne PW, Dubel J, Nasser GA, Ashizawa T, de Jong P. 1992. An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science (80-)* **255**: 1256–8. <http://www.ncbi.nlm.nih.gov/pubmed/1546326> (Accessed November 23, 2020).
- Fugier C, Klein A, Hammer C, Vassilopoulos S, Ivarsson Y, Toussaint A, Tosch V, Vignaud A, Ferry A, Messaddeq N, et al. 2011. Misregulated alternative splicing of BIN1 is associated with T tubule alterations and muscle weakness in myotonic dystrophy. *Nat Med* **17**: 720–726. <http://www.nature.com.gate2.inist.fr/articles/nm.2374.pdf> (Accessed May 25, 2018).
- Furling D, Lemieux D, Taneja K, Puymirat J. 2001. Decreased levels of myotonic dystrophy protein kinase (DMPK) and delayed differentiation in human myotonic dystrophy myoblasts. *Neuromuscul Disord* **11**: 728–735. www.elsevier.com/locate/nmd (Accessed January 8, 2019).
- Gabut M, Samavarchi-Tehrani P, Wang X, Slobodeniuc V, O’Hanlon D, Sung H-K, Alvarez M, Talukder S, Pan Q, Mazzoni EO, et al. 2011. An alternative splicing switch regulates embryonic stem cell pluripotency and reprogramming. *Cell* **147**: 132–46. <http://www.ncbi.nlm.nih.gov/pubmed/21924763> (Accessed December 7, 2020).
- García-Alcover I, Colonques-Bellmunt J, Garijo R, Tormo JR, Artero R, Álvarez-Abril MC, López Castel A, Pérez-Alonso M. 2014. Development of a *Drosophila melanogaster* spliceosensor system for in vivo high-throughput screening in myotonic dystrophy type 1. *Dis Model Mech* **7**: 1297–306. <http://www.ncbi.nlm.nih.gov/pubmed/25239918> (Accessed December 1, 2020).
- Gatchel JR, Zoghbi HY. 2005. Diseases of Unstable Repeat Expansion: Mechanisms and Common Principles. *Nat Rev Genet* **6**: 743–755. <http://www.nature.com/articles/nrg1691> (Accessed November 23, 2020).
- Gates DP, Coonrod LA, Berglund JA. 2011. Autoregulated splicing of muscleblind-like 1 (MBNL1) Pre-mRNA. *J Biol Chem* **286**: 34224–33. <http://www.ncbi.nlm.nih.gov/pubmed/21832083> (Accessed October 17, 2018).
- Geens M, Mateizel I, Sermon K, De Rycke M, Spits C, Cauffman G, Devroey P, Tournaye H, Liebaers I, Van de Velde H. 2009. Human embryonic stem cell lines derived from single blastomeres of two 4-cell stage embryos. *Hum Reprod* **24**: 2709–17. <http://www.ncbi.nlm.nih.gov/pubmed/19633307> (Accessed November 22, 2020).
- Gennarelli M, Pavoni M, Amicucci P, Angelini C, Menegazzo E, Zelano G, Novelli G, Dallapiccola B. 1999. Reduction of the DM-associated homeo domain protein (DMAHP) mRNA in different brain areas of myotonic dystrophy patients. *Neuromuscul Disord* **9**: 215–9. <http://www.ncbi.nlm.nih.gov/pubmed/10399747> (Accessed November 26, 2020).
- Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, Guimaraes C, Panning B, Ploegh HL, Bassik MC, et al. 2014. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159**: 647–61. <http://www.ncbi.nlm.nih.gov/pubmed/25307932> (Accessed November 23, 2020).
- Gomes-Pereira M, Foiry L, Nicole A, Huguet A, Junien C, Munnich A, Gourdon G. 2007. CTG trinucleotide repeat “big jumps”: large expansions, small mice. *PLoS Genet* **3**: e52. <http://www.ncbi.nlm.nih.gov/pubmed/17411343> (Accessed January 12, 2021).
- Gontier Y, Taivainen A, Fontao L, Sonnenberg A, van der Flier A, Carpen O, Faulkner G, Borradori L.

2005. The Z-disc proteins myotilin and FATZ-1 interact with each other and are connected to the sarcolemma via muscle-specific filamins. *J Cell Sci* **118**: 3739–49. <http://www.ncbi.nlm.nih.gov/pubmed/16076904> (Accessed February 8, 2021).
- González F, Boué S, Belmonte JCI. 2011. Methods for making induced pluripotent stem cells: reprogramming à la carte. *Nat Rev Genet* **12**: 231–242. <http://www.nature.com/articles/nrg2937> (Accessed November 22, 2020).
- Goodwin M, Mohan A, Batra R, Lee KY, Charizanis K, Fernández Gómez FJ, Eddarkaoui S, Sergeant N, Buée L, Kimura T, et al. 2015. MBNL Sequestration by Toxic RNAs and RNA Misprocessing in the Myotonic Dystrophy Brain. *Cell Rep* **12**: 1159–1168.
- Gourdon G, Meola G. 2017. Myotonic Dystrophies: State of the Art of New Therapeutic Developments for the CNS. *Front Cell Neurosci* **11**: 1–14. <http://journal.frontiersin.org/article/10.3389/fncel.2017.00101/full>.
- Guitart JR, Johnson JL, Chien WW. 2016. Research Techniques Made Simple: The Application of CRISPR-Cas9 and Genome Editing in Investigative Dermatology. *J Invest Dermatol* **136**: e87–e93. <https://www.sciencedirect.com/science/article/pii/S0022202X16313550> (Accessed November 23, 2020).
- Hale MA, Johnson NE, Berglund JA. 2019. Repeat-associated RNA structure and aberrant splicing. *Biochim Biophys Acta Gene Regul Mech* **1862**: 194405. <http://www.ncbi.nlm.nih.gov/pubmed/31323433> (Accessed November 26, 2020).
- Hallock PT, Chin S, Blais S, Neubert TA, Glass DJ. 2016. Sorbs1 and -2 Interact with CrkL and Are Required for Acetylcholine Receptor Cluster Formation. *Mol Cell Biol* **36**: 262–70. <http://www.ncbi.nlm.nih.gov/pubmed/26527617> (Accessed February 12, 2021).
- Han H, Irimia M, Ross PJ, Sung HK, Alipanahi B, David L, Golipour A, Gabut M, Michael IP, Nachman EN, et al. 2013. MBNL proteins repress ES-cell-specific alternative splicing and reprogramming. *Nature* **498**: 241–245. <http://dx.doi.org/10.1038/nature12270>.
- Hao M, Akrami K, Wei K, De Diego C, Che N, Ku J-H, Tidball J, Graves MC, Shieh PB, Chen F. 2008. Muscleblind-like 2 (Mbnl2) -deficient mice as a model for myotonic dystrophy. *Dev Dyn* **237**: 403–410. <http://doi.wiley.com/10.1002/dvdy.21428> (Accessed October 16, 2018).
- Harmon EB, Harmon ML, Larsen TD, Paulson AF, Perryman MB. 2008. Myotonic dystrophy protein kinase is expressed in embryonic myocytes and is required for myotube formation. *Dev Dyn* **237**: 2353–2366. <http://doi.wiley.com/10.1002/dvdy.21653> (Accessed November 26, 2020).
- Harper P. 2001. *Myotonic Dystrophy*. Oxford University Press, London https://books.google.fr/books?hl=fr&lr=&id=kbVHVPhFCOYC&oi=fnd&pg=PT6&ots=v7u3vzRL-&sig=SaW_BGEVPY7CK0CSHf3KbQXSK9I&redir_esc=y#v=onepage&q&f=false.
- Heatwole CR, Eichinger KJ, Friedman DI, Hilbert JE, Jackson CE, Logigian EL, Martens WB, McDermott MP, Pandya SK, Quinn C, et al. 2011. Open-label trial of recombinant human insulin-like growth factor 1/recombinant human insulin-like growth factor binding protein 3 in myotonic dystrophy type 1. *Arch Neurol* **68**: 37–44. <http://www.ncbi.nlm.nih.gov/pubmed/20837825> (Accessed December 3, 2020).
- Henry MP, Hawkins JR, Boyle J, Bridger JM. 2018. The Genomic Health of Human Pluripotent Stem Cells: Genomic Instability and the Consequences on Nuclear Organization. *Front Genet* **9**: 623. <http://www.ncbi.nlm.nih.gov/pubmed/30719030> (Accessed December 16, 2020).
- Hernández-Hernández O, Guiraud-Dogan C, Sicot G, Huguet A, Luillier S, Steidl E, Saenger S, Marciniak E, Obriot H, Chevarin C, et al. 2013. Myotonic dystrophy CTG expansion affects synaptic vesicle proteins, neurotransmission and mouse behaviour. *Brain* **136**: 957–70. <http://www.ncbi.nlm.nih.gov/pubmed/23404338> (Accessed December 1, 2020).
- Ho T, Savkur R, Poulos M, Mancini M, Swanson MS, Cooper TA. 2005. Colocalization of muscleblind with RNA foci is separable from mis-regulation of alternative splicing in myotonic dystrophy. *J Cell Sci* **118**: 2923–2933.

- <http://jcs.biologists.org.gate2.inist.fr/content/joces/118/13/2923.full.pdf> (Accessed April 18, 2018).
- Ho TH, Charlet-B N, Poulos MG, Singh G, Swanson MS, Cooper TA. 2004. Muscleblind proteins regulate alternative splicing. *EMBO J* **23**: 3103–3112. <http://emboj.embopress.org.gate2.inist.fr/content/embojnl/23/15/3103.full.pdf> (Accessed April 13, 2018).
- Hockemeyer D, Jaenisch R. 2016. Induced Pluripotent Stem Cells Meet Genome Editing. *Cell Stem Cell* **18**: 573–86. <http://www.ncbi.nlm.nih.gov/pubmed/27152442> (Accessed November 23, 2020).
- Hoffman EP, Kaminski HJ. 2012. Mexiletine for treatment of myotonia: a trial triumph for rare disease networks. *JAMA* **308**: 1377–8. <http://www.ncbi.nlm.nih.gov/pubmed/23032555> (Accessed December 3, 2020).
- Holt I, Jacquemin V, Fardaei M, Sewry CA, Butler-Browne GS, Furling D, Brook JD, Morris GE. 2009. Muscleblind-Like Proteins: Similarities and Differences in Normal and Myotonic Dystrophy Muscle. *Am J Pathol* **174**: 216–227. <https://www.ncbi.nlm.nih.gov.gate2.inist.fr/pmc/articles/PMC2631334/pdf/JPATH174000216.pdf> (Accessed April 23, 2018).
- Horrigan J, Gomes TB, Snape M, Nikolenko N, McMorn A, Evans S, Yaroshinsky A, Della Pasqua O, Oosterholt S, Lochmüller H. 2020. A Phase 2 Study of AMO-02 (Tideglusib) in Congenital and Childhood-Onset Myotonic Dystrophy Type 1 (DM1). *Pediatr Neurol* **112**: 84–93. <https://www.sciencedirect.com/science/article/pii/S0887899420302745> (Accessed December 4, 2020).
- Hsu M-N, Chang Y-H, Truong VA, Lai P-L, Nguyen TKN, Hu Y-C. 2019. CRISPR technologies for stem cell engineering and regenerative medicine. *Biotechnol Adv* **37**: 107447. <https://www.sciencedirect.com/science/article/abs/pii/S0734975019301478?via%3Dihub> (Accessed November 23, 2020).
- Hu JH, Miller SM, Geurts MH, Tang W, Chen L, Sun N, Zeina CM, Gao X, Rees HA, Lin Z, et al. 2018. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* **556**: 57–63. <http://www.nature.com.gate2.inist.fr/articles/nature26155.pdf> (Accessed June 20, 2018).
- Huangfu D, Osafune K, Maehr R, Guo W, Eijkelenboom A, Chen S, Muhlestein W, Melton DA. 2008. Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nat Biotechnol* **26**: 1269–1275. <http://www.nature.com/articles/nbt.1502> (Accessed November 22, 2020).
- Ilic D, Ogilvie C. 2017. Concise Review: Human Embryonic Stem Cells-What Have We Done? What Are We Doing? Where Are We Going? *Stem Cells* **35**: 17–25. <http://doi.wiley.com/10.1002/stem.2450> (Accessed November 22, 2020).
- Ishino Y, Krupovic M, Forterre P. 2018. History of CRISPR-Cas from Encounter with a Mysterious Repeated Sequence to Genome Editing Technology. *J Bacteriol* **200**. <http://www.ncbi.nlm.nih.gov/pubmed/29358495> (Accessed November 23, 2020).
- Jacinto F V., Link W, Ferreira BI. 2020. CRISPR/Cas9-mediated genome editing: From basic research to translational medicine. *J Cell Mol Med* **24**: 3766–3778. <https://onlinelibrary.wiley.com/doi/abs/10.1111/jcmm.14916> (Accessed November 23, 2020).
- Jansen G, Groenen PJ, Bächner D, Jap PH, Coerwinkel M, Oerlemans F, van den Broek W, Gohlsch B, Pette D, Plomp JJ, et al. 1996. Abnormal myotonic dystrophy protein kinase levels produce only mild myopathy in mice. *Nat Genet* **13**: 316–24. <http://www.ncbi.nlm.nih.gov/pubmed/8673131> (Accessed November 26, 2020).
- Jauvin D, Chrétien J, Pandey SK, Martineau L, Revillod L, Bassez G, Lachon A, McLeod AR, Gourdon G, Wheeler TM, et al. 2017. Targeting DMPK with Antisense Oligonucleotide Improves Muscle Strength in Myotonic Dystrophy Type 1 Mice. *Mol Ther - Nucleic Acids* **7**: 465–474.

