



**HAL**  
open science

# Développement d'une stratégie thérapeutique pour la dystrophie facio-scapulo-humérale

Anne-Charlotte Marsollier

► **To cite this version:**

Anne-Charlotte Marsollier. Développement d'une stratégie thérapeutique pour la dystrophie facio-scapulo-humérale. Biologie cellulaire. Université Pierre et Marie Curie - Paris VI, 2017. Français. NNT : 2017PA066035 . tel-02151743

**HAL Id: tel-02151743**

**<https://theses.hal.science/tel-02151743v1>**

Submitted on 10 Jun 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Université Pierre et Marie Curie

Ecole doctorale 515 - Complexité du vivant

Myologie, Centre de Recherche - UMRS974 UPMC - INSERM – AIM (Paris, France)  
UCL Great Ormond Street Institute of Child Health – Translational Myology Laboratory (Londres, UK)

## **Développement d'une stratégie thérapeutique pour la dystrophie facio-scapulo-humérale.**

Par Melle MARSOLLIER Anne-charlotte

Thèse de doctorat en Biologie Cellulaire et Moléculaire  
Dirigée par le Dr DUMONCEAUX Julie et le Dr MARIOT Virginie

Présentée et soutenue publiquement le 08 Juin 2017

Devant un jury composé de :

Dr MAGDINIER Frédérique	DR2 INSERM, Marseille	Rapporteur
Dr AIT-SI-ALI Slimane	DR2 CNRS, Paris	Rapporteur
Pr FEASSON Léonard	PU-PH Université Jean Monnet, St Etienne	Examinateur
Pr AGBULUT Onnik	PU UPMC, Paris	Président du Jury
Dr DUMONCEAUX Julie	DR AIM UCL, Paris/Londres	Directrice de thèse
Dr MARIOT Virginie	Research associate UCL, Londres	Co-Directrice de thèse



## RESUME

La dystrophie facio-scapulo-humérale (FSHD) est une maladie musculaire autosomique dominante rare. Cette pathologie est causée par la perte des marques épigénétiques répressives au macrosatellite D4Z4 en région subtélomérique du chromosome 4, ce qui conduit à la relaxation de la chromatine, l'expression aberrante du facteur de transcription DUX4 et la dérégulation de centaines de gènes. A ce jour, aucun traitement thérapeutique n'existe pour la FSHD. Le but de ce projet était de déterminer si cibler ou non par des oligonucléotides antisens (AOs) les séquences clés impliquées dans la polyadénylation des ARNm peut-être une stratégie thérapeutique pour inhiber l'expression de *DUX4* chez les patients FSHD. En effet, le clivage et la polyadénylation en 3' des ARNm sont des mécanismes fondamentaux de la maturation des ARNm nécessaires à leur export nucléaire, leur stabilité ou leur traduction efficace. Ces mécanismes représentent donc des cibles intéressantes pour une suppression de l'expression d'un gène dans des maladies à gain de fonction. Pour la première fois, nous avons pu montrer *in vitro* que l'utilisation d'AOs ciblant les séquences clés impliquées dans l'ajout d'une queue poly(A), notamment le signal de polyadénylation ou le site de clivage, conduit à une sous-expression de l'ARNm gène ciblé, et en particulier *DUX4*. Les AOs présentant *in vivo* une faible capacité à pénétrer les cellules et une forte clairance, les séquences des AOs les plus prometteurs ont été insérées dans un vecteur AAV sous promoteur U7. Les premiers résultats obtenus avec ces vecteurs sur un modèle murin sont prometteurs. Cette stratégie innovante apparaît comme une option thérapeutique pour la FSHD.

**Mots clés :** dystrophie facio-scapulo-humérale ; polyadénylation ; thérapie génique ; oligonucléotides antisens ; aav-u7 ; extinction génique.

## ABSTRACT

FacioScapuloHumeral Dystrophy (FSHD) is a rare autosomal dominant neuromuscular disorder. This disease is caused by a loss of epigenetic marks within the D4Z4 macrosatellite located in the subtelomeric region of chromosome 4 leading to chromatin relaxation, aberrant expression of the DUX4 transcription factor and a cascade of gene deregulations. So far, there is no curative treatment for FSHD. The aim of this project was to determine whether or not targeting 3'-end key sequences involved in the polyadenylation of mRNA by antisense oligonucleotides (AOs) can be used as an efficient strategy for *DUX4* gene silencing in FSHD. Indeed, cleavage and polyadenylation of the 3'-end of mRNAs are fundamental mechanisms of mRNAs maturation required for nuclear export, stability of the mRNAs and efficient translation and consequently could represent interesting targets for suppression of gene expression for gain of function diseases. For the first time, we demonstrated *in vitro* that targeting 3'-end key sequences involved in the addition of the poly(A) tail, such as the polyadenylation signal and the cleavage site, leads to an efficient extinction of the mRNA targeted and in particular *DUX4*. Because AOs have a weak cellular uptake and a rapid clearance *in vivo*, the sequences of the most promising AOs have been vectorised into an AAV vector under the control of the U7 promoter. The first results that we obtained with a FSHD mouse model treated with these vectors are promising. This innovating strategy appears as a therapeutic option for FSHD.

**Keywords:** facioscapulohumeral muscular dystrophy; polyadenylation; gene therapy; antisense oligonucleotides; aav-u7; gene silencing.

## REMERCIEMENTS

Mes premiers remerciements vont à mes deux directrices de thèse : Julie et Virginie, sans qui ce projet excitant n'aurait pas vu le jour. Vous avez été de super encadrantes scientifiquement et humainement. J'ai eu une chance énorme. Je ne pouvais pas rêver mieux !

Julie, merci pour tout. Merci de m'avoir fait confiance, dès le M2, en me donnant ce super sujet de thèse. Merci d'avoir été disponible (malgré ce que tu as traversé) quel que soit l'heure ou le jour pour répondre à mes questions et mes doutes durant ces trois années (cela a dû te paraître long !). Merci pour toutes les discussions et les *brainstorming* (il nous manque toujours ce fameux tableau blanc) sur le sujet passionnant qu'est la dystrophie facio-scapulo-humérale (mais pas que). Merci de m'avoir formée, je dirais même façonnée, scientifiquement. Merci de ta franchise, de ton enthousiasme, de ta joie de vivre, de ta bonne humeur (oui tu as bon caractère) et de ton soutien quotidien. Merci de m'avoir embarquée dans cette aventure londonienne, de m'avoir fait découvrir les bons et les mauvais côtés de la recherche, de m'avoir accordée chaque jour un peu plus de liberté que la vieille dans mes projets.

Virginie, également merci pour tout. Merci d'avoir (et d'être encore) mon « *coach* paillasse ». Merci de m'avoir encadrée et formée. Merci d'avoir pris le temps de répondre à mes questions. Merci de ta disponibilité (surtout quand il fallait mettre en place de nouvelles expériences). Merci d'avoir réfléchi avec moi sur mes différents projets, sur mes résultats. Merci pour ton soutien. Merci pour tes encouragements. Merci également pour ta confiance. Mille mercis à vous deux pour tout ça. J'ignore si nos chemins scientifiques se recroiseront, mais si c'est le cas, cela sera un plaisir pour moi de travailler de nouveau avec vous.

Je voudrais remercier le Pr Thomas VOIT de m'avoir accueillie tout d'abord dans son premier laboratoire à l'Institut de Myologie en tant que directeur de l'unité de recherche, puis dans un second temps dans son deuxième laboratoire *Translational Myology Laboratory* à UCL en tant que *Director Designate of Biomedical Research Centre*. Merci pour les discussions scientifiques toujours plus enrichissantes les unes que les autres.

Je tiens également à remercier les membres de mon jury de thèse : le Dr Frédérique MAGDINIER, le Dr Slimane AIT-SI-ALI, le Pr Léonard FEASSON et le Pr Onnik AGBULUT d'avoir accepté, avec beaucoup d'enthousiasme, de lire et de juger ce travail.

J'espère être à la hauteur de vos attentes avec ce manuscrit et lors de ma soutenance. Je remercie également les différents membres de mon comité de thèse : le Pr Sonia KARABINA, le Dr Aurélie GOYENVALLE et le Dr Delphine FRADIN pour les réflexions et les discussions sur mes différents projets (même si certains n'ont jamais abouti).

Je voudrais aussi remercier Romain. Tu as été un peu comme un grand frère pour moi durant cette thèse. A chaque jour sa bêtise ou sa taquinerie (j'avoue avoir vérifié, en arrivant certains matins, que tu ne m'avais pas préparé une quelconque farce !). La science, c'est parfois du bricolage, et on en a inventé des constructions toi et moi ! Merci pour toutes les discussions non scientifiques et les fous rires certains jours (Un petit extrait d'une scène quotidienne ? Romain : «  $2 + 4 = 8$  » ; Moi : « Euh non.  $2 + 4 = 6$  » ; Romain : « Mais non, j'avais déjà 2 ma tête » ; Moi « ok »). Pendant deux ans, je t'ai supporté et inversement ! J'ignore qui a le plus souffert !

J'adresse aussi un merci à tous ceux avec qui j'ai partagé de bons moments dans les différents laboratoires où je suis passée. Un merci en particulier à Jessy, pour les longues discussions (parfois absurdes) en salle de culture (où on aura chanté tant de belles chansons à nos cellules !) et tard le soir (surtout le jeudi soir) ! Un merci aux membres du 3<sup>ème</sup> étage du bâtiment 105 (même si mon passage a été de courte durée) pour les discussions endiablées à table, les échanges de recettes et la bonne humeur continuelle dans les couloirs. Un merci également aux « Anglais » (qui ne sont pas tous anglais, abus de langage !) pour votre accueil à Londres (dans les bureaux comme au laboratoire). Vous m'avez aidée à me sentir ici comme chez moi. Merci pour votre patience avec mon anglais, votre bonne humeur, votre sens du partage (même si j'ai encore du mal avec vos us et coutumes !).

Une thèse, c'est au minimum trois ans de dur labeur. Trois ans de hauts et de bas. Alors je tiens à remercier mes proches qui ont été présents et me soutiennent depuis bien avant cette thèse. En particulier je tiens à remercier une amie fidèle, Annabelle. On a parcouru un sacré bout de chemin depuis cette préparation à l'agrégation (je me rappelle comme si c'était hier de nos discussions par-dessus nos murs de livres dans la bibliothèque d'Orsay !). Merci à toi de m'avoir toujours soutenue durant ces longs mois, de m'avoir régulièrement poussée dans mes retranchements (en me posant mille questions) pour donner le meilleur de moi-même. Merci pour ton aide lors de la rédaction de ce manuscrit, tes relectures m'ont beaucoup apporté.

Un merci également à mes deux chères amies Cécile et Marie pour leur soutien. Merci d'avoir été des mamans pour moi à Paris et de m'avoir régulièrement remise dans le droit chemin. Même si je ne donnais pas l'impression d'écouter vos conseils, ils ont été appliqués bien plus souvent que ce que vous ne le croyez ! Merci d'avoir écouté mes doléances et d'avoir été là dans les moments les plus bas. Merci pour tous les fous rires. Merci de m'avoir aidée à m'évader régulièrement de cette thèse (par contre, le SAMBO, ce n'était peut-être pas une si bonne idée !). Merci pour ces merveilleuses années passées avec vous (et j'espère que cela va continuer).

Un merci aussi à une amie qui compte beaucoup, Anne-Marie. Nous nous sommes perdues de vue pour mieux nous retrouver. Merci d'être ce pilier sur lequel je peux compter en cas de besoin. J'ai parfois l'impression que tu me connais plus que moi-même. Merci pour tes encouragements permanents, pour les remontées de moral. Merci pour les longues, très longues discussions sur le sens de la vie (et on n'a toujours pas résolu le problème !), pour les conseils littéraires, musicaux, cinématographiques et pour toutes les petites choses du quotidien. Merci pour les joies partagées.

Je n'oublie pas non plus Ben (ton plan a presque fonctionné !), Olivier (un sacré coach), Marine (on a encore de longues années devant nous pour radoter comme des petites grand-mères sur nos élèves), les potes de promos mais pas que (les cachanais : Léa, Damien, Lucie, Clément, Rebecca et pleins d'autres ; le M2 Orsay : Hicham, Lucile, Franck, Amaury ; la prépa BCPST à Rennes : Carole, Marie, Katell ...). Merci à vous tous.

Une pensée particulière pour les enseignants exceptionnels qui ont contribué à ma réussite et mon orientation : Anne-Françoise Botrel et l'équipe pédagogique de l'ENS Cachan. Anne-Françoise, tu m'as transmis cette passion pour la biologie et le partage des connaissances. Si mon chemin n'était pas passé par ta salle de cours, je ne serais peut-être pas en train de faire le plus beau métier du monde.

Enfin, merci à mes parents et mon petit frère. Papa, maman, frangin, je vous aime et je vous dédie ce travail. Merci d'avoir attisé ma curiosité dès mon plus jeune âge. Merci de m'avoir permis de grandir dans un environnement sain et culturellement riche. Merci d'avoir cru continuellement en moi et mes capacités. Merci d'avoir été là durant mes grandes périodes de doutes (et la liste est longue). Merci de n'avoir jamais remis en question (ou juste un peu) mes

choix et de m'avoir encouragée à poursuivre mes rêves. Merci de m'accepter comme je suis (et de m'avoir supportée, notamment en période de concours). Maman, Papa, vous êtes formidables. Petit frère, je suis fière de toi et de tout ce que tu as accompli jusqu'ici. Ne t'arrête pas en si bon chemin.

Enfin, j'aimerais rendre un hommage aux souris sacrifiées (toujours dans le plus grand respect) pour les besoins de la science. Même si un long chemin reste à parcourir, votre sacrifice n'aura pas été vain. Bientôt nous atteindrons notre but. Bientôt.

Merci à tous ceux qui par leur aide scientifique, technique ou morale m'ont permis de réaliser ce travail.



# TABLE DES MATIERES

LISTE DES FIGURES ET TABLEAUX .....	12
ABREVIATIONS .....	15
<b>1. INTRODUCTION GENERALE .....</b>	<b>20</b>
<b>1.1. PREMIERES DESCRIPTIONS DE PATIENTS FSHD .....</b>	<b>21</b>
<b>1.2. DESCRIPTION CLINIQUE DE LA FSHD .....</b>	<b>25</b>
1.2.1. La FSHD, une pathologie du muscle strié squelettique .....	25
1.2.2. Atteintes musculaires chez les patients FSHD .....	27
1.2.3. Atteintes non musculaires chez les patients FSHD .....	29
1.2.4. Diagnostic clinique des patients FSHD .....	30
<b>1.3. DESCRIPTION GENETIQUE DE LA FSHD .....</b>	<b>31</b>
1.3.1. Microsatellites et macrosatellites : généralités .....	31
1.3.2. Localisation du locus de la FSHD .....	31
1.3.3. Identification des polymorphismes sur le chromosome 4 .....	34
1.3.4. Présence du motif D4Z4 sur le chromosome 10 .....	37
1.3.5. Modifications épigénétiques associées à la FSHD .....	39
1.3.5.1. Modifications de la condensation de la chromatine dans la FSHD .....	39
1.3.5.1.1. Modifications post-traductionnelles de la queue des histones dans la FSHD .....	41
1.3.5.1.1.1. Généralités .....	41
1.3.5.1.1.1.1. Perte des marques épigénétiques répressives aux histones chez les patients FSHD .....	42
1.3.5.1.2. Modifications du profil de méthylation de l'ADN dans la FSHD .....	44
1.3.5.1.2.1. Généralités .....	44
1.3.5.1.2.1.1. Hypométhylation aux régions D4Z4 chez les patients FSHD .....	44
1.3.5.1.2.2. Loci associés aux patients FSHD2 .....	46
1.3.5.1.3. Activité insulateur des répétitions D4Z4 .....	52
1.3.5.1.4. Localisation nucléaire des répétitions D4Z4 .....	53
1.3.5.1.5. Rôle de la taille des télomères dans la FSHD .....	57
1.3.5.1.6. Régulations épigénétiques au locus 4q35 par des complexes protéiques et des ARN non codants .....	58
1.3.6. Gènes candidats pour la FSHD .....	61
1.3.6.1. <i>FRG1</i> .....	62
1.3.6.2. <i>FRG2</i> .....	65
1.3.6.3. <i>ANTI</i> .....	66
1.3.6.4. <i>FAT1</i> .....	67
1.3.6.5. <i>DUX4</i> .....	70
1.3.6.1.1. Localisation de l'ORF de <i>DUX4</i> dans le motif D4Z4 .....	70
1.3.6.1.2. Différents variants de l'ARNm <i>DUX4</i> .....	71
1.3.6.1.3. Profil d'expression des différents transcrits <i>DUX4</i> .....	75
1.3.6.1.4. <i>DUX4</i> : un facteur de transcription .....	77
1.3.6.1.5. Dérégulations génétiques liées à l'expression de <i>DUX4</i> .....	79
1.3.6.1.6. Conséquences cellulaires de l'expression de <i>DUX4</i> .....	81
1.3.6.1.7. Faible expression de <i>DUX4</i> dans les cellules FSHD .....	83
1.3.6.1.8. Boucle de rétro-contrôle positif de l'expression de <i>DUX4</i> ? .....	84
1.3.7. Diagnostic génétique de la FSHD .....	86
1.3.7.1. Diagnostic génétique par southern blot .....	86
1.3.7.2. Diagnostic génétique par peignage moléculaire .....	89

<b>1.4. MODELES CELLULAIRES ET MODELES ANIMAUX FSHD.....</b>	<b>92</b>
1.4.1. Modèles cellulaires FSHD .....	92
1.4.2. Modèles animaux FSHD.....	93
1.4.2.1. Modèle du poisson zèbre et du xénope.....	93
1.4.2.2. Modèle de la drosophile.....	95
1.4.2.3. Modèles murins.....	96
1.4.2.3.1. Modèles murins exprimant d'autres gènes candidats de la FSHD .....	96
1.4.2.3.1.1. Modèle murin sur-exprimant <i>FRG1</i> .....	96
1.4.2.3.1.2. Modèle murin sur-exprimant <i>Pitx1</i> .....	97
1.4.2.3.1.3. Modèle murin muté pour <i>FAT1</i> .....	97
1.4.2.3.2. Modèles murins exprimant <i>DUX4</i> .....	98
1.4.2.3.2.1. Souris AAV6-DUX4.....	98
1.4.2.3.2.1. Souris iDUX4(2.7).....	99
1.4.2.3.2.1. Souris D4Z4(2.5) et (12.5).....	99
1.4.2.3.2.1. Souris FLExDUX4.....	100
1.4.2.3.3. Modèle murin humanisé.....	102
<b>1.5. STRATEGIES THERAPEUTIQUES .....</b>	<b>103</b>
1.5.1. Thérapie génique .....	103
1.5.1.1. Diversité des stratégies de thérapie génique.....	106
1.5.1.1.1. Complémentation et compensation.....	106
1.5.1.1.2. Manipulation du génome par les nucléases.....	107
1.5.1.1.2.1. Méganucléases.....	107
1.5.1.1.2.2. Nucléases à doigt de zinc .....	108
1.5.1.1.2.3. Transcription Activator-Like Effector Nucleases (TALENs).....	109
1.5.1.1.2.4. RNA Guided Nucleases (RGNs).....	110
1.5.1.1.2.5. Modifications du génome lors de la réparation de la cassure double-brin.....	112
1.5.1.1.3. Manipulation des transcrits par des oligonucléotides .....	113
1.5.1.1.3.1. siARN et miARN .....	114
1.5.1.1.3.2. snARN .....	116
1.5.1.1.3.3. Oligonucléotides antisens.....	117
1.5.1.2. Différentes méthodes de transfert de l'acide nucléique thérapeutique .....	122
1.5.1.2.1. Méthodes synthétiques et physiques .....	123
1.5.1.2.2. Méthodes virales.....	125
1.5.2. Stratégies thérapeutiques pour la FSHD .....	126
1.5.2.1. Traitements symptomatiques .....	126
1.5.2.2. Stratégies thérapeutiques avec des molécules pharmacologiques.....	127
1.5.2.3. Stratégies de thérapie génique ciblant <i>DUX4</i> .....	128
1.5.2.3. Stratégies de thérapie génique ciblant d'autres gènes candidats dans la FSHD .....	130
<b>1.6. PROJET DE THESE .....</b>	<b>130</b>
1.6.1. Mécanisme et rôle de la polyadénylation des ARNm.....	131
1.6.2. Hypothèse de travail.....	134
<b>2. RESULTATS .....</b>	<b>135</b>
<b>2.1. ETUDE I : DEVELOPPEMENT D'UNE STRATEGIE THERAPEUTIQUE IN VITRO BASEE SUR L'UTILISATION D'OLIGONUCLEOTIDES ANTISENS CIBLANT LES ELEMENTS CLES EN 3'UTR DE DUX4 .....</b>	<b>136</b>
<i>Publication : Antisens targeting of 3' end elements involved in DUX4 mRNA processing is an efficient therapeutic strategy for facioscapulohumeral dystrophy : a new gene-silencing approach .....</i>	138

<b>2.2. ETUDE II : EVALUATION IN VIVO DE L'EFFICACITE DES OLIGONUCLEOTIDES ANTISENS CIBLANT LES ELEMENTS CLES EN 3'UTR DE DUX4.....</b>	<b>154</b>
2.2.1. Introduction.....	154
2.2.1.1. Modifications du U7 small nuclear ribonucleoprotein (U7-snRNP) à des fins de thérapie génique.....	154
2.2.1.2. Utilisation des virus adéno-associés pour le transfert de l'acide nucléique.....	157
2.2.1.2.1. Généralités sur les virus adéno-associés.....	157
2.2.1.2.2. Utilisation des virus adéno-associés recombinants comme outils pour le transfert de l'acide nucléique.....	159
2.2.1.3. Hypothèse de travail.....	161
2.2.2. Matériels et méthodes.....	162
2.2.2.1. Constructions des vecteurs AAV-U7-PAS et pCINéoDUX4-ΔSV40.....	162
2.2.2.2. Production des AAV.....	163
2.2.2.3. Procédures animales.....	163
2.2.2.4. Extraction ARN et ADN, transcription reverse.....	164
2.2.2.5. PCR et PCR quantitative.....	165
2.2.2.6. Détection du U7-snARN.....	165
2.2.2.7. 3'RACE.....	166
2.2.3. Résultats préliminaires.....	168
2.2.3.1. Détermination de l'expression de l'ARNm <i>DUX4</i> et de ses gènes cibles <i>in vivo</i> dans les muscles de souris pCINéo-DUX4 traités avec le vecteur AAV-U7-PAS.....	168
2.2.3.2. Recherche d'une redirection du site de clivage <i>in vivo</i> dans les muscles de souris pCINéo-DUX4 traités avec le vecteur AAV-U7-PAS.....	170
2.2.3.3. Détermination de l'expression de l'ARNm <i>DUX4</i> et de ses gènes cibles <i>in vivo</i> dans les muscles de souris pCINéo-DUX4 ΔSV40 traités avec le vecteur AAV-U7-PAS.....	171
2.2.3.4. Détermination de l'expression de l'ARNm <i>DUX4 in vivo</i> dans les muscles de souris ACTA1-MCM,FLexD/+ traités avec le vecteur AAV-U7-PAS.....	174
<b>3. DISCUSSION ET PERSPECTIVES.....</b>	<b>177</b>
<b>3.1. NOUVELLE STRATEGIE THERAPEUTIQUE INNOVANTE POUR LA FSHD INHIBANT L'EXPRESSION DE DUX4.....</b>	<b>178</b>
3.1.1. Preuve de concept <i>in vitro</i> de la faisabilité d'une stratégie thérapeutique ciblant les séquences clés en 3'UTR du pré-ARNm <i>DUX4</i> .....	178
3.1.2. Mécanismes hypothétiques d'action des PMOs.....	179
3.1.3. Evaluation de l'efficacité de cette stratégie <i>in vivo</i> .....	180
3.1.4. Optimisation de la stratégie thérapeutique.....	182
3.1.5. Avantages et inconvénients de cette stratégie thérapeutique en comparaison aux autres stratégies proposées.....	183
<b>3.2. REDIRECTION DU SITE DE CLIVAGE EN PRESENCE D'OLIGONUCLEOTIDES ANTISENS CIBLANT LES ELEMENTS CLES EN 3' DU PRE-ARNM.....</b>	<b>184</b>
3.2.1. Mécanismes hypothétiques expliquant une redirection du site de clivage en présence d'oligonucléotides antisens.....	184
3.2.1.1. Mécanismes hypothétiques de redirection du site de clivage <i>in vitro</i> .....	185
3.2.1.1. Mécanismes hypothétiques de redirection du site de clivage <i>in vivo</i> .....	187
3.2.2. Conséquences d'une redirection du site de clivage sur la stabilité de l'ARNm.....	188
3.2.3. Utilisation d'oligonucléotides antisens induisant une redirection du site de clivage comme stratégie thérapeutique.....	191

<b>3.3. INHIBITION DES DIFFERENTS VARIANTS ARNm DE <i>DUX4</i> : LESQUELS ET QUAND ?</b>	<b>192</b>
3.3.1. Quelles conséquences de cette stratégie thérapeutique sur l'expression des différentes isoformes de <i>DUX4</i> ?.....	192
3.3.2. Quelle expression spatio-temporelle pour les différents variants ARNm <i>DUX4</i> ?.....	194
3.3.3. A partir de quand inhiber l'expression de <i>DUX4</i> dans les cellules somatiques ? .....	195
<b>3.4. COMBINAISON DE PLUSIEURS STRATEGIES DE THERAPIE GENIQUE POUR LA FSHD</b>	<b>196</b>
3.4.1. Faut-il combiner plusieurs stratégies thérapeutiques ciblant <i>DUX4</i> pour le traitement de la FSHD ? .....	196
3.4.2. Faut-il développer des stratégies thérapeutiques pour la FSHD ne ciblant que <i>DUX4</i> ? .....	198
<b>4. CONCLUSION GENERALE .....</b>	<b>199</b>
<b>BIBLIOGRAPHIE.....</b>	<b>201</b>

# LISTE DES FIGURES ET TABLEAUX

Les figures empruntées à d'autres auteurs sans modification sont indiquées par la mention « tiré de ». Si les figures ont subi des modifications ou si un élément provenant du travail d'un autre auteur est introduit dans une illustration, la mention « adapté de » est utilisée. Toutes les autres figures sont des figures que j'ai entièrement réalisées. Elles peuvent être reproduites, modifiées et diffusées librement pour toute utilisation autre que commerciale, sous réserve d'être publiées en me créditant de façon claire en indiquant mon nom, prénom, et la référence à ce manuscrit de thèse.

## FIGURES

<b>FIGURE 1.</b> Muscles affectés dans la FSHD.....	27
<b>FIGURE 2.</b> Photos de patients FSHD.....	28
<b>FIGURE 3.</b> Description du locus FSHD.....	32
<b>FIGURE 4.</b> Organisation des polymorphismes sur le chromosome 4q en région subtélomérique....	35
<b>FIGURE 5.</b> Organisation de la région subtélomérique du chromosome 10q homologue à la région subtélomérique du chromosome 4q.....	37
<b>FIGURE 6.</b> Structure du nucléosome.....	41
<b>FIGURE 7.</b> Modèle des changements chromatiniens aux répétitions D4Z4 liés à la perte des marques épigénétiques H3K9me3 chez les patients FSHD.....	43
<b>FIGURE 8.</b> Structure de la protéine SMCHD1.....	46
<b>FIGURE 9.</b> Mécanisme d'homodérivation des protéines SMCHD1.....	47
<b>FIGURE 10.</b> Spectre des mutations dans <i>SMCHD1</i> retrouvées chez les patients FSHD2.....	48
<b>FIGURE 11.</b> Modèle d'interaction entre les différents facteurs épigénétiques aux répétitions D4Z4 chez des individus contrôles et chez des patients FSHD.....	49
<b>FIGURE 12.</b> Structure de la protéine DNMT3B.....	50
<b>FIGURE 13.</b> Caractéristiques génétiques et épigénétiques des patients FSHD1 et FSHD2.....	52
<b>FIGURE 14.</b> Modèle de localisation nucléaire de la région subtélomérique 4q35 chez des individus contrôles et des patients FSHD.....	55
<b>FIGURE 15.</b> Modèle d'attachement de la chromatine à la membrane nucléaire interne au niveau de la région subtélomérique 4q35 chez des individus contrôles et des patients FSHD.....	56
<b>FIGURE 16.</b> Modèle de la régulation transcriptionnelle des gènes au locus 4q35 médiée par les protéines du complexe Polycomb et Trithorax chez des individus contrôles et des patients FSHD...	60
<b>FIGURE 17.</b> Modèle d'activation de l' <i>enhancer</i> aux répétitions D4Z4 par le facteur de transcription KLF15 au cours de la différenciation musculaire chez des patients FSHD ou des individus contrôles.....	61
<b>FIGURE 18.</b> Localisation des différents gènes en région subtélomérique 4q35 pouvant être impliqués dans le développement de la FSHD.....	62
<b>FIGURE 19.</b> Structure de la protéine FAT1.....	68
<b>FIGURE 20.</b> Structure génétique d'une répétition D4Z4.....	71
<b>FIGURE 21.</b> Structure génétique du promoteur lié à l'ORF présent dans chaque répétition D4Z4...	71
<b>FIGURE 22.</b> Structure du gène <i>DUX4</i> et localisation des différents exons de <i>DUX4</i> au sein des répétitions D4Z4.....	72

<b>FIGURE 23.</b> Modèle de la stabilité des ARNm <i>DUX4</i> transcrits à partir des répétitions D4Z4 internes ou de la dernière répétition D4Z4 sur le chromosome 4qA et 10qA.....	73
<b>FIGURE 24.</b> Récapitulatif des différents variants de l'ARNm <i>DUX4</i> transcrits dans les cellules somatiques ou les cellules germinales.....	74
<b>FIGURE 25.</b> Modèle de régulation de l'expression de <i>DUX4</i> par les <i>enhancers</i> myogénique DME1 et DME2.....	75
<b>FIGURE 26.</b> Modèle de diffusion de la protéine DUX4 au sein d'un myotube.....	83
<b>FIGURE 27.</b> Boucle de rétro-contrôle positif de l'expression de <i>DUX4</i> par le NMD.....	85
<b>FIGURE 28.</b> Schéma des profils attendus en <i>southern blot</i> après digestion par différentes enzymes de restriction de la région subtélomérique des chromosomes 4qA/B et 10qA, suivie d'une hybridation avec la sonde p13E-11.....	88
<b>FIGURE 29.</b> Schéma des profils attendus en peignage moléculaire pour les régions subtélomériques des chromosomes 4qA/B et 10qA/B.....	90
<b>FIGURE 30.</b> Récapitulatif des différentes étapes actuelles et futures dans le diagnostic de la FSHD.....	91
<b>FIGURE 31.</b> Répartition des essais cliniques de thérapie génique en cours (Avril 2017) par études de phase.....	105
<b>FIGURE 32.</b> Structure et mécanisme d'action des nucléases à doigt de zinc.....	109
<b>FIGURE 33.</b> Structure et mécanisme d'action des TALENs.....	110
<b>FIGURE 34.</b> Structure et mécanisme d'action du système CRISPR/Cas9.....	111
<b>FIGURE 35.</b> Diversité des mécanismes de réparation de la cassure double-brin générée par une action des nucléases.....	112
<b>FIGURE 36.</b> Voies d'action des siARNs et des miARNs.....	115
<b>FIGURE 37.</b> Structure des acides nucléiques chimiquement modifiés les plus communément utilisés.....	120
<b>FIGURE 38.</b> Mécanismes d'action des oligonucléotides antisens et de l'ARN interférence.....	121
<b>FIGURE 39.</b> Vecteurs utilisés dans les essais de thérapie génique en cours (Avril 2017).....	123
<b>FIGURE 40.</b> Mécanismes de clivage et de polyadénylation des ARNm en 3'.....	133
<b>FIGURE 41.</b> Mécanisme de clivage en 3' du pré-ARNm des histones.....	155
<b>FIGURE 42.</b> Organisation génomique de l'AAV et de l'AAVr.....	159
<b>FIGURE 43.</b> Séquence reconnue par le PMO-PAS et structure du vecteur AAV-U7-PAS.....	161
<b>FIGURE 44.</b> Schéma du protocole expérimental <i>in vivo</i> .....	164
<b>FIGURE 45.</b> Expression de <i>DUX4</i> et de ses gènes cibles ( <i>mTm7sf4</i> et <i>mWfdc3</i> ) dans les TA des souris pCINéo-DUX4 traités avec le vecteur AAV-U7-PAS ou AAV-Ctrl.....	169
<b>FIGURE 46.</b> Détection du U7-snARN endogène (U7-WT) ou du U7-PAS snARN transcrit à partir du vecteur AAV-U7-PAS dans les TA des souris pCINéo-DUX4 traités avec le vecteur AAV-U7-PAS ou AAV-Ctrl.....	169
<b>FIGURE 47.</b> Recherche d'une redirection du site de clivage pour l'ARNm <i>DUX4</i> en présence du U7-PAS snARN dans les muscles de souris exprimant <i>DUX4</i> à partir du plasmide pCINéo-DUX4..	171
<b>FIGURE 48.</b> Expression de <i>DUX4</i> et de deux de ses gènes cibles ( <i>mTm7sf4</i> et <i>mWfdc3</i> ) dans les TA des souris pCINéo-DUX4 $\Delta$ SV40 traités avec le vecteur AAV-U7-PAS ou AAV-Ctrl.....	172

<b>FIGURE 49.</b> Détection du U7-snARN endogène (U7-WT) ou du U7-PAS snARN transcrit à partir du vecteur AAV-U7-PAS dans les TA des souris pCINéo- DUX4 $\Delta$ SV40 traités avec le vecteur AAV-U7-PAS ou AAV-Ctrl.....	173
<b>FIGURE 50.</b> Recherche d'une redirection du site de clivage pour l'ARNm <i>DUX4</i> en présence du U7-PAS snARN dans les muscles de souris exprimant <i>DUX4</i> à partir du plasmide pCINéo-DUX4 $\Delta$ SV40.....	173
<b>FIGURE 51.</b> Expression de <i>DUX4</i> en présence de l'AAV-U7-PAS ou l'AAV-Ctrl dans des muscles de souris <i>ACTA1-MCM,FLexD/+</i> .....	175
<b>FIGURE 52.</b> Détection du U7 naturel (U7-WT) snARN ou du U7-PAS snARN transcrit à partir du vecteur AAV-U7-PAS dans les TA des souris <i>ACTA1-MCM,FLexD/+</i> traités avec le vecteur AAV-U7-PAS ou AAV-Ctrl.....	175
<b>FIGURE 53.</b> Recherche d'une redirection du site de clivage pour l'ARNm <i>DUX4</i> en présence ou en absence du U7-PAS snARN dans les muscles de souris <i>ACTA1-MCM,FLexD/+</i> .....	176
<b>FIGURE 54.</b> Séquences reconnues par les différents PMOs ciblant les éléments clés en région 3'UTR du pré-ARNm <i>DUX4</i> .....	179
<b>FIGURE 55.</b> Localisation des différents signaux de polyadénylation alternatifs sur une distance de 1kb en aval du signal de polyadénylation canonique de <i>DUX4</i> .....	185
<b>FIGURE 56.</b> Localisation des différents signaux de polyadénylation alternatifs en aval du signal de polyadénylation canonique de <i>DUX4</i> sur les plasmides pCINeo-DUX4 et pCINéo-DUX4 $\Delta$ SV40.....	188
<b>FIGURE 57.</b> Récapitulatif des différents variants de l'ARNm <i>DUX4</i> transcrits dans les cellules somatiques ou les cellules germinales.....	193

## TABLEAUX

<b>TABLEAU 1.</b> Description des différents haplotypes pour les variants 4qA et 4qB dans la population contrôle.....	36
<b>TABLEAU 2.</b> Description de différents haplotypes pour les variants 10qA dans la population générale.....	38
<b>TABLEAU 3.</b> Liste des gènes présents en région subtélomérique 4q35.....	62
<b>TABLEAU 4.</b> Tableau de conversion de la taille du fragment digéré en nombre de répétitions D4Z4 en fonction des enzymes utilisées.....	87
<b>TABLEAU 5.</b> Séquences hexamériques identifiées comme des signaux de polyadénylation chez l'Homme et la souris.....	132
<b>TABLEAU 6.</b> Amorces utilisées au cours de l'étude.....	167

## ABREVIATIONS

Les termes utilisés dans ce manuscrit sont ceux du langage courant de la communauté scientifique. La plupart de ces termes sont présentés en anglais, pour d'autres, une traduction en français est indiquée.

<b>2'-F</b>	2'fluoro
<b>2'-MOE</b>	2'O-methoxyethyl
<b>2'-O-Me</b>	2'O-mehtyl
<b>3'RACE</b>	3'Rapid Amplification of cDNA Ends
<b>AAP</b>	Assembly-Activating Protein
<b>AAV</b>	Virus adéno-associé
<b>ACTA1</b>	Actin Alpha1
<b>ADA-SCID</b>	Déficit immunitaire combiné sévère par déficit en adénosine désaminase
<b>ADD</b>	ATRX-DNMT3-DNMT3L
<b>ADN</b>	Acide désoxyribonucléique
<b>ADNc</b>	Acide désoxyribonucléique complémentaire
<b>ANT1</b>	Adenine Nucleotide Translocator 1 / ADP/ATP Translocator
<b>AO</b>	Oligonucléotides antisens (antisens oligonucleotide)
<b>AO PS</b>	Oligonucléotides antisens à phoshothioate
<b>ARN</b>	Acide ribonucléique
<b>ARNi</b>	interférence ARN
<b>ARNm</b>	Acide ribonucléique messenger
<b>ASH1L</b>	ASH1 like Histone Lysine Methyltransferase
<b>ATPase</b>	Adenosine triphosphate synthase
<b>BAH</b>	Bromo-Adjacent Homology
<b>BMD</b>	Dystrophie musculaire de Becker
<b>BZW1</b>	Basic Leucine Zipper And W2 Domains 1
<b>CAPN3</b>	Calpain3
<b>Cas9</b>	CRISPR associated 9
<b>CC</b>	Coiled-Coil region
<b>CCND1</b>	CyclineD1
<b>CDK4</b>	Cycline Dependante Kinase 4
<b>CF</b>	Cleavage Factor
<b>CpG</b>	Dinucléotide CG
<b>CPSF</b>	Cleavage and Polyadenylation Specificity Factor
<b>CPSF73</b>	Cleavage and Polyadenylation Specificity Factor 73
<b>CRISPR</b>	Clustered Regularly Interspaced Short Palindromic Repeats
<b>crRNA</b>	CRISPR ARN
<b>CstF</b>	Cleavage stimulating Factor
<b>CTCF</b>	CCCTC-binding Factor
<b>DBE</b>	D4Z4 Binding Element
<b>DBE-T</b>	D4Z4 Binding Element-Transcript
<b>DEFB103</b>	Defensin beta 103
<b>DMD</b>	Dystrophie musculaire de Duchenne
<b>DME1</b>	Double homeobox (chromosome) 4 myogenic enhancer 1
<b>DME2</b>	Double homeobox (chromosome) 4 myogenic enhancer 2
<b>DNMT3B</b>	DNA (cytosine-5)-methyltransferase 3 beta
<b>DSE</b>	Downstream sequence element
<b>DUX4</b>	Double homeobox (chromosome) 4
<b>DUX4-fl</b>	Double homeobox (chromosome) 4 full length
<b>DUX4-s</b>	Double homeobox (chromosome) 4 shorter
<b>EGF</b>	Epidermal Growth Factor



<b>eIF4A/G/E</b>	Eukaryotes Initiation Factor 4A/G/E
<b>EMG</b>	Electromyogramme
<b>EZH2</b>	Enhancer of Zeste Homolog 2
<b>FAT1</b>	FAT Atypical cadherin 1
<b>FDA</b>	Food and Drug Administration
<b>FN</b>	Nucléase FokI
<b>FOXP3</b>	Forkhead Box P3
<b>FRG1</b>	FSHD Region 1
<b>FRG2</b>	FSHD Region 2
<b>FSHD</b>	Dystrophie facio-scapulo-humérale (FacioScapuloHumeral Dystrophy)
<b>FSHD1</b>	Dystrophie facio-scapulo-humérale de type 1
<b>FSHD2</b>	Dystrophie facio-scapulo-humérale de type 2
<b>GHKL</b>	DNA Gyrase, Heat shock protein 90, Histidine kinase, DNA mismatch repair protein MutL
<b>H2A</b>	Histone 2A
<b>H2B</b>	Histone 2B
<b>H3</b>	Histone 3
<b>H3K27me3</b>	Triméthylation de la lysine 27 de l'histone H3
<b>H3K4me2</b>	Diméthylation de la lysine 4 de l'histone H3
<b>H3K9ac</b>	Acétylation de la lysine 9 de l'histone H3
<b>H3K9me3</b>	Triméthylation de la lysine 9 de l'histone H3
<b>H4</b>	Histone 4
<b>HBF</b>	Hairpin Binding Factor
<b>HBiX1</b>	Human inactive X Chromosome 1
<b>HD SMC</b>	Hinge domain SMC
<b>HDE</b>	Histone Downstream Element
<b>hhspm3</b>	hypomethylated sperm-specific sequence 3
<b>HMGB2</b>	High-Mobility Group Protein B2
<b>HP1</b>	Heterochromatine protein 1
<b>HPRT</b>	Hypoxanthine PhosphoRibosyl Transferase
<b>HR</b>	Homologous Recombinaison
<b>HSA</b>	Human Skeletal $\alpha$ -actin
<b>HSV1</b>	Virus de l'herpès
<b>hTERT</b>	human Telomerase Reverse Transcriptase
<b>IPEX</b>	Immune Dysfunction, Polyendocrinopathy, Enteropathy, X-linked
<b>iPS</b>	induced Pluripotent Stem cells
<b>IRM</b>	Imagerie par Résonance Magnétique
<b>ITR</b>	Inverted Terminal Repeat
<b>K.D.</b>	Knock Down
<b>K.I.</b>	Knock In
<b>K.O.</b>	Knock Out
<b>kb</b>	kilobases
<b>KLF15</b>	Krüppel-Like Factor 15
<b>LNA</b>	Locked Nucleic Acid
<b>LR PCR</b>	Long range Polymerase Chain Reaction
<b>LRIF1</b>	Ligand dependent nuclear receptor interacting factor 1
<b>Lsau</b>	LongSau3A
<b>LTR</b>	Long Terminal Repeat
<b>MAFbx</b>	Muscle atrophy F-box
<b>MaIR</b>	Mammalien apparent Retro-transposons
<b>MBD3L2</b>	Methyl-CpG Binding protein-like
<b>mCK</b>	muscle specific Creatine Kinase promoter
<b>MCL</b>	Mantle Cell Lymphoma
<b>MECP2</b>	Methyl Cytosine Binding Protein 2
<b>MeDIP</b>	Methylated DNA Immunoprecipitation

<b>mGapdh</b>	glycéraldéhyde-3-phosphate déshydrogénase (mouse)
<b>MHCK7</b>	hybrid $\alpha$ -myosin heavy chain enhancer/ MCK enhancer-promoter
<b>miARN</b>	micro ARN
<b>miRISC</b>	miRNA-Induced Silencing Complex
<b>mTm7sf4</b>	Transmembrane 7 superfamily member 4 (mouse)
<b>MTMR1</b>	Myotubularin-related Protein 1
<b>MurF1</b>	Muscle-specific RING Finger 1
<b>mWfdc3</b>	WAP Four-Disulfide Core Domain 3 (mouse)
<b>NAB</b>	Neutralizing AntiBodies
<b>NDE</b>	NonDeleted Element
<b>NHEJ</b>	Non-Homologous End Joining
<b>NLS</b>	Nuclear Localization Signal
<b>NMD</b>	Nonsense-Mediated Decay
<b>O.C.T.</b>	Optimum Cutting Temperature
<b>OPMD</b>	Dystrophie musculaire oculo-pharyngée
<b>ORF</b>	Cadre ouvert de lecture (Open Reading Frame)
<b>p53 (ou TP53)</b>	Tumor protein 53
<b>PABP</b>	Poly(A) Binding Protein
<b>PABPN1</b>	Poly(A) Binding Protein Nuclear 1
<b>PAP</b>	Poly(A) Polymerase
<b>PAS</b>	Signal de polyadénylation
<b>PAX3</b>	Protein paired box 3
<b>PAX7</b>	Protein paired box 7
<b>pb</b>	Paire de bases
<b>PCR</b>	Polymerase Chain Reaction
<b>PCRq</b>	Polymerase Chain Reaction quantitative
<b>PEI</b>	Polyethylenimine
<b>PITX1</b>	Paired like homeodomain 1
<b>PMO</b>	Phosphorodiamidate Morpholino Oligomers
<b>PNA</b>	Peptide Nucleic Acids
<b>Poly(A)</b>	Polyadénylé
<b>PTC</b>	Premature Termination codon
<b>RBFOX1</b>	RNA Binding Protein FOX1
<b>RGN</b>	RNA Guided Nuclease
<b>RT</b>	Reverse Transcription
<b>RVD</b>	Repeat-Variable Di-residues
<b>S/MAR</b>	Scaffold / Matrix Associated Region
<b>Scc2</b>	Sister Chromatid Cohesion protein 2
<b>sgARN</b>	Single guide ARN
<b>shARN</b>	short hairpin RNA
<b>siARN</b>	small interference ARN
<b>siRISC</b>	siRNA-Induced Silencing Complex
<b>SLBP</b>	Stem-Loop Binding Protein
<b>Sm OPT</b>	Originally Meaning Optimized
<b>SMA</b>	Amyotrophie spinale proximale
<b>SMC</b>	Structural Maintenance of chromosomes protein
<b>SMCHD1</b>	Structural Maintenance of Chromosomes flexible Hinge Domain containing protein 1
<b>SMN</b>	Survival of Motor Neurons
<b>snARN</b>	small nuclear RNA
<b>SNP</b>	Single Nucleotide Polymorphism
<b>SSLP</b>	Stable simple Sequence-Length Polymorphism
<b>SUV39H1</b>	Suppressor of Variegation 3-9 Homolog 1
<b>TA</b>	Tibialis anterior
<b>TAL</b>	Transcription Activator-Like

<b>TALENs</b>	Transcription Activator-Like Effector Nucleases
<b>TAP</b>	Transporter associated with Antigen Processing
<b>TNNT3</b>	Fast Skeletal Muscle Troponin T3
<b>tracrARN</b>	trans-acting crARN
<b>TRIM43</b>	Tripartite Motif-containing 43
<b>U7 snRNP</b>	U7 small nuclear RiboNucleoProtein
<b>UAS-GAL4</b>	GAL4-upstream activation sequence
<b>UPF-1</b>	Up-Frameshift Suppressor 1
<b>UTR</b>	Untranslated region
<b>X1-SCID</b>	Déficit immunitaire combiné sévère lié à l'X
<b>YY1</b>	Yin Yang 1
<b>ZF</b>	Doigt de zinc (Zinc finger)
<b>ZFN</b>	Nucléase à doigt de zinc (Zinc Finger Nuclease)
<b>ZFP100</b>	Zinc Finger Protein 100
<b>ZSCAN4</b>	Zinc Finger and SCAN domain containing 4

*« To treat a disease we must invest in basic science.  
Genetics is the most powerful tool we  
have today to understand human biology »*

Dr David Altshuler

# **1.INTRODUCTION GENERALE**

## 1.1. Premières descriptions de patients FSHD

Au milieu du XIX<sup>ième</sup> siècle, les cliniciens présentent un certain engouement pour les maladies musculaires. Le tissu musculaire représente entre 35 et 40 % de la masse totale d'un individu (plus de 600 muscles squelettiques sont dénombrés). De ce fait, une atteinte du muscle, notamment strié squelettique, peut être très invalidante. Pour la majorité des médecins de cette époque, ces pathologies sont uniquement dues à des atteintes neurologiques (plus précisément de la corne antérieure de la moelle). Mais tous ne sont pas convaincus. Parmi eux, Guillaume-Benjamin Duchenne de Boulogne, neurologue français. Duchenne est l'un des premiers médecins à effectuer une description précise de pathologies avec paralysies musculaires (Duchenne, 1858). Une série de documents, publiés dans les « Archives Générales de Médecine », rassemble ses observations cliniques, des résultats d'expérimentations électriques sur patients et des analyses histologiques de muscles. Une grande partie de ses écrits sont consacrés à la maladie musculaire pseudo-hypertrophique ou paralysie myo-sclérotique, communément appelée myopathie de Duchenne de nos jours. Ses travaux de recherche sur cette maladie (Duchenne, 1868), appuyés par des études *post-mortem* de patients réalisées par Albert Eulengurg et Julius Friedrich Cohnheim en 1866, lui permettent de montrer que les lésions musculaires peuvent être totalement indépendantes d'une atteinte de la corne antérieure de la moelle ou des nerfs.

Dans ses écrits, Duchenne effectue également une description précise d'une autre pathologie : « *l'atrophie musculaire graisseuse progressive de l'enfance* » (Duchenne, 1853, 1868). La première référence à cette maladie, mais sous le terme « *d'atrophie musculaire progressive avec transformation graisseuse* », se trouve dans un mémoire du 21 mai 1849, rédigé par Duchenne lui-même, et déposé à l'Académie des sciences. Les patients atteints de cette pathologie montrent dans un premier temps une atrophie asymétrique des muscles du visage et des membres supérieurs, et dans un second temps des muscles des membres inférieurs. En 1850, François-Amilcar Aran présente 11 cas cliniques atteints d'une affection « *non décrite* » du système musculaire qu'il nommera « *atrophie musculaire progressive* » (Aran, 1850). Dans son mémoire, sur les 11 patients rapportés, quatre d'entre eux sont des cas auparavant décrits par Duchenne, qu'Aran a reproduit avec le consentement de ce dernier. Quelques années plus tard, en 1851, Alexandre-André Thouvenet, sous la supervision de Jean Cruveilhier, soutient sa thèse de médecine à la Faculté de Médecine de Paris sur cette même

pathologie qu'il qualifia alors de « *paralysie musculaire atrophique* » en opposition à « *l'atrophie musculaire progressive* » d'Aran. En 1853, au cours d'une séance à l'Académie nationale de médecine, Cruveilhier rapporte le cas de trois patients dont celui de M. Lecomte, tous atteints de ce qu'il qualifie de « *paralysie musculaire progressive atrophique* » (ou paralysie graduelle du mouvement par atrophie musculaire) (Cruveilhier, 1852). La même année, dans un mémoire lu à la Société médico-chirurgicale de Paris, et en reprenant les différentes découvertes de cas effectuées par ces confrères, Duchenne y discute de la dénomination de cette maladie, qu'il finit par désigner sous les termes « *d'atrophie musculaire graisseuse progressive* » (Duchenne, 1853).

Aucune différence n'est alors faite entre des cas adultes et des cas infantiles, que Duchenne considère comme rares. Il faudra attendre l'un de ses articles, daté de 1868, et intitulé *De la paralysie musculaire pseudo-hypertrophique ou paralysie myo-sclérosique*, pour avoir la première distinction entre « *l'atrophie musculaire graisseuse progressive de l'adulte* » et « *l'atrophie musculaire graisseuse progressive de l'enfance* » (Duchenne, 1868). Cette dernière est décrite comme présentant un « *caractère singulier que je n'ai jamais rencontré dans l'atrophie musculaire de l'adulte* » du fait qu'elle débute « *par certains muscles de la face à laquelle elle donne une physionomie particulière, plusieurs années avant d'envahir les muscles moteurs des membres supérieurs et du tronc. Ce début par les muscles de la face peut donc être considéré comme un caractère propre à cette atrophie musculaire progressive de l'enfance [...]* ». En se basant sur l'étude d'une quinzaine de cas, dont certains concernent plusieurs membres d'une même famille, Duchenne décrit cette pathologie comme une maladie qui démarre « *entre 5 et 7 ans, par l'atrophie des mêmes muscles de la face que dans ces deux cas, et, en premier de l'orbiculaire des lèvres dont le défaut de contractilité occasionnait une épaisseur caractéristique des lèvres. Les désordres expressifs de la physionomie sont alors devenus les mêmes que dans les cas précédents, surtout pendant le rire qui n'était exécuté que par le buccinateur ou le risorius Santorini ; puis, après une période stationnaire, l'atrophie a envahie, de 9 à 14 ans, les membres supérieurs, le tronc, et, en dernier lieu, les membres inférieurs, dont elle a détruit les muscles progressivement, irrégulièrement ou capricieusement, laissant quelques-uns de ces derniers intacts* ». Au travers de ses études histologiques, Duchenne observe une transformation graisseuse des muscles atrophiés.

En ce qui concerne l'état du système nerveux des patients atteints d'atrophie musculaire graisseuse progressive de l'enfance, Duchenne se réfère aux résultats d'autopsie de M. Lecomte, un patient de Cruveilhier atteint d'atrophie musculaire progressive et que Duchenne a également pu suivre (Duchenne, 1853, 1868). L'autopsie, réalisée par Cruveilhier, révèle chez ce patient une conservation de l'intégrité du centre céphalo-rachidien mais une atrophie poussée des racines antérieures des nerfs spinaux et des nerfs hypoglosses (Aran, 1850). Prenant l'effet pour la cause, Duchenne en conclut que « *l'atrophie des racines antérieures de la moelle est, d'après les recherches de M. le professeur Cruveilhier, la lésion nerveuse anatomique de l'atrophie musculaire graisseuse progressive* » (Duchenne, 1868). Duchenne cite par ailleurs un histologiste et pathologiste anglais, Jacob Augustus Lockhart-Clarke, qui en 1867, découvre également une atteinte des cornes antérieures de la substance grise de la moelle chez un patient atteint d'atrophie musculaire progressive (Duchenne, 1868; Lockhart-Clarke, 1867). L'atrophie musculaire graisseuse progressive de l'enfance, tout comme l'atrophie musculaire graisseuse progressive de l'adulte sont alors communément considérées comme des maladies avec pour point de départ une atteinte de la corne antérieure de la moelle. Cependant, personne ne relève une seconde observation de Cruveilhier, datant de 1848, concernant un patient, M. Legrand, âgé de 18 ans, et atteint d'une forme sévère d'atrophie musculaire graisseuse progressive (Cruveilhier, 1852). Ce patient, également connu de Duchenne, présentait un amaigrissement extrême, avec une atrophie générale des muscles du tronc, des extrémités des membres inférieurs mais aussi supérieurs, ainsi que du visage. Après le décès de M. Legrand des conséquences d'une variole, une autopsie a été réalisée afin de rechercher une lésion du système nerveux pour expliquer la paralysie musculaire observée. Cependant alors que l'autopsie ne montre aucune atteinte du cerveau et de la moelle, cette observation n'est pas prise en considération, Cruveilhier justifiant une lacune dans l'observation du système nerveux périphérique au cours de l'autopsie, ce malgré un aspect normal aux premiers abords.

En Janvier 1884, Edmé Félix Alfred Vulpian présente au cours d'une séance de l'académie des sciences, les notes de deux médecins neurologues français, Théophile Joseph Louis Landouzy et Joseph Jules Dejerine, concernant une « *myopathie atrophique progressive (myopathie héréditaire débutant, dans l'enfance, par la face, sans altération du système nerveux)* » (Dejerine and Landouzy, 1884). La distinction est apportée entre l'atrophie musculaire graisseuse progressive de l'enfant et l'atrophie musculaire graisseuse progressive de l'adulte, aussi dénommée atrophie de l'adulte de type Aran-Duchenne. Dans la



symptomatologie, l'atrophie musculaire progressive de l'enfant et celle de l'adulte sont très similaires, exception faite que l'atrophie musculaire de l'enfant se caractérise d'abord par une atteinte des muscles du visage avant d'affecter, après une période plus ou moins prolongée, les muscles des membres supérieurs, du tronc et des extrémités inférieures. Une autre distinction est apportée entre ces deux maladies au niveau des altérations nerveuses. Pour cela, Dejerine et Landouzy ont suivi pendant plusieurs années un patient atteint d'une atrophie musculaire graisseuse progressive de l'enfant, qui présentait à trois ans uniquement une atrophie des muscles du visage, atrophie qui s'est répandue vers l'âge de 18 ans aux muscles des membres supérieurs puis à l'ensemble des muscles du corps, sans troubles paralytiques et sensitifs. Suite au décès du patient à l'âge de 24 ans de la tuberculose, une autopsie confirme une atrophie de tous les muscles du corps à l'exception des muscles de la langue, du pharynx, du larynx, de l'œil, du diaphragme, des intercostaux et des sous-scapulaires. Dejerine et Landouzy souligneront également une asymétrie dans l'atteinte et une absence d'hypertrophie musculaire. Par ailleurs, au cours de l'autopsie, aucune altération du système nerveux central et périphérique n'est détectée contrairement à ce qui est observé dans le cas de l'atrophie musculaire graisseuse progressive de l'adulte. Dejerine et Landouzy en concluent que l'atrophie musculaire progressive de l'enfance est une affection du système musculaire, qui débute généralement à l'enfance par une atteinte des muscles de la face, sans altération du système nerveux, et qu'elle se doit d'être dissociée de l'atrophie musculaire progressive myélopathique de l'adulte, type Aran-Duchenne. Le patient a une sœur et un père qui présentent des atteintes des muscles du visage, permettant d'établir une transmission selon un mode autosomique dominant, critère que Dejerine et Landouzy n'incluront pas dans leur description de la pathologie. Des cas sporadiques sont également observés. Un an plus tard, Dejerine et Landouzy rédigent un article très complet dans « Revue de médecine » sur « *La myopathie atrophique progressive ; myopathie sans neuropathie débutant d'ordinaire dans l'enfance, par la face* » (Landouzy and Dejerine, 1885). Entre 1884 et 1886, Dejerine et Landouzy rapporteront 24 cas, ce qui leur permettra d'ajuster leur description de la pathologie qu'ils nommeront alors « *myopathie progressive de type facioscapulohumérale* » (FSHD) ou dystrophie de Landouzy-Dejerine. D'autres études viendront compléter les données cliniques (Jacqueline, 2011).

Le concept qu'une « maladie primitive du muscle », et non une atteinte neurologique, puisse être la cause d'une atrophie musculaire progressive lente, n'a été accepté par la communauté scientifique qu'à la fin du XIX<sup>ième</sup> siècle. Un long débat sur la classification des myopathies

commencera alors, et durera une trentaine d'années. Il faudra attendre 1921 pour l'introduction d'une classification dite génétique en fonction du mode de transmission : autosomique dominant, autosomique récessif ou récessif lié au chromosome X (Padberg, 1982). La FSHD est considérée comme une maladie à transmission autosomique dominante, ce qui sera confirmé en 1950 par Tyler et Stephens lors d'une étude pluri-générationnelle sur 17 familles (Tyler and Stephens, 1950).

## **1.2. Description clinique de la FSHD**

La FSHD est la troisième myopathie héréditaire la plus commune après la dystrophie de Duchenne et la dystrophie myotonique. Sa prévalence est estimée à 1-9 naissances sur 100 000 (Orphanet, 2017). Les mutations *de novo* représentent 10 à 30 % des cas (Padberg, 1982; Padberg et al., 1995a; Zatz et al., 1995). La FSHD se développe entre 3 et 60 ans. Sa pénétrance est incomplète : 70% vers l'âge de 15 ans et 95% après 20 ans (Lunt et al., 1989). Les hommes sont typiquement plus affectés que les femmes (Zatz et al., 1998). Les critères cliniques de la FSHD ont été établis lors d'un consortium international en 1991 (Padberg et al., 1991). La FSHD se caractérise par une atrophie des certains muscles striés squelettiques, et par des atteintes non musculaires (cf review Pandya et al., 2008; Tawil, 2004, 2008).

### **1.2.1. La FSHD, une pathologie du muscle strié squelettique**

Le muscle est un organe composé majoritairement de tissu musculaire, lui-même formé de cellules contractiles différenciées ou fibres musculaires ou myocytes. En fonction de leur structure, trois types de fibres musculaires existent : les fibres musculaires lisses, localisées dans la paroi des viscères et des vaisseaux sanguins ; les fibres musculaires striées cardiaques, retrouvées dans le cœur ; les fibres musculaires striées squelettiques (ou rhabdomyocytes), qui composent les muscles associés au squelette impliqués dans la motilité et le maintien de la posture. La FSHD est une pathologie du muscle strié squelettique.

Une fibre musculaire striée squelettique est une cellule fusiforme de 20 à 100  $\mu\text{m}$  de diamètre pour une longueur pouvant atteindre 30 à 50 cm en fonction de la taille du muscle dans lequel elle se trouve. Les fibres musculaires striées squelettiques sont composées en leur centre de faisceaux de petites unités, parallèles, orientées selon le grand axe de la fibre appelées

myofibrilles. Les myofibrilles ont un diamètre de 1 à 2  $\mu\text{m}$  et correspondent à la répétition d'unités structurales et fonctionnelles de 4 à 6  $\mu\text{m}$  que sont les sarcomères. Ce sont les sarcomères qui confèrent la propriété de contraction volontaire à la fibre musculaire striée squelettique sous contrôle du système nerveux central (Frontera and Ochala, 2015).

La formation d'un muscle strié squelettique s'effectue en plusieurs étapes. Tout d'abord, au cours du développement fœtal, des cellules progénitrices totipotentes provenant des somites vont être soumises à un programme de détermination myogénique, les convertissant en myoblastes, des cellules déterminées mais qui peuvent encore proliférer. Par la suite, l'activation d'un programme de différenciation myogénique dans ces myoblastes permet l'expression de protéines spécifiques au muscle strié squelettique et leur fusion en myotubes, un syncytium. Ce processus de fusion concerne de trois à plus de 500 cellules. Enfin, la maturation de ces myotubes conduit à la formation de fibres musculaires striées squelettiques. L'ensemble de ces étapes se caractérisent par une modification de l'expression de nombreux gènes médiée par divers facteurs de transcription qui agissent de manière coordonnée dans le temps (Bentzinger et al., 2012; Braun and Gautel, 2011; Taylor, 2002).

Dans le tissu musculaire strié squelettique, des cellules mononuclées, quiescentes appelées cellules satellites, sont présentes entre la fibre musculaire et la lame basale de celle-ci (Mauro, 1961). Ces cellules, peu nombreuses, sont des cellules souches musculaires qui contribuent à la croissance et à la régénération du tissu musculaire (Hikida, 2011; Relaix and Zammit, 2012; Wang and Rudnicki, 2012). Suite à un dommage, les cellules satellites quiescentes sont activées par des facteurs myogéniques induisant leur prolifération et leur fusion avec des fibres musculaires. Les cellules satellites peuvent également fusionner entre elles pour former de nouvelles fibres musculaires (Bareja et al., 2014). La population de cellules satellites est maintenue par une division asymétrique : une des cellules filles s'engage vers la différenciation tandis que l'autre cellule fille retourne à un état de quiescence (Collins and Partridge, 2005; Kuang et al., 2007). Dans le cas des maladies neuromusculaires, les cellules satellites ne permettent pas de régénérer l'intégralité musculaire. Au fil du développement de la pathologie, l'atrophie s'accompagne alors peu à peu de la mise en place d'un tissu fibroadipeux.

### 1.2.2. Atteintes musculaires chez les patients FSHD

Le degré d'atrophie des muscles chez les patients FSHD est extrêmement variable d'un individu à un autre. Les premiers signes cliniques de la FSHD sont une atrophie sélective asymétrique des muscles du visage (plus de 50% des patients FSHD) suivie par une atteinte de la ceinture scapulaire et des muscles du haut des bras (biceps, triceps, dentelé antérieur ou grand dentelé, rhomboïde). Plus tardivement, les muscles abdominaux et les muscles des membres inférieurs peuvent être affectés (Mul et al., 2016; Pandya et al., 2008; Tawil, 2004, 2008) (Figure 1 et Figure 2).

Les muscles du visage initialement touchés sont le muscle orbiculaire de l'œil (*orbicularis oculi*), le muscle orbiculaire de la bouche (*orbicularis oris*) et les muscles zygomatiques (*zygomaticus*) (Figure 1 et Figure 2). Le muscle extraoculaire n'est jamais affecté. L'atrophie de ces muscles rend difficile le mouvement des paupières et des lèvres. De ce fait, les patients ne peuvent fermer entièrement les yeux et présentent des difficultés à sourire, à produire des sons labiaux, à siffler ou à boire avec une paille (Harold, 2006; Pandya et al., 2008) (Figure 2).

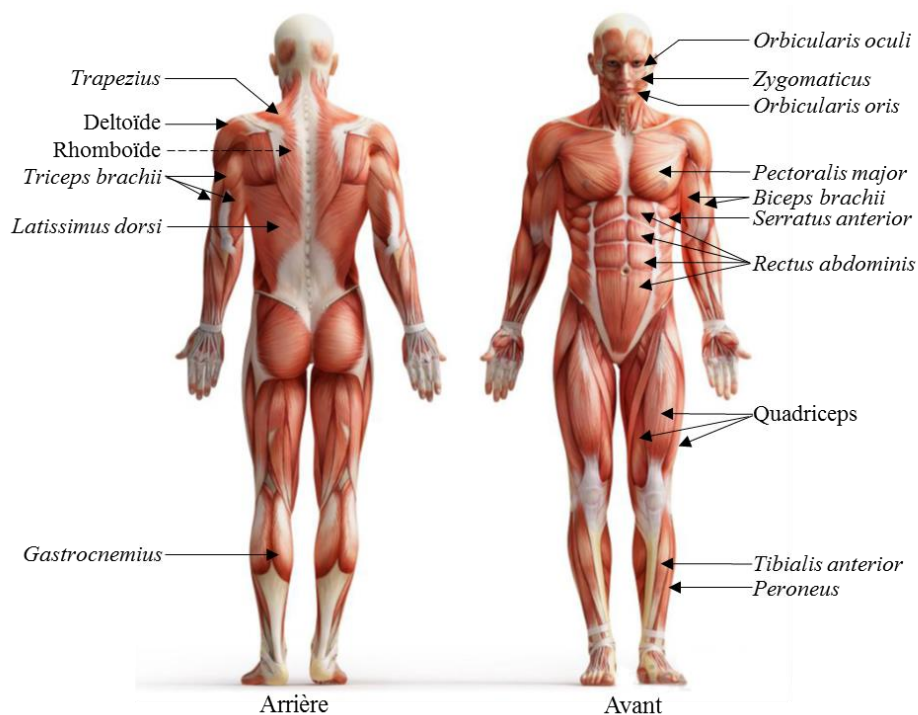


FIGURE 1. Muscles affectés dans la FSHD. La FSHD se caractérise dans un premier temps par une atrophie asymétrique de certains groupes de muscles du visage, de la ceinture scapulaire et des bras. Dans un second temps, les muscles abdominaux et des membres inférieurs peuvent être atteints.

Au niveau de la ceinture scapulaire, ce sont les muscles fixateurs de l'omoplate (*trapezius*, *latissimus dorsi*, *rhomboideus*, *serratus anterior*) qui sont les premiers touchés (Figure 1 et Figure 2). Le muscle grand pectoral peut être affecté plus tardivement, et uniquement dans certains cas. Le deltoïde, quant à lui, reste longtemps non affecté, mais finit souvent par présenter une atrophie partielle ou proximale. L'atteinte des muscles de la ceinture scapulaire rend difficile la levée horizontale du bras au-dessus de l'épaule. Les patients FSHD présentent également une saillie des omoplates et des épaules tombantes vers l'avant du fait du décollement scapulaire (Harold, 2006; Pandya et al., 2008) (Figure 2).



FIGURE 2. Photos de patients FSHD. A gauche, l'expression faciale est presque totalement perdue chez un patient FSHD. Les paupières ne peuvent plus se fermer. La bouche reste austère malgré la tentative de sourire. A droite, le patient FSHD présente un décollement de l'omoplate gauche et présente des difficultés à lever les bras (adapté de Wokke et al., 2013 et <https://www.mda.org/disease/facioscapulohumeral-muscular-dystrophy/signs-and-symptoms>)

En ce qui concerne les muscles proximaux des bras, le biceps et le triceps sont les muscles les plus atrophiés, donnant aux bras un aspect dit en « bras de Popeye » (Figure 1). La dégénérescence et la faiblesse musculaire peuvent aussi, dans certains cas, toucher le poignet et les muscles extenseurs des doigts (Harold, 2006; Pandya et al., 2008).

Dans un second temps, les muscles abdominaux (*rectus abdominis*) peuvent être également affectés conduisant à une protubérance de l'abdomen, à une lordose lombaire exagérée (courbure vers l'intérieur de la colonne vertébrale) et pour 90% des patients, à un signe de Beever positif (remontée de l'ombilic lors d'un exercice qui consiste à relever la tête en position allongée, ce qui indique une faiblesse des muscles inférieurs des abdominaux) (Awerbuch et al., 1990) (Figure 1). Il devient alors difficile pour les patients FSHD de passer d'une position allongée à assise ou de se tourner d'un côté à l'autre en position allongée. Certains muscles des jambes (*tibialis anterior*, *peroneus*, *gastrocnemius*), parfois les muscles de la ceinture pelvienne et certains muscles de la cuisse comme le quadriceps, peuvent

également être atrophiés rendant difficile la marche (Figure 1). Environ 10% des patients FSHD (20% des patients de plus de 50 ans) sont en fauteuils roulants (Harold, 2006; Pandya et al., 2008).

La progression de la maladie est souvent lente. Les patients présentent de longues périodes de stabilité suivies de phases de détérioration musculaire rapides, avec une grande variabilité inter-individus. Environ 30% des cas familiaux ne progressent pas au-delà d'une faiblesse de la ceinture scapulaire (Padberg, 1982).

### **1.2.3. Atteintes non musculaires chez les patients FSHD**

Des symptômes non musculaires peuvent être également associés à la FSHD. L'un des plus fréquents est la perte d'audition (25 à 65% des patients FSHD) (Brouwer et al., 1991; Harold, 2006; Lutz et al., 2013; Padberg et al., 1991, 1995b; Trevisan et al., 2008). Les premiers signes sont une surdité des fréquences élevées (sons aigües) qui progresse ensuite vers la perte d'audition de toutes les fréquences sonores. La sévérité de la surdité varie d'un individu à un autre.

Un autre symptôme non musculaire associé à la FSHD est une vasculopathie rétinienne asymptomatique et asymétrique (49 à 75% des patients FSHD) (Harold, 2006; Padberg et al., 1991, 1995b; Statland et al., 2013). Celle-ci peut évoluer dans de rares cas vers des télangiectasies rétinienne primaires (ou maladie de Coat) qui correspondent à des dilatations des vaisseaux rétiniens, des micro-anévrismes associés à des exsudations intra et sous-rétiniennes pouvant provoquer, dans les cas extrêmes, des décollements de la rétine voire des glaucomes néovasculaires et la perte de la vision (Fitzsimons et al., 1987; Lee et al., 2014; Pauleikhoff et al., 1992; Shields et al., 2007; Statland et al., 2013).

Des observations cliniques ont également rapporté que 5% des patients FSHD présentent des problèmes cardiaques : défauts dans la conduction intranodale ou arythmies supraventriculaires de type tachycardie (Baldwin et al., 1973; Faustmann et al., 1996; Finsterer et al., 2005; Galetta et al., 2005; Laforêt et al., 1998; Stevenson et al., 1990; Trevisan et al., 2006). Une insuffisance respiratoire, qui se déclare tardivement, est également observée chez 1% des patients FSHD alors que les muscles liés à la fonction respiratoire ne sont pas atrophiés (van Kernebeek et al., 2017; Wohlgemuth et al., 2004).

Par ailleurs, 30% à 75% des patients FSHD font état d'une douleur chronique modérée à sévère principalement localisée au niveau du cou ou des membres inférieurs. La majorité des patients FSHD attribuent cette douleur à l'effort physique ou à une mauvaise posture liée à l'atrophie musculaire (Jensen et al., 2008; Koetsier, 1997; van der Kooi et al., 2007; Urtizbera, 1997). Enfin, 61% des patients FSHD sont sujets à une grande fatigue s'exprimant par une diminution du niveau des activités physiques et une baisse de motivation (Kalkman et al., 2005, 2007, 2008). Un handicap mental et/ou de l'épilepsie chez des patients FSHD sévères a également été décrit (Funakoshi et al., 1998; Grosso et al., 2011; Miura et al., 1998; Saito et al., 2007; Sharma and Namrata, 2004). Les causes et mécanismes de ces atteintes non musculaires ne sont, pour le moment, pas toutes comprises.

#### **1.2.4. Diagnostic clinique des patients FSHD**

Le diagnostic clinique de la FSHD repose tout d'abord sur la concordance des observations des différents muscles atrophiés avec le tableau clinique de la pathologie et l'histoire naturelle familiale. Des tests complémentaires, basés sur des techniques non-invasives, peuvent être utilisés dans le diagnostic. Notamment, un électromyogramme (EMG), qui permet de mesurer l'activité électrique d'un muscle et des nerfs associés, peut être effectué. Bien que ce test permette de confirmer une myopathie, il n'est pas suffisant pour préciser le type de dystrophie (Dorobek et al., 2013; Stübgen, 2007). Une IRM (Imagerie par Résonance Magnétique) peut aussi être réalisée. Elle permet de détecter l'atrophie musculaire mais surtout de suivre la progression de la pathologie *via* l'évaluation de la fraction grasseuse infiltrée dans les muscles (Janssen et al., 2014a, 2014b; Leung et al., 2015; Tasca et al., 2014).

Par ailleurs, chez certains patients FSHD, le taux de créatine kinase circulante peut être légèrement plus élevé que pour un individu sain (3 à 5 fois le taux normal de créatine kinase plasmatique) (Hughes, 1971; Lunt and Harper, 1991; Zouvelou et al., 2009). Lors d'un diagnostic clinique, une variation du taux de créatine kinase peut donc être recherchée.

Le diagnostic clinique de la FSHD est souvent difficile. Ceci est notamment dû à la grande variabilité des symptômes entre les individus, à la fois au sein d'une même famille et entre les familles. Un diagnostic génétique est donc toujours nécessaire.

## **1.3. Description génétique de la FSHD**

### **1.3.1. Microsatellites et macrosatellites : généralités**

Une grande fraction du génome des eucaryotes correspond à des séquences répétées : chez l'Homme, plus de la moitié du génome en est constitué (Lander et al., 2001; Levy et al., 2007). Une catégorie de répétitions, les répétitions en tandem, sont formées de plusieurs copies de séquences ADN homologues, appelées unités répétées, disposées en continu selon une conformation tête à queue. En fonction de la taille des unités répétées, on distingue plusieurs types de répétitions en tandem : les microsatellites, les minisatellites et les macrosatellites. Les microsatellites, très abondants, sont des répétitions de courts motifs de 1 à 6 nucléotides. Le nombre de motifs répétés formant un microsatellite donné varie d'un individu à un autre mais aussi, chez un même individu, d'un allèle à un autre (on parle de polymorphisme). Du fait de ce polymorphisme important, les microsatellites sont très utilisés comme outils en biologie moléculaire, notamment pour localiser un locus impliqué dans une pathologie génétique ou lors d'analyses de liaison génétique (Holmes, 1994). Les minisatellites sont des répétitions de motifs de 6 à 100 bp. Les macrosatellites sont des répétitions de motifs d'une taille supérieure à 100 bp (Ames et al., 2008; Li et al., 2002; Vergnaud and Denoeud, 2000). Les macrosatellites peuvent contenir dans chaque unité répétée des éléments fonctionnels comme des exons ou des gènes entiers. Initialement considérées comme de l'ADN « poubelle », les répétitions en tandem sont impliquées dans la régulation de l'expression de gènes et de nombreuses pathologies (Brouwer et al., 2009; Gatchel and Zoghbi, 2005; Nakamura et al., 1998; Quilez et al., 2016).

### **1.3.2. Localisation du locus de la FSHD**

Durant de nombreuses années, la communauté scientifique a échoué à localiser le locus de la FSHD (Lucotte et al., 1989; Lunt et al., 1988; Wijmenga et al., 1990a). En 1989, une carte d'exclusion pour le locus de la FSHD est construite à partir des études de liaisons génétiques présentées en 1988 lors d'un consortium international (Sarfarazi et al., 1989). Les chercheurs s'accordent sur l'exclusion du rôle dans la FSHD de 80% du génome. En 1990, le locus associé à la FSHD est localisé sur le bras long du chromosome 4 par Wijmenga *et al.*, grâce à l'étude de la ségrégation de marqueurs microsatellites dans dix familles allemandes



(Wijmenga et al., 1990b). Cette localisation du locus de la maladie a été ensuite confirmée pour différentes familles avec l'utilisation d'autres marqueurs microsatellites, permettant également de préciser la position du locus de la FSHD en région subtélomérique 4q35 (Gilbert et al., 1992; Mathews et al., 1992; Mills et al., 1992; Sarfarazi et al., 1992; Upadhyaya et al., 1990, 1991; Weiffenbach et al., 1992; Wijmenga et al., 1991).