<http://dx.doi.org/10.1016/j.omtn.2017.05.007>.

- Jenquin JR, Coonrod LA, Silverglate QA, Pellitier NA, Hale MA, Xia G, Nakamori M, Berglund JA. 2018. Furamidine rescues myotonic dystrophy type I associated mis-splicing through multiple mechanisms. *ACS Chem Biol*. <http://pubs.acs.org> (Accessed August 20, 2018).
- Jenquin JR, Yang H, Huigens RW, Nakamori M, Berglund JA. 2019. Combination Treatment of Erythromycin and Furamidine Provides Additive and Synergistic Rescue of Mis-Splicing in Myotonic Dystrophy Type 1 Models. *ACS Pharmacol Transl Sci* **2**: 247–263. <http://www.ncbi.nlm.nih.gov/pubmed/31485578> (Accessed December 3, 2020).
- Jiang F, Doudna JA. 2017. CRISPR–Cas9 Structures and Mechanisms. *Annu Rev Biophys* **46**: 505–529. <http://www.annualreviews.org/doi/10.1146/annurev-biophys-062215-010822> (Accessed November 23, 2020).
- Jiang H, Mankodi A, Swanson MS, Moxley RT, Thornton CA. 2004. Myotonic dystrophy type 1 is associated with nuclear foci of mutant RNA, sequestration of muscleblind proteins and deregulated alternative splicing in neurons. *Hum Mol Genet* **13**: 3079–88. <http://www.ncbi.nlm.nih.gov/pubmed/15496431> (Accessed November 27, 2020).
- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**: 816–21. <http://www.ncbi.nlm.nih.gov/pubmed/22745249> (Accessed February 11, 2021).
- Johnson K, Jarjour R. 2006. Distinct Mechanisms Downstream of the Repeat Expansion Are Implicated in the Molecular Basis of Myotonic Dystrophy Type 1. In *Molecular mechanisms of muscular dystrophies* (ed. S.J. Winder), pp. 160–174 https://books.google.fr/books?id=5mS1DwAAQBAJ&pg=PA160&lpg=PA160&dq=Johnson+and+Jarjour+2006+dystrophy&source=bl&ots=esv0Uklk6h&sig=ACfU3U3jWxN-rqPaCxoHjfoTz6VGobWA&hl=fr&sa=X&ved=2ahUKewi_44P0pqDtAhVBahQKHaiEDpwQ6AEwDHoECAMQAg#v=onepage&q=Johnson+and (Accessed November 26, 2020).
- Jungbluth H, Treves S, Zorzato F, Sarkozy A, Ochala J, Sewry C, Phadke R, Gautel M, Muntoni F. 2018. Congenital myopathies: disorders of excitation-contraction coupling and muscle contraction. *Nat Rev Neurol* **14**: 151–167. <http://www.ncbi.nlm.nih.gov/pubmed/29391587> (Accessed November 30, 2020).
- Kaliman P, Llagostera E. 2008. Myotonic dystrophy protein kinase (DMPK) and its role in the pathogenesis of myotonic dystrophy 1. *Cell Signal* **20**: 1935–1941. <https://linkinghub.elsevier.com/retrieve/pii/S0898656808001502> (Accessed November 25, 2020).
- Kalsotra A, Singh RK, Gurha P, Ward AJ, Creighton CJ, Cooper TA. 2014. The Mef2 transcription network is disrupted in myotonic dystrophy heart tissue, dramatically altering miRNA and mRNA expression. *Cell Rep* **6**: 336–45. <http://www.ncbi.nlm.nih.gov/pubmed/24412363> (Accessed November 29, 2020).
- Kalsotra A, Xiao X, Ward A, Castle J, Johnson J, Burge C, Cooper T. 2008. A Postnatal Switch of CELF and MBNL Proteins Reprograms Alternative Splicing in the Developing Heart. *Proc Natl Acad Sci U S A* **105**. <https://pubmed.ncbi.nlm.nih.gov/19075228/> (Accessed May 6, 2020).
- Kanadia RN, Johnstone KA, Mankodi A, Lungu C, Thornton CA, Esson D, Timmers AM, Hauswirth WW, Swanson MS. 2003. A Muscleblind Knockout Model for Myotonic Dystrophy. *Science (80-)* **302**: 1978–1980.
- Kanadia RN, Shin J, Yuan Y, Beattie SG, Wheeler TM, Thornton CA, Swanson MS. 2006. Reversal of RNA missplicing and myotonia after muscleblind overexpression in a mouse poly(CUG) model for myotonic dystrophy. *Proc Natl Acad Sci* **103**: 11748–11753. <http://www.pnas.org/cgi/doi/10.1073/pnas.0604970103>.
- Kania A, Salzberg A, Bhat M, D'Evelyn D, He Y, Kiss I, Bellen HJ. 1995. P-element mutations affecting embryonic peripheral nervous system development in *Drosophila melanogaster*. *Genetics* **139**:

- 1663–78. <http://www.ncbi.nlm.nih.gov/pubmed/7789767> (Accessed January 13, 2021).
- Kanitz A, Syed AP, Kaji K, Zavolan M. 2019. Conserved regulation of RNA processing in somatic cell reprogramming. *BMC Genomics* **20**: 100. <http://www.ncbi.nlm.nih.gov/pubmed/30704403> (Accessed January 13, 2021).
- Ketley A, Chen CZ, Li X, Arya S, Robinson TE, Granados-Riveron J, Udosen I, Morris GE, Holt I, Furling D, et al. 2014. High-content screening identifies small molecules that remove nuclear foci, affect MBNL distribution and CELF1 protein levels via a PKC-independent pathway in myotonic dystrophy cell lines. *Hum Mol Genet* **23**: 1551–1562.
- Ketley A, Wojciechowska M, Ghidelli-Disse S, Bamborough P, Ghosh TK, Morato ML, Sedehizadeh S, Malik NA, Tang Z, Powalowska P, et al. 2020. CDK12 inhibition reduces abnormalities in cells from patients with myotonic dystrophy and in a mouse model. *Sci Transl Med* **12**. <http://www.ncbi.nlm.nih.gov/pubmed/32350131> (Accessed January 28, 2021).
- Kim D, Kim C-H, Moon J-I, Chung Y-G, Chang M-Y, Han B-S, Ko S, Yang E, Cha KY, Lanza R, et al. 2009. Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. *Cell Stem Cell* **4**: 472–6. <http://www.ncbi.nlm.nih.gov/pubmed/19481515> (Accessed November 22, 2020).
- Kim EY, Barefield DY, Vo AH, Gacita AM, Schuster EJ, Wyatt EJ, Davis JL, Dong B, Sun C, Page P, et al. 2019. Distinct pathological signatures in human cellular models of myotonic dystrophy subtypes. *JCI insight* **4**. <http://www.ncbi.nlm.nih.gov/pubmed/30730308> (Accessed October 22, 2020).
- Kim HK, Lee S, Kim Y, Park J, Min S, Choi JW, Huang TP, Yoon S, Liu DR, Kim HH. 2020. High-throughput analysis of the activities of xCas9, SpCas9-NG and SpCas9 at matched and mismatched target sequences in human cells. *Nat Biomed Eng* **4**: 111–124. <http://www.ncbi.nlm.nih.gov/pubmed/31937939> (Accessed January 28, 2021).
- Kim HK, Pham MHC, Ko KS, Rhee BD, Han J. 2018. Alternative splicing isoforms in health and disease. *Pflugers Arch* **470**: 995–1016. <http://www.ncbi.nlm.nih.gov/pubmed/29536164> (Accessed December 7, 2020).
- Kimura T, Nakamori M, Lueck JD, Pouliquin P, Aoike F, Fujimura H, Dirksen RT, Takahashi MP, Dulhunty AF, Sakoda S. 2005. Altered mRNA splicing of the skeletal muscle ryanodine receptor and sarcoplasmic/endoplasmic reticulum Ca²⁺-ATPase in myotonic dystrophy type 1. *Hum Mol Genet* **14**: 2189–2200.
- Kinoshita M, Takahashi R, Hasegawa T, Komori T, Nagasawa R, Hirose K, Tanabe H. 1996. (CTG)_n expansions in various tissues from a myotonic dystrophy patient. *Muscle Nerve* **19**: 240–242. [https://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1097-4598\(199602\)19:2%3C240::AID-MUS21%3E3.0.CO;2-X](https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1097-4598(199602)19:2%3C240::AID-MUS21%3E3.0.CO;2-X) (Accessed November 25, 2020).
- Kioka N, Ueda K, Amachi T. 2002. Vinexin, CAP/ponsin, ArgBP2: A novel adaptor protein family regulating cytoskeletal organization and signal transduction. *Cell Struct Funct* **27**: 1–7.
- Klein AF, Dastidar S, Furling D, Chuah MK. 2015. Therapeutic Approaches for Dominant Muscle Diseases: Highlight on Myotonic Dystrophy. *Curr Gene Ther* **15**: 329–37. <http://www.ncbi.nlm.nih.gov/pubmed/26122101> (Accessed December 1, 2020).
- Klein AF, Varela MA, Arandel L, Holland A, Naouar N, Arzumanov A, Seoane D, Revillod L, Bassez G, Ferry A, et al. 2019. Peptide-conjugated oligonucleotides evoke long-lasting myotonic dystrophy correction in patient-derived cells and mice. *J Clin Invest* **129**: 4739–4744. <http://www.ncbi.nlm.nih.gov/pubmed/31479430> (Accessed May 6, 2020).
- Kleinstiver BP, Pattanayak V, Prew MS, Tsai SQ, Nguyen NT, Zheng Z, Joung JK. 2016. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**: 490–5. <http://www.ncbi.nlm.nih.gov/pubmed/26735016> (Accessed January 28, 2021).
- Klimanskaya I, Chung Y, Becker S, Lu S-J, Lanza R. 2006. Human embryonic stem cell lines derived from single blastomeres. *Nature* **444**: 481–485. <http://www.nature.com/articles/nature05142>

(Accessed November 22, 2020).

- Kodaka Y, Rabu G, Asakura A. 2017. Skeletal Muscle Cell Induction from Pluripotent Stem Cells. *Stem Cells Int* **2017**: 1376151. <http://www.ncbi.nlm.nih.gov/pubmed/28529527> (Accessed November 23, 2020).
- Koehorst E, Ballester-Lopez A, Arechavala-Gomez V, Martínez-Piñero A, Nogales-Gadea G. 2020. The Biomarker Potential of miRNAs in Myotonic Dystrophy Type I. *J Clin Med* **9**: 3939. <https://www.mdpi.com/2077-0383/9/12/3939> (Accessed December 15, 2020).
- Konieczny P, Stepniak-Konieczna E, Sobczak K. 2018. MBNL expression in autoregulatory feedback loops. *RNA Biol* **15**: 1–8. <https://doi.org/10.1080/15476286.2017.1384119>.
- Konieczny P, Stepniak-Konieczna E, Sobczak K. 2014. MBNL proteins and their target RNAs, interaction and splicing regulation. *Nucleic Acids Res* **42**: 10873–10887. https://watermark.silverchair.com/gku767.pdf?token=AQECAHi208BE49Ooan9kkhW_Ercy7Dm3ZL_9Cf3qfKAc485ysgAAAagwggGkBgkqhkiG9w0BBwagggGVMIIbKQIBADCCAYoGCSqGSIB3DQEHATAeBglghkgBZQMEAS4wEQQMZEYyRKcyW2DSPJPTAgEQgIIBW6jN0x7jcyj0aSKvbQMdVxa2jK_3UjWJvfAqJaoMO9re4F4Q (Accessed April 18, 2018).
- Koonin E V, Makarova KS. 2019. Origins and evolution of CRISPR-Cas systems. *Philos Trans R Soc Lond B Biol Sci* **374**: 20180087. <http://www.ncbi.nlm.nih.gov/pubmed/30905284> (Accessed November 23, 2020).
- Koshelev M, Sarma S, Price RE, Wehrens XHT, Cooper TA. 2010. Heart-specific overexpression of CUGBP1 reproduces functional and molecular abnormalities of myotonic dystrophy type 1. *Hum Mol Genet* **19**: 1066–75. <http://www.ncbi.nlm.nih.gov/pubmed/20051426> (Accessed November 29, 2020).
- Krol J, Fiszler A, Mykowska A, Sobczak K, de Mezer M, Krzyzosiak WJ. 2007. Ribonuclease dicer cleaves triplet repeat hairpins into shorter repeats that silence specific targets. *Mol Cell* **25**: 575–86. <http://www.ncbi.nlm.nih.gov/pubmed/17317629> (Accessed November 30, 2020).
- Kumar A, Agarwal S, Pradhan S. 2014. Assessment of Premutation in Myotonic Dystrophy Type 1 Affected Family Members by TP-PCR and Genetic Counseling. *Case Rep Med* **2014**: 289643. <http://www.ncbi.nlm.nih.gov/pubmed/24715907> (Accessed November 23, 2020).
- Kuyumcu-Martinez NM, Wang G-S, Cooper TA. 2007. Increased steady-state levels of CUGBP1 in myotonic dystrophy 1 are due to PKC-mediated hyperphosphorylation. *Mol Cell* **28**: 68–78. <http://www.ncbi.nlm.nih.gov/pubmed/17936705> (Accessed November 29, 2020).
- Ladd AN, Charlet N, Cooper TA. 2001. The CELF family of RNA binding proteins is implicated in cell-specific and developmentally regulated alternative splicing. *Mol Cell Biol* **21**: 1285–96. <http://www.ncbi.nlm.nih.gov/pubmed/11158314> (Accessed November 28, 2020).
- Langlois M-A, Boniface C, Wang G, Alluin J, Salvaterra PM, Puymirat J, Rossi JJ, Lee NS. 2005. Cytoplasmic and nuclear retained DMPK mRNAs are targets for RNA interference in myotonic dystrophy cells. *J Biol Chem* **280**: 16949–54. <http://www.ncbi.nlm.nih.gov/pubmed/15722335> (Accessed February 11, 2021).
- Langlois M-A, Lee NS, Rossi JJ, Puymirat J. 2003. Hammerhead ribozyme-mediated destruction of nuclear foci in myotonic dystrophy myoblasts. *Mol Ther* **7**: 670–680. <https://www.sciencedirect.com/science/article/pii/S1525001603000686> (Accessed February 11, 2021).
- Laurent F-X, Sureau A, Klein AF, Trouslard F, Gasnier E, Furling D, Marie J. 2012. New function for the RNA helicase p68/DDX5 as a modifier of MBNL1 activity on expanded CUG repeats. *Nucleic Acids Res* **40**: 3159–71. <http://www.ncbi.nlm.nih.gov/pubmed/22156369> (Accessed December 10, 2020).
- Laustriat D, Gide J, Barrault L, Chautard E, Benoit C, Auboeuf D, Boland A, Battail C, Artiguenave F, Deleuze J-F, et al. 2015. In Vitro and In Vivo Modulation of Alternative Splicing by the Biguanide Metformin. *Mol Ther - Nucleic Acids* **4**. www.nature.com/mtna (Accessed May 25, 2018).