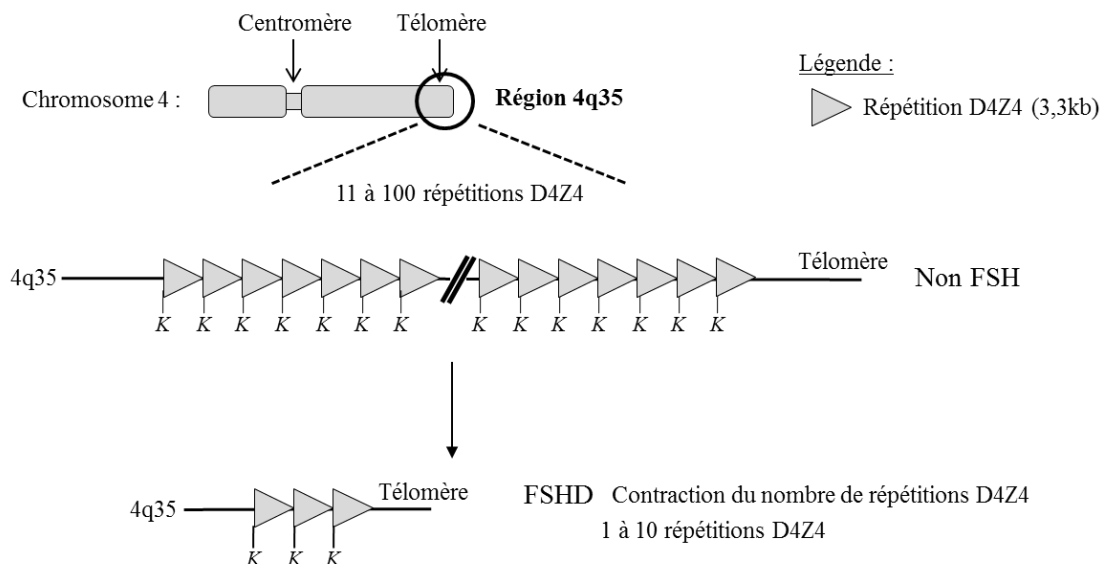


FIGURE 3. Description du locus FSHD. Le locus FSHD est localisé en région subtélomérique 4q35. Dans cette région située à 40-60kb du télomère, se trouve un macrosatellite d'un motif de 3,3 kb appelé D4Z4 (triangle gris), délimité par les sites de restriction reconnus par l'enzyme *KpnI* (K). Les individus contrôles possèdent dans cette région entre 11 et 150 répétitions D4Z4. Les patients FSHD, quant à eux, ne portent qu'entre 1 et 10 répétitions D4Z4.

Dans cette région, située à 40-60kb du télomère, se trouve un macrosatellite d'un motif de 3,3kb appelé D4Z4, délimité par des sites de restriction *KpnI* (Bengtsson et al., 1994; van Deutekom et al., 1993; Wijmenga et al., 1992; Wright et al., 1993a) (Figure 3). Chez les individus sains, ce macrosatellite, d'une longueur moyenne totale de 65kb selon une distribution multimodale, contient entre 11 et 150 répétitions D4Z4 (van Overveld et al., 2000) (Figure 3). La majorité des patients FSHD, quant à eux, présentent une délétion partielle et interne de ces répétitions. Ils ne comptent plus qu'entre 1 et 10 répétitions (van Deutekom et al., 1993; Hewitt et al., 1994; Lee et al., 1995; Wijmenga et al., 1992; Wright et al., 1993a) (Figure 3). Cependant, seuls les individus possédant au moins une répétition D4Z4 développent potentiellement la maladie. En effet, les individus avec une délétion totale des répétitions D4Z4 en région 4q35 ne présentent aucun signe clinique caractéristique de la FSHD (Tupler et al., 1996a).

Par ailleurs, environ 40% des cas de patients FSHD *de novo* sont dus à un réarrangement mitotique aux répétitions D4Z4 par conversion génique inter-chromatinienne (Lemmers et al., 2004b; van der Maarel et al., 2000). Ce réarrangement mitotique s'exprime chez le patient au travers d'un mosaïcisme cellulaire somatique (et parfois germlinal) (van der Maarel et al., 2000). Autrement dit, ces patients FSHD ont subi une contraction post-zygotique des répétitions D4Z4. Ainsi, au sein de l'organisme (ou du tissu), au moins deux populations de cellules génotypiquement différentes uniquement au locus 4q35 coexistent : une population de cellules sans contraction des répétitions D4Z4 et une population de cellules FSHD avec une contraction des répétitions D4Z4 (Lemmers et al., 2004c; van der Maarel et al., 2000). Le pourcentage des diverses populations (ou degré de mosaïcisme) dépend de l'instant auquel la mutation est mise en place au cours de l'embryogenèse. Le fait que le pourcentage de mosaïcisme détecté dans le sang soit généralement représentatif de celui retrouvé dans les cellules du muscle ou de la peau, indique que les réarrangements mitotiques aux répétitions D4Z4 ont lieu très tôt au cours de l'embryogenèse (Lemmers et al., 2004c; van der Maarel et al., 2000; Tonini et al., 2006). Chez certains porteurs asymptomatiques, le mosaïcisme est généralement inférieur à 25%, ce qui traduit de réarrangements mitotiques plus tardifs. Ainsi, le pourcentage de mosaïcisme est un modulateur de la sévérité de la pathologie (van der Maarel et al., 2000). Les patients FSHD mosaïques sont très utiles pour la compréhension de la pathologie puisque la seule différence qui subsiste entre les cellules non FSHD et les cellules FSHD est la contraction des répétitions D4Z4 (Krom et al., 2012).

Une corrélation entre le nombre de répétitions et la sévérité ou l'âge de développement de la FSHD n'est pas clairement établie. En effet, la progression de la maladie est variable entre patients de différentes familles mais également au sein d'une même famille, d'une génération à une autre ou entre jumeaux (Griggs et al., 1995; Tawil and Van Der Maarel, 2006; Tupler et al., 1998). Il semblerait cependant que les patients portant un faible nombre de répétitions D4Z4 (entre une et trois répétitions) développent précocement la maladie avec un tableau clinique très sévère. Plus précisément, pour des patients portant entre une à trois répétitions D4Z4, la pénétrance est complète à l'âge de 50 ans tandis que pour les patients possédant entre quatre et sept répétitions, le risque de développer les symptômes est de 55% à 63% à l'âge de 50 ans (Ricci et al., 2013). Pour les patients ayant entre 7 et 11 répétitions D4Z4, l'âge de développement des premiers signes de la FSHD et le tableau clinique sont très variables selon l'individu (Lunt et al., 1995a, 1995b; Ricci et al., 1999; Tawil et al., 1996; Zatz et al., 1995). A noter que les porteurs asymptomatiques, qui ne développent aucun signe

clinique de la FSHD, sont des individus possédant généralement entre cinq et dix répétitions (Galluzzi et al., 1999; Lunt et al., 1989; Ricci et al., 1999; Tonini et al., 2004). La contraction des répétitions D4Z4 n'est donc pas suffisante pour expliquer la FSHD.

### 1.3.3. Identification des polymorphismes sur le chromosome 4

Puisque certains patients avec une contraction des répétitions D4Z4 ne présentent pas de phénotype FSHD, il a été proposé que la présence de séquences à proximité des répétitions D4Z4 soient nécessaires au développement de la pathologie (Ricci et al., 1999; Tonini et al., 2004).

Différents polymorphismes ont été décrits au locus 4q35. Deux variations subtélomériques homologues à 92% et localisées en aval de la dernière répétition D4Z4 ont été identifiées : l'allèle 4qA et l'allèle 4qB (van Geel et al., 2002) (Figure 4).

Ces variants sont en proportions équivalentes dans la population générale (Lemmers et al., 2002). L'allèle 4qA est constitué, juste en aval de la dernière répétition D4Z4, d'une région de 2,9kb délimitée par les sites de restriction *KpnI-EcoRI*, immédiatement suivie, sur 8kb, par des répétitions de 68bp de type  $\beta$ -satellite et, sur 1kb, d'une séquence répétée divergente (TTAGGG)<sub>n</sub> (Figure 4). Cette région de 2,9kb délimitée par les sites de restriction *KpnI-EcoRI*, nommée pLAM, correspond à une unité D4Z4 divergente et contient une séquence consensus ATTAAA identifiée comme un signal de polyadénylation (PAS) (Dixit et al., 2007; Snider et al., 2009) (Figure 4). Une insertion d'une séquence *Alu* dans la région subtélomérique de l'allèle 4qA a également été identifiée (van Deutekom et al., 1993; van Geel et al., 2002) (Figure 4). Aucune de ces structures n'est retrouvée au niveau de l'allèle 4qB. L'allèle 4qB se caractérise par la présence des premières 570bp d'une répétition D4Z4 et d'une insertion d'une séquence LTR (*Long Terminal Repeat*) en région subtélomérique (van Geel et al., 2002) (Figure 4). Cette divergence de séquences, probablement due à une absence d'échanges entre les deux allèles, est importante, puisque la FSHD est associée uniquement au variant 4qA (Lemmers et al., 2004a, 2002; Rossi et al., 2007; Thomas et al., 2007).

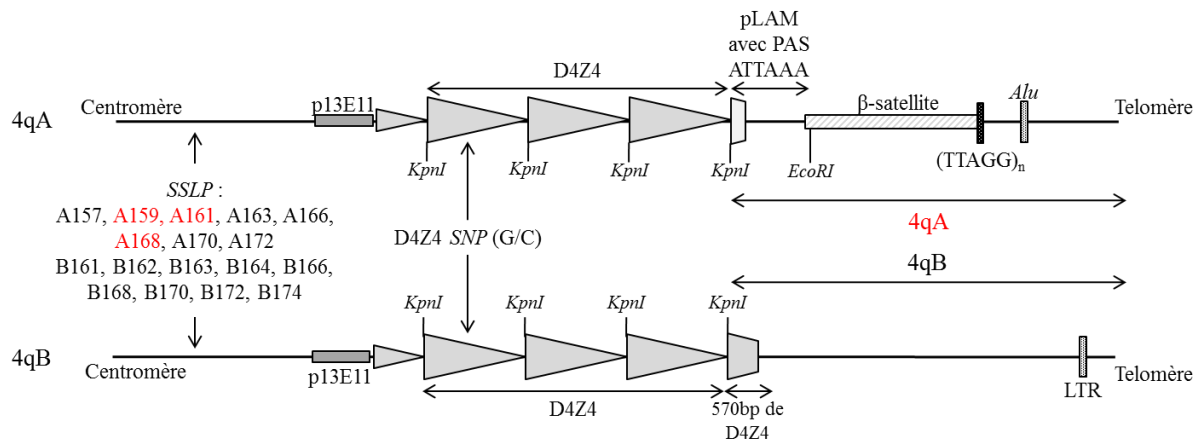


FIGURE 4. Organisation des polymorphismes sur le chromosome 4q en région subtélomérique. Deux polymorphismes, 4qA et 4qB, en aval de la dernière répétition D4Z4 sont présents dans la population générale. L'allèle 4qA est constitué d'une région *KpnI-EcoRI* pLAM de 2,9kb correspondant à une unité D4Z4 divergente, immédiatement suivie sur 8kb par des répétitions de type  $\beta$ -satellite, puis sur 1kb d'une séquence répétée divergente  $(TTAGG)_n$  et enfin d'une séquence *Alu*. L'allèle 4qB est, quant à lui, constitué, en aval de la dernière répétition D4Z4 complète, des premières 570bp d'une unité D4Z4 et d'une séquence LTR. A 3,5kb de la première répétition D4Z4 se trouve un *SSLP* d'une taille de 157bp à 182bp et caractérisé par deux courtes séquences répétées du microsatellite  $(CA)_n$ -MA- $(CA)_n$  et du microsatellite  $(CT)_n$ , de deux polymorphismes nucléotidiques A/C et G/T ainsi que de la présence ou de l'absence d'une séquence de huit nucléotides. Des SNPs sont également présents au sein des répétitions D4Z4. Les haplotypes permissifs associés à la FSHD sont indiqués en rouge.

Par ailleurs, à 3,5kb de la première répétition D4Z4 se trouve une séquence polymorphe stable (*SSLP Stable simple Sequence-Length Polymorphism*). Ce *SSLP*, d'une taille variant de 157bp à 182bp, se caractérise par deux courtes séquences répétées du microsatellite  $(CA)_n$ -MA- $(CA)_n$  et du microsatellite  $(CT)_n$ , de deux polymorphismes nucléotidiques (*SNP Single Nucleotide Polymorphism*) A/C et G/T ainsi que de la présence ou de l'absence d'une séquence de huit nucléotides (Tableau 1, Figure 4). La combinaison de ces marqueurs définit un nombre limité d'haplotypes. L'allèle 4qA est subdivisé en huit haplotypes (4qA157, 4qA159, 4qA161, 4qA163, 4qA166, 4qA168, 4qA170 et 4qA172), et l'allèle 4qB en neuf haplotypes (4qB161, 4qB162, 4qB163, 4qB164, 4qB166, 4qB168, 4qB170, 4qB172 et 4qB174) (Tableau 1). Seuls certains de ces haplotypes sont associés à la FSHD, notamment l'haplotype 4qA159, l'haplotype 4qA161, qui est le plus fréquent dans la population ou encore l'haplotype 4qA168 (Lemmers et al., 2007, 2010a, 2010b; Scionti et al., 2012; Spurlock et al., 2010) (Figure 4).

TABLEAU 1. Description des différents haplotypes pour les variants 4qA et 4qB dans la population contrôle. Les caractéristiques des séquences *SSLP* pour chaque haplotype sont indiquées : taille du microsatellite (CA)<sub>n</sub>-MA-(CA)<sub>n</sub> et du microsatellite (CT)<sub>n</sub>, variant majoritaire au SNP C/A ou au SNP G/T (M = A ou C), et présence (+) ou absence (-) d'une séquence de huit nucléotides. La colonne D4Z4 *SNP* indique la présence majoritaire du variant G ou C au SNP localisé dans chaque répétition D4Z4. N.A.= non analysé (adapté de Lemmers et al., 2010b).

Haplotypes	Séquences <i>SSLP</i>				D4Z4 <i>SNP</i>	
	(CA) <sub>n</sub> -MA-(CA) <sub>n</sub>	G/T	(CT) <sub>n</sub>	8nt	(G/C)	
4qA	4qA157	N.A.			N.A.	
	4qA159	(CA) <sub>10</sub> -AA-(CA) <sub>9</sub>	G	CT <sub>6</sub>	-	G
	4qA161	(CA) <sub>10</sub> -AA-(CA) <sub>10</sub>	G	CT <sub>6</sub>	-	G
	4qA163	N.A.			G	
	4qA166	(CA) <sub>9</sub> -AA-(CA) <sub>8</sub>	T	CT <sub>7</sub>	+	G
	4qA168	(CA) <sub>9</sub> -AA-(CA) <sub>9</sub>	T	CT <sub>7</sub>	+	C
	4qA170	N.A.			N.A.	
	4qA172	N.A.			N.A.	
4qB	4qB161	(CA) <sub>10</sub> -AA-(CA) <sub>10</sub>	G	CT <sub>6</sub>	-	G
	4qB162	(CA) <sub>16</sub>	T	CT <sub>7</sub>	+	C
	4qB163	(CA) <sub>10</sub> -AA-(CA) <sub>11</sub>	G	CT <sub>6</sub>	-	G
	4qB164	N.A.			C	
	4qB166	(CA) <sub>18</sub>	T	CT <sub>7</sub>	+	C
	4qB168	(CA) <sub>19</sub>	T	CT <sub>7</sub>	+	C
	4qB170	(CA) <sub>20</sub>	T	CT <sub>7</sub>	+	C
	4qB172	(CA) <sub>21</sub>	T	CT <sub>7</sub>	+	C
	4qB174	N.A.			N.A.	

Les patients FSHD se caractérisent donc par une contraction des répétitions D4Z4 (une à dix répétitions) sur le chromosome d'haplotype 4qA, associés à l'un des *SSLP* : 4qA159 ou 4qA161 ou 4qA168. On parle alors de chromosome permissif. Cependant, l'association d'une contraction des répétitions D4Z4 avec l'un des haplotypes lié à la pathologie n'est pas toujours nécessaire au développement de la FSHD (Scioni et al., 2012). En effet, des patients FSHD ne présentant pas de contraction des répétitions D4Z4 ont été observés. Néanmoins, ces patients appelés FSHD2 (5% des patients FSHD), en opposition aux patients FSHD1 (95% des patients FSHD) porteurs d'une contraction des répétitions D4Z4, ne peuvent développer la pathologie qu'en présence du chromosome permissif 4qA et des *SSLP* associés (Gilbert et al., 1993; de Greef et al., 2009, 2010). La relation entre phénotype et génotype est complexe.

### 1.3.4. Présence du motif D4Z4 sur le chromosome 10

Des répétitions D4Z4 homologues ont été identifiées un peu partout dans le génome au niveau de régions associées à de l'hétérochromatine (Hewitt et al., 1994; Lyle et al., 1995; Winokur et al., 1994). Cependant, cette homologie ne se restreint pas toujours uniquement aux répétitions D4Z4. En effet, la région subtélomérique du chromosome 4q35 présente un homologue à plus de 98% en région subtélomérique du chromosome 10q26. Cette homologie débute en amont des motifs D4Z4 et se poursuit jusqu'aux répétitions télomériques (Bakker et al., 1995; Cacurri et al., 1998; Deidda et al., 1995; Hewitt et al., 1994) (Figure 5). Plus précisément, la séquence homologue débute environ 42kb en amont de la première répétition D4Z4 par une répétition D4Z4 inversée sur le chromosome 4 tandis que cette répétition D4Z4 inversée est partielle sur le chromosome 10q (van Geel et al., 2002) (Figure 5).

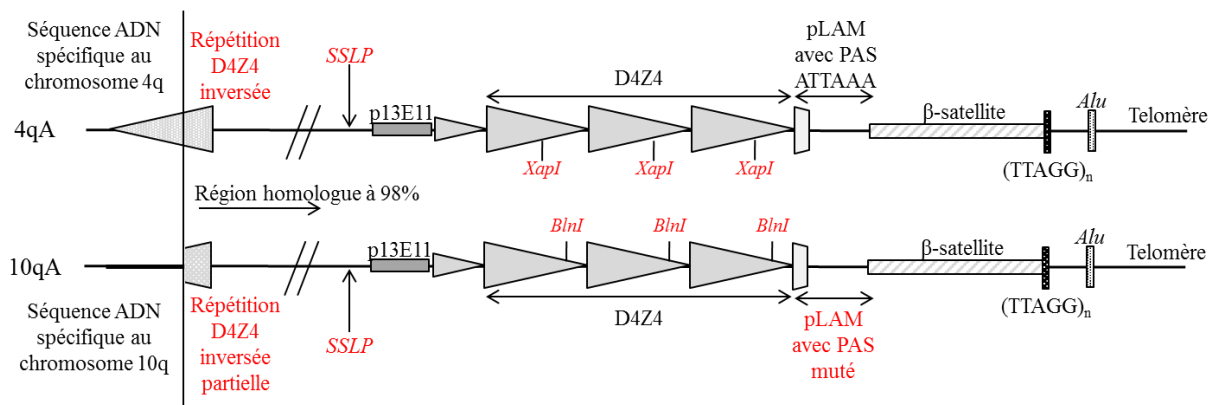


FIGURE 5. Organisation de la région subtélomérique du chromosome 10q homologue à la région subtélomérique du chromosome 4q. L'homologie débute environ 42kb en amont de la première répétition D4Z4 par une répétition D4Z4 inversée sur le chromosome 4q et une répétition D4Z4 inversée partielle sur le chromosome 10q. En région subtélomérique du chromosome 10qA, les répétitions D4Z4 sont suivies d'une région *KpnI-EcoRI* pLAM de 2,9kb correspondant à une unité D4Z4 divergente, associée sur 8kb à des répétitions de type  $\beta$ -satellite, puis sur 1kb à une séquence répétée divergente  $(TTAGGG)_n$  et enfin à une séquence *Alu*. Des polymorphismes permettent de différencier les unités D4Z4 du chromosome 10q de celles du chromosome 4q. Les unités D4Z4 du chromosome 4q sont sensibles à une digestion par *XapI* mais résistantes à une digestion par *BlnI* ( $B^-/X^+$ ), alors que les unités D4Z4 du chromosome 10q sont sensibles à une digestion par *BlnI* mais résistantes à une digestion par *XapI* ( $X/B^+$ ).

Les répétitions D4Z4 sont homologues à 99% entre les deux chromosomes. Les unités répétées D4Z4 du chromosome 4q et 10q peuvent tout de même être discriminées grâce à des *SNP* qui modifient des sites de restriction pour les enzymes *XapI* et *BlnI*. Ainsi, les unités D4Z4 du chromosome 4q sont sensibles à une digestion par *XapI* mais résistantes à une digestion par *BlnI* ( $B^-/X^+$ ). A l'inverse, les unités D4Z4 du chromosome 10q sont sensibles à

digestion par *BlnI* mais résistantes à une digestions par *XapI* (*X/B*<sup>+</sup>) (Deidda et al., 1996) (Figure 5). Le nombre de répétitions D4Z4 de 3,3kb sur le chromosome 10q varie de 3 à 85 (Rossi et al., 2007).

Deux allèles pour le chromosome 10q existent dans la population : l'allèle 10qA et l'allèle 10qB, rare (Lemmers et al., 2010b). Au niveau de l'allèle 10qA, à la manière de l'allèle 4qA, les répétitions D4Z4 sont immédiatement suivies d'une région *KpnI-EcoRI* pLAM, de répétitions de 68bp de type  $\beta$ -satellites ainsi que d'une séquence *Alu* (van Geel et al., 2002) (Figure 5). Cependant, la région pLAM du chromosome 10qA présente un polymorphisme dans le signal de polyadénylation conduisant à la perte de fonctionnalité de celui-ci (Lemmers et al., 2010a) (Figure 5). En ce qui concerne l'allèle 10qB, il est identique à l'allèle 4qB (Lemmers et al., 2010b) : présence des premières 570pb d'une répétition D4Z4 et insertion d'une séquence LTR en région subtélomérique (van Geel et al., 2002) (Figure 4).

TABLEAU 2. Description de différents haplotypes pour les variants 10qA dans la population générale. Les caractéristiques des séquences *SSLP* pour chaque haplotype sont indiquées : taille du microsatellite (CA)<sub>n</sub>-MA-(CA)<sub>n</sub> et du microsatellite (CT)<sub>n</sub>, variant majoritaire au SNP C/A ou au SNP G/T (M = A ou C), et présence (+) ou absence (-) d'une séquence de huit nucléotides. La colonne D4Z4 SNP indique la présence majoritaire du variant G ou C au SNP localisé dans chaque répétition D4Z4. N.A.= non analysé (adapté de Lemmers et al., 2010b).

Haplotypes	Séquences SSLP				D4Z4 SNP	
	(CA) <sub>n</sub> -MA-(CA) <sub>n</sub>	G/T	(CT) <sub>n</sub>	8nt	(G/C)	
10qA	10qA161	N.A.				N.A.
	10qA162	(CA) <sub>9</sub> -AA-(CA) <sub>8</sub>	T	CT <sub>7</sub>	+	G
	10qA164	(CA) <sub>9</sub> -AA-(CA) <sub>8</sub>	T	CT <sub>7</sub>	+	G
	10qA166	(CA) <sub>9</sub> -AA-(CA) <sub>8</sub>	T	CT <sub>7</sub>	+	G
	10qA174	N.A.				C
	10qA176	(CA) <sub>9</sub> -AA-(CA) <sub>13</sub>	T	CT <sub>7</sub>	+	C
	10qA180	(CA) <sub>9</sub> -AA-(CA) <sub>15</sub>	T	CT <sub>7</sub>	+	C

Le chromosome 10q présente également un *SSLP* à 3,5kb de la première répétition D4Z4, subdivisant l'allèle 10qA en plusieurs haplotypes : 10qA161, 10qA162, 10qA164, 10qA166, 10qA174, 10qA176 et 10qA180 (Lemmers et al., 2007, 2010a, 2010b) (Tableau 2).

Du fait de cette forte homologie, de nombreuses translocations totales ou partielles des régions subtélomériques entre le chromosome 4q et le chromosome 10q sont observées dans différentes populations (van Deutekom et al., 1996a; Lemmers et al., 1998; van der Maarel et

al., 2000; Matsumura et al., 2002; van Overveld et al., 2000; Wu et al., 2004). La prévalence des individus qui ont un (monosomie) ou trois (trisomie) chromosomes avec uniquement des répétitions D4Z4 de type 4q est d'environ 20% (Lemmers et al., 1998). Des chromosomes hybrides avec un mélange de répétitions D4Z4 issues du chromosome 4q et du chromosome 10q sont également observés dans la population (van Deutekom et al., 1996a; Lemmers et al., 1998; Matsumura et al., 2002).

La contraction des répétitions D4Z4 en aval d'une région subtélomérique de type 10qA ou 10qB n'est pas associée à la FSHD, même si ces répétitions proviennent du chromosome 4qA. Or l'une des principales différences entre le variant subtélomérique 4qA et 10qA réside dans la perte du signal de polyadénylation en 10qA, démontrant l'importance de cette séquence pour le développement de la FSHD (Bakker et al., 1995; Cacurri et al., 1998; Lemmers et al., 1998).

### **1.3.5. Modifications épigénétiques associées à la FSHD**

La contraction des répétitions D4Z4 n'est observée que chez 95% des patients, appelés FSHD1. Pour les 5% des patients FSHD restants, appelés FSHD2, qui développent le même phénotype clinique que les patients FSHD1, le nombre de répétitions D4Z4 est similaire à celui d'individus contrôles (Gilbert et al., 1993; de Greef et al., 2009, 2010). Le facteur génétique déclencheur de la maladie n'est donc pas uniquement lié à la contraction des répétitions D4Z4 mais reste dépendant de la présence du chromosome permissif 4qA avec un signal de polyadénylation fonctionnel (Gilbert et al., 1993; de Greef et al., 2009, 2010). Associé au fait que la contraction des répétitions D4Z4 sur un chromosome permissif 4qA avec un signal de polyadénylation fonctionnel ne soit pas suffisant pour déclarer la FSHD, comme le démontre des patients asymptomatiques, ces observations suggèrent qu'une composante épigénétique intervient dans le développement de la FSHD (Galluzzi et al., 1999; Lunt et al., 1989; Ricci et al., 1999; Tonini et al., 2004).

#### **1.3.5.1. Modifications de la condensation de la chromatine dans la FSHD**

La chromatine est une structure nucléoprotéique stable et hautement organisée qui permet d'empaqueter environ 2m d'ADN dans un noyau d'un diamètre de 10µm en moyenne. Cette



structuration est permise par l'interaction de l'ADN avec des histones, des protéines non histones et des ARNs. La chromatine est une structure dynamique, non homogène dans le noyau interphasique. En effet, en microscopie électronique, des régions de chromatine plus denses aux électrons se différencient de régions moins denses. Ces régions denses correspondent à de la chromatine sous un état très condensé ou hétérochromatine, en opposition à de la chromatine sous un état plus relâché ou euchromatine (Heitz, 1928). La région 4q35, incluant les répétitions D4Z4, correspond à de l'hétérochromatine (Masny et al., 2004; Tam et al., 2004).

L'hétérochromatine se décompose en une hétérochromatine constitutive et une hétérochromatine facultative. L'hétérochromatine constitutive est une chromatine hautement condensée, stable, et généralement localisée au niveau de régions très riches en répétitions. Ces régions d'hétérochromatine constitutive sont, en théorie, identiques d'une cellule à une autre au sein d'un même organisme (Dimitri et al., 2009; Saksouk et al., 2015). L'hétérochromatine facultative, quant à elle, peut être transitoire et évoluer vers des états chromatinien plus relâchés (Trojer and Reinberg, 2007).

La fréquence de transcription d'un gène dépend de la présence de facteurs de transcription mais aussi de sa localisation dans une région d'hétérochromatine ou d'euchromatine (Feinberg, 2007). De façon simplifiée, l'euchromatine est compatible avec une fréquence de transcription élevée à l'inverse de l'hétérochromatine, dans laquelle les promoteurs des gènes sont difficilement accessibles (Dillon, 2004). Néanmoins, des changements dans l'état de condensation de la chromatine peuvent modifier la fréquence de transcription d'un gène, l'état de condensation de la chromatine étant dynamique. Ces changements d'états chromatinien conduisent donc à des modifications fonctionnelles de l'expression des gènes, héréditaires par mitose et méiose mais qui n'impliquent pas de modifications de la séquence nucléotidique de l'ADN. Autrement dit, ce sont des modifications épigénétiques. Deux mécanismes majeurs interagissent pour modifier l'état de condensation de la chromatine : les modifications post-traductionnelles de la queue des histones et la méthylation des dinucléotides CG (ou CpG).

### 1.3.5.1.1. Modifications post-traductionnelles de la queue des histones dans la FSHD

#### 1.3.5.1.1.1. Généralités

L'ADN, chargé négativement par son squelette sucre-phosphate, est associé dans la cellule à des protéines histones, formant ainsi le nucléosome (Alberts et al., 2002). Un nucléosome se compose d'un noyau protéique de deux exemplaires de chacune des histones H2A, H2B, H3 et H4, autour duquel s'enroule 146pb d'ADN (Kornberg and Lorch, 1999; Ridgway and Almouzni, 2001) (Figure 6). Chaque cœur de nucléosome est séparé du suivant par une région d'ADN dit de liaison, et dont la longueur varie de quelques nucléotides à 80 nucléotides. La succession des nucléosomes forme un nucléofilament ou « collier de perles » de 11nm de diamètre (Kornberg and Lorch, 1999; Oudet et al., 1975).

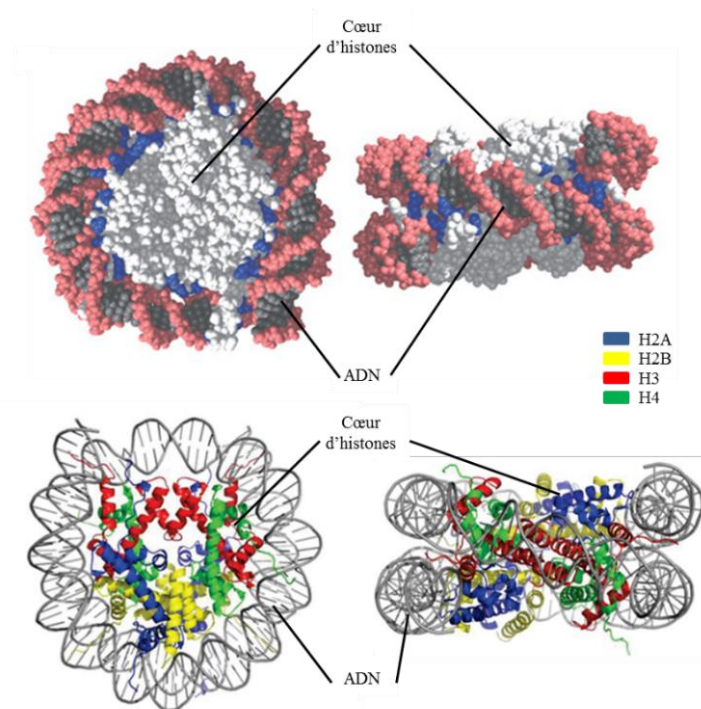


FIGURE 6. Structure du nucléosome. Un nucléosome est composé d'un noyau protéique de deux exemplaires de chacune des histones (en gris sur l'image du haut ou en couleur sur l'image du bas) H2A (en bleu), H2B (en jaune), H3 (en rouge) et H4 (en vert), autour duquel s'enroule 146pb d'ADN (en rouge et noir sur l'image du haut et en gris sur l'image du bas) (adapté de Jiang and Pugh, 2009; Melters et al., 2015).

Les histones, comme de nombreuses protéines, peuvent subir des modifications post-traductionnelles. Ces modifications peuvent être des méthylations, des acétylations, des phosphorylations, des ubiquitylations ou encore des SUMOylations (Tan et al., 2011). Les groupements chimiques sont ajoutés ou retirés par des enzymes sur les résidus lysine,

arginine, sérine, thréonine, proline ou tyrosine présents au niveau de l'extrémité N-terminale non structurée des histones qui pointe à l'extérieur du nucléosome (Portela and Esteller, 2010). Certaines de ces modifications post-traductionnelles sont retrouvées associées de manière spécifique à des régions d'hétérochromatine comme la triméthylation de la lysine 9 de l'histone H3 (H3K9me3). Au contraire, l'acétylation de la lysine 9 de l'histone H3 (H3K9ac) constitue une marque de l'euchromatine.

#### **1.3.5.1.1.1. Perte des marques épigénétiques répressives aux histones chez les patients FSHD**

Les régions contenant les répétitions D4Z4 sont formées à la fois d'euchromatine (principalement au niveau de la première répétition proximale) et d'hétérochromatine (Zeng et al., 2009). En effet, des modifications épigénétiques propres aux deux états chromatiniens ont été identifiées (Bernstein et al., 2005; Das and Chadwick, 2016; Peters et al., 2003; Rice et al., 2003; Vakoc et al., 2005; Zeng et al., 2009, 2014) :

- diméthylation de H3K4 (H3K4me2) et acétylation de H3K9 (H3K9Ac) pour l'euchromatine ;
- triméthylation de H3K9 (H3K9me3) et triméthylation de H3K27 (H3K27me3) pour l'hétérochromatine.

Chez les patients FSHD (FSHD1 et FSHD2), la marque d'hétérochromatine répressive H3K9me3 est perdue au niveau des motifs D4Z4 sur le chromosome 4q. Au contraire, les autres marques épigénétiques comme H3K27me3 ne sont pas quantitativement modifiées (Cabianca et al., 2012; Zeng et al., 2009). Cette perte de la marque H3K9me3 chez les patients FSHD concerne également les motifs D4Z4 du chromosome 10q. Cette modification est indépendante du nombre de répétitions D4Z4, et est retrouvée à l'identique chez les patients FSHD dans différents types cellulaires dont les fibroblastes, les myoblastes et les lymphoblastes (Cabianca et al., 2012; Zeng et al., 2009).

Par ailleurs, les marques H3K9me3 permettent un recrutement direct de protéines HP1 (*Heterochromatine protein 1*) aux régions d'hétérochromatine. Ces protéines sont connues pour leur rôle dans la répression de la transcription (Bannister et al., 2001; Lachner et al., 2001). Chez les individus contrôles, au niveau des répétitions D4Z4, c'est plus précisément le variant HP1 $\gamma$  qui est associé à l'hétérochromatine (Zeng et al., 2009, 2014). Cette protéine

HP1 $\gamma$ , par l'intermédiaire d'une interaction avec le facteur *Scs2* (*Sister Chromatid Cohesion protein 2*), recrute ensuite la cohésine. Les cohésines peuvent interagir entre elles, ce qui conduit à la formation de boucles à l'échelle de la chromatine, difficilement accessibles pour les facteurs de transcription (Bernard et al., 2001; Ciosk et al., 2000; Huichalaf et al., 2014; LeRoy et al., 2009; Nonaka et al., 2002; Zeng et al., 2009). De ce fait, chez les individus contrôles, HP1 $\gamma$  et la cohésine pourraient contribuer à la mise en place d'une interaction physique entre des segments de répétitions D4Z4 ainsi que d'autres régions génomiques distantes en *cis* ou *trans* (Figure 7). Par ce mécanisme, l'hétérochromatine diffuse sur plusieurs centaines de kb (Gruber et al., 2003; Horike et al., 2005; Li et al., 2003; Lomvardas et al., 2006; Spilianakis et al., 2005; Tolhuis et al., 2002). Une perte des marques H3K9me3 aux répétitions D4Z4 compromettrait le recrutement de HP1 $\gamma$  et de la cohésine dans les cellules musculaires. Il en résulterait une perte des interactions de la chromatine et une levée de l'inhibition de la transcription dans cette région (Zeng et al., 2009, 2014) (Figure 7).

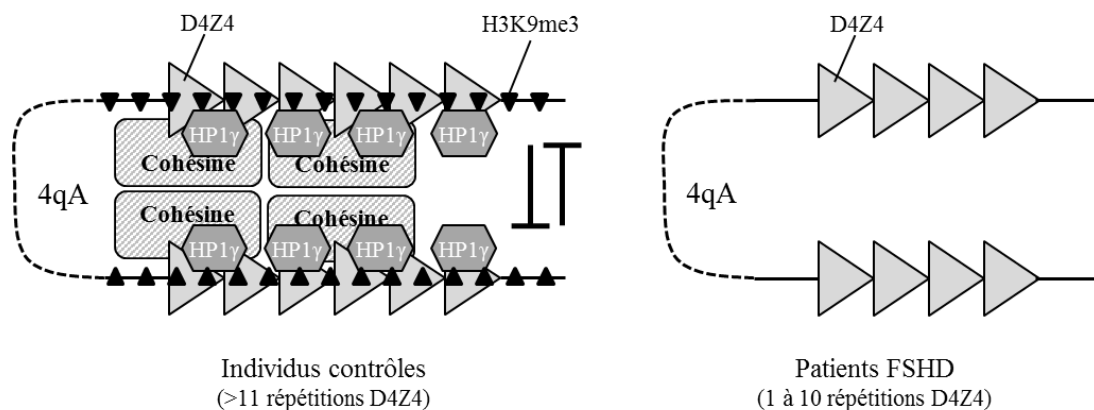


FIGURE 7. Modèle des changements chromatinien aux répétitions D4Z4 liés à la perte des marques épigénétiques H3K9me3 chez les patients FSHD. Les marques épigénétiques H3K9me3 (triangle noir) chez les individus contrôles permettent le recrutement du variant protéique HP1 $\gamma$  (hexagone gris) et de la cohésine (rectangle gris clair) aux répétitions D4Z4 (triangle gris). HP1 $\gamma$  et la cohésine pourraient contribuer à la mise en place d'une interaction physique entre les régions D4Z4 et d'autres régions génomiques distantes en *cis* ou en *trans*, permettant ainsi une diffusion de l'hétérochromatine. Chez les patients FSHD, les marques épigénétiques H3K9me3 sont perdues aux répétitions D4Z4 sur le chromosome 4q et 10q. Les protéines HP1 $\gamma$  et cohésines ne sont plus recrutées. Les interactions entre les différentes régions chromatinien sont perdues, conduisant à l'expression des gènes localisés dans cette région (adapté de Zeng et al., 2009).

La triméthylation de H3K9 au niveau des répétitions D4Z4 est mise en place par l'histone méthyltransférase H3K9 SUV39H1 (*Suppressor of Variegation 3-9 Homolog 1*) (Rea et al., 2000; Zeng et al., 2009, 2014). Cependant, cette enzyme n'est pas mutée chez les patients FSHD et le niveau global de H3K9me3 (*i.e.* dans le génome) est similaire entre patients FSHD et individus contrôles (de Greef et al., 2007; Zeng et al., 2009). La perte des marques

H3K9me3 au niveau des répétitions D4Z4 est donc probablement due à un défaut dans un cofacteur de SUV39H1 ou dans un autre facteur de modification des marques épigénétiques associées aux histones au niveau des répétitions D4Z4.

### **1.3.5.1.2. Modifications du profil de méthylation de l'ADN dans la FSHD**

#### **1.3.5.1.2.1. Généralités**

La méthylation de l'ADN est une modification épigénétique fondamentale dans de nombreux processus physiologiques dont le développement embryonnaire (Cantone and Fisher, 2013; Li et al., 1993; Sharp et al., 2011; Smith and Meissner, 2013). Des variations dans le profil de méthylation de l'ADN peuvent être associées à diverses pathologies (Bergman and Cedar, 2013; Robertson, 2005; Robertson and Wolffe, 2000; Wilson et al., 2007). Cette modification épigénétique, catalysée par des ADN méthyl-transférases, consiste en l'ajout d'un groupe méthyle sur le carbone en position 5 de la cytosine d'un dinucléotide CpG (Goll and Bestor, 2005).

Dans le génome humain, entre 70 à 90% des CpG sont méthylés, selon le type cellulaire. La répartition des CpG et leur méthylation ne sont pas aléatoires dans le génome (Lander et al., 2001). Par exemple, les sites d'initiation de la transcription au niveau de la majorité des promoteurs (70%), sont des régions très denses en CpG, tout comme les séquences répétées (Illingworth and Bird, 2009; Lander et al., 2001; Larsen et al., 1992; Saxonov et al., 2006). Ainsi, la fréquence de transcription d'un gène peut être régulée par la méthylation ou non de l'ADN au niveau de sa séquence promotrice (Moore et al., 2013). La méthylation de l'ADN est une marque associée à la répression de la transcription (Cedar and Bergman, 2012; Robertson, 2005).

#### **1.3.5.1.2.1. Hypométhylation aux régions D4Z4 chez les patients FSHD**

Le locus de la FSHD représente l'un des plus grands îlots CpG présent dans le génome (Neguembor and Gabellini, 2010). En effet, les répétitions D4Z4 sont très riches en GC (ratio C+G de 70% contre 42% en moyenne dans le génome) avec une forte fréquence de dinucléotides CpG (10% de CpG pour une fréquence moyenne de 1% dans le génome). Pas moins de 325 CpG par répétitions D4Z4 sont dénombrés. De ce fait, les motifs D4Z4 sont des

bons candidats pour une régulation épigénétique par méthylation de l'ADN (Gabriëls et al., 1999; Hewitt et al., 1994; Jabbari and Bernardi, 2004; Tsien et al., 2001).

Chez les patients FSHD1 et les patients FSHD2, les répétitions D4Z4 sont hypométhylées comparées à celles des individus contrôles (Gaillard et al., 2014; de Greef et al., 2007, 2009, 2010; Hartweck et al., 2013; Jones et al., 2014, 2015, van Overveld et al., 2003, 2005; Sacconi et al., 2012). Ces changements du profil de méthylation entre individus contrôles et patients FSHD sont limités aux répétitions D4Z4. Ils sont, de plus, observés à l'identique dans diverses cellules somatiques dont les cellules musculaires, les cellules mononuclées du sang périphérique ou encore les fibroblastes (de Greef et al., 2007, 2009; Hartweck et al., 2013; Jones et al., 2014; Lemmers et al., 2015a). Chez les patients FSHD1, cette hypométhylation ne concerne que le chromosome porteur d'une contraction des répétitions D4Z4, indépendamment de l'haplotype : on parle d'hypométhylation contraction dépendante. Au contraire, chez les patients FSHD2, qui ne possèdent pas de contraction des répétitions D4Z4 sur le chromosome permissif, ce sont les chromosomes 4q et 10q qui sont hypométhylés, indépendamment donc du nombre de répétitions : on parle d'hypométhylation contraction indépendant. Ceci implique, pour les patients FSHD2, l'existence d'un défaut génétique responsable de la perte de méthylation aux répétitions D4Z4 (de Greef et al., 2007, 2009, van Overveld et al., 2003, 2005). Par ailleurs, la méthylation n'est pas uniforme au sein d'une répétition D4Z4 ou entre les répétitions D4Z4 (Das and Chadwick, 2016; Gaillard et al., 2014; de Greef et al., 2009; Hartweck et al., 2013; Huichalaf et al., 2014). En résumé, que le patient FSHD soit porteur ou non d'une contraction des répétitions D4Z4 (respectivement, qu'il soit FSHD1 ou FSHD2), une hypométhylation est toujours observée au niveau des répétitions D4Z4, au moins sur le chromosome permissif 4qA.

Il n'y a pas de corrélation clairement avérée entre le nombre de répétitions et le niveau de méthylation (Gaillard et al., 2014; de Greef et al., 2009; Lemmers et al., 2015a; van Overveld et al., 2005). En effet, les porteurs asymptomatiques, dont les répétitions D4Z4 sont contractées sur le chromosome permissif 4qA, présentent un niveau de méthylation des répétitions D4Z4 comparable à celui d'individus contrôles (Gaillard et al., 2014). Cependant, ces observations suggèrent un lien entre le niveau de méthylation des répétitions D4Z4 et le développement des symptômes de la FSHD, permettant d'expliquer en partie la pénétrance de cette pathologie (Gaillard et al., 2014; de Greef et al., 2009; Lemmers et al., 2015a; van Overveld et al., 2005).

### 1.3.5.1.2.2. Loci associés aux patients FSHD2

#### *SMCHD1*

Afin d'expliquer la cause de l'hypométhylation contraction indépendante observée chez les patients FSHD2, des défauts dans des facteurs génétiques ont été recherchés. Des mutations dans un modulateur épigénétique, *SMCHD1* (*Structural Maintenance of Chromosomes flexible Hinge Domain containing protein 1*), localisé sur le chromosome 18, ont été identifiées (Lemmers et al., 2012a, 2015a; Mitsuhashi et al., 2013b).

*SMCHD1* est un gène de 48 exons de la famille des protéines SMC (*Structural maintenance of chromosomes protein*). Il code une protéine de 2 005 acides aminés pour un poids moléculaire de 230kDa. D'un point de vue structural, la protéine *SMCHD1* est composée, en région C-terminale, d'un domaine charnière SMC (*HD SMC Hinge domain*), entouré de part et d'autre de régions en double-hélice (*CC coiled-coil region*) et en région N-terminale, d'un motif GHKL ATPase (*DNA Gyrase, Heat shock protein 90, Histidine kinase, DNA mismatch repair protein MutL*) et d'un domaine homologue au domaine BAH (*Bromo-Adjacent Homology*) (Brideau et al., 2015; Dutta and Inouye, 2000) (Figure 8).

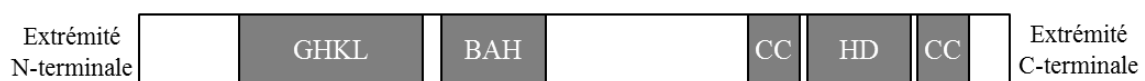


FIGURE 8. Structure de la protéine *SMCHD1*. *SMCHD1* est composée en région C-terminale d'un domaine charnière SMC (*HD SMC Hinge domain*), entouré de part et d'autre de régions en double-hélice (*CC coiled-coil region*). *SMCHD1* est composée en région N-terminale d'un motif GHKL ATPase (*DNA Gyrase, Heat shock protein 90, histidine kinase, DNA mismatch repair protein MutL*) et d'un domaine homologue au domaine *bromo-adjacent homolgy* (BAH) (adapté de Brideau et al., 2015).

D'un point de vue mécanistique, *SMCHD1* s'homodimérise dans un premier temps par une interaction entre les domaines SMC (Figure 9). Dans un second temps, l'homodimère se positionne en formant une sorte de pince autour de l'ADN (Figure 9). Ce positionnement est réversible et contrôlé par l'activité ATPase du motif GHKL (Brideau et al., 2015; Chen et al., 2016b; Hirano, 2006).

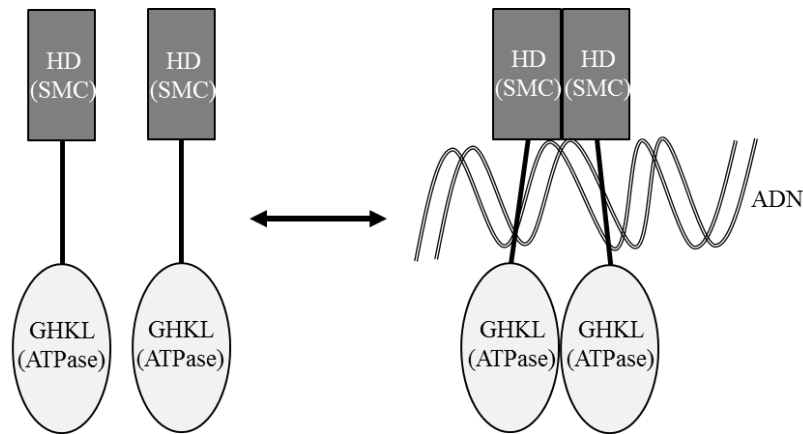


FIGURE 9. Mécanisme d'homodérisation des protéines SMCHD1. Les protéines SMCHD1 s'homodimérisent par une interaction de leur domaine charnière HD (*HD SMC Hinge domain*). L'homodimère vient ensuite se positionner sur l'ADN à la manière d'une pince. Cette fixation à l'ADN est contrôlée par l'activité ATPase du motif GHKL (adapté de Brideau et al., 2015).

SMCHD1 intervient dans différents processus cellulaires dont la mise en place et le maintien de la méthylation de l'ADN de certains sous-ensembles d'îlots CpG lors de l'inactivation du chromosome X, au niveau de séquences répétées ou encore au niveau de gènes autosomaux exprimés de manière monallélique (empreinte parentale). SMCHD1 intervient également dans la réparation des cassures double-brin de l'ADN (Blewitt et al., 2008; Coker and Brockdorff, 2014; Gendrel et al., 2013; Mould et al., 2013; Nozawa et al., 2013; Tang et al., 2014).

Chez 80% des patients FSHD2, *SMCHD1* est muté pour l'un des allèles (Blewitt et al., 2008; van den Boogaard et al., 2015; Larsen et al., 2015; Lemmers et al., 2012a, 2015a, 2015b; Mitsunashi et al., 2013b; Sacconi et al., 2013; Winston et al., 2014). Les mutations répertoriées chez les patients FSHD2, et associées à une hypométhylation des répétitions D4Z4, sont des mutations faux-sens, des mutations non-sens ou de petites délétions qui parfois affectent les sites d'épissages (Figure 10). Ces mutations sont réparties dans tout le gène *SMCHD1*. Lorsque les mutations affectent le cadre ouvert de lecture (*ORF Open Reading Frame*), les transcrits mutants générés sont instables et dégradés. De ce fait, la quantité de transcrits *SMCHD1* totale (comprenant les transcrits naturels ou *wild-type* et quelques transcrits mutants non dégradés) est diminuée chez les patients FSHD2 comparés aux individus contrôles. Ainsi, le nombre potentiel de monomères SMCHD1 fonctionnels est réduit, suggérant un mécanisme d'haploinsuffisance lorsque les mutations affectent le cadre ouvert de lecture. Les mutations faux-sens, quant à elles, ne modifient pas la stabilité des transcrits mutants. Des quantités équivalentes de transcrits issus de l'allèle mutant et de transcrits issus de l'allèle *wild-type* sont présentes dans la cellule. Les protéines mutantes



peuvent donc s'homodimériser et former des hétérodimères avec les protéines *wild-type*. Il en résulte une activité moindre de SMCHD1 ou un mode d'action selon un mécanisme très délétère de dominants négatifs (Lemmers et al., 2015a).

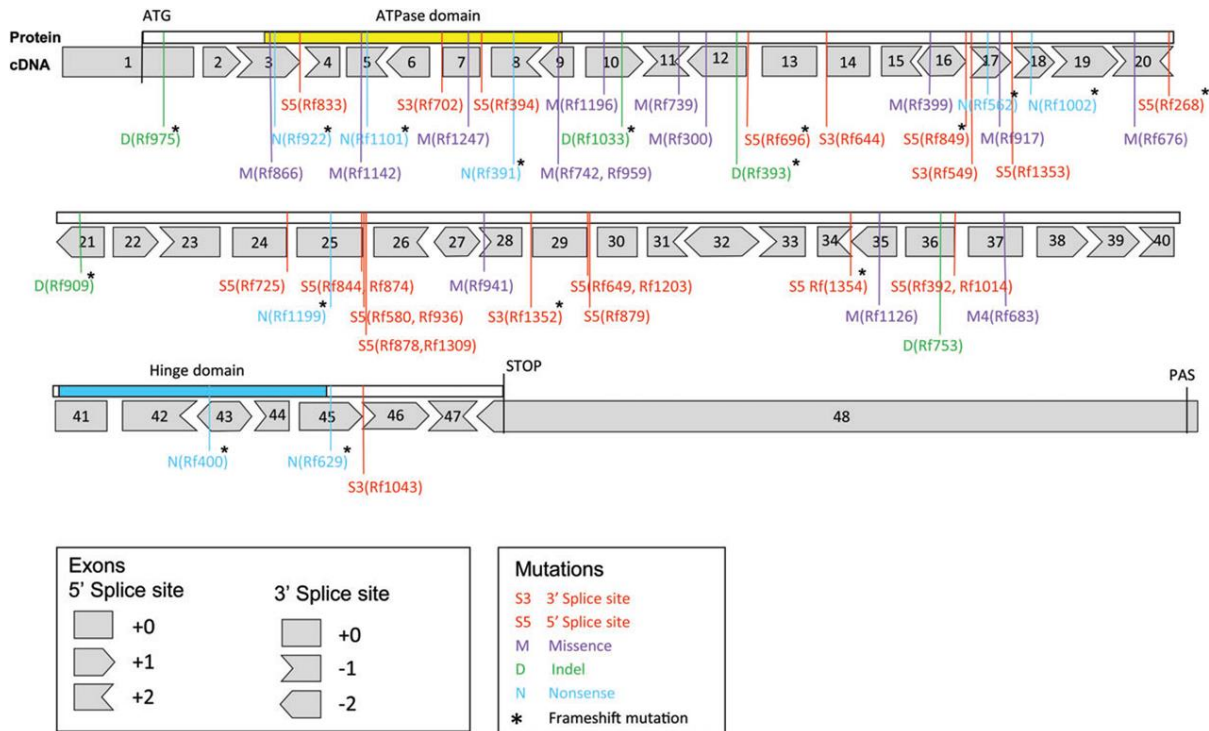


FIGURE 10. Spectre des mutations dans *SMCHD1* retrouvées chez les patients FSHD2. Les mutations peuvent être de type non-sens (en bleu), faux-sens (en violet), délétions (en vert), dans des sites d'épissages en 3' ou 5' (en rouge) et induire un décalage du cadre de lecture (astérisque). Les 48 exons de *SMCHD1* sont numérotés (1 à 48). La localisation des régions codantes pour les domaines APTase et HD est indiquée au-dessus des exons (tiré de Lemmers et al., 2015a).

*SMCHD1* est donc un bon candidat pour expliquer les modifications épigénétiques observées chez les patients FSHD2 portant entre 10 et 20 répétitions D4Z4 (Lemmers et al., 2015a). En effet, bien que son rôle dans l'établissement et le maintien de la chromatine dans un état réprimé ne soit à ce jour pas connu, il a été montré que SMCHD1 interagit indirectement avec les marques épigénétiques H3K9me3, par l'intermédiaire des protéines LRIF1 (*Ligand dependent nuclear receptor interacting factor 1*) ou HBiX1 (*Human inactive X chromosome 1*) et HP1γ (Brideau et al., 2015; Eberl et al., 2013; Hayakawa et al., 2003; Minc et al., 2000; Nikolov et al., 2011; Nozawa et al., 2013; Zeng et al., 2014) (Figure 11). Or, les loci préférentiellement régulés par LRIF1, HP1γ et la marque épigénétique H3K9me3 sont les télomères et les répétitions D4Z4 (Canudas et al., 2011; Déjardin and Kingston, 2009; Lemmers et al., 2012a; Zeng et al., 2001) (Figure 11). Ainsi, dans le cas d'un mécanisme

d'haploinsuffisance ou de dominants négatifs, la quantité de protéines SMCHD1 actives pouvant se fixer au niveau des répétitions D4Z4 est probablement insuffisante pour maintenir la chromatine dans un état réprimé (Figure 11). Malgré tout, 19% des patients FSHD2 avec une mutation dans *SMCHD1*, une hypométhylation des répétitions D4Z4 et un haplotype permissif sont asymptomatiques (Lemmers et al., 2012a).

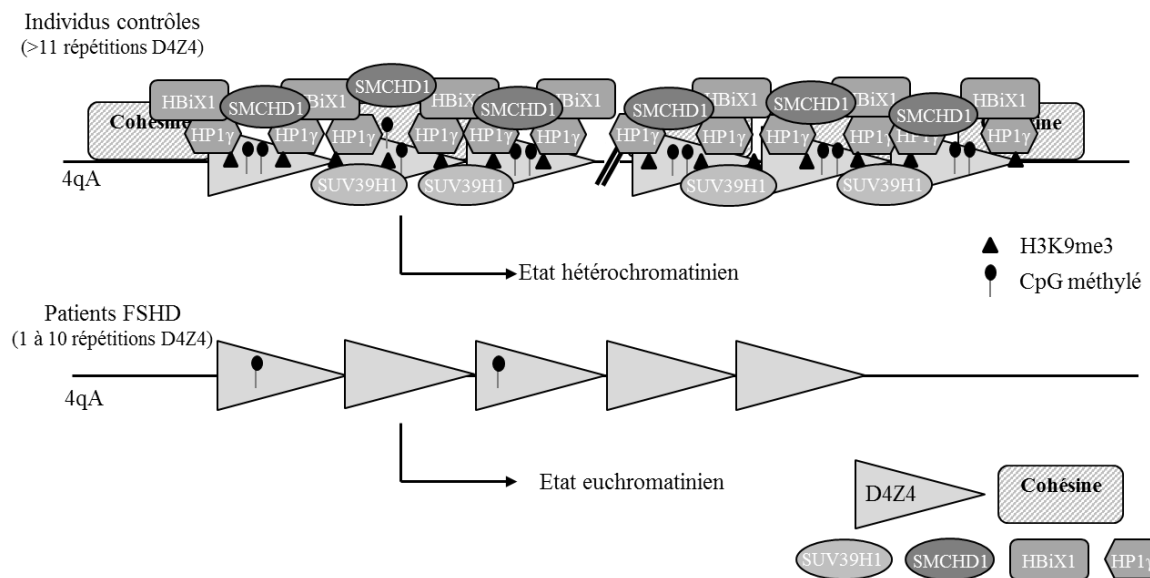


FIGURE 11. Modèle d'interaction entre les différents facteurs épigénétiques aux répétitions D4Z4 chez des individus contrôles et chez des patients FSHD. Chez les individus contrôles, les marques épigénétiques de type H3K9me3 (triangle noir) et CpG méthylé (rond noir), permettent le recrutement de SMCHD1 par l'intermédiaire des protéines LRIF1 (ou HBiX1) et HP1γ. Ces facteurs maintiennent la chromatine dans un état condensé. Chez les patients FSHD, ces marques épigénétiques sont perdues, ce qui compromet le recrutement des différentes protéines énoncées précédemment. La perte de ces marques épigénétiques chez certains patients FSHD2 peut être liée à une mutation dans *SMCHD1* (adapté de Himeda et al., 2015).

Par ailleurs, *SMCHD1* est un modificateur de la sévérité de la pathologie dans les familles FSHD1 à 7-10 répétitions D4Z4. En effet, la combinaison d'une contraction des répétitions D4Z4 et d'une mutation dans *SMCHD1* est associée à un développement précoce et une progression plus rapide de la pathologie (Larsen et al., 2015; Sacconi et al., 2013).

Récemment, des mutations dans *SMCHD1* ont été montrées comme impliquées dans le développement de l'arhinie (absence congénitale partielle ou totale de nez) et le syndrome de Bosma (Gordon et al., 2017; Shaw et al., 2017). Bien que certaines des mutations soient identiques à celles identifiées chez des patients FSHD, aucune atteinte musculaire n'est rapportée dans les cas d'arhinie ou de syndrome de Bosma. Inversement, une absence congénitale partielle ou totale de nez n'est pas décrite chez les patients FSHD. Ces

observations soulèvent donc des questions quant au rôle exact de SMCHD1 dans le développement de ces pathologies au phénotype très différent.

### ***DNMT3B***

Uniquement 80% des cas de FSHD2 peuvent s'expliquer par des mutations dans *SMCHD1*. D'autres défauts dans des facteurs génétiques ont donc été recherchés. Récemment, des mutations dans le modulateur épigénétique *DNMT3B* (*DNA (cytosine-5)-methyltransferase 3 beta*) localisé sur le chromosome 20, ont été mises en évidence chez des patients FSDH2 (van den Boogaard et al., 2016).

*DNMT3B* est un gène de 23 exons appartenant à la famille des ADN méthyl-transférases 3 (DNMT3, qui compte les variants DNMT3A, DNMT3B et DNMT3L) et codant une protéine de 853 acides aminés pour un poids moléculaire de 95kDa. DNMT3B possède un domaine N-terminal régulateur composé d'un motif PWWP et d'une région en doigt de zinc riche en cystéine appelée domaine ADD (ATRX-DNMT3-DNMT3L) (Figure 12). Le domaine C-terminal est un domaine catalytique méthyltransférase avec six motifs conservés (I, IV, VI, VIII, IX et X) dans la famille des protéines DNMT (Bestor, 2000; Chen et al., 2004; Cheng and Blumenthal, 2008; Goll and Bestor, 2005; Lukasik et al., 2006; Qiu et al., 2002; Zhang et al., 2010) (Figure 12).

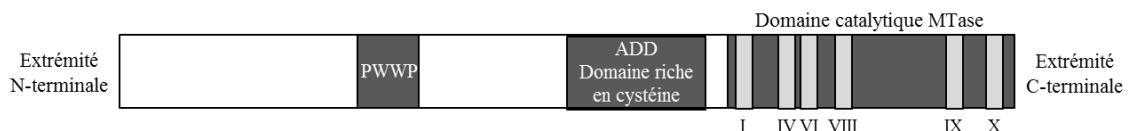


FIGURE 12. Structure de la protéine DNMT3B. DNMT3B possède un domaine N-terminal régulateur composé d'un motif PWWP et d'une région en doigt de zinc riche en cystéines appelée domaine ADD (ATRX-DNMT3-DNMT3L). Le domaine C-terminal est un domaine catalytique méthyltransférase avec six motifs conservés (I, IV, VI, VIII, IX et X) dans la famille des protéines DNMT.

Ces différentes structures sont impliquées dans la mise en place de la méthylation *de novo* des CpG (Goll and Bestor, 2005; Okano et al., 1999). Ce rôle est particulièrement important lors des étapes précoces de l'embryogenèse, notamment lors de la reprogrammation épigénétique et la mise en place de l'empreinte parentale (Geiman and Muegge, 2010; Kim et al., 2009a; Li et al., 1992, 1993). Le patron de méthylation établit par l'action de DNMT3A et DNMT3B est ensuite maintenu au cours des divisions cellulaires par DNMT1, une autre ADN méthyl-transférase (Goll and Bestor, 2005; Song et al., 2012). Cependant, DNMT3A et DNMT3B participent également au maintien de la méthylation, notamment aux loci des gènes de la

lignée germinale ou des éléments répétés (Chen et al., 2003; Goll and Bestor, 2005; Liang et al., 2002; Walton et al., 2011). De manière intéressante, l'un des mécanismes décrit permettant de cibler spécifiquement certaines régions génomiques lors de la mise en place de la méthylation *de novo*, est une interaction directe de DNMT3A et DNMT3B avec la protéine HP1 liée aux marques épigénétiques H3K9me3 (Lehnertz et al., 2003; Noh et al., 2016).

Chez certains patients FSHD2, des mutations dans ce modulateur épigénétique *DNMT3B* sont associées à une hypométhylation des répétitions D4Z4 (van den Boogaard et al., 2016; Das and Chadwick, 2016; Huichalaf et al., 2014). *DNMT3B* est donc aussi un bon candidat pour expliquer les modifications épigénétiques observées chez les patients FSHD2 avec un faible nombre de répétitions D4Z4 (entre 10 et 30). *DNMT3B* jouerait par ailleurs un rôle de modificateur de la sévérité de la pathologie chez les patients FSHD1. En effet, des patients avec neuf répétitions D4Z4 et porteurs d'une mutation dans *DNMT3B* ont un score de sévérité clinique corrigé par l'âge, plus important que des patients portant uniquement neuf répétitions D4Z4. Cependant, une mutation dans *DNMT3B* avec un faible nombre de répétitions D4Z4 sur le chromosome permissif 4qA n'est pas toujours suffisante pour développer la FSHD (van den Boogaard et al., 2016).

Parmi les deux principales modifications épigénétiques, à savoir la méthylation de l'ADN et la modification post-traductionnelle de la queue des histones, il est difficile à ce jour de savoir laquelle prédomine dans la FSHD. En effet, les modifications spécifiques des histones servent généralement de marques pour le recrutement d'ADN méthyltransférases qui permettent alors la méthylation de l'ADN. Mais dans certains cas, la méthylation de l'ADN peut également promouvoir la modification post-traductionnelle spécifique de la queue des histones (Cedar and Bergman, 2012; Jones et al., 1998; Lachner et al., 2001; Nan et al., 1998; Sarraf and Stancheva, 2004). La perte de l'une, de l'autre, voire de ces deux marques, pourrait être suffisante pour lever une répression transcriptionnelle au locus FSHD (Das and Chadwick, 2016; Li, 2002). Ces hypothèses restent à confirmer.

En résumé, il existe deux classes de FSHD génétiquement différentes mais phénotypiquement identiques. Les patients FSHD1 (95% des patients) présentent une contraction des répétitions D4Z4 sur le chromosome permissif 4qA, associée à une perte des marques épigénétiques répressives au niveau de ce macrosatellite (Figure 13). Les patients FSHD2 (5% des patients), quant à eux, ne présentent pas de contraction des répétitions D4Z4, mais sont hypométhylés et

ont perdu la marque répressive H3K9me3 aux niveaux des motifs D4Z4, en raison d'une mutation dans le modulateur épigénétique *SMCHD1* ou le modulateur épigénétique *DNMT3B* ou d'autres facteurs encore non identifiés (Figure 13). Le point commun entre la FSHD1 et la FSHD2 est donc une perte des marques épigénétiques répressives aux répétitions D4Z4 sur le chromosome permissif 4qA (Figure 13).

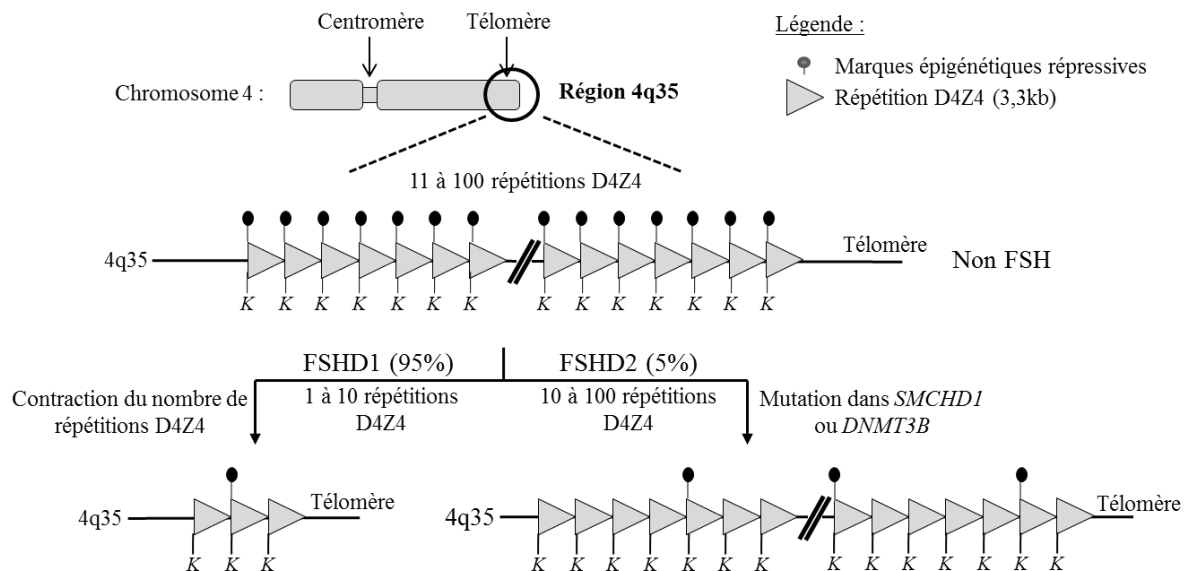


FIGURE 13. Caractéristiques génétiques et épigénétiques des patients FSHD1 et FSHD2. Le locus FSHD est localisé en région subtélomérique 4q35. Dans cette région à 40-60kb du télomère, se trouve un macrosatellite d'un motif de 3,3 kb appelé D4Z4 (représenté par un triangle gris), délimité par les sites de restriction reconnue par l'enzyme *KpnI* (K). Les individus contrôles possèdent dans cette région entre 11 et 150 répétitions D4Z4. Les patients FSHD1 possèdent, quant à eux, entre 1 et 10 répétitions D4Z4. Par ailleurs, cette contraction des répétitions D4Z4 est associée à une perte des marques épigénétiques répressives (méthylation de l'ADN, H3K9me3). Les patients FSHD2 ne sont pas contractés au niveau des répétitions D4Z4 mais ont également perdu les marques épigénétiques répressives en raison de mutations dans les modificateurs épigénétiques *SMCHD1* ou *DNMT3B* ou d'autres facteurs non identifiés.

### 1.3.5.1.3. Activité insulateur des répétitions D4Z4

Les insulateurs sont des séquences d'ADN avec deux fonctions distinctes non-exclusives. Lorsque l'insulateur est situé entre un *enhancer* et un promoteur, il permet de bloquer l'action de l'*enhancer* sur le promoteur. Les *enhancers* sont des séquences d'ADN allant d'une dizaine à quelques centaines de nucléotides, sur lesquels se fixent des protéines qui permettent de stimuler la transcription de gènes, parfois localisés à une distance de plus de 100kb du gène. De ce fait, bloquer l'action d'un *enhancer* sur un promoteur a pour conséquence de réguler négativement la transcription du gène en aval. Dans d'autres situations, l'insulateur

peut jouer le rôle d'une « barrière chromatinienne » en empêchant la diffusion de l'hétérochromatine le long de la chromatine (West et al., 2002). Les répétitions D4Z4 ont ces propriétés d'insulateur (Ottaviani et al., 2009a).

Plus précisément, chaque répétition D4Z4 contient un site de fixation pour la protéine CTCF (*CCCTC-binding Factor*) (position 382 à 814 à partir du site de restriction *KpnI* de chaque répétition D4Z4). CTCF est une protéine multifonctionnelle intervenant dans la répression ou l'activation de la transcription. Elle est impliquée dans les mécanismes insulateurs et favorise l'attachement de séquences spécifiques à la périphérie nucléaire (Filippova, 2008; Gaszner and Felsenfeld, 2006; Yusufzai et al., 2004). La fixation de CTCF sur une séquence ADN est altérée par la méthylation de l'ADN (Bell and Felsenfeld, 2000; Hark et al., 2000; Kanduri et al., 2000; Wang et al., 2012). La protéine CTCF est nécessaire mais pas suffisante à la régulation de l'activité insulateur des répétitions D4Z4. En effet, une autre interaction, avec les lamines A/C le long des répétitions D4Z4, à proximité de la séquence reconnue par CTCF, intervient également dans la régulation de cette activité insulateur du motif D4Z4 (Ottaviani et al., 2009a, 2009b, 2010). Ainsi, il est supposé que dans un contexte non contracté et hyperméthylé, la protéine CTCF n'est pas liée aux répétitions D4Z4. Les répétitions D4Z4 pourraient alors jouer un rôle de répresseur. Au contraire, une réduction du nombre de répétitions D4Z4, associée à une hypométhylation, favorise la fixation de CTCF et des lamines A/C. La fonction biologique des répétitions D4Z4 est modifiée. Les motifs D4Z4 jouent alors le rôle d'un insulateur et peuvent protéger une partie de la région 4q35 de la diffusion de l'hétérochromatine associée à la région subtélomérique (Ottaviani et al., 2009a, 2010).

#### **1.3.5.1.4. Localisation nucléaire des répétitions D4Z4**

Un autre niveau de régulation de l'expression des gènes est celui de l'architecture nucléaire, c'est-à-dire ici la localisation de la chromatine dans le noyau mais également son interaction dynamique avec d'autres composants nucléaires (Andrulis et al., 1998; Cremer and Cremer, 2001; Csink and Henikoff, 1996; Francastel et al., 2000; Jackson, 2003; Marshall, 2003). Le noyau est structuré en différents compartiments dans lesquels une fonction bien spécifique, comme la transcription ou la réplication, est réalisée (Spector, 2001). Ainsi, un chromosome ou une région chromosomique occupe un espace distinct et défini dans le noyau en fonction de sa densité en séquences codantes, de son activité transcriptionnelle, mais aussi de sa taille

(Sun et al., 2000; Tanabe et al., 2002). De façon simplifiée, les régions riches en hétérochromatine sont principalement retrouvées à la périphérie nucléaire au contraire des régions riches en euchromatine dispersées dans le noyau (Cockell and Gasser, 1999; Finlan et al., 2008; Kalverda et al., 2008; Reddy et al., 2008; Shopland et al., 2003). L'accès des facteurs de transcription à un gène est alors fonction de la position relative du gène dans le noyau (Baxter et al., 2002; Kurz et al., 1996). Cette organisation tridimensionnelle du noyau dépend d'une interaction entre des segments chromosomiques et la lamina qui recouvre la surface interne de l'enveloppe nucléaire (Guelen et al., 2008; Malhas et al., 2007; Marshall, 2003; Mattout-Drubezki and Gruenbaum, 2003). Ainsi, il a été proposé que la contraction des répétitions D4Z4 pourrait altérer la localisation de la région 4q35 dans le noyau et *de facto* modifier le profil transcriptionnel des gènes localisés dans cette région chromosomique (Bickmore et al., 2003; Masny et al., 2004; Tam et al., 2004; Winokur et al., 2003a).

La région subtélomérique 4q35, chez les patients FSHD et les individus contrôles, est localisée à proximité de l'enveloppe nucléaire dans un compartiment dit d'hétérochromatine. Au contraire, la région homologue subtélomérique 10q35 présente une localisation nucléaire plus centrale (Masny et al., 2004; Ottaviani et al., 2009b; Tam et al., 2004). Cette localisation de la région 4q35 à proximité de l'enveloppe nucléaire chez les patients FSHD et les individus contrôles, dépend d'une interaction avec les lamines A/C constituant la lamina (Guelen et al., 2008; Masny et al., 2004; Ottaviani et al., 2009b, 2010; Tam et al., 2004). Elle est commune à différents types cellulaires (Masny et al., 2004; Tam et al., 2004).

Cependant, bien que la région 4q35 soit localisée à proximité de l'enveloppe nucléaire chez les patients FSHD et les individus contrôles, une contraction des répétitions D4Z4 modifie légèrement cette localisation nucléaire du locus 4q35 à proximité de l'enveloppe nucléaire. En effet, un locus 4q35, avec un grand nombre de répétitions D4Z4, présente une localisation un peu moins périphérique qu'un locus 4q35 avec peu de répétitions D4Z4 (Ottaviani et al., 2009b) (Figure 14). Plus précisément, le locus 4q35 des individus contrôles est attaché à l'enveloppe nucléaire par une interaction entre les lamines A/C et une séquence localisée en amont des répétitions D4Z4 (Guelen et al., 2008; Masny et al., 2004; Ottaviani et al., 2009b, 2010; Tam et al., 2004) (Figure 14). Ce système d'ancrage à l'enveloppe nucléaire confère une certaine mobilité à la région 4q terminale. Chez les patients FSHD, les répétitions D4Z4 sont contractées et hypométhylées, ce qui permet le recrutement de CTCF (Ottaviani et al., 2009a, 2009b, 2010; van Overveld et al., 2003). Or la protéine CTCF, outre son implication

dans l'activité insulateur de D4Z4, et de par son association aux lamines A/C, contribue également à l'ancrage à l'enveloppe nucléaire des motifs D4Z4 (Ottaviani et al., 2009b, 2010) (Figure 14). Ainsi, chez les patients FSHD, la combinaison d'un ancrage médié par une séquence localisée en amont des répétitions D4Z4 et d'un ancrage médié par la protéine CTCF, le tout dépendant des lamines A/C, limite la diffusion du locus 4q35 vers le centre du noyau (Ottaviani et al., 2009b, 2010) (Figure 14). Ce changement d'environnement périphérique pourrait impacter l'expression des gènes avoisinants.