- Lee E, Marcucci M, Daniell L, Pypaert M, Weisz OA, Ochoa G-C, Farsad K, Wenk MR, De Camilli P. 2002. Amphiphysin 2 (Bin1) and T-tubule biogenesis in muscle. *Science* **297**: 1193–6. <http://www.ncbi.nlm.nih.gov/pubmed/12183633> (Accessed November 30, 2020).
- Lee JE, Cooper TA. 2009. Pathogenic mechanisms of myotonic dystrophy. *Biochem Soc Trans* **37**: 1281–6. <http://www.ncbi.nlm.nih.gov/pubmed/19909263> (Accessed November 29, 2020).
- Lee K-S, Smith K, Amieux PS, Wang EH. 2008a. MBNL3/CHCR prevents myogenic differentiation by inhibiting MyoD- dependent gene transcription. *Differentiation* **76**: 299–309. https://ac-els-cdn-com.gate2.inist.fr/S0301468109600748/1-s2.0-S0301468109600748-main.pdf?_tid=f5202fe7-d9ae-4617-8501-2bef96febb94&acdnat=1527600092_2cb7dc27679a535fde293b8876cbea5c (Accessed May 29, 2018).
- Lee K-Y, Chang H-C, Seah C, Lee L-J. 2019. Deprivation of Muscleblind-Like Proteins Causes Deficits in Cortical Neuron Distribution and Morphological Changes in Dendritic Spines and Postsynaptic Densities. *Front Neuroanat* **13**: 75. <https://www.frontiersin.org/article/10.3389/fnana.2019.00075/full> (Accessed August 19, 2019).
- Lee KY, Li M, Manchanda M, Batra R, Charizanis K, Mohan A, Warren SA, Chamberlain CM, Finn D, Hong H, et al. 2013. Compound loss of muscleblind-like function in myotonic dystrophy. *EMBO Mol Med* **5**: 1887–1900.
- Lee S, Oh S, Kim D, Han J, Moon S, Lim J. 2006. Serum replacement with a growth factor-free synthetic substance in culture medium contributes to effective establishment of mouse embryonic stem cells of various origins. *Fertil Steril* **86**: 1137–1145. <https://linkinghub.elsevier.com/retrieve/pii/S0015028206010788> (Accessed November 22, 2020).
- Lee Y, Rudell J, Yechikhov S, Taylor R, Swope S, Ferns M. 2008b. Rapsyn carboxyl terminal domains mediate muscle specific kinase-induced phosphorylation of the muscle acetylcholine receptor. *Neuroscience* **153**: 997–1007. <http://www.ncbi.nlm.nih.gov/pubmed/18436384> (Accessed January 21, 2021).
- Lefort N, Perrier AL, Laâbi Y, Varela C, Peschanski M. 2009. Human embryonic stem cells and genomic instability. *Regen Med* **4**: 899–909. <http://www.ncbi.nlm.nih.gov/pubmed/19903007> (Accessed January 13, 2021).
- Légaré C, Overend G, Guay S-P, Monckton DG, Mathieu J, Gagnon C, Bouchard L. 2019. DMPK gene DNA methylation levels are associated with muscular and respiratory profiles in DM1. *Neurol Genet* **5**: e338. <http://ng.neurology.org/lookup/doi/10.1212/NXG.0000000000000338> (Accessed November 25, 2020).
- Leikina E, Gamage DG, Prasad V, Goykhberg J, Crowe M, Diao J, Kozlov MM, Chernomordik L V, Millay DP. 2018. Myomaker and Myomerger Work Independently to Control Distinct Steps of Membrane Remodeling during Myoblast Fusion. *Dev Cell* **0**. <http://www.ncbi.nlm.nih.gov/pubmed/30197239> (Accessed September 17, 2018).
- Lin X, Miller JW, Mankodi A, Kanadia RN, Yuan Y, Moxley RT, Swanson MS, Thornton CA. 2006. Failure of MBNL1-dependent post-natal splicing transitions in myotonic dystrophy. *Hum Mol Genet* **15**: 2087–2097.
- Liquori CL, Ricker K, Moseley ML, Jacobsen JF, Kress W, Naylor SL, Day JW, Ranum LP. 2001. Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science (80-)* **293**: 864–7. <http://www.ncbi.nlm.nih.gov/pubmed/11486088> (Accessed November 23, 2020).
- Liu Q, Segal DJ, Ghiara JB, Barbas CF. 1997. Design of polydactyl zinc-finger proteins for unique addressing within complex genomes. *Proc Natl Acad Sci* **94**: 5525–5530. <https://www.pnas.org/content/94/11/5525> (Accessed November 23, 2020).
- Liu X, Yu T, Sun Y, Wang H. 2018. Characterization of novel alternative splicing variants of Oct4 gene expressed in mouse pluripotent stem cells. *J Cell Physiol* **233**: 5468–5477. <http://www.ncbi.nlm.nih.gov/pubmed/29266259> (Accessed January 13, 2021).

- Llagostera E, Catalucci D, Marti L, Liesa M, Camps M, Ciaraldi TP, Kondo R, Reddy S, Dillmann WH, Palacin M, et al. 2007. Role of Myotonic Dystrophy Protein Kinase (DMPK) in Glucose Homeostasis and Muscle Insulin Action ed. B. McCabe. *PLoS One* **2**: e1134. <https://dx.plos.org/10.1371/journal.pone.0001134> (Accessed November 26, 2020).
- Llamusi B, Artero R. 2008. Molecular Effects of the CTG Repeats in Mutant Dystrophia Myotonica Protein Kinase Gene. *Curr Genomics* **9**: 509–16. <http://www.ncbi.nlm.nih.gov/pubmed/19516957> (Accessed November 26, 2020).
- Loh Y-H, Hartung O, Li H, Guo C, Sahalie JM, Manos PD, Urbach A, Heffner GC, Grskovic M, Vigneault F, et al. 2010. Reprogramming of T Cells from Human Peripheral Blood. *Cell Stem Cell* **7**: 15. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2913590/> (Accessed November 22, 2020).
- López-Martínez A, Soblechero-Martín P, de-la-Puente-Ovejero L, Nogales-Gadea G, Arechavala-Gomez V. 2020. An Overview of Alternative Splicing Defects Implicated in Myotonic Dystrophy Type I. *Genes (Basel)* **11**: 1109. <https://www.mdpi.com/2073-4425/11/9/1109> (Accessed October 6, 2020).
- López-Morató M, Brook JD, Wojciechowska M. 2018. Small Molecules Which Improve Pathogenesis of Myotonic Dystrophy Type 1. *Front Neurol* **9**. <https://www.ncbi.nlm.nih.gov.proxy.insermbiblio.inist.fr/pmc/articles/PMC5968088/> (Accessed May 6, 2020).
- Macdermot V. 1961. THE HISTOLOGY OF THE NEUROMUSCULAR JUNCTION IN DYSTROPHIA MYOTONICA. *Brain* **84**: 75–84. <https://academic.oup.com/brain/article-lookup/doi/10.1093/brain/84.1.75> (Accessed February 2, 2021).
- Maffioletti SM, Gerli MFM, Ragazzi M, Dastidar S, Benedetti S, Loperfido M, VandenDriessche T, Chuah MK, Tedesco FS. 2015. Efficient derivation and inducible differentiation of expandable skeletal myogenic cells from human ES and patient-specific iPSC cells. *Nat Protoc* **10**: 941–958. <http://www.nature.com/articles/nprot.2015.057> (Accessed November 23, 2020).
- Maffioletti SM, Sarcar S, Henderson ABH, Mannhardt I, Pinton L, Moyle LA, Steele-Stallard H, Cappellari O, Wells KE, Ferrari G, et al. 2018. Three-Dimensional Human iPSC-Derived Artificial Skeletal Muscles Model Muscular Dystrophies and Enable Multilineage Tissue Engineering. *Cell Rep* **23**: 899–908. <http://www.ncbi.nlm.nih.gov/pubmed/29669293> (Accessed May 6, 2020).
- Magana JJ, Cisneros B. 2012. Myotonic Dystrophy Type 1 (DM1): From the Genetics to Molecular Mechanisms. In *Muscular Dystrophy*, InTech <http://www.intechopen.com/books/muscular-dystrophy/myotonic-dystrophy-type-1-dm1-from-the-genetics-to-molecular-mechanisms> (Accessed November 28, 2020).
- Mahadevan M, Tsilfidis C, Sabourin L, Shutler G, Amemiya C, Jansen G, Neville C, Narang M, Barceló J, O'Hoy K, et al. 1992. Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. *Science (80-)* **255**: 1253–1255. <http://science.sciencemag.org.gate2.inist.fr/content/sci/255/5049/1253.full.pdf> (Accessed April 18, 2018).
- Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. 2013. RNA-guided human genome engineering via Cas9. *Science (80-)* **339**: 823–6. <http://www.ncbi.nlm.nih.gov/pubmed/23287722> (Accessed November 23, 2020).
- Malik N, Rao MS. 2013. A Review of the Methods for Human iPSC Derivation. *Methods Mol Biol* **997**: 23. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4176696/> (Accessed December 16, 2020).
- Mallon BS, Park K-Y, Chen KG, Hamilton RS, McKay RDG. 2006. Toward xeno-free culture of human embryonic stem cells. *Int J Biochem Cell Biol* **38**: 1063–75. <http://www.ncbi.nlm.nih.gov/pubmed/16469522> (Accessed November 22, 2020).
- Mandai K, Nakanishi H, Satoh A, Takahashi K, Satoh K, Nishioka H, Mizoguchi A, Takai Y. 1999. Ponsin/SH3P12: an I-afadin- and vinculin-binding protein localized at cell-cell and cell-matrix adherens junctions. *J Cell Biol* **144**: 1001–17. <http://www.ncbi.nlm.nih.gov/pubmed/10085297> (Accessed January 21, 2021).