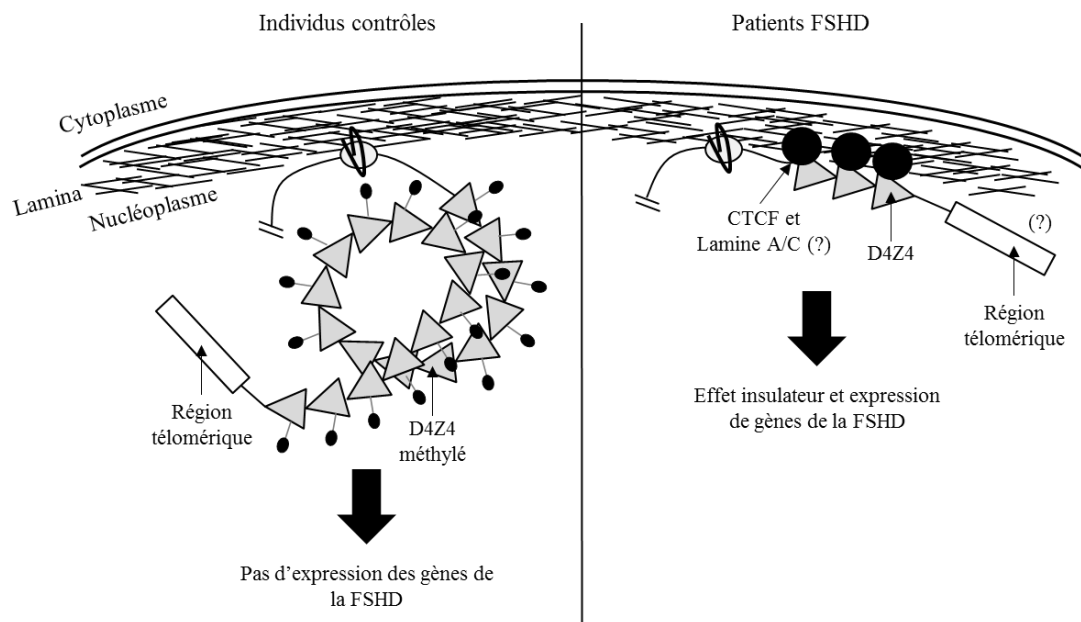


FIGURE 14. Modèle de localisation nucléaire de la région subtélomérique 4q35 chez des individus contrôles et des patients FSHD. Chez les individus contrôles, le locus 4q35 est attaché à l'enveloppe nucléaire par une interaction entre les lamines A/C et une séquence (rond blanc) localisée en amont des répétitions D4Z4 (rectangle gris). Chez les patients FSHD, la perte des marques épigénétiques permet le recrutement de CTCF (rond noir) à chaque répétition D4Z4. Outre le rôle insulateur de CTCF, l'association de CTCF et des lamines A/C contribue à la mise en place d'un second ancrage à l'enveloppe nucléaire des motifs D4Z4. Ce changement de localisation nucléaire peut induire des dérégulations géniques au locus 4q35 (adapté de Ottaviani et al., 2009b, 2010).

Par ailleurs, dans le noyau interphasique, la chromatine est organisée en boucles attachées au squelette ou à la matrice nucléaire *via* des séquences d'ADN spécifiques appelées S/MARs (*Scaffold / Matrix Associated Regions*) (Buongiorno-Nardelli et al., 1982; Cockerill and Garrard, 1986; McCready et al., 1979; Mirkovitch et al., 1984; Vassetzky et al., 2000a). Ces séquences d'ADN de 200 à 1000bp, généralement riches en AT ou GC, correspondent souvent à des fragments d'ADN courbes qui peuvent s'ouvrir localement (Fiorini et al., 2006; Kohwi-Shigematsu and Kohwi, 1990; Liebich et al., 2002; Razin, 2001; Vassetzky et al., 2000a, 2000b). Les S/MAR sont localisés dans des régions non transcrites, et plus



précisément dans les introns, à proximité de régions codantes ou de séquences régulatrices (Cockerill and Garrard, 1986; Gasser and Laemmli, 1986; Käs and Chasin, 1987; Keaton et al., 2011; Vassetzky et al., 2000a). Certains S/MARs peuvent avoir un rôle d'insulateur (Antes et al., 2001; Bode et al., 2000; Gombert et al., 2003; Vassetzky et al., 2000a; West et al., 2002; Yusufzai et al., 2004).

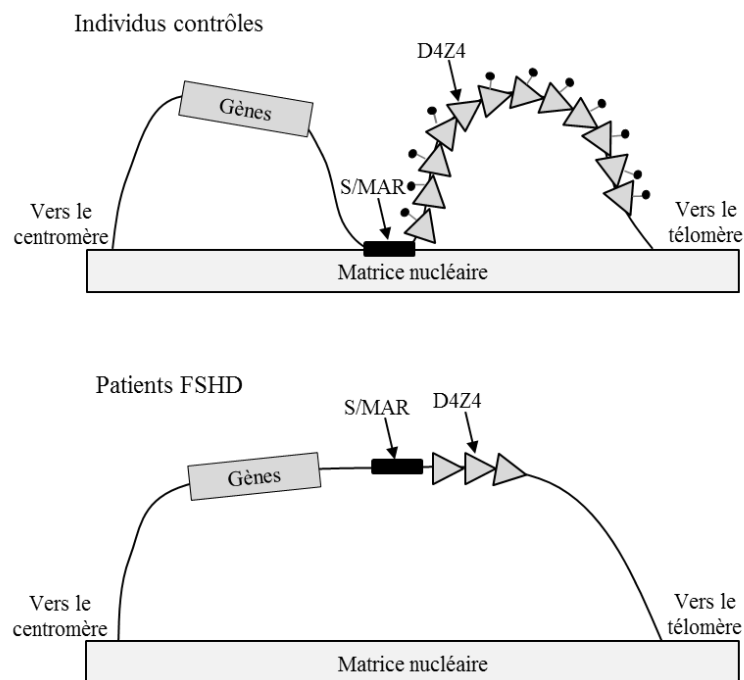


FIGURE 15. Modèle d'attachement de la chromatine à la membrane nucléaire interne au niveau de la région subtélomérique 4q35 chez des individus contrôles et des patients FSHD. Une séquence d'ADN spécifique S/MAR (rectangle noir) localisée juste en amont des répétitions D4Z4 (triangle gris) permet une interaction avec la matrice nucléaire. Cette interaction isole les répétitions D4Z4 au sein d'une boucle chromosomique. Chez les patients FSHD, la modification des marques épigénétiques (et la contraction des répétitions D4Z4) induit la perte de l'interaction entre la séquence S/MAR et la matrice nucléaire. Les répétitions D4Z4 ne sont plus isolées. Ce changement de localisation de la chromatine peut conduire à une dérégulation génique dans cette région (adapté de Petrov et al., 2006).

Le locus 4q35 contient une séquence S/MAR juste en amont des répétitions D4Z4 (Figure 15). Ainsi, dans des cellules non musculaires et des myoblastes contrôles, les répétitions D4Z4 sont isolées au sein d'une boucle (Petrov et al., 2006) (Figure 15). Chez les patients FSHD, l'interaction entre le S/MAR en amont des répétitions D4Z4 et la matrice nucléaire est diminuée d'environ 30%. Les répétitions D4Z4 et les gènes avoisinants du locus 4q35 sont alors regroupés sur la même boucle (Petrov et al., 2006) (Figure 15). L'absence d'interaction entre la séquence S/MAR au locus 4q35 et la matrice nucléaire chez les patients FSHD serait due aux modifications épigénétiques locales. En effet, la perte des marques épigénétiques répressives aux histones et l'hypométhylation de l'ADN réduisent le recrutement de protéines

associées aux S/MAR et impliquées dans l'interaction avec la matrice nucléaire, dont MECP2 (*Methyl Cytosine Binding Protein 2*) (Horike et al., 2005; Kisseljova et al., 2014; Klose et al., 2005; van Overveld et al., 2003; Strätling and Yu, 1999). Ce changement dans l'organisation des domaines de la chromatine pourrait conduire à une dérégulation de l'expression des gènes en *cis* au locus FSHD (Petrov et al., 2006, 2008).

#### **1.3.5.1.5. Rôle de la taille des télomères dans la FSHD**

En raison de la localisation des répétitions D4Z4 en région subtélomérique, un effet de position télomérique a également été recherché (Stadler et al., 2013). Les télomères, ces séquences d'ADN non codantes répétées en tandem (TTAGGG chez l'Homme) situées à l'extrémité de la chromatine et très conservées, voient leur longueur diminuer au cours des cycles cellulaires (Harley, 1991; Harley et al., 1990; Hastie et al., 1990; de Lange, 2004; Meyne et al., 1989; Moyzis et al., 1988; Shay and Wright, 2011). Lorsque les télomères atteignent une taille critique, la cellule rentre dans un état de quiescence ou en sénescence (Harley, 1991; Shay, 2003). Le raccourcissement des télomères peut être contré par l'action d'une enzyme, la télomérase, qui permet de maintenir la longueur des télomères dans les cellules germinales et certaines cellules somatiques cancéreuses (Collins and Mitchell, 2002; Greider, 1991; Greider and Blackburn, 1985; Shay and Wright, 2011). Les télomères ont un rôle protecteur contre la perte de l'information génétique, mais aussi contre les exonucléases et évitent que les extrémités chromatiniennes ne soient reconnues comme des cassures double-brin (Shay, 2003; Shay and Wright, 2011). Une autre conséquence du raccourcissement des télomères est la modification de l'expression des gènes environnants. En effet, les télomères sont des régions fortement condensées. L'hétérochromatine télomérique, en l'absence de barrières subtélomériques, peut se propager et induire une répression transcriptionnelle des gènes à proximité (Ottaviani et al., 2008). C'est l'effet de position télomérique (Baur et al., 2001; Gottschling et al., 1990; Koering et al., 2002; Lou et al., 2009; Pedram et al., 2006; Zakian, 2012). Logiquement, plus le gène sera éloigné de la région télomérique, moins la réduction de la taille du télomère aura d'impact sur son niveau d'expression.

Au niveau du chromosome 4q, la longueur des télomères, dans un système artificiel, est inversement corrélée au niveau d'expression des gènes localisés en 4q35 (Robin et al., 2015; Stadler et al., 2013). La réduction de la longueur des télomères pourrait donc aider à la

modification de l'expression des gènes en 4q35 dans la FSHD (Robin et al., 2015; Stadler et al., 2013). En effet, et malgré le fait qu'aucune étude n'ait encore démontré une différence dans la longueur des télomères entre patients FSHD et individus contrôles, la réduction des télomères est accélérée dans certaines dystrophies musculaires en raison de la forte division des cellules satellites au cours des processus de régénération, plus nombreux. Cela pourrait être également le cas dans la FSHD (Decary et al., 2000). De plus, la variation de la longueur des télomères inter-individus pourrait être un élément explicatif de la pénétrance de la FSHD (Stadler et al., 2013). En effet, des individus avec des télomères courts pourraient développer la FSHD plus tôt et avec une plus grande probabilité que des individus avec des télomères plus longs.

#### **1.3.5.1.6. Régulations épigénétiques au locus 4q35 par des complexes protéiques et des ARN non codants**

Un autre niveau de régulation de la transcription des gènes fait intervenir des facteurs protéiques qui reconnaissent des séquences spécifiques d'ADN. Ces facteurs protéiques peuvent activer ou au contraire empêcher l'expression de gènes. De ce fait, un des mécanismes proposé pour la FSHD est une modification des interactions protéiques aux répétitions D4Z4 impactant le profil d'expression des gènes du locus 4q35 (Gabellini et al., 2002).

Chaque répétition D4Z4 contient une séquence de 27bp appelée DBE (*D4Z4 Binding Element*) (CCATTCATGAAGGGGTGGAGCCTGCCT) (Gabellini et al., 2002). Cette séquence est un élément de répression transcriptionnelle. Elle permet le recrutement de protéines du complexe Polycomb dont notamment le facteur de transcription YY1 (*Yin Yang 1*), la protéine EZH2 (*Enhancer of Zeste Homolog 2*), la protéine associée à la chromatine non-histone HMGB2 (*High-Mobility Group Protein B2*) et la nucléoline qui compose plusieurs complexes de facteurs de transcription (Cabianca et al., 2012; Déjardin and Kingston, 2009; Gabellini et al., 2002; Hanakahi et al., 1997; Infante et al., 2015; Margueron and Reinberg, 2011; Schulz et al., 2001; Thomas and Seto, 1999; Thomas and Travers, 2001; Ying et al., 2000). De manière intéressante, HMGB2 peut interagir avec HP1 pour former un complexe de répression transcriptionnelle (Lehming et al., 1998). Les protéines du complexe Polycomb sont également retrouvées enrichies en amont des répétitions D4Z4, au niveau d'un

élément appelé *NDE* (*NonDeleted Element*) (Cabianca et al., 2012). Le complexe Polycomb, et le complexe antagoniste Trithorax, constituent des groupes d'éléments qui modifient l'architecture de la chromatine et par conséquent l'expression des gènes à proximité de leurs sites de fixation (Ringrose and Paro, 2004; Schuettengruber et al., 2007). La présence de protéines du complexe Polycomb aux répétitions D4Z4 et en amont des répétitions D4Z4 chez les individus contrôles induit une répression des gènes avoisinants par mise en place d'hétérochromatine (Cabianca et al., 2012; Gabellini et al., 2002).

Cependant, le motif DBE n'est pas uniquement un site de recrutement des protéines du complexe Polycomb. Chez les patients FSHD, la séquence démarrant au motif DBE et se poursuivant au-delà du site NDE localisé en amont des répétitions D4Z4, est transcrite en un long ARN non codant (dit antisens) nommé *DBE-T* (*D4Z4 Binding Element-Transcript*) (Block et al., 2012; Cabianca et al., 2012) (Figure 16). *DBE-T* joue un rôle positif sur la transcription des gènes en *cis* au locus 4q35. En effet, il permet de modifier la structure chromatinienne en recrutant des protéines du complexe Trithorax, dont *ASH1L* (*ASH1 Like histone lysine methyltransferase*) (Cabianca et al., 2012; Gregory et al., 2007; Schuettengruber et al., 2007). Par ailleurs, dans un schéma de boucle rétro-positive, *ASH1L* active la transcription de *DBE-T* (Cabianca et al., 2012).

Ainsi, il est supposé que chez les individus contrôles, le recrutement de protéines du complexe Polycomb principalement aux séquences DBE de chaque répétition D4Z4, ainsi qu'au motif NDE, induit une répression de la transcription des gènes au locus 4q35 par établissement d'une structure d'hétérochromatine (Cabianca et al., 2012; Gabellini et al., 2002) (Figure 16). Chez les patients FSHD, la réduction du nombre de répétitions D4Z4 a pour conséquence une réduction de la fixation du nombre de complexes répresseurs totaux et une transcription de l'ARN non codant *DBE-T* (Figure 16). La production de *DBE-T* permet alors le recrutement de protéines du complexe Trithorax et donc une levée de la répression des gènes de la région 4q35 par perte des marques épigénétiques répressives (Bodega et al., 2009; Cabianca et al., 2012; Gabellini et al., 2002) (Figure 16). Cependant, un rôle de *DBE-T* dans le développement de la FSHD est à moduler. En effet, des patients FSHD présentant des délétions majeures juste en amont des répétitions D4Z4, incluant la séquence NDE, ont été décrits (Deak et al., 2007; Lemmers et al., 2003). Chez ces patients, un rôle de *DBE-T* dans le développement de la FSHD est donc exclu.

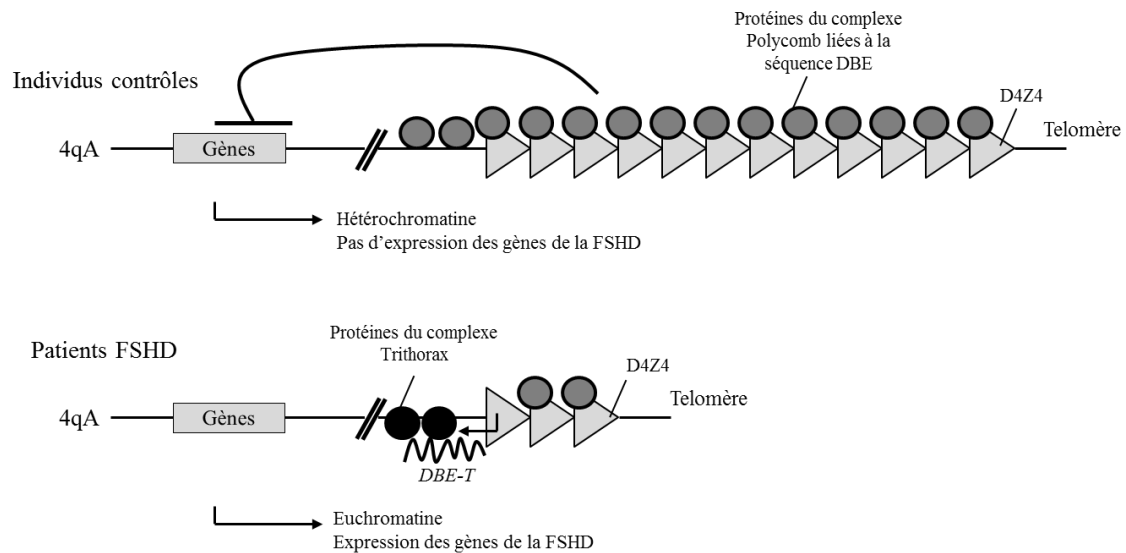


FIGURE 16. Modèle de la régulation transcriptionnelle des gènes au locus 4q35 médiée par les protéines du complexe Polycomb et Trithorax chez des individus contrôles et des patients FSHD. Chez les individus contrôles, des protéines du complexe Polycomb (rond gris foncé) sont recrutées aux séquences DBE présentes au niveau de chaque répétition D4Z4 (triangle gris). Le complexe Polycomb permet une répression transcriptionnelle des gènes du locus 4q35. Chez les patients FSHD, la contraction des répétitions D4Z4 réduit le recrutement du nombre de complexes Polycomb totaux et permet une transcription de l'ARN non codant *DBE-T*. La production de *DBE-T* induit le recrutement de protéines du complexe Trithorax (rond noir). La répression transcriptionnelle des gènes en 4q35 est levée (adapté de Cabianca et al., 2012; Gabellini et al., 2002).

La conséquence de la perte des marques épigénétiques associées à l'hétérochromatine est donc une potentielle levée de la répression transcriptionnelle des gènes localisés à proximité du locus FSHD. Cette idée est confortée par la découverte d'un *enhancer* au niveau du motif D4Z4 (position 1 à 170) (Petrov et al., 2008). L'activité de cet *enhancer* est dépendante de la fixation du facteur de transcription KLF15 (*Krippel-Like Factor 15*) (Dmitriev et al., 2011). L'expression de KLF15 est contrôlée par les facteurs de transcription myogéniques comme MyoD. Par conséquent, KLF15 est exprimé durant la différenciation musculaire. Par ailleurs, les facteurs de transcription myogéniques sont plus exprimés dans les cellules de patients FSHD comparées aux cellules d'individus contrôles (Broucqsault et al., 2013; Dmitriev et al., 2011). Il en résulte une augmentation plus importante de la fréquence de transcription de KLF15 dans les myoblastes et myotubes de patients FSHD, en comparaison aux myoblastes et myotubes d'individus contrôles. L'activité *enhancer* des répétitions D4Z4 est alors plus importante dans les cellules de patients FSHD (Dmitriev et al., 2011).

Ainsi, il est supposé qu'au cours de la différenciation musculaire, en présence de marques épigénétiques répressives (méthylation de l'ADN, modification post-traductionnelles de la queue des histones) (associées ou non à la contraction des répétitions D4Z4), l'*enhancer* n'est

pas accessible à la faible quantité de protéines KLF15. De ce fait, la transcription des gènes au locus FSHD n'est pas favorisée (Figure 17). A l'inverse, chez les patients FSHD, une forte expression de KLF15, associée à la perte des marques épigénétiques répressives aux répétitions D4Z4, permet la fixation de KLF15 à l'*enhancer*, et par conséquent, l'activation transcriptionnelle des gènes en 4q35 dont la dérégulation confère un phénotype FSHD (Dmitriev et al., 2011) (Figure 17).

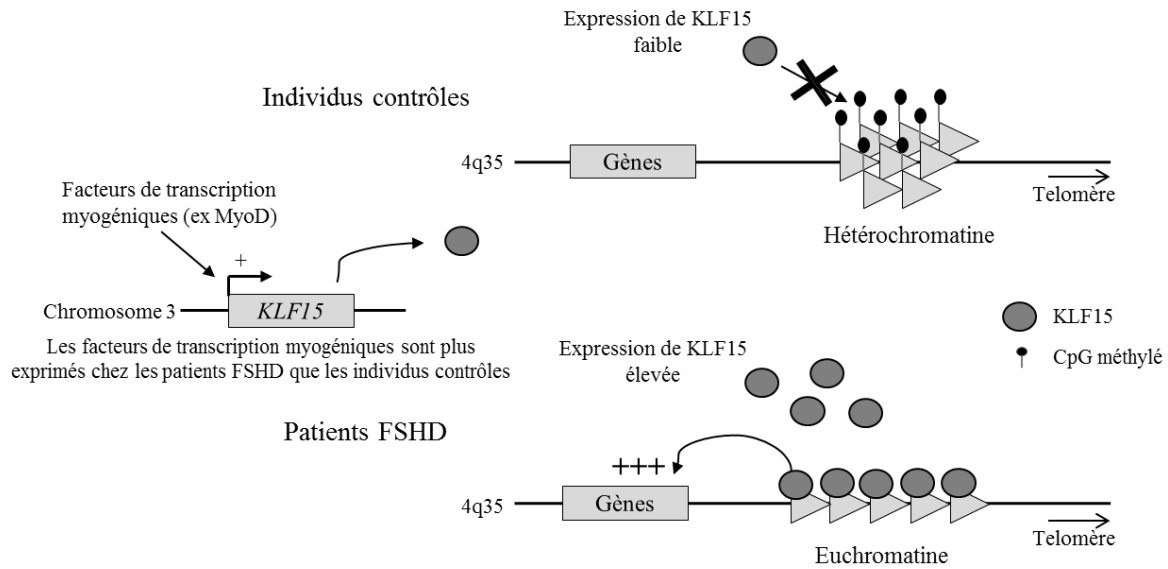


FIGURE 17. Modèle d'activation de l'*enhancer* aux répétitions D4Z4 par le facteur de transcription KLF15 au cours de la différenciation musculaire chez des patients FSHD ou des individus contrôles. L'expression de KLF15 est contrôlée par les facteurs de transcription myogéniques. Chez les individus contrôles, peu de protéines KLF15 sont produites. De plus, les marques chromatiniennes répressives empêchent la fixation de KLF15 aux séquences *enhancer* présentes au niveau de chaque répétition D4Z4. La transcription des gènes du locus FSHD n'est pas favorisée. A l'inverse, chez les patients FSHD, la fréquence de transcription de *KLF15* est plus élevée en raison d'une expression plus importante des facteurs myogéniques. Cette forte expression de KLF15, associée à une perte des marques épigénétiques répressives, permet la fixation de KLF15 à l'*enhancer*. Il en résulte une activation transcriptionnelle des gènes en 4q35 et un phénotype FSHD (adapté de Dmitriev et al., 2011)

### 1.3.6. Gènes candidats pour la FSHD

L'une des priorités de la recherche sur la FSHD est de déterminer quels sont les gènes dérégulés dans la FSHD et de comprendre comment ces dérégulations peuvent conduire au phénotype FSHD. Les modèles précédemment décrits supposent qu'une combinaison entre la contraction des répétitions D4Z4 et la perte des marques épigénétiques répressives pourrait induire une modification globale de la structure de la chromatine conduisant, par effet de position, à un changement du profil de transcription des gènes à proximité des répétitions

D4Z4 (Hewitt et al., 1994; Jiang et al., 2003; Winokur et al., 1994). Bien que plusieurs gènes soient localisés en aval des répétitions D4Z4, seuls certains sont proposés comme dérégulés dans la FSHD : *ANT1* (ou *SLC25A4*), *FAT1*, *FRG1*, *FRG2* (Tableau 3, Figure 18).

TABLEAU 3. Liste des gènes présents en région subtélomérique 4q35. Pour chaque gène, son nom ainsi que sa position en kb par rapport à la première répétition D4Z4 sont indiqués. Les gènes potentiellement dérégulés dans la FSHD sont indiqués en gras (adapté de Masny et al., 2010).

Gène	Nom	Distance en kb des répétitions D4Z4
<b><i>FRG2</i></b>	<b><i>FSHD region gene 2</i></b>	<b>55</b>
<i>TUBB4Q</i>	<i>Tubulin 4q pseudogene</i>	97
<i>DUX4c</i>	<i>Double homeobox protein 4c</i>	42
<b><i>FRG1</i></b>	<b><i>FSHD region gene 1</i></b>	<b>141</b>
<b><i>FAT1</i></b>	<b><i>FAT tumor suppressor homolog</i></b>	<b>3508</b>
<i>F11</i>	<i>Coagulation factor XI</i>	3816
<i>KLKB1</i>	<i>Kallikrein B</i>	3855
<i>CYP4V2</i>	<i>Cytochrome P450, family 4</i>	3890
<i>TLR3</i>	<i>Toll-like receptor 3</i>	4012
<i>SORBS2 (ARGP2)</i>	<i>Sorbin and SH3 domain containing 2</i>	4540
<i>ALP (PDLIM3)</i>	<i>PDZ and LIM domain 3</i>	4581
<i>LRP2BP</i>	<i>LRP2-binding protein</i>	4717
<i>SNX25</i>	<i>Sorting nexin 25</i>	4888
<b><i>ANT1 (SLC25A4)</i></b>	<b><i>Adenine Nucleotide Translocase 1</i></b>	<b>4936</b>
<i>HELT</i>	<i>Hey-like transcriptional repressor</i>	5060
<i>IRF2</i>	<i>Interferon regulatory factor 2</i>	5702
<i>ING2</i>	<i>Inhibitor of growth family, member 2</i>	6574

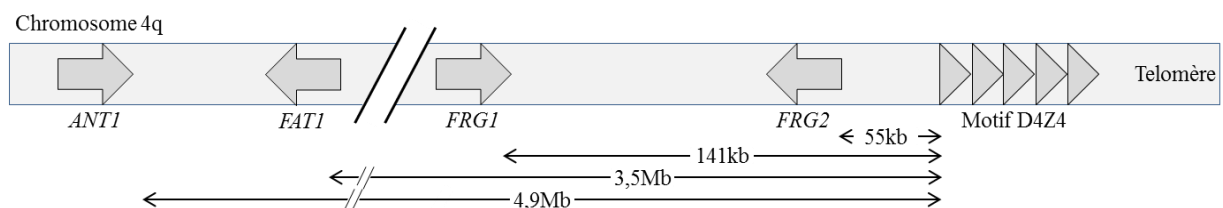


FIGURE 18. Localisation des différents gènes en région subtélomérique 4q35 pouvant être impliqués dans le développement de la FSHD. Pour chaque gène (représenté par une flèche grise), sa distance vis-à-vis de la première répétition D4Z4 (triangle gris) est reportée en kb ou Mb (adapté de Masny et al., 2010).

### 1.3.6.1. *FRG1*

*FRG1* (*FSHD Region 1*) a été proposé pour la première fois en 1996 comme un gène candidat de la FSHD (van Deutekom et al., 1996b). Ce gène est localisé à 141kb en amont des

répétitions D4Z4 sur le chromosome 4q (Masny et al., 2010). L'ORF *FRG1*, composé de neuf exons, produit un transcrit de 1042bp traduit en une protéine de 258 acides aminés (van Deutekom et al., 1996b). FRG1 (29kDa) possède deux NLS (*Nuclear Localization Signal*), un motif lipocaline, impliqué dans le transport de petites particules hydrophobes et un domaine central *fascin-like*, permettant une interaction avec le cytosquelette d'actine (Edwards and Bryan, 1995; Flower, 1996; Kureishy et al., 2002; Liu et al., 2010).

FRG1 est une protéine multifonctionnelle à localisation nucléaire et cytoplasmique (Hanel et al., 2011; Liu et al., 2010; Sun et al., 2011). Au niveau nucléaire, la protéine FRG1 est retrouvée aux sites de synthèse des ARNs (nucléoles, corps de Cajal) où elle intervient comme activateur du complexe d'épissage des ARNs (Bessonov et al., 2010; Jurica et al., 2002; van Koningsbruggen et al., 2004, 2007; Makarov et al., 2002; Rappsilber et al., 2002). FRG1 interagit également avec TAP (*Transporter associated with Antigen Processing*), une protéine impliquée dans l'export nucléaire, dont celui des ARNs (Grüter et al., 1998). Au niveau cytoplasmique, FRG1 interagit avec le cytosquelette d'actine (Liu et al., 2010; Sun et al., 2011). Ainsi, il est supposé que dans le noyau, FRG1 s'associe à l'ARNm en cours de synthèse pour permettre dans un premier temps son épissage puis dans un second temps son export par la voie TAP-NXT1. Une fois dans le cytoplasme, le complexe ARNm-FRG1 se fixe au cytosquelette d'actine pour rejoindre un lieu de traduction de l'ARNm (Sun et al., 2011). Dans la cellule musculaire, la protéine FRG1 est en plus retrouvée localisée au niveau des sarcomères, plus particulièrement des disques Z (Hanel et al., 2011; Liu et al., 2010).

Une surexpression de *FRG1* dans plusieurs modèles animaux (murin, *Xenopus* ou *Caenorhabditis elegans*) montre que FRG1 intervient dans le développement musculaire et vasculaire (Gabellini et al., 2006; Hanel et al., 2009; Liu et al., 2010; Wuebbles et al., 2009, 2010). Ces modèles permettent de récapituler la majorité des signes cliniques, musculaires et vasculaires, associés à FSHD à l'exception de l'asymétrie (Gabellini et al., 2006; Hanel et al., 2009; Liu et al., 2010; Wuebbles et al., 2009). Plus précisément, dans un modèle de souris surexprimant *FRG1*, des analyses histologiques ont montré une augmentation de la taille des fibres musculaires, la présence de noyaux centraux caractérisant une régénération musculaire suite à une lésion et une nécrose des fibres (Gabellini et al., 2006). Une surexpression de *FRG1* altère la myogenèse, ainsi que la prolifération et la migration des cellules satellites ou des myoblastes (Chen et al., 2011; Hanel et al., 2009; Wuebbles et al., 2009). Par ailleurs, une



surexpression de *FRG1* conduit à une augmentation de l'angiogenèse et une dilatation des vaisseaux sanguins (Wuebbles et al., 2009).

Ces modèles ont également permis d'étudier le rôle de *FRG1* dans l'apparition des symptômes dystrophiques, et ce à un niveau moléculaire. Il a notamment été montré que *FRG1*, en diminuant la stabilité de l'ARNm, régule négativement l'expression du facteur d'épissage *RBFOX1* (*RNA Binding Protein FOX1*) exprimé dans le cerveau, le muscle squelettique et le cœur (Damianov and Black, 2010; Fukumura et al., 2007; Kuroyanagi, 2009; McKee et al., 2005; Pistoni et al., 2013; Tang et al., 2009). Une conséquence de la diminution de *RBFOX1*, et donc de la surexpression de *FRG1*, est une augmentation du transcrit *CAPN3* (*Calpain 3*) où l'exon 6 est exclu (Pistoni et al., 2013). *CAPN3* appartient à la famille des protéases calcium-dépendantes solubles et non lysosomales. Cette protéine permet de maintenir l'intégrité du muscle et donc sa fonction, notamment au cours de la différenciation (Beckmann and Spencer, 2008). L'exon 6 de *CAPN3* code une séquence d'acides aminés impliqués dans l'auto-inhibition de la protéine (Beckmann and Spencer, 2008). Cette isoforme de *CAPN3* où l'exon 6 est exclu, n'est pas présente dans le muscle adulte contrôle mais semble intervenir au cours du développement et de la régénération musculaire. Une surexpression de *CAPN3* sans exon 6 dans un muscle adulte résulte en un développement musculaire anormal (Pistoni et al., 2013). A noter que les souris exprimant cette isoforme de *CAPN3* présentent le même phénotype que des souris surexprimant *FRG1*. De plus, le niveau d'expression de *CAPN3* sans exon 6 est corrélé au degré de dystrophie dans les différents muscles et ce à différents âges. Cela pourrait expliquer les différences de susceptibilité des muscles à la surexpression de *FRG1*, et donc la progression variable de la pathologie (Pistoni et al., 2013).

Dans le modèle de souris surexprimant *FRG1*, d'autres ARNm présentant des épissages aberrants ont été identifiés et peuvent expliquer le phénotype dystrophique (Buj-Bello et al., 2002; Chaudhuri et al., 2005; Gabellini et al., 2006; Galińska-Rakoczy et al., 2008; Gordon et al., 2000; Perry, 1998; Sancisi et al., 2014):

- *TNNT3* (*Fast Skeletal Muscle Troponin T3*) impliqué dans la contractibilité musculaire
- *MTMR1* (*Myotubularin-related Protein 1*) qui régule l'atrophie musculaire.

Ces transcrits aberrants ont également été détectés dans des cellules dérivées de patients FSHD sans pour autant pouvoir effectuer un lien direct avec une surexpression de *FRG1* (Gabellini et al., 2006).

Néanmoins, l'implication de *FRG1* dans la FSHD n'est pas clairement établie. En effet, ces modèles animaux constituent des cas de surexpression d'une protéine dans des cellules musculaires, ce qui peut induire un phénotype dystrophique indépendamment de la fonctionnalité réelle du gène. Par ailleurs, bien que certaines études aient montré une surexpression de *FRG1* dans des biopsies musculaires ou des cellules dérivant de patients FSHD, notamment au cours de la différenciation (Bodega et al., 2009; Broucqsault et al., 2013; Gabellini et al., 2002), de nombreuses autres études contredisent ces résultats, empêchant toutes conclusions quant au rôle de *FRG1* dans la FSHD (Arashiro et al., 2009; Celegato et al., 2006; van Deutekom et al., 1996b; Jiang et al., 2003; Klooster et al., 2009; Masny et al., 2010; Osborne et al., 2007; Winokur et al., 2003a). L'obtention de ces résultats contradictoires pourrait être due à la présence de plusieurs homologues de *FRG1* un peu partout dans le génome (Grewal et al., 1999). Le signal *FRG1* détecté dans ces études pourrait donc correspondre à des transcrits issus de ces séquences homologues, éliminant un effet en *cis* de la contraction des répétitions D4Z4 sur la transcription du gène *FRG1* localisé sur le chromosome 4 (Masny et al., 2010). Une action en *trans* sur l'expression de *FRG1* reste néanmoins une possibilité.

### **1.3.6.2. *FRG2***

*FRG2* (*FSHD Region 2*), localisé à 55kb en amont des répétitions D4Z4 sur le chromosome 4q, a été aussi proposé comme un gène candidat pour la FSHD (Masny et al., 2010; Rijkers et al., 2004). *FRG2*, composé de quatre exons, est transcrit en un ARNm de 2084bp traduit en une protéine de 278 acides aminés d'une masse de 30kDa (Rijkers et al., 2004). La protéine *FRG2* présente une localisation nucléaire mais sa fonction au noyau n'est pas connue à ce jour (Rijkers et al., 2004).

Le rôle de *FRG2* dans le développement de la FSHD est controversé. Certaines études ont détecté une expression de *FRG2* uniquement dans des muscles de patients FSHD ou dans des myoblastes dérivants de patients FSHD en cours de différenciation (Gabellini et al., 2002; Rijkers et al., 2004). Ces études mettent également en évidence une corrélation inverse entre le niveau d'expression de *FRG2* et le nombre de répétitions D4Z4, permettant d'avancer l'hypothèse d'un rôle inhibiteur des répétitions D4Z4 sur la transcription de *FRG2* (Gabellini et al., 2002; Klooster et al., 2009; Rijkers et al., 2004; Thijssen et al., 2014). Cependant,

*FRG2* possède de nombreux homologues un peu partout dans le génome, dont un homologue sur le chromosome 10q qui diffère seulement de quatre acides aminés. Aucun de ces travaux n'a confirmé que le transcrit détecté provenait bien du chromosome 4q et non du chromosome 10q (Gabellini et al., 2002; Rijkers et al., 2004). A l'inverse, certains auteurs montrent une absence de surexpression de *FRG2* dans des biopsies ou des myoblastes dérivant de patients FSHD (Broucqsault et al., 2013; Masny et al., 2010; Osborne et al., 2007). Par ailleurs, des individus avec des délétions majeures, jusqu'à 60kb en amont des répétitions D4Z4, incluant *FRG2*, développent tout de même un phénotype FSHD, excluant *FRG2* comme gène candidat majeur de la FSHD (Deak et al., 2007; Lemmers et al., 2003).

### **1.3.6.3. *ANT1***

*ANT1* (*Adenine Nucleotide Translocator 1 / ADP/ATP translocator*), localisé à 4,9 Mb en amont des répétitions D4Z4 sur le chromosome 4q, a été le premier gène candidat proposé pour la FSHD (Haraguchi et al., 1993; Li et al., 1989; Masny et al., 2010; Wijmenga et al., 1993). Ce gène contenant quatre exons, est transcrit en un ARNm de 1300bp et code une protéine de 298 acides aminés appartenant à la famille des protéines ANT (Doerner et al., 1997; Li et al., 1989). Les protéines ANT existent sous quatre isoformes, ANT1, ANT2, ANT3 et ANT4, homologues à 77-79% (Doerner et al., 1997; Dolce et al., 2005; Li et al., 1989; Stepien et al., 1992). Les protéines ANT, qui sont les protéines les plus abondantes dans la mitochondrie, catalysent l'échange entre les molécules d'ADP et d'ATP produites au cours de la respiration mitochondriale au niveau de la membrane interne de la mitochondrie (Klingenberg, 2008; Li et al., 1989). L'activité des protéines ANT est essentielle au maintien de l'activité de l'ATP synthase (Brand et al., 2005). L'expression des isoformes ANT est tissu-spécifique. La protéine ANT1 est exprimée uniquement dans le cerveau, le cœur et le tissu musculaire, ANT2 est abondante dans les cellules en prolifération, ANT3 est produite de manière ubiquitaire et la protéine ANT4 est retrouvée spécifiquement dans les cellules germinales (Brower et al., 2007; Doerner et al., 1997; Haraguchi et al., 1993). Néanmoins, dans un même tissu, plusieurs isoformes ANT peuvent être exprimées séquentiellement. Par exemple, le gène *ANT1* est fortement transcrit au cours des dernières étapes de la différenciation des myoblastes en myocytes alors qu'auparavant, c'est l'isoforme ANT2, absente des tissus différenciés, qui prédomine dans le myoblaste (Doerner et al., 1997; Stepien et al., 1992).

Des mutations dans *ANT1* ont été répertoriées dans plusieurs pathologies neuromusculaires mitochondriales dont l'ophtalmoplégie externe progressive (myopathie tubulaire mitochondriale) ou le syndrome de Senger (syndrome de cataracte congénitale-cardiomyopathie hypertrophique-myopathie mitochondriale) (Liu and Chen, 2013; Sharer, 2005). Ces pathologies ont permis de mettre en évidence les rôles directs ou indirects de *ANT1* dans la régulation de l'apoptose et la production d'espèces réactives à l'oxygène (Liu and Chen, 2013; Sharer, 2005).

L'implication de *ANT1* dans le développement de la FSHD est très controversée. En effet, alors que des études ont détecté une surexpression de *ANT1* dans des biopsies ou des myoblastes dérivants de patients FSHD (Gabellini et al., 2002; Laoudj-Chenivresse et al., 2005), d'autres travaux ne montrent aucune variation de l'expression de *ANT1* chez des patients FSHD comparés à des individus contrôles (Jiang et al., 2003; Klooster et al., 2009; Masny et al., 2010; Osborne et al., 2007).

#### **1.3.6.4. *FAT1***

*FAT1* (*FAT Atypical cadherin 1*), proposé également comme gène candidat de la FSHD, est localisé à 3,5Mb en amont de la première répétition D4Z4 sur le chromosome 4q (Dunne et al., 1995; Masny et al., 2010). Composé de 28 exons répartis sur une région de 139kb, *FAT1* est transcrit en un long ARNm de 15kb et traduit en une protéine transmembranaire de 4 594 acides aminés pour un poids moléculaire de 506kDa (Dunne et al., 1995). Trois autres gènes *FAT* existent : *FAT2*, *FAT3* et *FAT4* (Dunne et al., 1995; Hoeng et al., 2004; Katoh and Katoh, 2006; Mitsui et al., 2002; Nakayama et al., 2002). Les protéines *FAT* appartiennent à la famille des super-cadhérines en raison de leur grand nombre de répétitions cadhérines (Nollet et al., 2000). La protocadherine *FAT1* possède dans sa région extracellulaire 34 répétitions cadhérines associées à cinq motifs EGF (*Epidermal Growth Factor*) et un domaine laminine A-G localisé entre le 1<sup>er</sup> et le 2<sup>nd</sup> motif EGF (Dunne et al., 1995; Mahoney et al., 1991; Ponassi et al., 1999) (Figure 19). Ces motifs permettent des interactions cellule-cellule dépendantes du calcium et l'adhésion cellulaire (Davis, 1990; Dunne et al., 1995; Shapiro et al., 1995; Takeichi, 1995; Tepass et al., 2000). Une unique région transmembranaire est suivie d'un domaine cytoplasmique contenant des domaines PDZ et deux sites de fixation des  $\beta$ -

caténines qui permettent une liaison au cytosquelette (Cox et al., 2000; Dunne et al., 1995; Kemler, 1993; Ponassi et al., 1999; Provost and Rimm, 1999) (Figure 19).

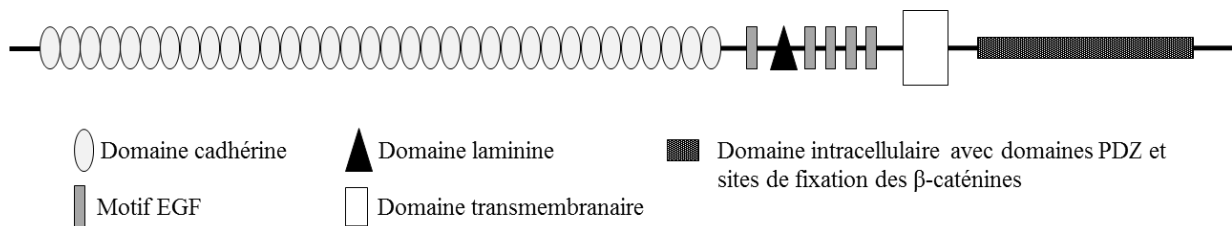


FIGURE 19. Structure de la protéine FAT1. La protocadherine FAT1 possède dans sa région extracellulaire, 34 répétitions cadhérines (rond gris clair) associées à cinq motifs EGF (*Epidermal Growth Factor*) (rectangle gris) et un domaine laminine A-G (triangle noir). Une unique région transmembranaire (rectangle blanc) est suivie d'un domaine cytoplasmique (rectangle gris foncé) contenant des domaines PDZ et deux sites de fixation des  $\beta$ -caténines, qui permettent une liaison au cytosquelette.

FAT1 est une protéine modulatrice de la polarité planaire. Elle joue un rôle dans la morphogénèse tissulaire de par son implication dans la mobilité cellulaire, dans la polarité cellulaire, dans la voie de signalisation Hippo (contrôle de la croissance des organes au cours du développement) ou encore dans le contact cellule-cellule (Bennett and Harvey, 2006; Cho et al., 2006; Harumoto et al., 2010; Hou and Sibinga, 2009; Moeller et al., 2004; Simons and Mlodzik, 2008; Skouloudaki et al., 2009; Sopko and McNeill, 2009; Tanoue and Takeichi, 2004). La majorité de ces fonctions implique une régulation de la dynamique de l'actine ou des microtubules (Tanoue and Takeichi, 2004). *FAT1* est exprimé au cours du développement dans les muscles et les tissus épithéliaux, dont la peau, le rein, les poumons ou encore le neuro-épithélium (Caruso et al., 2013; Cox et al., 2000; Ponassi et al., 1999; Smith et al., 2007). La protéine FAT1 est également exprimée dans les fibres musculaires différenciées, enrichie au niveau de l'interface entre les tubules-T et les unités contractiles, à proximité du récepteur dihydropyridine, un canal calcium voltage-dépendant (Caruso et al., 2013). Des études ont montré que le niveau d'expression de *FAT1* varie au cours du temps. Très élevée au début du stade fœtal, son expression diminue d'environ 90% à la fin du développement (Mariot et al., 2015).

Une délétion constitutive de *FAT1* chez la souris entraîne une mortalité néonatale, confirmant l'importance de cette protéine au cours du développement (Ciani et al., 2003). Plus précisément, chez des souris homozygotes *Fat1*<sup>-/-</sup>, la léthalité est probablement due à la présence de podocytes, cellules épithéliales au niveau du glomérule rénal, anormaux. Ce défaut de développement du rein peut être associé à une holoprosencéphalie (défaut de

développement du cerveau antérieur) et une microphthalmie-anophthalmie (défaut de développement des yeux) (Ciani et al., 2003). Un autre modèle murin existe : des souris homozygotes hypomorphiques *Fat1* avec une délétion d'une partie des répétitions cadhérines, du domaine lamine A-G et des motifs EFG. Ce modèle est un peu moins sévère et présente un taux de mortalité de l'ordre de 50% dans les trois mois suivant la naissance (Caruso et al., 2013). De manière intéressante, ces souris avec des anomalies développementales rénales, récapitulent le phénotype musculaire et non musculaire de la FSHD, y compris l'asymétrie et la rétinopathie. *FATI* a donc probablement un rôle dans le développement de la FSHD (Caruso et al., 2013). L'expression d'une protéine hypomorphique *Fat1* module la polarité migratoire des myoblastes au stade fœtus chez la souris. Aux stades postnataux, les souris présentent alors une réduction sélective de la masse musculaire, au niveau des muscles des épaules et de la face. Aux stades plus avancés, une dystrophie musculaire étendue est observée (Caruso et al., 2013).

Par ailleurs, des études montrent un niveau d'expression de *FATI* plus faible chez les patients FSHD comparés à des individus contrôles (Caruso et al., 2013; Mariot et al., 2015). En effet, le niveau de *FATI* est réduit de cinq fois dans le deltoïde de certains patients FSHD comparé à des individus contrôles (Caruso et al., 2013). A cela s'ajoute une variabilité intra-individuelle. *FATI* est exprimé, dans des fœtus contrôles, plus faiblement dans les muscles précocement atteints comme le deltoïde *versus* les muscles tardivement atteints comme le quadriceps (Mariot et al., 2015). Les faibles niveaux de *FATI* peuvent être corrélés au niveau de sévérité de la pathologie (Caruso et al., 2013; Mariot et al., 2015). L'ensemble de ces résultats suggèrent un rôle de *FATI* dans la FSHD. Cette hypothèse est appuyée par la découverte d'un polymorphisme délétère dans une potentielle séquence régulatrice intronique dans le locus *FATI* et qui ségrége avec la FSHD (Caruso et al., 2013). De plus, récemment, un groupe de patients FSHD ne présentant pas de contraction des répétitions D4Z4 mais possédant des protéines *FAT1* mutées a été décrit (Puppo et al., 2015).

Il est donc proposé que des individus avec un plus faible niveau de *FATI* comparé au niveau moyen quantifié dans la population générale, soient plus susceptibles de développer la FSHD (Mariot et al., 2015). Les muscles dans lesquels *FATI* est faiblement exprimé, pourraient être plus sensibles à une dérégulation des gènes. Cependant, plusieurs zones d'ombres demeurent. Notamment la raison pour laquelle un groupe musculaire précis est affecté par la perte de *FATI* n'est pas claire.

Par ailleurs, la souris hypomorphique *Fat1* présente des atteintes du rein, ce qui n'est pas observé chez les patients FSHD (Caruso et al., 2013). Deux hypothèses sont envisageables (Caruso et al., 2013; Reynolds et al., 2010) :

- 1) l'effet de l'absence de *FAT1* au niveau du rein chez les patients FSHD est contrecarré par une autre protéine ;
- 2) les patients avec une perte sévère de la fonction de *FAT1* et des dysfonctionnements rénaux décèdent probablement avant l'apparition des symptômes musculaires, et ne sont ainsi pas classifiés FSHD.

La FSHD ne peut pas être uniquement considérée comme une pathologie due à une mutation dans *FAT1*, mais *FAT1* semble tout de même jouer un rôle dans le développement de cette pathologie.

### **1.3.6.5. *DUX4***

Les unités D4Z4 jouent un rôle fondamental dans le développement de la pathologie puisque pour développer la FSHD, il faut posséder au moins une répétition D4Z4 (Tupler et al., 1996b).

#### **1.3.6.1.1. Localisation de l'ORF de *DUX4* dans le motif D4Z4**

Chaque motif D4Z4, dont les extrémités sont délimitées par les sites de restriction *KpnI*, contient plusieurs motifs génétiques (Figure 20). Les répétitions D4Z4, identiques entre elles, sont caractérisées par la présence d'une séquence *LSau* (*LongSau3A* unités répétées sur 400 à 500bp) riche en GC et associée à de l'hétérochromatine (position 1 à 340), d'une région A (position 869 à 1071), d'une séquence *hhspm3* (*hypomethylated sperm-specific sequence 3*) (position 1313 à 1780bp) et de deux séquences homéoboites localisées au sein d'un ORF (position 1792 à 3063) (Gabriëls et al., 1999; Hewitt et al., 1994; Lee et al., 1995; Meneveri et al., 1993; Ottaviani et al., 2009a; Zhang et al., 1987) (Figure 20). Un promoteur lié à cet ORF a été détecté. Il est composé des éléments régulateurs que sont une boîte CATT, une boîte GC (à 179bp en 5' de l'ORF), une boîte TACAA (au lieu de TATAA) (à 149bp en 5' de l'ORF) et une boîte E contenant une séquence initiateur GCACCTG (Dixit et al., 2007; Gabriëls et

al., 1999) (Figure 21). La région pLAM contient également une séquence *LSau*, une séquence *hhspm3* et un petit ORF avec une homéoboîte (Gabriëls et al., 1999; Hewitt et al., 1994).

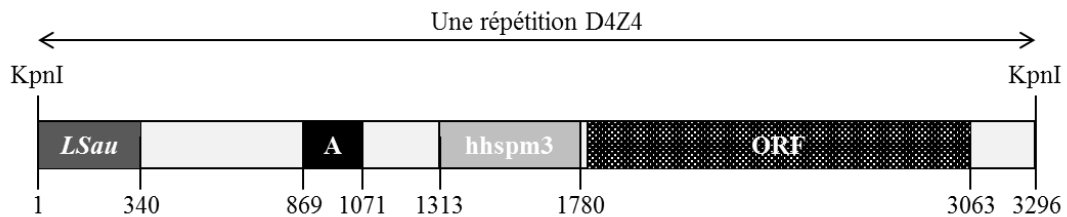


FIGURE 20. Structure génétique d'une répétition D4Z4. Le motif D4Z4 de 3,3kb, délimité par les sites de restriction *KpnI*, se compose d'un motif *LSau*, d'une région A, d'une séquence *hhspm3* et d'un ORF contenant deux séquences homéoboîtes. La position de chacune de ces régions est indiquée. La position +1 correspond au début de la répétition D4Z4 (adapté de Jiang et al., 2003).

L'ORF, présent dans chaque unité D4Z4, code une protéine à double homéodomaine appelée DUX4 (*Double homeobox chromosome 4*) (Gabriëls et al., 1999; Hewitt et al., 1994; Holland et al., 2007; Lee et al., 1995; Lyle et al., 1995; Winokur et al., 1997). Le gène *DUX4*, apparenté aux rétrogènes, serait issu de la rétrotransposition de l'ARNm *DUXC* (Clapp et al., 2007).

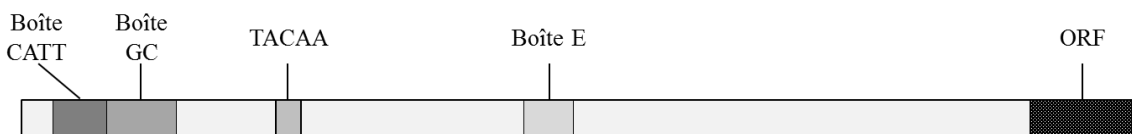


FIGURE 21. Structure génétique du promoteur lié à l'ORF présent dans chaque répétition D4Z4. Ce promoteur est composé d'éléments régulateurs que sont la boîte CATT, la boîte GC, la boîte TACAA et la boîte E.

### 1.3.6.1.2. Différents variants de l'ARNm *DUX4*

L'ARNm *DUX4* se compose de trois exons dans les cellules somatiques (Dixit et al., 2007; Snider et al., 2009). Chaque répétition D4Z4 contient les séquences des deux premiers exons : l'exon 1 (1466bp) dans lequel est retrouvé l'ORF *DUX4* et l'exon 2 (92bp) (Figure 22). L'ORF *DUX4* démarre soit au codon start MKG (Met-Lys-Gly), soit au codon start initialement identifié MAL (Met-Ala-Leu) (Dixit et al., 2007; Snider et al., 2009, 2009) (Figure 22). L'exon 3 (198bp) est localisé, quant à lui, en aval de la dernière répétition D4Z4, dans la région pLAM, sur le chromosome permissif 4qA uniquement (Dixit et al., 2007; Lemmers et al., 2004a, 2002, 2010a; Rossi et al., 2007; Snider et al., 2009; Thomas et al., 2007) (Figure 22).



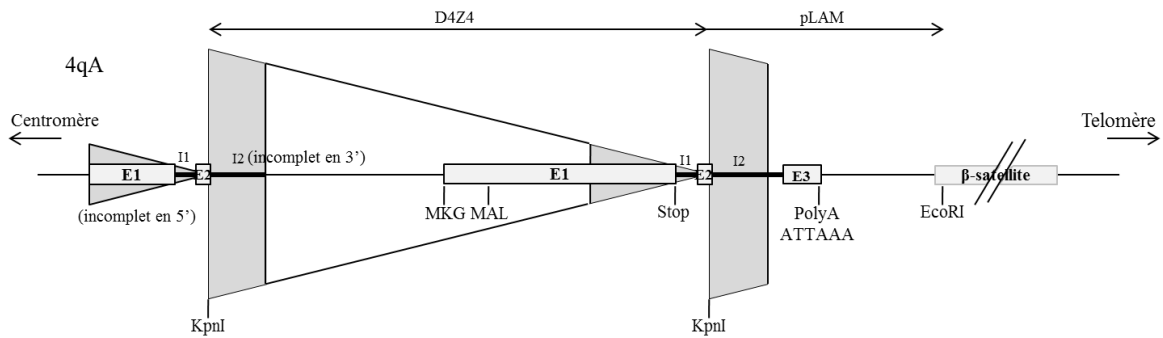


FIGURE 22. Structure du gène *DUX4* et localisation des différents exons de *DUX4* au sein des répétitions D4Z4. Chaque répétition D4Z4 contient l'exon 1 et 2 (E1 et E2 respectivement) du gène codant pour le facteur de transcription DUX4. L'ORF *DUX4* n'est cependant présent que dans l'exon 1. L'exon 3 est localisé en aval de la dernière répétition D4Z4, dans la région pLAM, et contient le signal de polyadénylation ATTAAA pour les variants 4qA.

Ce dernier exon, l'exon 3, joue un rôle fondamental car il contient un signal de polyadénylation consensus ATTAAA, qui permet l'ajout d'une queue polyadénylée (poly(A)), et donc une stabilisation de l'ARNm (Dixit et al., 2007; Sachs, 1990; Snider et al., 2009). Par conséquent, et malgré le fait que chaque répétition D4Z4 possède le promoteur et l'ORF de *DUX4*, il est supposé que l'ARNm *DUX4* transcrit à partir des répétitions D4Z4 internes est instable car celui-ci n'est pas polyadénylé. A l'inverse, l'ARNm *DUX4* transcrit à partir de la dernière répétition D4Z4 directement suivie de l'exon 3 contenant le signal de polyadénylation, est stable. Il est ensuite traduit en protéine DUX4 (Dixit et al., 2007; Lemmers et al., 2010a) (Figure 23).

L'importance de ce signal de polyadénylation a été démontrée grâce à l'étude du chromosome 10qA. En effet, le chromosome 10qA contient également cette séquence pLAM en aval de la dernière répétition D4Z4 (Lemmers et al., 2010b). Pour autant, une contraction des répétitions D4Z4 sur le chromosome 10qA ne conduit pas à la pathologie (Bakker et al., 1995; Cacurri et al., 1998; Lemmers et al., 1998). Ceci est dû à un polymorphisme dans la séquence du signal de polyadénylation (ATTAAA devient ATCAAA ou ATTTAA) du chromosome 10qA, conduisant à la perte de fonctionnalité de celui-ci (Beaudoing et al., 2000; Lemmers et al., 2010a; Tian et al., 2005) (Figure 23). Ainsi, aucun ARNm *DUX4* stable ne peut être transcrit à partir des répétitions internes ou de la dernière répétition D4Z4 du chromosome 10qA (Lemmers et al., 2010a) (Figure 23). L'échange entre le chromosome 4qA et le chromosome 10qA, d'un fragment comprenant la dernière répétition D4Z4 et se poursuivant au-delà de la séquence pLAM, conduit au développement de la FSHD chez des individus porteurs d'une contraction sur le chromosome 10q, normalement non pathologique, alors qu'ils ne présentent

aucune contraction des répétitions D4Z4 sur le chromosome 4qA (Lemmers et al., 2010a). Dans ce cas, la translocation du chromosome 4qA vers le chromosome 10qA apporte un signal de polyadénylation fonctionnel permettant de stabiliser les transcrits *DUX4* issus du chromosome 10qA.

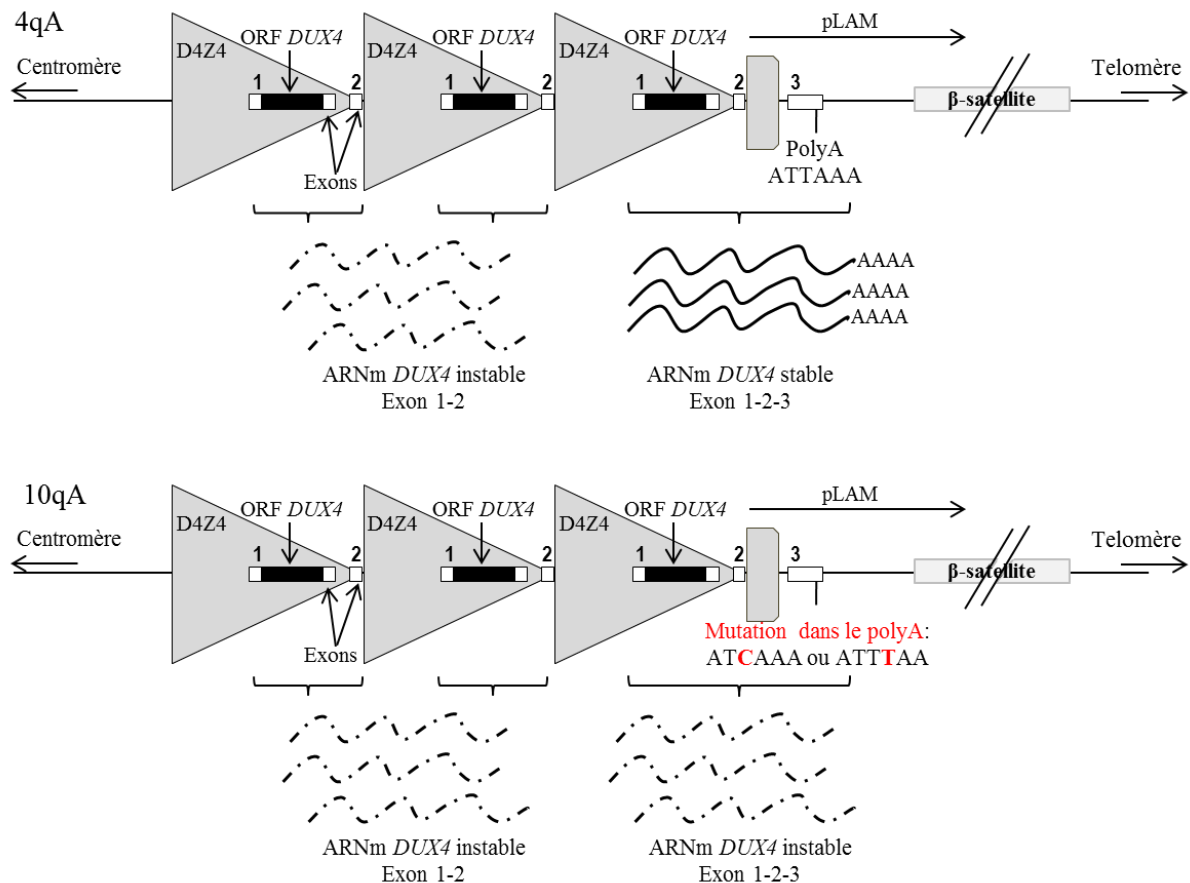


FIGURE 23. Modèle de la stabilité des ARNm *DUX4* transcrits à partir des répétitions D4Z4 internes ou de la dernière répétition D4Z4 sur le chromosome 4qA et 10qA. Bien que chaque répétition D4Z4 possède le promoteur et l'ORF de *DUX4*, l'ARNm transcrit à partir des répétitions D4Z4 internes est instable car celui-ci n'est pas polyadénylé. A l'inverse, l'ARNm *DUX4* transcrit à partir de la dernière répétition D4Z4, qui est directement suivie de l'exon 3 contenant un signal de polyadénylation, est stable. Sur le chromosome 10qA, le signal de polyadénylation est muté. Aucun ARNm *DUX4* stable n'est transcrit à partir du chromosome 10qA.

Différents variants de l'ARNm *DUX4* ont été décrits, classés en deux catégories (Dixit et al., 2007; Snider et al., 2009, 2010). Deux variants longs, les ARNm *DUX4-fl*, qui donnent tous deux naissance à une protéine *DUX4* entière et fonctionnelle (Figure 24). Ces deux variants, composés des exons 1-2-3, diffèrent par l'épissage ou non de l'intron 1 (136bp) en 3'UTR (*Untranslated region*) de l'ARNm (Figure 24). L'intron 2 (362bp) est quant à lui toujours épissé (Figure 24). Le rôle de ces deux variants qui codent pour une même protéine n'est pas établi à ce jour. Un autre variant court, l'ARNm *DUX4-s*, donne naissance à une protéine

tronquée, DUX4-s. Ce variant se caractérise par l'utilisation d'un site d'épissage cryptique situé dans l'exon 1 donnant un ARNm avec les exons 1s-2-3 (1s signifiant ici que l'exon 1 est tronqué) (Dixit et al., 2007; Snider et al., 2009, 2010) (Figure 24). La protéine DUX4-s possède toujours son domaine N-terminal de fixation à l'ADN mais ne présente plus de domaine C-terminal, domaine de transactivation (Bosnakovski et al., 2008; Kawamura-Saito et al., 2006; Snider et al., 2010). La protéine DUX4-s n'est pas toxique pour les cellules (Geng et al., 2011).

Dans la lignée germinale, deux autres variants de l'ARNm *DUX4* sont exprimés (Geng et al., 2012; Snider et al., 2010). Ces variants font intervenir des exons supplémentaires notés 4-5-6-7 et localisés à plus de 6000bp en aval de l'exon 3 sur les chromosomes 4qA et 10qA. Les deux ARNm *DUX4* transcrits utilisant ces exons sont : un ARNm composé de l'exon 1-2-6-7 et un ARNm composé de l'exon 1-2-4-5-7 (Figure 24). L'exon 3 est toujours exclu. Le signal de polyadénylation utilisé lors de la transcription de ces deux ARNm est localisé dans l'exon 7. Ainsi un ARNm *DUX4* long et stable est transcrit à partir du chromosome 4qA ou 10qA en quantité équivalente. Le signal de polyadénylation situé dans l'exon 7 ne semble pas être utilisé dans les cellules somatiques (Snider et al., 2010).

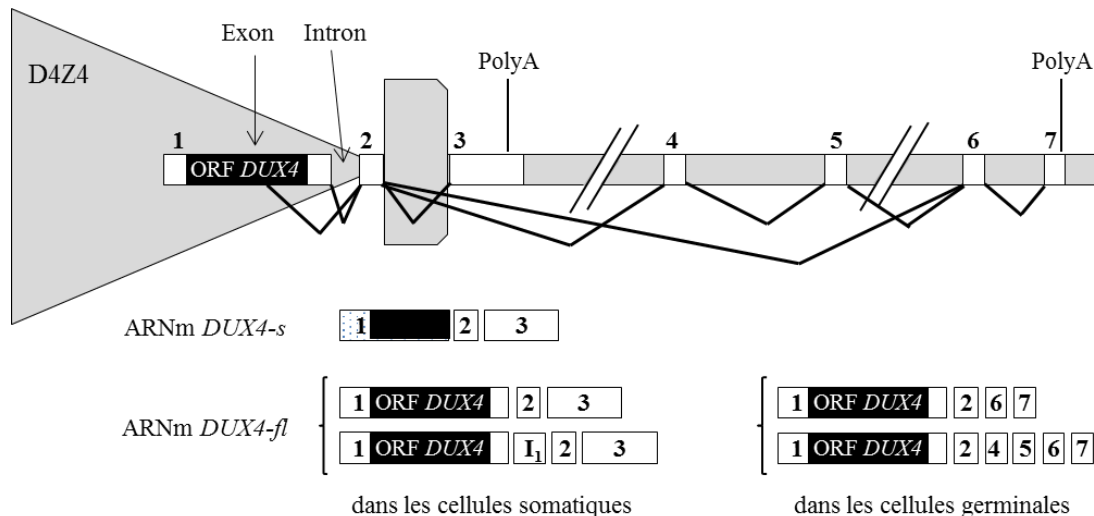


FIGURE 24. Récapitulatif des différents variants de l'ARNm *DUX4* transcrits dans les cellules somatiques ou les cellules germinales. Dans les cellules somatiques, trois ARNm *DUX4* peuvent être générés à partir de la dernière unité D4Z4. Tous les ARNm contiennent l'exon 2 et l'exon 3. L'ARNm *DUX4-s* résulte d'un épissage cryptique démarrant dans l'exon 1. Il est exprimé dans les myoblastes contrôles et dans les tissus somatiques. Les deux ARNm *DUX4-fl* contiennent l'exon 1 en entier. Ces deux ARNm *DUX4-fl* diffèrent par la présence ou non de l'intron 1. Cependant, ces deux variants sont tous deux associés à la FSHD et sont exprimés dans les myotubes d'haplotype 4qA. Dans les cellules germinales, deux autres ARNm *DUX4* sont transcrits. Ces ARNm sont composés de l'exon 1-2-6-7 ou 1-2-4-5-6-7. L'exon 3 est toujours exclu. Le signal de polyadénylation utilisé ici est localisé dans l'exon 7.

### 1.3.6.1.3. Profil d'expression des différents transcrits *DUX4*

Les expressions des différents transcrits *DUX4*, *DUX4-fl* et *DUX4-s*, ont été étudiées dans des biopsies d'adultes, des cultures de cellules musculaires issues de patients FSHD (FSHD1 et FSHD2) ou d'individus contrôles porteurs du chromosome permissif (Dixit et al., 2007; Jones et al., 2012; Snider et al., 2010; Tsumagari et al., 2011). Une expression de *DUX4-fl* est détectée dans toutes les cultures de cellules musculaires de patients FSHD, et dans 50 à 80% des biopsies de patients FSHD. De manière surprenante, et selon les mêmes études, *DUX4-fl* est également exprimé sporadiquement dans 0 à 40% des cultures de cellules musculaires d'individus contrôles, et dans 0 à 10% des biopsies d'individus contrôles. Le transcrit *DUX4-s* est quant à lui détecté dans 100% des biopsies ou cultures de cellules musculaires d'individus contrôles et à hauteur de 30 à 40% dans des biopsies ou cultures de cellules musculaires de patients FSHD (Jones et al., 2012; Snider et al., 2010; Tsumagari et al., 2011). D'autres études, réalisées sur des biopsies fœtales FSHD et contrôles, ont permis de mettre en évidence une expression de *DUX4-fl* dans les cellules fœtales FSHD et contrôles (Broucqsault et al., 2013; Ferreboeuf et al., 2014a).

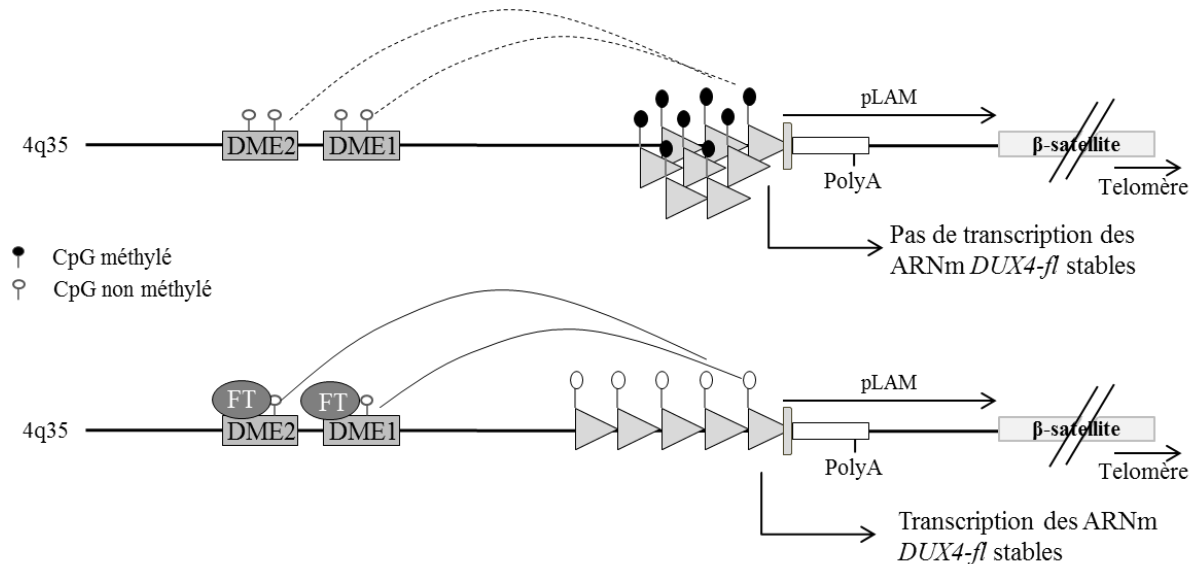


FIGURE 25. Modèle de régulation de l'expression de *DUX4* par les *enhancers* myogénique DME1 et DME2. Dans la majorité des myocytes, les régions DME1 et DME2 sont dans un état dit permissif (i.e. hypométhylées mais associées à de nombreux nucléosomes). Ces régions ne permettent pas la transcription de *DUX4*. Dans certaines cellules, et en réponse à un signal, des facteurs de transcription (FT) se fixent à ces *enhancers*, les activant. DME1 et DME2 induisent alors la transcription de *DUX4* (adapté de Himeda et al., 2014).

Au sein de la cellule musculaire, l'expression de *DUX4-fl* est induite durant la différenciation cellulaire, supposant une régulation par des *enhancers* myogéniques (Ferreboeuf et al., 2014a; Himeda et al., 2014). Deux régions en amont des répétitions D4Z4 pouvant jouer ce rôle, DME1 (*DUX4 myogenic enhancer 1*) et DME2 (*DUX4 myogenic enhancer 2*), ont été identifiées (Figure 25).

Les régions DME1 et DME2 sont actives seulement dans les cellules musculaires en différenciation. Ces régions contiennent de nombreux éléments caractéristiques des *enhancers* ainsi que des sites reconnus par les facteurs de transcription comme MyoD, myogénine ou encore PAX (*Paired box*), impliqués dans le développement précoce et la spécification-différenciation du muscle strié squelettique. Dans la majorité des myocytes, ces régions DME1 et DME2 sont dans un état dit permissif : hypométhylées, elles sont associées à de nombreux nucléosomes et ne peuvent interagir avec les promoteurs des gènes (Figure 25). Dans de rares cellules, en réponse à un signal encore non déterminé, des facteurs de transcription se fixent à ces *enhancers* permissifs qui deviennent actifs (Figure 25). Par un changement de conformation de la chromatine, ces *enhancers* interagissent alors avec le promoteur de *DUX4* hypométhylé et induisent l'expression de *DUX4-fl* (Himeda et al., 2014) (Figure 25).

Malgré tout, l'expression de *DUX4-fl* n'est pas limitée au tissu musculaire et peut concerner également, à faible fréquence, des tissus non affectés (Broucqsault et al., 2013; Das and Chadwick, 2016; Geng et al., 2012; Snider et al., 2010).

Les fibroblastes contrôles expriment *DUX4-s*. Lorsque ces cellules contrôles sont reprogrammées en iPS (*induced Pluripotent Stem cells*), elles n'expriment plus *DUX4-s* mais *DUX4-fl*. Après différenciation en corps embryoides, ces cellules expriment de nouveau *DUX4-s*. Les fibroblastes FSHD, reprogrammés en iPS, puis différenciés en corps embryoides, expriment, quant à eux, uniquement *DUX4-fl* (Snider et al., 2010). Cette variation observée dans l'expression des variants de l'ARNm *DUX4* en fonction de l'état de la cellule est corrélée avec une modification des marques épigénétiques répressives à la chromatine (de type H3K9me3) au niveau des répétitions D4Z4 (Snider et al., 2010). Or, il a été montré que l'épissage alternatif peut-être régulé par des modifications des marques épigénétiques (Luco et al., 2010). Ainsi dans les fibroblastes contrôles ou les cellules contrôles différenciées en corps embryoides, une modification des marques épigénétiques

répressives post-traductionnelles au niveau de la queue des histones pourrait favoriser un épissage vers la forme *DUX4-s* (Snider et al., 2010).

*DUX4-fl* est donc exprimé très tôt, dès le développement fœtal, majoritairement dans les tissus musculaires. En conséquence, les événements moléculaires conduisant à un phénotype FSHD sembleraient se mettre en place durant le développement fœtal. Cependant, la FSHD est décrit comme une myopathie de l'adulte avec l'apparition de symptômes au cours de l'adolescence tardive. Plusieurs hypothèses peuvent être proposées pour expliquer ce décalage temporel (Ferreboeuf et al., 2014a) :

- la fibre musculaire ne pourrait tolérer qu'une certaine quantité de DUX4 uniquement pendant une durée de temps limitée ;

- DUX4 aurait un rôle positif sur le développement du muscle squelettique au stade fœtal. Par la suite, son expression doit être réduite *via* des mécanismes épigénétiques, ce sans quoi l'individu développerait un phénotype FSHD ;

- la présence ou l'absence d'un gène qui facilite ou empêche l'action néfaste de DUX4, rôle qui pourrait être joué par *FAT1*.

#### **1.3.6.1.4. DUX4 : un facteur de transcription**

DUX4 est un facteur de transcription de 424 acides aminés pour un poids moléculaire de 52kDa (Gabriëls et al., 1999; Geng et al., 2011). Le domaine C-terminal de la protéine est un activateur transcriptionnel. Le domaine N-terminal de la protéine contient quant à lui deux homéodomains similaires entre eux. Ces homéodomains sont également homologues à ceux des protéines à un seul homéodomaine PAX3 (*Protein paired box 3*) et PAX7 (*Protein paired box 7*), impliquées dans le développement du muscle squelettique et dans le maintien des cellules satellites (Bosnakovski et al., 2008; Buckingham, 2006; Holland et al., 2007; Kawamura-Saito et al., 2006). Les homéodomains (environ 60 acides aminés) sont codés par des séquences nucléotidiques de 180bp appelées homéoboîtes. Les homéodomains contiennent un NLS et interagissent avec l'ADN par l'intermédiaire de trois hélices alpha qui forment un motif en hélice-tour-hélice (Ostlund et al., 2005; Scott et al., 1989). Les homéodomains sont caractéristiques d'un certain nombre de facteurs de transcription.

En tant que facteur de transcription, DUX4 reconnaît et se fixe à la séquence consensus TAAYBBAATCA (nomenclature IUPAC ; Y = C ou T, B = C ou G ou T) (Geng et al., 2012; Zhang et al., 2016). Cette séquence, détectée au niveau de séquences non répétées mais aussi des séquences associées aux MaLR (*Mammalien apparent Retro-transposons*), contient le motif consensus de fixation des homéodomaines (TAAT), en deux copies séparées par un nucléotide. L'affinité et l'activité de DUX4 sont les plus élevées pour une fixation à la séquence TAATCTAATCA. Par ailleurs, les gènes cibles de DUX4 possédant un seul motif reconnu par DUX4 dans leur promoteur, présentent une activation transcriptionnelle relativement faible comparée à celle des gènes cibles qui possèdent au moins deux motifs reconnus par DUX4 (Zhang et al., 2016).

Les protéines avec un homéodomaine de la famille *Paired box*, dont font partie PAX3/7, se lient généralement à l'ADN sous forme de dimères (Wilson et al., 1993). Chaque homéodomaine interagit avec la séquence TAAT selon une configuration où le dimère est en position symétrique par rapport à l'axe de l'ADN. Par conséquent, le deuxième homéodomaine du dimère se fixe à la séquence complémentaire ATTA. Les protéines avec un homéodomaine de la famille *Paired box* reconnaissent donc les séquences palindromiques TAAT-(N)<sub>x</sub>-ATTA. Pour PAX7, le nombre de nucléotides entre les deux séquences TAAT et ATTA est de deux, alors qu'il est de trois pour PAX3 (Czerny and Busslinger, 1995; Soleimani et al., 2012). Les homéodomaines de DUX4 sont physiquement très proches (séparés de 15-20 acides aminés) impliquant probablement une fixation de la protéine sur le même brin d'ADN, d'où la reconnaissance de la séquence non palindromique TAAT-(N)<sub>x</sub>-TAAT avec ici X=1 (Anseau et al., 2016; Ding et al., 1998; Zhang et al., 2016). L'étude de la structure du complexe DUX4-ADN permettra de confirmer ou d'infirmer cette hypothèse.

Par ailleurs, du fait de la forte similitude entre les homéodomaines de DUX4 et ceux de PAX3/7, ces facteurs de transcription peuvent entrer en compétition lors de la modulation de l'expression de gènes cibles communs. Ainsi, DUX4 pourrait venir interférer avec PAX7 dans la fonction de maintien des cellules satellites ou encore altérer l'expression de certains gènes cibles de PAX3/7, dont MyoD, et inversement (Bosnakovski et al., 2008; Relaix et al., 2006; Tajbakhsh et al., 1997).

### 1.3.6.1.5. Dérégulations génétiques liées à l'expression de DUX4

Lors de la surexpression de DUX4-fl (ou DUX4-s) dans des myoblastes contrôles, environ 710 gènes sont dérégulés (38 pour DUX4-s) (Geng et al., 2012). Plus précisément, 466 gènes voient leur expression augmentée tandis que 244 gènes voient leur expression réduite (37 et 1 respectivement pour DUX4-s) (Geng et al., 2012). D'autres études sont venues compléter celle-ci (Rickard et al., 2015; Sharma et al., 2013; Yao et al., 2014). Au final, seul un peu plus de 90 de ces gènes ont été démontrés comme dérégulés dans des biopsies ou des cellules dérivant de patients FSHD (Bosnakovski et al., 2008; Geng et al., 2012; Rickard et al., 2015; Sharma et al., 2013; Tsumagari et al., 2011; Yao et al., 2014). Parmi ces gènes, certains sont impliqués dans le développement de la lignée germinale et des cellules souches, d'autres jouent un rôle dans la modification des ARNs comme *PABPN1* (*Poly(A) Binding Protein Nuclear 1*) qui est une protéine de liaison à la queue poly(A). Certains des gènes inhibés appartiennent à la voie de l'ubiquitine ou sont engagés dans l'immunité et la défense innée comme *DEFB103* (*Defensin beta 103*) qui inhibe la réponse immunitaire innée et la différenciation musculaire. D'autres gènes négativement régulés interviennent, quant à eux, dans le contrôle du cycle cellulaire, ou encore dans la réponse au stress et la transduction des signaux (Bosnakovski et al., 2008; Geng et al., 2012). Les gènes surexprimés en présence de DUX4 et couramment utilisés comme marqueurs de la FSHD sont :

- *ZSCAN4* (*Zinc Finger and SCAN domain containing 4*) qui a un rôle dans la stabilité du génome et le contrôle de la longueur des télomères ;
- *MBD3L2* (*Methyl-CpG Binding protein-like*) qui code une protéine se liant aux îlots CpG méthylés ;
- *TRIM43* (*Tripartite Motif-containing 43*) qui est impliqué dans l'implantation de l'embryon.

Un autre gène dérégulé dans les biopsies de patients FSHD est *PITX1* (*Paired like homeodomain 1*), un membre des facteurs de transcription à homéodomaine (Dixit et al., 2007; Geng et al., 2012; Tassin et al., 2013; Wallace et al., 2011a). *Pitx1* est un gène impliqué dans l'identité des membres postérieurs et dans la mise en place de la symétrie droite – gauche chez la souris (Cole et al., 2003; Lanctôt et al., 1999; Logan and Tabin, 1999; Marcil et al., 2003; Pandey et al., 2012; Shapiro et al., 2004; Szeto et al., 1999; Tanaka et al., 2005). Des souris surexprimant *Pitx1* présentent une réduction de leur masse musculaire due à une atrophie musculaire (Pandey et al., 2012). En plus de son rôle dans le développement



musculaire, PITX1 induit également l'apoptose par une voie dépendante du facteur p53 dont il active la transcription (Liu and Lobie, 2007; Yamaguchi et al., 2010). *PITX1* pourrait donc jouer un rôle dans le développement de la FSHD et être impliqué dans l'atteinte asymétrique des muscles. Des expériences de gènes rapporteurs fusionnés au promoteur *Pitx1* de la souris, montrent une activation transitoire du gène rapporteur en présence de DUX4 (Dixit et al., 2007). Chez la souris, la séquence qui pourrait être reconnue par DUX4 dans le promoteur de *Pitx1* est CGGATGCTGTCTTCTAATTAGTTTGGACCC (Dixit et al., 2007). Cette séquence ne contient pas le motif en tandem TAAT mais possède les deux séquences palindromiques TAAT et ATTA en position chevauchante (*i.e.* TAATTA) qui permettrait une reconnaissance du promoteur *Pitx1* par DUX4 (Dixit et al., 2007; Wallace et al., 2011a). Ainsi la dérégulation de *PITX1* observée dans les biopsies de patients FSHD pourrait être DUX4 dépendante. Cependant, aucune interaction directe entre DUX4 et le promoteur de *Pitx1* n'a été confirmée (Zhang et al., 2016). De plus, il est difficile de généraliser à l'humain le mécanisme de dérégulation de *Pitx1* qui est observé chez la souris. En effet, la séquence promotrice (et notamment la séquence potentiellement reconnue par DUX4 chez la souris) est très peu conservée entre les deux espèces (Zhang et al., 2016). Un éventuel rôle de *PITX1* dans la FSHD n'est donc pas clairement établi à ce jour.

Au contraire de DUX4-fl, la protéine DUX4-s n'a pas ou peu d'impact sur l'expression des gènes dans le myoblastes : DUX4-s n'est pas toxique pour la cellule. Par ailleurs, la co-expression de DUX4-fl et DUX4-s permet de supprimer l'activation des gènes cibles de DUX4-fl (Geng et al., 2012; Mitsuhashi et al., 2013a). En effet, DUX4-s partage avec DUX4-fl les acides aminés 1 à 159 en région N-terminale, incluant les deux homéodomains, mais est dépourvu du domaine C-terminal (acides aminés 160 à 424). Or la fixation de DUX4 au promoteur de ses gènes cibles s'effectue par le domaine N-terminal alors que l'activation transcriptionnelle est quant à elle permise par le domaine C-terminal (Kawamura-Saito et al., 2006). Ainsi, DUX4-s peut entrer en compétition avec DUX4-fl en se fixant aux mêmes séquences promotrices, sans pour autant pouvoir activer la transcription des gènes. DUX4-s agit à la manière d'un dominant négatif (Geng et al., 2012; Mitsuhashi et al., 2013a).

En résumé, DUX4 est un facteur de transcription qui dérégule l'expression d'un grand nombre de gènes, dont certains codent eux-mêmes pour d'autres facteurs de transcription.

### 1.3.6.1.6. Conséquences cellulaires de l'expression de *DUX4*

*DUX4*, exprimé dans les muscles affectés ou les lignées cellulaires de patients FSHD, semble donc jouer un rôle dans le développement de la FSHD en modifiant l'expression de centaines de gènes. Cependant, à ce jour, aucun lien précis n'est établi entre son expression ou son activité de facteur de transcription et le phénotype FSHD observé.

Une surexpression de *DUX4-fl* semble être toxique pour la cellule (Bosnakovski et al., 2008; Kowaljow et al., 2007; Mitsuhashi et al., 2013a; Vanderplanck et al., 2011; Wuebbles et al., 2010). En effet, dans des cellules surexprimant *DUX4-fl*, et parfois dans les muscles de patients FSHD, de nombreux marqueurs spécifiques de l'apoptose sont observés, incluant une activation des caspases 3/6/7/8/9, une fragmentation de la chromatine, et un marquage à l'annexine V positif (Bosnakovski et al., 2008; Kowaljow et al., 2007; Sandri et al., 2001, 2001; Winokur et al., 2003a; Wuebbles et al., 2010). A cela s'ajoute une désorganisation de l'enveloppe nucléaire marquée par une redistribution de l'émérine (Kowaljow et al., 2007). Le cycle cellulaire est également perturbé, avec une augmentation de p21, un inhibiteur des complexes cycline/kinase dépendante des cyclines, et une diminution de la cycline E (Bosnakovski et al., 2008; Winokur et al., 2003b). Cependant, ces dérégulations moléculaires sont peu probables *in vivo* au cours des premiers stades de la pathologie, puisque *DUX4* est faiblement exprimé dans des muscles atteints de patients FSHD (Dixit et al., 2007; Winokur et al., 2003b). Or, même si une faible expression de *DUX4* n'induit pas une mort cellulaire dans un premier temps, cette expression peut néanmoins réduire la prolifération cellulaire et modifier la morphologie des cellules. En effet, il a notamment été observé une réduction de la taille des cellules musculaires dans la FSHD : on parle d'atrophie musculaire (Barro et al., 2010; Bosnakovski et al., 2008; Vanderplanck et al., 2011). Le processus d'atrophie musculaire peut-être médié par plusieurs voies cellulaires. L'une d'elle consiste en l'activation de la transcription de deux gènes qui codent pour des E3 ubiquitines ligases spécifiques du muscle : *Atrogin1/MAFbx* (*Muscle atrophy F-box*) et *MurF1* (*Muscle-specific RING Finger 1*) (Bodine and Baehr, 2014; Bodine et al., 2001; Foletta et al., 2011; Gomes et al., 2001; Lecker et al., 1999). Ces ubiquitines ligases reconnaissent et permettent l'ubiquitinylation de protéines myofibrillaires, dont la myogénine ou MyoD qui sont impliquées dans la différenciation des myoblastes et la formation du myotube (Lagrand-Cantaloube et al., 2009; Tintignac et al., 2005). L'ubiquitinylation des protéines myofibrillaires aboutit à leur dégradation par le protéasome. Il en résulte alors l'apparition

d'un phénotype atrophique, notamment par inhibition de la fusion et de la différenciation en myocytes (Barro et al., 2010; Bodine and Baehr, 2014; Bosnakovski et al., 2008; Lagirand-Cantaloube et al., 2009; Tintignac et al., 2005). Or, une surexpression de *DUX4* induit une activation de l'expression de *MurF1* et de *Atrogin1/MAFbx*, ce qui permet de faire un lien entre l'expression de *DUX4* et l'atrophie musculaire observée (Vanderplanck et al., 2011). Par ailleurs, une surexpression de *DUX4* est associée à une activation de p53 (ou TP53, *Tumor protein 53*) (Fox et al., 2014; Vanderplanck et al., 2011; Wallace et al., 2011a). Le suppresseur de tumeur p53 a différentes fonctions. Chez la souris, son expression stable et chronique induit notamment une atrophie musculaire (Tyner et al., 2002). Le suppresseur de tumeur p53 est également un facteur pro-apoptotique (Ozaki and Nakagawara, 2011). L'apoptose détectée lors d'une expression de *DUX4* pourrait donc être dépendante de p53. Cependant, aucun lien direct entre l'activation de p53 et *DUX4* n'a pour le moment été établi (Wallace et al., 2011a). Un mécanisme potentiel impliquerait une activation de p53 par PITX1, qui serait lui-même activé (directement ou indirectement) par *DUX4* (Dixit et al., 2007; Liu and Lobie, 2007; Pandey et al., 2012; Vanderplanck et al., 2011; Wallace et al., 2011a).

Une expression de *DUX4*, même faible, est donc associée à une diminution de l'expression de certains facteurs myogéniques comme *MyoD*, *Pax7*, la *myogénine* ou encore la *desmine* (Bosnakovski et al., 2008; Celegato et al., 2006; Mitsushashi et al., 2013a; Vanderplanck et al., 2011; Winokur et al., 2003b). Concernant *MyoD*, il ne s'agit pas uniquement d'un facteur de transcription impliqué dans la formation et la différenciation de la fibre musculaire. Ce facteur de transcription régule également l'expression de certains gènes jouant un rôle dans la réponse au stress, notamment oxydant (Guttridge et al., 2000). Ainsi une diminution de *MyoD* médiée par *DUX4*, induit également une diminution du niveau d'expression des protéines de réponse au stress dont des enzymes impliquées dans le métabolisme redox du glutathion (Bosnakovski et al., 2008). De ce fait, les cellules exprimant *DUX4* sont plus sensibles à un stress oxydant, ce qui pourrait également expliquer la mort cellulaire observée (Winokur et al., 2003b).

L'ensemble des processus dérégulés par *DUX4*, *i.e.* la myogenèse et le stress oxydant, sont des processus biologiques affectés chez les patients FSHD (Barro et al., 2010; Celegato et al., 2006; Dellavalle et al., 2007; Laoudj-Chenivresse et al., 2005; Macaione et al., 2007; Winokur et al., 2003b). Cependant, les mécanismes moléculaires précis qui font le lien entre l'expression de *DUX4* et le phénotype FSHD ne sont pas totalement élucidés à ce jour. De

plus, tous les types cellulaires ne présentent pas la même sensibilité à différents niveaux d'expression de *DUX4* (Bosnakovski et al., 2008). Il est tout de même important de souligner que d'autres études contredisent les résultats précédents. En effet, des travaux ont montré, dans un contexte de surexpression de *DUX4*, une augmentation de l'expression de *MyoD* et aucune altération de la différenciation ou de la fusion des myoblastes (Broucqsault et al., 2013; Krom et al., 2012; Sharma et al., 2013; Tsumagari et al., 2011).

Malgré tout, et au vu de l'ensemble de ces données, *DUX4* semble être un gène candidat légitime dans le développement de la FSHD, son expression induisant potentiellement des changements moléculaires majeurs incompatibles avec le maintien d'une intégrité musculaire normale.

### 1.3.6.1.7. Faible expression de *DUX4* dans les cellules FSHD

La protéine *DUX4* est très faiblement exprimée chez les patients FSHD : la protéine est détectée dans 0,5 à 9% des noyaux, tandis que l'ARNm est seulement exprimé dans 1 noyau sur 1000 (Block et al., 2013; Ferreboeuf et al., 2014a; Snider et al., 2010; Tassin et al., 2013). Peu de laboratoires arrivent à détecter *DUX4* par PCR (*Polymerase Chain Reaction*). Or, l'expression de *DUX4*, même faible, induit une dérégulation quantifiable d'un certain nombre de gènes cibles (Bosnakovski et al., 2008; Geng et al., 2012; Rickard et al., 2015; Yao et al., 2014).

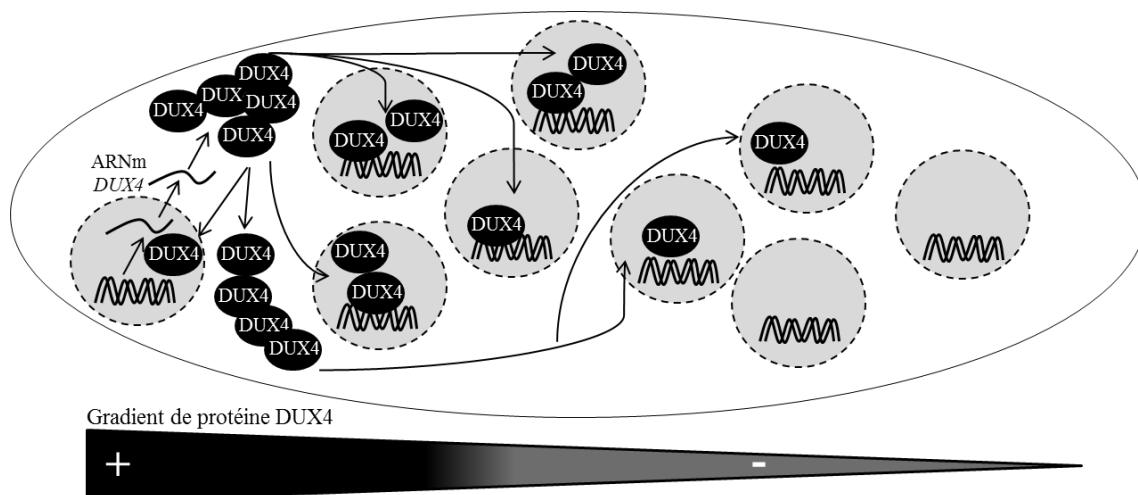


FIGURE 26. Modèle de diffusion de la protéine *DUX4* au sein d'un myotube. L'ARNm *DUX4* est transcrit dans un noyau. Cet ARNm est traduit localement dans le cytoplasme à proximité du noyau émetteur. La protéine diffuse ensuite dans la fibre avant d'être importée dans les nombreux noyaux avoisinants où elle induit des dérégulations génétiques majeures. Ainsi le long de la fibre musculaire, un gradient de protéines *DUX4* est observé.

Des expériences de co-cultures de myoblastes murins et de myoblastes humains dérivant de patients FSHD, ainsi que le suivi de la protéine par fluorescence, ont permis d'expliquer ces résultats. L'ARNm *DUX4* est exprimé à partir d'un noyau (Figure 26). Cet ARNm est traduit localement dans le cytoplasme à proximité du noyau émetteur (Figure 26). Les protéines *DUX4* diffusent ensuite dans la fibre musculaire avant d'être importées dans les nombreux noyaux avoisinants où elles induisent des dérégulations génétiques majeures (Ehrlich and Lacey, 2012; Ferreboeuf et al., 2014b; Rickard et al., 2015; Tassin et al., 2013) (Figure 26). Le phénotype FSHD diffuse ainsi dans toute la fibre (Figure 26). Par ailleurs, la forte activation transcriptionnelle des gènes cibles de *DUX4* observée dans toute la fibre peut se faire selon deux modèles : i) un modèle de forte activation transcriptionnelle des gènes cibles dans un nombre limité de noyaux ou ii) un modèle de faible activation transcriptionnelle des gènes cibles dans plusieurs noyaux. Le dernier modèle semble être privilégié (Ferreboeuf et al., 2014b).

#### **1.3.6.1.8. Boucle de rétro-contrôle positif de l'expression de *DUX4* ?**

Le *nonsense-mediated decay* (NMD) est un mécanisme de contrôle qualité des ARNm. Lorsqu'un ARNm est considéré comme aberrant, il est pris en charge par les protéines effectrices du NMD et dégradé. Les substrats du NMD sont principalement les ARNm avec un long 3'UTR ou possédant des introns après un codon stop et les ARNm présentant un codon de terminaison précoce, PTC (*Premature Termination Codon*), ce qui conduit à la synthèse d'une protéine tronquée. Les PTC, localisés plus de 50 nucléotides en aval de la dernière jonction exon-exon, ont plusieurs origines : mutations du gène, épissages aberrants, réarrangements géniques ou insertions de retrotransposons (Chang et al., 2007; Kervestin and Jacobson, 2012; Nagy and Maquat, 1998; Nicholson et al., 2010).

Dans des cellules exprimant *DUX4*, l'efficacité du NMD est réduite (Feng et al., 2015). *DUX4* induit une dégradation de UPF-1 (*Up-Frameshift Suppressor 1*) par le protéasome. UPF-1 est une protéine effectrice du NMD. Sa dégradation a pour conséquence l'accumulation d'ARNm aberrants, toxiques pour la cellule. Le mécanisme médié par *DUX4* qui provoque la dégradation de UPF-1 n'est pas connu. Cependant, *DUX4* pourrait induire des dysfonctionnements du système ubiquitine-protéasome, dysfonctionnements qui auraient pour conséquence la dégradation de UPF-1 (Feng et al., 2015).

De manière intéressante, la longue région 3'UTR composée d'introns des variants *DUX4-fl* font d'eux des substrats du NMD. Plus précisément, le codon stop de *DUX4* est situé : soit à 92 nucléotides de la dernière jonction exon-exon, lorsque l'intron 1 est épissé, soit à 236 nucléotides de la dernière jonction exon-exon, lorsque l'intron 1 n'est pas épissé. Or, il est établi qu'un ARNm présentant un codon stop localisé plus de 50-55 nucléotides en amont de la jonction exon-exon la plus en 3', est pris en charge par le NMD (Nagy and Maquat, 1998). Ceci est confirmé par l'observation d'une augmentation de l'expression de *DUX4* dans des cellules UPF-1 *knock-down* (K.D.). Ainsi, les variants *DUX4-fl* sont normalement éliminés par le NMD (Feng et al., 2015).

En résumé, la protéine *DUX4* inhibe le NMD tandis que le NMD dégrade l'ARNm *DUX4* (Figure 27). Cette boucle de régulation complexe permettrait à *DUX4*, en inhibant le NMD, de stabiliser son propre ARNm localement mais aussi au niveau des noyaux avoisinants suite à la diffusion de la protéine dans la fibre musculaire (Figure 27). Le phénotype FSHD se propage ainsi dans la fibre musculaire. Cependant, il n'est pas établi si l'augmentation de l'expression de la quantité relative d'ARNm *DUX4* est la conséquence d'une augmentation du nombre de transcrits *DUX4* au noyau initialement émetteur ou une augmentation du nombre de noyaux produisant le transcrit *DUX4* (Feng et al., 2015) (Figure 27).

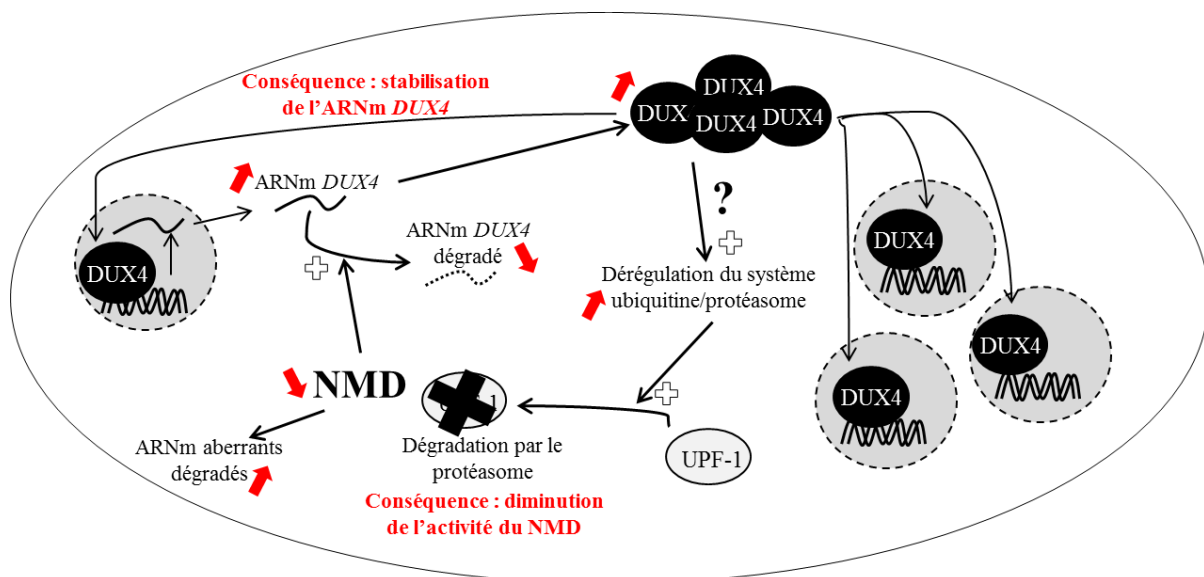


FIGURE 27. Boucle de rétro-contrôle positif de l'expression de *DUX4* par le NMD. La protéine *DUX4* inhibe le NMD en augmentant, selon un mécanisme non déterminé, la dégradation par le protéasome de UPF-1, une protéine effectrice du NMD. Cette inhibition du NMD permet une stabilisation de l'ARNm *DUX4*, une cible du NMD. La stabilisation de l'ARNm *DUX4* a pour conséquence une augmentation de la quantité de protéines *DUX4*. Ces protéines diffusent et propagent le phénotype FSHD dans la fibre musculaire.

En conclusion, bien que plusieurs gènes candidats aient été proposés pour la FSHD, et malgré le fait que l'expression de *DUX4* n'ait pas encore été directement liée au phénotype des patients FSHD, il est admis par la communauté scientifique que *DUX4* joue un rôle majeur dans cette pathophysiologie car (i) il faut au moins une répétition D4Z4 pour développer la maladie, (ii) seuls les allèles dits permissifs de type 4qA, possédant le signal de polyadénylation de *DUX4* localisé en aval de la dernière répétition D4Z4, sont associés à la FSHD, (iii) la contraction des répétitions D4Z4 homologues sur le chromosome 10qA, qui porte en aval de la dernière répétition un signal de polyadénylation muté, ne conduit pas à la FSHD, (iv) certains individus asymptomatiques portent les contractions D4Z4 mais sans la perte des marques épigénétiques répressives et ne présentent pas de signes de la maladie, (v) l'expression de gènes en aval de *DUX4* est une signature moléculaire de la FSHD et (vi) l'expression de *DUX4* est le point commun entre la FSHD1 et la FSHD2.

### **1.3.7. Diagnostic génétique de la FSHD**

Le diagnostic de la FSHD est, à l'heure actuelle, avant tout clinique. Un diagnostic génétique peut ensuite être réalisé pour confirmer ou infirmer le diagnostic clinique. Le diagnostic génétique est basé, à ce jour, uniquement sur la détermination du nombre de répétitions D4Z4 en région subtélomérique du chromosome 4.

#### **1.3.7.1. Diagnostic génétique par *southern blot***

La méthode la plus couramment utilisée dans le diagnostic génétique de la FSHD est le *southern blot*. L'ADN génomique est digéré par l'enzyme de restriction *EcorI* (ou le couple d'enzymes de restriction *EcorI* et *HindIII*) qui permet de couper de part et d'autre de la région formée de répétitions D4Z4. Les fragments d'ADN sont ensuite séparés par électrophorèse sur gel ou par électrophorèse en champ pulsé. L'électrophorèse sur gel permet d'étudier des fragments d'ADN entre 3 et 50kb, alors que l'électrophorèse en champ pulsé, plus résolutive, permet de séparer des fragments de quelques centaines de kb. Le nombre de répétitions D4Z4 est déterminé après transfert sur une membrane de nylon et hybridation avec la sonde p13E-11. La sonde p13E-11 reconnaît une région de 0,1kb immédiatement en amont de la première répétition D4Z4 (Hewitt et al., 1994; Wright et al., 1993b). Les fragments obtenus ont une

taille variant entre 10kb et 38kb pour les patients FSHD1 et entre 40kb et 350kb pour les individus contrôles (Ehrlich et al., 2007; Lemmers et al., 2012b) (Tableau 4).

La détermination du nombre de répétitions D4Z4 sur le chromosome 4q est compliquée du fait de la forte homologie entre la région subtélomérique du chromosome 4q et celle du chromosome 10q, l'homologie incluant la région d'hybridation de la sonde p13E-11. Pour discriminer le chromosome 4q du chromosome 10q, des digestions supplémentaires des fragments d'ADN sont réalisées. En raison de polymorphismes nucléotidiques, des sites de restriction différents entre les deux chromosomes sont présents dans les répétitions D4Z4 (Figure 28). Ainsi, une double digestion avec le couple d'enzymes de restriction *EcoRI* et *BlnI* fragmente les répétitions D4Z4 du chromosome 10q mais laisse celles du chromosome 4q intactes (Figure 28). Au contraire, une double digestion avec le couple d'enzymes de restriction *EcoRI* et *XapI* fragmente les répétitions D4Z4 du chromosome 4q et laisse celles du chromosome 10q intactes (cf review Lemmers et al., 2012b) (Figure 28).

TABLEAU 4. Tableau de conversion de la taille du fragment digéré en nombre de répétitions D4Z4 en fonction des enzymes utilisées. La première colonne indique le nombre de répétitions. La deuxième, la troisième et la quatrième colonne donnent la taille attendue du fragment D4Z4 en *southern blot* après une digestion par *EcoRI* ou le couple *EcoRI/HindIII* ou le couple *EcoRI/BlnI* respectivement (adapté de Lemmers, 2017).

Nombre de répétitions D4Z4	<i>EcoRI</i> (en kb)	<i>EcoRI/HindIII</i> (en kb)	<i>EcoRI/BlnI</i> (en kb)
1	10.2	8.1	7.0
2	13.5	11.4	10.3
3	16.8	14.7	13.6
4	20.1	18.0	16.9
5	23.4	21.3	20.2
6	26.7	24.6	23.5
7	30.0	27.9	26.8
8	33.3	31.2	30.1
9	36.6	34.5	33.4
10	39.9	37.8	36.7
11	43.2	41.4	40.0
12	46.5	44.4	43.3
13	49.8	47.7	46.6

Le nombre de répétitions est alors calculé selon les formules suivantes (Lemmers, 2017) (Tableau 4, Figure 28) :

$$\text{Digestion } EcoRI : \text{Nombre de répétitions} = [(\text{Taille du fragment} - 6900) / 3300]$$



Digestion *EcoRI/HindIII* : Nombre de répétitions =  $[(\text{Taille du fragment} - 4800) / 3300]$

Digestion *EcoRI/BlnI* : Nombre de répétitions =  $[(\text{Taille du fragment} - 3700) / 3300]$

Cependant, dans 20% des cas, des échanges subtélomériques entre le chromosome 4q et 10q sont observés (van Deutekom et al., 1996a; Lemmers et al., 1998; Matsumura et al., 2002). Ainsi, des unités D4Z4 issues du chromosome 10q et *XapI* résistantes, peuvent être transloquées en place et lieu de répétitions D4Z4 *BlnI* résistantes du chromosome 4q. A cela s'ajoute, dans 1 à 3% des cas, une délétion dans la région en amont des répétitions D4Z4, ne permettant plus à la sonde p13E-11 de s'hybrider et conduisant à l'obtention de faux négatifs (Deak et al., 2007; Lemmers et al., 2003, 2004c). L'interprétation du *southern blot* devient alors compliquée. D'autres sondes, en complément de la sonde p13E-11, peuvent être utilisées pour pallier à certains de ces problèmes (Ehrlich et al., 2007).

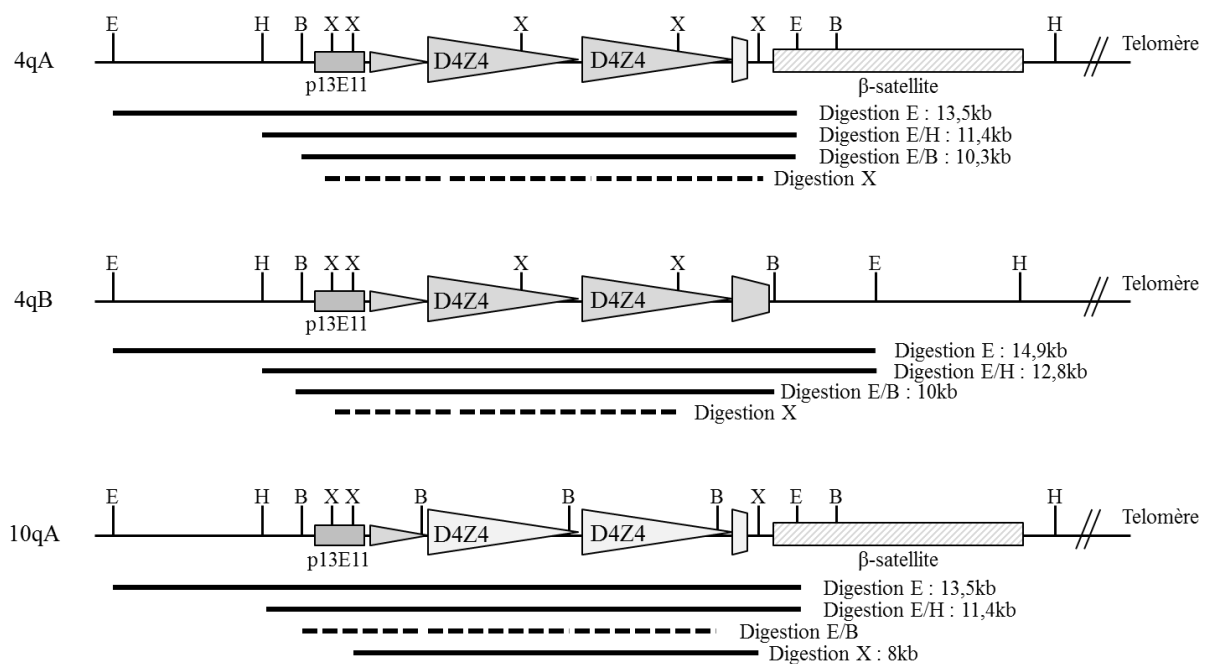


FIGURE 28. Schéma des profils attendus en *southern blot* après digestion par différentes enzymes de restriction de la région subtélomérique des chromosomes 4qA/B et 10qA, suivie d'une hybridation avec la sonde p13E-11. Les profils de digestion présentés sont ceux attendus pour une digestion des chromosomes 4qA/B et 10qA avec deux répétitions D4Z4 de 3,3kb. Les différents sites de restriction utilisés sont positionnés avec *E* : *EcoRI*, *H* : *HindIII*, *B* : *BlnI* et *X* : *XapI*. La taille de chaque produit de digestion est indiquée en kb (adapté de Ehrlich et al., 2007).

### 1.3.7.2. Diagnostic génétique par peignage moléculaire

Une autre technique utilisée aujourd'hui dans le diagnostic de la FSHD est celle du peignage moléculaire. Le peignage moléculaire permet de visualiser directement l'organisation génomique d'une molécule d'ADN à l'aide de sondes fluorescentes. Sa résolution est de 1kb, donnant lieu à des cartographies précises (Bensimon et al., 1994; Lebofsky and Bensimon, 2003). D'un point de vue technique, l'ADN préparé est déposé sur une lame en verre traitée avec du silane (SiH<sub>4</sub>), ce qui favorise la fixation de l'ADN par l'une de ses extrémités à la lame, et ce avec une haute affinité. Par action mécanique, qui consiste à retirer la solution dans laquelle se trouvait l'ADN à une vitesse constante de 300µm/s, les molécules d'ADN sont étirées, séparées et alignées sur la lame. Enfin, des sondes fluorescentes de plusieurs couleurs sont utilisées simultanément pour une hybridation à des régions spécifiques sur ces molécules d'ADN (Bensimon et al., 1994).

Dans le cadre de la FSHD, plusieurs sondes sont utilisées pour différencier visuellement les allèles des chromosomes 4qA, 4qB, 10qA et 10qB (Figure 29). Deux sondes bleues s'hybrident à la région de 42kb en amont des répétitions D4Z4, région commune aux chromosomes 4q et 10q (Figure 29). En amont de cette région de 42kb, des sondes rouges s'hybrident sur les régions spécifiques du chromosome 4q et 10q, permettant ainsi leur discrimination (sondes de 10kb pour le chromosome 4q et de 5kb pour le chromosome 10q) (Figure 29). D'autres sondes s'hybrident à la région subtélomérique. Plus précisément, une sonde rouge spécifique de l'haplotype 4/10qA et une sonde bleue spécifique de l'haplotype 4/10qB sont utilisées (Figure 29). Le télomère est quant à lui marqué par une sonde rouge (Figure 29). Des sondes vertes s'hybrident à chacune des répétitions D4Z4 (Figure 29). La longueur du signal vert est proportionnel au nombre de répétitions D4Z4 (Nguyen et al., 2011; Vasale et al., 2015) (Figure 29).

Le peignage moléculaire permet donc de visualiser les différents haplotypes des chromosomes 4q et 10q en une seule expérience, alors que de multiples digestions sont nécessaires avec la technique du *southern blot*. Par ailleurs, le peignage moléculaire permet d'effectuer un diagnostic même en cas de délétions ou de réarrangements chromosomiques des régions en amont, en aval ou des répétitions D4Z4, du fait d'une observation directe des chromosomes. De plus, 40% des cas *de novo* de FSHD sont des patients mosaïques (van der Maarel et al.,

2000). Dans le cadre de l'utilisation d'un *southern blot*, le mosaïcisme est détectable pour un minimum de 15% de cellules FSHD (van der Maarel et al., 2000). Le peignage moléculaire, plus sensible, peut identifier un mosaïcisme inférieur à 10% (Nguyen et al., 2011).

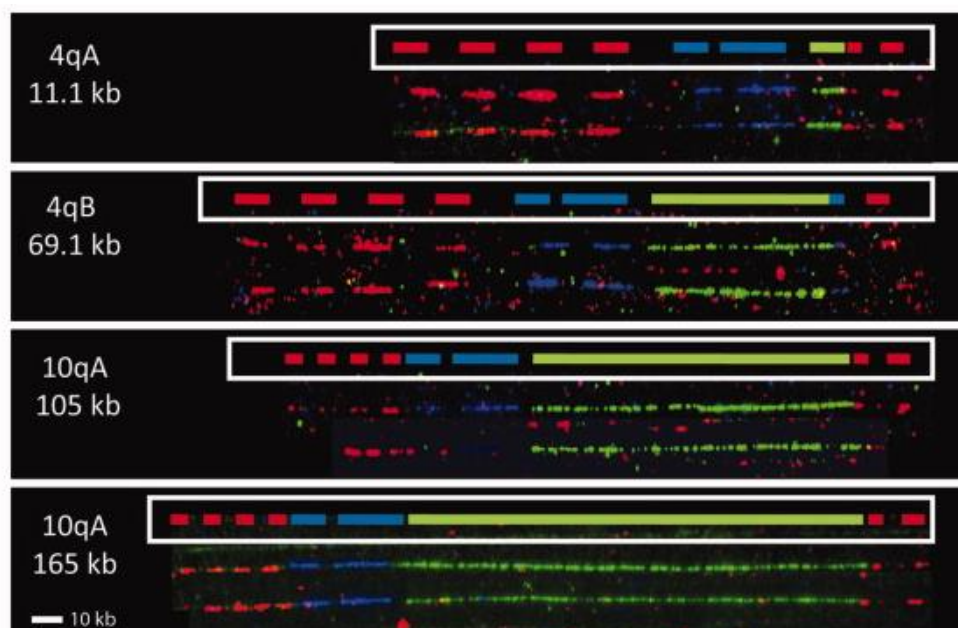
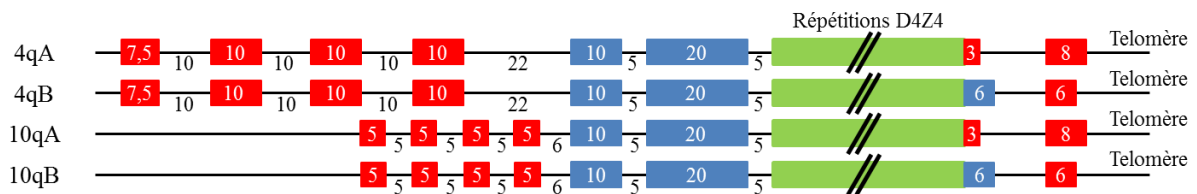


FIGURE 29. Schéma des profils attendus en peignage moléculaire pour les régions subtélocériques des chromosomes 4qA/B et 10qA/B. Les chromosomes 4qA/B et 10qA/B sont différenciés par la combinaison de plusieurs sondes : deux sondes bleues qui s'hybrident à la région de 42kb en amont des répétitions D4Z4, qui est commune à tous les chromosomes ; des sondes rouges qui s'hybrident en amont des sondes bleues et sont spécifiques de chaque chromosome (sondes de 10kb pour le chromosome 4q et de 5kb pour le chromosome 10q) ; une sonde rouge spécifique de l'haplotype 4/10qA et une sonde bleue spécifique de l'haplotype 4/10qB en aval des répétitions D4Z4 ; une sonde rouge télomérique ; des sondes vertes qui s'hybrident à chaque répétition D4Z4 et qui génèrent un signal dont la longueur est proportionnelle au nombre de répétitions D4Z4. Un exemple de peignage moléculaire obtenu pour différents chromosomes est présenté, avec, en encadré blanc, la visualisation théorique attendue. La taille en kb de la région formée par l'ensemble des répétitions D4Z4 pour chaque échantillon testé est indiquée (adapté de Nguyen et al., 2011; Vasale et al., 2015).

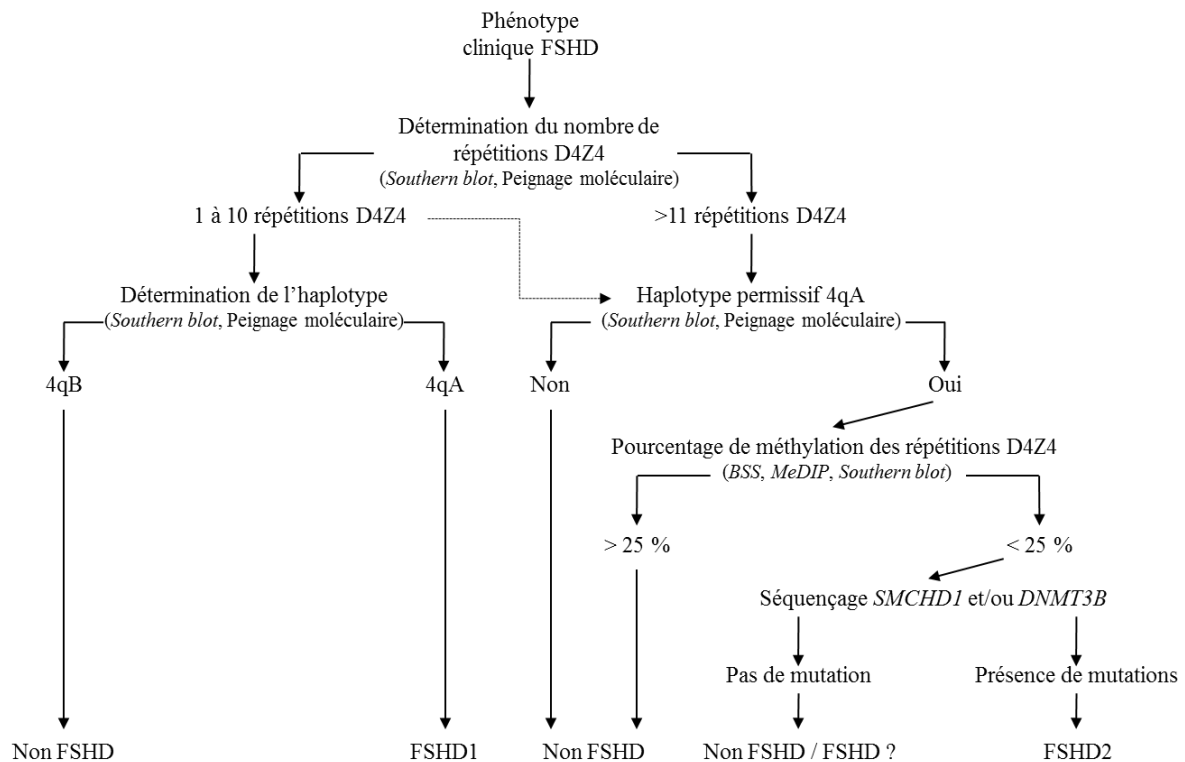


FIGURE 30. Récapitulatif des différentes étapes actuelles et futures dans le diagnostic de la FSHD (adapté de Larsen et al., 2015).

Les précédentes techniques de diagnostic sont adaptées uniquement pour le diagnostic de patients FSHD avec des contractions des répétitions D4Z4, ce qui n'est pas le cas des patients FSHD2. Il est donc proposé de développer, dans les années à venir, un diagnostic génétique évaluant également le niveau de méthylation de l'ADN au niveau des répétitions D4Z4 pour confirmer le diagnostic clinique de patients FSHD2. Plusieurs méthodes, utilisées en recherche fondamentale, sont possibles pour quantifier le niveau de méthylation des répétitions D4Z4 : séquençage après traitement bisulfite, immunoprécipitation de l'ADN méthylé (MeDIP *Methylated DNA Immunoprecipitation*) ou encore *southern blot* après une digestion par des enzymes de restriction sensibles à la méthylation (Gaillard et al., 2014; Jones et al., 2015; Larsen et al., 2015). Ainsi, pour un individu présentant les signes cliniques de la FSHD, un allèle permissif 4qA, mais aucune contraction des répétitions D4Z4, si le taux de méthylation moyen au niveau de ses répétitions D4Z4 est inférieur à 25% comparé à celui des individus contrôles, le patient pourra être considéré comme FSHD2. Un séquençage de *SMCHD1* et/ou *DNMT3B* pourra éventuellement être réalisé afin de rechercher des mutations dans ces modificateurs épigénétiques et venir ainsi compléter le diagnostic génétique (van den Boogaard et al., 2016; Larsen et al., 2015) (Figure 30).

## 1.4. Modèles cellulaires et modèles animaux FSHD

Les modèles cellulaires et animaux sont des outils essentiels pour la compréhension des facteurs génétiques et des mécanismes moléculaires impliqués dans les processus pathologiques, mais également pour le développement de stratégies thérapeutiques.

### 1.4.1. Modèles cellulaires FSHD

Les myoblastes humains FSHD ou les biopsies FSHD sont de bons modèles pour l'étude des dérégulations génétiques et métaboliques chez les patients FSHD ainsi que le test de stratégies thérapeutiques *in vitro*. Les myoblastes primaires FSHD sont obtenus par isolation et purification de cellules, sortant des explants préparés à partir de biopsies de patients FSHD. Par la suite, des lignées cellulaires immortelles FSHD peuvent être créées par une transduction rétrovirale des myoblastes primaires FSHD avec le gène *hTERT* (*human Telomerase Reverse Transcriptase*) et le gène *CDK4* (*Cycline Dependante Kinase 4*) (Krom et al., 2012; Mamchaoui et al., 2011; Stadler et al., 2011; Zhu et al., 2007). Ce protocole a notamment été utilisé pour générer plusieurs lignées immortelles de cellules isogéniques contractées et non contractées aux répétitions D4Z4 à partir de la biopsie d'un patient FSHD mosaïque (Krom et al., 2012). Ces lignées ont l'avantage de posséder le même génome à l'exception du nombre de répétitions D4Z4. Ces lignées permettent également de s'affranchir de toutes les variations entre individus, qui pourraient biaiser les études des mécanismes menant au développement de la FSHD. Par ailleurs, ces cellules proviennent d'une même biopsie, donc d'un même muscle, ce qui permet cette fois-ci d'éliminer toutes les variations liées au degré d'atteinte variable entre les muscles (Krom et al., 2012). Le désavantage de ces modèles cellulaires humains FSHD est le faible taux d'expression de *DUX4* associé à des difficultés de détection de la protéine DUX4, rendant l'expérimentation parfois compliquée (Tassin et al., 2013).

Des modèles artificiels de cellules (humaines, murines) peuvent être générés par la transfection ou la transduction de vecteurs exogènes exprimant *DUX4*. Les vecteurs inductibles ou non couramment utilisés sont : des variants du plasmide pCINéo-DUX4 (déposé par Alexandra Belayew), qui se caractérise par la présence d'une répétition D4Z4 entière suivie de la région pLAM ; et des variants du plasmide pCS2-mkgDUX4 (déposé par

Stephen Tapscott), qui contient uniquement l'ORF de *DUX4* (exon 1) immédiatement suivi du signal de polyadénylation du virus simien 40 (Anseau et al., 2009; Dixit et al., 2007; Snider et al., 2009). L'utilisation de ces vecteurs induit une expression de *DUX4* non physiologique, bien plus importante que ce qui peut être observé dans des cellules humaines primaires FSHD, pouvant conduire à des erreurs d'interprétations quant au rôle de *DUX4* dans la cellule. Par ailleurs, l'expression de *DUX4* dans des cellules non humaines peut également augmenter les erreurs d'interprétations quant au rôle de *DUX4* dans le développement de la pathologie. La raison est simple : les répétitions D4Z4 contenant le rétrogène *DUX4* sont des éléments spécifiques des primates de l'ancien monde (les catarrhiniens) et des *Afrotheria* (Clapp et al., 2007; Leidenroth and Hewitt, 2010; Leidenroth et al., 2012). Utiliser un modèle cellulaire non humain FSHD n'exprimant pas *DUX4* de manière endogène, n'a de sens que si l'expression exogène de *DUX4* est régulée par le même réseau et active les mêmes gènes que ce qui est observé dans les cellules des primates de l'ancien monde. Or il a été montré qu'une expression exogène de *DUX4* dans des cellules humaines (rhabdomyosarcomes) ou murines (C2C12) de la lignée musculaire n'induit pas la même dérégulation à la fois qualitative et quantitative des gènes dans ces cellules (Sharma et al., 2013). Par ailleurs, cette surexpression est souvent associée à une forte toxicité et de l'apoptose (Vanderplanck et al., 2011). Le choix du modèle cellulaire doit donc être pensé en fonction de la question biologique posée.

## **1.4.2. Modèles animaux FSHD**

La création de modèles animaux FSHD est un réel défi (Lek et al., 2015). Les raisons sont les mêmes que celles évoquées précédemment, c'est-à-dire notamment la présence uniquement de l'ORF *DUX4* chez les primates de l'ancien monde et les *Afrotheria* (Clapp et al., 2007). Le travail avec une espèce animale de laboratoire non transgénique est donc impossible. La solution consiste en la génération de modèles FSHD non-primates par introduction d'une construction génétique exogène exprimant *DUX4*, sortant ainsi le gène de son réseau naturel.

### **1.4.2.1. Modèle du poisson zèbre et du xénope**

Un modèle de poisson zèbre (*Danio reio*) FSHD a été généré par l'injection *in vitro* dans des œufs fécondés au stade une cellule d'ARNm *DUX4-fl* (Mitsuhashi et al., 2013a). Une trop grande quantité d'ARNm *DUX4-fl* ( $>1.10^7$  copies) est cytotoxique dans les 24h suivant

l'injection pour plus de 95% des œufs fécondés. Une injection d'environ 1 à  $2.10^5$  copies, donnant approximativement une molécule d'ARN par cellule après dilution au cours du développement avant somitogenèse, est cytotoxique pour 67% des embryons à quatre jours après fertilisation. Les embryons survivants ont une durée de vie très courte. L'expression de *DUX4-fl* est donc extrêmement toxique dans ces organismes. Pour les plus faibles quantités d'ARNm injectées, les embryons de poisson zèbre montrent des anomalies asymétriques, à sévérité variable, au niveau des yeux, des oreilles, des nageoires (équivalent des bras), de la face et/ou du tronc musculaire, à l'image de ce qui est observé dans la FSHD. A cela s'ajoute une réduction de la capacité de nage. Par ailleurs, les cellules myogéniques présentent une localisation aberrante dans la région céphalique. A l'échelle cellulaire, le muscle squelettique est totalement désorganisé, surtout aux niveaux des sarcomères où les disques Z indistincts et les *myosepta* sont anormalement alignés. Aucune de ces anomalies n'est obtenue avec une expression de *DUX4-s*. La fonction pathogénique de *DUX4-fl* est donc conservée au moins dans certains organismes non primates (Mitsuhashi et al., 2013a).

Un autre modèle de poisson zèbre, cette fois-ci transgénique, a également été généré (Wallace et al., 2011a). Une séquence contenant *DUX4* sous un promoteur muscle spécifique MHCK7 (*hybrid  $\alpha$ -myosin heavy chain enhancer/ MCK enhancer-promoter*) est intégré au génome du poisson zèbre, par utilisation du système de transposon Tol2 dans des œufs fécondés au stade une cellule. Le promoteur MHCK7 est décrit comme permettant une expression du gène en aval, trois jours après l'injection, ce qui limite l'étude des effets de l'expression de *DUX4* aux stades plus tardifs de développement, en comparaison au modèle précédent. Chez ces poissons zèbres, quatre jours après fécondation, 47% des embryons ont une grossière malformation du corps et 81% ont des anomalies musculaires importantes. Le phénotype de ces poissons zèbres est similaire au modèle précédemment décrit, auquel s'ajoute une hypertrophie cardiaque pour certains individus (Wallace et al., 2011a).

Ces modèles de poisson zèbre, qui récapitulent les principaux symptômes associés à la FSHD dont l'atteinte musculaire asymétrique, sont des outils utiles pour la compréhension des mécanismes physiopathologiques médiés par *DUX4* et pour le test de stratégies thérapeutiques (Mitsuhashi et al., 2013a; Wallace et al., 2011a).

De la même manière, des modèles embryonnaires de *Xenopus laevis* FSHD existent (Wuebbles et al., 2010). Ces modèles sont générés par injection au stade deux cellules de

l'ARNm *DUX4* ou *PITX1* (dont un orthologue est présent chez le xénope). L'expression de l'un ou l'autre de ces deux gènes conduit à une perte cellulaire massive non restreinte au muscle du côté de l'injection. Une surexpression des protéines DUX4 et PITX1 est toxique, la cytotoxicité de DUX4 étant plus importante que celle de PITX1 (Wuebbles et al., 2010). Les modèles xénopes FSHD sont très peu utilisés en raison de l'incompatibilité de l'expression de DUX4 ou PITX1 avec le développement embryonnaire.

#### **1.4.2.2. Modèle de la drosophile**

La drosophile (*Drosophila melanogaster*) est un organisme modèle déjà utilisé avec succès pour étudier les mécanismes pathogéniques de plusieurs maladies neuromusculaires (Allikian et al., 2007; Chan et al., 2003; Chartier et al., 2006; de Haro et al., 2006; Pantoja and Ruohola-Baker, 2013). Récemment, un modèle de drosophile pour la FSHD a été mis au point pour étudier les voies de signalisation affectées par l'expression de *DUX4* ou *FRG1* (Jones et al., 2016). Ces lignées de drosophiles transgéniques expriment, sous contrôle du système UAS-GAL4 tissu-spécifique (*GAL4-upstream activation sequence*), *FRG1* dans les muscles adultes ou *DUX4-fl* uniquement dans les yeux (Duffy, 2002; Jones et al., 2016). En effet, une expression ubiquitaire de *DUX4-fl* au cours du développement chez la drosophile résulte en une mortalité de 100%. De plus, l'induction de l'expression de *DUX4-fl* plus tardivement et spécifiquement dans des muscles primaires chez l'adulte conduit également à une mortalité de 100%. Il est donc impossible de générer des drosophiles exprimant *DUX4-fl* au cours du développement ou spécifiquement dans les muscles, même chez l'adulte. Cependant, une expression de *DUX4-fl* restreinte aux yeux, et ce dès les premières étapes du développement, génère une mortalité moins prononcée chez les mâles, tandis que la grande majorité des femelles sont viables. Ainsi, certaines interactions géniques de *DUX4-fl* peuvent être étudiées avec cette lignée de drosophiles transgéniques, à condition que le réseau d'interaction naturel de *DUX4-fl* soit aussi présent chez la drosophile et plus particulièrement au niveau des yeux (Jones et al., 2016). L'utilisation de ce modèle FSHD *DUX4* reste tout de même restreinte.

Afin d'identifier les voies métaboliques modifiées par une surexpression de *FRG1*, des lignées de drosophiles transgéniques exprimant *FRG1* spécifiquement dans les muscles thoraciques adultes ont été générées (Jones et al., 2016). Ces lignées transgéniques sont viables. Elles se caractérisent par une réduction de la capacité à voler malgré un



développement normal des ailes. Ceci s'explique par une perturbation de l'organisation et de l'intégrité de la musculature, comparable à ce qui est observé dans d'autres modèles. Ce modèle transgénique *FRG1* représente donc également un outil pour la compréhension des voies développementales perturbées dans le muscle par l'expression de *FRG1* (Jones et al., 2016).

### **1.4.2.3. Modèles murins**

#### **1.4.2.3.1. Modèles murins exprimant d'autres gènes candidats de la FSHD**

Plusieurs gènes ont été proposés comme impliqués dans le développement de la FSHD. Afin d'étudier leur rôle dans la pathologie, plusieurs modèles murins sur-exprimant ou sous-exprimant ces gènes ont été créés (Lek et al., 2015).

##### **1.4.2.3.1.1. Modèle murin sur-exprimant *FRG1***

Des souris transgéniques sur-exprimant fortement *FRG1* sous contrôle du promoteur HSA (*Human Skeletal  $\alpha$ -actin*) développent un phénotype sévère de cyphose (déformation de la colonne vertébrale) indépendamment d'une déformation du squelette, suggérant une faiblesse musculaire pour origine (Gabellini et al., 2006). Cette courbure de la colonne vertébrale est aussi retrouvée chez les patients FSHD. Ces souris se caractérisent également par un phénotype dystrophique progressif avec une nécrose des fibres musculaires striées squelettiques, une infiltration de tissu conjonctif accompagnée d'une réduction du poids des souris et d'une difficulté à la réalisation d'exercices. Par ailleurs, l'atrophie musculaire chez la souris *FRG1* est variable selon les muscles considérés, comme c'est le cas chez les patients FSHD (Gabellini et al., 2006).

Au niveau moléculaire, les ARNm *TNNT3* et *MTMRI* sont anormalement épissés chez les souris exprimant *FRG1*. *TNNT3* et *MTMRI* régulant respectivement la contractilité musculaire et l'atrophie musculaire, ces deux gènes peuvent contribuer au développement d'une atrophie musculaire dans ce modèle murin mais aussi dans la FSHD (Buj-Bello et al., 2002; Chaudhuri et al., 2005; Gabellini et al., 2006; Galińska-Rakoczy et al., 2008; Gordon et al., 2000; Perry, 1998; Sancisi et al., 2014). D'autres ARNm sont potentiellement affectés, et

pourraient expliquer le phénotype exclusivement musculaire observé chez des souris exprimant fortement *FRG1* (Gabellini et al., 2006).

L'intérêt scientifique pour ce modèle de souris est tout de même mitigé, compte tenu du fait qu'une surexpression de *FRG1* chez les patients FSHD n'a pas pu être démontrée par de nombreux laboratoires (Arashiro et al., 2009; Celegato et al., 2006; van Deutekom et al., 1996b; Jiang et al., 2003; Klooster et al., 2009; Masny et al., 2010; Osborne et al., 2007; Winokur et al., 2003a). Cependant, *FRG1* étant potentiellement une cible transcriptionnelle de DUX4, son rôle dans la FSHD ne peut être totalement exclu (Ferri et al., 2015).

#### **1.4.2.3.1.2. Modèle murin sur-exprimant *Pitx1***

Pour vérifier le potentiel rôle pathogénique d'une surexpression de *PITX1* dans la FSHD, des souris transgéniques sur-exprimant conditionnellement *Pitx1* sous promoteur mCK (*muscle specific Creatine Kinase promoter*) en l'absence d'administration de doxycycline, ont été générées (Pandey et al., 2012, 2014). Une sur-expression de *Pitx1* pendant cinq semaines chez des souris âgées de 10 semaines conduit à une réduction de la masse de certains groupes musculaires associée à une diminution de la force, qui s'observe, elle, trois semaines après l'induction de l'expression de *Pitx1*. L'histologie des muscles révèle une atrophie des fibres musculaires, de la nécrose et des infiltrations inflammatoires, caractéristiques retrouvées chez les patients FSHD. Au niveau moléculaire, une expression de p53 est détectée dans les muscles affectés de la souris *Pitx1* (Pandey et al., 2012, 2014).

#### **1.4.2.3.1.3. Modèle murin muté pour *FAT1***

Un modèle murin pouvant être utile à la compréhension de la FSHD est une souris *knock out* (K.O.) pour le gène codant la protocadhérine FAT1 (Caruso et al., 2013). Une inactivation constitutive de *Fat1* chez la souris entraîne une mortalité néonatale confirmant l'importance de cette protéine au cours du développement. Des podocytes anormaux sont probablement à l'origine de cette létalité importante chez les souris homozygotes *Fat1*<sup>-/-</sup>. Ce défaut de développement du rein peut être associé à un défaut de développement du cerveau antérieur et un défaut de développement des yeux (Ciani et al., 2003). Un autre modèle murin, des souris homozygotes hypomorphiques *Fat1* (délétion d'une partie des répétitions cadhérines, du domaine lamine A-G et des motifs EFG), présente un taux de mortalité de 50% dans les trois

mois suivant la naissance. Ces souris montrent également des anomalies développementales au niveau des reins et récapitulent le phénotype musculaire et non musculaire de la FSHD : aux stades postnataux, ces souris se caractérisent par une réduction sélective et asymétrique de la masse musculaire, au niveau des muscles des épaules et de la face, associée parfois à une rétinopathie. Aux stades plus avancés, une dystrophie musculaire étendue est observable. A l'échelle cellulaire, les souris hypomorphiques *Fat1* présentent des signes de nécrose des fibres musculaires, des infiltrations inflammatoires et des sarcomères désorganisés. L'expression de *FAT1* étant réduite chez les patients FSHD, ce modèle murin est un outil très utile pour la compréhension des mécanismes cellulaires indépendants de DUX4 pouvant être impliqués dans le développement de la FSHD (Caruso et al., 2013; Mariot et al., 2015).

#### **1.4.2.3.2. Modèles murins exprimant *DUX4***

##### **1.4.2.3.2.1. Souris AAV6-DUX4**

Plusieurs modèles murins exprimant *DUX4* ont été développés ces dernières années. Un premier modèle, transitoire, est la souris AAV6-DUX4 (Wallace et al., 2011a). Ce modèle murin consiste en l'injection du vecteur adéno-associé (AAV) non-intégratif AAV6-DUX4 dans un tissu musculaire précis de souris, permettant d'initier une expression contrôlée de *DUX4* dans le temps et l'espace. Ce modèle a été le premier à montrer qu'une surexpression de *DUX4* dans le muscle cause des dommages importants, dont une dégénération des fibres musculaires en lien avec des infiltrations de cellules mononucléaires et l'expression de marqueurs pro-apoptotiques. Ces effets de *DUX4* passent par une voie dépendante du suppresseur de tumeur p53. Ce modèle est transitoire : les souris montrent une faible force lors de tests d'agrippement, et ce durant les deux semaines suivants l'injection, avant de retrouver une force normale dans les trois semaines. Ceci s'explique par le remplacement des fibres musculaires endommagées exprimant *DUX4* par des fibres saines (Wallace et al., 2011a). Ces résultats laissent penser que le phénotype FSHD induit par l'expression de *DUX4* pourrait être potentiellement inversé par une stratégie thérapeutique inhibant l'expression de *DUX4*.

#### 1.4.2.3.2.1. Souris iDUX4(2.7)

Un second modèle murin transgénique, nommé iDUX4(2.7), a été généré par intégration d'un transgène *DUX4*, inductible par la doxycycline. Le transgène est intégré en amont du gène *HPRT* (*Hypoxanthine PhosphoRibosyl Transferase*), au niveau d'une région d'euchromatine sur le chromosome X (Dandapat et al., 2014). Le transgène correspond à un fragment d'ADN génomique de 2,7kb contenant l'ORF de *DUX4* et une partie de sa région 3'UTR. Le signal de polyadénylation est apporté par le signal de polyadénylation SV40.

Dans ce modèle, en absence de doxycycline, une faible expression de *DUX4* est détectée dans divers tissus, à l'exception du muscle, conduisant à une létalité embryonnaire chez les mâles. Les rares mâles survivants, non distinguables des souris contrôles à la naissance, développent rapidement un phénotype extrêmement sévère avec une espérance de vie courte (au maximum deux mois). A six semaines, les souris mâles iDUX4(2.7) sont plus petits et ont moins de force que les souris mâles contrôles, en raison de muscles plus petits avec moins de fibres musculaires. Malgré ce phénotype musculaire particulier, les souris mâles iDUX4(2.7) ne sont pas clairement dystrophiques. Par ailleurs, *DUX4* est également exprimé dans les testicules et la rétine de ces souris, conduisant à des défauts dans la gamétogenèse et des télangiectasies rétiniennes. Les femelles, quant à elles, ne présentent pas d'atteintes musculaires, ce qui est biaisé par l'inactivation aléatoire du chromosome X, dont potentiellement le chromosome X porteur du transgène. Par ailleurs, l'induction de l'expression de *DUX4* par la doxycycline dans des cellules en culture prélevées chez des souris mâles iDUX4(2.7) est délétère. Par conséquent, la transplantation de cellules satellites iDUX4(2.7) dans une souris dont le muscle *tibialis anterior* a été irradié, ne permet pas de régénérer un nouveau muscle en présence de doxycycline (Dandapat et al., 2014). Ce modèle murin iDUX4(2.7) n'est donc pas idéal pour le test de thérapies géniques du fait du phénotype extrêmement sévère et de la courte durée de vie des souris, mais des cellules dérivant de ces souris peuvent être utilisées dans l'étude de la fonction de *DUX4*.

#### 1.4.2.3.2.1. Souris D4Z4(2.5) et (12.5)

Les modèles murins transgéniques pathologique D4Z4(2.5) et contrôle D4Z4(12.5) sont générés par introduction dans le génome murin d'un large segment du génome humain contenant respectivement 2,5 ou 12,5 répétitions D4Z4 associées à l'haplotype permissif 4qA (Krom et al., 2013). Des transcrits *DUX4* sont détectés en abondance dans les testicules des

souris D4Z4(2.5) et des souris D4Z4(12.5), en accord avec ce qui est observé quant à l'expression de *DUX4* dans la lignée germinale humaine (Snider et al., 2010). Cependant, *DUX4* est exprimé uniquement chez la souris D4Z4(2.5) mais faiblement et de manière extrêmement variable dans l'ensemble de l'organisme. Une expression un peu plus importante de *DUX4* dans les muscles communément affectés chez les patients FSHD, c'est-à-dire les muscles des membres, du tronc et de la tête, peut-être observée. *DUX4* est également détecté dans le cervelet, les yeux, la rate ou le rein de la souris D4Z4(2.5). Au contraire, les souris D4Z4(12.5) expriment *DUX4* que dans le *tibialis anterior* et les muscles du groupe pectoral. D'un point de vue épigénétique, les répétitions D4Z4 chez les souris D4Z4(12.5) sont hyperméthylées aux sites CpG, tandis que les répétitions D4Z4 chez les souris D4Z4(2.5) sont hypométhylées aux sites CpG dans l'ensemble de l'organisme. Ces modèles murins récapitulent le profil épigénétique des individus FSHD comparé à celui des individus contrôles (Krom et al., 2013).

La souris D4Z4(2.5) ne présente aucun signe de faiblesse, aucune dégénérescence musculaire et aucune anomalie morphologique ou histologique. Le seul phénotype parfois observé chez ces souris D4Z4(2.5) est une kératite progressive qui peut aller jusqu'à la cécité et qui apparaît à l'âge de 8-12 semaines (Krom et al., 2013). Ce phénotype n'est pas retrouvé chez des patients FSHD. Bien que ces modèles murins permettent, dans la majorité des cas, de récapituler les marqueurs génétiques et épigénétiques de la FSHD, l'absence d'un phénotype musculaire ou dystrophique ne fait pas de ce modèle un modèle idéal pour des tests thérapeutiques nécessitant des mesures de la force du muscle.

#### **1.4.2.3.2.1. Souris FLExDUX4**

Depuis peu, un nouvel modèle murin, développé par l'équipe de Peter Jones, est disponible. Cette souris transgénique nommée FLExDUX4, se caractérise par la présence en aval du promoteur ROSA26, de la séquence en sens inverse de *DUX4* (comprenant les exons 1, 2 et 3 plus quelques bases ainsi que les intron 1 et 2). La séquence de *DUX4* est encadrée de sites de recombinaison incompatibles *loxP* et *lox511* (Srinivas et al., 2001; Zambrowicz et al., 1997). Le gène *DUX4* intégré contient quatre mutations dans le site d'épissage en 5' (ou site donneur) cryptique de l'exon 1 pour la forme *DUX4-s*. De ce fait, les souris ne peuvent exprimer que les isoformes *DUX4-fl*. L'induction par le tamoxifène de l'expression de la recombinaise Cre chez ces souris, au préalable croisées avec des souris Tg(ACTA1-

cre/Esr1\*)2Kesr/J, permet l'inversion de la séquence *DUX4* et ainsi l'expression du gène. A noter qu'en l'absence d'une expression de la Cre, les souris hémizygotés et homozygotés expriment tout de même faiblement *DUX4-fl*. La présence de la recombinaise Cre permet d'augmenter cette expression. Il est donc possible d'induire une expression durable et contrôlée de *DUX4* chez des souris à différents âges, permettant ainsi de s'affranchir de la létalité importante caractérisant certains des autres modèles murins.

Aucun article publié établit à l'heure actuelle le phénotype de ces souris. Néanmoins, une description succincte de ce modèle murin est présente sur le site *The Jackson Laboratory*, qui les commercialise. Les souris FLExDUX4 sont viables, fertiles, avec une espérance de vie normale et des portées nombreuses. Elles se caractérisent phénotypiquement par une alopecie, des selles molles, de l'inflammation et une faiblesse musculaire. Les homozygotés sont plus affectés que les hémizygotés, chez les mâles comme chez les femelles.

Un autre modèle murin, similaire en bien des points, a été développé par l'équipe de Scott Haper. Cette souris, non commercialisée à ce jour, se caractérise par la présence en aval du promoteur ROSA26 du transgène *DUX4* (comprenant le signal de polyadénylation localisé dans l'exon 3) étiqueté avec l'épitope V5. Un gène de résistance à la néomycine encadré de sites de recombinaison *lox*, est situé entre le promoteur ROSA26 et le transgène *DUX4*, empêchant une expression de *DUX4*. L'induction par le tamoxifène de l'expression de la recombinaise Cre chez ces souris, au préalable croisées avec des souris HSA promoter-CRE-ERT2, permet de retirer le gène de résistance à la néomycine et d'induire l'expression du transgène. En l'absence d'une administration de tamoxifène, ces souris transgéniques sont viables et n'expriment pas *DUX4*. En fonction de la dose de tamoxifène injectée, le phénotype de ces souris est variable : pour de fortes doses de tamoxifène administrées, les souris présentent dans les quelques jours suivant l'injection une incapacité à bouger ; pour de faibles doses de tamoxifène administrées, les souris présentent des difficultés à se déplacer, une diminution de la force du *tibialis anterior* et une réduction de l'activité globale. Aucune information complémentaire sur ces souris n'est encore disponible.

En résumé, diverses techniques ont été utilisées pour générer des modèles murins FSHD exprimant *DUX4*. Chaque modèle possède ces propres avantages et permet d'apporter des réponses quant aux rôles cellulaires de *DUX4* dans le développement de la FSHD. Cependant, ces modèles ont de nombreux inconvénients. Dans les modèles transgéniques murins

présentés, l'insertion du transgène s'effectue dans une autre région que la région subtélomérique du chromosome 4q. De plus, certains de ces transgènes *DUX4* sont exprimés de manière constitutive et à des doses non physiologiques. Par ailleurs, le gène *DUX4* n'est pas présent initialement chez ces êtres vivants, tout comme potentiellement son réseau de régulation. Il en résulte une perte du contexte génétique et épigénétique naturel de régulation de l'expression de *DUX4*. Un autre point est qu'aucun de ces modèles ne récapitule simultanément tous les symptômes de la FSHD. Ainsi, l'étude de ces modèles murins peut apporter des informations sur les mécanismes médiés par *DUX4* à l'origine du développement de la FSHD, mais ces informations peuvent être biaisées.

Enfin, ces modèles murins (notamment les modèles AAV6-*DUX4*, i*DUX4*(2.7) et D4Z4(2,5)) ne permettent pas une étude de l'efficacité de stratégies thérapeutiques sur un long terme. A noter également que l'utilisation de modèles murins AAV-*DUX4* est à exclure dans les essais de thérapie génique basées sur l'utilisation de virus AAV, en raison des risques de réponses immunes contre le vecteur viral suite à une première exposition à un virus AAV lors de l'injection de l'AAV-*DUX4*. Le nouveau modèle murin FLE*x**DUX4* permettra peut-être de répondre aux exigences techniques du développement d'une stratégie thérapeutique.

#### **1.4.2.3.3. Modèle murin humanisé**

Une autre alternative pour étudier *in vivo* le rôle de *DUX4* dans le développement de la FSHD, consiste à venir greffer en place et lieu du *tibialis anterior*, dans la loge antérieure de la patte arrière de souris immunodéficientes, des biopsies de patients FSHD. Rapidement, les muscles humains greffés montrent une vascularisation, une innervation et survivent au moins 40 jours après transplantation (Zhang et al., 2014). De plus, les greffons sont capables de se contracter de manière calcium-dépendante. Le profil d'expression des gènes du tissu greffé est identique au profil des biopsies de patients FSHD. L'avantage de ce modèle est que le contexte génétique et épigénétique des gènes impliqués dans le développement de la FSHD n'est pas modifié. Par ailleurs, les cellules greffées peuvent provenir de patients FSHD1 ou FSHD2. Le désavantage de ce modèle est qu'il a, pour l'instant, une durée de vie limitée et ne peut être étendu à tout un organisme. De plus, la mise en place de ces xéngreffes nécessite des biopsies musculaires, qui ne sont pas toujours disponibles en grande quantité. Malgré tout, ce modèle représente un outil très intéressant pour l'étude *in vivo* de la FSHD, puisqu'il

permet de conserver, sans aucune exception, tous les mécanismes génétiques et épigénétiques associés à la FSHD.

De nombreux modèles cellulaires et animaux dits FSHD sont donc à la disposition de la recherche pour comprendre les mécanismes conduisant à l'atrophie sélective asymétrique de certains groupes musculaires dans la FSHD mais aussi pour développer des stratégies thérapeutiques curatives. Chaque modèle possède ces avantages et ces inconvénients. Le choix du modèle dépend alors de la question scientifique posée.

## **1.5. Stratégies thérapeutiques**

### **1.5.1. Thérapie génique**

La dystrophie musculaire de Duchenne (DMD), la dystrophie myotonique, la dystrophie musculaire oculo-pharyngée (OPMD) et de nombreuses autres maladies forment un groupe hétérogène de pathologies musculaires monogéniques, c'est-à-dire causées par une ou plusieurs mutations dans un seul des nombreux gènes qui composent notre génome (Emery, 2002). Ces pathologies sont qualifiées de rares car leur prévalence est inférieure à un seuil de cinq personnes touchées sur 10 000 (selon le règlement (CE) N° 141/200 du Parlement européen et du Conseil du 16 décembre 1999). En conséquence, les centres de recherche et les compagnies pharmaceutiques ont délaissé pendant longtemps ces maladies pour lesquelles le développement d'un traitement est peu rentable. De ce fait, ces pathologies sont qualifiées de pathologies orphelines en raison de l'absence de traitements curatifs. Les patients ont uniquement accès à un traitement qui réduit les symptômes ou ralentit la progression de la maladie. Cela permet d'améliorer leur qualité de vie ainsi que leur longévité. Aujourd'hui, ces pathologies rares font l'objet de nombreux programmes de thérapies géniques, souvent soutenus par des associations de patients ou des organisations académiques.

Le concept de thérapie génique date des années 1970, lorsque Stanfield Roger, puis Theodore Friedmann et Richard Roblin évoquent la possibilité d'utiliser de l'ADN exogène pour remplacer un ADN défectueux chez des individus atteints de pathologies génétiques. Le but est de traiter la cause de la pathologie plutôt que les symptômes. Aujourd'hui, la thérapie génique se définit comme le transfert d'une substance active qui contient ou consiste en de



l'acide nucléique recombinant exogène, en vue de réguler, corriger, remplacer, réparer, supprimer ou ajouter un gène ou une séquence génique dans une cellule, un tissu ou chez un individu, et ce uniquement à des fins thérapeutiques (Wirth et al., 2013). Il existe deux types de thérapie génique : la thérapie génique germinale et la thérapie génique somatique. La différence entre ces deux types de thérapie repose sur les cellules ciblées. Dans le cas de la thérapie génique somatique, le matériel exogène est transféré uniquement dans des cellules somatiques, sans transmission à la descendance. Au contraire, la thérapie génique germinale est appliquée aux premiers stades embryonnaires (lorsque les cellules ne sont pas différenciées) ou aux cellules germinales adultes, ce qui implique une transmission à la génération suivante de la modification génétique (Wirth et al., 2013). Cette différence est importante puisqu'en raison de problèmes éthiques majeurs, et de crainte de dérives eugéniques, la législation dans de nombreux pays ne permet pas une thérapie génique germinale (UNESCO – *Report on Human Gene Therapy* 1994). Cette interdiction est reportée en France dans la loi bioéthique n° 94-653 du 29 juillet 1994 relative au respect du corps humain, Chapitre 2, Article 16-4 « *Sans préjudice des recherches tendant à la prévention et au traitement des maladies génétiques, aucune transformation ne peut être apportée aux caractères génétiques dans le but de modifier la descendance de la personne.* ».

Deux stratégies de thérapie génique somatique existent : *ex vivo* ou *in vivo* (Crystal, 1999). Schématiquement, dans la thérapie génique *ex vivo*, la première étape consiste à prélever des cellules de patients pour les mettre en culture. Puis, l'acide nucléique thérapeutique est transféré *ex vivo* dans ces cellules maintenues en culture. Seules les cellules qui ont été « corrigées » lors du transfert de l'acide nucléique thérapeutique sont sélectionnées et réimplantées chez le patient (Crystal, 1999; Herrero et al., 2012; Suhonen et al., 2006). Dans la thérapie génique *in vivo*, le transfert de l'acide nucléique thérapeutique est direct, par injection systémique ou locale au niveau du tissu ou de l'organe cible (Crystal, 1999; Suhonen et al., 2006).