- Mankodi A, Logigian E, Callahan L, McClain C, White R, Henderson D, Krym M, Thornton C. 2000. Myotonic Dystrophy in Transgenic Mice Expressing an Expanded CUG Repeat. *Science (80-)* **289**: 1769–1773. <http://science.sciencemag.org.gate2.inist.fr/content/sci/289/5485/1769.full.pdf> (Accessed April 18, 2018).
- Mankodi A, Takahashi MP, Jiang H, Beck CL, Bowers WJ, Moxley RT, Cannon SC, Thornton CA. 2002. Expanded CUG repeats trigger aberrant splicing of CIC-1 chloride channel pre-mRNA and hyperexcitability of skeletal muscle in myotonic dystrophy. *Mol Cell* **10**: 35–44.
- Mankodi A, Teng-Umuay P, Krym M, Henderson D, Swanson M, Thornton CA. 2003. Ribonuclear inclusions in skeletal muscle in myotonic dystrophy types 1 and 2. *Ann Neurol* **54**: 760–8. <http://www.ncbi.nlm.nih.gov/pubmed/14681885> (Accessed November 27, 2020).
- Marteyn A, Maury Y, Gauthier MM, Lecuyer C, Vernet R, Denis JA, Pietu G, Peschanski M, Martinat C. 2011. Mutant Human Embryonic Stem Cells Reveal Neurite and Synapse Formation Defects in Type 1 Myotonic Dystrophy. *Cell Stem Cell* **8**: 434–444. [https://www.cell.com/cell-stem-cell/pdf/S1934-5909\(11\)00055-5.pdf](https://www.cell.com/cell-stem-cell/pdf/S1934-5909(11)00055-5.pdf) (Accessed June 12, 2018).
- Martin GR. 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A* **78**: 7634. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC349323/> (Accessed November 22, 2020).
- Martineau L, Racine V, Benichou SA, Puymirat J. 2018. Lymphoblastoids cell lines – Derived iPSC line from a 26-year-old myotonic dystrophy type 1 patient carrying (CTG)200 expansion in the DMPK gene: CHUQi001-A. *Stem Cell Res* **26**: 103–106. <https://www.sciencedirect.com/science/article/pii/S1873506117302696?via%3Dihub> (Accessed November 23, 2020).
- Martorell L, Johnson K, Boucher CA, Baiget M. 1997. Somatic instability of the myotonic dystrophy (CTG)_n repeat during human fetal development. *Hum Mol Genet* **6**: 877–80. <http://www.ncbi.nlm.nih.gov/pubmed/9175733> (Accessed November 25, 2020).
- Maruyama T, Dougan SK, Truttmann MC, Bilate AM, Ingram JR, Ploegh HL. 2015. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat Biotechnol* **33**: 538–42. <http://www.ncbi.nlm.nih.gov/pubmed/25798939> (Accessed November 23, 2020).
- Masuda A, Andersen HS, Koed Doktor T, Okamoto T, Ito M, Andresen BS, Ohno K. 2012. CUGBP1 and MBNL1 preferentially bind to 3'UTRs and facilitate mRNA decay. *Sci Rep* **2**. <http://www.nature.com.gate2.inist.fr/articles/srep00209.pdf> (Accessed April 18, 2018).
- Mateizel I, De Temmerman N, Ullmann U, Cauffman G, Sermon K, Van de Velde H, De Rycke M, Degreef E, Devroey P, Liebaers I, et al. 2006. Derivation of human embryonic stem cell lines from embryos obtained after IVF and after PGD for monogenic disorders. *Hum Reprod* **21**: 503–11. <http://www.ncbi.nlm.nih.gov/pubmed/16284066> (Accessed November 22, 2020).
- Mateos-Aierdi AJ, Goicoechea M, Aiausti A, Fernández-Torrón R, Garcia-Puga M, Matheu A, Munain AL de. 2015. Muscle wasting in myotonic dystrophies: a model of premature aging. *Front Aging Neurosci* **7**. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4496580/> (Accessed December 14, 2020).
- Matloka M, Klein AF, Rau F, Furling D. 2018. Cells of Matter—In Vitro Models for Myotonic Dystrophy. *Front Neurol* **9**: 361. <https://www.frontiersin.org/article/10.3389/fneur.2018.00361/full> (Accessed April 16, 2020).
- Matynia A, Ng CH, Dansithong W, Chiang A, Silva AJ, Reddy S. 2010. Muscleblind1, but Not Dmpk or Six5, Contributes to a Complex Phenotype of Muscular and Motivational Deficits in Mouse Models of Myotonic Dystrophy ed. H. Okazawa. *PLoS One* **5**: e9857. <http://dx.plos.org/10.1371/journal.pone.0009857> (Accessed July 24, 2018).
- Maury Y, Côme J, Piskorowski RA, Salah-Mohellibi N, Chevaleyre V, Peschanski M, Martinat C, Nedelec

- S. 2015. Combinatorial analysis of developmental cues efficiently converts human pluripotent stem cells into multiple neuronal subtypes. *Nat Biotechnol* **33**: 89–96. <http://www.nature.com/articles/nbt.3049> (Accessed May 6, 2020).
- Maury Y, Poydenot P, Brinon B, Lesueur L, Gide J, Roquevière S, Côme J, Polvèche H, Auboeuf D, Alexandre Denis J, et al. 2019. Pluripotent Stem Cell-Based Drug Screening Reveals Cardiac Glycosides as Modulators of Myotonic Dystrophy Type 1. *iScience* **11**: 258–271. <http://www.ncbi.nlm.nih.gov/pubmed/30639849> (Accessed May 6, 2020).
- Meola G, Cardani R. 2015. Myotonic dystrophies: An update on clinical aspects, genetic, pathology, and molecular pathomechanisms. *Biochim Biophys Acta - Mol Basis Dis* **1852**: 594–606. <https://www.sciencedirect.com/science/article/pii/S0925443914001471?via%3Dihub> (Accessed November 23, 2020).
- Merkle FT, Ghosh S, Kamitaki N, Mitchell J, Avior Y, Mello C, Kashin S, Mekhoubad S, Ilic D, Charlton M, et al. 2017. Human pluripotent stem cells recurrently acquire and expand dominant negative P53 mutations. *Nature* **545**: 229–233. <http://www.nature.com/articles/nature22312> (Accessed January 13, 2021).
- Michalowski S, Miller JW, Urbinati CR, Paliouras M, Swanson MS, Griffith J. 1999. Visualization of double-stranded RNAs from the myotonic dystrophy protein kinase gene and interactions with CUG-binding protein. *Nucleic Acids Res* **27**: 3534–42. <http://www.ncbi.nlm.nih.gov/pubmed/10446244> (Accessed November 29, 2020).
- Michel L, Huguet-Lachon A, Gourdon G. 2015. Sense and Antisense DMPK RNA Foci Accumulate in DM1 Tissues during Development. *PLoS One* **10**. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4560382/> (Accessed May 6, 2020).
- Millay DP, O'Rourke JR, Sutherland LB, Bezprozvannaya S, Shelton JM, Bassel-Duby R, Olson EN. 2013. Myomaker is a membrane activator of myoblast fusion and muscle formation. *Nature* **499**: 301–5. <http://www.ncbi.nlm.nih.gov/pubmed/23868259> (Accessed February 2, 2021).
- Miller JW, Urbinati CR, Thornton CA, Swanson MS. 2000. Recruitment of human muscleblind proteins to (CUG)_n expansions associated with myotonic dystrophy. *EMBO J* **19**: 4439–4448.
- Mishra SK, Singh S, Lee B, Khosa S, Moheb N, Tandon VA. 2018. Dystrophia Myotonica and the Legacy of Hans Gustav Wilhelm Steinert. *Ann Indian Acad Neurol* **21**: 116–118. <http://www.ncbi.nlm.nih.gov/pubmed/30122835> (Accessed November 23, 2020).
- Mondragon-Gonzalez R, Perlingeiro RCR. 2018. Recapitulating muscle disease phenotypes with myotonic dystrophy 1 induced pluripotent stem cells: a tool for disease modeling and drug discovery. *Dis Model Mech* **11**. <http://www.ncbi.nlm.nih.gov/pubmed/29898953> (Accessed October 22, 2020).
- Montes M, Sanford BL, Comiskey DF, Chandler DS. 2019. RNA Splicing and Disease: Animal Models to Therapies. *Trends Genet* **35**. <https://doi.org/10.1016/j.tig.2018.10.002> (Accessed January 12, 2021).
- Mounsey JP, John JE, Helmke SM, Bush EW, Gilbert J, Roses AD, Perryman MB, Jones LR, Moorman JR. 2000. Phospholemman is a substrate for myotonic dystrophy protein kinase. *J Biol Chem* **275**: 23362–7. <http://www.ncbi.nlm.nih.gov/pubmed/10811636> (Accessed November 26, 2020).
- Mulders SAM, van Horssen R, Gerrits L, Bennink MB, Pluk H, de Boer-van Huizen RT, Croes HJE, Wijers M, van de Loo FAJ, Fransen J, et al. 2011. Abnormal actomyosin assembly in proliferating and differentiating myoblasts upon expression of a cytosolic DMPK isoform. *Biochim Biophys Acta - Mol Cell Res* **1813**: 867–877. <https://www.sciencedirect.com/science/article/pii/S0167488911000322#bb0040> (Accessed November 26, 2020).
- Nakamori M, Hamanaka K, Thomas JD, Wang ET, Hayashi YK, Takahashi MP, Swanson MS, Nishino I, Mochizuki H. 2017. Aberrant Myokine Signaling in Congenital Myotonic Dystrophy. *Cell Rep* **21**: 1240–1252. <http://www.ncbi.nlm.nih.gov/pubmed/29091763> (Accessed September 19, 2018).

- Nakamori M, Kimura T, Fujimura H, Takahashi MP, Sakoda S. 2007. Altered mRNA splicing of dystrophin in type 1 myotonic dystrophy. *Muscle Nerve* **36**: 251–257. <http://doi.wiley.com/10.1002/mus.20809> (Accessed September 19, 2018).
- Nakamori M, Sobczak K, Puwanant A, Welle S, Eichinger K, Pandya S, Dekdebrun J, Heatwole CR, McDermott MP, Chen T, et al. 2013. Splicing biomarkers of disease severity in myotonic dystrophy. *Ann Neurol* **74**: 862–872.
- Nakamori M, Taylor K, Mochizuki H, Sobczak K, Takahashi MP. 2016. Oral administration of erythromycin decreases RNA toxicity in myotonic dystrophy. *Ann Clin Transl Neurol* **3**: 42–54. <http://www.ncbi.nlm.nih.gov/pubmed/26783549> (Accessed December 3, 2020).
- Napierała M, Krzyzosiak WJ. 1997. CUG repeats present in myotonin kinase RNA form metastable slippery hairpins. *J Biol Chem* **272**: 31079–85. <http://www.ncbi.nlm.nih.gov/pubmed/9388259> (Accessed November 26, 2020).
- Nikonova E, Kao S-Y, Ravichandran K, Wittner A, Spletter ML. 2019. Conserved functions of RNA-binding proteins in muscle. *Int J Biochem Cell Biol* **110**: 29–49. <http://www.ncbi.nlm.nih.gov/pubmed/30818081> (Accessed October 22, 2020).
- Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, Ishitani R, Zhang F, Nureki O. 2014. Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell* **156**: 935–949. [https://www.cell.com/fulltext/S0092-8674\(14\)00156-1](https://www.cell.com/fulltext/S0092-8674(14)00156-1) (Accessed November 23, 2020).
- Nishimasu H, Shi X, Ishiguro S, Gao L, Hirano S, Okazaki S, Noda T, Abudayyeh OO, Gootenberg JS, Mori H, et al. 2018. Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science* **361**: 1259–1262. <http://www.ncbi.nlm.nih.gov/pubmed/30166441> (Accessed January 28, 2021).
- Nutter CA, Bubenik JL, Oliveira R, Ivankovic F, Sznajder ŁJ, Kidd BM, Pinto BS, Otero BA, Carter HA, Vitriol EA, et al. 2019. Cell-type-specific dysregulation of RNA alternative splicing in short tandem repeat mouse knockin models of myotonic dystrophy. *Genes Dev*. <http://genesdev.cshlp.org/lookup/doi/10.1101/gad.328963.119> (Accessed October 21, 2019).
- Orengo JP, Chambon P, Metzger D, Mosier DR, Snipes GJ, Cooper TA. 2008. Expanded CTG repeats within the DMPK 3' UTR causes severe skeletal muscle wasting in an inducible mouse model for myotonic dystrophy. *Proc Natl Acad Sci U S A* **105**: 2646–51. <http://www.ncbi.nlm.nih.gov/pubmed/18272483> (Accessed November 29, 2020).
- Osborne RJ, Lin X, Welle S, Sobczak K, O'Rourke JR, Swanson MS, Thornton CA. 2009. Transcriptional and post-transcriptional impact of toxic RNA in myotonic dystrophy. *Hum Mol Genet* **18**: 1471–81. <http://www.ncbi.nlm.nih.gov/pubmed/19223393> (Accessed October 22, 2020).
- Panaite P-A, Kuntzer T, Gourdon G, Lobrinus JA, Barakat-Walter I. 2013. Functional and histopathological identification of the respiratory failure in a DMSXL transgenic mouse model of myotonic dystrophy. *Dis Model Mech* **6**: 622–31. <http://www.ncbi.nlm.nih.gov/pubmed/23180777> (Accessed February 11, 2021).
- Pandey SK, Wheeler TM, Justice SL, Kim A, Younis HS, Gattis D, Jauvin D, Puymirat J, Swayze EE, Freier SM, et al. 2015. Identification and Characterization of Modified Antisense Oligonucleotides Targeting DMPK in Mice and Nonhuman Primates for the Treatment of Myotonic Dystrophy Type 1. *J Pharmacol Exp Ther* **355**: 310–321. <http://jpet.aspetjournals.org/cgi/doi/10.1124/jpet.115.226969>.
- Park S, Phukan PD, Zeeb M, Martinez-Yamout MA, Dyson HJ, Wright PE. 2017. Structural Basis for Interaction of the Tandem Zinc Finger Domains of Human Muscleblind with Cognate RNA from Human Cardiac Troponin T. *Biochemistry* **56**: 4154–4168. <http://www.ncbi.nlm.nih.gov/pubmed/28718627> (Accessed December 8, 2020).
- Pascual M, Lidon V, Artero MR. 2006. The Muscleblind family of proteins: an emerging class of regulators of developmentally programmed alternative splicing. *Differentiation*. https://ac-els-cdn-com.gate2.inist.fr/S0301468109601973/1-s2.0-S0301468109601973-main.pdf?_tid=06101d1e-c243-459e-b105-

2dd15e64d284&acdnat=1523610799_7273b430745488fc25976fc6f7c5de98 (Accessed April 13, 2018).

- Paul S, Dansithong W, Kim D, Rossi J, Webster NJG, Comai L, Reddy S. 2006. Interaction of muscleblind, CUG-BP1 and hnRNP H proteins in DM1-associated aberrant IR splicing. *EMBO J* **25**: 4271–83. <http://www.ncbi.nlm.nih.gov/pubmed/16946708> (Accessed February 17, 2021).
- Paulson H. 2018. Repeat expansion diseases. *Handb Clin Neurol* **147**: 105–123. <http://www.ncbi.nlm.nih.gov/pubmed/29325606> (Accessed November 23, 2020).
- Pearson CE. 2011. Repeat associated non-ATG translation initiation: one DNA, two transcripts, seven reading frames, potentially nine toxic entities! *PLoS Genet* **7**: e1002018. <http://www.ncbi.nlm.nih.gov/pubmed/21423665> (Accessed November 30, 2020).
- Pedrotti S, Giudice J, Dagnino-Acosta A, Knoblauch M, Singh RK, Hanna A, Mo Q, Hicks J, Hamilton S, Cooper TA. 2015. The RNA-binding protein Rbfox1 regulates splicing required for skeletal muscle structure and function. *Hum Mol Genet* **24**: 2360–74. <http://www.ncbi.nlm.nih.gov/pubmed/25575511> (Accessed October 22, 2020).
- Pénisson-Besnier I, Devillers M, Porcher R, Orlikowski D, Doppler V, Desnuelle C, Ferrer X, Bes M-CA, Bouhour F, Tranchant C, et al. 2008. Dehydroepiandrosterone for myotonic dystrophy type 1. *Neurology* **71**: 407–12. <http://www.ncbi.nlm.nih.gov/pubmed/18678823> (Accessed December 3, 2020).
- Pettersson OJ, Aagaard L, Jensen TG, Damgaard CK. 2015. Molecular mechanisms in DM1 - a focus on foci. *Nucleic Acids Res* **43**: 2433–41. <http://www.ncbi.nlm.nih.gov/pubmed/25605794> (Accessed December 10, 2020).
- Pinto BS, Saxena T, Oliveira R, Méndez-Gómez HR, Cleary JD, Denes LT, McConnell O, Arboleda J, Xia G, Swanson MS, et al. 2017. Impeding Transcription of Expanded Microsatellite Repeats by Deactivated Cas9. *Mol Cell* **68**: 479-490.e5. <http://www.ncbi.nlm.nih.gov/pubmed/29056323> (Accessed November 23, 2020).
- Pistoni M, Ghigna C, Gabellini D. 2010. Alternative splicing and muscular dystrophy. *RNA Bio* **7**: 441–52. <http://www.ncbi.nlm.nih.gov/pubmed/20603608> (Accessed December 7, 2020).
- Poulos MG, Batra R, Charizanis K, Swanson MS. 2011. Developments in RNA Splicing and Disease. *Cold Spring Harb Perspect Biol* **3**. <http://cshperspectives.cshlp.org/> (Accessed January 8, 2019).
- Poulos MG, Batra R, Li M, Yuan Y, Zhang C, Darnell RB, Swanson MS. 2013. Progressive impairment of muscle regeneration in muscleblind-like 3 isoform knockout mice. *Hum Mol Genet* **22**: 3547–3558.
- Pratte A, Prévost C, Puymirat J, Mathieu J. 2015. Anticipation in myotonic dystrophy type 1 parents with small CTG expansions. *Am J Med Genet Part A* **167**: 708–714. <http://doi.wiley.com/10.1002/ajmg.a.36950> (Accessed November 24, 2020).
- Provenzano C, Cappella M, Valaperta R, Cardani R, Meola G, Martelli F, Cardinali B, Falcone G. 2017. CRISPR/Cas9-Mediated Deletion of CTG Expansions Recovers Normal Phenotype in Myogenic Cells Derived from Myotonic Dystrophy 1 Patients. *Mol Ther - Nucleic Acids* **9**: 337–348. <https://doi.org/10.1016/j.omtn.2017.10.006>.
- Pulecio J, Verma N, Mejía-Ramírez E, Huangfu D, Raya A. 2017. CRISPR/Cas9-Based Engineering of the Epigenome. *Cell Stem Cell* **21**: 431–447. <http://www.ncbi.nlm.nih.gov/pubmed/28985525> (Accessed November 23, 2020).
- Qawasmi L, Braun M, Guberman I, Cohen E, Naddaf L, Mellul A, Matilainen O, Roitenberg N, Share D, Stupp D, et al. 2019. Expanded CUG Repeats Trigger Disease Phenotype and Expression Changes through the RNAi Machinery in *C. elegans*. *J Mol Biol* **431**: 1711–1728. <http://www.ncbi.nlm.nih.gov/pubmed/30878478> (Accessed December 1, 2020).
- Querido E, Gallardo F, Beaudoin M, Ménard C, Chartrand P. 2011. Stochastic and reversible aggregation of mRNA with expanded CUG-triplet repeats. *J Cell Sci* **124**. <http://jcs.biologists.org.gate2.inist.fr/content/joces/124/10/1703.full.pdf> (Accessed April 13,

2018).

- Raaijmakers RHL, Ripken L, Ausems CRM, Wansink DG. 2019. CRISPR/Cas Applications in Myotonic Dystrophy: Expanding Opportunities. *Int J Mol Sci* **20**: 3689. <https://www.mdpi.com/1422-0067/20/15/3689> (Accessed November 23, 2020).
- Ran FA, Hsu PD, Lin C-Y, Gootenberg JS, Konermann S, Trevino A, Scott DA, Inoue A, Matoba S, Zhang Y, et al. 2013. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**: 1380. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3856256/> (Accessed November 23, 2020).
- Ranum LP, Rasmussen PF, Benzow KA, Koob MD, Day JW. 1998. Genetic mapping of a second myotonic dystrophy locus. *Nat Genet* **19**: 196–8. <http://www.ncbi.nlm.nih.gov/pubmed/9620781> (Accessed November 23, 2020).
- Rau F, Freyermuth F, Fugier C, Villemin J-P, Fischer M-C, Jost B, Dembele D, Gourdon G, Nicole A, Duboc D, et al. 2011. Misregulation of miR-1 processing is associated with heart defects in myotonic dystrophy. *Nat Struct Mol Biol* **18**: 840–846. <http://www.nature.com/gate2.inist.fr/articles/nsmb.2067.pdf> (Accessed April 18, 2018).
- Rau F, Lainé J, Ramanoudjame L, Ferry A, Arandel L, Delalande O, Jollet A, Dingli F, Lee KY, Peccate C, et al. 2015. Abnormal splicing switch of DMD's penultimate exon compromises muscle fibre maintenance in myotonic dystrophy. *Nat Commun* **6**.
- Ravel-Chapuis A, Bélanger G, Yadava RS, Mahadevan MS, DesGroseillers L, Côté J, Jasmin BJ. 2012. The RNA-binding protein Staufen1 is increased in DM1 skeletal muscle and promotes alternative pre-mRNA splicing. *J Cell Biol* **196**: 699–712. <http://www.ncbi.nlm.nih.gov/pubmed/22431750> (Accessed February 11, 2021).
- Reddy K, Jenquin JR, Cleary JD, Berglund JA. 2019. Mitigating RNA Toxicity in Myotonic Dystrophy using Small Molecules. *Int J Mol Sci* **20**. <http://www.ncbi.nlm.nih.gov/pubmed/31426500> (Accessed December 3, 2020).
- Reddy S, Smith DB, Rich MM, Leferovich JM, Reilly P, Davis BM, Tran K, Rayburn H, Bronson R, Cros D, et al. 1996. Mice lacking the myotonic dystrophy protein kinase develop a late onset progressive myopathy. *Nat Genet* **13**: 325–35. <http://www.ncbi.nlm.nih.gov/pubmed/8673132> (Accessed November 26, 2020).
- Rees HA, Liu DR. 2018. Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat Rev Genet* **19**: 770. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6535181/> (Accessed November 23, 2020).
- Renault V, Piron-Hamelin G, Forestier C, DiDonna S, Decary S, Hentati F, Saillant G, Butler-Browne GS, Mouly V. 2000. Skeletal muscle regeneration and the mitotic clock. *Exp Gerontol* **35**: 711–9. <http://www.ncbi.nlm.nih.gov/pubmed/11053661> (Accessed November 30, 2020).
- Ribon V, Herrera R, Kay BK, Saltiel AR. 1998. A role for CAP, a novel, multifunctional Src homology 3 domain-containing protein in formation of actin stress fibers and focal adhesions. *J Biol Chem* **273**: 4073–80. <http://www.ncbi.nlm.nih.gov/pubmed/9461600> (Accessed February 8, 2021).
- Ricker K, Koch MC, Lehmann-Horn F, Pongratz D, Otto M, Heine R, Moxley RT. 1994. Proximal myotonic myopathy: a new dominant disorder with myotonia, muscle weakness, and cataracts. *Neurology* **44**: 1448–52. <http://www.ncbi.nlm.nih.gov/pubmed/8058147> (Accessed November 23, 2020).
- Rivera T, Zhao Y, Ni Y, Wang J. 2020. Human-Induced Pluripotent Stem Cell Culture Methods Under cGMP Conditions. *Curr Protoc Stem Cell Biol* **54**. <https://onlinelibrary.wiley.com/doi/10.1002/cpsc.117> (Accessed January 13, 2021).
- Roberts R, Timchenko NA, Miller JW, Reddy S, Caskey CT, Swanson MS, Timchenko LT. 1997. Altered phosphorylation and intracellular distribution of a (CUG)_n triplet repeat RNA-binding protein in patients with myotonic dystrophy and in myotonin protein kinase knockout mice. *Proc Natl Acad Sci U S A* **94**: 13221–6. <http://www.ncbi.nlm.nih.gov/pubmed/9371827> (Accessed

November 29, 2020).

- Rowe RG, Daley GQ. 2019. Induced pluripotent stem cells in disease modelling and drug discovery. *Nat Rev Genet* **20**: 377. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6584039/> (Accessed November 22, 2020).
- Runfola V, Sebastian S, Dilworth FJ, Gabellini D. 2015. Rbfox proteins regulate tissue-specific alternative splicing of Mef2D required for muscle differentiation. *J Cell Sci* **128**: 631–7. <http://www.ncbi.nlm.nih.gov/pubmed/25609712> (Accessed October 22, 2020).
- Sakurai H, Sakaguchi Y, Shoji E, Nishino T, Maki I, Sakai H, Hanaoka K, Kakizuka A, Sehara-Fujisawa A. 2012. In Vitro Modeling of Paraxial Mesodermal Progenitors Derived from Induced Pluripotent Stem Cells ed. Q. Xiao. *PLoS One* **7**: e47078. <https://dx.plos.org/10.1371/journal.pone.0047078> (Accessed November 23, 2020).
- Sampath SC, Millay DP. 2018. Myoblast fusion confusion: the resolution begins. *Skelet Muscle* **8**: 3. <http://www.ncbi.nlm.nih.gov/pubmed/29386054> (Accessed October 22, 2020).
- Sanjana NE, Cong L, Zhou Y, Cunniff MM, Feng G, Zhang F. 2012. A transcription activator-like effector toolbox for genome engineering. *Nat Protoc* **7**: 171–92. <http://www.ncbi.nlm.nih.gov/pubmed/22222791> (Accessed November 23, 2020).
- Sarkar PS, Han J, Reddy S. 2004. In situ hybridization analysis of Dmpk mRNA in adult mouse tissues. *Neuromuscul Disord* **14**: 497–506. https://www-sciencedirect-com.gate2.inist.fr/science/article/pii/S0960896604000999?_rdoc=1&_fmt=high&_origin=gateway&_docanchor=&md5=b8429449ccfc9c30159a5f9aeaa92ffb&ccp=y (Accessed July 24, 2018).
- Saunders A, Faiola F, Wang J. 2013. Concise review: pursuing self-renewal and pluripotency with the stem cell factor Nanog. *Stem Cells* **31**: 1227–36. <http://www.ncbi.nlm.nih.gov/pubmed/23653415> (Accessed November 22, 2020).
- Savić Pavićević D, Miladinović J, Brkušanić M, Šviković S, Djurica S, Brajušković G, Romac S. 2013. Molecular Genetics and Genetic Testing in Myotonic Dystrophy Type 1. *Biomed Res Int* **2013**: 1–13. <http://www.hindawi.com/journals/bmri/2013/391821/> (Accessed July 31, 2020).
- Savkur RS, Philips A V, Cooper TA. 2001. Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy. *Nat Genet* **29**: 40–7. <http://www.ncbi.nlm.nih.gov/pubmed/11528389> (Accessed November 29, 2020).
- Savouret C, Brisson E, Essers J, Kanaar R, Pastink A, te Riele H, Junien C, Gourdon G. 2003. CTG repeat instability and size variation timing in DNA repair-deficient mice. *EMBO J* **22**: 2264–73. <http://www.ncbi.nlm.nih.gov/pubmed/12727892> (Accessed November 25, 2020).
- Schoner W, Scheiner-Bobis G. 2007. Endogenous and exogenous cardiac glycosides: their roles in hypertension, salt metabolism, and cell growth. *Am J Physiol Physiol* **293**: C509–C536. <https://www.physiology.org/doi/10.1152/ajpcell.00098.2007> (Accessed November 23, 2020).
- Schoer B, Timchenko L. 2010. Myotonic dystrophies 1 and 2: complex diseases with complex mechanisms. *Curr Genomics* **11**: 77–90. <http://www.ncbi.nlm.nih.gov/pubmed/20885816> (Accessed November 23, 2020).
- Scoles DR, Minikel E V, Pulst SM. 2019. Antisense oligonucleotides: A primer. *Neurol Genet* **5**: e323. <http://www.ncbi.nlm.nih.gov/pubmed/31119194> (Accessed December 3, 2020).
- Scrudato M Lo, Poulard K, Sourd C, Tomé S, Klein AF, Corre G, Huguet A, Furling D, Gourdon G, Buj-Bello A. 2019. Genome Editing of Expanded CTG Repeats within the Human DMPK Gene Reduces Nuclear RNA Foci in the Muscle of DM1 Mice. *Mol Ther* **27**: 1372. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6697452/> (Accessed November 23, 2020).
- Shadiack AM, Nitkin RM. 1991. Agrin induces α -actinin, filamin, and vinculin to co-localize with AChR clusters on cultured chick myotubes. *J Neurobiol* **22**: 617–628.
- Shamblott MJ, Axelman J, Wang S, Bugg EM, Littlefield JW, Donovan PJ, Blumenthal PD, Huggins GR, Gearhart JD. 1998. Derivation of pluripotent stem cells from cultured human primordial germ