La thérapie génique a été appliquée pour la première fois à l'Homme en 1990 par l'équipe de French Anderson aux Etats-Unis sur deux enfants atteints d'un déficit en adénosine désaminase (ADA-SCID), une pathologie monogénique conduisant à une immunodéficience sévère (Bordignon et al., 1995; Ginn et al., 2013; Wirth et al., 2013). Pour cette thérapie, la technique employée est une technique de transfert *ex vivo*, sur des cellules lymphocytaires, du gène manquant qui est l'*adenosine deaminase* (Bordignon et al., 1995). Une des patientes a

montré une réponse temporaire à la thérapie génique tandis que chez la deuxième patiente, les effets ont été moindres (Blaese et al., 1995; Wirth et al., 2013). En Europe, le premier essai clinique de thérapie génique a été mené en 1992, sur des patients également atteints de ADA-SCID (Bordignon et al., 1995; Wirth et al., 2013). Ces essais cliniques, qui n'ont pas donné les résultats escomptés, ont conduit en 2000 à un premier succès d'envergure pour une thérapie génique chez des patients atteints d'un déficit immunitaire combiné sévère lié à l'X (X1-SCID) (Cavazzana-Calvo et al., 2000; Hacein-Bey-Abina et al., 2002). A ce jour, plus de 2 000 essais cliniques de thérapie génique sont en cours dans différents domaines allant de la cancérologie aux maladies cardiovasculaires en passant par les maladies infectieuses ou encore les maladies monogéniques (Ginn et al., 2013; Wirth et al., 2013). Ces essais cliniques sont majoritairement des études de phase I/II, qui évaluent la sécurité et l'efficacité du traitement. Les essais de phase III, qui permettent d'établir les bénéfices et les risques du nouveau traitement comparé à un traitement déjà sur le marché ou un placebo, sont encore peu nombreux (<http://www.abedia.com/wiley/indications.php> et <https://clinicaltrials.gov/>) (Figure 31).

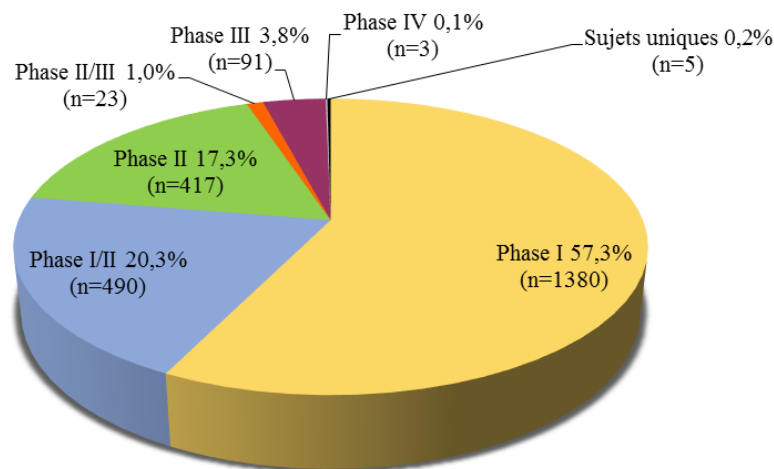


FIGURE 31. Répartition des essais cliniques de thérapie génique en cours (Avril 2017) par études de phase. Les études de phase I/II évaluent la sécurité et l'efficacité du traitement. Les essais de phase III permettent d'établir les bénéfices et les risques du nouveau traitement comparé à un traitement déjà sur le marché ou un placebo. Les études de phase IV sont des suivis sur le long terme du traitement quand celui-ci est autorisé sur le marché (adapté de <http://www.abedia.com/wiley/indications.php>).

Rares sont les médicaments de thérapie génique qui ont passé les différents obstacles du développement clinique. Le premier pays à avoir approuvé la mise sur le marché d'un médicament de thérapie génique est la Chine en 2003 avec la Gendicine (Peng, 2005; Wilson, 2005). La Gendicine est un vecteur adénoviral portant l'ADN complémentaire (ADNc) du gène humain p53, un suppresseur de tumeur. Cette molécule thérapeutique est utilisée pour le

traitement des carcinomes de la tête et du cou. Bien qu'aucun essai clinique de phase III n'ait été mis en place pour le test de la Gendicine, plus de 10 000 personnes ont reçu à ce jour ce médicament sans effets secondaires notables. En Europe, le premier médicament de thérapie génique, le Glybera, a été approuvé en 2012 (Wirth et al., 2013). Le Glybera est un vecteur de type AAV, conçu pour exprimer la lipoprotéine lipase dans les tissus musculaires chez des patients déficients pour cette enzyme. De manière intéressante, le Glybera a reçu trois fois des recommandations négatives avant d'obtenir une autorisation de mise sur le marché en Europe (Miller, 2012; Wirth et al., 2013; Ylä-Herttuala, 2012). Le développement d'un médicament de thérapie génique est un processus extrêmement long, très coûteux et rarement un succès. Cependant l'évolution des technologies en biologie offre à ce jour de nouvelles possibilités tant dans le développement de nouvelles stratégies thérapeutiques que dans l'élaboration de médicaments efficaces et sans risque pour les patients.

### **1.5.1.1. Diversité des stratégies de thérapie génique**

#### **1.5.1.1.1. Complémentation et compensation**

Les premières approches de thérapie génique consistent :

- à apporter une copie fonctionnelle du gène muté (complémentation) sans venir pour autant modifier sa séquence ou son expression ;
- à apporter une copie d'un gène différent mais qui, de par son expression, peut atténuer les effets du gène muté (compensation).

Dans les deux cas, le gène muté n'est pas corrigé.

La thérapie par complémentation est une stratégie réalisable dans le cas de maladies récessives à perte de fonction. Cependant, cette technique ne permet pas de contrôler l'expression du transgène. Par ailleurs, la complémentation par un gène de grande taille, souvent administré sous la forme d'ADNc, est difficilement réalisable. Pour des gènes extrêmement longs, il est envisageable de raccourcir la séquence codante du gène sans pour autant modifier la fonction de la protéine. Cette technique a été utilisée dans le cadre de la DMD. La DMD est une pathologie génétique sévère du muscle qui affecte les garçons avec une prévalence de 1 à 9 naissances sur 100 000 (Orphanet, 2017). Le défaut génétique qui conduit à la DMD est une mutation dans le gène codant pour la dystrophine, une protéine qui connecte les filaments d'actine au sarcolemme (Koenig et al., 1989; Monaco et al., 1988).

Cette mutation abroge l'expression de la dystrophine. Une stratégie thérapeutique consiste donc à apporter un gène codant pour la dystrophine. Cependant, l'ADNc de la dystrophine étant de grande taille (14 kb), un ADNc réduit de la dystrophine est utilisé, conduisant à la production d'une micro-dystrophine fonctionnelle qui allège le phénotype DMD (Rodino-Klapac et al., 2010, 2013; Wang et al., 2000). Des essais cliniques basés sur ce principe sont actuellement en cours.

La thérapie par complémentation est plus difficile à mettre en place. En effet, il faut déterminer dans un premier temps les gènes pouvant compenser indirectement le défaut génétique dans la pathologie. Généralement, cette stratégie intervient dans les mécanismes moléculaires en aval de la déficience, ce qui permet une application à toute une catégorie de pathologies aux symptômes similaires. Des stratégies de complémentation pour la DMD ont été développées avec l'apport d'un gène codant pour la follistatine, un inhibiteur de la myostatine ou l'apport d'un gène codant l'utrophine (Bish et al., 2011; Gillis, 2002; Haidet et al., 2008; Mendell et al., 2015; Qiao et al., 2008; Tinsley et al., 1998).

Les thérapies par complémentation et compensation peuvent parfois être combinées (Rodino-Klapac et al., 2013).

#### **1.5.1.1.2. Manipulation du génome par les nucléases**

Pendant de nombreuses années, les scientifiques ont recherché des moyens de modifier ou de manipuler directement la séquence ADN des gènes chez les plantes et les animaux. L'édition du génome par des nucléases qui reconnaissent des séquences spécifiques d'ADN est devenue une réalité et semble être une stratégie prometteuse pour le traitement de maladies génétiques dont les pathologies musculaires (Bengtsson et al., 2017; Carroll, 2008; Long et al., 2014, 2016; Nelson et al., 2016; Urnov et al., 2005).

##### **1.5.1.1.2.1. Méganucléases**

Les premiers outils qui ont permis de répondre à cette demande sont les méganucléases (Silva et al., 2011; Stoddard, 2005). Les méganucléases sont des endonucléases qui reconnaissent de larges séquences spécifiques d'ADN (>14 bp) au niveau desquelles elles catalysent des coupures double-brin (Chevalier and Stoddard, 2001; Thierry and Dujon, 1992). Des

centaines de méganucléases naturelles existent. Cependant, et ce malgré le développement en parallèle de techniques de production de méganucléases artificielles, la probabilité de trouver une méganucléase reconnaissant le site d'intérêt à cibler est faible (Arnould et al., 2006; Ashworth et al., 2006; Chevalier and Stoddard, 2001; Smith et al., 2006). Ainsi, les méganucléases sont peu ou pas utilisées en thérapie génique.

#### **1.5.1.1.2.2. Nucléases à doigt de zinc**

La stratégie d'édition du génome a pris un nouveau tournant il y a quelques années avec la découverte des domaines de liaisons spécifiques à l'ADN de type doigt de zinc (ZF). Les ZF sont retrouvés dans de nombreuses protéines chez les eucaryotes. Un module ZF est composé d'environ 30 acides aminés organisés en une hélice  $\alpha$  et un feuillet  $\beta$ . Le feuillet  $\beta$  se compose de deux brins antiparallèles stabilisés par des interactions hydrophobes et par une liaison, *via* une chaîne de deux résidus cystéines et de deux résidus histidines, à un atome de zinc. Les acides aminés à la surface de l'hélice  $\alpha$  sont en contact avec les bases du grand sillon de l'ADN et reconnaissent une séquence spécifique de trois paires de bases (Dreier et al., 2001; Lee et al., 1989; Miller et al., 1985; Pavletich and Pabo, 1991; Wolfe et al., 2000). L'assemblage de plusieurs modules ZF permet une fixation spécifique à des séquences d'ADN de tailles variables (Beerli et al., 1998, 2000; Desjarlais and Berg, 1993; Kim and Pabo, 1998; Liu et al., 1997; Wolfe et al., 1999). La fusion d'une combinaison appropriée de modules ZF et du domaine catalytique non spécifique de la nucléase *FokI* (FN) a permis de créer des nucléases à doigt de zinc (ZFN *Zinc Finger Nuclease*) capables de couper l'ADN en une région spécifique (Kim et al., 1996; Sugisaki and Kanazawa, 1981; Wah et al., 1997) (Figure 32).

Cependant la nucléase *FokI* se doit d'être sous la forme d'un dimère pour cliver l'ADN (Smith et al., 2000; Vanamee et al., 2001) (Figure 32). Par conséquent, deux sites spécifiques indépendants, dans une conformation queue à queue, séparés de cinq à sept bases, doivent être reconnus par les ZFN, pour favoriser le rapprochement de deux unités *FokI* (Bibikova et al., 2001) (Figure 32). La dimérisation de *FokI* conduit alors à la formation d'une coupure double-brin au niveau de la région d'intérêt (Bibikova et al., 2001; Kim et al., 1996; Vanamee et al., 2001) (Figure 32).

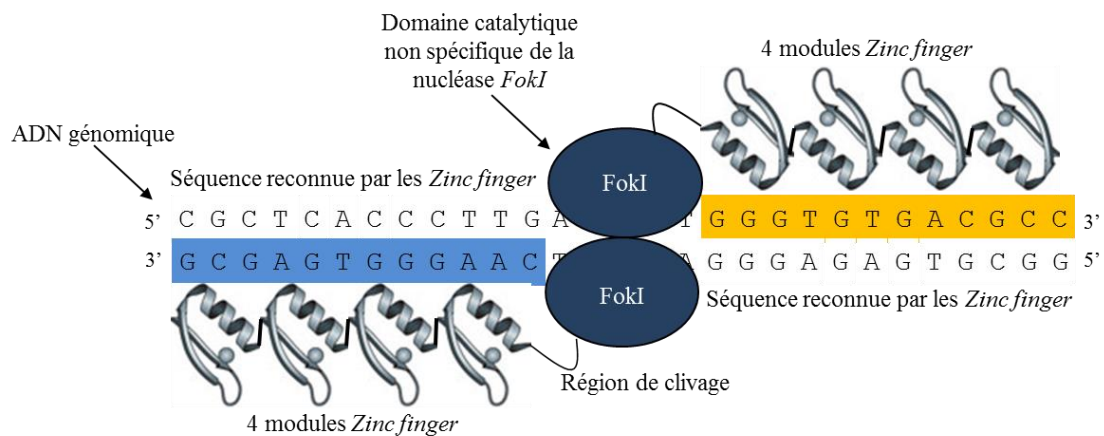


FIGURE 32. Structure et mécanisme d'action des nucléases à doigt de zinc. Un module ZF est composé d'une hélice  $\alpha$  et d'un feuillet  $\beta$  de deux brins antiparallèles. Les acides aminés à la surface de l'hélice  $\alpha$  reconnaissent une séquence spécifique d'ADN de trois paires de bases. L'association de plusieurs modules ZF permet d'augmenter la taille de la séquence nucléotidique reconnue (ici 12 nucléotides en bleu ou orange sont reconnus par une association de quatre modules ZF). Les modules ZF sont fusionnés au domaine catalytique non spécifique de la nucléase *FokI*, formant une nucléase à doigt de zinc (ZFN). La nucléase *FokI* ne présentant une activité nucléase que sous la forme d'un dimère, deux sites indépendants, séparés par 5 à 7 bases, sont nécessairement reconnus par deux ZFN pour permettre une dimérisation de *FokI* et la formation d'une coupure double-brin au site ciblé (adapté de Urnov et al., 2010).

#### 1.5.1.1.2.3. *Transcription Activator-Like Effector Nucleases (TALENs)*

Récemment, de nouvelles classes de nucléases séquences spécifiques ont émergé. Parmi elles, les TALENs (*Transcription Activator-Like Effector Nucleases*) (Joung and Sander, 2013; Sanjana et al., 2012). Les TALENs sont composées en partie de l'effecteur TAL (*Transcription Activator-Like*) qui se fixe à des séquences spécifiques de l'ADN (Boch et al., 2009; Moscou and Bogdanove, 2009). La protéine TAL, issue d'une bactérie phytopathogène du genre *Xanthomonas*, contient en son domaine central des répétitions (7 à 34) en tandem d'un motif de 33 à 35 acides aminés (Figure 33). Chaque répétition permet de reconnaître un nucléotide : au niveau de chacune de ces répétitions, deux acides aminés variables en position 12 et 13, qui forment le domaine RVD (*Repeat-Variable Di-residues*), confèrent la spécificité de fixation à l'ADN (Figure 33). Les nucléotides C, T, A et G sont reconnus respectivement par les paires d'acides aminés 12/13 HD, NG, NI et NN (avec ici H : Histidine, D : Acide aspartique, G : Glycine, I : Isoleucine, N : Asparagine) (Boch et al., 2009; Deng et al., 2012; Mak et al., 2012; Moscou and Bogdanove, 2009) (Figure 33). Le nombre de répétitions en tandem des motifs de 33 à 35 acides aminés (idéalement entre 15 et 20), le type de résidus en position 12/13 dans chaque motif et l'ordre des répétitions des motifs déterminent la

spécificité de reconnaissance de la séquence ADN cible (Bogdanove and Voytas, 2011; Miller et al., 2011; Mussolino et al., 2011). De manière identique aux ZFN, l'activité nucléase est apporté par le domaine catalytique non spécifique à activité nucléase *FokI* (Christian et al., 2010; Li et al., 2011b; Mahfouz et al., 2011; Miller et al., 2011). La fusion d'un TAL et du domaine catalytique non spécifique de la nucléase *FokI* conduit à la production d'un TALEN (Figure 33). Les TALENs fonctionnent également sous la forme d'un dimère. La fixation d'un couple de TALENs au niveau de deux séquences cibles d'ADN, dans une conformation queue à queue et séparées de 10 à 30 bases, permet une dimérisation des sous unités *FokI* et la catalyse d'une cassure double-brin au site d'intérêt (Christian et al., 2010) (Figure 33).

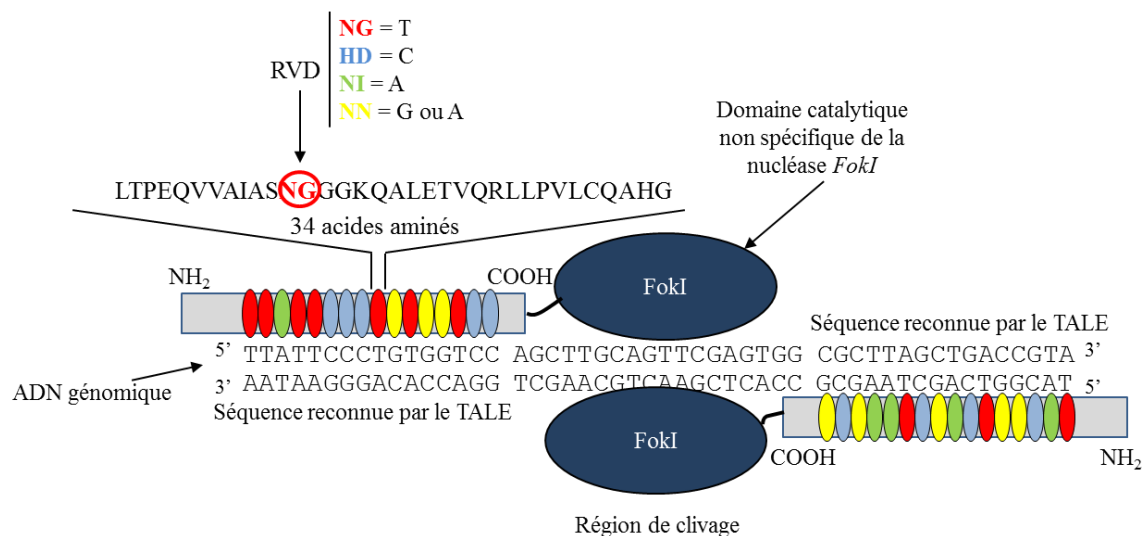


FIGURE 33. Structure et mécanisme d'action des TALENs. Les TALENs sont composés pour partie de l'effecteur TAL, qui permet une fixation spécifique à une séquence d'ADN. Le domaine central de la protéine TAL est composé de répétitions (7 à 34) en tandem d'un motif de 33 à 35 acides aminés. Chacune de ces répétitions permet de reconnaître spécifiquement un nucléotide par l'intermédiaire du domaine RDV. Le domaine RVD se caractérise par la présence de deux acides aminés variables en position 12 et 13 (entouré en rouge ici). Les nucléotides C, T, A et G sont reconnus respectivement par les paires d'acides aminées 12/13 HD, NG, NI et NN (avec H : Histidine, D : Acide aspartique, G : Glycine, I : Isoleucine, N : Asparagine). Les effecteurs TALs sont fusionnés au domaine catalytique non spécifique de la nucléase *FokI*, formant un TALEN capable de générer une cassure double-brin d'ADN. La nucléase *FokI* ne présentant une activité nucléase que sous la forme d'un dimère, deux sites indépendants, séparés par 10 à 30 bases, sont nécessairement reconnus par deux TALENs pour permettre une dimérisation de *FokI* et la formation d'une coupure double-brin au site ciblé (adapté de Sanjana et al., 2012).

#### 1.5.1.1.2.4. RNA Guided Nucleases (RGNs)

Plus récemment, les RGNs (*RNA-Guided Nucleases*) ont été décrites comme des outils de thérapie génique (Bassett et al., 2013; Gratz et al., 2013; Hsu et al., 2014; Yu et al., 2013). Les RGNs sont basées sur le système CRISPR/Cas9 (*Clustered Regularly Interspaced Short Palindromic Repeats / CRISPR associated 9*) identifié pour la première fois en 1987 chez la

bactérie *Escherichia coli*. Ce système a été démontré, plus tard, comme étant à la base du système de défense des bactéries et des archées contre l'invasion d'acides nucléiques étrangers, dont ceux des virus (Barrangou et al., 2007; Gasiunas et al., 2012; Ishino et al., 1987; Jinek et al., 2012; Wiedenheft et al., 2012). Dans le système CRISPR/Cas9 de type II, le locus *CRISPR* code pour un long ARN composé d'une série de séquences uniques et répétées. Ce long ARN est ensuite transformé en petits ARNs CRISPR (*crRNA*). Chaque *crRNA* consiste en des séquences répétées associées à un court fragment ARN nommé « *spacer* » et complémentaire d'une séquence ADN double-brin exogène. De nouvelles séquences « *spacer* » peuvent être créées par l'intégration de fragments d'ADN exogènes dans les loci *CRISPR* (Amitai and Sorek, 2016). Dans un second temps, le *crRNA* s'apparie par ses séquences répétées à un autre ARN, le *tracrRNA* (*Trans-acting crRNA*). Le *tracrRNA* permet, quant à lui, une interaction avec l'endonucléase Cas9 (Figure 34).

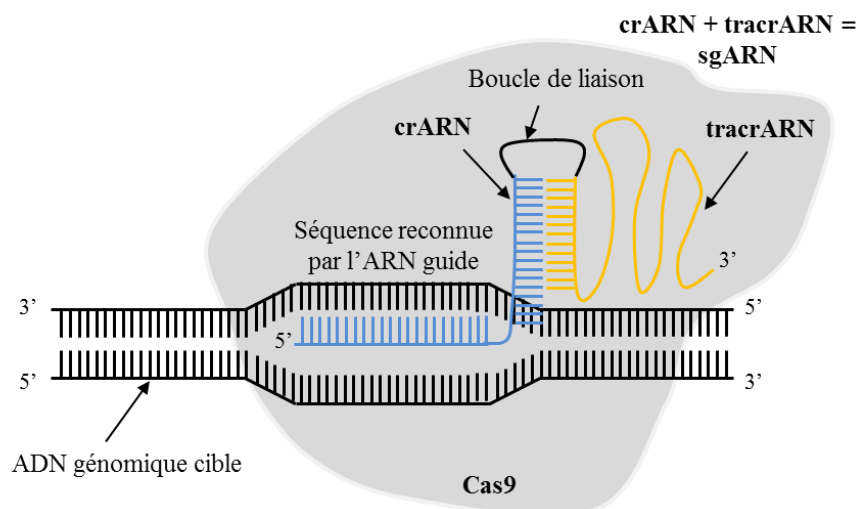


FIGURE 34. Structure et mécanisme d'action du système CRISPR/Cas9. Un long ARN composé d'une série de séquences uniques et répétées transcrit à partir du locus CRISPR est transformé en petits ARNs CRISPR (*crRNA* en bleu). Ce *crRNA* se décompose en une région de séquences répétées et un court fragment d'ADN (« *spacer* ») complémentaire d'une séquence d'ADN double-brin. Cette région de séquences répétées permet une interaction avec un autre ARN, le *trans-acting crRNA* (*tracrRNA* en jaune), qui permet, lui, une interaction avec l'endonucléase Cas9. Ce complexe *crRNA/tracrRNA/Cas9* permet de générer une cassure double-brin au niveau d'une séquence ADN reconnue par le *crRNA* (adapté de Agrotis and Ketteler, 2015).

Le complexe *crRNA/tracrRNA/Cas9* permet de cliver spécifiquement un ADN double-brin exogène complémentaire de la région *spacer* du *crRNA* (Jinek et al., 2012) (Figure 34). Les fonctions du *crRNA* et du *tracrRNA* peuvent être combinées dans un seul ARN chimérique appelé *sgRNA* (*single guide RNA*) (Jinek et al., 2012) (Figure 34). En se basant sur ces données et en modifiant une séquence d'une vingtaine de paires de bases dans le *sgRNA* (au



niveau du crARN), le système CRISPR/Cas9 est utilisé pour cibler des séquences spécifiques du génome eucaryote (Chang et al., 2013; Hwang et al., 2013; Shen et al., 2013).

### 1.5.1.1.2.5. Modifications du génome lors de la réparation de la cassure double-brin

L'utilisation de l'ensemble de ces nucléases conduit à la génération d'une cassure double-brin en un site spécifique du génome. Cette cassure double-brin est ensuite reconnue et réparée par recombinaison non-homologue (*NHEJ Non-Homologous End Joining*) ou par recombinaison homologue (*HR Homologous Recombination*) (Lieber et al., 2003; Sonoda et al., 2006).

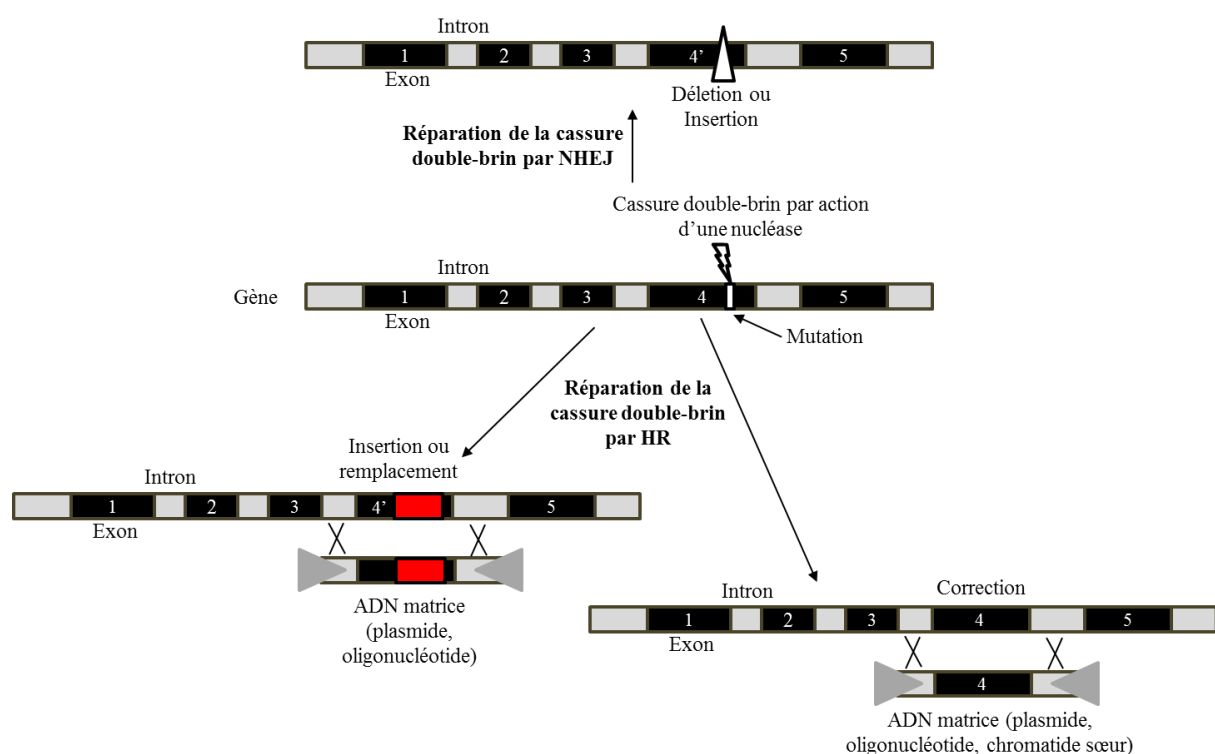


FIGURE 35. Diversité des mécanismes de réparation de la cassure double-brin générée par une action des nucléases. Une cassure double-brin peut être réparée par recombinaison non homologue (*NHEJ Non-Homologous End Joining*) ou recombinaison homologue (*HR Homologous Recombination*). Une réparation par NHEJ conduit à une délétion ou une insertion au site de cassure. Une réparation par HR en présence d'une quantité importante d'ADN matrice (plasmide, oligonucléotides) contenant des séquences homologues aux régions localisées de part et d'autre du site de cassure, permet d'insérer une séquence d'ADN ou de remplacer une séquence d'ADN par une autre ou d'effectuer une correction de la mutation (adapté de Silva et al., 2011).

Une réparation d'une cassure double-brin par le mécanisme de NHEJ se caractérise généralement par des insertions ou des délétions de différentes tailles (Cathomen and Joung, 2008; Lieber et al., 2003; Xiao et al., 2013). Dans un gène, des décalages du cadre de lecture sont souvent observés, entraînant, s'ils affectent un exon, la production d'une protéine

tronquée ou non fonctionnelle (*K.O.*) (Santiago et al., 2008) (Figure 35). La réparation de la cassure double-brin par HR peut avoir lieu en présence d'une grande quantité d'un fragment exogène d'ADN. Pour cela, le fragment exogène d'ADN doit contenir des séquences homologues aux régions localisées de part et d'autre de la cassure double-brin générée par les nucléases (Figure 35). Ainsi, ce fragment exogène d'ADN servira de matrice lors de la recombinaison. Il peut être apporté par un plasmide donneur ou un oligonucléotide simple-brin de petite taille (Bedell et al., 2012; Bonner and Kmiec, 2009; Urnov et al., 2005). Ce mécanisme permet de corriger une mutation présente dans une séquence (*K.I. Knock In*) ou d'ajouter une séquence dans un fragment d'ADN (Cathomen and Joung, 2008) (Figure 35).

Les nucléases peuvent donc être utilisées dans le cadre de maladies génétiques à gain de fonction, en venant éteindre l'expression d'un gène anormalement exprimé, mais aussi dans le cadre de maladies à perte de fonction, en venant corriger une mutation dans un gène déficient. D'une manière avantageuse, ces modifications ne sont pas transitoires mais perdurent dans la cellule modifiée et les cellules filles qui en sont issues. De plus, les nucléases ne modifient pas le contexte d'expression ou de régulation du gène puisque seul l'ADN génomique est ciblé. Cependant, les informations portant notamment sur la toxicité et les *off-targets* (modification de l'expression de gènes non ciblés) générés par une expression à long terme de ces nucléases dans des cellules ou un organisme sont incomplètes (Cornu and Cathomen, 2010; Pattanayak et al., 2011). Néanmoins, les nucléases représentent une technologie idéale pour le traitement de maladies génétiques comme le démontrent des études *in vivo* et quelques rares études cliniques de phase I (Cox et al., 2015; Genovese et al., 2014; Hwang et al., 2013; Li et al., 2011a; Nelson et al., 2016; Tebas et al., 2014).

#### **1.5.1.1.3. Manipulation des transcrits par des oligonucléotides**

Une autre stratégie thérapeutique consiste à modifier quantitativement et/ou qualitativement les transcrits par un mécanisme d'interférence ARN (*ARNi*) via l'utilisation des siARN (*small interference RNA*), des miARN (*micro ARN*) ou plus récemment, des oligonucléotides antisens synthétiques simple-brin (*AOs*). Les siARN, les miARN et les AOs partagent le même principe de fonctionnement : l'oligonucléotide, un court segment d'acide nucléique (ADN ou ARN), s'hybride à sa cible ARN et interfère dans sa maturation, sa stabilité ou sa traduction (DeVos and Miller, 2013; Dias and Stein, 2002; Watts and Corey, 2012).

### 1.5.1.1.3.1. siARN et miARN

Le rôle de petits ARNs dans la régulation de l'expression de gènes a été mis en évidence en 1990 par des travaux chez le pétunia, puis en 1993 chez le nématode *Caenorhabditis elegans* (Lee et al., 1993; Napoli et al., 1990; Wightman et al., 1993). Le mécanisme de fonctionnement de ces petits ARNs a été décrypté en 1998 (Fire et al., 1998).

Les miARN sont des petits ARNs monocaténaire de 20 à 24 nucléotides codés par le génome (Ha and Kim, 2014; Kim et al., 2009b; Krol et al., 2010; Rodriguez et al., 2004). Les séquences codantes pour des miARNs sont localisées à plus de 90% dans des introns. La transcription d'une séquence miARN par la polymérase II conduit à la production d'un long précurseur monocaténaire, le pri-miARN (microARN primaire), qui se replie en partie pour arborer une structure double-brin en épingle à cheveux (Lee et al., 2004) (Figure 36). Ce pri-miARN est reconnu dans le noyau par un complexe composé notamment de DROSHA, une RNase III, et de son co-facteur protéique DGCR8. Ces protéines permettent de cliver le pri-miARN en un pré-miARN (miARN précurseur) de 70 nucléotides présentant une structure en épingle à cheveux et possédant une extrémité simple-brin de deux nucléotides en 3' (Ha and Kim, 2014; Kim et al., 2009b; Krol et al., 2010; Lee et al., 2003; Yeom et al., 2006) (Figure 36). Le pré-miARN est alors exporté hors du noyau par l'exportine 5 (Yi et al., 2003) (Figure 36). Dans le cytoplasme, le pré-miARN est de nouveau clivé, mais cette fois-ci par DICER, une endoribonucléase. Un miARN double-brin, de 20 à 24 nucléotides, sans boucle et avec deux extrémités simple-brin de deux nucléotides en 3', est libéré (Ha and Kim, 2014; Kim et al., 2009b; Krol et al., 2010) (Figure 36). Ce miARN double-brin est pris en charge par la protéine ARGONAUTE, ainsi que d'autres protéines, afin de former le complexe multiprotéique miRISC (*miRNA-Induced Silencing Complex*) (Gan and Gunsalus, 2015; Meister, 2013) (Figure 36). Lorsque le miARN double-brin est associé à RISC, un seul des deux brins est conservé pour donner le miARN mature (Meijer et al., 2014; Okamura et al., 2009) (Figure 36). Le complexe miRISC induit, en fonction de son pourcentage de complémentarité avec la séquence cible, soit une répression de la traduction, soit une dégradation de l'ARNm (Filipowicz et al., 2008) (Figure 36). Un seul miARN peut avoir plus de 100 gènes cibles.

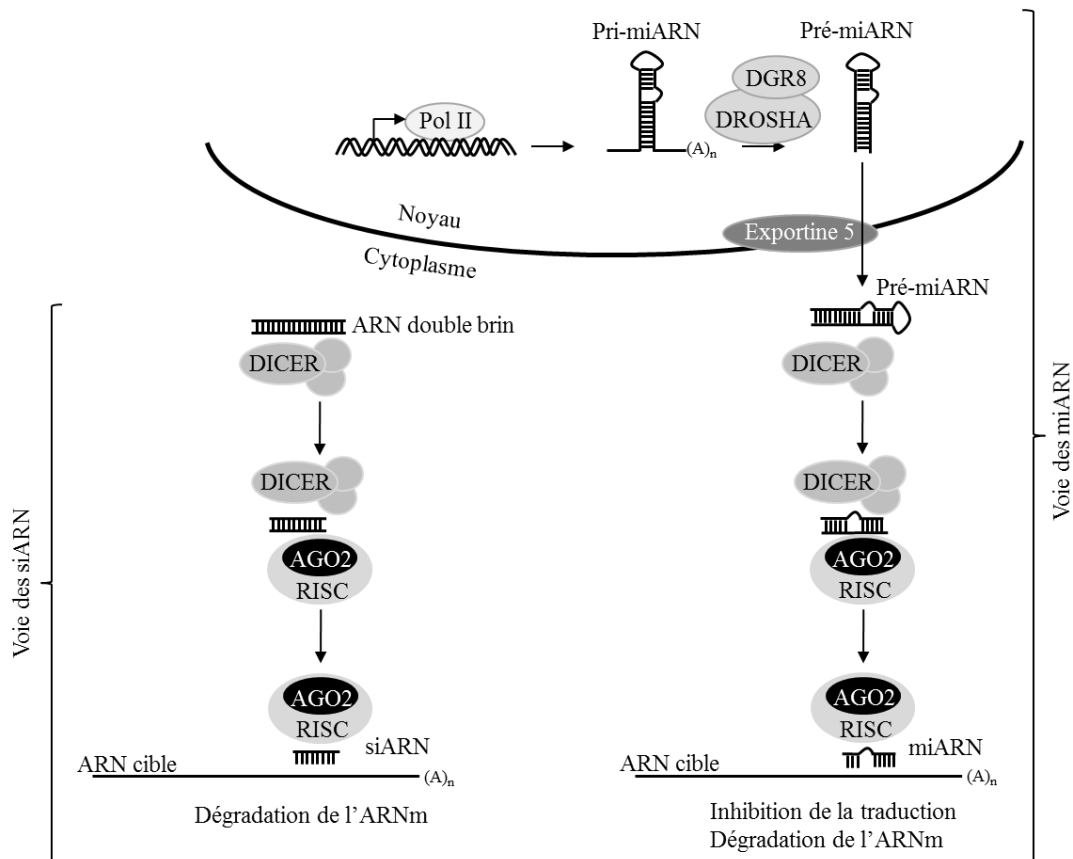


FIGURE 36. Voies d'action des siARNs et des miARNs. Les miARNs sont de petits ARNs monocaténares de 20 à 24 nucléotides codés par le génome. La transcription d'une séquence miARN conduit à la production d'un long précurseur monocaténaire, le pri-miARN (microARN primaire), qui se replie en partie pour arborer une structure double-brin en épingle à cheveux. Ce pri-miARN est reconnu dans le noyau par un complexe composé de DROSHA. DROSHA clive le pri-miARN en un pré-miARN (miARN précurseur) de 70 nucléotides organisés en structure en épingle à cheveux et possédant une extrémité simple brin de deux nucléotides en 3'. Le pré-miARN est ensuite exporté hors du noyau par l'exportine 5 puis pris en charge par DICER. Suite à un clivage par DICER, un miARN double-brin de 20 à 24 nucléotides sans boucle et avec deux extrémités simple-brin de deux nucléotides en 3' est généré. Le miARN double-brin est pris en charge par la protéine ARGONAUTE, et d'autres protéines, afin de former le complexe multiprotéique miRISC où un seul des deux brins du miARN est conservé. Ce complexe miRISC induit, en fonction de son pourcentage de complémentarité avec la séquence cible, soit une répression de la traduction, soit une dégradation de l'ARNm. Les siARNs sont des ARNs monocaténares de 20 à 26 nucléotides, d'origine exogène ou endogène. Le siARN se présente tout d'abord sous la forme d'un ARN double brin (ARNdb) avec parfois une structure en tige boucle. Cet ARNdb est pris en charge par la protéine DICER qui réalise deux coupures enzymatiques en 3' et 5' de l'ARNdb et génère un produit de clivage avec deux extrémités simple-brin de deux nucléotides en 3'. Ce siARN est alors pris en charge par ARGONAUTE, et d'autres protéines, pour former le complexe siRISC, où un seul des deux brins du siARN est conservé. Le complexe siRISC induit le clivage d'une cible ARN complémentaire du siARN (adapté de Farrar et al., 2012).

Les siARN sont quant à eux des ARNs monocaténares de 20 à 26 nucléotides, d'origine exogène ou endogène (ARNm dégradé ou produit par des éléments mobiles en répétitions inversés) (Kim et al., 2009b). Dans un premier temps, le siARN se présente sous la forme d'un ARN double-brin (ARNdb) avec parfois une structure en tige boucle (Figure 36). Cet

ARNdb est pris en charge par la protéine DICER qui réalise alors deux coupures enzymatiques en 3' et 5' de l'ARNdb (Bernstein et al., 2001; Hutvagner and Zamore, 2002) (Figure 36). Le produit du clivage, qui possède deux extrémités simple-brin de deux nucléotides en 3', s'associe ensuite à la protéine ARGONAUTE, ainsi que d'autres protéines, pour former le complexe multiprotéique siRISC (*siRNA Induced Silencing Complex*) (Figure 36). Dans ce complexe siRISC, un seul des deux brins du siARN est conservé, donnant le siARN mature (Noland et al., 2011; Wei et al., 2009). Le complexe siRISC induit le clivage d'une cible ARN complémentaire du siARN (Elbashir et al., 2001) (Figure 36).

En thérapie génique, les siARNs, tout comme les miARNs, sont utilisés pour dégrader ou inhiber la traduction d'un ARNm cible. Un mécanisme artificiel d'interférence ARN est induit par l'utilisation soit d'un siARN/miARN simple-brin ou double-brin synthétisé, soit d'un vecteur ADN portant une séquence transcrite en une autre forme de siARN, le shARN (*short hairpin RNA*) (Cullen, 2005; Rao et al., 2009). Un shARN consiste en deux séquences ARNs de 19 à 22 paires de bases complémentaires liées par une boucle de 5 à 11 nucléotides. Le shARN présente une structure en épingle à cheveux. A la différence du siARN, la séquence du shARN peut être intégrée dans le génome. Une fois le shARN transcrit dans le noyau, il est exporté dans le cytoplasme où il est pris en charge par DICER pour être clivé en un siARN. De la même manière que le siARN synthétique exogène, le shARN mature, ou siARN synthétique endogène, est incorporé dans le complexe RISC et induit la dégradation de l'ARNm dont il est complémentaire (Moore et al., 2010; Rao et al., 2009). Le choix de l'utilisation de l'une ou l'autre des molécules dépend de plusieurs facteurs dont le type cellulaire ciblé (en lien avec l'efficacité de transfection), du temps à disposition et du besoin d'une expression transitoire ou stable de l'ARN interférent.

#### **1.5.1.1.3.2. snARN**

Une autre classe de petits ARNs est utilisée à des fins de thérapie génique. Il s'agit des snARNs (*small nuclear RNA*) ou U-ARNs. Ces petits ARNs ont une localisation et une action exclusivement nucléaire. Leur taille varie autour de 50 à 300 nucléotides. Associés à diverses protéines, ils sont impliqués dans l'épissage des ARNm (U1 à U6 snARN) et dans le recrutement de facteurs de clivage lors de la maturation des ARNs des histones (U7 snARN) (Dominski and Marzluff, 2007; Patel and Bellini, 2008; Schümperli and Pillai, 2004). Les snARNs sont issus de gènes transcrits par l'ARN polymérase II ou III et subissent un

processus de maturation en plusieurs étapes (Dahlberg and Lund, 1987; Patel and Bellini, 2008; Paule and White, 2000). Pour une utilisation en thérapie génique, la séquence du U-ARN peut être partiellement remplacée par la séquence d'un oligonucléotide antisens. Ceci permet alors de cibler des séquences spécifiques de pré-ARNm. A la différence des siARN/miARN, les U-ARN permettent d'intervenir au cours de processus nucléaires précoces de maturation des ARNm (Gorman et al., 1998). Cette technologie, qui ne peut être apportée que par un vecteur ADN, a été utilisée dans le cadre de la DMD. Pour rappel, le défaut génétique qui conduit à la DMD est, dans la majorité des cas, une délétion causant un décalage du cadre de lecture dans le gène codant pour la dystrophine (Koenig et al., 1989; Monaco et al., 1988). Il en résulte une absence d'expression de la dystrophine, ce qui se caractérise par la rupture des fibres musculaires. Dans une forme plus légère de la DMD, la dystrophie musculaire de Becker (BMD), les mutations dans le gène de la dystrophine ne changent pas le cadre de lecture. La dystrophine exprimée chez les patients BMD est tronquée mais reste partiellement fonctionnelle. Une des stratégies thérapeutiques développée pour la DMD est donc de restaurer le cadre de lecture du gène de la dystrophine par saut d'exon afin de produire une protéine tronquée mais fonctionnelle comme celle produite dans la BMD (Muntoni et al., 2003). Pour cela, une partie de la séquence du U7-ARN ou U1-ARN a été remplacée par la séquence d'un oligonucléotide antisens capable de masquer les signaux accepteurs d'épissage de l'exon à exclure pour restaurer le cadre de lecture du gène. L'expression de cet U7-ARN ou U1-ARN modifié, chez des souris DMD, a permis l'expression de la dystrophine, certes tronquée mais fonctionnelle. Les souris DMD traitées présentent un phénotype dystrophique plus léger comparées aux souris DMD non traitées (Brun et al., 2003; Denti et al., 2006; Goyenvalle et al., 2004, 2009, 2012).

#### **1.5.1.1.3.3. Oligonucléotides antisens**

L'un des points faibles des siARN/miARN en thérapie génique est qu'ils sont rapidement dégradés par les nucléases circulantes. Leur durée de vie excède rarement quelques minutes et leur intégration dans les organes ou les tissus cibles est souvent peu efficace (DeVincenzo et al., 2008; Song et al., 2003). Des modifications chimiques de petites séquences (environ 20-30 nucléotides) simple-brin ADN ou ARN ont alors été envisagées pour résoudre ces problèmes (Kole et al., 2012).

Depuis les premiers oligonucléotides antisens (AOs) ADN utilisés en 1978, de nombreuses modifications chimiques au niveau du squelette ribose-phosphate de l'acide nucléique ont été proposées (DeVos and Miller, 2013; Kole et al., 2012; Watts and Corey, 2012; Zamecnik and Stephenson, 1978). Un AO typique, sans aucune modification chimique, est composé d'une succession de nucléotides liés par une liaison phosphodiester. La première catégorie de modifications des AOs porte sur cette liaison phosphodiester. Certains AOs sont construits avec des nucléotides liés par une liaison phosphorothioate (AO PS), c'est-à-dire qu'ici l'un des oxygènes non impliqué dans la liaison est remplacé par un soufre (Eckstein, 2000) (Figure 37). Ces modifications chimiques confèrent une stabilité à l'AO PS et notamment une résistance à la dégradation par les nucléases. Ces modifications autorisent également le recrutement de l'enzyme RNaseH1, nécessaire à l'action de l'AO (Figure 38). En effet, l'enzyme RNaseH1, par reconnaissance des hétéroduplexes ARNm/AO, clive l'ARNm cible. L'AO n'est pas affecté par l'action de la RNaseH1 et est libéré pour une hybridation avec un autre ARNm cible (Cerritelli and Crouch, 2009; Stein et al., 1988). D'autres AOs sont produits par un remplacement du groupement phosphate par un groupement thiophosphoramidate (N3'→P5' phosphoramidate). Cette modification chimique confère également à l'AO une résistance à la dégradation par des nucléases (Figure 37). Cependant, elle abroge l'activité RNaseH1, ce qui peut être utile lorsque l'ARNm cible ne doit pas être dégradé (Pongracz and Gryaznov, 1998).

La deuxième catégorie de modifications chimiques porte sur le sucre (ribose ou désoxyribose) de l'acide nucléique. Les modifications couramment utilisées sont le 2'-O-méthyl (2'-O-Me), le 2'-O-méthoxyéthyl (2'-MOE), le 2'-fluoro (2'-F) et le *locked nucleic acid* (LNA) (Figure 37). Cette dernière modification correspond à un branchement entre les carbones en position 2' et 4' du sucre de l'AO (Figure 37). Ces modifications n'augmentent pas uniquement la résistance de l'AO à des dégradations par les nucléases ainsi que son affinité pour la cible ARNm, mais réduisent également sa fixation non spécifique à des protéines, et donc sa toxicité (excepté pour les LNA) (Teplova et al., 1999; Wengel, 1999). Cependant, la majorité de ces modifications des sucres diminuent le recrutement de la RNaseH1. Une solution consiste alors en la production de *gapmer*, c'est-à-dire d'AOs avec une région centrale ne portant au mieux que des modifications du phosphate, encadrée par deux à cinq résidus modifiés à la fois au groupement phosphate et au groupement ribose (ou désoxyribose). Ces *gapmer* présentent une forte affinité de liaison à leur cible ARNm, une résistance aux

nucléases et une capacité de recrutement de la RNaseH1 (Monia et al., 1993; Wahlestedt et al., 2000).

Les AOs modifiés présentés jusqu'ici sont chargés négativement. Le remplacement du ribose ou du désoxyribose par un anneau morpholine (hétérocycle saturé porteur d'une fonction éther et d'une fonction amine secondaire) et de la liaison phosphodiester par une liaison phosphorodiamidate, résulte en un AO neutre, stable, extrêmement résistant à la dégradation et peu toxique, appelé PMO (*Phosphorodiamidate Morpholino Oligomers*) (Figure 37). Des charges positives peuvent être ajoutées à ce PMO par association avec des résidus comme la piperazine (PMOplus) ou un peptide riche en arginine (PPMO, Pip6-PMO) (Kole et al., 2012). Les PMOs ne permettent pas le recrutement de la RNase H1, mais forment un hybride stable avec leur cible ARNm, ce qui a pour conséquence d'interférer dans la maturation ou la traduction de l'ARNm et ainsi modifier l'expression du gène (Kole et al., 2012) (Figure 38).

Une autre catégorie d'AOs stables ne possédant pas de squelette ribose-phosphate a été créée. Ce sont les *peptide nucleic acids* (PNA), des analogues des acides nucléiques, non chargés, flexibles, avec un squelette polyamide comprenant des répétitions d'unités N-(2-aminoethyl)glycine auxquelles des bases sont attachées par une liaison carbonyle méthylène (Egholm et al., 1993; Nielsen et al., 1991) (Figure 37). Ces PNA peuvent former des duplex ou des triplex extrêmement stables avec les acides nucléiques (simple-brin ou double-brin ADN ou ARN) (Jensen et al., 1997). De la même manière que les PMOs, les PNA ne permettent pas de recruter la RNaseH1. Leur mode d'action implique plutôt un encombrement stérique (Dean, 2000; Dias and Stein, 2002) (Figure 38).



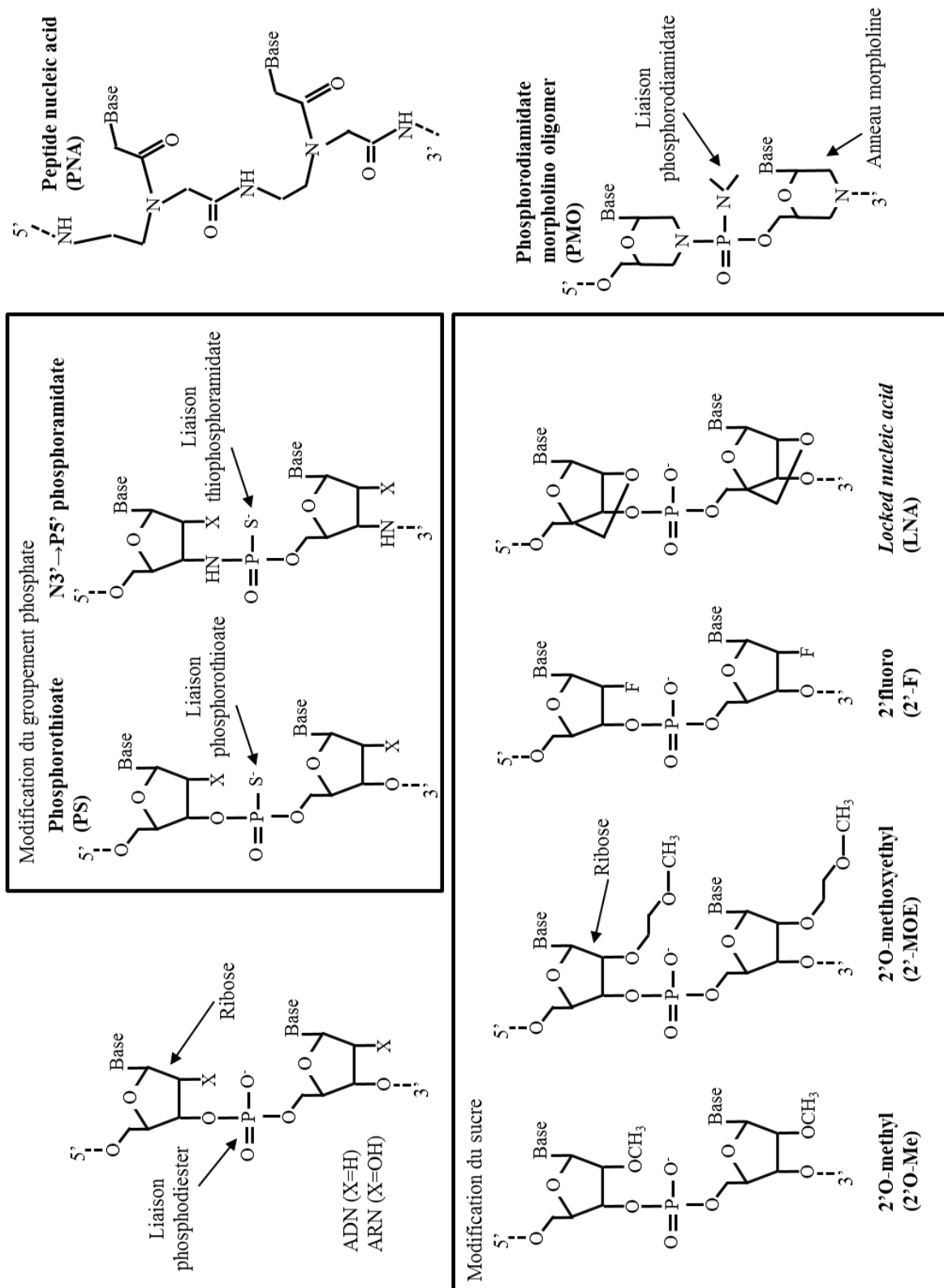


FIGURE 37. Structure des acides nucléiques chimiquement modifiés les plus communément utilisés. Ces modifications concernent aussi bien le sucre (ribose ou désoxyribose) que le groupement phosphate (adapté de Watts and Corey, 2012).

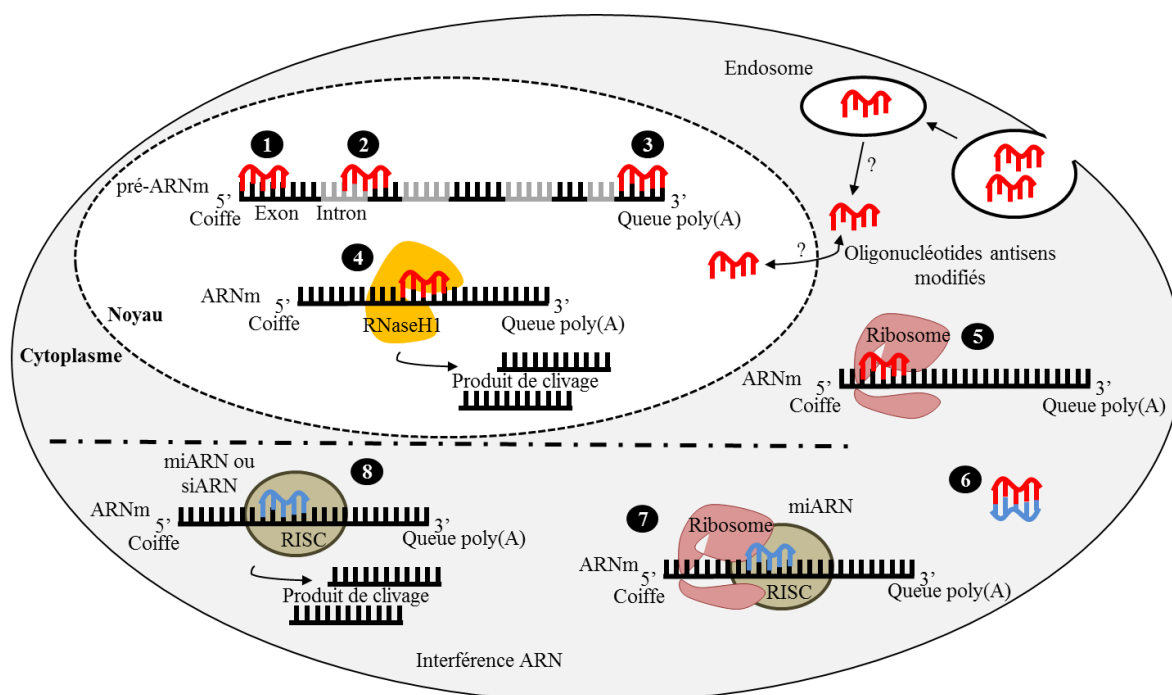


FIGURE 38. Mécanismes d'action des oligonucléotides antisens et de l'ARN interférence. Les oligonucléotides antisens permettent 1) d'empêcher la formation de la coiffe en 5' du pré-ARNm, 2) de moduler l'épissage du pré-ARNm, 3) de modifier l'ajout d'une queue poly(A) sur le pré-ARNm, 4) de dégrader le pré-ARNm ou l'ARNm en recrutant la nucléase RNaseH1, 5) d'inhiber la traduction d'un ARNm, 6) de cibler des miARN. L'ARN interférence (siARN et miARN) permet, par l'intermédiaire du complexe RISC, et uniquement dans le cytoplasme 7) d'inhiber la traduction ou 8) de cliver l'ARNm cible (adaptée de DeVos and Miller, 2013).

A la différence des siARN/miARN ou des snARN pris individuellement, les AOs présentent un champ d'action plus large (DeVos and Miller, 2013; Kole et al., 2012; Watts and Corey, 2012) (Figure 38). Parce qu'ils sont indépendants d'une interaction avec le complexe RISC, leur activité porte aussi bien sur des ARNm localisés dans le cytoplasme que sur des pré-ARNm localisés dans le noyau. Les AOs ont été décrits comme permettant notamment de dégrader l'ARNm ou le pré-ARNm cible par recrutement de l'activité de l'enzyme RNaseH1, d'empêcher la formation de la coiffe en 5' du pré-ARNm, de moduler l'épissage du pré-ARNm, de modifier l'ajout d'une queue de polyadénylation sur le pré-mRNA, de cibler des miARN ou encore d'inhiber la traduction (DeVos and Miller, 2013; Watts and Corey, 2012) (Figure 38).

Les oligonucléotides antisens sont donc des outils très prometteurs pour le développement de thérapies dans le cadre de maladies à gain de fonction mais aussi de maladies à perte de fonction, notamment en raison de leur capacité à augmenter ou à diminuer le niveau en ARNm cible. Des preuves quant à leurs fonctionnalités ont été apportées lors d'études *in vitro*

et *in vivo* de l'efficacité de stratégies thérapeutiques utilisant des oligonucléotides antisens pour diverses pathologies incluant des cancers, la thrombose mais également des pathologies musculaires telles que la DMD ou l'Amyotrophie Musculaire Spinale (SMA). Dans les stratégies thérapeutiques pour des pathologies musculaires, les oligonucléotides antisens sont très utilisés pour générer un saut ou une inclusion d'exons (Betts et al., 2012; Büller et al., 2015; Kaczmarek et al., 2015; Ma et al., 2012; Mitropant et al., 2013; Odermatt et al., 2016; Porensky and Burghes, 2013; Zhou et al., 2013). A ce jour, plusieurs essais cliniques utilisant des oligonucléotides antisens sont en cours. Récemment, la FDA (*Food and Drug Administration*) aux Etats-Unis a approuvé un PMO développé par Sarepta, l'Eteplirsén (ou Exondys 51), qui induit un saut d'exon au niveau de l'ARNm de la dystrophine chez des patients DMD (Mendell et al., 2016). Cependant, l'expression des oligonucléotides antisens est transitoire (à l'exception des shRNA), ce qui implique des injections locales ou systémiques répétées pour obtenir une activité à long terme de l'oligonucléotides antisens. Ces injections répétées d'oligonucléotides antisens peuvent être à l'origine d'effets secondaires indésirables comme des réponses immunitaires. Malgré les grands progrès pour augmenter la stabilité des oligonucléotides antisens et réduire leur toxicité, de nombreux développements sont encore à prévoir afin d'améliorer leur efficacité thérapeutique.

### **1.5.1.2. Différentes méthodes de transfert de l'acide nucléique thérapeutique**

L'efficacité de transfert de l'acide nucléique thérapeutique dépend d'au moins deux facteurs (Herweijer and Wolff, 2003; Nayerossadat et al., 2012) :

- de l'organe ou des cellules à cibler,
- de la forme de l'acide nucléique thérapeutique : nu ou associé à des vecteurs physiques, chimiques/synthétiques ou viraux.

En effet, avant d'atteindre sa cible, l'acide nucléique aura plusieurs barrières biologiques à franchir, dont la membrane plasmique cellulaire (Kaneda, 2001). Associer l'acide nucléique à des vecteurs physiques, chimiques ou viraux impactera positivement sa capacité à franchir les différentes barrières biologiques. Les acides nucléiques nus ont généralement une action transitoire, sont plus sensibles à une action des nucléases et sont rapidement éliminés par l'organisme. La durée de demi-vie d'un plasmide ADN est estimée à 10 min lors d'une

injection intraveineuse chez la souris (Kawabata et al., 1995). Les méthodes virales permettent d'obtenir un effet sur un plus long terme et sont de ce fait favorisées dans les essais cliniques comparées aux méthodes non virales (Figure 39). Malgré tout, ces dernières présentent également des limites. Parfois difficiles à produire, les virus modifiés sont des fois associés à une carcinogénèse et/ou une immunogénicité. De plus, les virus modifiés ont une capacité limitée quant à la taille de l'acide nucléique qu'ils peuvent prendre en charge et présentent un tropisme quelque fois très restreint (Mountain, 2000; Yin et al., 2014).

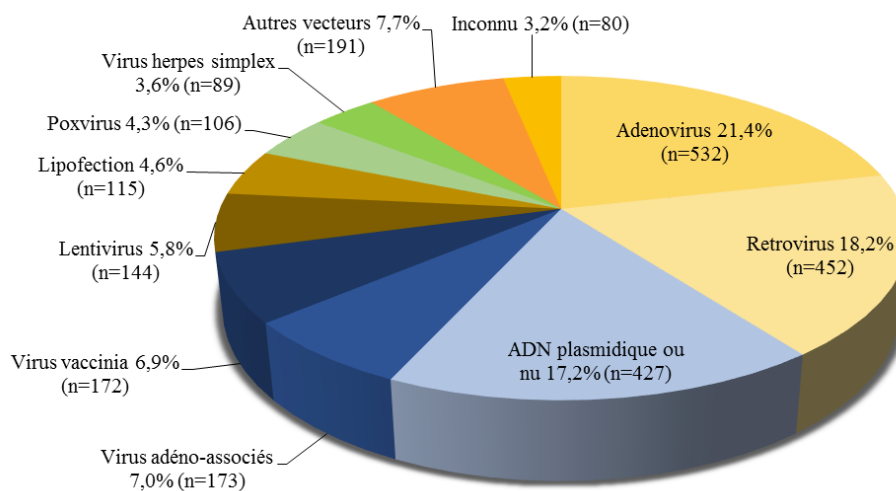


FIGURE 39. Vecteurs utilisés dans les essais de thérapie génique en cours (Avril 2017) (<http://www.abedia.com/wiley/indications.php>).

### 1.5.1.2.1. Méthodes synthétiques et physiques

Plusieurs méthodes physiques améliorent l'efficacité de transfert d'un acide nucléique en favorisant son cheminement du milieu extracellulaire à sa cible localisée dans le cytoplasme ou le noyau (Mountain, 2000; Yin et al., 2014). Le transfert d'un acide nucléique seul, dont la production par des bactéries est facilement réalisable et représente un faible coût, est peu efficace *in vitro*, et encore moins *in vivo*. Afin d'augmenter l'efficacité de transfert *in vivo*, l'acide nucléique nu peut être délivré en systémique selon une procédure hydrodynamique, qui consiste à injecter rapidement un très large volume d'acides nucléiques en solution aqueuse au niveau de la veine de la queue d'une souris ou d'un rat (Zhang et al., 2004). Cette technique est principalement appliquée dans le cas où l'organe cible est le foie. Une autre alternative pour un transfert *in vivo* ou *in vitro* est l'électroporation. L'application d'un champ électrique crée des modifications transitoires de la perméabilité membranaire (création de

pores transitoires) permettant à l'acide nucléique de franchir la membrane plasmique (André et al., 2008; Li, 2004; Rosazza et al., 2016). Cette procédure est difficilement utilisable pour une application en thérapie génique, en partie parce qu'elle rend la majorité des cellules non viables même si des progrès ont été obtenus (Gehl, 2014). L'acide nucléique ne peut également être transféré aux cellules cibles à l'aide de la technique de *gene gun* ou biolistique. Des microbilles de métal enrobées d'ADN sont projetées à grande vitesse sur les cellules pour qu'elles traversent leur membrane plasmique et délivrent ainsi l'acide nucléique (Goudy et al., 2008; O'Brien and Lummis, 2006). La biolistique, qui peut être utilisée *in vitro* et *in vivo*, est très délétère pour des cellules déjà fragilisées au niveau de leur membrane plasmique. La dernière méthode physique de transfert des acides nucléiques est la sonoporation ultrason. Les ultrasons génèrent un stimulus de perméabilisation de la membrane favorisant l'entrée des acides nucléiques dans la cellule (Li et al., 2009; Newman et al., 2001). Ces méthodes physiques sont envisageables dans des stratégies de thérapies géniques mais fragilisent souvent les cellules. Par ailleurs, ces techniques ne peuvent être appliquées que pour une surface restreinte à traiter, excluant notamment les maladies musculaires.

Des méthodes synthétiques de transfert de gènes *in vivo* et *in vitro*, moins brutales que les méthodes physiques, ont été développées. Ces méthodes utilisent des vecteurs chimiques répartis en deux grandes familles : les lipides cationiques et les polymères cationiques (Tros de Ilarduya et al., 2010). Les lipides cationiques sous la forme de liposomes, se lient aux acides nucléiques par des interactions électrostatiques et forment naturellement des complexes moléculaires (*lipoplexes*). Les lipoplexes offrent une protection transitoire à l'acide nucléique contre la dégradation. Les lipoplexes sont chargés positivement et peuvent donc interagir avec des molécules chargées négativement, ce qui réduit l'efficacité de transfert de l'acide nucléique à la cellule cible (Junquera and Aicart, 2014; Lonz et al., 2008; Zhang et al., 2012). Les polymères cationiques, incluant le PEI (*polyethylenimine*), s'associent également à l'acide nucléique par interactions électrostatiques pour former des complexes moléculaires (*polyplexes*) (Sun and Zhang, 2010). L'efficacité de transfert de l'acide nucléique par ces deux types de vecteurs synthétiques est généralement faible. Cette efficacité dépend de la taille de la macromolécule ou de l'acide nucléique à transférer, de la charge globale du lipoplexe ou du polyplexe et du type cellulaire ciblé. Une nouvelle génération de vecteurs synthétiques, les polylipoplexes (combinaison de lipides cationiques et de polymères cationiques (PEI)), ont une meilleure efficacité de transfert de l'acide nucléique à la cellule cible (García et al., 2012; Pelisek et al., 2006; Rezaee et al., 2016; Yamazaki et al., 2000).

Dans l'ensemble, les vecteurs chimiques permettent le transfert puis l'expression de nombreux types d'acides nucléiques (simple-brin, double-brin, ARN, ADN AO modifié) de différentes tailles, et ce avec une toxicité variable en fonction de la chimie du vecteur synthétique et de la dose injectée (Argyros et al., 2008; Ferrer et al., 2004; Jiao et al., 1992; Mountain, 2000; Yew and Scheule, 2005; Yin et al., 2014). Le mécanisme exact de transfert de l'acide nucléique par ces vecteurs chimiques n'est pas totalement élucidé, même s'il est admis que les complexes moléculaires sont internalisés par endocytose suite à une interaction avec la membrane plasmique (Medina-Kauwe et al., 2005; Midoux et al., 2008; Pichon et al., 2010). Les vecteurs synthétiques sont peu utilisés en thérapie génique car ils ne permettent que trop rarement un taux d'expression élevé et sur le long terme du transgène (Li et al., 2009; Yin et al., 2014).

#### **1.5.1.2.2. Méthodes virales**

Les méthodes virales reposent sur la capacité des virus à transférer leur matériel génétique dans la cellule cible. Les principaux vecteurs viraux utilisés en thérapie génique sont des adénovirus, des rétro/lentivirus, des virus adéno-associés (AAV), des poxivirus ou encore le virus de l'herpès (HSV1) (Berto et al., 2005; Büning et al., 2008; Escors and Breckpot, 2010; Manservigi et al., 2010; Moroziewicz and Kaufman, 2005; Mountain, 2000; Vetrini and Ng, 2010; Yi et al., 2011). Ces virus ont subi des modifications afin d'inhiber leur pathogénicité et d'atténuer leur immunogénicité, tout en gardant ou modifiant leur tropisme cellulaire ainsi que leur persistance, et en remplaçant les gènes viraux ou de réplication par le transgène d'intérêt (Bouard et al., 2009). Cependant, les vecteurs viraux modifiés peuvent être encore carcinogènes ou génotoxiques, et présenter une certaine immunogénicité (Baum et al., 2006; Bessis et al., 2004). Malgré tout, les vecteurs viraux sont les vecteurs de transfert d'acides nucléiques les plus couramment utilisés dans les essais de thérapie génique (Figure 39). A noter que le choix du vecteur viral dépend de la stratégie thérapeutique à mener et de l'organe visé. En effet, la capacité de transfert de l'acide nucléique, pouvant aller de quelques kb à plusieurs dizaines de kb, ainsi que la structure ou la chimie de l'acide nucléique pris en charge, varie d'un vecteur viral modifié à un autre. Le type de vecteur viral a également une influence sur la stabilité de l'expression (long terme pour les AAV et les lentivirus), ou sur la forme de maintien du transgène *in vivo* (épisomale pour la majorité des vecteurs et intégré pour les lentivirus).

## **1.5.2. Stratégies thérapeutiques pour la FSHD**

### **1.5.2.1. Traitements symptomatiques**

A ce jour, malgré tous les outils de thérapie génique et tous les progrès réalisés dans la compréhension des facteurs génétiques et des mécanismes physiopathologiques impliqués dans la FSHD, aucun traitement curatif pour cette maladie n'existe. Actuellement, les seuls traitements proposés aux patients FSHD sont des traitements symptomatiques qui visent à prévenir raideurs et douleurs articulaires ainsi qu'à améliorer la qualité de vie. Ces traitements se résument principalement à des exercices physiques, des soins de kinésithérapie et éventuellement l'administration d'antalgiques (Attarian et al., 2012; Tawil et al., 2015). Les exercices musculaires, lorsqu'ils peuvent être réalisés, jouent un rôle dans la prévention d'excès de fatigue et limitent l'atrophie musculaire ou le déconditionnement cardio-respiratoire et métabolique. Trois études basées sur un programme d'entraînement d'endurance et deux études basées sur un programme d'entraînement de force ont été menées chez des patients FSHD (Andersen et al., 2015; Colson et al., 2010; van der Kooi et al., 2004; Olsen et al., 2005; Voet et al., 2014). Seuls les entraînements d'endurance donnent des résultats positifs et permettent d'améliorer la qualité de vie. Cependant, la combinaison des composantes force et endurance dans le même entraînement semble être plus efficace qu'un entraînement avec la composante endurance seule (Bankolé et al., 2016; Cup et al., 2007; Voet et al., 2013). Récemment, une étude d'une durée de six mois menée par le groupe de Léonard Féasson, avec 16 patients FSHD, a montré des effets bénéfiques d'un programme d'entraînement combinant à la fois force et endurance (avec une alternance de phases de faible intensité et de forte intensité) sans perception accrue d'une fatigue ou de douleurs. Ce programme, compatible avec les activités professionnelles et sociales de la vie quotidienne des patients FSHD, conduit à des bénéfices fonctionnels comme une amélioration des capacités de la marche ou de la force maximale, et cela sans compromettre le tissu musculaire (Bankolé et al., 2016). Dans le traitement des autres symptômes de la FSHD, en cas de complications respiratoires, un respirateur artificiel est proposé au patient. Par ailleurs, des orthèses peuvent être utilisées pour maintenir une posture perdue en raison de la faiblesse musculaire. Enfin, une intervention chirurgicale orthopédique, qui consiste à fixer l'omoplate à la cage thoracique, peut être programmée, dans le but d'améliorer l'amplitude des mouvements des membres supérieurs (Giannini et al., 2007).

### 1.5.2.2. Stratégies thérapeutiques avec des molécules pharmacologiques

Malgré l'absence de traitements thérapeutiques pour la FSHD, plusieurs stratégies sont proposées dans la littérature. Parmi ces stratégies, certaines ont pour but de modifier la progression de la maladie en utilisant des molécules pharmacologiques comme l'albuterol (également appelé salbutamol) ou encore le resolaris (ATYR1940, *aTyr Pharma*) (Kissel et al., 1998, 2001, van der Kooi et al., 2004, 2007; Payan et al., 2009). L'albuterol est un agoniste bêta-2-adrénergique dont les effets sur le muscle ne sont pas connus. Une hypothèse est que cette molécule, en tant qu'agoniste bêta-2-adrénergique, permet de diminuer la fatigabilité musculaire en augmentant la force et la masse du muscle (Martineau et al., 1992; Yang and McElligott, 1989). Des études ont associé entraînements de force et administrations d'albutérol dans le but de réduire le handicap lié à la pathologie et d'améliorer la qualité de vie. Les résultats de ces études n'ont pas montré d'effets positifs ou négatifs nets de la combinaison de l'entraînement de force avec l'albutérol sur la douleur, la fatigue, la détresse psychologique, ou encore la force et la masse musculaire (Kissel et al., 1998, 2001, van der Kooi et al., 2004, 2007; Payan et al., 2009). Le Resolaris, quant à lui, est une physiocrine immunomodulatrice (protéine endogène humaine dérivant des ARNt synthétases qui agit comme une molécule de signalisation extracellulaire dans plusieurs voies physiologiques). Cette molécule est actuellement testée par la compagnie *aTyr Pharma* dans un essai clinique de phase Ib/II chez des patients FSHD et des patients atteints de la myopathie des ceintures de type 2B où une composante inflammatoire dans le développement de la pathologie a été mise en évidence (Arahata et al., 1995). Les résultats de cet essai clinique, démarré en 2015, n'ont pas encore été présentés.

Récemment, une étude de criblage haut-débit de petites molécules a été menée sur un modèle cellulaire murin C2C12 surexprimant de manière conditionnelle *DUX4*. Cette étude a permis d'identifier des molécules inhibant la toxicité ou la mort cellulaire médiée par une surexpression de *DUX4* (Bosnakovski et al., 2014). Ces molécules peuvent directement interférer dans l'activité de la protéine *DUX4* ou interférer dans les voies de signalisation en aval de *DUX4* qui conduisent à la mort cellulaire. Bien que le mécanisme d'action de la majorité de ces molécules ne soit pas connu, certaines d'entre elles sont caractérisées comme intervenant en tant qu'antioxydants dans la voie de réponse au stress oxydant (Bosnakovski et al., 2014). Ces résultats sont appuyés par des études montrant que la réponse au stress oxydant



est altérée dans la FSHD. En effet, une expression des marqueurs du stress oxydant et une dérégulation des gènes impliqués dans la réponse à ce stress oxydant ont été détectées dans des myoblastes FSHD, rendant la cellule plus susceptible aux agents oxydants (Cheli et al., 2011; Turki et al., 2012; Winokur et al., 2003a, 2003b). En accord avec ces résultats, une stratégie visant à augmenter la réponse cellulaire ou à protéger la cellule d'un stress oxydant pourrait permettre de ralentir le développement de la FSHD. Cependant, une étude, de 17 semaines chez des patients FSHD, testant l'effet d'une supplémentation en antioxydants (Vitamine C, Vitamine E, gluconate de zinc et sélénométhionine) n'a permis de montrer que de faibles impacts positifs de ces molécules sur la force et l'endurance du quadriceps des patients FSHD traités (Passerieux et al., 2015).

### **1.5.2.3. Stratégies de thérapie génique ciblant *DUX4***

Des stratégies de thérapie génique ont également été développées pour la FSHD. Néanmoins, aucune d'entre elles n'a atteint à ce jour les essais cliniques (Himeda et al., 2016a; Mitsuhashi et al., 2013a; Vanderplanck et al., 2011; Wallace et al., 2012).

Dans le cas des maladies à perte de fonction, les stratégies thérapeutiques ont pour but de venir corriger ou compléter un gène muté. Or la FSHD n'est pas une maladie à perte de fonction mais une maladie à gain de fonction, qui se caractérise par l'expression aberrante du facteur de transcription *DUX4* et de ses gènes cibles. Les stratégies thérapeutiques, dans ce cas, doivent donc être basées sur l'extinction du gène *DUX4*. L'ORF de *DUX4* étant localisé uniquement dans l'exon 1, il est extrêmement difficile de développer une stratégie venant perturber le cadre de lecture par des oligonucléotides antisens.

Diverses études *in vitro* et *in vivo* ont validé les stratégies induisant la dégradation ou déstabilisant l'ARNm *DUX4*. Ces stratégies permettent de réverser un grand nombre des phénotypes moléculaires liés à la FSHD (Vanderplanck et al., 2011; Wallace et al., 2012). En effet, des siARN ou des oligonucléotides antisens 2'O-Me masquant soit un site accepteur d'épissage aux jonctions introns-exons, soit la région 3'UTR de *DUX4*, altèrent la stabilité de l'ARN et empêchent l'expression de la protéine *DUX4* (Vanderplanck et al., 2011). Des siARN ciblant la région en amont du site de début de transcription de *DUX4* suppriment également l'expression de *DUX4*, augmentent les marques épigénétiques H3K9me3 et induisent le recrutement d'ARGONAUTE2. De manière surprenante, l'utilisation de *gapmer*

ciblant les mêmes séquences ne permet pas une répression de l'expression de *DUX4*, indiquant un mécanisme répresseur dépendant, dans cette stratégie, de DICER et d'ARGONAUTE2 (Lim et al., 2015). Une autre stratégie consiste à venir inhiber l'expression de *DUX4* par des miARN artificiels sous promoteur U6 ciblant l'ORF *DUX4* (Wallace et al., 2012). Une réduction de l'ARNm *DUX4* est observée, associée à une diminution ou une absence d'expression des symptômes moléculaires de la FSHD. La limite des approches d'ARN interférence proposées est la haute toxicité générée en raison d'effets *off-targets* (Grimm et al., 2006; McBride et al., 2008).