- cells. *Proc Natl Acad Sci U S A* **95**: 13726–31. <http://www.ncbi.nlm.nih.gov/pubmed/9811868> (Accessed November 22, 2020).
- Shelton M, Metz J, Liu J, Carpenedo RL, Demers S-P, Stanford WL, Skerjanc IS. 2014. Derivation and expansion of PAX7-positive muscle progenitors from human and mouse embryonic stem cells. *Stem cell reports* **3**: 516–29. <http://www.ncbi.nlm.nih.gov/pubmed/25241748> (Accessed November 23, 2020).
- Shmakov S, Smargon A, Scott D, Cox D, Pyzocha N, Yan W, Abudayyeh OO, Gootenberg JS, Makarova KS, Wolf YI, et al. 2017. Diversity and evolution of class 2 CRISPR-Cas systems. *Nat Rev Microbiol* **15**: 169–182. <http://www.ncbi.nlm.nih.gov/pubmed/28111461> (Accessed November 23, 2020).
- Siboni RB, Nakamori M, Wagner SD, Struck AJ, Coonrod LA, Harriott SA, Cass DM, Tanner MK, Berglund JA. 2015. Actinomycin D Specifically Reduces Expanded CUG Repeat RNA in Myotonic Dystrophy Models. *Cell Rep* **13**: 2386–2394. <http://dx.doi.org/10.1016/j.celrep.2015.11.028>.
- Sicot G, Gomes-Pereira M. 2013. RNA toxicity in human disease and animal models: From the uncovering of a new mechanism to the development of promising therapies. *Biochim Biophys Acta - Mol Basis Dis* **1832**: 1390–1409. <https://www.sciencedirect.com/science/article/pii/S0925443913000768?via%3Dihub> (Accessed November 27, 2020).
- Sicot G, Gourdon G, Gomes-Pereira M. 2011. Myotonic dystrophy, when simple repeats reveal complex pathogenic entities: new findings and future challenges. *Hum Mol Genet* **20**: R116–R123. <https://academic.oup.com/hmg/article-lookup/doi/10.1093/hmg/ddr343> (Accessed January 8, 2019).
- Singh AM, Reynolds D, Cliff T, Ohtsuka S, Mattheyses AL, Sun Y, Menendez L, Kulik M, Dalton S. 2012. Signaling network crosstalk in human pluripotent cells: a Smad2/3-regulated switch that controls the balance between self-renewal and differentiation. *Cell Stem Cell* **10**: 312–26. <http://www.ncbi.nlm.nih.gov/pubmed/22385658> (Accessed November 22, 2020).
- Singh G, Charlet-B N, Han J, Cooper TA. 2004. ETR-3 and CELF4 protein domains required for RNA binding and splicing activity in vivo. *Nucleic Acids Res* **32**: 1232–41. <http://www.ncbi.nlm.nih.gov/pubmed/14973222> (Accessed November 28, 2020).
- Singhal N, Martin PT. 2011. Role of extracellular matrix proteins and their receptors in the development of the vertebrate neuromuscular junction. *Dev Neurobiol* **71**: 982–1005. <http://www.ncbi.nlm.nih.gov/pubmed/21766463> (Accessed February 8, 2021).
- Skoglund G, Lainé J, Darabi R, Fournier E, Perlingeiro R, Tabti N. 2014. Physiological and ultrastructural features of human induced pluripotent and embryonic stem cell-derived skeletal myocytes in vitro. *Proc Natl Acad Sci U S A* **111**: 8275–80. <http://www.ncbi.nlm.nih.gov/pubmed/24843168> (Accessed January 28, 2021).
- Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang F. 2016. Rationally engineered Cas9 nucleases with improved specificity. *Science (80-)* **351**: 84–88. <https://science.sciencemag.org/content/351/6268/84> (Accessed January 28, 2021).
- Sobczak K, Wheeler TM, Wang W, Thornton CA. 2013. RNA interference targeting CUG repeats in a mouse model of myotonic dystrophy. *Mol Ther* **21**: 380–387. <http://dx.doi.org/10.1038/mt.2012.222>.
- Souidi A, Zmojdian M, Jagla K. 2018. Dissecting Pathogenetic Mechanisms and Therapeutic Strategies in Drosophila Models of Myotonic Dystrophy Type 1. *Int J Mol Sci* **19**. <http://www.ncbi.nlm.nih.gov/pubmed/30567354> (Accessed December 1, 2020).
- Staerk J, Dawlaty MM, Gao Q, Maetzel D, Hanna J, Sommer CA, Mostoslavsky G, Jaenisch R. 2010. Reprogramming of human peripheral blood cells to induced pluripotent stem cells. *Cell Stem Cell* **7**: 20–4. <http://www.ncbi.nlm.nih.gov/pubmed/20621045> (Accessed November 22, 2020).
- Suenaga K, Lee K-Y, Nakamori M, Tatsumi Y, Takahashi MP, Fujimura H, Jinnai K, Yoshikawa H, Du H, Ares M, et al. 2012. Muscleblind-Like 1 Knockout Mice Reveal Novel Splicing Defects in the

- Myotonic Dystrophy Brain ed. T.H. Gillingwater. *PLoS One* **7**: e33218. <http://dx.plos.org/10.1371/journal.pone.0033218> (Accessed October 17, 2018).
- Suwinska A, Ciemerych MA. 2011. Factors Regulating Pluripotency and Differentiation in Early Mammalian Embryos and Embryo-derived Stem Cells. *Vitam Horm* **87**: 1–37. <https://www.sciencedirect.com/science/article/pii/B9780123860156000226?via%3Dihub> (Accessed November 22, 2020).
- Sznajder LJ, Michalak M, Taylor K, Cywoniuk P, Kabza M, Wojtkowiak-Szlachcic A, Matłoka M, Konieczny P, Sobczak K. 2016. Mechanistic determinants of MBNL activity. *Nucleic Acids Res* **44**: 10326–10342.
- Sznajder LJ, Swanson MS. 2019. Short Tandem Repeat Expansions and RNA-Mediated Pathogenesis in Myotonic Dystrophy. *Int J Mol Sci* **20**. <http://www.ncbi.nlm.nih.gov/pubmed/31323950> (Accessed February 11, 2021).
- Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S. 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**: 861–72. <http://www.ncbi.nlm.nih.gov/pubmed/18035408> (Accessed November 22, 2020).
- Takahashi K, Yamanaka S. 2016. A decade of transcription factor-mediated reprogramming to pluripotency. *Nat Rev Mol Cell Biol* **17**: 183–193. <http://www.nature.com/articles/nrm.2016.8> (Accessed November 22, 2020).
- Takahashi K, Yamanaka S. 2006. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* **126**: 663–676. [https://www.cell.com/cell/pdf/S0092-8674\(06\)00976-7.pdf](https://www.cell.com/cell/pdf/S0092-8674(06)00976-7.pdf) (Accessed June 12, 2018).
- Talukdar I, Sen S, Urbano R, Thompson J, Yates JR, Webster NJG. 2011. hnRNP A1 and hnRNP F Modulate the Alternative Splicing of Exon 11 of the Insulin Receptor Gene ed. Z.-M. Zheng. *PLoS One* **6**: e27869. <https://dx.plos.org/10.1371/journal.pone.0027869> (Accessed November 27, 2020).
- Taneja KL, McCurrach M, Schalling M, Housman D, Singer RH. 1995. Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. *J Cell Biol* **128**: 995–1002. <http://www.ncbi.nlm.nih.gov/pubmed/7896884> (Accessed November 26, 2020).
- Tang ZZ, Yarotsky V, Wei L, Sobczak K, Nakamori M, Eichinger K, Moxley RT, Dirksen RT, Thornton CA. 2012. Muscle weakness in myotonic dystrophy associated with misregulated splicing and altered gating of CaV1.1 calcium channel. *Hum Mol Genet* **21**: 1312–1324. <https://academic.oup.com/hmg/article-lookup/doi/10.1093/hmg/ddr568> (Accessed September 19, 2018).
- Tannenbaum SE, Tako Turetsky T, Singer O, Aizenman E, Kirshberg S, Ilouz N, Gil Y, Berman-Zaken Y, Perlman TS, Geva N, et al. 2012. Derivation of Xeno-Free and GMP-Grade Human Embryonic Stem Cells – Platforms for Future Clinical Applications ed. M. Pera. *PLoS One* **7**: e35325. <https://dx.plos.org/10.1371/journal.pone.0035325> (Accessed November 22, 2020).
- Taylor K, Sznajder LJ, Cywoniuk P, Thomas JD, Swanson MS, Sobczak K. 2018. MBNL splicing activity depends on RNA binding site structural context. *Nucleic Acids Res* **46**: 9119–9133. <http://www.ncbi.nlm.nih.gov/pubmed/29955876> (Accessed December 4, 2020).
- Teplova M, Patel DJ. 2008. Structural insights into RNA recognition by the alternative-splicing regulator muscleblind-like MBNL1. *Nat Struct Mol Biol* **15**: 1343–1351. <http://www.nature.com/gate2.inist.fr/articles/nsmb.1519.pdf> (Accessed April 18, 2018).
- Teplova M, Song J, Gaw HY, Teplov A, Patel DJ. 2010. Structural Insights into RNA Recognition by the Alternate-Splicing Regulator CUG-Binding Protein 1. *Structure* **18**: 1364–1377. <https://www.sciencedirect.com/science/article/pii/S0969212610003059> (Accessed November 29, 2020).
- Terenzi F, Ladd AN. 2010. Conserved developmental alternative splicing of muscleblind-like (MBNL) transcripts regulates MBNL localization and activity. *RNA Biol* **7**: 43–55.

www.landesbioscience.com (Accessed April 18, 2018).

- Theadom A, Rodrigues M, Roxburgh R, Balalla S, Higgins C, Bhattacharjee R, Jones K, Krishnamurthi R, Feigin V. 2014. Prevalence of Muscular Dystrophies: A Systematic Literature Review. *Neuroepidemiology* **43**: 259–268. www.karger.com/ned (Accessed November 24, 2020).
- Thomas JD, Oliveira R, Sznajder ŁJ, Swanson MS. 2018. Myotonic Dystrophy and Developmental Regulation of RNA Processing. In *Comprehensive Physiology*, Vol. 8 of, pp. 509–553, John Wiley & Sons, Inc., Hoboken, NJ, USA <http://doi.wiley.com/10.1002/cphy.c170002> (Accessed November 23, 2020).
- Thomas JD, Sznajder ŁJ, Bardhi O, Aslam FN, Anastasiadis ZP, Scotti MM, Nishino I, Nakamori M, Wang ET, Swanson MS. 2017. Disrupted prenatal RNA processing and myogenesis in congenital myotonic dystrophy. *Genes Dev* **31**: 1122–1133. <http://genesdev.cshlp.org/content/31/11/1122.full.pdf> (Accessed May 29, 2018).
- Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM. 1998. Embryonic Stem Cell Lines Derived from Human Blastocysts. *Science (80-)* **282**: 1145–1147. <https://science.sciencemag.org/content/282/5391/1145.long> (Accessed November 22, 2020).
- Thornton CA, Griggs RC, Moxley RT. 1994. Myotonic dystrophy with no trinucleotide repeat expansion. *Ann Neurol* **35**: 269–272. <http://doi.wiley.com/10.1002/ana.410350305> (Accessed November 23, 2020).
- Thornton CA, Wymer JP, Simmons Z, McClain C, Moxley RT. 1997. Expansion of the myotonic dystrophy CTG repeat reduces expression of the flanking DMAHP gene. *Nat Genet* **16**: 407–409. <http://www.nature.com/articles/ng0897-407> (Accessed November 26, 2020).
- Tian B, Manley JL. 2013. Alternative cleavage and polyadenylation: the long and short of it. *Trends Biochem Sci* **38**: 312–20. <http://www.ncbi.nlm.nih.gov/pubmed/23632313> (Accessed December 8, 2020).
- Tiburcy M, Hudson JE, Balfanz P, Schlick S, Meyer T, Chang Liao M-L, Levent E, Raad F, Zeidler S, Wingender E, et al. 2017. Defined Engineered Human Myocardium With Advanced Maturation for Applications in Heart Failure Modeling and Repair. *Circulation* **135**: 1832–1847. <http://www.ncbi.nlm.nih.gov/pubmed/28167635> (Accessed January 28, 2021).
- Timchenko LT, Timchenko NA, Caskey CT, Roberts R. 1996. Novel proteins with binding specificity for DNA CTG repeats and RNA CUG repeats: implications for myotonic dystrophy. *Hum Mol Genet* **5**: 115–21. <http://www.ncbi.nlm.nih.gov/pubmed/8789448> (Accessed November 28, 2020).
- Tintignac LA, Brenner H-R, Rüegg MA. 2015. Mechanisms Regulating Neuromuscular Junction Development and Function and Causes of Muscle Wasting. *Physiol Rev* **95**: 809–852. <https://www.physiology.org/doi/10.1152/physrev.00033.2014> (Accessed February 3, 2021).
- Todd PK, Ackall FY, Hur J, Sharma K, Paulson HL, Dowling JJ. 2014. Transcriptional changes and developmental abnormalities in a zebrafish model of myotonic dystrophy type 1. *Dis Model Mech* **7**: 143–55. <http://www.ncbi.nlm.nih.gov/pubmed/24092878> (Accessed December 1, 2020).
- Tomasovic A, Kurrle N, Banning A, Tikkanen R. 2012. *Role of Cbl-associated protein/ponsin in receptor tyrosine kinase signaling and cell adhesion*. <https://www.jmolbiochem.com/index.php/JmolBiochem/article/view/60/35> (Accessed January 21, 2021).
- Tomé S, Gourdon G. 2020. DM1 Phenotype Variability and Triplet Repeat Instability: Challenges in the Development of New Therapies. *Int J Mol Sci* **21**: 457. <https://www.mdpi.com/1422-0067/21/2/457> (Accessed January 21, 2020).
- Tosoni D, Cestra G. 2009. CAP (Cbl associated protein) regulates receptor-mediated endocytosis. *FEBS Lett* **583**: 293–300. <http://doi.wiley.com/10.1016/j.febslet.2008.12.047> (Accessed January 21, 2021).
- Tran H, Gourrier N, Lemercier-Neuillet C, Dhaenens C-M, Vautrin A, Fernandez-Gomez FJ, Arandel L,