Une autre stratégie thérapeutique, n'utilisant pas d'antisens oligonucléotides, réside dans le fait que l'isoforme courte de la protéine DUX4, DUX4-s, peut également se lier aux séquences cibles reconnues par l'isoforme longue et toxique DUX4-fl, sans pour autant pouvoir activer la transcription des gènes, en raison de l'absence du domaine de transactivation (Geng et al., 2012; Kawamura-Saito et al., 2006; Mitsuhashi et al., 2013a). Ainsi, la surexpression de la protéine DUX4-s, qui entre en compétition avec DUX4-fl et agit donc à la manière d'un dominant négatif, permet de diminuer l'activation des gènes cibles de DUX4-fl et de réverser le phénotype FSHD (Geng et al., 2012; Mitsuhashi et al., 2013a).

Récemment, des CRISPR/Cas-9 déficients pour l'activité nucléase et fusionnés à un répresseur transcriptionnel KRAB (dCas9-KRAB) ont été testés dans le cadre du développement d'une autre stratégie thérapeutique pour la FSHD (Himeda et al., 2016a, 2016b; Kearns et al., 2015; Mandegar et al., 2016; Parsi et al., 2017; Thakore et al., 2015). Ce système, en se fixant à la région promotrice ou au niveau de l'exon 1 de *DUX4*, réprime spécifiquement la transcription de *DUX4*. Une diminution de l'expression des gènes cibles de *DUX4* dans des cellules FSHD est également observée. Le complexe dCas9-KRAB permet le recrutement du co-répresseur KAP1/TRIM28 qui engendre à son tour le recrutement de la protéine HP1 $\gamma$ . Il en résulte, aux sites de fixation du complexe dCas9-KRAB, de légers changements épigénétiques avec la mise en place de marques associées à un état de répression transcriptionnelle (Himeda et al., 2016a, 2016b).

### **1.5.2.3. Stratégies de thérapie génique ciblant d'autres gènes candidats dans la FSHD**

Le gène *DUX4* n'ayant pas été proposé comme le seul gène candidat dans la FSHD, des stratégies thérapeutiques ciblant *FRG1* avec des miARN délivrés par AAV ou des shARN ont également été proposées (Bortolanza et al., 2011; Wallace et al., 2011b). L'inhibition de l'expression de *FRG1* dans un modèle murin surexprimant ce gène améliore notamment la masse musculaire, la force et réduit les anomalies histopathologiques associées au phénotype de cette souris (Bortolanza et al., 2011; Wallace et al., 2011b). Par ailleurs, l'expression de *DUX4* conduisant à de multiples dérégulations géniques dans le muscle, une stratégie ciblant cette fois-ci un gène potentiellement activé par *DUX4*, *PITX1*, a été testée *in vivo* (Dixit et al., 2007; Pandey et al., 2012, 2014). Dans cette étude des morpholinos dessinés pour inhiber la traduction de *PITX1* sont injectés dans le modèle murin surexprimant ce gène. Les résultats indiquent une amélioration du phénotype musculaire de la souris et démontrent la faisabilité de cette stratégie (Pandey et al., 2014). Cependant, une activation par *DUX4* de *PITX1*, tout comme un rôle de *PITX1* dans la FSHD reste à confirmer.

En résumé, même si notre compréhension générale des mécanismes génétiques, épigénétiques et moléculaires conduisant au phénotype FSHD a rapidement progressé ces dernières années, le développement, plus lent, de stratégies thérapeutiques, n'est pas encore suffisant pour offrir un traitement curatif aux patients.

## **1.6. Projet de thèse**

L'essor de nouvelles technologies en chimie et biologie permet aujourd'hui de modifier positivement ou négativement l'expression d'un gène en agissant à différents niveaux. De ce fait, de nombreuses stratégies de thérapie génique totalement différentes peuvent être proposées pour une même pathologie, ce qui est le cas pour la FSHD. Le développement de plusieurs stratégies thérapeutiques distinctes pour la FSHD ne doit pas être considéré comme une dispersion de la recherche et des ressources. Bien au contraire, la multiplicité des propositions, par le biais de technologies variées et ciblant différents éléments génétiques ou épigénétiques impliqués dans la FSHD, augmente la probabilité que l'une de ces stratégies atteigne les essais cliniques dans les quelques années à venir.

### 1.6.1. Mécanisme et rôle de la polyadénylation des ARNm

Au cours de ma thèse, nous avons développé une stratégie thérapeutique innovante pour la FSHD. Nous avons décidé de cibler les éléments clés en 3' de l'ARNm *DUX4* afin de le déstabiliser et d'inhiber la synthèse protéique. En effet, le clivage et la polyadénylation des ARNm en 3' sont des mécanismes essentiels à la maturation de la majorité des ARNm des eucaryotes (Danckwardt et al., 2008; Millevoi and Vagner, 2010; Proudfoot, 2011; Zhao et al., 1999). Chez l'Homme, ces réactions ont lieu dans le noyau. Ce sont des étapes co-transcriptionnelles qui sont dirigées pas plus de 80 protéines et par des séquences régulatrices en *cis* (Danckwardt et al., 2008; Mandel et al., 2008; Nunes et al., 2010; Shi and Manley, 2015). La région de clivage est définie par trois séquences clés en 3' de l'ARNm qui sont : (i) une séquence hexamérique qui correspond au signal de polyadénylation, (ii) un site de clivage positionné 20 à 30 nucléotides en aval du signal de polyadénylation, et (iii) une région riche en résidus G/U ou U (*Downstream sequence element, DSE*) localisée plus de 30 nucléotides après le site de clivage (Danckwardt et al., 2008) (Tableau 5 et Figure 40).

D'un point de vue mécanistique, le complexe protéique CPSF (*Cleavage and Polyadenylation Specificity Factor*) reconnaît tout d'abord la séquence consensus du signal de polyadénylation (Danckwardt et al., 2008; Keller et al., 1991) (Figure 40). Chez l'humain et la souris, plusieurs séquences hexamériques avec des fréquences d'apparition variables en 3' des ARNm ont été identifiées comme des signaux de polyadénylation. Les séquences AAUAAA et AUUAAA sont les plus communes (Beaudoing et al., 2000; Tian et al., 2005) (Tableau 5). Le complexe protéique CstF (*Cleavage stimulating Factor*) se fixe ensuite au niveau de la région riche en résidus G/U et U, alors que les facteurs de clivage CF (*Cleavage Factor*) se positionnent au niveau de la future région de clivage (Danckwardt et al., 2008; MacDonald et al., 1994) (Figure 40). Le recrutement de l'ensemble de ces protéines conduit à un clivage endonucléolytique. Deux fragments sont générés : un fragment avec un groupe hydroxyle sortant en 3' et un fragment avec un groupe phosphate sortant en 5' (Figure 40). Le fragment avec le groupe phosphate sortant en 5' et lié au complexe CstF est immédiatement dégradé par des exonucléases. La polymérase Poly(A) (*PAP Poly(A) Polymerase*) ajoute alors environ 250 adénosines, par catalyse de la réaction de polymérisation des adénosines à partir des adénosines triphosphates, au niveau du fragment possédant le groupe hydroxyle libre en 3' (Figure 40). Cette queue poly(A) est ensuite recouverte par la protéine de liaison poly(A)

PABP (*Poly(A) Binding Protein*) (Figure 40). La fixation de PABP ne peut se faire que si la queue poly(A) est composée d'au moins dix adénosines. Par ailleurs, PABP stimule la réaction de polyadénylation catalysée par PAP et régule la taille de la queue poly(A) (Kerwitz et al., 2003; Kühn et al., 2009). Une interaction entre PABP et les facteurs d'initiation de la traduction en 5' de l'ARNm induit une circularisation de l'ARNm (Danckwardt et al., 2008; Mandel et al., 2008; Zhao et al., 1999) (Figure 40). Cette maturation est généralement nécessaire à l'export de l'ARNm, et joue un rôle dans la stabilité et l'efficacité de la traduction de l'ARNm (Danckwardt et al., 2008; Mangus et al., 2003; Millevoi and Vagner, 2010).

TABLEAU 5. Séquences hexamériques identifiées comme des signaux de polyadénylation chez l'Homme et la souris. Ces signaux de polyadénylation sont classés par fréquence d'utilisation (adapté de Beadoing et al., 2000; Tian et al., 2005).

Séquence de six nucléotides	Fréquences (Tian et al., 2005)		Fréquences (Beadoing et al., 2000)
	Humain	Souris	Humain
AAUAAA	53.18	59.16	58.2
AUUAAA	16.78	16.11	14.9
UAUAAA	4.37	3.79	3.2
AGUAAA	3.72	3.28	2.7
AAGAAA	2.99	2.15	1.1
AAUAUA	2.13	1.71	1.7
AAUACA	2.03	1.65	1.2
CAUAAA	1.92	1.80	1.3
GAUAAA	1.75	1.16	1.3
AAUGAA	1.56	0.90	0.8
UUUAAA	1.20	1.08	1.2
ACUAAA	0.93	0.64	0.6
AAUAGA	0.60	0.36	0.7
AAAAAG	-	-	0.8
AAAACA	-	-	0.5
GGGGCU	-	-	0.3

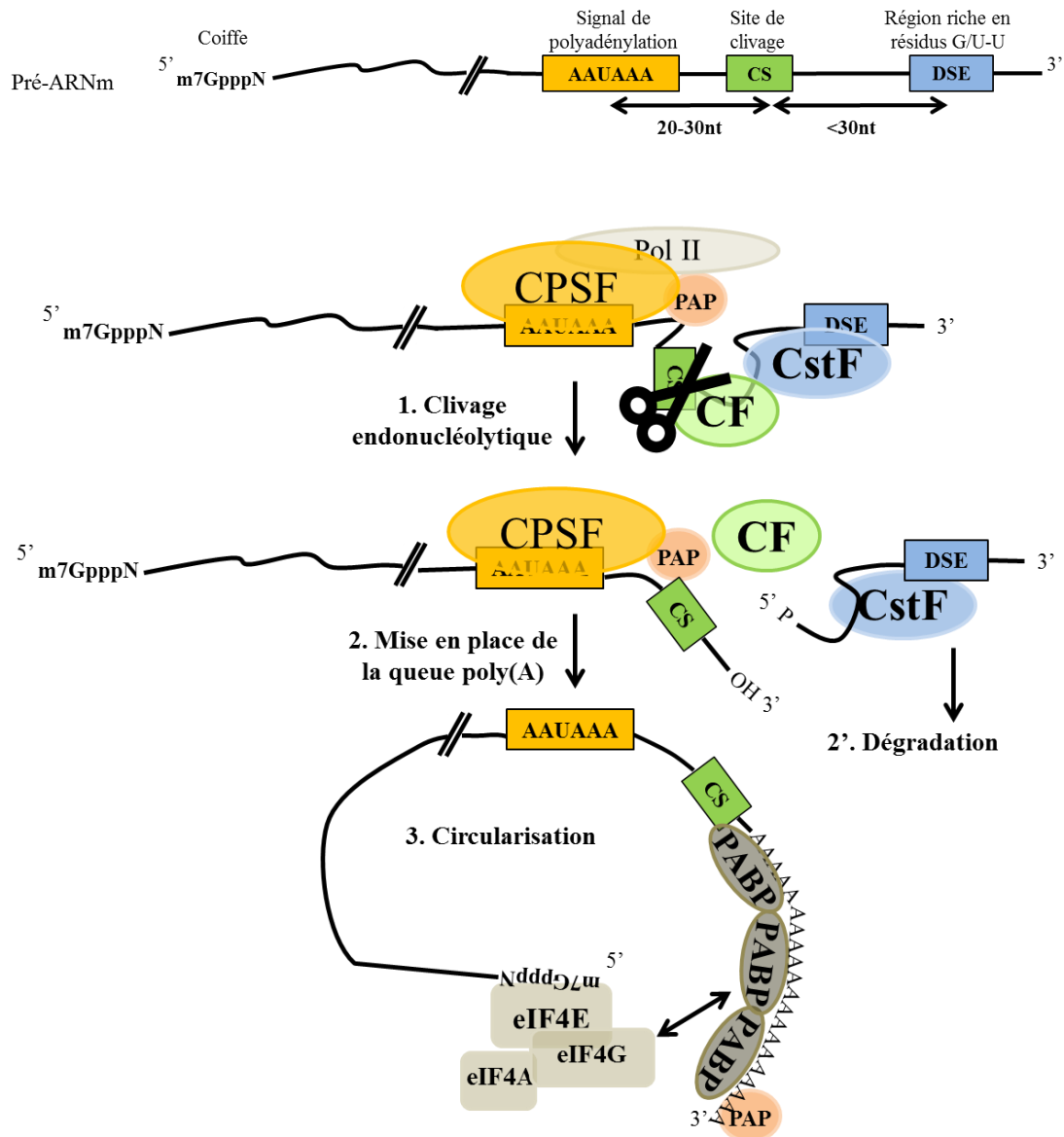


FIGURE 40. Mécanismes de clivage et de polyadénylation des ARNm en 3'. La région de clivage est définie par trois séquences clés en 3' de l'ARNm qui sont : une séquence hexamérique qui correspond au signal de polyadénylation (AAUAAA ici), un site de clivage (CS) positionné 20 à 30 nucléotides en aval du signal de polyadénylation, et une région riche en résidus G/U ou U (*DSE Downstream sequence element*) localisée plus de 30 nucléotides après le site de clivage. La première étape du mécanisme de maturation en 3' du pré-ARNm implique la fixation du complexe multiprotéique CPSF sur le signal de polyadénylation, du complexe CstF sur le signal DSE et des protéines CF au niveau du site de clivage. Le recrutement de l'ensemble de ces protéines induit un clivage endonucléolytique. Cette étape est suivie par l'ajout d'une queue poly(A) d'environ 250 adénosines en 3' libre de l'ARNm par la polymérase PAP. Cette queue poly(A) est recouverte par les protéines PABP, qui permettent, par interaction avec les facteurs d'initiation de la traduction en 5', la circularisation de l'ARNm. AAUAAA : signal de polyadénylation ; DSE : *Downstream sequence element* ; CS : Site de clivage ; CstF : *Cleavage stimulating Factor*; CPSF : *Cleavage/Polyadenylation specificity factor* ; PolII : *RNA Polymerase II* ; PAP : *Poly(A) polymerase* ; PABP : *Poly(A) Binding Protein* ; eIF4A/G/E : *eukaryotes initiation factor 4A/G/E* (adapté de Danckwardt et al., 2008).

Comme pour tous les ARNm, la stabilité de l'ARNm *DUX4* est modulée par l'ajout d'une queue poly(A). Ceci a été mis en évidence par l'étude de la région en aval de la dernière

répétition sur le chromosome 4qA et 10qA, puis la découverte de patients FSHD avec des réarrangements mitotiques entre ces deux chromosomes (Lemmers et al., 2004a, 2010a, 2012a). Les séquences qui gouvernent l'ajout d'une queue poly(A) sont situées dans l'exon 3 de l'ARNm *DUX4* et sont donc uniquement associées à l'haplotype 4/10qA. Cependant, une contraction des répétitions D4Z4 sur le chromosome 10qA ne permet pas la production d'un ARNm stable (Lemmers et al., 2010a). En effet, sur le chromosome 4qA, l'exon 3 de *DUX4* contient le signal de polyadénylation ATTA AAA, tandis que sur le chromosome 10qA, un polymorphisme affecte la séquence du signal de polyadénylation, qui devient alors AUCAAA ou AUUUAA. Ces séquences hexamériques ne sont pas décrites comme des signaux de polyadénylation (Beaudoing et al., 2000; Lemmers et al., 2010a; Tian et al., 2005). Ainsi, les transcrits *DUX4* issus des répétitions D4Z4 présentes sur le chromosome 10qA ne peuvent être stabilisés par l'ajout d'une queue poly(A) en l'absence d'un signal de polyadénylation fonctionnel et sont donc dégradés.

L'importance de ce signal de polyadénylation est également démontrée par le cas d'un individu atteint de FSHD, porteur d'une contraction des répétitions D4Z4 sur le chromosome 10qA mais pas sur le chromosome 4qA. Chez cet individu, des transcrits *DUX4* stables sont produits à partir du chromosome 10qA. Ceci est possible suite à une translocation du chromosome 4 vers le chromosome 10 d'un fragment démarrant au niveau de la dernière répétition D4Z4 et se poursuivant jusqu'au télomère. Cette translocation apporte un signal de polyadénylation fonctionnel sur le chromosome 10qA en aval de la dernière répétition D4Z4 (Lemmers et al., 2010a).

### **1.6.2. Hypothèse de travail**

Notre hypothèse de travail était que cibler, à l'aide d'oligonucléotides antisens, les éléments clés impliqués dans la maturation de l'ARNm *DUX4*, empêchant ainsi l'ajout d'une queue poly(A), pourrait permettre l'inhibition de *DUX4* et être utilisé comme une stratégie thérapeutique dans le cas de la FSHD.

## **2.RESULTATS**



## **2.1. Etude I : Développement d'une stratégie thérapeutique *in vitro* basée sur l'utilisation d'oligonucléotides antisens ciblant les éléments clés en 3'UTR de *DUX4***

Publication : *Antisens targeting of 3' end elements involved in DUX4 mRNA processing is an efficient therapeutic strategy for facioscapulohumeral dystrophy : a new gene-silencing approach.*

Anne-Charlotte Marsollier, Lukasz Ciszewski, Virginie Mariot, Linda Popplewell, Thomas Voit, George Dickson, Julie Dumonceaux.

Published in : *Human Molecular Genetics*, 2016, Vol. 25, No. 8. 1468-78.

## Résumé

L'objectif du projet était de déterminer si cibler des séquences clés impliquées dans la polyadénylation des ARNm, en utilisant des oligonucléotides antisens, peut-être une stratégie thérapeutique d'inhibition de l'expression de *DUX4* dans le cas de la FSHD.

Dans les cellules musculaires, seul le signal de polyadénylation AUUAAA a été décrit, sans étudier séparément les deux isoformes de l'ARNm *DUX4*. Afin de dessiner des oligonucléotides antisens spécifiques de la région 3'UTR de *DUX4*, la détermination précise des sites de clivage et des signaux de polyadénylation associés à chaque isoforme de *DUX4* est nécessaire. Une première expérience de 3'RACE-PCR nous a permis de montrer que les deux isoformes *DUX4-fl* partagent la même région de clivage et le même signal de polyadénylation. Au total, cinq PMOs ciblant le signal de polyadénylation et/ou le site de clivage et/ou le DSE ont été dessinés et testés *in vitro*, à différentes concentrations, sur des clones FSHD immortalisés à deux jours de différenciation. L'efficacité de chaque PMO a été évaluée par détermination de la quantité d'ARNm *DUX4* résiduels et le niveau d'expression de trois des gènes cibles de *DUX4*, dans des extraits d'ARNs totaux de myotubes à quatre jours de différenciation, lorsque l'expression de *DUX4* est la plus forte. Tous les PMOs sont efficaces et induisent une sous-expression de *DUX4* et de ses gènes cibles. Néanmoins, les PMOs ciblant le site de clivage et/ou le signal de polyadénylation (PMO-PAS et PMO-CS3) semblent être particulièrement prometteurs puisqu'ils induisent une diminution de l'ARNm *DUX4* de plus de 40% pour les deux isoformes.

Un changement de signal de polyadénylation en présence de ces PMOs a également été recherché. En effet, les données de la littérature rapportent que 54% des gènes humains présentent plus d'un signal de polyadénylation. Des événements de polyadénylation alternatifs peuvent ainsi avoir lieu lorsque le signal de polyadénylation est défaillant dans le dernier exon. Une redirection du site de clivage n'a été observée qu'en présence du PMO-CS3 ciblant le signal de polyadénylation et le site de clivage. En présence de ce PMO, un clivage en amont du signal de polyadénylation initial est réalisé. Un encombrement stérique généré par la présence du PMO peut être à l'origine de cette redirection de la région de clivage.

En conclusion, nous avons montré pour la première fois que l'utilisation d'oligonucléotides antisens ciblant les séquences clés impliquées dans l'ajout d'une queue poly(A) peut être une

stratégie thérapeutique efficace dans le cadre de la FSHD, mais également dans le cadre de maladies génétiques caractérisées par un gain de fonction.

## ORIGINAL ARTICLE

# Antisense targeting of 3' end elements involved in *DUX4* mRNA processing is an efficient therapeutic strategy for facioscapulohumeral dystrophy: a new gene-silencing approach

Anne-Charlotte Marsollier<sup>1</sup>, Lukasz Ciszewski<sup>2,†</sup>, Virginie Mariot<sup>1,†</sup>, Linda Popplewell<sup>2</sup>, Thomas Voit<sup>1,‡</sup>, George Dickson<sup>2</sup> and Julie Dumonceaux<sup>1,\*</sup>

<sup>1</sup>Sorbonne Universités UPMC Univ Paris 06, Inserm, CNRS, Centre de Recherche en Myologie (CRM), GH Pitié Salpêtrière, 47 bld de l'hôpital, Paris 13, France and <sup>2</sup>Centre of Biomedical Sciences, School of Biological Sciences, Royal Holloway University of London, Surrey TW20 0EX, UK

\*To whom correspondence should be addressed at: Sorbonne Universités UPMC Univ Paris 06, Inserm, CNRS, Centre de Recherche en Myologie (CRM), GH Pitié Salpêtrière, 47 bld de l'hôpital, Paris 13, France. Tel: +33 142165716; Fax: +33 142165700; Email: julie.dumonceaux@upmc.fr

## Abstract

Defects in mRNA 3' end formation have been described to alter transcription termination, transport of the mRNA from the nucleus to the cytoplasm, stability of the mRNA and translation efficiency. Therefore, inhibition of polyadenylation may lead to gene silencing. Here, we choose facioscapulohumeral dystrophy (FSHD) as a model to determine whether or not targeting key 3' end elements involved in mRNA processing using antisense oligonucleotide drugs can be used as a strategy for gene silencing within a potentially therapeutic context. FSHD is a gain-of-function disease characterized by the aberrant expression of the Double homeobox 4 (*DUX4*) transcription factor leading to altered pathogenic deregulation of multiple genes in muscles. Here, we demonstrate that targeting either the mRNA polyadenylation signal and/or cleavage site is an efficient strategy to down-regulate *DUX4* expression and to decrease the abnormally high-pathological expression of genes downstream of *DUX4*. We conclude that targeting key functional 3' end elements involved in pre-mRNA to mRNA maturation with antisense drugs can lead to efficient gene silencing and is thus a potentially effective therapeutic strategy for at least FSHD. Moreover, polyadenylation is a crucial step in the maturation of almost all eukaryotic mRNAs, and thus all mRNAs are virtually eligible for this antisense-mediated knockdown strategy.

## Introduction

The cleavage and polyadenylation of the 3' end of pre-mRNAs are fundamental processing steps for the maturation of the vast majority of eukaryotic mRNAs. In human cells, these reactions are governed by >80 RNA-binding proteins and by regulatory cis-acting RNA sequence elements (1–3). The key element dictating

the cleavage is a six-nucleotide (nt) motif called the poly(A) signal (PAS). Most of the mammalian mRNAs contain the PAS consensus AAUAAA or AUUAAA hexamer or close variants (4–6), which is recognized by the multimeric set of cleavage and polyadenylation factors. This RNA-protein interaction determines the site of cleavage which usually occurs 10–30 nt downstream of the PAS. The second important element is a Uracil/Guanine and

<sup>†</sup>L.C. and V.M. contributed equally.

<sup>‡</sup>Present address: NIHR Biomedical Research Centre, Institute of Child Health, University College London, 30 Guilford Street, London WC1N1EH, UK.

Received: December 4, 2015. Revised and Accepted: January 14, 2016

© The Author 2016. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com

uracil (U/GU)-rich sequence [downstream sequence element (DSE)] contacted by the cleavage stimulation factor and located 30–60 nt downstream the PAS motif (7). In most cases, these co-transcriptional maturations are required for nuclear export and stability of mRNAs, and their efficient translation (8), and thus consequently represent attractive and interesting drug and/or genetic targets for suppression of gene expression. Indeed, the functional importance of the 3' end mRNA processing has been highlighted by the discovery of mutations in the PAS cis-element causing or contributing to human diseases including haematological diseases, neurodegenerative diseases or cancer (9).

The aim of our study was to determine whether or not targeting elements involved in pre-mRNA maturation (such as PAS or DSE) using antisense oligonucleotide (AO) drugs can be a strategy for gene silencing within a potentially therapeutic context. We focused on facioscapulohumeral dystrophy (FSHD), which is a rare autosomal dominant neuromuscular disorder with an incidence of 1:14 000–1:20 000 characterized by the atrophy of specific groups of muscles (10). This pathology is caused by a loss of epigenetic marks within the D4Z4 macrosatellite located in the sub-telomeric region of chromosome 4 leading to chromatin relaxation (11). In 95% of the FSHD patients (named FSHD1), this chromatin relaxation is associated with a contraction of the D4Z4 array (12,13). The remaining 5% of the FSHD patients (named FSHD2) do not present a contraction of D4Z4, but the vast majority of them carry a mutation in the epigenetic modifier gene *SMCHD1* (14,15). This loss of epigenetic marks, when associated with a permissive chromosome 4 [i.e. carrying the 4qA haplotype containing the ATTAAA PAS commonly found in human (16)], leads to the aberrant transcription of a double-homeobox transcription factor named Double homeobox 4 (*DUX4*) whose ORF is present in each D4Z4 repeat (17). *DUX4* protein and mRNA have been robustly detected in adult and fetal FSHD1 and FSHD2 cells and biopsies whereas they can be found at very low levels in control (18–21). *DUX4* mRNA is only expressed in 1/1000 nuclei, but the *DUX4* protein can be found in 0.5–9% of nuclei (18,22), suggesting that although *DUX4* is not transcribed in most nuclei, the resulting protein may spread further (23). *DUX4* is a transcription factor and its overexpression is described to disturb several cellular pathways (23–30). Moreover, it was shown that even if *DUX4* expression has not been directly linked to patient's phenotype, *DUX4* may play a major role in the pathophysiology of FSHD because: (i) it has been shown that at least one D4Z4 repeat is needed for FSHD onset (31), (ii) only alleles with the 4qA type (containing the AUUAAA PAS for *DUX4* mRNA) are associated with FSHD (32,33), (iii) contraction of the D4Z4 array on chromosome 10, which carries a mutated PAS (AUCAAA), does not lead to FSHD (16), (iv) *DUX4*-induced gene expression is the major molecular signature in FSHD skeletal muscles (34) and (v) *DUX4* expression is the common point between FSHD1 and FSHD2 patients (16).

Several therapeutic strategies targeting *DUX4* expression have been proposed in the literature: RNA interference, 2'-O-methyl AOs targeting intron–exon junctions or over-expression of truncated *DUX4* (27,35–37). Here, we describe a new therapeutic AO-based therapeutic approach for FSHD using phosphorodiamidate morpholino oligomers (PMOs) targeting the key elements of 3' end processing of the *DUX4* transcript and which represent the key disease-permissive feature of the FSHD1 and 2 locus. Indeed, during the past 10 years, synthetic AOs have emerged as a promising strategy for drug therapy of genetic disorders affecting skeletal muscles such as the muscular dystrophies and motor-neuron diseases. PMOs have been mainly described to bind

with high affinity the targeted pre-mRNAs and mRNAs and to form a stable hybrid that efficiently blocks the access of splicing or translation machineries (so-called steric blocking AOs). PMOs can also form a structure mimicking a tRNA precursor, which is recognized by RNaseP leading to mRNA cleavage (38,39). Numerous studies investigating the therapeutic potential of antisense technology have been already performed and several clinical trials are now in progress for muscular dystrophies (40–43).

In this article, we choose to target the 3' key elements of *DUX4* pre-mRNA, focusing on the disease-permissive PAS, the cleavage site and the U/GU-rich DSE sequence. We demonstrated *in vitro* that targeting the 3' end elements of mRNA with antisense drugs can be an efficient therapeutic strategy for a genetic disease. We observed that targeting *DUX4* 3' key elements leads to an efficient extinction of *DUX4*, may redirect the mRNA cleavage site and prevents aberrant expression of genes downstream of the *DUX4* transcription factor.

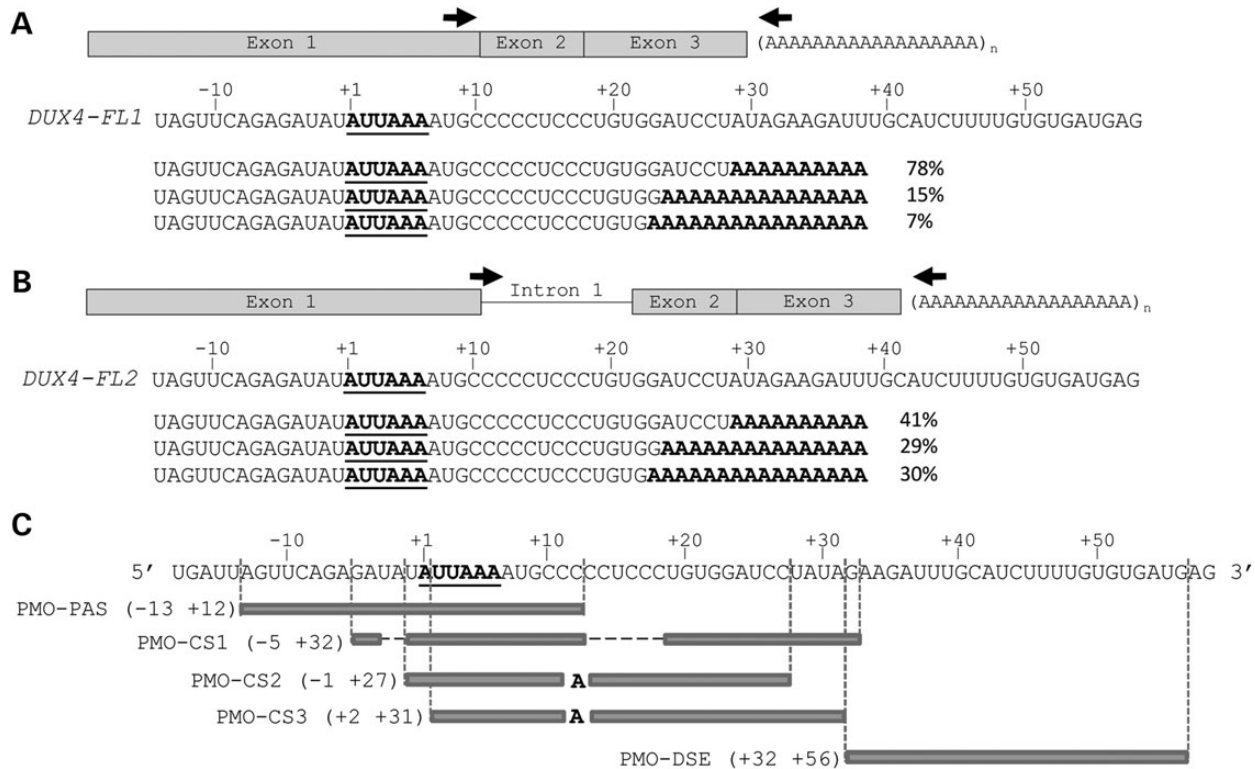
## Results

### Determination of 3' end key elements of *DUX4* mRNA and PMO design

For muscle tissue, one PAS (AUUAAA) has been described for *DUX4* mRNA, located 766 bp downstream of the stop codon in genomic DNA (16). Because the *DUX4* gene gives rise to two *DUX4*-FL isoforms (*DUX4*-FL1 with spliced intron 1 and *DUX4*-FL2 with non-spliced intron 1), the cleavage site of *DUX4* was precisely determined by RT-3' rapid amplification of cDNA-ends by polymerase chain reaction (3'RACE-PCR) using primers allowing the detection of each isoform. Total RNAs were extracted from FSHD myotubes at Day 4 of differentiation when *DUX4* expression is the highest (21). The sequences of the PCR products revealed the presence of three different cleavage sites located 16–22 bp after the PAS for both *DUX4*-FL1 and -FL2 *DUX4* isoforms (Fig. 1A and B). However, these cleavage sites are not used at the same frequencies. The cleavage site located 22 bases downstream of the PAS represents 78% of the *DUX4*-FL1 isoforms, but only 41% of *DUX4*-FL2 isoforms thus showing the interplay between mRNA cleavage site and splicing (Fig. 1A and B). In order to investigate the therapeutic potential of AONs targeting 3' key elements of *DUX4* mRNA, we designed AONs covering either the PAS (PMO-PAS), the cleavage sites and the polyA (PMO-CS1–3) or the U/GU-rich (PMO-DSE) sequence (Fig. 1C). Briefly, PMOs were designed to (i) contain little or no self-complementarity, (ii) form no >16 contiguous intra-strand hydrogen bonds, (iii) contain no >9 total guanines (>36% G) or >3 contiguous guanines to ensure good water solubility and (iv) where possible to target >1 of the 3' key elements with 1 PMO. For these reasons, the cleavage site PMOs contains either a single mismatch to target (PMO-CS2 and PMO-CS3) or do not target contiguous bases (PMO-CS1). These PMOs were of varying lengths (25–30 mers) (Table 1).

### PMOs induce a down expression of *DUX4* mRNA

The efficacy of each PMO was evaluated in a dose-dependent manner after transfection into differentiated immortalized FSHD clones. Total RNAs were extracted from myotubes and RT-PCR allowing the detection of either *DUX4*-all (Fig. 2A), or *DUX4*-FL1 and -FL2 (Fig. 2B) was performed. No modification of *DUX4*-all mRNA was observed with PMO-control compared to unlipofected cells thus showing that introduction of PMO-control does not modify *DUX4* expression. With PMO-CS1, -CS2 and -DSE, 27% ± 5, 25% ± 17 and 36% ± 10 reduction of *DUX4*-all



**Figure 1.** 3' end key elements of *DUX4* mRNA and PMO design. (A and B) To determine the cleavage sites of *DUX4* mRNA, a RT using an oligo-dT adapter primer was realized on total mRNAs isolated from FSHD myotubes at Day 4 of differentiation. The 3'RACE-PCR was performed using forward primer matching Exon 1 and reverse primer recognizing the adapter. The PCR products were cloned into a TOPO-TA plasmid and 27 and 17 clones were sequenced for *DUX4-FL1* (A) and -FL2 (B), respectively. The frequencies of the cleavage site usages are indicated. Arrows indicate the location of the primers used. The poly(A) site is bolded and underlined and the poly(A) tail is in bold. (C) Sequences recognized by the different PMOs targeting *DUX4*. The position +1 is defined as the first base of the poly(A) site. The sequences recognized by the PMOs are indicated as well as their position on the mRNA. Deletions are represented by hyphen (-). Nucleotidic substitutions are indicated.

**Table 1.** Summary of the characteristics of the PMOs

PMO name	PMO sequence 5'-3'	Leash sequence 5'-3'	Length (bp)	%GC
PMO-PAS (-13 +12)	GGGCATTTTAATATATCTCTGAACT	gattgGATATATTTAAAATGCCGtgat	25	0.32
PMO-CS1 (-5 +32)	CTATAGATCCACAGGGCATTTTAATATC	gttacCCCTGTGGATCCTATAGtgat	29	0.38
PMO-CS2 (-1 +27)	GGATCCACAGGGAGGAGGCATTTTAATA	gattgTCCTCCCTGTGGATCCgtgat	28	0.46
PMO-CS3 (+2 +31)	TATAGGATCCACAGGGAGGAGGCATTTTAA	gaatgTCCTGTGGATCCTATAgat	30	0.43
PMO-DSE (+32 +56)	CATCACACAAAAGATGCAAAATCTTC	gattgGCATCTTTTGTGTGATGtgat	25	0.36
PMO-control	CCTCTACCTCAGTTACAATTATA	gattgTAACTGAGGTAAGAGGtgat	25	0.32

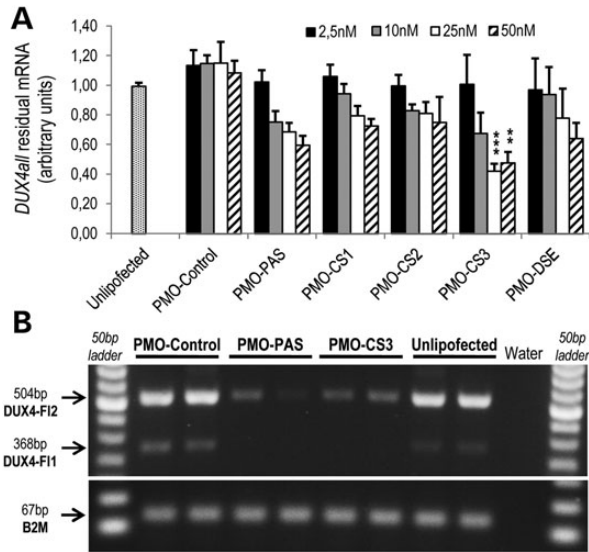
The locations, the PMO sequences and the leash sequences (uppercases indicate the complementary region between leash and PMO, lowercases indicates the free ends), the percentage of GC and the length are shown.

mRNA was observed, respectively, at the highest PMO concentration used compared with PMO-control (Fig. 2A). The highest efficiencies were obtained with the PMO-PAS and -CS3 for which a dose-dependent reduction of *DUX4-all* levels was observed with 40% ± 6 and 52% ± 7 inhibition at 50 nM, respectively (Fig. 2A). These two AOs were investigated further to ascertain if both *DUX4* isoforms were identically targeted by them. A decreased expression of both isoforms was observed at 50 nM when primers allowing the detection of both *DUX4-FL1* and *FL-2* [*DUX4-3'* untranslated region (3'UTR) oligonucleotides] were used (Fig. 2B).

#### PMO-CS3 induces a redirection of cleavage region

Because 22–54% of human mRNA have more than one poly(A) sites (4,5), and because redirections of the PAS have been described in the presence of AOs targeting the poly(A) of other mRNAs (44,45),

in silico analyses were performed which revealed the presence of putative alternative PAS downstream of Exon 3 (Fig. 3A). A redirection of the poly(A) and/or cleavage sites was thus investigated in the presence of the different PMOs at the highest concentration by 3'RACE nested PCR using forward primers located in Exon 3. A switch in cleavage site or poly(A) usage was not observed with any of the PMOs except PMO-CS3 (Fig. 3B). The sequence of this supplemental band revealed that the cleavage site of the residual *DUX4* mRNA in the presence of PMO-CS3 was ~40 nt upstream of the canonical cleavage site (Fig. 3C), thus suggesting that an alternative PAS was used to generate this new cleavage site. However, in silico analysis of the genomic region upstream of this alternative cleavage site (up to *DUX4* Exon 2) did not reveal the presence of any of the 13 PAS previously described (4). This new cleavage site is used in ~20% of the residual *DUX4* mRNAs.

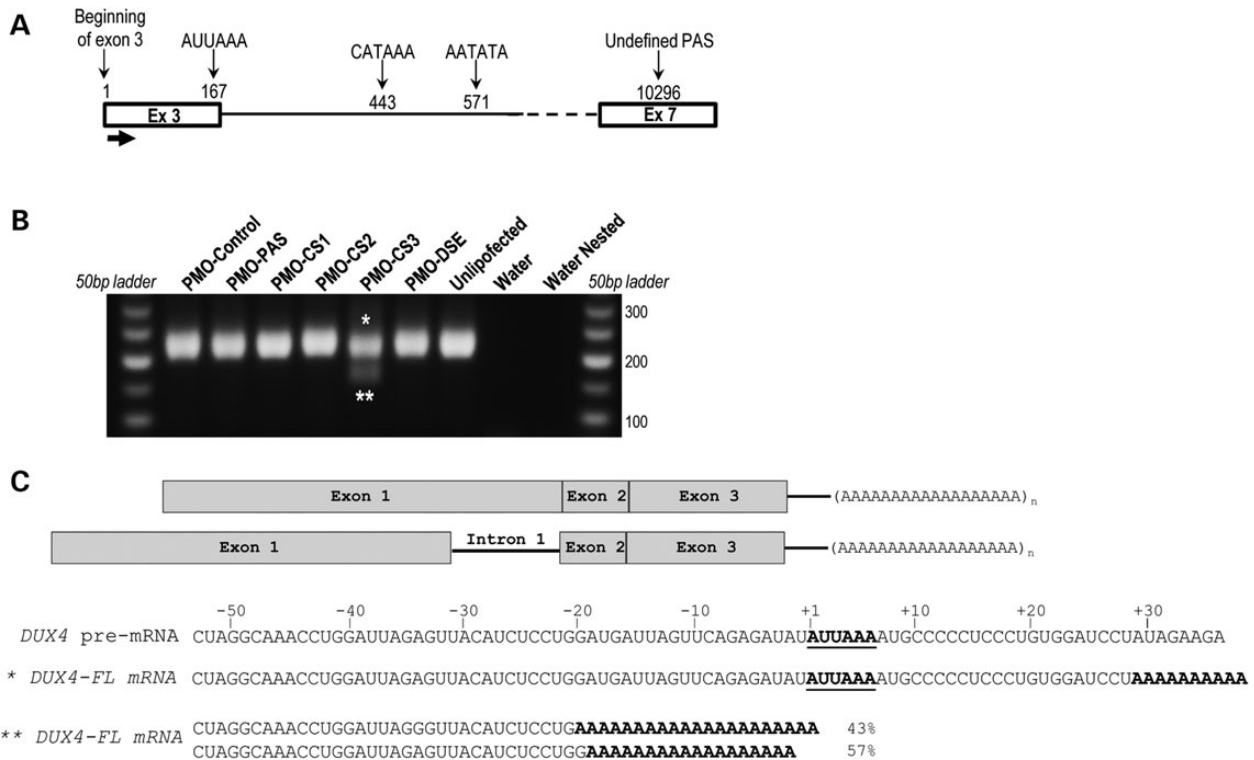


**Figure 2.** DUX4 expression in the presence of the PMOs. Immortalized FSHD myotubes were transfected with PMOs in a dose-dependent manner. (A) The percentage of residual DUX4-all mRNA was measured by semi-quantitative PCR using the ImageJ software on gel scans. Independent transfections for PMO-control, -PAS, -CS1, -CS2, CS3 and -DSE ( $n \geq 6$ ) were performed. Values are compared with unlipofected conditions. All data represent mean  $\pm$  standard error of mean (SEM). A MANOVA and a Newman-Keuls post hoc test were performed ( $^*P < 0.05$ ;  $^{**}P < 0.01$  and  $^{***}P < 0.001$ ). (B) Representative RT-PCR analysis using DUX4-3'UTR primers in the presence of PMO-control, -PAS and -CS3 at 50 nM. As a negative control, a PMO targeting a human  $\beta$ -globin mutation that causes  $\beta$ -thalassemia was used. B2M was used as the reference (housekeeping) gene.

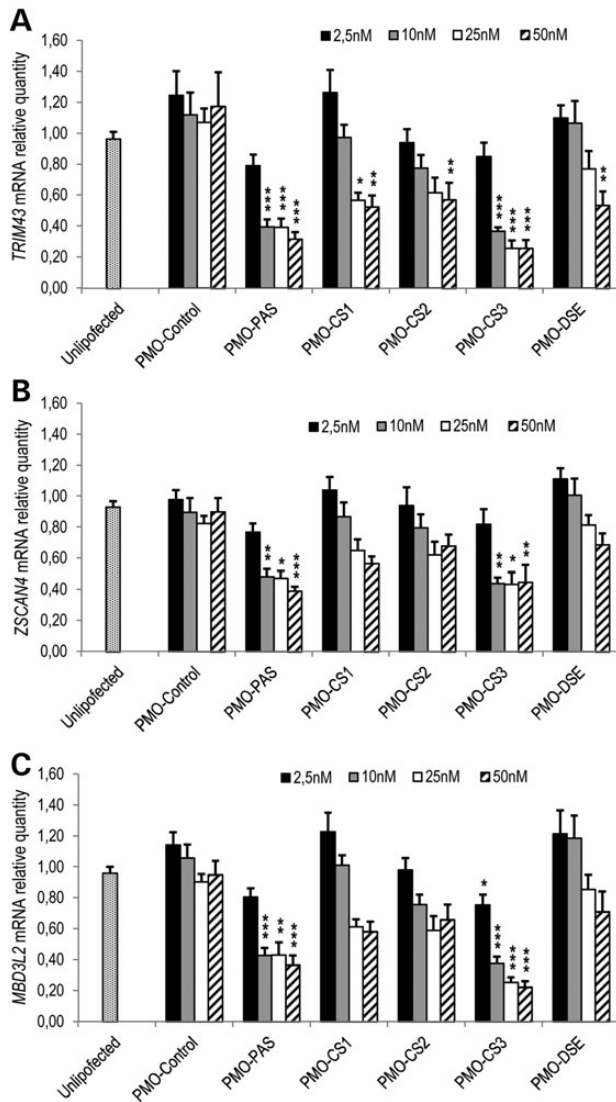
### PMO-PAS and -CS3 induce a knockdown in expression of genes downstream of DUX4 in FSHD cells

Because the DUX4 protein is very hard to detect, expression levels of DUX4 biomarkers were investigated. Indeed, DUX4 being a transcription factor, the expression levels of several genes downstream of DUX4 have been described to be modulated after DUX4 expression among them TRIM43, MBD3L2 and ZSCAN4 (21,27,46). The expression of these three genes was determined by RT-qPCR in the presence of the PMOs in a dose-dependent manner. No significant modulation of their expressions was observed in the presence of PMO-control compared with unlipofected condition. PMO-CS1, -CS2 and -DSE moderately modulate expression of genes downstream of DUX4 at the lowest concentrations used, but a down expression of TRIM43 is observed at the highest concentration. The best modulation of the genes downstream of DUX4 was observed in the presence of PMO-PAS and -CS3 in a dose-dependent manner (Fig. 4).

Since we have previously published that DUX4 is able to spread from one FSHD nucleus to neighboring nuclei thus inducing an aberrant activation of genes downstream of DUX4 in many nuclei (23), one possible explanation for the lower expression of TRIM43, MBD3L2 and ZSCAN4 observed in the presence of the PMO-PAS and -CS3 could have been an inhibition of myotube formation mediated by the PMOs. PMO-transfected cells were immunostained with the MF20 antibody recognizing all the myosin heavy chains (Fig. 5A) and fusion indexes were calculated by counting the number of nuclei in MF20-positive myotubes containing three or more nuclei as the percentage of the total number of nuclei (Fig. 5B). No statistically significant differences were observed with PMO treatment thus showing that the down-



**Figure 3.** PMO-CS3 induces a switch in cleavage site usage. A redirection of poly(A) usage was investigated in the presence of the different PMOs. (A) *In silico* analyses and literature (18) reveal the presence of APA site in the subtelomeric region of chromosome 4. (B) 3'RACE nested PCR using forward primers located in Exon 3 shows a switch in cleavage site only in the presence of PMO-CS3. The bands with (double asterisks) or without (single asterisk) a redirection of the cleavage site are indicated. (C) The sequence of the most abundant mRNA carrying the redirected cleavage site (DUX4 pre-mRNA). The sequence of poly(A) site is underlined and bolded. The poly(A) tail is in bold. The frequencies of each variant showing alternative cleavage site usage are indicated (14 analyzed sequences).

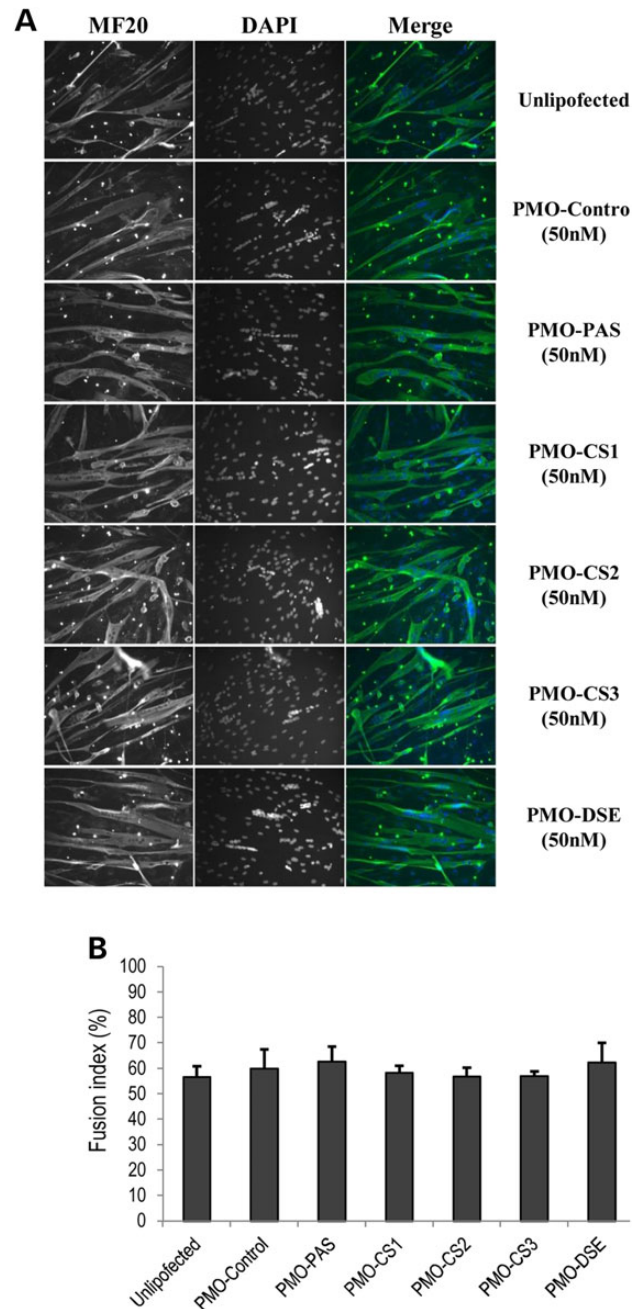


**Figure 4.** Expression of the genes downstream of DUX4 in the presence of the PMOs. Expression levels of TRIM43 (A), ZSCAN4 (B) and MBD3L2 (C) were measured by RT-qPCR in FSHD myotubes (4 days of differentiation) in the presence of the PMOs at different concentrations and compared with the PMO control at the same concentration. Values are compared with unlipofected conditions. B2M was used as a normalizer. Data represent mean  $\pm$  SEM from independent transfections for PMO-control, -PAS, -CS1, -CS2, CS3 and -DSE ( $n \geq 6$ ). A MANOVA and a Newman-Keuls post hoc test were performed (\* $P < 0.05$ ; \*\* $P < 0.01$  and \*\*\* $P < 0.001$ ).

regulation of the genes downstream of DUX4 in the presence of the PMOs was not due to myotube formation impairment, but rather to lowered DUX4 expression.

#### No significant off-targets are predicted

The potential off-target effects were also evaluated, PMOs targeting the DUX4 transcript were submitted to short-input-sequence basic alignment search tool (BLAST) searches against coding and non-coding human genome-nucleotide collection databases (nr/nr and nr/nt: see Table 2). According to BLAST a very strong match is considered to have  $E$ -value  $< 1 \times 10^{-4}$ , whereas a poor match has  $E$ -value  $> 1 \times 10^{-3}$ . For each PMO, the two BLAST returns showing highest alignment scores (lowest  $E$ -values) were



**Figure 5.** Fusion index of immortalized FSHD myotubes in the presence of the PMOs. Immortalized FSHD myotubes were transfected at Day 2 of differentiation with the different PMOs at 50 nM. At Day 4 of differentiation, cells were stained with MF20 antibody recognizing all the myosin heavy chain and counterstained with 4',6-diamidino-2-phenylindole (A). Fusion indexes were calculated by counting the number of nuclei in myotubes containing more than two nuclei as the percentage of the total number of nuclei (B). The transfections were performed six times for the unlipofected, four times for PMO-control, -PAS -DSE and two times for PMO-CS1, -CS2 and -CS3. In each experiment, 1000 nuclei were counted. All data represent mean  $\pm$  standard deviation.

selected for further analysis. With the exception of PMO-CS1, which contains specifically selected deletions, all designed PMOs returned a very strong match to the desired DUX4 transcript ( $E$ -value  $\leq 2 \times 10^{-5}$ ), whereas all of the off-target gene candidates show high  $E$ -values  $> 0.5$ . Predicted off-target sequence matches that showed homology on their negative or forward



**Table 2.** Prediction of the anti-DUX4 off target candidates

PMO name	Target name	E-value	% identity	No. of base overlap	Predicted homology with (+/-) strand	Coding/non-coding	Overall free binding energy ( $\Delta G_i$ )
PMO-PAS	DUX4	4.00E-05	100	25/25	-	Non-coding	-32.04
	IL1RAPL2	0.58	72	18/25	+	N.A.	N.A.
	GYS2	2.3	68	17/25	-	Non-coding	-16.62
PMO-CS1	C1ORF101	0.91	72	21/29	+	N.A.	N.A.
	DUX4	3.6	59	17/29	-	Non-coding	-21.82
	PTPRT	3.6	59	17/29	-	Non-coding	-6.60
PMO-CS2	DUX4	2.00E-05	96	27/28	-	Non-coding	-38.02
	OTOF	0.82	75	21/28	+	N.A.	N.A.
	POR	0.82	64	18/28	+	N.A.	N.A.
PMO-CS3	DUX4	2.00E-05	97	29/30	-	Non-coding	-37.93
	OTOF	0.99	70	21/30	+	N.A.	N.A.
	POR	0.99	60	18/30	+	N.A.	N.A.
PMO-DSE	DUX4	4.00E-05	100	25/25	-	Non-coding	-27.41
	TCRA/TC	0.58	72	18/25	+	N.A.	N.A.
	KCNJ6	2.3	68	17/25	+	N.A.	N.A.

N.A., non-applicable; IL1RAPL2, interleukin 1 receptor accessory protein-like 2; GYS2, glycogen synthase 2; C1ORF101, chromosome 1 open reading frame 101; PTPRT, protein tyrosine phosphatase receptor type T; OTOF, otoferlin; POR, P450 cytochrome oxidoreductase; TCRA/TCRD, T-cell receptor alpha-delta locus; KCNJ6, potassium channel inwardly rectifying subfamily J member.

Each PMO sequence was run through the BLAST software, optimized for short-input sequences and limited to human and nt/nr database. Overall the binding energy was calculated only for the predicted targets that show sequence homology with the PMO on their negative strand (-). The units of  $\Delta G_i$  are expressed in kcal/mol.

strand, corresponding to RNA transcript, were selected for the calculation of total free energy of binding using the RNAup Server (Table 2; Supplementary Material, Fig. S1). The only potential off-target interactions on the RNA transcript (i.e. negative strand) was GYS2 for PMO-PAS and PTPRT for PMO-CS1, but overall free energies of binding were low (-16.62 and -6.60, respectively), and with major mismatches (8/25 and 12/29, respectively). These data suggest a low probability of interference of any of the designed PMOs targeting DUX4 with off target RNAs.

## Discussion

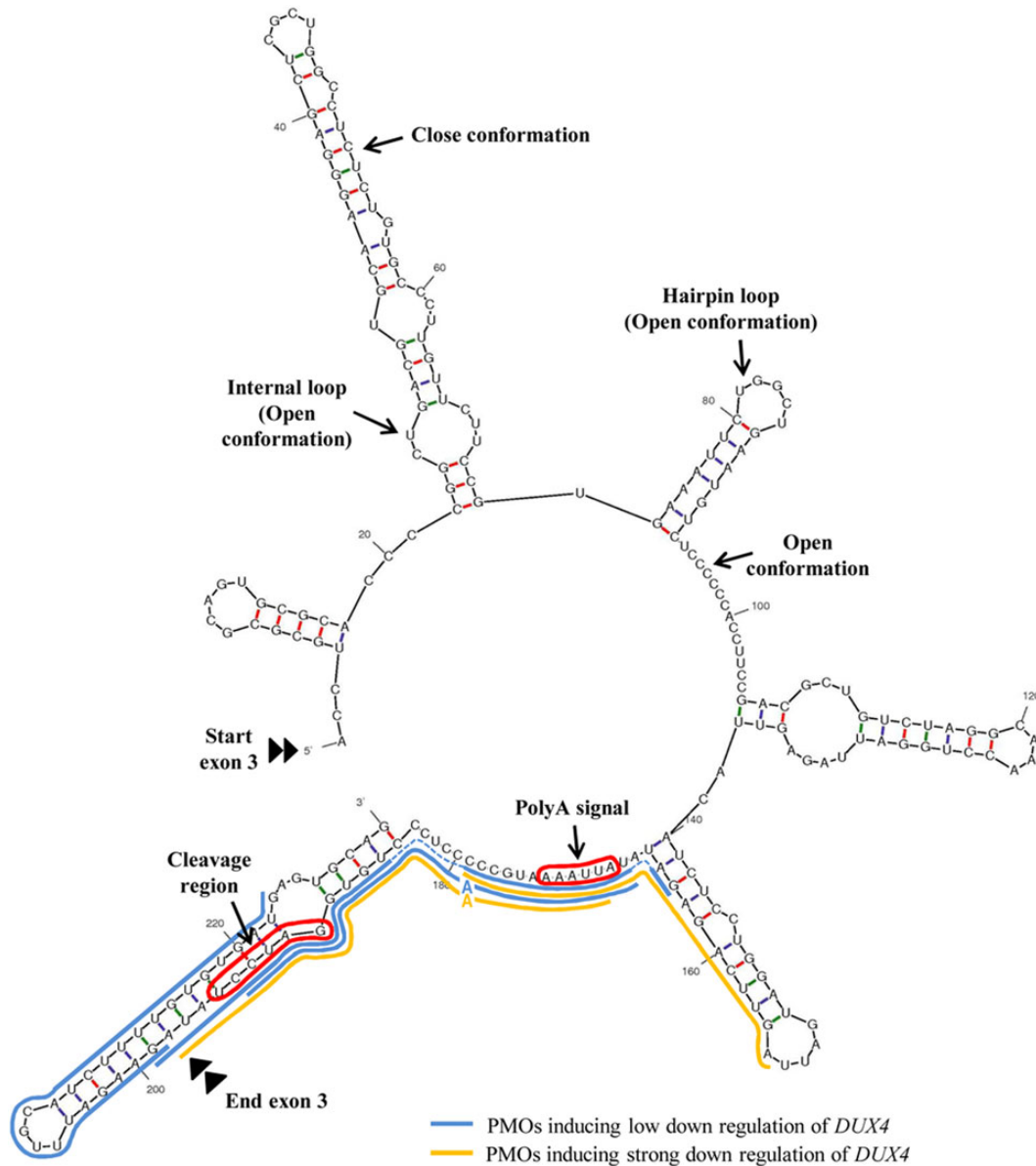
Several strategies using AOs have been described during the last decade. Upon binding, AOs can alter the original function of the targeted mRNA through an array of different mechanisms. In the literature, AOs have been described to (i) prevent 5' cap formation, (ii) bind the pre-mRNA to modulate splicing by masking the splicing sequences to force exon inclusion or exclusion, (iii) induce mRNA degradation after the recognition of the duplex AO-mRNA (depending on the chemical modification used), (iv) inhibit ribosome access to the mRNA leading to the suppression of protein translation and (v) modulate polyadenylation selection in transcripts with more than one poly(A) in their 3'UTR (39). However, to our knowledge, the possibility to target the 3'UTR key elements of the mammalian pre-mRNA to inhibit the polyadenylation in genetic diseases has not been studied.

As previously described, defects in mRNA 3' end formation can result in disruption of gene expression processes in a number of ways, including: transcription termination; transport of the mRNA from the nucleus to the cytoplasm; stability of the mRNA; translation efficiency and the inhibition of polyadenylation, which in turn might lead to gene silencing (47). In this study, we investigate the possibility of interference in the 3' end mRNA processing as a therapeutic approach in the FSHD model. Since FSHD is a gain-of-function disease, characterized by the aberrant over-expression of the DUX4 gene, the therapeutic strategies should therefore be based on gene silencing rather than gene rescue. Here, we show that targeting the 3' end elements

involved in mRNA processing is an effective way to decrease DUX4 expression and prevent aberrant expression of genes downstream of DUX4.

Among the five AONs tested, only two of them lead to a significant decrease of DUX4 mRNA levels under the conditions evaluated. PMOs length could play an important role in their efficacies, since PMO-CS3 is the only 30-mer used in this study and is the AO showing the best efficiency. The number of mismatches could also play a role affecting their relative activity when compared with a completely homologous sequence of equal length. Secondary structures may also have affected effectiveness of the PMOs and secondary structures of Exon 3 of DUX4 or the AON were analyzed *in silico* using thermodynamic methods. Indeed, the probability of interaction between the AON and the targeted mRNA can be impacted by the presence of closed or opened conformations at the site of hybridization (47). When calculating the minimum free energy of a folded mRNA molecule using mfold (<http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form2.3>, date last accessed, January 13, 2016) (48,49), the Exon 3 of DUX4 was predicted to form highly stable secondary structure with free energies of -56.3 kcal/mol (Fig. 6). The resulting outputs show that DUX4 poly(A) is located within the region of an open conformation, whereas the cleavage site and the DSE lay within a closed region. The PAS, thus, appears more accessible than the DSE and this could explain the low efficacy of the PMO-DSE which totally overlap a closed region.

Different mechanisms may be involved in the down-regulation of the targeted mRNA in the presence of the PMOs. The first one is that the duplex formation of the PMO with the mRNA 3'UTR could prevent poly(A) tail elongation. The second one could be the removal of the entire tail mediated by a deadenylase-dependent cutting of the poly(A) tail as previously described in zebrafish (50), thus affecting mRNA stability. Another possibility is that spatial rearrangements of the 3'UTR in the presence of the PMO could have modulated the nonsense mediated decay (NMD) pathway (51). On the one hand, it was previously shown that NMD is an endogenous suppressor of DUX4 mRNA that contributes to the low expression of DUX4 in FSHD cells



**Figure 6.** Secondary structure predictions of *DUX4* mRNA. Mfold analysis was performed on the 3' end of *DUX4* (starting first nucleotide of Exon 3 and ending 230 nucleotides downstream). The PAS and the cleavage region are boxed in red. The positions of the sequences recognized by the PMOs are in yellow for the PMOs inducing a strong down-regulation of *DUX4*, whereas the others are in blue. Examples of opened and closed RNA secondary structures are indicated.

(30). But on the other hand, *DUX4* expression was also described to inhibit NMD by decreasing Up-frameshift protein 1 (UPF-1) levels (30). The PMO-*DUX4* duplex formation, causing a destabilization of *DUX4* mRNA by at least one of the mechanisms described earlier, could have thus increased NMD activity, enhancing *DUX4* repression.

Recent discoveries have revealed that many human genes use more than one poly(A) and cleavage sites, indicating that alternative polyadenylation (APA) is a widespread phenomenon. The use of APA has been described to play important roles in several human diseases such as (i) in oculopharyngeal muscular dystrophy, in which mutations in PolyA binding protein nuclear 1 (*PABPN1*) result in a global enhancement of the usage of proximal cleavage sites, leading to an unbalanced formation of alternative mRNA 3' ends and to a modification of gene expression through aberrant microRNA regulation (52); (ii) in  $\beta$ -thalassemia, in which

a loss of function of globin 3' end processing linked to an alteration of the poly(A) site is one of the causes of this disease, resulting in the generation of an unstable transcript that is ~900 nt longer (1,7); (iii) a single-nucleotide polymorphism in the U/GU rich region downstream of an alternative PAS in the fibrinogen gamma haplotype 2 gene has been shown to affect the usage of this poly(A) and has been associated with an increased risk of deep-venous thrombosis (53). Considering that the use of AOs targeting the poly(A) site can lead to the use of APA and cleavage site (44,45), one important point raised in our experiments was to determine if such a redirection was observed in the presence of the PMOs. Although in the presence of PMO-CS3 the cleavage site was detected to be redirected by ~40 nt upstream of its normal position, it was still considerably distal to any of the previously described alternative PASs (4). Interestingly, this finding does not fit with the current consensus describing that there

Table 3. Primers used in this study

	Targeted gene	Primer	Accession number	Sequence	Location	Size (bp)	qPCR efficiency (calculated from slope)
RT-oligonucleotides		Oligo_dT_3'RACE	N.A.	GCTGTCAACGATACGCTACGTAACGGCATGA CAGTGTTTTTTTTTTTTTTTTTTTTTTTTTT	N.A.	N.A.	N.A.
		Oligo_NVT	N.A.	GCGAGCTCCGCGCCGCGTTTTTTTTTTTVN	N.A.	N.A.	N.A.
PCR oligonucleotides	<b>DUX4-all</b>	DUX4-all_fw	ENSG00000260596	CCCAGGTACCAGCAGACC	Spanning Exon 2-Exon 3	164 bp	N.A.
		DUX4-all_rev		TCCAGGAGATGTAACCTCTAATCCA	Exon 3		
	<b>DUX4-3'UTR</b>	DUX4-3'UTR_fw	HQ266760 ( <i>DUX4-fl1</i> ) and	AGGCGCAACCTCTCCTAGAAAC	Exon 1	368 bp and	N.A.
		DUX4-3'UTR_rev	HQ266761 ( <i>DUX4-fl2</i> )	TCCAGGAGATGTAACCTCTAATCCA	Exon 3	504 bp	
qPCR oligonucleotides	<b>B2M</b>	B2M_fw	NM_004048.2	CTCTCTTTCTGGCCTGGAGG	Exon 1	67 bp	100%
		B2M_rev		TGCTGGATGACGTGAGTAAACC	Exon 2		
	<b>MBD3L2</b>	MBD3L2_fw	NM_144614.3	CGTTCACCTCTTTTCCAAGC	Exon 1	142 bp	106%
		MBD3L2_rev		AGTCTCATGGGAGAGCAGA	Exon 2		
	<b>TRIM 43</b>	TRIM43_fw	NM_138800.1	ACCCATCACTGGACTGGTGT	Exon 6	100 bp	109%
		TRIM43_rev		CACATCCTCAAAGAGCCTGA	Exon 7		
	<b>ZSCAN4</b>	ZSCAN4_962U20_fw	NM_152677.2	CTGGAGCAGTTTATGATTGG	Exon 3	162 bp	98%
		ZSCAN4_rev		AGCTTCCTGTCCCTGCATGT	Exon 4		
3'RACE-PCR oligonucleotides		3RACE_3'_primer_rev	N.A.	GCTGTCAACGATACGCTACGTAACG	N.A.	N.A.	N.A.
		3RACE_3'_nested_rev	N.A.	CGCTACGTAACGGCATGACAGTG	N.A.	N.A.	N.A.
	<b>DUX4</b>	3RACE_Exon1_fw	ENSG00000260596	CTCTCCTAGAAACGGAGGCCCCCG	Exon 1	N.A.	N.A.
		3RACE_Exon1_nested_fw		CTCGTGGAAGCACCCCTCAGC	Exon 1	N.A.	N.A.
		3RACE_3'_Exon3_fw	ENSG00000260596	CGCACCCCGGCTGACGTGCAAG	Exon 3	N.A.	N.A.
		3RACE_3'_Exon3_nested_fw		CGCTGGCCTCTCTGTGCCCTTG	Exon 3	N.A.	N.A.

N.A., non-applicable.

are four general classes of APA, including: tandem 3'UTR APA; alternative terminal exon APA; intronic APA and internal exon APA (1). This could be explained by the steric congestion at the PAS and cleavage site provided by the PMO-CS3 and the cellular proteins involved in the polyadenylation process, respectively. The cleavage factors would be then forced to cleave the pre-mRNA upstream of the PAS, resulting in a new cleavage site, but still utilizing the original poly(A) site. Importantly, in the 14 sequences analyzed, the same cleavage region was observed, strongly suggesting a specificity to the target sequence. Alternatively, a novel APA was used that has not been previously described (4).

A redirection of the cleavage site associated to the alternative poly(A) site located in Exon 7 of *DUX4* was not observed either. It was shown that chromosomes 4 and 10 produce *DUX4* mRNA in human testes (18). These transcripts can use alternative 3' exons with a PAS in Exon 7, leading to a *DUX4* mRNA carrying Exons 1, 2, 6 and 7 or Exons 1, 2, 4–7. Exons 3 and 7, thus, seem to be mutually exclusive. It is, therefore, not surprising to not observe a redirection of the cleavage site to Exon 7 when Exon 3 is targeted.

Targeting the 3' end elements of mRNA could thus represent an alternative approach for gene silencing compared with classic AO approaches leading to exon skipping, translation inhibition or mRNA destruction into the cytoplasm. The strategy to interfere with the mRNA 3' end processing provides a number of advantages. First, polyadenylation is a crucial step required for the maturation of almost all eukaryotic mRNAs (except for histones), meaning that virtually any mRNA can be subjected to this strategy. Secondly, genes with only 1 exon can be targeted using AON hybridizing the 3' end key elements. Thirdly, AON targeting the PASs may be investigated in a therapeutic context for some disease caused by utilization of inadequate APAs. For example, some mantle cell lymphoma patients express isoforms of *cyclinD1* mRNA with short 3'UTR compared with healthy individuals. Shortening of the 3'UTR is caused by the mutations in the *CCND1* gene leading to the creation of a premature PAS. This mutated mRNA lacks most of the 3'UTR region that contains mRNA destabilizing elements, making the *CCND1* mRNA abnormally stable (54). Therefore, targeting a premature poly(A) with AON to promote the use of the distal poly(A) may also serve as an promising approach to treat disease.

## Materials and Methods

PMO were manufactured and supplied by Gene Tools (LLC, Philomath, OR, USA). The DNA leashes for PMO transfection were synthesized by Eurogentec and their sequences are indicated in Table 1. Before transfection, the PMOs (2.5  $\mu$ l at 1 mM) were annealed with the leash (25  $\mu$ l at 100  $\mu$ M) in the final volume of 50  $\mu$ l (12.5  $\mu$ l of PBS10 $\times$  and 10  $\mu$ l of RNase-free water) (final concentration of leashed PMO 50  $\mu$ M) using the following protocol: 95°C for 5 min, 85°C for 1 min, 75°C for 1 min, 65°C for 5 min, 55°C for 1 min, 45°C for 1 min, 35°C for 5 min, 25°C for 1 min and then held at 15°C. Efficient annealing was verified using gel electrophoresis band shift as previously described (55). Leashed PMOs were stored at 4°C.

Immortalized FSHD clones have been characterized previously (56). Cells were cultivated in the proliferation medium [4 vols of Dulbecco's modified Eagle medium (DMEM), 1 vol of 199 medium, fetal bovine serum 20%, gentamycin 50  $\mu$ g/ml (Life Technologies, Saint Aubin, France)] supplemented with insulin 5  $\mu$ g/ml, dexamethasone 0.2  $\mu$ g/ml,  $\beta$ -fibroblast growth factor ( $\beta$ -FGF) 0.5 ng/ml, human epidermal growth factor (hEGF) 5 ng/ml and fetuine 25  $\mu$ g/ml. The differentiation was induced by replacing the proliferation medium by DMEM supplemented with insulin (10  $\mu$ g/ml). Two days after the induction of differentiation, the transfections

were realized using Lipofectamine<sup>®</sup> RNAiMax Reagent (Life Technologies) according to the manufacturer's instructions. Cells were harvested at Day 4 of differentiation.

RNA extraction from cultured cells was performed using Trizol according the manufacturer's protocol (Life Technologies, Saint Aubin, France). RNA concentration was determined using a nanodrop ND-1000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA). Reverse transcription was performed on 1  $\mu$ g of total RNA in a 10  $\mu$ l final volume (Roche Transcriptor First Strand cDNA Synthesis kit, Roche, Meylan, France) with a NVT primer (according to the International Union of Pure and Applied Chemistry, N = any base; V = adenine or cytosine or guanine and T = thymine) (GCGAGCTCCGCGCCGCGTTTTTTTTTTT) (10  $\mu$ M) according to the manufacturer instructions. The PCR for *DUX4*-all was performed on 1  $\mu$ l of cDNA using 2X Reddy Mix PCR Master Mix (Thermo Scientific) according the manufacturer's instructions. The PCR program was the following: 94°C for 5 min, followed by 35 cycles at 94°C for 20 s and 60°C for 20 s and 72°C for 20 s, finished with 72°C for 7 min. Quantitative PCR (qPCR) was designed according to the minimum of information for publication of quantitative (MIQE) standards (57). The primers used for both PCR and qPCR are indicated in Table 3. Quantitative PCRs were performed in a final volume of 9  $\mu$ l with 4  $\mu$ l of RT product, 0.18  $\mu$ l of each forward and reverse primers at 20  $\mu$ M, and 4.5  $\mu$ l of SYBR<sup>®</sup> Green mastermix 2 $\times$  (Roche, Meylan, France) in a lightcycler LC480 (Roche). The qPCR cycling conditions were 94°C for 5 min, followed by 50 cycles at 94°C for 10 s and 60°C for 5 s and 72°C for 5 s. All the amplicons were validated using sequencing. *B2M* was the most appropriate gene for normalization: a multi-parametric analysis of variance (MANOVA) and a Newman-Keuls *post hoc* test did not reveal any correlation between *B2M* expression and PMO addition (at different concentrations).

For the 3'RACE, the RT was realized with an oligo-dT adapter primer: (GCTGTCAACGATACGCTACGTAACGGCATGACAGTGT TTTTTTTTTTTTTTTTTTTTTT) as described in Table 3. The primary PCRs were performed in a final volume of 15  $\mu$ l with 1  $\mu$ l of RT product, 3  $\mu$ l of reverse primer at 20  $\mu$ M (Table 3) and 1  $\mu$ l of primer at 20  $\mu$ M, 7.5  $\mu$ l 2X Reddy Mix PCR Master Mix (Thermo Scientific). The PCR cycling conditions were 94°C for 5 min, followed by 5 cycles at 94°C for 20 s and 72°C for 50 s, then 5 cycles at 94°C for 20 s and 70°C for 30 s and 72°C for 20 s and 25 cycles at 94°C for 20 s and 68°C for 20 s and 72°C for 30 s, finished with 72°C for 7 min. The nested PCRs were realized in a final volume of 15  $\mu$ l with 1  $\mu$ l of primary PCR product, 1  $\mu$ l of reverse and/or forward primers 20  $\mu$ M, 7.5  $\mu$ l 2X Reddy Mix PCR Master Mix (Thermo Scientific). The PCR cycling conditions were 94°C for 5 min, followed by 35 cycles at 94°C for 20 s and 60°C for 20 s and 72°C for 30 s, finished with 72°C for 7 min. Nested PCR products were purified according to NucleoSpin Gel and PCR Clean-Up manufacturer's instructions (Macherey-Nagel). Fragments were cloned into TOPO-TA vector with TOPO-TA Cloning kit (Life Technologies). Cleavage sites were determined by sequencing at least 10 colony-forming units.

## Supplementary Material

Supplementary Material is available at HMG online.

*Conflict of Interest statement.* None declared.

## Funding

This study was financially supported by the French Association against Myopathies (AFM-Téléthon, France) and the National

Research Agency (FSHDecipher, ANR-13-BSV1-0004) and The Rosetrees Trust, UK.

## References

- Elkon, R., Ugalde, A.P. and Agami, R. (2013) Alternative cleavage and polyadenylation: extent, regulation and function. *Nat. Rev.*, **14**, 496–506.
- Nunes, N.M., Li, W., Tian, B. and Furger, A. (2010) A functional human Poly(A) site requires only a potent DSE and an A-rich upstream sequence. *EMBO J.*, **29**, 1523–1536.
- Proudfoot, N.J. (2011) Ending the message: poly(A) signals then and now. *Genes Dev.*, **25**, 1770–1782.
- Tian, B., Hu, J., Zhang, H. and Lutz, C.S. (2005) A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res.*, **33**, 201–212.
- Beaudoing, E., Freier, S., Wyatt, J.R., Claverie, J.M. and Gautheret, D. (2000) Patterns of variant polyadenylation signal usage in human genes. *Genome Res.*, **10**, 1001–1010.
- Gruber, A.R., Martin, G., Keller, W. and Zavolan, M. (2014) Means to an end: mechanisms of alternative polyadenylation of messenger RNA precursors. *Wiley Interdiscip. Rev.*, **5**, 183–196.
- Danckwardt, S., Hentze, M.W. and Kulozik, A.E. (2008) 3' end mRNA processing: molecular mechanisms and implications for health and disease. *EMBO J.*, **27**, 482–498.
- Sachs, A. (1990) The role of poly(A) in the translation and stability of mRNA. *Curr. Opin. Cell Biol.*, **2**, 1092–1098.
- Hollerer, I., Grund, K., Hentze, M.W. and Kulozik, A.E. (2014) mRNA 3' end processing: A tale of the tail reaches the clinic. *EMBO Mol. Med.*, **6**, 16–26.
- Tawil, R., van der Maarel, S.M. and Tapscott, S.J. (2014) Facioscapulohumeral dystrophy: the path to consensus on pathophysiology. *Skeletal Muscle*, **4**, 12.
- van der Maarel, S.M., Miller, D.G., Tawil, R., Filippova, G.N. and Tapscott, S.J. (2012) Facioscapulohumeral muscular dystrophy: consequences of chromatin relaxation. *Curr. Opin. Neurol.*, **25**, 614–620.
- van Deutekom, J.C., Wijmenga, C., van Tienhoven, E.A., Gruter, A.M., Hewitt, J.E., Padberg, G.W., van Ommen, G.J., Hofker, M.H. and Frants, R.R. (1993) FSHD associated DNA rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit. *Hum. Mol. Genet.*, **2**, 2037–2042.
- Wijmenga, C., Hewitt, J.E., Sandkuijl, L.A., Clark, L.N., Wright, T.J., Dauwerse, H.G., Gruter, A.M., Hofker, M.H., Moerer, P., Williamson, R. et al. (1992) Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. *Nat. Genet.*, **2**, 26–30.
- Lemmers, R.J., Tawil, R., Petek, L.M., Balog, J., Block, G.J., Santen, G.W., Amell, A.M., van der Vliet, P.J., Almomani, R., Straasheijm, K.R. et al. (2012) Digenic inheritance of an SMCHD1 mutation and an FSHD-permissive D4Z4 allele causes facioscapulohumeral muscular dystrophy type 2. *Nat. Genet.*, **44**, 1370–1374.
- Lemmers, R.J., van den Boogaard, M.L., van der Vliet, P.J., Donlin-Smith, C.M., Nations, S.P., Ruivenkamp, C.A., Heard, P., Bakker, B., Tapscott, S., Cody, J.D. et al. (2015) Hemizyosity for SMCHD1 in facioscapulohumeral muscular dystrophy Type 2: consequences for 18p deletion syndrome. *Hum. Mutat.*, **36**, 679–683.
- Lemmers, R.J., van der Vliet, P.J., Klooster, R., Sacconi, S., Camano, P., Dauwerse, J.G., Snider, L., Straasheijm, K.R., van Ommen, G.J., Padberg, G.W. et al. (2010) A unifying genetic model for facioscapulohumeral muscular dystrophy. *Science* **329**, 1650–1653.
- Snider, L., Asawachaicharn, A., Tyler, A.E., Geng, L.N., Petek, L.M., Maves, L., Miller, D.G., Lemmers, R.J., Winokur, S.T., Tawil, R. et al. (2009) RNA transcripts, miRNA-sized fragments and proteins produced from D4Z4 units: new candidates for the pathophysiology of facioscapulohumeral dystrophy. *Hum. Mol. Genet.*, **18**, 2414–2430.
- Snider, L., Geng, L.N., Lemmers, R.J., Kyba, M., Ware, C.B., Nelson, A.M., Tawil, R., Filippova, G.N., van der Maarel, S.M., Tapscott, S.J. et al. (2010) Facioscapulohumeral dystrophy: incomplete suppression of a retrotransposed gene. *PLoS Genet.*, **6**, e1001181.
- Jones, T.I., Chen, J.C., Rahimov, F., Homma, S., Arashiro, P., Beermann, M.L., King, O.D., Miller, J.B., Kunkel, L.M., Emerson, C.P. Jr. et al. (2012) Facioscapulohumeral muscular dystrophy family studies of DUX4 expression: evidence for disease modifiers and a quantitative model of pathogenesis. *Hum. Mol. Genet.*, **21**, 4419–4430.
- Broucqsaault, N., Morere, J., Gaillard, M.C., Dumonceaux, J., Torrents, J., Salort-Campana, E., Maues de Paula, A., Bartoli, M., Fernandez, C., Chesnais, A.L. et al. (2013) Dysregulation of 4q35- and muscle-specific genes in fetuses with a short D4Z4 array linked to Facio-Scapulo-Humeral Dystrophy. *Hum. Mol. Genet.*, **22**, 4206–4214.
- Ferreboeuf, M., Mariot, V., Bessieres, B., Vasiljevic, A., Attie-Bitach, T., Collardeau, S., Morere, J., Roche, S., Magdinier, F., Robin-Ducellier, J. et al. (2014) DUX4 and DUX4 downstream target genes are expressed in fetal FSHD muscles. *Hum. Mol. Genet.*, **23**, 171–181.
- Block, G.J., Narayanan, D., Amell, A.M., Petek, L.M., Davidson, K.C., Bird, T.D., Tawil, R., Moon, R.T. and Miller, D.G. (2013) Wnt/beta-catenin signaling suppresses DUX4 expression and prevents apoptosis of FSHD muscle cells. *Hum. Mol. Genet.*, **22**, 390–396.
- Ferreboeuf, M., Mariot, V., Furling, D., Butler-Browne, G., Mouly, V. and Dumonceaux, J. (2014) Nuclear protein spreading: implication for pathophysiology of neuromuscular diseases. *Hum. Mol. Genet.*, **23**, 4125–4133.
- Kowaljow, V., Marcowycz, A., Anseau, E., Conde, C.B., Sauvage, S., Matteotti, C., Arias, C., Corona, E.D., Nunez, N.G., Leo, O. et al. (2007) The DUX4 gene at the FSHD1A locus encodes a proapoptotic protein. *Neuromuscul. Disord.*, **17**, 611–623.
- Dixit, M., Anseau, E., Tassin, A., Winokur, S., Shi, R., Qian, H., Sauvage, S., Matteotti, C., van Acker, A.M., Leo, O. et al. (2007) DUX4, a candidate gene of facioscapulohumeral muscular dystrophy, encodes a transcriptional activator of PITX1. *Proc. Natl Acad. Sci. USA*, **104**, 18157–18162.
- Wallace, L.M., Garwick, S.E., Mei, W., Belayew, A., Coppee, F., Ladner, K.J., Guttridge, D., Yang, J. and Harper, S.Q. (2011) DUX4, a candidate gene for facioscapulohumeral muscular dystrophy, causes p53-dependent myopathy in vivo. *Ann. Neurol.*, **69**, 540–552.
- Geng, L.N., Yao, Z., Snider, L., Fong, A.P., Cech, J.N., Young, J. M., van der Maarel, S.M., Ruzzo, W.L., Gentleman, R.C., Tawil, R. et al. (2012) DUX4 activates germline genes, retroelements, and immune mediators: implications for facioscapulohumeral dystrophy. *Dev. Cell*, **22**, 38–51.
- Xu, H., Wang, Z., Jin, S., Hao, H., Zheng, L., Zhou, B., Zhang, W., Lv, H. and Yuan, Y. (2014) Dux4 induces cell cycle arrest at G1 phase through upregulation of p21 expression. *Biochem. Biophys. Res. Commun.*, **446**, 235–240.
- Rickard, A.M., Petek, L.M. and Miller, D.G. (2015) Endogenous DUX4 expression in FSHD myotubes is sufficient to cause cell

- death and disrupts RNA splicing and cell migration pathways. *Hum. Mol. Genet.*, **24**, 5901–5914.
30. Feng, Q., Snider, L., Jagannathan, S., Tawil, R., van der Maarel, S.M., Tapscott, S.J. and Bradley, R.K. (2015) A feedback loop between nonsense-mediated decay and the retrogene DUX4 in facioscapulohumeral muscular dystrophy. *eLife*, **4**, doi: 10.7554/eLife.04996.
  31. Tupler, R., Berardinelli, A., Barbierato, L., Frants, R., Hewitt, J. E., Lanzi, G., Maraschio, P. and Tiepolo, L. (1996) Monosomy of distal 4q does not cause facioscapulohumeral muscular dystrophy. *J. Med. Genet.*, **33**, 366–370.
  32. Lemmers, R.J., de Kievit, P., Sandkuijl, L., Padberg, G.W., van Ommen, G.J., Frants, R.R. and van der Maarel, S.M. (2002) Facioscapulohumeral muscular dystrophy is uniquely associated with one of the two variants of the 4q subtelomere. *Nat. Genet.*, **32**, 235–236.
  33. Thomas, N.S., Wiseman, K., Spurlock, G., MacDonald, M., Ustek, D. and Upadhyaya, M. (2007) A large patient study confirming that facioscapulohumeral muscular dystrophy (FSHD) disease expression is almost exclusively associated with an FSHD locus located on a 4qA-defined 4qter subtelomere. *J. Med. Genet.*, **44**, 215–218.
  34. Yao, Z., Snider, L., Balog, J., Lemmers, R.J., Van Der Maarel, S. M., Tawil, R. and Tapscott, S.J. (2014) DUX4-induced gene expression is the major molecular signature in FSHD skeletal muscle. *Hum. Mol. Genet.*, **23**, 5342–5352.
  35. Wallace, L.M., Liu, J., Domire, J.S., Garwick-Coppens, S.E., Guckes, S.M., Mendell, J.R., Flanigan, K.M. and Harper, S.Q. (2012) RNA interference inhibits DUX4-induced muscle toxicity in vivo: implications for a targeted FSHD Therapy. *Mol. Ther.*, **20**, 1417–1423.
  36. Mitsuhashi, H., Mitsuhashi, S., Lynn-Jones, T., Kawahara, G. and Kunkel, L.M. (2013) Expression of DUX4 in zebrafish development recapitulates facioscapulohumeral muscular dystrophy. *Hum. Mol. Genet.*, **22**, 568–577.
  37. Vanderplanck, C., Anseau, E., Charron, S., Stricwant, N., Tassin, A., Laoudj-Chenivresse, D., Wilton, S.D., Coppee, F. and Beleyew, A. (2011) The FSHD atrophic myotube phenotype is caused by DUX4 expression. *PLoS One*, **6**, e26820.
  38. Kole, R., Krainer, A.R. and Altman, S. (2012) RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nat. Rev. Drug Discov.*, **11**, 125–140.
  39. DeVos, S.L. and Miller, T.M. (2013) Antisense oligonucleotides: treating neurodegeneration at the level of RNA. *Neurotherapeutics*, **10**, 486–497.
  40. Cirak, S., Arechavala-Gomez, V., Guglieri, M., Feng, L., Torelli, S., Anthony, K., Abbs, S., Garralda, M.E., Bourke, J., Wells, D.J. et al. (2011) Exon skipping and dystrophin restoration in patients with Duchenne muscular dystrophy after systemic phosphorodiamidate morpholino oligomer treatment: an open-label, phase 2, dose-escalation study. *Lancet*, **378**, 595–605.
  41. Mendell, J.R., Rodino-Klapac, L.R., Sahenk, Z., Roush, K., Bird, L., Lowes, L.P., Alfano, L., Gomez, A.M., Lewis, S., Kota, J. et al. (2013) Eteplirsen for the treatment of Duchenne muscular dystrophy. *Ann. Neurol.*, **74**, 637–647.
  42. Jarmin, S., Kymalainen, H., Popplewell, L. and Dickson, G. (2013) New developments in the use of gene therapy to treat Duchenne muscular dystrophy. *Expert Opin. Biol. Ther.*, **14**, 209–230.
  43. Voit, T., Topaloglu, H., Straub, V., Muntoni, F., Deconinck, N., Campion, G., De Kimpe, S.J., Eagle, M., Guglieri, M., Hood, S. et al. (2014) Safety and efficacy of drisapersen for the treatment of Duchenne muscular dystrophy (DEMAND II): an exploratory, randomised, placebo-controlled phase 2 study. *Lancet*, **13**, 987–996.
  44. Raz, V., Buijze, H., Raz, Y., Verwey, N., Anvar, S.Y., Aartsma-Rus, A. and van der Maarel, S.M. (2014) A novel feed-forward loop between ARIH2 E3-ligase and PABPN1 regulates aging-associated muscle degeneration. *Am. J. Pathol.*, **184**, 1119–1131.
  45. Vickers, T.A., Wyatt, J.R., Burckin, T., Bennett, C.F. and Freier, S.M. (2001) Fully modified 2' MOE oligonucleotides redirect polyadenylation. *Nucleic Acids Res.*, **29**, 1293–1299.
  46. Mariot, V., Roche, S., Hourde, C., Portilho, D., Sacconi, S., Puppo, F., Duguez, S., Rameau, P., Caruso, N., Delezoide, A.L. et al. (2015) Correlation between low FAT1 expression and early affected muscle in facioscapulohumeral muscular dystrophy. *Ann. Neurol.*, **78**, 387–400.
  47. Millevoi, S. and Vagner, S. (2010) Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation. *Nucleic Acids Res.*, **38**, 2757–2774.
  48. Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
  49. Wu, J.C., Gardner, D.P., Ozer, S., Gutell, R.R. and Ren, P. (2009) Correlation of RNA secondary structure statistics with thermodynamic stability and applications to folding. *J. Mol. Biol.*, **391**, 769–783.
  50. Wada, T., Hara, M., Taneda, T., Qingfu, C., Takata, R., Moro, K., Takeda, K., Kishimoto, T. and Handa, H. (2012) Antisense morpholino targeting just upstream from a poly(A) tail junction of maternal mRNA removes the tail and inhibits translation. *Nucleic Acids Res.*, **40**, e173.
  51. Eberle, A.B., Stalder, L., Mathys, H., Orozco, R.Z. and Muhlemann, O. (2008) Posttranscriptional gene regulation by spatial rearrangement of the 3' untranslated region. *PLoS Biol.*, **6**, e92.
  52. Jenal, M., Elkon, R., Loayza-Puch, F., van Haften, G., Kuhn, U., Menzies, F.M., Oude Vrielink, J.A., Bos, A.J., Drost, J., Rooijers, K. et al. (2012) The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. *Cell*, **149**, 538–553.
  53. Uitte de Willige, S., Rietveld, I.M., De Visser, M.C., Vos, H.L. and Bertina, R.M. (2007) Polymorphism 10034C>T is located in a region regulating polyadenylation of FGG transcripts and influences the fibrinogen gamma'/gammaA mRNA ratio. *J. Thromb. Haemost.*, **5**, 1243–1249.
  54. Wiestner, A., Tehrani, M., Chiorazzi, M., Wright, G., Gibellini, F., Nakayama, K., Liu, H., Rosenwald, A., Muller-Hermelink, H. K., Ott, G. et al. (2007) Point mutations and genomic deletions in CCND1 create stable truncated cyclin D1 mRNAs that are associated with increased proliferation rate and shorter survival. *Blood*, **109**, 4599–4606.
  55. Popplewell, L.J., Graham, I.R., Malerba, A. and Dickson, G. (2011) Bioinformatic and functional optimization of antisense phosphorodiamidate morpholino oligomers (PMOs) for therapeutic modulation of RNA splicing in muscle. *Methods Mol. Biol.*, **709**, 153–178.
  56. Krom, Y.D., Dumonceaux, J., Mamchaoui, K., den Hamer, B., Mariot, V., Negroni, E., Geng, L.N., Martin, N., Tawil, R., Tapscott, S.J. et al. (2012) Generation of isogenic D4Z4 contracted and non-contracted immortal muscle cell clones from a mosaic patient: a cellular model for FSHD. *Am. J. Pathol.*, **181**, 1387–1401.
  57. Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L. et al. (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.*, **55**, 611–622.

## Supplementary Materials

### No significant off-targets are predicted

To evaluate potential off-target effects, PMOs targeting the *DUX4* transcript were submitted to short-input-sequence BLAST searches against coding and non-coding human genome nucleotide collection databases (nr/nr and nr/nt: see Table Sx). According to BLAST a very strong match is considered to have E-value of  $<1e-04$ , whereas a poor match has  $E >1e-03$ . For each PMO, the two BLAST returns showing highest alignment scores (lowest E-values) were selected for further analysis. With the exception of CS1 which contains specifically selected mismatches, all designed PMOs returned a very strong match to the desired *DUX4* transcript (E-value  $\leq 2e-05$ ), whereas all of the off-target gene candidates show high E-values  $>0.5$ . Predicted off-target sequence matches that showed homology on their negative or forward strand, corresponding to RNA transcript, were selected for calculation of total free energy of binding using the *RNAup Server* (Table Sx; Supplementary figure S1). The only potential off-target interactions on the RNA transcript (ie negative strand) was *GYS2* for PMO-PAS and *PTPRT* for PMO-CS1, but overall free energies of binding were low (-16.62 and -6.60 respectively), and with major mismatches (8/15 and 12/29, respectively). These data suggest a low probability of interference of any of the designed PMOs targeting *DUX4* with off target RNAs.

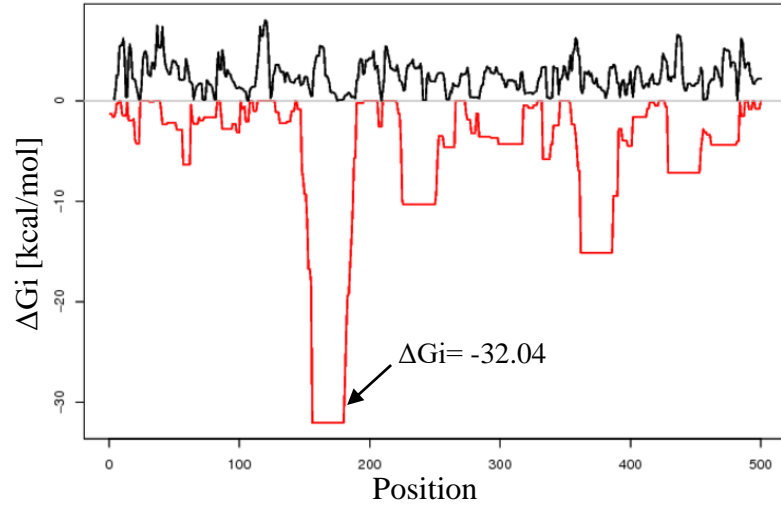
PMO name	Target name	E-value	% identity	No. of base overlap	Predicted homology with (+/-) strand	Coding/non-coding	Overall free binding energy ( $\Delta G_i$ )
PAS	DUX4	4e-05	100	25/25	-	Non-coding	-32.04
	IL1RAPL2	0.58	72	18/25	+	N.A.	N.A.
	GYS2	2.3	68	17/25	-	Non-coding	-16.62
CS1	C1ORF101	0.91	72	21/29	+	N.A.	N.A.
	DUX4	3.6	59	17/29	-	Non-coding	-21.82
	PTPRT	3.6	59	17/29	-	Non-coding	-6.60
CS2	DUX4	2e-05	96	27/28	-	Non-coding	-38.02
	OTOF	0.82	75	21/28	+	N.A.	N.A.
	POR	0.82	64	18/28	+	N.A.	N.A.
CS3	DUX4	2e-05	97	29/30	-	Non-coding	-37.93
	OTOF	0.99	70	21/30	+	N.A.	N.A.
	POR	0.99	60	18/30	+	N.A.	N.A.
DSE	DUX4	4e-05	100	25/25	-	Non-coding	-27.41
	TCRA/TC	0.58	72	18/25	+	N.A.	N.A.
	KCNJ6	2.3	68	17/25	+	N.A.	N.A.

**Table 1. Prediction of the anti-DUX4 off target candidates.** Each PMO sequence was run through the BLAST software, optimised for short input sequences and limited to human and nt/nr database. Overall binding energy was calculated only for the predicted targets that show sequence homology with the PMO on their negative strand (-). The units of  $\Delta G_i$  are expressed in kcal/mol. of Abbreviations: PMO, phosphorodiamidate morpholino oligomer; CS, cleavage site; DUX4, double homeobox 4; IL1RAPL2, interleukin 1 receptor accessory protein-like 2; GYS2, glycogen synthase 2; C1ORF101, chromosome 1 open reading frame 101; PTPRT, protein tyrosine phosphatase receptor type T; OTOF, otoferlin; POR, P450 cytochrome oxidoreductase; TCRA/TCRD, T cell receptor alpha delta locus; KCNJ6, potassium channel inwardly rectifying subfamily J member.

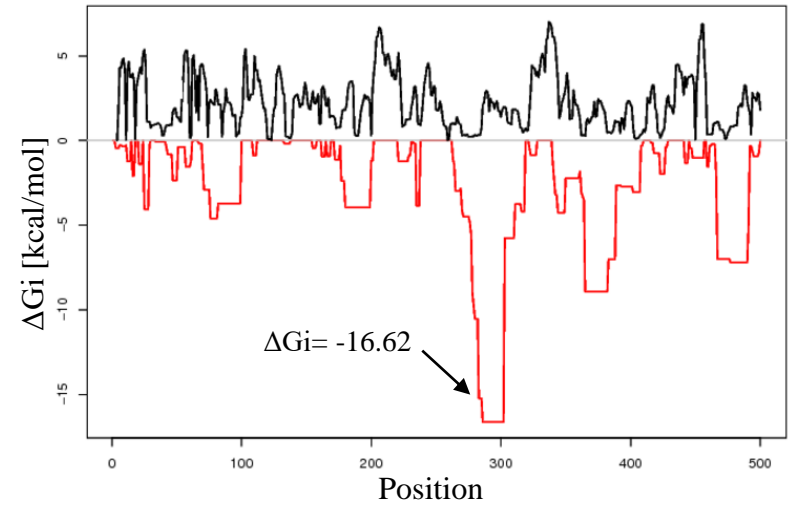


PMO-PAS

*DUX4*

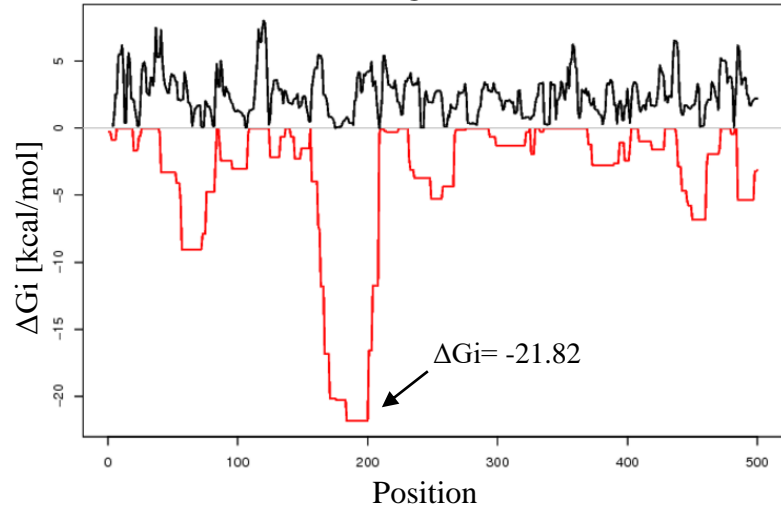


*GYS2*

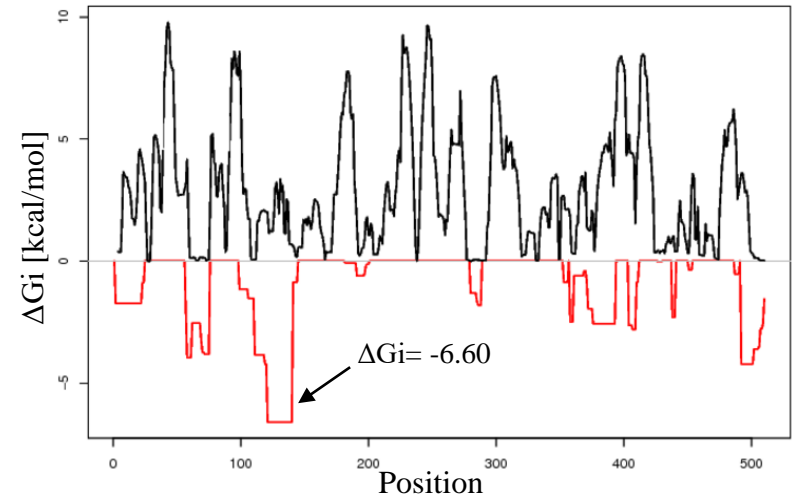


PMO-CS1

*DUX4*

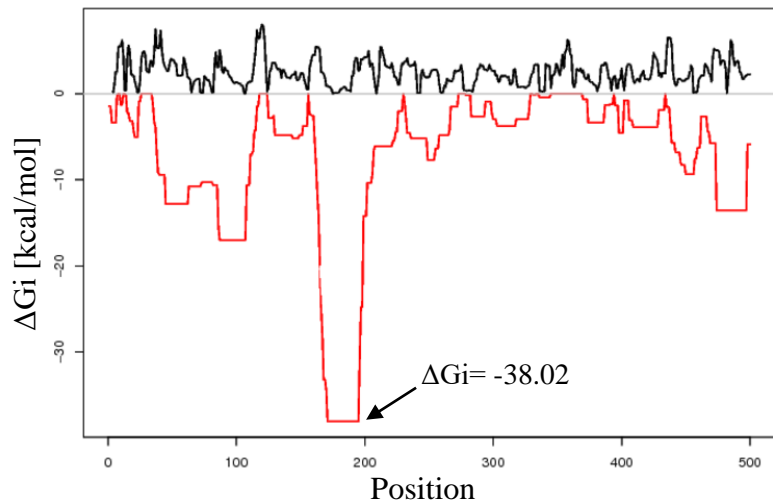


*PTPRT*



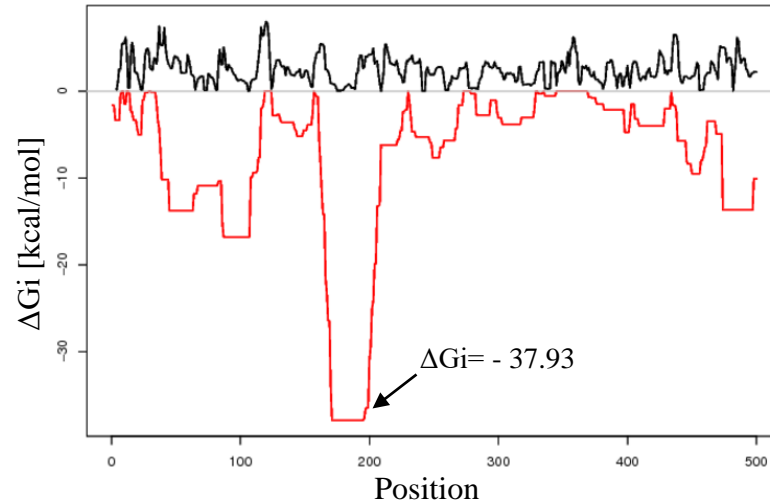
PMO-CS2

*DUX4*



PMO-CS3

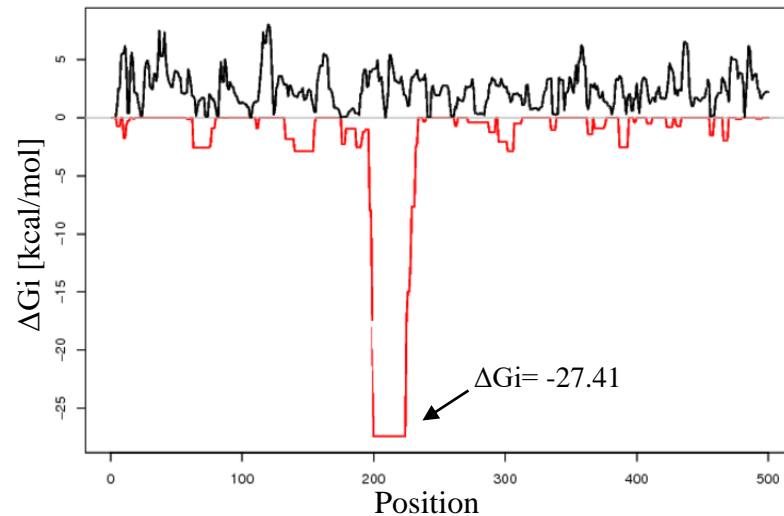
*DUX4*



**Supplementary figure 1. Overall energy of binding of the predicted PMO candidates.** The PMO overall free energy ( $\Delta G_i$ ) of binding was calculated using the RNAup server, only for the candidate genes that show sequence homology on the reverse strand (-)(Table 1). The RNAup derived plot shows the interaction free energy of the PMO along the target sequence (RED)  $\Delta G_i$  and the energy needed to open structures in the longer sequence (BLACK). On the Y-axis nucleotide position of the sequence is indicated.

PMO-DSE

*DUX4*



## **2.2. Etude II : Evaluation *in vivo* de l'efficacité des oligonucléotides antisens ciblant les éléments clés en 3'UTR de *DUX4***

### **2.2.1. Introduction**

Notre étude *in vitro* fournit les preuves de la faisabilité d'une approche thérapeutique pour la FSHD, basée sur l'utilisation d'oligonucléotides antisens ciblant les éléments clés en 3'UTR de l'ARNm *DUX4*. Par ailleurs, ultérieurement et indépendamment, l'efficacité du PMO-PAS a été confirmée par une équipe américaine dans un modèle *in vitro* de cellules primaires FSHD (Chen et al., 2016a).

Au vu de l'ensemble de ces résultats, et toujours dans l'optique du développement d'une stratégie thérapeutique pour la FSHD, l'efficacité des PMOs les plus prometteurs (ici le PMO-PAS, qui, à la différence du PMO-CS3 n'induit pas de redirection du site clivage) ont été testés *in vivo* dans différents modèles de souris FSHD. Les PMOs, tel que nous les avons utilisés dans la première partie de notre étude, présentent *in vivo* une forte clairance et une faible efficacité à franchir la membrane plasmique des cellules. De ce fait, des administrations répétées de fortes doses de l'oligonucléotide antisens sont nécessaires pour espérer obtenir des effets thérapeutiques notables. L'utilisation de vecteurs performants pour un transfert efficace de l'oligonucléotide antisens s'impose donc. La stratégie choisie dans notre étude est une vectorisation de la séquence du PMO-PAS dans un virus adéno-associé (AAV) sous contrôle du promoteur U7. En effet, ce système est, à ce jour, le plus adapté pour une expression efficace et durable de l'oligonucléotide antisens exclusivement dans le noyau.

#### **2.2.1.1. Modifications du *U7 small nuclear ribonucleoprotein* (U7-snRNP) à des fins de thérapie génique**

Les U7-snARN sont impliqués dans la maturation nucléaire en 3' des ARNm des histones (Dominski and Marzluff, 1999, 2007; Schümperli and Pillai, 2004). A la différence des autres ARNm, les ARNm des histones ne sont pas polyadénylés mais uniquement clivés en 3'. Le site de clivage des ARNm des histones est caractérisé par la présence : (i) d'une structure très

conservée en tige boucle (appariement sur 6 pb et boucle de 4 nucléotides) localisée 14 à 50 pb en aval d'un codon stop et (ii) d'une séquence HDE (*Histone Downstream Element*) riche en purines, située à environ 15 nucléotides en aval de la structure en tige boucle (Figure 41). Le clivage endonucléolytique est réalisé par l'endonucléase CPSF73 (*Cleavage and Polyadenylation Specificity Factor 73*). Il a lieu majoritairement après les cinq nucléotides ACCCA en aval de la tige boucle (Dominski and Marzluff, 2007; Dominski et al., 2005; Furger et al., 1998) (Figure 41). Deux facteurs sont requis pour cette spécificité de la réaction de clivage : (i) le U7-snRNP (*U7 small nuclear RiboNucleoProtein*) qui reconnaît la séquence HDE et (ii) les protéines SLBP (*Stem-Loop Binding Protein*) ou HBF (*Hairpin Binding Factor*) qui reconnaissent la structure en tige boucle. Les protéines SLBP, *via* une interaction avec la protéine ZFP100 (*Zinc Finger Protein 100*), stabilisent le duplex formé entre le U7 snRNP et la séquence HDE (Dominski and Marzluff, 2007; Dominski et al., 1999, 2002; Galli et al., 1983; Strub et al., 1984) (Figure 41).

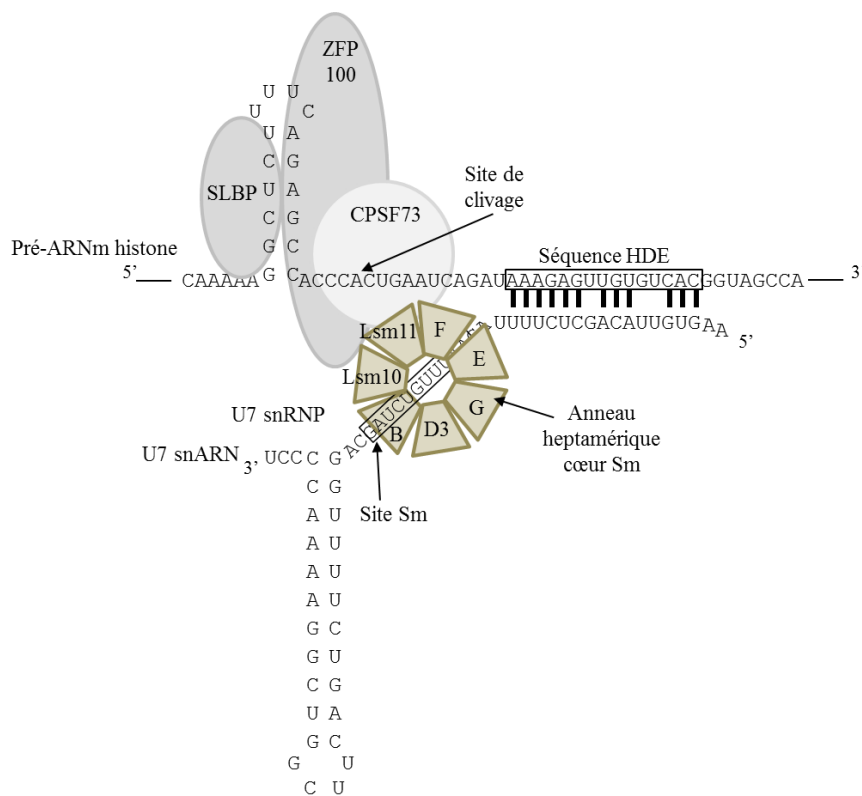


FIGURE 41. Mécanisme de clivage en 3' du pré-ARNm des histones. Le site de clivage du pré-ARNm des histones est localisé entre la tige boucle et la séquence riche en purines HDE (*Histone Downstream Element*). La spécificité du site de clivage est déterminée par une interaction, *via* ZFP100, entre les protéines SLBP et le U7-snRNP. Le U7-snRNP est constitué d'un U7-ARN associé en son site Sm (encadré) à un anneau héptamérique cœur Sm (protéines Sm B, Sm D3, Sm G, Sm E, Sm F, Lsm10 et Lsm11). La région 5' riche en purines du U7-snARN est complémentaire de la séquence HDE et forme un duplex avec le pré-ARNm de l'histone. Le clivage est réalisé par l'endonucléase CPSF73 et a lieu après les cinq nucléotides ACCCA en aval de la tige boucle (adapté de Dominski and Marzluff, 2007).

Après le clivage, la protéine SLBP reste liée à la tige-boucle de l'ARNm mature des histones et stimule son export dans le cytoplasme ainsi que sa traduction. La région en 3' clivée du pré-ARNm des histones est quant à elle dégradée par une exonucléase 5'-3', libérant le U7-snRNP (Gorgonie et al., 2005; Sánchez and Marzluff, 2002; Walther et al., 1998; Williams et al., 1994).

Le U7-snRNP est composé d'un U7-snARN de 57 à 71pb transcrit par la polymérase II (Galli et al., 1983; Strub et al., 1984). Ce U7-snARN comporte une région 5' riche en purines et complémentaire de la séquence HDE, une région 3' repliée en une tige-boucle et une région centrale contenant un site Sm de séquence consensus (AAUUUGUCUAG) (Dominski and Marzluff, 1999, 2007; Kolev and Steitz, 2006). Autour du site Sm, sous contrôle des protéines SMN (*Survival of Motor Neurons*), sept protéines (Sm B, Sm D3, Sm G, Sm E, Sm F, Lsm10 et Lsm11) s'assemblent pour former un anneau heptamérique cœur Sm (Pillai et al., 2001, 2003). Ces structures sont essentielles au fonctionnement du U7-snRNP car elles permettent la reconnaissance de la séquence HDE du pré-ARNm des histones et ainsi la détermination précise du site de clivage (Dominski and Marzluff, 1999, 2007; Schümperli and Pillai, 2004) .

Le U7-snRNP, du fait de sa localisation et de son action exclusivement nucléaire, peut être modifié à des fins de thérapie génique pour cibler des pré-ARNm dans le noyau (Vacek et al., 2003). Pour cela, la séquence Sm du U7-snARN est remplacée par une séquence Sm dite optimisée (Sm OPT *Originally Meaning Optimized*). Il en résulte une accumulation plus importante quantitativement de U7-snRNP dans le noyau. A noter que cette modification engendre une perte de la fonctionnalité du U7-snRNP dans l'action de clivage des pré-ARNm des histones (Grimm et al., 1993). En effet, la séquence Sm OPT est spécifique des U-snARN impliqués dans l'épissage des ARNm. Elle est, de ce fait, associée à un anneau heptamérique différent où les protéines Sm D1/D2 se substituent aux protéines Lsm10/11. Ainsi, le U7-snRNP Sm OPT peut se lier à sa cible ARNm, de la même manière qu'un U7-snRNP Sm, mais l'absence de Lsm10/11 ne permet pas le recrutement de l'endonucléase CPSF73, essentielle au clivage endonucléolytique (Schümperli and Pillai, 2004). Le U7-snARN Sm OPT est ensuite modifié au niveau de sa séquence HDE pour cibler un pré-ARNm spécifique. Par exemple, dans le cadre d'une stratégie thérapeutique pour la DMD, la séquence HDE d'un U7-snARN Sm OPT a été remplacée par la séquence d'un oligonucléotide antisens capable de masquer les signaux accepteurs d'épissage d'un exon à exclure dans l'ARNm de la dystrophine. Cet U7-snRNP Sm OPT modifié n'induit pas de clivage de l'ARNm mais

permet, par saut d'exon, de restaurer le cadre de lecture de la dystrophine et ainsi de produire une protéine tronquée mais fonctionnelle (Brun et al., 2003; De Angelis et al., 2002; Goyenvalle et al., 2004, 2009, 2012).

### **2.2.1.2. Utilisation des virus adéno-associés pour le transfert de l'acide nucléique**

#### **2.2.1.2.1. Généralités sur les virus adéno-associés**

Le gène codant pour un U7-snARN Sm OPT est de petite taille (environ 600pb avec le promoteur). Il peut donc être facilement inséré dans différents types de vecteurs viraux dont les rétrovirus, les lentivirus ou encore les virus adéno-associés (AAV) afin d'améliorer son transfert aux cellules cibles (De Angelis et al., 2002; Goyenvalle et al., 2004; Liu et al., 2004; Vacek et al., 2003). Plusieurs éléments sont à prendre en compte dans le choix du vecteur viral : sa capacité à se lier puis à entrer dans la cellule cible ; sa capacité à transférer efficacement l'acide nucléique qu'il transporte au noyau, associée à une expression du transgène pendant une longue période ; sa toxicité. Les AAV répondent avec succès à ces critères. Notamment, et à la différence des lentivirus, les AAV permettent de transduire efficacement tous les muscles d'un organisme, aussi bien chez de petits que de gros animaux, incluant l'Homme, faisant d'eux des outils très utilisés dans les stratégies thérapeutiques ciblant des cellules musculaires (Wang et al., 2014).

Les AAVs sont de petits virus de la famille des *Parvoviridae*, genre *Dependovirus*, non pathogènes et pour lesquels 12 sérotypes (AAV1 à AAV12) et plus de 100 variants ont été identifiés (Daya and Berns, 2008; Gao et al., 2011). Le matériel génétique de l'AAV, sous forme d'un ADN linéaire simple-brin de 4,7kb, est contenu dans une capsidie icosaédrique non enveloppée d'un diamètre de 26nm (Daya and Berns, 2008; DiMattia et al., 2012; Xie et al., 2002). Le tropisme tissulaire ou cellulaire ainsi que la biodistribution de l'AAV sont déterminés par les protéines de la capsidie et varient d'un sérotype à un autre. Le génome est constitué de deux gènes, *Rep* et *Cap* (Grieger and Samulski, 2012; Sonntag et al., 2010) (Figure 42). Par utilisation de différents sites d'initiation de la transcription et par épissages alternatifs, le gène *Rep* code pour quatre protéines Rep (Rep78, Rep 68, Rep52 et Rep40) impliquées dans le cycle viral tandis que le gène *Cap* code pour trois protéines de la capsidie

(VP1, VP2 et VP3) ainsi que pour une protéine impliquée dans la formation active de la capsid (AAP *Assembly-Activating Protein*) (Grieger and Samulski, 2012; Sonntag et al., 2010; Wang et al., 2014). Ces gènes sont encadrés de deux séquences répétées inversées ou ITR (*Inverted Terminal Repeat*) de 145bp qui servent d'origines de réplication. Le génome de l'AAV, ne codant pas pour les protéines nécessaires à sa réplication, est dépendant pour se répliquer de la machinerie de la cellule hôte et de l'aide d'autres virus comme les adénovirus, le virus « de la vaccine » ou le virus de l'Herpès simplex de type 1 (Buller et al., 1981; Hoggan et al., 1966; Schlehofer et al., 1986). En l'absence de l'aide d'autres virus, l'AAV reste à l'état de latence dans la cellule hôte infectée par intégration de son génome majoritairement dans une région de 4kb sur le chromosome 19q13 (région AAVS1) selon un mécanisme Rep dépendant (Daya and Berns, 2008; Kotin et al., 1990).

L'infection de la cellule hôte par un AAV s'effectue en plusieurs étapes (Daya and Berns, 2008). Tout d'abord, le virus se fixe à des récepteurs et co-récepteurs membranaires (Akache et al., 2006; Kashiwakura et al., 2005; Murlidharan et al., 2014; Qing et al., 1999; Summerford and Samulski, 1998). Une fois fixé à la membrane de la cellule hôte, l'AAV est principalement internalisé selon un mécanisme d'endocytose (Bartlett et al., 2000; Ding et al., 2005; Duan et al., 1999; Nonnenmacher and Weber, 2012). Les événements cellulaires suivant l'entrée de la particule virale ne sont aujourd'hui pas totalement caractérisés. Il semblerait cependant que l'AAV s'échappe de l'endosome par lipolyse de la membrane, grâce à l'activité phospholipase A2 des protéines de capsid VP1 et VP2 (Bartlett et al., 2000; Douar et al., 2001; Girod et al., 2002; Nonnenmacher and Weber, 2012; Stahnke et al., 2011). Une faible fraction des particules atteindra le noyau en se déplaçant le long du cytosquelette (Sanlioglu et al., 2000). Le génome de l'AAV peut être alors libéré à plusieurs endroits : (i) lors d'une décapsidation dans le cytoplasme avant un import nucléaire, (ii) lors d'une décapsidation au moment de l'entrée dans le noyau par les pores nucléaires ou (iii) lors d'une décapsidation dans le noyau après entrée du virion probablement par les pores nucléaires. En présence d'adénovirus, de nombreuses particules d'AAV sont détectées dans le noyau, suggérant plutôt une libération du génome de l'AAV lors d'une décapsidation dans le noyau (Bartlett et al., 2000; Johnson and Samulski, 2009). A cette étape, et dans le cas d'une co-infection avec un virus auxiliaire, l'AAV peut se répliquer, après conversion de son ADN simple-brin en ADN double-brin, et produire de nouvelles particules virales. En l'absence d'une co-infection avec un virus auxiliaire, l'AAV reste à l'état de latence (Daya and Berns, 2008).

### 2.2.1.2.2. Utilisation des virus adéno-associés recombinants comme outils pour le transfert de l'acide nucléique

Pour une application en thérapie génique, des vecteurs AAV dits recombinants (AAVr) ont été générés à partir des sérotypes et variants naturels de l'AAV (Hermonat and Muzyczka, 1984; Muzyczka, 1992; Tratschin et al., 1984; Yang et al., 2011). Seules les séquences virales ITR ont été conservées dans le génome de ces AAVr (Figure 42). Les gènes *Rep* et *Cap* sont remplacés par la cassette d'expression d'intérêt (promoteur, gène, séquence poly(A)) du ou des transgènes d'une taille maximale de 4,5kb (Figure 42).

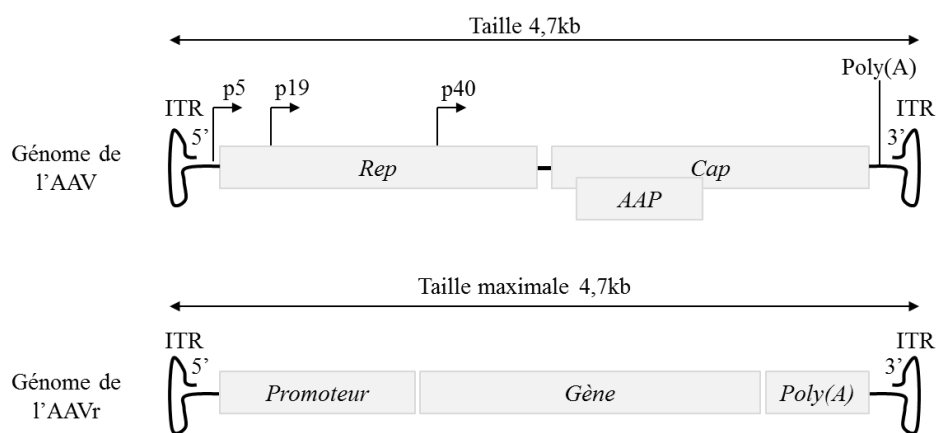


FIGURE 42. Organisation génomique de l'AAV et de l'AAVr. Le matériel génétique de l'AAV, sous forme d'un ADN linéaire simple-brin de 4,7kb, est constitué de deux gènes : *Rep* et *Cap* codant respectivement pour des protéines impliquées dans le cycle viral et des protéines de la capsid (par utilisation de plusieurs sites d'initiation de la transcription p5, p19 ou p40 et épissages alternatifs). Dans le génome de l'AAVr, les gènes *Rep* et *Cap* sont remplacés par une cassette d'expression d'intérêt (promoteur, gène, poly(A)). Les séquences virales ITR sont conservées (adapté de Zacchigna et al., 2014).

La production de ces vecteurs AAVr nécessite la présence dans les mêmes cellules des protéines Rep, Cap ainsi que des facteurs fournis par le virus auxiliaire. Généralement, la production des particules virales s'effectue dans des cellules embryonnaires humaines de rein 293 (HEK293) par co-transfection de trois plasmides : un plasmide contenant le génome AAVr, un plasmide contenant les gènes *Rep* et *Cap* sans les séquences virales ITR et un plasmide contenant les facteurs auxiliaires (Ayuso et al., 2010). Le génome de l'AAVr étant dépourvu de gènes viraux, il s'intègre très rarement dans le génome de la cellule hôte et persiste *in vivo*, pendant plusieurs années, sous la forme d'épisomes (Afione et al., 1996; Niemeyer et al., 2009; Nowrouzi et al., 2012; Rivera et al., 2005; Schnepf et al., 2003, 2005). Les étapes limitantes dans l'expression du transgène sont principalement le trafic intracellulaire (peu de particules virales atteignent le noyau) et l'efficacité de conversion de



l'ADN simple-brin en double-brin (Ferrari et al., 1996; Hauck et al., 2004). De nouvelles technologies ou concepts tendent à s'affranchir de ces contraintes.

En comparaison à d'autres vecteurs viraux, les AAVr présentent de nombreux avantages pour le transfert des acides nucléiques *in vivo* dans des tissus post-mitotiques. Une seule injection locale ou systémique des vecteurs AAVr permet une expression forte du transgène et ce pendant parfois plusieurs années. De plus, les AAVr, en fonction du sérotype et de la voie d'administration, sont peu immunogènes comparés à la majorité des vecteurs viraux (Brockstedt et al., 1999; Mátrai et al., 2010; Mingozzi and High, 2011; Wang et al., 2007; Zaiss et al., 2002). En effet, l'un des problèmes de l'utilisation des vecteurs viraux est l'induction de réponses immunitaires contre les protéines de la capsid. Généralement, une forte inflammation est observée au site d'injection. Au contraire, l'administration d'AAVr ne cause aucune inflammation significative. Les réponses immunitaires induites sont uniquement des réponses adaptives. Plus précisément, la majorité des individus sont exposés, au cours de leur vie, à des AAV naturels. De plus, chez l'Homme, l'infection par un AAV est aidée par un adénovirus. Il en résulte une importante réponse inflammatoire et une production de cellules T-CD8+ mémoires spécifiques des protéines de la capsid (Gao et al., 2009; Mays et al., 2009; Mingozzi et al., 2007). Par ailleurs, des anticorps neutralisant les virus (NAB *Neutralizing AntiBodies*) sont aussi produits (Chirmule et al., 1999; Mingozzi and High, 2013). Ainsi, l'injection d'AAVr peut entraîner une réactivation de ces cellules T-CD8+ et une production de NAB si le sérotype de l'AAVr est le même que celui des AAV auxquels l'individu a été confrontés. Dans ce cas, l'efficacité de transduction de certains types cellulaires par l'AAVr est réduite. De plus, les cellules T-CD8+ peuvent induire la mort des cellules transduites, supprimant ainsi les effets thérapeutiques positifs de l'acide nucléique apporté par l'AAVr.

Il est possible d'anticiper les effets néfastes d'une réponse immunitaire par une administration transitoire d'un traitement immunosuppresseur (Halbert et al., 1998; Manning et al., 1998; Nathwani et al., 2011). Un autre avantage des AAVr est que la grande diversité des sérotypes permet de transduire efficacement et de cibler préférentiellement de nombreux tissus, dont le muscle strié squelettique ou cardiaque, le foie, la rétine et même le système nerveux central (Bish et al., 2011; Bockstael et al., 2011; Qiao et al., 2011; Sands, 2011; Stieger et al., 2011; Wu et al., 2006; Zaiss et al., 2002). Pour une transduction des cellules du muscle strié

squelettique, les AAVr de sérotypes 1/6/7/8/9 sont les plus adaptés (Gregorevic et al., 2004; Wang et al., 2005; Wu et al., 2006; Zincarelli et al., 2008).

En résumé, les vecteurs AAVr sont donc à ce jour les vecteurs viraux les plus efficaces pour le transfert d'acides nucléiques *in vivo*, et présentent le moins de risques pour l'individu. Ils sont actuellement utilisés dans de nombreux essais cliniques, dont des essais pour la DMD ou la dystrophie musculaire des ceintures.

### 2.2.1.3. Hypothèse de travail

Pour tester l'efficacité du PMO-PAS *in vivo*, nous avons donc décidé d'utiliser au cours de cette étude une construction U7-SmOPT contenant la séquence du PMO-PAS (Figure 43). La séquence du U7, modifié uniquement dans sa région 5', a été introduite dans des vecteurs AAV-8 recombinants pour un transfert efficace *in vivo* du transgène dans les cellules du muscle strié squelettique chez deux modèles de souris FSHD. Notre hypothèse est que les fibres musculaires transduites expriment suffisamment d'ARN U7-PAS fonctionnels permettant de masquer le signal de polyadénylation sur le pré-ARNm *DUX4*. Ceci empêcherait l'ajout d'une queue poly(A) et induirait *de facto* la dégradation de l'ARNm *DUX4*.

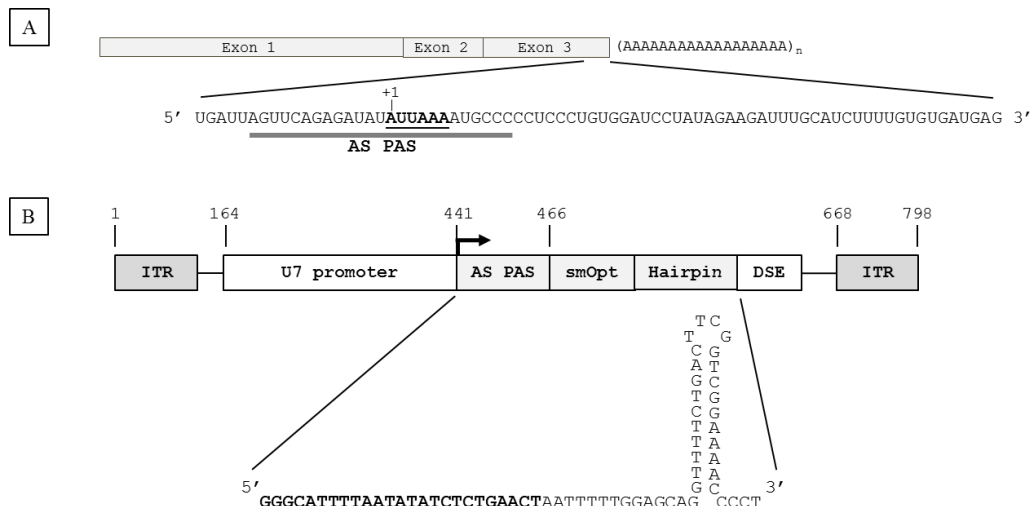


FIGURE 43. Séquence reconnue par le PMO-PAS et structure du vecteur AAV-U7-PAS. A) Séquence reconnue par le PMO-PAS (trait noir AS-PAS) sur l'ARNm *DUX4*. La position +1 détermine la première base du signal de polyadénylation localisé dans l'exon 3. Le signal de polyadénylation est en gras et surligné. B) Structure du vecteur AAV-U7-PAS. Le promoteur U7 est directement suivi de la séquence de l'oligonucléotide antisens PMO-PAS (boîte AS-PAS et en gras), de la séquence Sm optimisée (boîte Sm Opt et surlignée), de la région en tige-boucle (boîte *Hairpin*) et d'une région DSE (*Downstream sequence element*). Cette cassette d'expression, qui inclut le promoteur U7 (boîte *U7 promoter*), est encadrée par des répétitions terminales inversées (ITR).

## 2.2.2. Matériels et méthodes

### 2.2.2.1. Constructions des vecteurs AAV-U7-PAS et pCINéoDUX4-ΔSV40

La région SD23/BP22 ((BP22: 5' AAATAGAAGTTCATTTACTACTAAC 3' and SD23: 5' GGCCAAACCTCGGCTTACCT 3') du vecteur pU7-DTex23 est remplacée par la séquence du PMO-PAS (Goyenville et al., 2004; Marsollier et al., 2016). Pour cela, deux fragments PCR ont été générés indépendamment par amplification du vecteur pU7-DTex23 avec la GoTaq® G2 Hot Start Green Master Mix 2X (Promega), selon les instructions du fabricant, et à l'aide des couples d'amorces PCR1 et PCR2 suivants : PCR 1 - (U7\_PCR1\_fw: 5' CTACGTAGCCATGCTCTAGATAAC 3' et U7\_PAS\_PCR1\_rev: 5' AGAGATATATATAAAATGCCCTTGCGGAAG 3') et PCR 2 - (U7\_PAS\_PCR2\_fw: 5' GGCATTTTAATATATCTCTGAACTAATTTTTGG 3' et U7\_PCR2\_rev: 5' GCTCTAGCCACATACGCGTTTC 3'). Les conditions de PCR sont : 2 min à 95°C, puis 40 cycles à 95°C pendant 30s, 55°C pendant 30s, 72°C pendant 30s, et une élongation finale à 72°C pendant 5min. Les produits PCR sont purifiés avec le kit Wizard® SV Gel and PCR Clean-up Kit (Promega) selon les instructions du fabricant. Le fragment U7-PAS, contenant la séquence du PMO-PAS sous promoteur U7 en substitution de la séquence SD23/BP22, est généré par une PCR sur les deux produits obtenus précédemment, en utilisant la GoTaq® G2 Hot Start Green Master Mix 2X (Promega), selon les instructions du fabricant, et à l'aide des amorces suivantes : U7\_PCR1\_fw: 5' CTACGTAGCCATGCTCTAGATAAC 3' et U7\_PCR2\_rev: 5' GCTCTAGCCACATACGCGTTTC 3'. Les conditions de PCR sont : 2 min à 95°C, puis 40 cycles à 95°C pendant 30s, 55°C pendant 30s, 72°C pendant 30s, et une élongation finale à 72°C pendant 5min. Le fragment U7-PAS produit est introduit dans le vecteur pU7-DTex23, en place et lieu du U7-SD23/BP22, par utilisation des sites de restriction *XbaI* et *MluI*.

Le plasmide pCINéo-DUX4 (déposé par Alexandra Belayew) contient une répétition D4Z4 immédiatement suivie de la région pLAM (avec le signal de polyadénylation canonique ATTTAAA) et d'une séquence de polyadénylation du virus simien 40 (SV40-PolyA). Le plasmide pCINéo-DUX4-ΔSV40 est généré à partir du plasmide pCINéo-DUX4 par élimination du fragment SV40-PolyA à l'aide des enzymes de restriction *HpaI* et *EcoRV*.

### 2.2.2.2. Production des AAV

Les vecteurs AAV (AAV8-U7-PAS et AAV1-Ctrl) sont produits dans des cellules embryonnaires humaines de rein 293 (HEK293) par triple transfection au PEI du plasmide pXX6 codant une séquence adénovirale essentielle à la production de l'AAV, du plasmide pRepCap codant pour les protéines du cycle viral et de la capside (capside 8 pour le vecteur AAV8-U7-PAS et capside 1 pour le vecteur AAV1-Ctrl) et le génome viral portant la séquence U7-PAS. Les virus sont ensuite purifiés par gradient iodixanol et concentrés sur des colonnes Amicon Ultra-15 (Merk Millipore, Molsheim, France).

### 2.2.2.3. Procédures animales

Toutes les procédures animales ont été conduites sous le numéro d'agrément APAFIS#5555-2016062713055186 v3 en accord avec la réglementation appliquée en France et en Europe.

Deux modèles murins ont été utilisés pour ces expériences (Figure 44) :

1. Des souris femelles de six semaines C57BL/6 ont été anesthésiées par injection intra-péritonéale d'un mélange en PBS de kétamine (100mg/mL) et de xylazin (10mg/mL). Une préparation de 40µL de vecteur AAVr (AAV-U7-PAS à un titre de  $5,53 \cdot 10^{12}$  vg/mL ou AAV-Ctrl à un titre de  $2,9 \cdot 10^{12}$  vg/mL) est tout d'abord injectée dans le muscle du *tibialis anterior* (TA) de ces souris. Deux semaines plus tard, 40µL de plasmide pCINéo-DUX4 ou pCINéo-DUX4  $\Delta$ SV40 (2µg) dilué dans une solution à 50:50 de PBS 1X et d'eau sans RNase, sont injectés dans le TA des souris C57/BL6. Immédiatement après l'injection, le plasmide est électroporé dans les fibres musculaires avec les paramètres suivants : MODE LV, Voltage 200V/cm (90V), Pulse length 20ms, Pulses 08, Interval 500ms, Polarity Unipolar (électrodes : Tweezertrodes 7mm, BTX Harvard Apparatus, générateur : ECM 830 Square Wave Electroporator, BTX Harvard Apparatus). Les souris C57BL/6 sont euthanasiées quatre jours après l'électroporation (Figure 44).

2. Des souris mâles de neufs semaines *ACTA1-MCM,FLexD/+* (souris FLexDUX4 croisées avec des souris ACTA1-cre/Esr1) sont injectées avec 50µL de préparation AAV-U7-PAS (à un titre de  $2,16 \cdot 10^{12}$ vg/mL) dans le TA gauche et 50µL de préparation AAV-Ctrl (à un titre de  $2,9 \cdot 10^{12}$ vg/mL) dans le TA droite sous anesthésie isoflurane. Deux semaines plus tard, de

faibles doses de tamoxifène (5mg/kg) sont injectées en intra-péritonéal aux souris pour induire l'expression de *DUX4*. Les souris *ACTA1-MCM,FLexD/+* sont sacrifiées deux semaines après l'induction de *DUX4* (par administration de tamoxifène) (Figure 44). Ces expériences ont été réalisées aux USA, par l'équipe du Dr Peter Jones, propriétaire de ces souris.

Dans les deux cas, les muscles sont récupérés. La partie supérieure du TA, utilisée en histologie, est montée sur bouchon de liège entourée d'O.C.T. (*Optimum Cutting Temperature*) puis plongée dans de l'isopentane refroidi à l'azote liquide. Les TA montés sont conservés à  $-80^{\circ}\text{C}$ . La partie inférieure du TA est également conservée à  $-80^{\circ}\text{C}$  pour une extraction ARN et ADN.

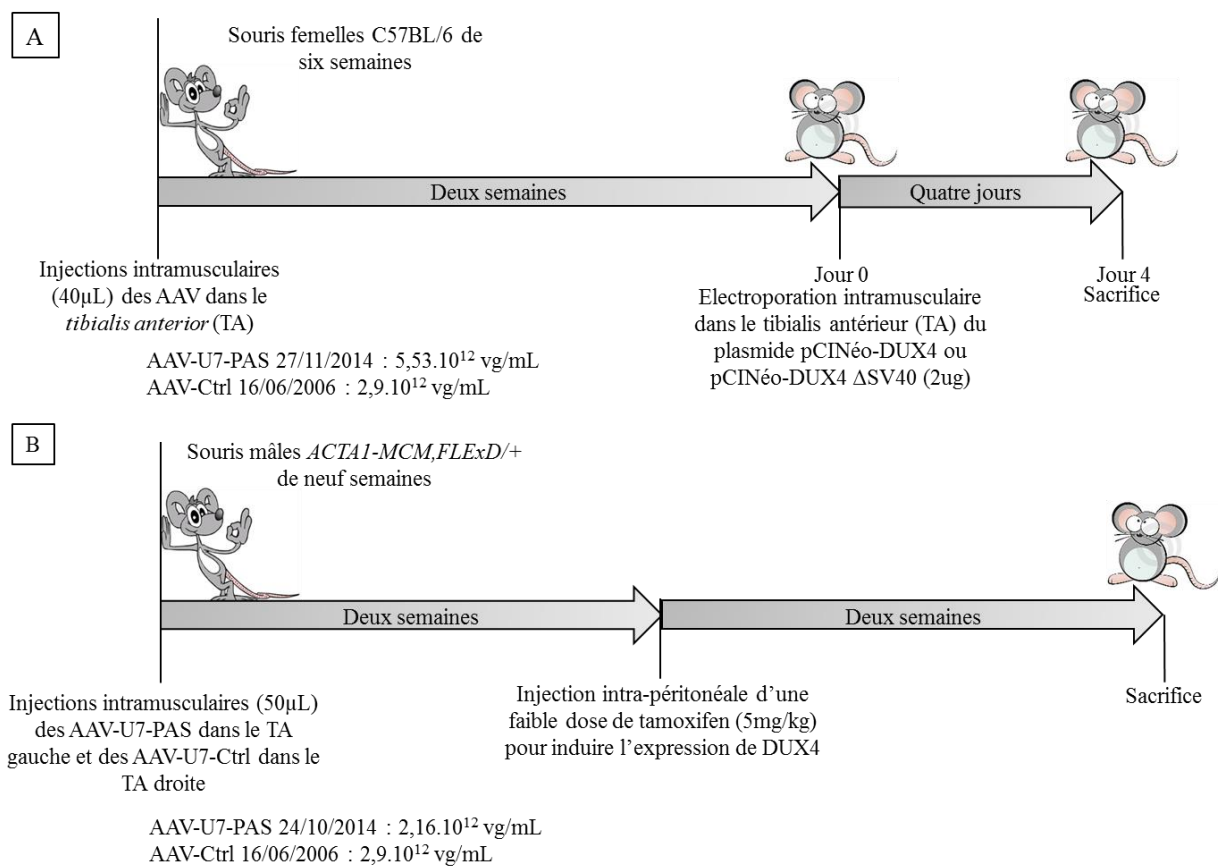


FIGURE 44. Schéma du protocole expérimental *in vivo*. A) Protocole appliqué aux souris femelles C57BL/6 de six semaines. B) Protocole appliqué aux souris mâles *ACTA1-MCM,FLexD/+* de neuf semaines.

#### 2.2.2.4. Extraction ARN et ADN, transcription reverse

L'extraction d'ADN et d'ARN des échantillons de TA est réalisée en utilisant du Trizol selon les instructions du fabricant (Life Technologies, Saint Aubin, France). La concentration en

ADN ou ARN est déterminée avec le nanodrop ND-1000 spectrophotometer (Thermo Scientific, Wilmington, USA). La transcription inverse (RT *Reverse transcription*) est réalisée sur 1µg d'ARN total traité par DNase Rnase-free (RQ1 RNase-free DNase Product, Promega, France) dans un volume final de 10,5µL avec le kit Roche Transcriptor First Strand cDNA Synthesis Kit (Roche, Meylan, France) selon les instructions du fabricant.

### **2.2.2.5. PCR et PCR quantitative**

La PCR et la PCR quantitative (PCRq) sont conçues en accord avec les critères MIQE (*Quantitative Real-Time PCR Experiments*) (Bustin et al., 2009). Les amorces utilisées pour la PCR ou la PCRq sont reportées dans le Tableau 6. Les PCR (*DUX4*, *mGapdh*) sont réalisées sur 1µL de produit de RT avec la GoTaq® G2 Hot Start Green Master Mix 2X (Promega) selon les instructions du fabricant. Le programme PCR est le suivant : 94°C pendant 2min, puis 38 cycles à 94°C pendant 10s, 60°C pendant 10s, 72°C pendant 10s, et une élongation finale de 5min à 72°C. Les PCRq sont réalisées avec le Lighcycler LC480 (Roche) dans un volume final de 9µL avec 4µL de produit de RT, 0,18µL pour chaque amorce sens et antisens à 20µM et 4,5µL de SYBR® Green Mastermix 2X (Roche, Meylan, France). Les conditions de PCRq sont : 94°C pendant 5 min puis 50 cycles à 94°C pendant 10s, 60°C pendant 7s et 72°C pendant 7s. Tous les produits d'amplification ont été validés par séquençage. Le gène *mGapdh* était le gène le plus approprié pour la normalisation : un test multiparamétrique d'analyse de variance (MANOVA) et un test post-hoc Newman-Keuls n'a pas révélé de corrélation entre le niveau d'expression de *mGapdh* et les injections.

### **2.2.2.6. Détection du U7-snARN**

Pour la détection du U7 naturel (*wild type*) et du U7-PAS dans les muscles de souris, une RT a été réalisée sur 1µg d'ARN total traité par DNase Rnase-free (RQ1 RNase-free DNase Product, Promega, France) en présence d'hexamères aléatoires à 100µM, avec le kit Roche Transcriptor First Strand cDNA Synthesis Kit (Roche, Meylan, France), dans un volume final de 10,5µL et selon les instructions du fabricant. L'ADNc U7 est amplifié à partir de 1µL de produit de RT en utilisant la GoTaq® G2 Hot Start Green Master Mix 2X (Promega) selon les instructions du fabricant. Les amorces utilisées sont les amorces sens suivantes (5' AAGTGTTACAGCTCTTTTAG 3' pour le U7 *wild type* et 5'

AAGGGCATTTTAATATATCTCTGAAC 3' pour le U7-PAS) et la même amorce antisens localisée dans la tige boucle du U7 snARN (5' AGGGGTTTTCCGACCGAAG 3') (Tableau 6). Le programme PCR est le suivant : 95°C pendant 2min, puis 40 cycles à 95°C pendant 20s, 54°C pendant 20s, 72°C pendant 20s, et une élongation finale à 72°C pendant 5min. Les produits PCR sont analysés sur un gel d'agarose à 2,5%.

### **2.2.2.7. 3'RACE**

Pour la 3'RACE (*3' Rapid amplification of cDNA Ends*), la RT est réalisée avec l'amorce adaptatrice oligo dT :

(GCTGTCAACGATACGCTACGTAACGGCATGACAGTGTTTTTTTTTTTTTTTTTTTTTTT TTTT) (Tableau 6) sur 1µg d'ARN total dans un volume final de 10µL avec le kit Roche Transcriptor First Strand cDNA Synthesis Kit (Roche, Meylan, France) selon les instructions du fabricant. Une première PCR est réalisée dans un volume final de 15µL avec 1µL de produit de RT, 1µL d'amorce sens (20µM), 3µL d'amorce antisens (20µM), 7,5µL de GoTaq® G2 Hot Start Green Master Mix 2X (Promega) (Tableau 6). Le programme PCR appliqué est le suivant : 95°C pendant 2min, puis 5 cycles à 95°C pendant 20s, 72°C pendant 1min20s, puis 5 cycles à 95°C pendant 20s, 70°C pendant 20s, 72°C pendant 1min et enfin 25 cycles à 95°C pendant 20s, 68°C pendant 20s, 72°C pendant 1min avant une élongation finale à 72°C pendant 5min. La PCR nichée est réalisée dans un volume final de 15µL avec 1µL du produit de la première PCR, 1µL de chaque amorce sens et antisens à 20µM et 7,5µL de GoTaq® G2 Hot Start Green Master Mix 2X (Promega). Les conditions PCR sont les suivantes : 95°C pendant 2min, puis 40 cycles à 95°C pendant 20s, 60°C pendant 20s, 72°C pendant 1min, avant une élongation finale à 72°C pendant 5min. Les produits de la PCR nichée sont analysés sur gel d'agarose à 1% et purifiés avec le kit Wizard® SV Gel and PCR Clean-up Kit (Promega) selon les instructions du fabricant. Les fragments PCR sont clonés en vecteur TOPO TA avec le kit TOPO TA Clonong (Lifetechnologies) selon les instructions du fabricant. Les sites de clivage sont déterminés par séquençage d'au moins 10 unités formant des colonies (UFC).

TABLEAU 6. Amorces utilisées au cours de l'étude. fw : amorce sens ; rev : amorce antisens.

Gènes ciblés	Amorces	Numéro d'entrée	Séquence	Localisation	Taille (bp)
<i>U7 (clonage)</i>	U7 PCR1_fw		CTACGTAGCCATGCTCTAGATAAC	N.A.	298bp
	U7 PAS PCR1_rev		AGAGATAATATAAAAATGCCCTTGCGGAAG	N.A.	
	U7 PAS PCR2_fw		GGCAATTTAATAATATCTCTGAACATAATTTTGG	N.A.	181bp
	U7 PAS PCR2_rev		GCTCTAGCCACATAACGGCTTTC	N.A.	
Oligonucleotides PCR	U7 cribre PAS_rev		AGTTCAGAGATATATAAAATGCC	N.A.	
	U7_WT_detection_fw		AAGTGTACAGCTCTTTTAG	N.A.	60bp
	U7_PAS_detection_fw		AAGGGCAATTTAATAATATCTCTGAAC	N.A.	69bp
<i>U7 (détection)</i>	U7_detection_rev		AGGGGTTTTCCGACCGAAG	N.A.	
	DUX4-all_fw	ENSG00000260596	CCCAGGTACACGAGAGACC	Exon2-Exon3	164bp
DUX4-all_rev		TCCAGGAGATGTAACCTCTAATCCA	Exon3		
<i>mGapdh</i>	mGapdh_fw	NM_001289726.1	CACCCACCCAGCAAGGA	Exon7	98bp
	mGapdh_rev		ATGGGGTCTGGGATGGAAA	Exon7	
<i>mTm7sf4</i>	mTm7sf4_fw	NM_029422.4	TATCGGCTCATCTCCTCCAT	Exon1	98bp
	mTm7sf4_rev		ACTCCTTGGGTTCCCTTGCCT	Exon2	
<i>mWfdc3</i>	mWfdc3_fw	NM_027961.1	GGTAGCTGCAGGAGAGACAG	Exon2	94bp
	mWfdc3_rev		CTGGGACAGGATTCGTCTC	Exon3	
<i>DUX4</i>	DUX4_1797_fw	ENSG00000260596	ATGGCCCTCCCGACACCCCT	Exon1	138bp
	DUX4_1906_rev		GATGCCCGGTACGGGTTCCGCTCAAAGC	Exon1	
<i>Neomycine</i>	Neo_Kan_fw	NM_	TGCCTGCTTGCCGAATATCA	ORF	114bp
	Neo_Kan_rev		ATATCAGGGTAGCCAACGC	ORF	
<i>mTitine</i>	mTitine_fw	NM_011652.3	ACCCCTGACICTGTGGAAAGT		133bp
	mTitine_rev		AAATGGTACAAGCCTCCCCG		
Oligonucleotides 3RACE-PCR	Oligo_dT_Reverse		GCTGTCAACGATACGCTACGTAACGGCATGACA	N.A.	N.A.
	3RACE_Primer_rev		GTGTTTTTTTTTTTTTTTTTTTTTTT	N.A.	N.A.
	3RACE_Nested_rev		GCTGTCAACGATACGCTACGTAACG	N.A.	N.A.
	3RACE_Exon1_fw		CGTACGTAAACGGCATGACAGTGT	N.A.	N.A.
<i>DUX4</i>	3RACE_Exon1_fw	ENSG00000260596	CCTCTCCTAGAAACGGAGGCCCG	Exon1	N.A.
	3RACE_Exon1_Nested_fw		CTCGCTGGAAGCACCCCTCAGC	Exon1	N.A.
	3RACE_Exon3_fw	ENSG00000260596	CGCACCCCGGCTGACGTGCAAG	Exon3	N.A.
	3RACE_Exon3_Nested_fw		CGCTGGCCTCTCTGTGCCCTTG	Exon3	N.A.
<i>pCiNeo DUX4</i>	3RACE_pCiNeo_fw		CAGGGATAAGTCAATAAAGCCTCCTG	N.A.	N.A.

N.A. : Non applicable



## 2.2.3. Résultats préliminaires

### 2.2.3.1. Détermination de l'expression de l'ARNm *DUX4* et de ses gènes cibles *in vivo* dans les muscles de souris pCINéo-*DUX4* traités avec le vecteur AAV-U7-PAS

L'efficacité *in vivo* du U7-PAS snARN a d'abord été testée dans des *tibialis anterior* (TA) de souris C57BL/6 exprimant *DUX4*, suite à une électroporation intramusculaire du plasmide pCINéo-*DUX4*. Le plasmide pCINéo-*DUX4* contient une répétition D4Z4 suivie de la région pLAM. La présence de cette région pLAM, qui contient les éléments clés impliqués dans la maturation de l'ARNm *DUX4*, est essentielle à notre étude.

Les vecteurs AAV-U7-PAS ou AAV-Ctrl sont injectés dans le TA des souris femelles C57BL/6 âgées de 6 semaines (Figure 44). Deux semaines après la transduction des muscles avec les vecteurs AAV-U7-PAS ou AAV-Ctrl, l'électroporation intramusculaire du plasmide pCINéo-*DUX4* est effectuée (Figure 44). Quatre jours après l'électroporation, les souris sont sacrifiées et les TA prélevés (Figure 44). Les ARNs totaux sont extraits des muscles traités et non traités. Le niveau d'expression de *DUX4* et de deux de ces gènes cibles identifiés chez la souris, *mTm7sf4* et *mWfdc3*, sont évalués par PCRq à partir de ces extraits ARNs (Ferreboeuf et al., 2014b). Puisque l'efficacité d'électroporation varie d'une souris à une autre, le niveau d'expression de *DUX4*, *mTm7sf4* et *mWfdc3* est rapporté au nombre de copies de plasmide pCINéo-*DUX4* par génome diploïde de la cellule hôte. Pour cela, la quantité relative du nombre de copies du gène *Néomycine*, présent exclusivement sur le plasmide pCINéo-*DUX4*, et la quantité relative du nombre de copies du gène *mTitine*, présent exclusivement en deux copies dans le génome murin, sont déterminés par PCRq à partir des ADNs extraits des TA.

Les ARNm *DUX4*, *mTm7sf4* et *mWfdc3* sont détectés dans les TA des souris pCINéo-*DUX4* traités avec l'AAV-Ctrl (Figure 45). La quantité relative d'ARNm *DUX4* est de 1,75 +/- 1,64 dans les TA traités AAV-Ctrl et de 1,68 +/- 0,77 dans les TA traités AAV-U7-PAS (Figure 45). La quantité relative d'ARNm *mTm7sf4* est de 0,88 +/- 0,63 dans les TA traités AAV-Ctrl et de 0,89 +/- 0,38 dans les TA traités AAV-U7-PAS (Figure 45). La quantité relative d'ARNm *mWfdc3* est de 0,68 +/- 0,46 dans les TA traités AAV-Ctrl et de 0,64 +/- 0,18 dans les TA traités AAV-U7-PAS (Figure 45). Aucune modification significative dans l'expression

de *DUX4* ou de ses gènes cibles n'est donc observée dans les TA traités avec le vecteur AAV-U7-PAS comparés aux TA traités avec le vecteur AAV-Ctrl (p-value des t-test pour des échantillons appariés = 0,89 ; 0,98 ; 0,81 pour *DUX4*, *mTm7sf4* et *mWfdc3* respectivement) (Figure 45). Ce résultat n'est pas corrélé à une absence d'expression dans les muscles du U7-PAS snARN. En effet, en comparaison au U7 snARN endogène, le U7-PAS snARN est détecté de manière robuste dans les TA près de trois semaines après l'injection de vecteurs AAV-U7-PAS (Figure 46).