- Carpentier C, Obriot H, Eddarkaoui S, et al. 2011. Analysis of Exonic Regions Involved in Nuclear Localization, Splicing Activity, and Dimerization of Muscleblind-like-1 Isoforms. *J Biol Chem* **286**. <http://www.jbc.org.gate2.inist.fr/content/286/18/16435.full.pdf?with-ds=yes> (Accessed April 13, 2018).
- Truebestein L, Leonard TA. 2016. Coiled-coils: The long and short of it. *Bioessays* **38**: 903–16. <http://www.ncbi.nlm.nih.gov/pubmed/27492088> (Accessed January 24, 2021).
- Turner C, Hilton-Jones D. 2014. Myotonic dystrophy: diagnosis, management and new therapies. *Curr Opin Neurol* **27**: 599–606. <http://www.ncbi.nlm.nih.gov/pubmed/25121518> (Accessed February 11, 2021).
- Turner C, Hilton-Jones D. 2010. The myotonic dystrophies: diagnosis and management. *J Neurol Neurosurg Psychiatry* **81**: 358–67. <http://www.ncbi.nlm.nih.gov/pubmed/20176601> (Accessed November 23, 2020).
- Uddin F, Rudin CM, Sen T. 2020. CRISPR Gene Therapy: Applications, Limitations, and Implications for the Future. *Front Oncol* **10**: 1387. <http://www.ncbi.nlm.nih.gov/pubmed/32850447> (Accessed January 28, 2021).
- Ueki J, Nakamori M, Nakamura M, Nishikawa M, Yoshida Y, Tanaka A, Morizane A, Kamon M, Araki T, Takahashi MP, et al. 2017. Myotonic dystrophy type 1 patient-derived iPSCs for the investigation of CTG repeat instability. *Sci Rep* **7**: 1–7.
- Ule J, Blencowe BJ. 2019. Molecular Cell Review Alternative Splicing Regulatory Networks: Functions, Mechanisms, and Evolution. *Mol Cell* **76**. <https://doi.org/10.1016/j.molcel.2019.09.017> (Accessed January 12, 2021).
- Vallier L, Mendjan S, Brown S, Chng Z, Teo A, Smithers LE, Trotter MWB, Cho CH-H, Martinez A, Rugg-Gunn P, et al. 2009. Activin/Nodal signalling maintains pluripotency by controlling Nanog expression. *Development* **136**: 1339–49. <http://www.ncbi.nlm.nih.gov/pubmed/19279133> (Accessed November 22, 2020).
- van Agtmaal EL, André LM, Willemse M, Cumming SA, van Kessel IDG, van den Broek WJAA, Gourdon G, Furling D, Mouly V, Monckton DG, et al. 2017. CRISPR/Cas9-Induced (CTG·CAG)_n Repeat Instability in the Myotonic Dystrophy Type 1 Locus: Implications for Therapeutic Genome Editing. *Mol Ther* **25**: 24–43. <http://www.ncbi.nlm.nih.gov/pubmed/28129118> (Accessed November 23, 2020).
- van Cruchten RTP, Wieringa B, Wansink DG. 2019. Expanded CUG repeats in DMPK transcripts adopt diverse hairpin conformations without influencing the structure of the flanking sequences. *RNA* **25**: 481–495. <http://www.ncbi.nlm.nih.gov/pubmed/30700578> (Accessed November 26, 2020).
- van den Broek WJAA, Nelen MR, Wansink DG, Coerwinkel MM, te Riele H, Groenen PJTA, Wieringa B. 2002. Somatic expansion behaviour of the (CTG)_n repeat in myotonic dystrophy knock-in mice is differentially affected by Msh3 and Msh6 mismatch-repair proteins. *Hum Mol Genet* **11**: 191–8. <http://www.ncbi.nlm.nih.gov/pubmed/11809728> (Accessed November 25, 2020).
- Waddington CH. 1957. *The Strategy of the Genes*. Routledge <https://www.taylorfrancis.com/books/9781315765471> (Accessed January 13, 2021).
- Wagner SD, Struck AJ, Gupta R, Farnsworth DR, Mahady AE, Eichinger K, Thornton CA, Wang ET, Berglund JA. 2016. Dose-Dependent Regulation of Alternative Splicing by MBNL Proteins Reveals Biomarkers for Myotonic Dystrophy. *PLoS Genet* **12**: 1–24.
- Wakimoto H, Maguire CT, Sherwood MC, Vargas MM, Sarkar PS, Han J, Reddy S, Berul CI. 2002. Characterization of Cardiac Conduction System Abnormalities in Mice with Targeted Disruption of Six5 Gene. *J Interv Card Electrophysiol* **7**: 127–135. <http://link.springer.com/10.1023/A:1020881520353> (Accessed November 30, 2020).
- Wang ET, Cody NAL, Jog S, Biancolella M, Wang TT, Treacy DJ, Luo S, Schroth GP, Housman DE, Reddy S, et al. 2012. Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. *Cell* **150**: 710–724.

- Wang ET, Sandberg R, Luo S, Khrebtkova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**: 470–476.
- Wang ET, Treacy D, Eichinger K, Struck A, Estabrook J, Olafson H, Wang TT, Bhatt K, Westbrook T, Sedehizadeh S, et al. 2019a. Transcriptome alterations in myotonic dystrophy skeletal muscle and heart. *Hum Mol Genet* **28**: 1312–1321. <http://www.ncbi.nlm.nih.gov/pubmed/30561649> (Accessed October 22, 2020).
- Wang ET, Ward AJ, Cherone JM, Giudice J, Wang TT, Treacy DJ, Lambert NJ, Freese P, Saxena T, Cooper TA, et al. 2015. Antagonistic regulation of mRNA expression and splicing by CELF and MBNL proteins. *Genome Res* **25**: 858–871. <https://genome-cshlp-org.gate2.inist.fr/content/25/6/858.full.pdf> (Accessed April 18, 2018).
- Wang G-S, Kearney DL, De Biasi M, Taffet G, Cooper TA. 2007. Elevation of RNA-binding protein CUGBP1 is an early event in an inducible heart-specific mouse model of myotonic dystrophy. *J Clin Invest* **117**: 2802–11. <http://www.ncbi.nlm.nih.gov/pubmed/17823658> (Accessed November 29, 2020).
- Wang M, Weng W-C, Stock L, Lindquist D, Martinez A, Gourdon G, Timchenko N, Snape M, Timchenko L. 2019b. Correction of Glycogen Synthase Kinase 3 β in Myotonic Dystrophy 1 Reduces the Mutant RNA and Improves Postnatal Survival of DMSXL Mice. *Mol Cell Biol* **39**. <http://www.ncbi.nlm.nih.gov/pubmed/31383751> (Accessed December 3, 2020).
- Wang PY, Lin YM, Wang LH, Kuo TY, Cheng SJ, Wang GS. 2017. Reduced cytoplasmic MBNL1 is an early event in a brain-specific mouse model of myotonic dystrophy. *Hum Mol Genet* **26**: 2247–2257.
- Wang Y, Hao L, Wang H, Santostefano K, Thapa A, Cleary J, Li H, Guo X, Terada N, Ashizawa T, et al. 2018a. Therapeutic Genome Editing for Myotonic Dystrophy Type 1 Using CRISPR/Cas9. *Mol Ther* **0**. <http://www.ncbi.nlm.nih.gov/pubmed/30274788> (Accessed October 11, 2018).
- Wang Y, Wang Z, Sun H, Shi C, Yang J, Liu Y, Liu H, Zhang S, Zhang L, Xu Y, et al. 2018b. Generation of induced pluripotent stem cell line(ZZUi006-A)from a patient with myotonic dystrophy type 1. *Stem Cell Res* **32**: 61–64. <https://www.sciencedirect.com/science/article/pii/S1873506118302058> (Accessed October 26, 2020).
- Wansink DG, van Herpen REMA, Coerwinkel-Driessen MM, Groenen PJTA, Hemmings BA, Wieringa B. 2003. Alternative splicing controls myotonic dystrophy protein kinase structure, enzymatic activity, and subcellular localization. *Mol Cell Biol* **23**: 5489–501. <http://www.ncbi.nlm.nih.gov/pubmed/12897125> (Accessed November 25, 2020).
- Ward AJ, Rimer M, Killian JM, Dowling JJ, Cooper TA. 2010. CUGBP1 overexpression in mouse skeletal muscle reproduces features of myotonic dystrophy type 1. *Hum Mol Genet* **19**: 3614–22. <http://www.ncbi.nlm.nih.gov/pubmed/20603324> (Accessed November 29, 2020).
- Warf MB, Diegel J V, von Hippel PH, Berglund JA. 2009. The protein factors MBNL1 and U2AF65 bind alternative RNA structures to regulate splicing. *Proc Natl Acad Sci U S A* **106**: 9203–8. <http://www.ncbi.nlm.nih.gov/pubmed/19470458> (Accessed December 8, 2020).
- Watanabe K, Ueno M, Kamiya D, Nishiyama A, Matsumura M, Wataya T, Takahashi JB, Nishikawa S, Nishikawa S, Muguruma K, et al. 2007. A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nat Biotechnol* **25**: 681–686. <http://www.nature.com/articles/nbt1310> (Accessed November 22, 2020).
- Weninger S, Montagnese F, Schoser B. 2018. Core Clinical Phenotypes in Myotonic Dystrophies. *Front Neurol* **9**: 303. <http://www.ncbi.nlm.nih.gov/pubmed/29770119> (Accessed November 23, 2020).
- Westerlaken JHAM, Van der Zee CEEM, Peters W, Wieringa B. 2003. The DMWD protein from the myotonic dystrophy (DM1) gene region is developmentally regulated and is present most prominently in synapse-dense brain areas. *Brain Res* **971**: 116–27.

- <http://www.ncbi.nlm.nih.gov/pubmed/12691844> (Accessed November 26, 2020).
- Weyn-Vanhenteryck SM, Feng H, Ustianenko D, Duffié R, Yan Q, Jacko M, Martinez JC, Goodwin M, Zhang X, Hengst U, et al. 2018. Precise temporal regulation of alternative splicing during neural development. *Nat Commun* **9**: 2189. <http://www.nature.com/articles/s41467-018-04559-0> (Accessed May 7, 2020).
- Wheeler TM, Krym MC, Thornton CA. 2007a. Ribonuclear foci at the neuromuscular junction in myotonic dystrophy type 1. *Neuromuscul Disord* **17**: 242–7. <http://www.ncbi.nlm.nih.gov/pubmed/17306536> (Accessed July 20, 2018).
- Wheeler TM, Leger AJ, Pandey SK, MacLeod AR, Nakamori M, Cheng SH, Wentworth BM, Bennett CF, Thornton CA. 2012. Targeting nuclear RNA for in vivo correction of myotonic dystrophy. *Nature* **488**: 111–5. <http://www.ncbi.nlm.nih.gov/pubmed/22859208> (Accessed December 3, 2020).
- Wheeler TM, Lueck JD, Swanson MS, Dirksen RT, Thornton CA. 2007b. Correction of CIC-1 splicing eliminates chloride channelopathy and myotonia in mouse models of myotonic dystrophy. *J Clin Invest* **117**: 3952–7. <http://www.ncbi.nlm.nih.gov/pubmed/18008009> (Accessed December 3, 2020).
- Wheeler TM, Sobczak K, Lueck JD, Osborne RJ, Lin X, Dirksen RT, Thornton CA. 2009. Reversal of RNA dominance by displacement of protein sequestered on triplet repeat RNA. *Science* **325**: 336–9. <http://www.ncbi.nlm.nih.gov/pubmed/19608921> (Accessed December 3, 2020).
- Wiedenheft B, Sternberg SH, Doudna JA. 2012. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **482**: 331–338. <http://www.nature.com/articles/nature10886> (Accessed November 23, 2020).
- Williams LA, Davis-Dusenbery BN, Eggan KC. 2012. SnapShot: directed differentiation of pluripotent stem cells. *Cell* **149**: 1174–1174.e1. <http://www.ncbi.nlm.nih.gov/pubmed/22632979> (Accessed January 13, 2021).
- Wohrle D, Kennerknecht I, Wolf M, Enders H, Schwemmle S, Steinbach P. 1995. Heterogeneity of DM kinase repeat expansion in different fetal tissues and further expansion during cell proliferation in vitro: evidence for a causal involvement of methyl-directed DNA mismatch repair in triplet repeat stability. *Hum Mol Genet* **4**: 1147–1153. <https://academic.oup.com/hmg/article-lookup/doi/10.1093/hmg/4.7.1147> (Accessed November 25, 2020).
- Wood L, Bassez G, Bleyenheuft C, Campbell C, Cossette L, Jimenez-Moreno AC, Dai Y, Dawkins H, Manera JAD, Dogan C, et al. 2018. Eight years after an international workshop on myotonic dystrophy patient registries: case study of a global collaboration for a rare disease. *Orphanet J Rare Dis* **13**: 155. <https://ojrd.biomedcentral.com/articles/10.1186/s13023-018-0889-0> (Accessed November 24, 2020).
- Wu S-S, Li Q-C, Yin C-Q, Xue W, Song C-Q. 2020. Advances in CRISPR/Cas-based Gene Therapy in Human Genetic Diseases. *Theranostics* **10**: 4374. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7150498/> (Accessed November 23, 2020).
- Xia G, Ashizawa T. 2015. Dynamic changes of nuclear RNA foci in proliferating DM1 cells. *Histochem Cell Biol* **143**: 557–64. <http://www.ncbi.nlm.nih.gov/pubmed/25715678> (Accessed November 26, 2020).
- Xia G, Santostefano KE, Goodwin M, Liu J, Subramony SH, Swanson MS, Terada N, Ashizawa T. 2013. Generation of Neural Cells from DM1 Induced Pluripotent Stem Cells As Cellular Model for the Study of Central Nervous System Neuropathogenesis. *Cell Reprogram* **15**: 166. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3616452/> (Accessed November 23, 2020).
- Xing G, Jing H, Zhang L, Cao Y, Li L, Zhao K, Dong Z, Chen W, Wang H, Cao R, et al. 2019. A mechanism in agrin signaling revealed by a prevalent Rapsyn mutation in congenital myasthenic syndrome. *Elife* **8**. <https://elifesciences.org/articles/49180> (Accessed January 21, 2021).
- Xing G, Xiong W-C, Mei L. 2020. Rapsyn as a signaling and scaffolding molecule in neuromuscular junction formation and maintenance. *Neurosci Lett* **731**: 135013.

- <http://www.ncbi.nlm.nih.gov/pubmed/32344108> (Accessed January 22, 2021).
- Yadava RS, Frenzel-McCardell CD, Yu Q, Srinivasan V, Tucker AL, Puymirat J, Thornton CA, Prall OW, Harvey RP, Mahadevan MS. 2008. RNA toxicity in myotonic muscular dystrophy induces NKX2-5 expression. *Nat Genet* **40**: 61–8. <http://www.ncbi.nlm.nih.gov/pubmed/18084293> (Accessed November 28, 2020).
- Yamashita Y, Matsuura T, Shinmi J, Amakusa Y, Masuda A, Ito M, Kinoshita M, Furuya H, Abe K, Ibi T, et al. 2012. Four parameters increase the sensitivity and specificity of the exon array analysis and disclose 25 novel aberrantly spliced exons in myotonic dystrophy. *J Hum Genet* **57**: 368–374. <http://www.nature.com/articles/jhg201237> (Accessed November 29, 2020).
- Yang G, Huang X. 2019. Methods and applications of CRISPR/Cas system for genome editing in stem cells. *Cell Regen* **8**: 33–41. <http://www.ncbi.nlm.nih.gov/pubmed/31666940> (Accessed November 23, 2020).
- Yanovsky-Dagan S, Bnaya E, Diab MA, Handal T, Zahdeh F, Broek WJAA van den, Epsztejn-Litman S, Wansink DG, Eiges R. 2019. Deletion of the CTG Expansion in Myotonic Dystrophy Type 1 Reverses DMPK Aberrant Methylation in Human Embryonic Stem Cells but not Affected Myoblasts. *bioRxiv* 631457. <https://www.biorxiv.org/content/10.1101/631457v1> (Accessed November 23, 2020).
- Yotova V, Labuda D, Zietkiewicz E, Gehl D, Lovell A, Lefebvre J-F, Bourgeois S, Lemieux-Blanchard E, Labuda M, Vézina H, et al. 2005. Anatomy of a founder effect: myotonic dystrophy in Northeastern Quebec. *Hum Genet* **117**: 177–87. <http://www.ncbi.nlm.nih.gov/pubmed/15883838> (Accessed February 11, 2021).
- Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R, et al. 2007. Induced pluripotent stem cell lines derived from human somatic cells. *Science (80-)* **318**: 1917–20. <http://www.ncbi.nlm.nih.gov/pubmed/18029452> (Accessed June 12, 2018).
- Yu Z, Teng X, Bonini NM. 2011. Triplet repeat-derived siRNAs enhance RNA-mediated toxicity in a Drosophila model for myotonic dystrophy. *PLoS Genet* **7**: e1001340. <http://www.ncbi.nlm.nih.gov/pubmed/21437269> (Accessed November 30, 2020).
- Yuan Y, Compton SA, Sobczak K, Stenberg MG, Thornton CA, Griffith JD, Swanson MS. 2007. Muscleblind-like 1 interacts with RNA hairpins in splicing target and pathogenic RNAs. *Nucleic Acids Res* **35**: 5474–5486. https://watermark.silverchair.com/gkm601.pdf?token=AQECAHi208BE49Ooan9kkhW_Ercy7Dm3ZL_9Cf3qfKAc485ysgAAaYwggGiBgkqhkiG9w0BBwagggGTMIIBjwIBADCCAYgGCSqGSIb3DQEHATAeBgIghkgBZQMEAS4wEQQMSH7BLRUVpx1dA2gYAgEQgIIBWxeRleV9v1Bh8JoXSYS4rRM7Zky6SPuKLG8k6Qu24bZy3cOL (Accessed April 13, 2018).
- Zakrzewski W, Dobrzyński M, Szymonowicz M, Rybak Z. 2019. Stem cells: past, present, and future. *Stem Cell Res Ther* **10**: 68. <https://stemcellres.biomedcentral.com/articles/10.1186/s13287-019-1165-5> (Accessed November 22, 2020).
- Zhang F, Bodycombe NE, Haskell KM, Sun YL, Wang ET, Morris CA, Jones LH, Wood LD, Pletcher MT. 2017a. A flow cytometry-based screen identifies MBNL1 modulators that rescue splicing defects in myotonic dystrophy type I. *Hum Mol Genet* **26**: 3056–3068.
- Zhang M, Kimura A, Saltiel AR. 2003. Cloning and characterization of Cbl-associated protein splicing isoforms. *Mol Med* **9**: 18–25. <http://www.ncbi.nlm.nih.gov/pubmed/12765336> (Accessed January 21, 2021).
- Zhang M, Liu J, Cheng A, DeYoung SM, Chen X, Dold LH, Saltiel AR. 2006. CAP interacts with cytoskeletal proteins and regulates adhesion-mediated ERK activation and motility. *EMBO J* **25**: 5284–5293. <http://emboj.embopress.org/cgi/doi/10.1038/sj.emboj.7601406> (Accessed January 21, 2021).
- Zhang M, Liu J, Cheng A, DeYoung SM, Saltiel AR. 2007. Identification of CAP as a Costameric Protein

- that Interacts with Filamin C ed. Y. Wang. *Mol Biol Cell* **18**: 4731–4740. <https://www.molbiolcell.org/doi/10.1091/mbc.e07-06-0628> (Accessed February 3, 2021).
- Zhang X-H, Tee LY, Wang X-G, Huang Q-S, Yang S-H. 2015. Off-target Effects in CRISPR/Cas9-mediated Genome Engineering. *Mol Ther - Nucleic Acids* **4**: e264. <https://www.sciencedirect.com/science/article/pii/S216225311630049X#bib8> (Accessed January 28, 2021).
- Zhang Y. 2008. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**: 40. <http://www.ncbi.nlm.nih.gov/pubmed/18215316> (Accessed February 12, 2021).
- Zhang Z, Zhang Y, Gao F, Han S, Cheah KS, Tse H-F, Lian Q. 2017b. CRISPR/Cas9 Genome-Editing System in Human Stem Cells: Current Status and Future Prospects. *Mol Ther Nucleic Acids* **9**: 230–241. <http://www.ncbi.nlm.nih.gov/pubmed/29246302> (Accessed November 23, 2020).
- Zu T, Gibbens B, Doty NS, Gomes-Pereira M, Huguet A, Stone MD, Margolis J, Peterson M, Markowski TW, Ingram MAC, et al. 2011. Non-ATG-initiated translation directed by microsatellite expansions. *Proc Natl Acad Sci U S A* **108**: 260–5. <http://www.ncbi.nlm.nih.gov/pubmed/21173221> (Accessed November 30, 2020).

Title

Study of the function of MBNL proteins during development using human induced pluripotent stem cells

Abstract

Alternative splicing has emerged as a fundamental mechanism not only for the diversification of protein isoforms but also for the spatiotemporal control of development. Therefore, a better understanding of how this mechanism is regulated has the potential not only to elucidate fundamental biological principles, but also to decipher pathological mechanisms involved in diseases where normal splicing networks are mis-regulated. As part of this thesis, we took advantage of human pluripotent stem cells to decipher during human myogenesis the role of MBNL proteins, a family of tissue-specific splicing regulators whose loss of function is associated with Myotonic Dystrophy type 1 (DM1), an inherited neuromuscular disease. Thanks to the CRISPR/Cas9 technology, we generated human-induced pluripotent stem cells (hiPSCs) depleted in MBNL proteins and evaluated the molecular and functional consequences of this loss on the generation of skeletal muscle cells. Our results indicated that MBNL proteins are specifically required for the late myogenic maturation but not for early myogenic commitment. By a transcriptomic analysis, we were able to highlight the molecular pathways regulated by these proteins during myogenesis, as well as the compensatory effects between MBNL paralogs. This study also allowed us to identify a new alternative splicing defect in DM1, regulated by MBNL proteins, which leads to structural abnormalities of the muscular post-synaptic compartment. Together, our results reveal the temporal requirement of MBNL proteins in human myogenesis and allow the identification of new molecular pathways regulated by these proteins that could be involved in the development of DM1. In the longer term, the tools developed in this study should also facilitate the identification of new therapeutic strategies capable to cope with the loss of function of these proteins.

Keywords

Human pluripotent stem cells, CRISPR/Cas9, Myotonic dystrophy type 1.

Titre

Étude de la fonction des protéines MBNL au cours du développement à l'aide de cellules souches humaines induites à la pluripotence

Résumé

L'épissage alternatif est apparu comme un mécanisme fondamental non seulement pour la diversification des isoformes des protéines mais aussi pour la régulation spatio-temporelle du développement. Par conséquent, une meilleure compréhension de la manière dont ce mécanisme est régulé permettrait non seulement d'élucider les principes biologiques fondamentaux, mais aussi de déchiffrer les mécanismes pathologiques impliqués dans les maladies dans lesquelles ces processus moléculaires sont dérégulés. Dans le cadre de cette thèse, nous avons utilisé des cellules souches pluripotentes humaines pour déchiffrer, au cours de la myogenèse humaine, le rôle des protéines MBNL, une famille de régulateurs d'épissage spécifiques aux tissus dont la perte de fonction est associée à la dystrophie myotonique de type 1 (DM1), une maladie neuromusculaire héréditaire. Grâce à la technologie CRISPR/Cas9, nous avons généré des cellules souches pluripotentes d'origine humaine (hiPSC) déplétées en protéines MBNL et évalué les conséquences moléculaires et fonctionnelles de cette perte sur la génération de cellules musculaires squelettiques.

Nos résultats ont indiqué que les protéines MBNL sont spécifiquement requises pour la maturation myogénique tardive mais pas pour l'engagement myogénique précoce. Par une analyse transcriptomique, nous avons pu mettre en évidence les voies moléculaires régulées par ces protéines durant la myogenèse, ainsi que les phénomènes de compensation entre les paralogues MBNL. Cette étude nous a également permis d'identifier un nouveau défaut d'épissage alternatif dans la DM1, régulé par les protéines MBNL, et qui aboutit à des anomalies structurelles du compartiment post-synaptique musculaire. Ensemble, nos résultats révèlent à la fois la temporalité de l'action des protéines MBNL dans la myogenèse humaine et permettent également d'identifier de nouvelles voies moléculaires régulées par ces protéines et pouvant être impliquées dans le développement de la DM1. A plus long terme, les outils développés dans cette étude devraient également faciliter l'identification de nouvelles stratégies thérapeutiques capables de faire face à la perte de fonction de ces protéines.

Mots clés

Cellules souches pluripotentes humaines, CRISPR/Cas9, Dystrophie myotonique de type 1.