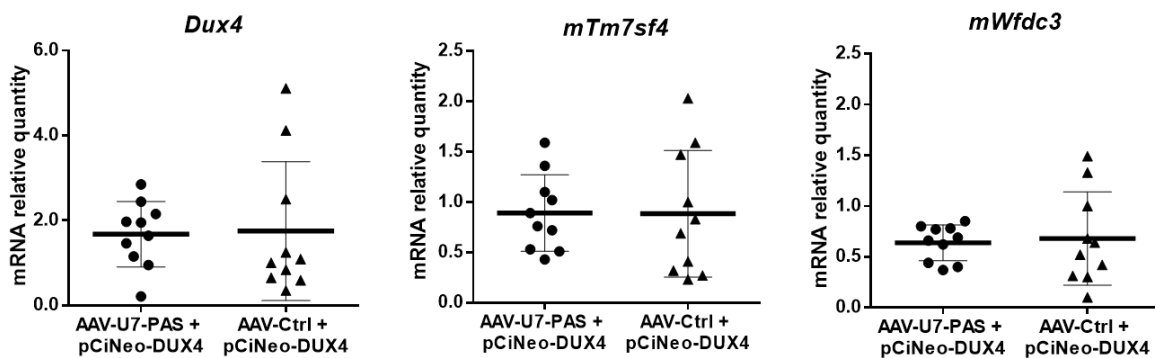


FIGURE 45. Expression de *DUX4* et de ses gènes cibles (*mTm7sf4* et *mWfdc3*) dans les TA des souris pCINéo-DUX4 traités avec le vecteur AAV-U7-PAS ou AAV-Ctrl. Deux semaines avant l'électroporation du plasmide d'expression de *DUX4* (pCINéo-DUX4) dans les deux TA des souris C57BL/6, les vecteurs AAV-U7-PAS (ronds noirs) ou AAV-Ctrl (triangles noirs) sont injectés en intramusculaire dans les TA. La quantité relative d'ARNm *DUX4* et de deux de ses gènes cibles, *mTmSf4* et *mWfdc3*, est mesurée par RT-PCR quantitative. *mGapdh* est utilisé comme gène de référence. Le niveau d'expression de ces gènes est rapporté à la quantité de plasmides pCINéo-DUX4 présents par noyaux cellulaires. Les données représentées correspondent à la moyenne (barre en gras noir) +/- écart-type standard (n=10). Un t-test pour des échantillons appariés a été réalisée sur ces données (\*P>0,05 ; \*\*P<0,01 ; \*\*\*P<0,001).

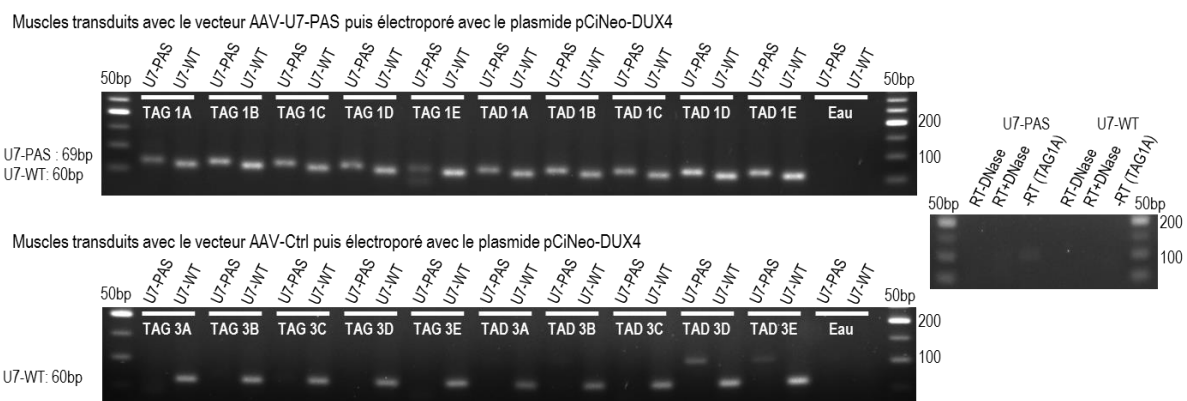


FIGURE 46. Détection du U7-snARN endogène (U7-WT) ou du U7-PAS snARN transcrit à partir du vecteur AAV-U7-PAS dans les TA des souris pCINéo-DUX4 traités avec le vecteur AAV-U7-PAS ou AAV-Ctrl. Les ARN totaux extraits des deux TA de chaque souris sont analysés par RT-PCR. Les produits d'amplification d'une taille de 69bp et de 60bp correspondent respectivement au U7-PAS snARN et U7-WT snARN.

### **2.2.3.2. Recherche d'une redirection du site de clivage *in vivo* dans les muscles de souris pCINéo-DUX4 traités avec le vecteur AAV-U7-PAS**

La littérature rapporte que des évènements de polyadénylation alternatifs peuvent avoir lieu lorsqu'un signal de polyadénylation canonique est non fonctionnel dans le dernier exon. Par ailleurs, nous avons montré au cours de notre étude précédente, que l'utilisation d'oligonucléotides antisens masquant un signal de polyadénylation et un site de clivage peut induire une redirection du site de clivage (Marsollier et al., 2016). Ainsi, nous avons recherché, par analyses *in silico* de la séquence du plasmide pCINéo-DUX4, la présence de signaux de polyadénylation alternatifs en aval de l'exon 3 de *DUX4*, pouvant être utilisés en présence du U7-PAS snARN (Beaudoing et al., 2000; Marsollier et al., 2016; Tian et al., 2005). Plusieurs signaux de polyadénylation alternatifs ont été mis en évidence (Figure 47). La majorité de ces signaux ont une faible fréquence d'utilisation chez la souris, à l'exception du signal AAUAAA, qui est le signal le plus détecté en région 3' des ARNm murins et humains (Beaudoing et al., 2000; Tian et al., 2005). Ce signal de polyadénylation fort est apporté par la séquence de polyadénylation du virus simien 40 (SV40-PolyA) (Figure 47).

Suite à ces analyses *in silico*, une redirection du signal de polyadénylation ou du site de clivage en présence du U7-PAS snARN a été examinée par 3'RACE PCR nichée. Deux amorces sens différentes de par leur localisation ont été utilisées : i) l'amorce sens 3RACE\_Exon3\_Nested\_fw, localisée dans l'exon 3, ii) l'amorce sens 3RACE\_pCINeo\_fw localisée entre le signal de polyadénylation canonique de *DUX4* et la séquence SV40-PolyA (Figure 47). Deux produits d'amplification de taille différente sont obtenus, indiquant un changement d'utilisation du signal de polyadénylation et du site de clivage pour certains ARNm *DUX4* en présence du U7-PAS snARN (Figure 47). Le séquençage de ces produits révèle que le signal de polyadénylation utilisé pour certains des ARNm *DUX4* est le signal de polyadénylation fort, apporté par la séquence SV40-PolyA. De manière intéressante, cette redirection du signal de polyadénylation est observée également en l'absence de U7-PAS snARN (Figure 47). Deux types d'ARNm *DUX4* stables sont donc produits à partir du plasmide pCINéo-DUX4, et ce indépendamment de la présence d'oligonucléotides antisens. Un ARNm *DUX4* utilise le signal de polyadénylation localisé dans l'exon 3 tandis que le second ARNm *DUX4* utilise le signal de polyadénylation localisé dans la séquence SV40-PolyA (Figure 47). Ces deux ARNm peuvent coder pour une protéine DUX4 fonctionnelle

puisque l'ORF est exclusivement présent dans l'exon 1 de *DUX4*. Le modèle de souris pCINéo-DUX4 ne nous permet donc pas de tester l'efficacité *in vivo* du U7-PAS snARN, dessiné pour cibler uniquement le signal de polyadénylation canonique localisé dans l'exon 3 de *DUX4*.

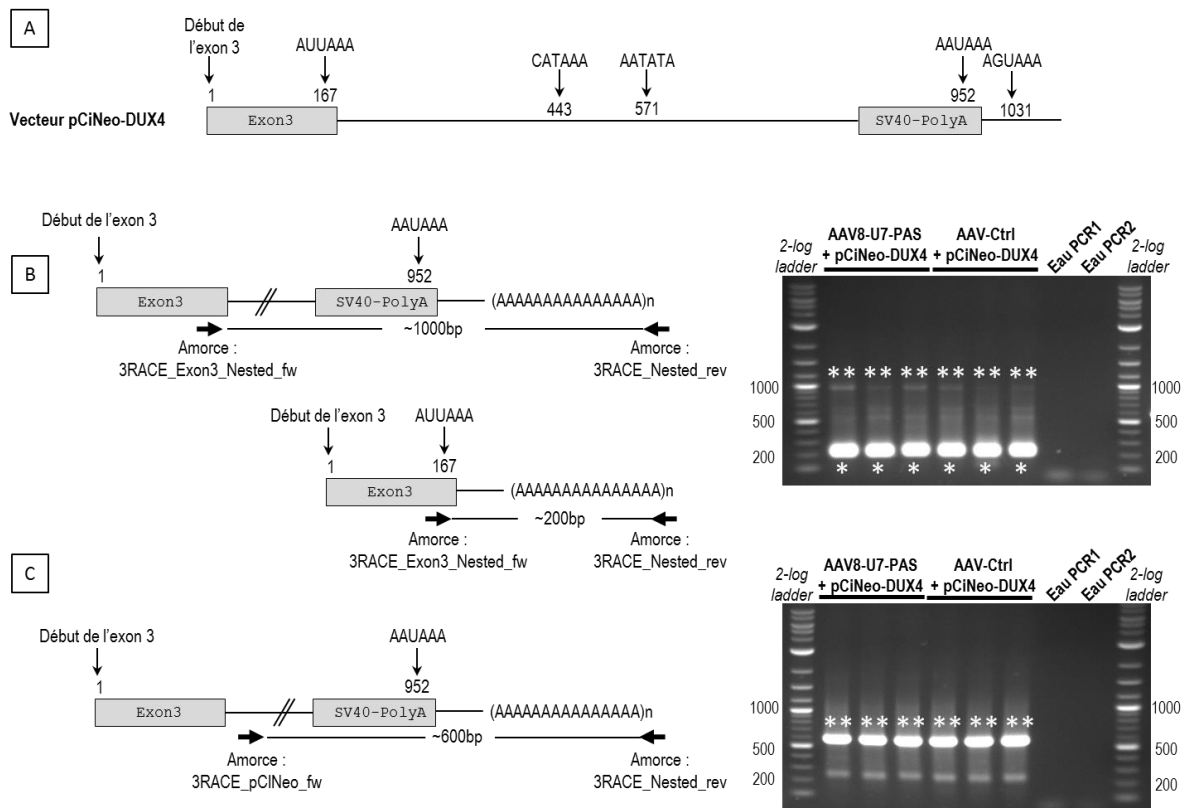


FIGURE 47. Recherche d'une redirection du site de clivage pour l'ARNm *DUX4* en présence du U7-PAS snARN dans les muscles de souris exprimant *DUX4* à partir du plasmide pCINéo-DUX4. A) Des analyses *in silico* du plasmide pCINéo-DUX4 ont révélé la présence de plusieurs signaux de polyadénylation alternatifs dont le signal de polyadénylation fort (AAUAAA) apporté par la séquence de polyadénylation du virus simien 40 (SV40 – PolyA). B) et C) Une redirection du site de clivage a été mise en évidence par 3'RACE PCR nichée en utilisant une amorce sens (3RACE-Exon3-Nested-fw) dans l'exon 3 de *DUX4* (B) ou une amorce sens (3RACE-pCiNeo-fw) localisée en aval du signal de polyadénylation canonique dans l'exon 3 de *DUX4* et la séquence SV40 – PolyA (C). Les bandes annotées d'un astérisque correspondent à une utilisation du signal de polyadénylation localisé dans l'exon 3 de *DUX4*. Les bandes annotées de deux astérisques correspondent à une redirection du signal de clivage associée à l'utilisation du signal de polyadénylation SV40-PolyA.

### 2.2.3.3. Détermination de l'expression de l'ARNm *DUX4* et de ses gènes cibles *in vivo* dans les muscles de souris pCINéo-DUX4 $\Delta$ SV40 traités avec le vecteur AAV-U7-PAS

Afin de tester l'efficacité *in vivo* du U7-PAS snARN à inhiber l'expression de *DUX4*, un autre modèle de souris a été généré cette fois-ci par électroporation intramusculaire du plasmide

pCINéo-DUX4  $\Delta$ SV40. Ce plasmide ne possède pas la séquence SV40-polyA en aval du gène *DUX4* et permet, en théorie, uniquement la transcription de la forme de l'ARNm *DUX4* utilisant le signal de polyadénylation localisé dans l'exon 3 de *DUX4*. Le protocole appliqué est le même que précédemment (Figure 44). Tout d'abord, la suppression de la séquence SV40-polyA du plasmide pCINéo-DUX4 n'affecte pas l'expression de *DUX4*. En effet, les ARNm *DUX4*, *mTm7sf4* et *mWfdc3* sont détectés dans les TA des souris pCINéo-DUX4  $\Delta$ SV40 traités avec l'AAV-Ctrl (Figure 48). La quantité relative d'ARNm *DUX4* est de 3,66 +/- 4,80 dans les TA traités AAV-Ctrl et de 1,63 +/- 1,90 dans les TA traités AAV-U7-PAS (Figure 48). La quantité relative d'ARNm *mTm7sf4* est de 4,87 +/- 7,73 dans les TA traités AAV-Ctrl et de 3,48 +/- 2,50 dans les TA traités AAV-U7-PAS (Figure 48). La quantité relative d'ARNm *mWfdc3* est de 6,12 +/- 10,30 dans les TA traités AAV-Ctrl et de 2,03 +/- 2,42 dans les TA traités AAV-U7-PAS (Figure 48). Le niveau d'expression de *DUX4* et de ses deux gènes cibles, *mTm7sf4* et *mWfdc3*, est donc inchangé entre les TA traités avec le vecteur AAV-U7-PAS et les TA traités avec le vecteur AAV-Ctrl (p-value des t-test pour des échantillons appariés = 0,31 ; 0,65 ; 0,32 pour *DUX4*, *mTm7sf4* et *mWfdc3* respectivement) bien que le U7-PAS snARN soit toujours exprimé fortement dans les muscles traités avec l'AAV-U7-PAS (Figure 48 et Figure 49).

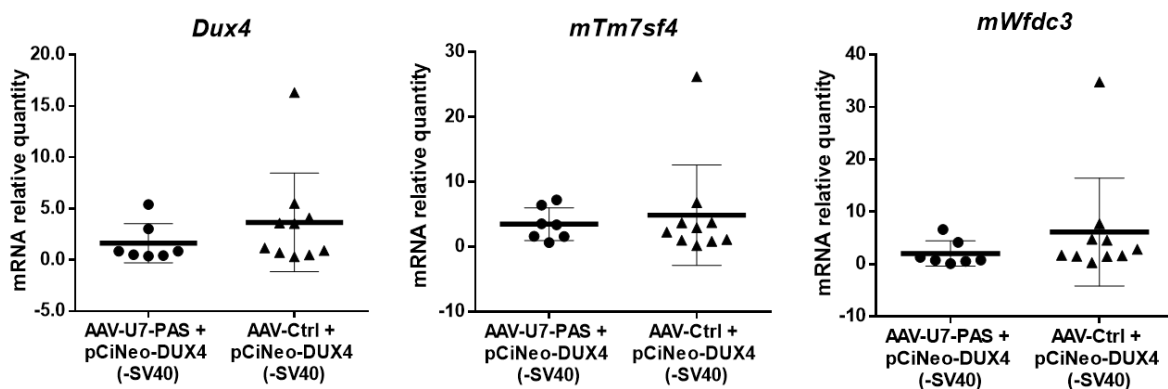


FIGURE 48. Expression de *DUX4* et de deux de ses gènes cibles (*mTm7sf4* et *mWfdc3*) dans les TA des souris pCINéo-DUX4  $\Delta$ SV40 traités avec le vecteur AAV-U7-PAS ou AAV-Ctrl. Les vecteurs AAV-U7-PAS (ronds noirs) et AAV-Ctrl (triangles noirs) sont injectés en intramusculaire dans les TA de souris C57BL/6. Deux semaines après l'injection, le plasmide d'expression de *DUX4*, pCINéo-DUX4  $\Delta$ SV40, sans la séquence codante pour le signal de polyadénylation du virus simien 40, est électroporé dans les deux TA des souris. La quantité relative d'ARNm *DUX4* et de deux de ses gènes cibles, *mTm7sf4* et *mWfdc3*, est mesurée par RT-PCR quantitative. *mGapdh* est utilisé comme gène de référence. Le niveau d'expression de ces gènes est rapporté à la quantité de plasmides pCINéo-DUX4  $\Delta$ SV40 présents par noyaux cellulaires. Les données représentées correspondent à la moyenne (barre en gras noir) +/- écart-type standard (n=8 souris injectées AAV-U7-PAS et n=10 souris injectées AAV-Ctrl). Un t-test pour des échantillons appariés a été réalisée sur ces données (\*P>0,05 ; \*\*P<0,01 ; \*\*\*P<0,001).

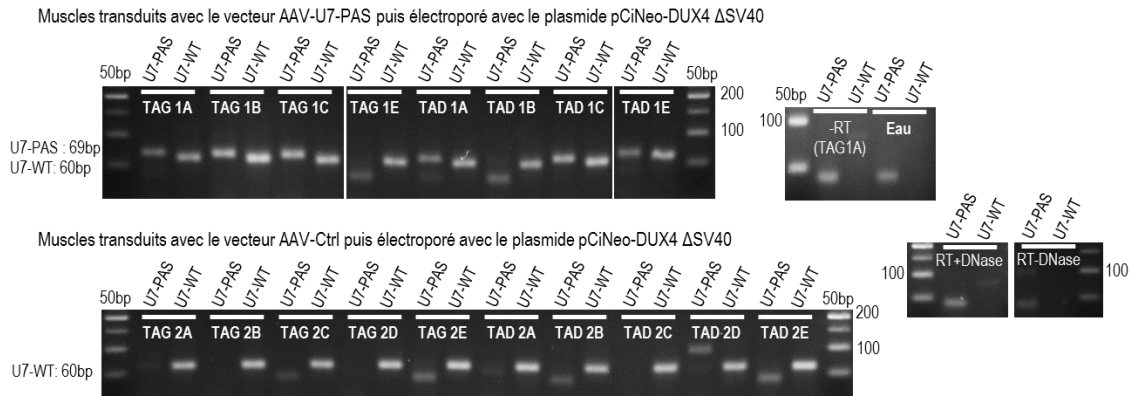


FIGURE 49. Détection du U7-snARN endogène (U7-WT) ou du U7-PAS snARN transcrit à partir du vecteur AAV-U7-PAS dans les TA des souris pCINéo- DUX4 ΔSV40 traités avec le vecteur AAV-U7-PAS ou AAV-Ctrl. Les ARN totaux extraits des deux TA de chaque souris sont analysés par RT-PCR. Les produits d'amplification d'une taille de 69bp et de 60bp correspondent respectivement au U7-PAS snARN et U7-WT snARN.

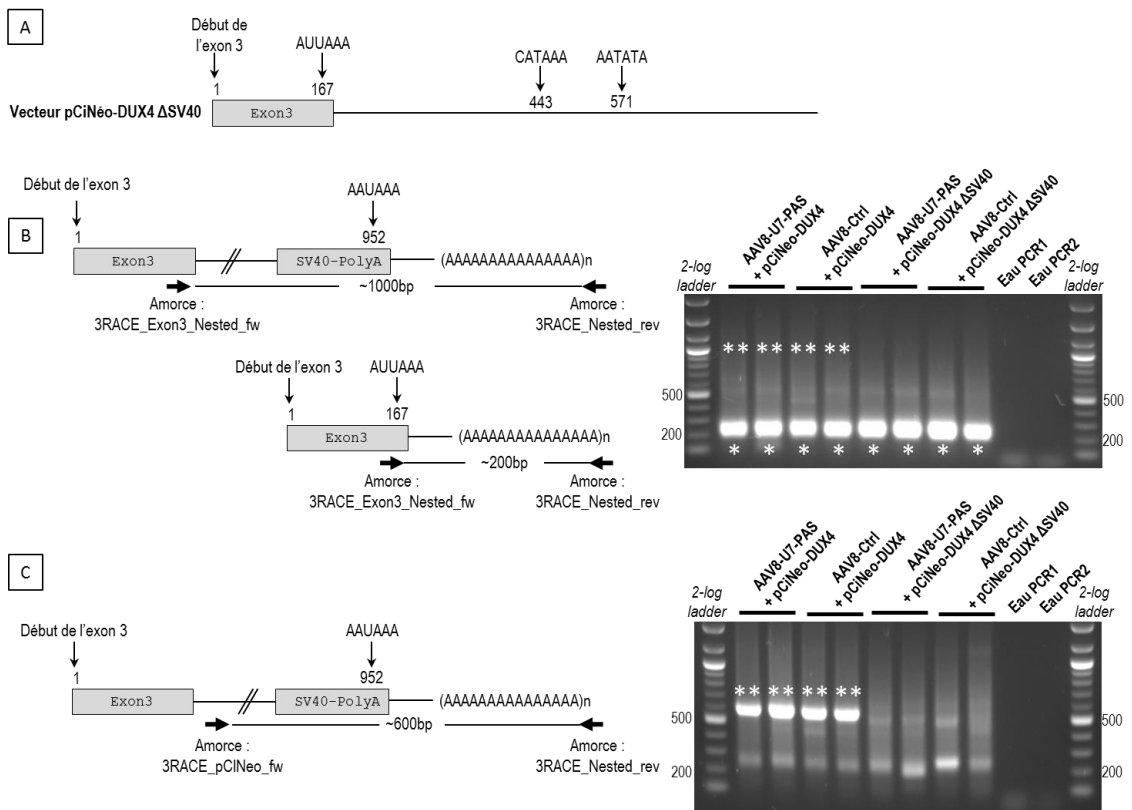


FIGURE 50. Recherche d'une redirection du site de clivage pour l'ARNm *DUX4* en présence du U7-PAS snARN dans les muscles de souris exprimant *DUX4* à partir du plasmide pCINéo-DUX4 ΔSV40. A) Des analyses *in silico* du plasmide pCINéo-DUX4 ΔSV40 ont révélé la présence de deux signaux de polyadénylation alternatifs associés à la région 3'UTR du gène *DUX4*. B) et C) Aucune redirection du site de clivage dans les muscles de souris pCINéo-DUX4 ΔSV40 n'a été mise en évidence par 3'RACE PCR nichée en utilisant une amorce sens (3RACE-Exon3-Nested-fw) dans l'exon 3 de *DUX4* (B) ou une amorce sens (3RACE-pCINeo-fw) localisée en aval du signal de polyadénylation canonique entre l'exon 3 de *DUX4* et la séquence SV40 – PolyA (C). Les bandes annotées d'un astérisque correspondent à une utilisation du signal de polyadénylation localisé dans l'exon 3 de *DUX4*. Les bandes annotées de deux astérisques correspondent à une redirection du signal de clivage associée à l'utilisation du signal de polyadénylation SV40-PolyA, observée uniquement dans des muscles de souris pCINéo-DUX4.

Sur le plasmide pCINéo-DUX4  $\Delta$ SV40, deux signaux de polyadénylation alternatifs faibles persistent à moins de 600bp en aval de l'exon 3 de *DUX4*. Ces signaux de polyadénylation alternatifs sont associés à la région 3'UTR de *DUX4*. Une redirection du signal de polyadénylation vers ces signaux de polyadénylation alternatifs en présence du U7-PAS snARN a été étudiée. Un produit unique d'amplification est obtenu par 3'RACE PCR nichée sur les extraits d'ARN totaux de TA de souris pCINéo-DUX4  $\Delta$ SV40 traités par le vecteur AAV-U7-PAS ou le vecteur AAV-Ctrl (Figure 50). Ce produit correspond à un ARNm *DUX4* utilisant le signal de polyadénylation localisé dans l'exon 3 (Figure 50). Aucune redirection du site de clivage n'est donc observée pour les ARNm *DUX4* en présence du U7-PAS snARN dans des muscles de souris pCINéo-DUX4  $\Delta$ SV40.

A la vue de l'ensemble de ces résultats, deux conclusions sont possibles : (i) le U7-PAS snARN transcrit n'est pas fonctionnel et ne permet pas de masquer efficacement le signal de polyadénylation pour empêcher l'ajout d'une queue poly(A) en 3' de l'ARNm *DUX4*, (ii) le U7-PAS snARN est fonctionnel mais n'est pas suffisamment exprimé pour inhiber l'ensemble des pré-ARNm *DUX4* transcrits à partir de la grande quantité de plasmides pCINéo-DUX4 +/- SV40 présents dans les TA des souris.

#### **2.2.3.4. Détermination de l'expression de l'ARNm *DUX4* *in vivo* dans les muscles de souris *ACTA1-MCM,FLexD/+* traités avec le vecteur AAV-U7-PAS**

En collaboration avec l'équipe du Dr Peter Jones, l'efficacité du U7-PAS snARN a également été testée dans le modèle de souris *ACTA1-MCM,FLexD/+*. Chez ces souris, une seule copie du transgène *DUX4*, comprenant les exons 1 à 3 et les introns 1 à 2, est présente sous contrôle du promoteur constitutif *ROSA26*. L'induction de l'expression de la recombinaise Cre par administration du tamoxifène permet une inversion de la séquence de *DUX4* par recombinaison aux sites *loxP* et *lox511* encadrant le transgène, et de ce fait, une expression de *DUX4*. L'expression de *DUX4* pouvant varier d'une souris à une autre, les vecteurs AAV-U7-PAS sont injectés dans le TA gauche tandis que les vecteurs AAV-Ctrl sont injectés dans le TA droit de souris mâles âgées de neuf semaines. Deux semaines plus tard, l'expression de *DUX4* est induite par injection intra-péritonéale de tamoxifène. Deux semaines après l'administration de tamoxifène, les souris sont sacrifiées et les TA prélevés (Figure 44). Le

niveau d'expression de *DUX4* en présence ou en absence du U7-PAS snARN est déterminé par PCR sur les ARN totaux extraits des TA. Une faible diminution de *DUX4*, associée à une expression du U7-PAS snARN, est observée pour certaines souris dans les TA traités avec les vecteurs AAV-U7-PAS en comparaison aux TA traités avec les vecteurs AAV-Ctrl (Figure 51 et Figure 52). Les résultats préliminaires obtenus chez ce modèle murin FSHD nous indiquent un potentiel effet du U7-PAS snARN *in vivo*.

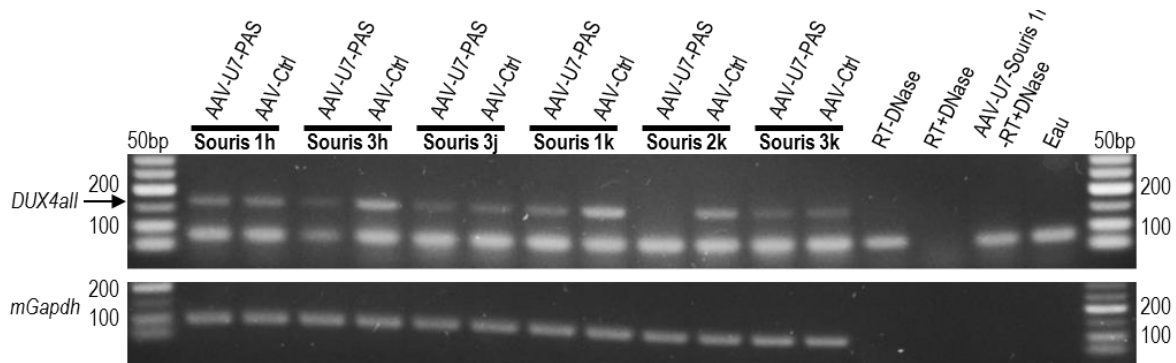


FIGURE 51. Expression de *DUX4* en présence de l'AAV-U7-PAS ou l'AAV-Ctrl dans des muscles de souris *ACTA1-MCM, FLexD/+*. Le niveau d'expression de l'ARNm *DUX4* a été étudié par PCR en utilisant les amorces *DUX4-all*. *mGapdh* est utilisé comme gène référence.

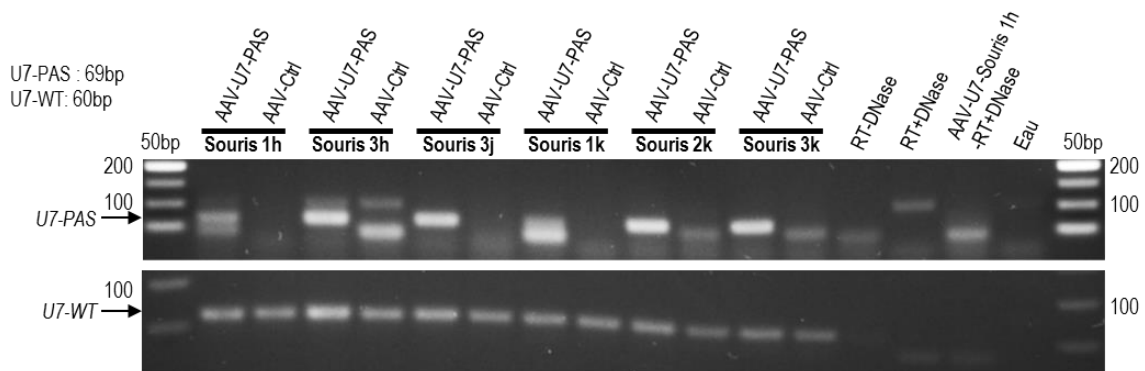


FIGURE 52. Détection du U7 naturel (U7-WT) snARN ou du U7-PAS snARN transcrit à partir du vecteur AAV-U7-PAS dans les TA des souris *ACTA1-MCM, FLexD/+* traités avec le vecteur AAV-U7-PAS ou AAV-Ctrl. Les ARN totaux extraits des deux TA de chaque souris sont analysés par RT-PCR. Les produits d'amplification d'une taille de 69bp et de 60bp correspondent respectivement au U7-PAS snARN et U7-WT snARN.

La séquence du transgène *DUX4* dans ce modèle murin s'arrête 235bp en aval du signal de polyadénylation canonique de *DUX4*. Les deux signaux de polyadénylation alternatifs CATAAA et AATATA, situés dans la région 3'UTR de *DUX4* dans le génome humain, ne sont donc pas présents chez ces souris (Figure 50). Cependant, il n'est pas exclu que des signaux de polyadénylation alternatifs soient localisés juste en aval du site d'insertion du transgène. Ne connaissant pas le site exact d'insertion du transgène, nous n'avons pas pu



réaliser une étude *in silico* pour rechercher la présence de signaux de polyadénylation alternatifs. Une redirection du site de clivage en présence ou en absence du U7-PAS snARN dans les TA traités avec le vecteur AAV-U7-PAS ou AAV-Ctrl a tout de même été étudiée par 3'RACE PCR nichée, avec l'amorce sens 3RACE\_Exon3\_Nested\_fw positionnée dans l'exon 3. Un seul produit d'amplification est obtenu, n'indiquant pas de redirection du site de clivage en présence du U7-PAS snARN (Figure 53).

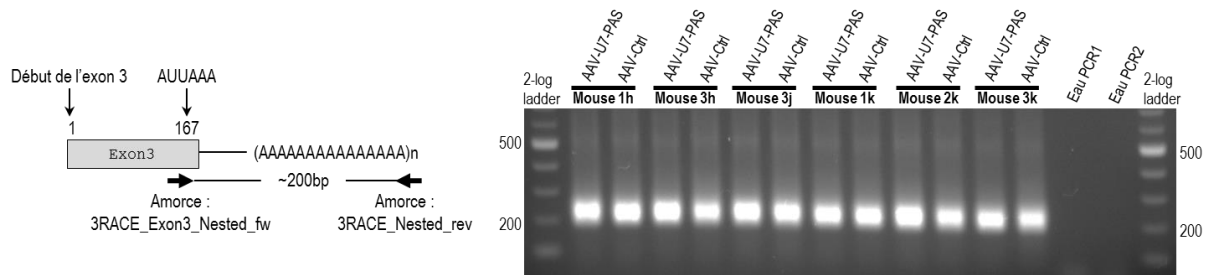


FIGURE 53. Recherche d'une redirection du site de clivage pour l'ARNm *DUX4* en présence ou en absence du U7-PAS snARN dans les muscles de souris *ACTA1-MCM, FLexD/+*. Le site de clivage utilisé pour l'ARNm *DUX4* dans ce modèle murin a été mis en évidence par 3'RACE PCR nichée en utilisant une amorce (3RACE-Exon3-Nested-fw) sens dans l'exon 3 de *DUX4*. Les produits d'amplification indiquent une utilisation du signal de polyadénylation localisé dans l'exon 3 de *DUX4*.

### **3.DISCUSSION ET PERSPECTIVES**

### **3.1. Nouvelle stratégie thérapeutique innovante pour la FSHD inhibant l'expression de *DUX4***

L'expression de *DUX4* est un point commun avéré entre les patients FSHD1 et FSHD2. De ce fait, plusieurs stratégies thérapeutiques visant l'expression de *DUX4* dans la FSHD ont été proposées, mais, à ce jour, aucune n'a encore atteint les essais cliniques (Mitsuhashi et al., 2013a; Vanderplanck et al., 2011; Wallace et al., 2012).

#### **3.1.1. Preuve de concept *in vitro* de la faisabilité d'une stratégie thérapeutique ciblant les séquences clés en 3'UTR du pré-ARNm *DUX4***

Au cours de ma thèse, nous avons développé *in vitro* une nouvelle stratégie innovante ciblant les séquences clés en 3'UTR du pré-ARNm de *DUX4* (Marsollier et al., 2016). L'objectif de cette stratégie est d'inhiber le mécanisme de polyadénylation. En effet, des défauts dans la maturation en 3' des ARNm conduisent à une perturbation de l'expression du gène concerné en impactant la stabilité, l'export nucléaire ou encore la traduction de l'ARNm (Danckwardt et al., 2008; Mangus et al., 2003; Millevoi and Vagner, 2010). Ainsi, empêcher la maturation en 3' du pré-ARNm *DUX4* permettrait sa déstabilisation et conduirait *in fine* à une inhibition de son expression.

Notre étude *in vitro* met en évidence la faisabilité d'une telle approche. L'utilisation d'oligonucléotides antisens ciblant les éléments clés en 3', c'est-à-dire le signal de polyadénylation et/ou le site de clivage et/ou le DSE, impliqués dans le processus de maturation des ARNm, entraîne une inhibition de l'expression de *DUX4* et de ses gènes cibles.

Sur les cinq PMO testés, deux engendrent une diminution importante de l'expression de l'ARNm *DUX4* pour les conditions expérimentales testées (Figure 54) :

- le PMO-PAS qui cible uniquement le signal de polyadénylation,
- le PMO-CS3 qui cible à la fois le signal de polyadénylation et le site de clivage.

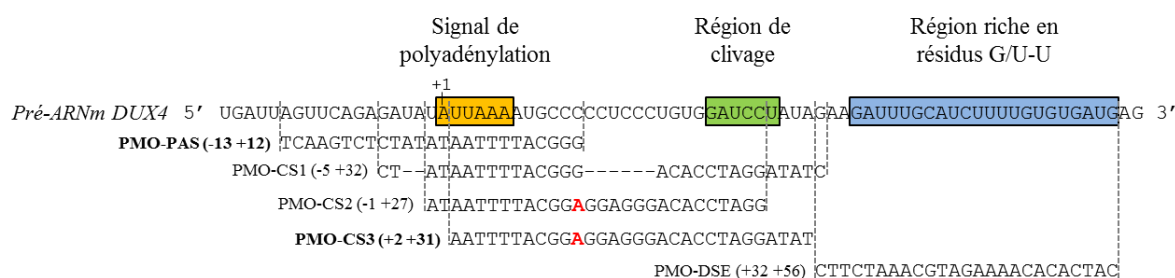


FIGURE 54. Séquences reconnues par les différents PMOs ciblant les éléments clés en région 3'UTR du pré-ARNm *DUX4*. Le signal de polyadénylation, la région de clivage, la région riche en résidus G/U-U (DSE), la séquence de chaque PMO, ainsi que sa position, sont indiqués. Les délétions sont représentées par des tirets (-). Les substitutions nucléotidiques sont signalées en rouge.

Chen et al. ont également confirmé l'efficacité du PMO-PAS (renommé FM10 dans leur étude, mais totalement identique à notre AO publié) (Chen et al., 2016a). En effet, ils ont observé *in vitro* une forte réduction de l'expression de *DUX4* dans des cellules primaires FSHD traitées à J0 de différenciation avec le PMO-PAS. Néanmoins, dans cette étude, les quantités de PMO-PAS employées pour obtenir un effet sur l'expression de *DUX4* sont 20 fois supérieures à celles que nous avons utilisées. Cette différence majeure pourrait être due à la méthodologie utilisée. Alors que nous avons privilégié un transfert par vecteur chimique (lipide cationique) des PMOs hybridés avec une « laisse » (oligonucléotide chargé), Chen et al. ont opté pour un transfert de PMOs nus, sans vecteur chimique ou physique. Les PMOs seuls présentent une faible capacité à pénétrer dans les cellules. De ce fait, pour qu'une proportion suffisante de PMOs atteigne le noyau et qu'un effet soit potentiellement observé, une grande quantité de PMOs nus doit nécessairement être utilisée pour traiter les cellules, d'où les différences de réponses observées entre nos deux études.

### 3.1.2. Mécanismes hypothétiques d'action des PMOs

Les PMOs, ciblant les éléments clés impliqués dans la maturation en 3' des ARNm, peuvent agir *via* plusieurs mécanismes pour inhiber l'expression du gène cible. Les PMOs forment un hybride stable avec l'ARNm cible, mais, de par leur chimie, ne recrutent pas la RNaseH1 (Kole et al., 2012). De ce fait, un clivage des ARNm hybridés à des PMOs est exclu. Le mécanisme le plus probable est que la mise en place de ce duplex stable PMO-ARNm empêche ou diminue, par encombrement stérique et/ou modification de la structure secondaire de l'ARNm, le recrutement des facteurs de clivage et/ou de polyadénylation. De cette manière, le PMO entrave le clivage de l'ARNm au site consensus et/ou l'ajout de la queue poly(A), ce qui déstabilise l'ARNm.

Une autre possibilité est que les PMOs n'entravent pas l'ajout de la queue poly(A), mais induisent son retrait (Filipowicz et al., 2008; Wada et al., 2012). Il en résulte alors une déstabilisation de l'ARNm et une inhibition de la traduction. Ce retrait de la queue poly(A) est dépendant de déadénylases et d'exonucléases qui reconnaissent le duplex AO-ARNm formé. A noter que ce mécanisme, qui a nécessairement lieu après la maturation en 3' de l'ARNm cible, n'est pas le mécanisme impliqué dans la diminution de l'ARNm *DUX4* en présence du PMO-DSE. En effet, le PMO-DSE s'hybride à la région riche en résidus G/U et U. Or cette région est éliminée juste après le clivage du pré-ARNm, c'est-à-dire avant l'ajout de la queue poly(A). Ainsi, le PMO-DSE ne peut pas former de complexe stable avec l'ARNm *DUX4* polyadénylé et ne peut donc pas induire le retrait de la queue poly(A) de celui-ci.

Les PMOs peuvent également impacter la stabilité de l'ARNm *DUX4* en modulant indirectement la voie du NMD. L'ARNm *DUX4*, de par sa longue région 3'UTR, est un substrat du NMD. Cependant, une expression de *DUX4* réduit l'efficacité du NMD, en induisant, selon un mécanisme inconnu, la dégradation d'UPF-1, une protéine effectrice du NMD (Feng et al., 2015). L'inhibition du NMD par *DUX4* permet alors une stabilisation de son propre ARNm. De ce fait, une diminution, même faible, de la quantité d'ARNm *DUX4*, conduit à une diminution en protéines *DUX4*, à une augmentation en protéines UPF-1 et donc à une augmentation de l'efficacité du NMD qui peut alors déstabiliser l'ARNm *DUX4*. L'effet déstabilisateur du PMO sur l'ARNm *DUX4* pourrait ainsi être amplifié.

### **3.1.3. Evaluation de l'efficacité de cette stratégie *in vivo***

Au cours de leur étude, Chen et al. ont également démontré la faisabilité d'une telle approche *in vivo* sur 14 jours, par électroporation des PMO-PAS nus dans des xéno greffes de tissus musculaires FSHD chez des souris immunodéficientes. Une diminution de l'expression de *DUX4* et de deux de ses gènes cibles est observée. Cependant, le choix du mode de transfert de l'AO n'est probablement pas le plus adapté pour une application de cette approche thérapeutique chez l'Homme. En effet, tout d'abord, les PMOs sont rapidement éliminés par l'organisme. De plus, la FSHD est une pathologie du muscle, tissu qui représente environ 35 à 40% de la masse totale d'un individu. Un traitement de thérapie génique pour la FSHD implique d'apporter la molécule thérapeutique dans l'ensemble des muscles, ou au moins dans

les muscles atteints. De ce fait, pour espérer obtenir un effet thérapeutique sur le long terme dans le cas de pathologies musculaires, l'utilisation de PMOs pour le traitement d'êtres vivants de grande taille nécessite des injections répétées et conséquentes d'AOs. Outre le coût des PMOs, ce mode de transfert des AOs peut s'avérer très contraignant pour le patient. Pour éviter aux patients des injections répétées et obtenir un effet sur le long terme de l'AO, nous évaluons donc *in vivo* une approche alternative consistant à vectoriser la séquence de l'oligonucléotide antisens PMO-PAS dans un AAV sous contrôle du promoteur U7. Ce vecteur AAV-U7-PAS a été injecté dans deux modèles murins :

1-) L'injection de l'AAV-U7-PAS dans un modèle de souris exprimant *DUX4* par électroporation intramusculaire du plasmide pCINéo-DUX4 (+/- SV40) n'induit pas de diminution de l'expression de *DUX4* et de ses gènes cibles. Ces résultats laissent penser que l'AAV-U7-PAS n'est pas efficace *in vivo* bien qu'il soit exprimé. Cette inefficacité pourrait s'expliquer par une expression insuffisante du U7-PAS snARN, par une faible reconnaissance de l'ARNm *DUX4* par le U7-PAS snARN ou par une faible stabilité du duplex ARNm-U7-PAS-snARN.

2-) Néanmoins, dans le modèle de souris *ACTA1-MCM,FLexD/+*, l'AAV-U7-PAS induit une légère diminution de l'expression de *DUX4*. Ces observations préliminaires doivent être confirmées par l'étude des gènes cibles de *DUX4*. Par ailleurs, un plus grand nombre de répliques est nécessaire, afin de réduire au minimum les biais expérimentaux dus à une expression de *DUX4* variable d'une souris à une autre, car induite par le tamoxifène. Dans le cas d'une efficacité avérée du U7-PAS snARN, des études histologiques et fonctionnelles sont également prévues pour déterminer les répercussions musculaires positives et négatives d'une expression du U7-PAS snARN.

Les deux modèles murins utilisés pour éprouver cette stratégie *in vivo* présentent une différence majeure pouvant expliquer ces résultats contradictoires. En effet, *DUX4* est exprimé à partir d'un transgène, présent sous une forme différente entre ces deux modèles murins, ce qui peut notamment impacter sa fréquence de transcription. Le modèle de souris pCINéo-DUX4 (+/- SV40) exprime *DUX4* à partir d'un plasmide. L'électroporation permet le transfert de plusieurs copies de plasmide pCINéo-DUX4 par noyaux de cellules musculaires (et dans les cellules avoisinantes). Bien que des plasmides ne soient pas présents dans toutes les cellules musculaires, il en résulte tout de même dans ce modèle murin une expression

robuste de *DUX4* détectée par PCRq. A l'inverse, dans le modèle de souris *ACTA1-MCM,FLexD/+*, un seul transgène *DUX4*, sous régulation du promoteur ubiquitaire *ROSA26*, est présent par génome et ce dans tous les noyaux de la fibre musculaire. De plus, l'expression de *DUX4* est dépendante d'une expression de la recombinaise Cre qui, dans ce modèle murin, est sous régulation du promoteur muscle-spécifique *ACTA1 (Actin Alpha1)*. Il en résulte, chez ces souris, une faible expression de *DUX4*. On suppose donc que la quantité relative d'ARNm *DUX4* transcrit dans les cellules murines de souris *ACTA1-MCM,FLexD/+* est moins importante que ce qui est observé chez la souris pCINéo-*DUX4 (+/- SV40)*. Une comparaison directe entre ces deux modèles de l'expression de *DUX4* et de ses gènes cibles en l'absence de traitement permettra de confirmer ou non cette hypothèse. Si cette supposition est avérée, il est possible que la quantité d'ARNm *DUX4* présente dans les cellules musculaires du modèle murin pCINéo-*DUX4* soit trop importante pour une bonne efficacité du U7-PAS snARN.

### 3.1.4. Optimisation de la stratégie thérapeutique

Afin d'améliorer cette stratégie thérapeutique pour la FSHD, plusieurs modifications de l'AAV-U7-PAS sont à envisager.

Tout d'abord, le niveau d'expression du U7-PAS, probablement insuffisant, peut être amélioré soit en augmentant le nombre de vecteurs par cellule, soit en augmentant sa fréquence de transcription en modifiant par exemple la séquence promotrice U7 ou les séquences régulatrices.

De plus, nous n'avons testé *in vivo* qu'un seul oligonucléotide antisens, l'un des plus performants *a priori* parmi les cinq PMO éprouvés *in vitro*. Le même protocole pourrait être appliqué au PMO-CS3 pour lequel nous avons obtenu *in vitro* une importante diminution de l'expression de *DUX4* et de ses gènes cibles (mais une redirection du site de clivage pour l'ARNm *DUX4*). De plus, les séquences des AOs utilisés *in vitro* ne sont peut-être pas les plus compétitives. Plusieurs facteurs peuvent impacter la stabilité du duplex pré-ARNm-AO : la longueur de l'AO, le nombre de mésappariements entre ce dernier et la séquence cible, des délétions dans la séquence de l'AO ou encore la localisation de l'AO sur l'ARN vis-à-vis des structures secondaires (tiges boucles) présentes. Une étude avec pour objectif une optimisation de la séquence des AOs (décalage de la séquence de l'AO de quelques bases en

5' ou 3', diminution ou augmentation de sa taille, insertion de mésappariements) est prévue. Enfin, afin d'augmenter l'efficacité de cette stratégie, il est envisagé d'injecter plusieurs AOs (de préférence non chevauchants) ciblant les différentes séquences régulatrices.

### **3.1.5. Avantages et inconvénients de cette stratégie thérapeutique en comparaison aux autres stratégies proposées**

Une stratégie thérapeutique avec des AOs ciblant les éléments clés en 3'UTR de l'ARNm présente plusieurs avantages comparés aux approches classiques (*i.e.* saut d'exons par AOs, inhibition de la traduction ou dégradation de l'ARNm par miARN ou siARN). Bien que cette stratégie ne puisse pas être utilisée pour réparer une mutation dans un gène ou modifier un ARNm mais uniquement utilisée pour inhiber l'expression génique, elle permet, à la différence des autres approches, de cibler la majorité des ARNm puisque le clivage et la polyadénylation sont des mécanismes essentiels à la maturation en 3' des ARNm des eucaryotes (Danckwardt et al., 2008; Millevoi and Vagner, 2010; Proudfoot, 2011; Zhao et al., 1999). Cette stratégie permet également de réguler l'expression de gènes avec un seul exon, ce qui n'est pas le cas avec une stratégie ciblant les sites donneurs ou accepteurs d'épissage. Par ailleurs, dans le cas de la FSHD, l'ORF de *DUX4* est localisé uniquement dans l'exon 1. Il est donc extrêmement difficile de développer une stratégie qui perturbe le cadre de lecture du gène avec des oligonucléotides antisens. Une stratégie thérapeutique ciblant les éléments clés en 3'UTR de l'ARNm est donc une alternative intéressante.

A la différence de stratégies utilisant des miARN ou siARN, le mécanisme d'action des AOs ciblant les éléments clés en 3'UTR est indépendant des protéines intracellulaires. Alors que l'action des miARN ou siARN ne peut s'effectuer que dans le cytoplasme, l'action des AOs dans cette stratégie est nucléaire et éventuellement cytoplasmique. En effet, il est envisageable que certains des PMOs (mais pas le U7 snARN) testés dans cette stratégie puissent s'hybrider dans le cytoplasme avec l'ARNm cible et ainsi soit induire le retrait de sa queue poly(A), soit inhiber sa traduction. De plus, les stratégies utilisant les miARN et siARN sont souvent associées à de nombreux effets *off-targets* (modification de l'expression de gènes non ciblés) (Kole et al., 2012). Or, au cours de notre étude, et de celle de Chen et al., aucun effet *off-target* n'a été associé à l'utilisation des PMO.



Au cours de leur étude, Chen et al. ont également testé des AOs masquant un site accepteur d'épissage (à la jonction intron 1 - exon 2). Les résultats obtenus montrent que les AOs masquant un site accepteur d'épissage sont moins efficaces pour diminuer l'expression de *DUX4* que les AOs ciblant les éléments clés en 3'UTR de l'ARNm (Chen et al., 2016a). Une comparaison de l'efficacité entre des AOs ciblant les sites accepteurs d'épissage à la jonction intron 2 - exon 3 et des AOs ciblant les éléments clés en 3'UTR de l'ARNm serait intéressante pour déterminer la meilleure stratégie à adopter pour le traitement de la FSHD.

En conclusion, nos résultats *in vitro* démontrent la faisabilité d'une approche thérapeutique pour la FSHD en ciblant les éléments clés en 3' de l'ARNm *DUX4*. Jusque-là, l'absence d'un modèle animal robuste exprimant *DUX4* sur le long terme rendait difficile le développement de cette approche *in vivo*. La commercialisation depuis fin février 2017 du modèle murin *ACTA1-MCM,FLexD* exprimant *DUX4* de manière stable nous donne l'opportunité d'éprouver l'efficacité de l'oligonucléotide antisens PMO-PAS. La compilation de nos résultats préliminaires dans ce modèle murin avec ceux obtenus par Chen et al. indiquent une potentielle capacité de l'AO PMO-PAS à diminuer l'expression de *DUX4 in vivo*.

## **3.2. Redirection du site de clivage en présence d'oligonucléotides antisens ciblant les éléments clés en 3' du pré-ARNm**

### **3.2.1. Mécanismes hypothétiques expliquant une redirection du site de clivage en présence d'oligonucléotides antisens**

La littérature rapporte que plus de 50% des ARNm ont au moins deux signaux de polyadénylation (Beaudoing et al., 2000; Tian et al., 2005). Les événements de polyadénylation alternatifs constituent un mécanisme de contrôle de l'expression des gènes dans l'espace et le temps (Danckwardt et al., 2008; Hall-Pogar et al., 2005; Zhao and Manley, 1996). Pour les ARNm contenant plusieurs signaux de polyadénylation, le choix du site de clivage est influencé par : i) la composition et la taille des séquences régulatrices, dont le signal de polyadénylation et le DSE, ii) la concentration ou l'activité des facteurs de polyadénylation aux différentes séquences régulatrices et iii) la présence ou non de facteurs régulateurs spécifiques exprimés à des temps ou dans des tissus particuliers (Barabino and

Keller, 1999; Danckwardt et al., 2008). Par exemple, la protéine PABPN1 intervient dans les événements de polyadénylation alternatifs en supprimant l'usage de signaux de polyadénylation alternatifs faibles (Jenal et al., 2012). Par ailleurs, des événements de polyadénylation alternatifs peuvent avoir lieu lorsque le signal de polyadénylation est défaillant dans le dernier exon.

*DUX4* appartient à la famille des ARNm possédant plusieurs signaux de polyadénylation. Sur une distance de 1kb en aval du signal de polyadénylation canonique ATTTAAA, huit signaux de polyadénylation alternatifs sont identifiés (Figure 55) (Beaudoing et al., 2000; Tian et al., 2005). Cependant, aucun phénomène de polyadénylation alternatif n'a été décrit pour *DUX4* à ce jour. De plus, des mutations dans le signal de polyadénylation canonique, comme c'est le cas au niveau du chromosome 10qA, n'induisent pas d'évènements de polyadénylation alternatifs, ni la production d'un ARNm *DUX4* stable (Lemmers et al., 2010a).

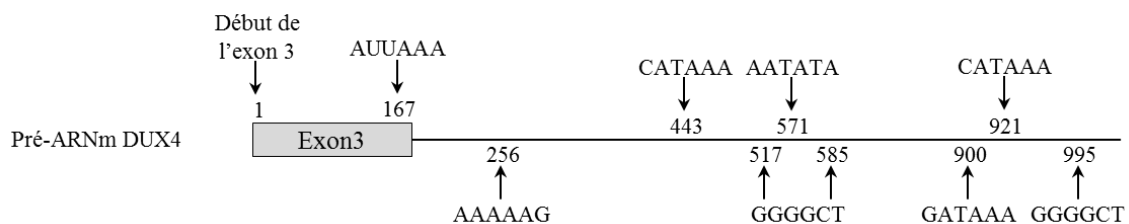


FIGURE 55. Localisation des différents signaux de polyadénylation alternatifs sur une distance de 1kb en aval du signal de polyadénylation canonique de *DUX4*.

### 3.2.1.1. Mécanismes hypothétiques de redirection du site de clivage *in vitro*

L'utilisation d'AOs ciblant les éléments clés en 3' du pré-ARNm ne conduit pas toujours à une déstabilisation de l'ARNm. En effet, au cours de nos études, nous avons observé que les AOs pouvaient induire une redirection du site de clivage. L'ARNm alors produit est stable et présente une queue poly(A). Le mécanisme à l'origine de cette redirection en présence des AOs n'a pas encore été élucidé. Nous supposons que l'AO génère un encombrement stérique ou modifie la structure secondaire du pré-ARNm au niveau du signal de polyadénylation et/ou du site de clivage. Les facteurs de polyadénylation et/ou de clivage sont alors forcés d'utiliser un autre signal de polyadénylation ou site de clivage généralement localisé en aval des signaux canoniques.

L'utilisation du PMO-CS3 *in vitro* induit, pour une certaine fraction d'ARNm *DUX4*, une redirection vers un site de clivage à environ 40 nucléotides en amont du signal de polyadénylation canonique, alors que dans les quelques articles décrivant ce phénomène, la redirection du site de clivage s'effectue en aval (Danckwardt et al., 2008; Raz et al., 2014; Vickers et al., 2001). Par ailleurs, ce nouveau site de clivage n'est associé à aucun signal de polyadénylation alternatif décrit (Beaudoing et al., 2000; Tian et al., 2005). Le PMO-CS3 masque en partie le signal de polyadénylation canonique et le site de clivage. Une hypothèse est qu'en l'absence d'un signal de polyadénylation alternatif fort, c'est-à-dire pour lequel les facteurs de polyadénylation et de clivage présentent une forte affinité, une redirection du site de clivage pourrait s'effectuer vers un nouveau site localisé quelques bases en aval ou en amont du signal de polyadénylation canonique. Cette redirection de quelques bases implique tout de même, en présence de l'AO, une reconnaissance des séquences régulatrices canoniques de polyadénylation de l'ARNm *DUX4* par les facteurs protéiques impliqués dans la maturation en 3' des ARNm. Bien que dans cette hypothèse l'AO n'empêche pas le recrutement des protéines (en raison d'une hybridation partielle de l'AO à l'ARNm cible par exemple), la présence de l'AO au site de clivage empêche le clivage endonucléolytique (modification de la structure secondaire, encombrement stérique). Une réorganisation des protéines permettrait alors d'effectuer un clivage en amont (ou en aval) du signal de polyadénylation. Le choix de ce nouveau site de clivage serait guidé : i) par la structure secondaire du pré-ARNm ou ii) par une interaction entre l'ARN polymérase, les différents facteurs de clivage et de polyadénylation et/ou d'autres facteurs. Cette hypothèse peut être étudiée en modifiant un signal de polyadénylation alternatif de *DUX4* faible en un signal de polyadénylation fort (ATTAAA par exemple). Ainsi, si l'absence de signal de polyadénylation fort est la condition qui induit une redirection du site de clivage en amont du signal de polyadénylation canonique, en présence du PMO-CS3 et d'un signal de polyadénylation fort alternatif sur l'ARNm *DUX4*, une redirection vers ce nouveau signal devrait être détectée.

Une redirection du signal de polyadénylation n'est observée qu'avec le PMO-CS3. Or, le PMO-CS1 et le PMO-CS2, comme le PMO-CS3, masquent le signal de polyadénylation et la région de clivage. L'efficacité de ces PMOs à diminuer l'expression de *DUX4* est moins importante que pour le PMO-CS3. Il est ainsi envisageable que : i) ces PMOs permettent, comme tous les PMOs testés, uniquement une inhibition de l'ajout qu'une queue poly(A) ou

la déadénylation de l'ARNm, et n'induisent pas de redirection du site de clivage (en raison d'une plus faible stabilité du duplex PMO-ARNm par exemple) ou ii) ces PMOs, en plus d'une inhibition de la polyadénylation ou une activation de la déadénylation de l'ARNm *DUX4*, induisent une redirection du site de clivage pour une très faible fraction d'ARNm *DUX4*, ce qui n'est pas détecté par 3'RACE-PCR dans nos conditions expérimentales.

A l'inverse des PMO précédents, le PMO-PAS et le PMO-DSE ne masquent qu'une seule des trois séquences régulatrices : le signal de polyadénylation ou la région riche en G/U ou U respectivement. Le PMO-DSE présente une faible efficacité *in vitro*. Une redirection du site de clivage en présence de ce PMO n'est peut-être pas détectée. Cependant, il est difficile d'imaginer qu'une redirection du site de clivage puisse se produire avec cet AO. En effet, la région riche en G/U ou U est une séquence régulatrice dont les limites ne sont pas aussi bien déterminées que pour le signal de polyadénylation ou le site de clivage. Ainsi, il est possible qu'un AO ciblant en partie cette région n'induisse qu'un repositionnement du facteur CstF sans compromettre le clivage au site canonique. En ce qui concerne le PMO-PAS, qui présente une efficacité *in vitro* à diminuer l'expression de *DUX4* similaire à celle du PMO-CS3, l'absence de redirection du site de clivage en présence de ce PMO peut s'expliquer par sa position sur l'ARNm cible. En effet, à la différence de tous les autres PMOs, le PMO-PAS couvre entièrement le signal de polyadénylation ainsi que 13 nucléotides en 5' et six nucléotides en 3' du signal de polyadénylation. De ce fait, la formation d'un duplex PMO-PAS-ARNm stable abolirait totalement le recrutement du facteur CPSF au signal de polyadénylation. Le PMO-PAS agirait alors de la même manière qu'une mutation dans le signal de polyadénylation canonique, conduisant à la production d'un ARNm *DUX4* instable, non polyadénylé.

### **3.2.1.1. Mécanismes hypothétiques de redirection du site de clivage *in vivo***

*In vivo*, dans le modèle murin pCINéo-DUX4 (+SV40), une redirection du site de clivage est également observée pour une fraction des ARNm *DUX4*. Cette redirection du site de clivage se caractérise par une utilisation du signal de polyadénylation (AATAAA) du SV40-PolyA, localisé en aval, à 791pb, du signal de polyadénylation canonique (ATTA AAA) (Figure 56). Cependant, cette redirection est indépendante de la présence du U7-PAS snARN. Le signal de polyadénylation AATAAA est le signal de polyadénylation le plus fréquemment utilisé pour

la détermination du site de clivage des ARNm. Le signal de polyadénylation canonique de *DUX4* est le deuxième signal de polyadénylation le plus représenté (Beaudoing et al., 2000; Tian et al., 2005). De ce fait, lors de la production du pré-ARNm *DUX4* à partir du plasmide pCiNeo-DUX4 (+SV40), les facteurs impliqués dans la maturation en 3' des ARNm reconnaissent et sont probablement concentrés, dans des proportions équivalentes, au niveau de la séquence régulatrice de polyadénylation canonique du pré-ARNm *DUX4* et au niveau de la séquence régulatrice de polyadénylation du SV40-poly(A). Par un mécanisme encore inconnu, dépendant sûrement de l'interaction avec d'autres facteurs protéiques, le clivage et la polyadénylation sont alors réalisés en amont de l'un ou l'autre des signaux de polyadénylation.

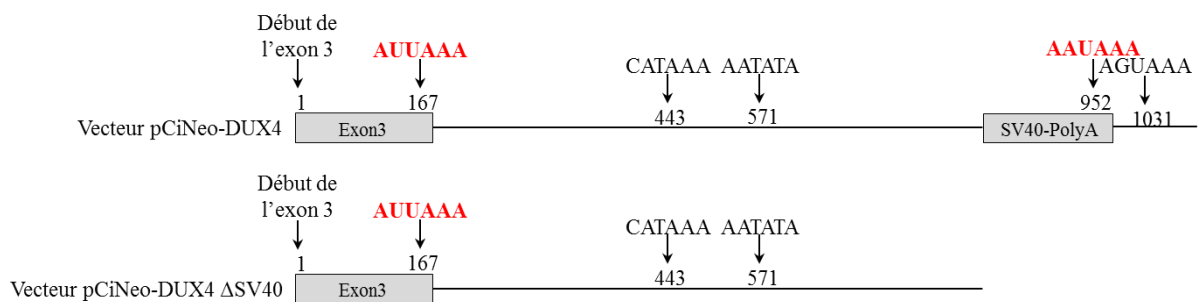


FIGURE 56. Localisation des différents signaux de polyadénylation alternatifs en aval du signal de polyadénylation canonique de *DUX4* sur les plasmides pCiNeo-DUX4 et pCiNeo-DUX4  $\Delta$ SV40. Les signaux de polyadénylation utilisés sont indiqués en rouge.

En l'absence de la séquence SV40-poly(A) (modèle murin pCiNeo-DUX4  $\Delta$ SV40 et *ACTA1-MCM, FLexD/+*), et malgré la présence de signaux de polyadénylation alternatifs (dont les signaux CATAAA et AATATA), mais faibles, localisés immédiatement en aval du signal de polyadénylation canonique, aucune redirection du site de clivage n'est observée en présence ou absence du U7-PAS snARN (Figure 56).

### 3.2.2. Conséquences d'une redirection du site de clivage sur la stabilité de l'ARNm

Cette redirection du site de clivage induite par les AOs est un point important à prendre en compte dans le cadre du développement d'une approche thérapeutique pour inhiber l'expression d'un gène en utilisant des AOs ciblant les éléments clés en 3' impliqués dans le mécanisme de polyadénylation. En effet, une redirection du site de clivage dans une région

3'UTR conduit à la production de plusieurs isoformes d'ARNm qui varient de par la longueur de leur séquence 3'UTR mais codent généralement pour une même protéine (Di Giammartino et al., 2011; Lutz and Moreira, 2011). Dans les différents cas de redirection du site de clivage observés au cours de notre étude, l'ARNm *DUX4* produit contient toujours les exons 1 et 2 ainsi que l'exon 3 modifié. L'exon 3 en présence du PMO-CS3 est délété de quelques bases (environ 40) tandis qu'il est rallongé de plusieurs centaines de bases dans le modèle murin pCINéo-DUX4 (+SV40). Or l'ORF codant pour la protéine DUX4 est localisé dans l'exon 1, et n'est donc pas affecté par les redirections du site de clivage listées précédemment. Ces ARNm sont alors tous traduits en une protéine DUX4 fonctionnelle. Au contraire, si l'exon 3 codait également pour la séquence polypeptidique, la redirection du site de clivage générée par la présence du PMO-CS3 aurait conduit à la production d'une protéine tronquée.

La taille de la région 3'UTR impacte la stabilité de l'ARNm et l'efficacité de la traduction. En effet, il semble qu'un ARNm avec une longue région 3'UTR soit moins stable que le même ARNm avec une région 3'UTR plus courte (Di Giammartino et al., 2011). Cette hypothèse est confirmée par l'étude de patients atteints du syndrome IPEX (*Immune Dysfunction, Polyendocrinopathy, Enteropathy, X-linked*) qui présentent un dysfonctionnement des cellules T régulatrices. Chez ces patients, le gène *FOXP3* (*Forkhead Box P3*), qui code pour un facteur de transcription, présente une mutation dans le signal de polyadénylation canonique (AAUAAA → AAUGAA). Cette mutation n'abolit pas le signal de polyadénylation mais le transforme en un signal de polyadénylation plus faible. Il en résulte alors une redirection du site de clivage avec utilisation du signal de polyadénylation alternatif suivant, localisé à 5,1 kb en aval du signal de polyadénylation canonique. L'ARNm *FOXP3* généré possède une longue région 3'UTR et est très instable. La protéine FOXP3 est alors peu exprimée chez les patients IPEX comparés aux individus contrôles (Bennett et al., 2001). L'observation inverse est réalisée chez des patients avec un lymphome à cellules du manteau (*MCL Mantle Cell Lymphoma*). Une isoforme de l'ARNm de *CCND1* (*CyclineD1*) avec une région 3'UTR plus courte est détectée chez les patients mais n'est pas retrouvée chez des individus contrôles. La réduction de la région 3'UTR de cet ARNm est due à l'utilisation d'un nouveau signal alternatif de polyadénylation fort. Ce signal est localisé en amont du signal de polyadénylation canonique et est généré par une mutation. La réduction de la région 3'UTR conduit à la perte d'éléments déstabilisateurs de l'ARNm et permet donc sa stabilisation (Wiestner et al., 2007). Cependant, ceci n'est pas toujours avéré. En effet, l'ARNm *BZWI* (*Basic Leucine Zipper And W2 Domains 1*), qui code pour un facteur de transcription, existe

sous trois isoformes créées par utilisation de différents signaux de polyadénylation alternatifs. Des études ont montré que contrairement à ce qui est attendu, c'est l'isoforme la plus courte qui présente l'efficacité de traduction la plus faible (Sood et al., 2006; Tanguay and Gallie, 1996; Yu et al., 2006). Par ailleurs la taille de la région 3'UTR régule la localisation de l'ARNm dans la cellule, ce qui a pour conséquence de modifier son expression (Andreassi and Riccio, 2009; Kislauskis and Singer, 1992). Au vu de ces observations, il est difficile de prédire la stabilité ou l'efficacité de traduction des différents ARNm *DUX4* obtenus dans nos études suite à une redirection du site de clivage.

Les longues régions 3'UTR contiennent généralement de nombreux sites de fixation des miARN ou des protéines de liaison à l'ARN, permettant une régulation post-transcriptionnelle de l'expression du gène (Di Giammartino et al., 2011; Pasquinelli, 2012). Dans l'OPMD, les ARNm produits sont raccourcis en 3'UTR suite à l'utilisation de sites de polyadénylation non canoniques en amont du signal de polyadénylation canonique. Ces ARNm, en raison du raccourcissement de leur région 3'UTR, ont perdu des sites de fixation des miARN ou de protéines permettant une régulation post-transcriptionnelle. La dérégulation générale de la polyadénylation des ARNm observée est due à une mutation dans la protéine PABPN1. Une courte expansion des triplets (GCG) dans la séquence codante de la protéine PABPN1 conduit à une courte expansion d'acides aminés alanine dans la partie N-terminale de la protéine et une agrégation des protéines dans le noyau. Une perte de fonction des protéines PABPN1 résulte en une augmentation de l'usage de sites de polyadénylation alternatifs proximaux au détriment des sites de polyadénylation alternatifs distaux (Jenal et al., 2012).

Du fait de la longue région 3'UTR de *DUX4*, une redirection du site de clivage peut modifier les sites de fixation des miARN. Une analyse *in silico* (<http://www.mirbase.org/search.shtml>) de l'ARNm *DUX4* délété de 40pb en 3' (généralisé lors de la redirection du site de clivage en présence du PMO-CS3), ne montre pas de modification des sites de fixation de potentiels miARN. Ainsi cette redirection du site de clivage ne modifie pas la régulation post-transcriptionnelle de *DUX4* par un mécanisme dépendant des miARN.

Par ailleurs, une redirection du site de clivage ou des changements de conformation de la région 3'UTR de *DUX4* lors de la formation du duplex PMO-ARNm peuvent moduler le NMD (Eberle et al., 2008). En effet, une diminution de la distance physique entre le codon normal de terminaison et la queue poly(A) inhibe la déstabilisation de l'ARNm par le NMD

(Eberle et al., 2008). Ainsi, une redirection du site de clivage en amont du signal canonique, ce qui diminue la distance physique entre le codon stop normal de *DUX4*, localisé dans l'exon 1, et la queue poly(A), peut stabiliser un peu plus l'ARNm *DUX4*. A l'inverse, cette redirection et la formation d'un duplex stable PMO-ARNm peut abolir la structure secondaire de l'ARNm, éloignant la queue poly(A) du codon stop normal. La perte de cette structure secondaire conduit alors à la déstabilisation de l'ARNm par le NMD. Il serait intéressant d'étudier la durée de demi-vie des différents variants ARNm *DUX4* pour établir une corrélation entre la taille de la région 3'UTR et la stabilité de l'ARNm.

En conclusion, lors du développement de nouveaux oligonucléotides antisens ciblant les éléments clés en 3'UTR du pré-ARNm *DUX4*, il est important de vérifier qu'aucune redirection du site de clivage ne s'opère, conduisant potentiellement à la production d'un ARNm plus stable ou plus efficacement traduit, ce qui serait délétère pour la cellule.

### **3.2.3. Utilisation d'oligonucléotides antisens induisant une redirection du site de clivage comme stratégie thérapeutique**

Ce phénomène de polyadénylation alternatif est impliqué dans le développement de plusieurs pathologies humaines (Danckwardt et al., 2008; Elkon et al., 2013; Jenal et al., 2012; Uitte de Willige et al., 2007). Ces phénomènes délétères peuvent être induits pas des mutations dans les séquences régulatrices canoniques (béta-thalassémie, syndrome IPEX, maladie de Fabry, thromboses) mais également par des mutations dans les facteurs protéiques (OPMD, cancers) impliqués dans la maturation en 3' des ARNm.

Ainsi, dans le cadre du développement de stratégies thérapeutiques pour ces pathologies caractérisées par la production d'ARNm utilisant un signal de polyadénylation alternatif au lieu du signal de polyadénylation canonique, cette capacité de redirection du site de clivage en présence d'un AO peut être exploitée afin de forcer les différents facteurs protéiques à utiliser le signal de polyadénylation canonique.



### **3.3. Inhibition des différents variants ARNm de *DUX4* : lesquels et quand ?**

#### **3.3.1. Quelles conséquences de cette stratégie thérapeutique sur l'expression des différentes isoformes de *DUX4* ?**

Trois isoformes de *DUX4* sont produites dans les cellules somatiques à partir du chromosome 4qA : deux variants longs *DUX4-fl* (1 et 2) composés des exons 1-2-3 et qui diffèrent par l'épissage ou non de l'intron 1, et un variant court *DUX4-s*, qui se caractérise par l'utilisation d'un site d'épissage cryptique situé dans l'exon 1, donnant un ARNm exon 1s-2-3 (Figure 57). Les deux variants *DUX4-fl* codent pour une protéine *DUX4-fl* fonctionnelle. Cependant, il n'est pas exclu que les isoformes *DUX4-fl* codent pour des protéines différentes. En effet, trois codons initiateurs MKG, MAL et MQG, pourraient permettre la traduction de trois protéines *DUX4* (Snider et al., 2009). La plupart des ARNm *DUX4* commencent en amont du codon initiateur MAL. Néanmoins, l'isolation de la ou des protéines *DUX4* représente un défi technique majeur du fait des faibles expressions protéiques dans les conditions physiologiques. Pour le moment, aucune étude n'a encore pu isoler de potentielles isoformes protéiques *DUX4* dans des cellules FSHD, et encore moins déterminer leur rôle. A l'inverse *DUX4-s* est traduit en une protéine *DUX4-s* présentant le domaine de fixation à l'ADN mais pas le domaine de transactivation nécessaire à l'activité du facteur de transcription (Bosnakovski et al., 2008; Dixit et al., 2007; Kawamura-Saito et al., 2006; Snider et al., 2009, 2010). L'expression de *DUX4-s* n'est pas toxique pour la cellule au contraire de l'expression de *DUX4-fl* (Geng et al., 2011).

Ces trois variants ARNm de *DUX4* possèdent l'exon 3. Ainsi, l'utilisation d'oligonucléotides antisens ciblant les séquences impliquées dans le mécanisme de polyadénylation de l'ARNm *DUX4*, toutes localisées dans l'exon 3 ou en aval de celui-ci, pourrait réduire l'expression de ces trois isoformes de *DUX4*. A ce jour, il n'existe aucune information quant aux rôles des différentes isoformes protéiques *DUX4* ou des variants ARNm *DUX4*. Il est donc difficile d'établir si une stratégie thérapeutique inhibant l'expression de *DUX4* doit nécessairement réduire l'expression de toutes ces isoformes.

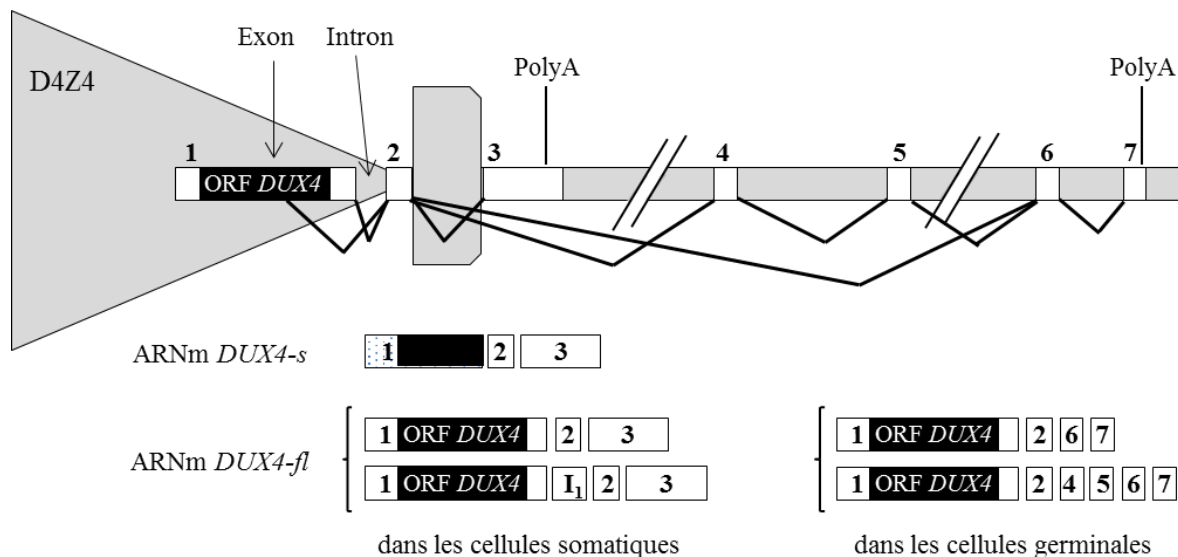


FIGURE 57. Récapitulatif des différents variants de l'ARNm *DUX4* transcrits dans les cellules somatiques ou les cellules germinales. Dans les cellules somatiques, trois ARNm *DUX4* peuvent être générés à partir de la dernière unité D4Z4. Tous les ARNm contiennent l'exon 2 et l'exon 3. L'ARNm *DUX4-s* résulte d'un épissage cryptique démarrant dans l'exon 1. Il est exprimé dans les myoblastes contrôles et dans les tissus somatiques. Les deux ARNm *DUX4-fl* contiennent l'exon 1 en entier. Ces deux ARNm *DUX4-fl* diffèrent par la présence ou non de l'intron 1. Cependant, ces deux variants sont tous deux associés à la FSHD et sont exprimés dans les myotubes d'haplotype 4qA. Dans les cellules germinales, deux autres ARNm *DUX4* sont transcrits. Ces ARNm sont composés de l'exon 1-2-6-7 ou 1-2-4-5-6-7. L'exon 3 est toujours exclu. Le signal de polyadénylation utilisé ici est localisé dans l'exon 7.

La séquence codante du rétrotransposon *DUX4* est conservée depuis plus de 100 millions d'années chez les *Catarrhini* (primates de l'ancien monde) et les *Afrotheria* (Clapp et al., 2007). D'un point de vue évolutif, une pression de sélection empêche probablement la divergence de la séquence codante *DUX4* dans ces clades en raison d'une potentielle fonction physiologique de *DUX4*. L'expression de *DUX4* dans des cellules musculaires est toxique. Cette toxicité pourrait être médiée par l'activation, sur une longue période de temps, par *DUX4*, de l'expression de centaines de gènes impliqués dans diverses fonctions biologiques (Geng et al., 2012; Snider et al., 2010). Par ailleurs, *DUX4* s'exprime dans les cellules de la lignée germinale testiculaire et dans des iPS (Snider et al., 2010). Le profil d'expression de *DUX4* dans la lignée germinale est similaire à celui observé pour d'autres protéines à deux homéodomaines dont la fonction n'est pas encore connue (Booth and Holland, 2007; Wu et al., 2014). L'ensemble de ces observations suggère que *DUX4* pourrait avoir un rôle au cours du développement précoce mais aussi dans les fonctions biologiques de la reproduction. Chez l'adulte, *DUX4* est ensuite réprimé uniquement dans les cellules somatiques dont celles du muscle squelettique. L'absence d'une répression de *DUX4* chez les patients FSHD conduit à l'activation transcriptionnelle entre autres de gènes de la lignée germinale et du

développement précoce, et dont l'expression pourrait être toxique pour la cellule somatique (Geng et al., 2012; Snider et al., 2010). Bien que ceci reste une hypothèse à confirmer, si *DUX4*, ou certains variants *DUX4*, sont nécessaires au développement précoce ou aux fonctions biologiques de reproduction, le traitement thérapeutique ne doit pas inhiber l'expression de *DUX4* au cours du développement ou dans la lignée germinale.

Or, dans les cellules germinales, deux autres isoformes faisant intervenir les exons 4-5-6-7 sont exprimées à partir des chromosomes 4 et 10 (Figure 57). Une isoforme correspond à l'association des exons 1-2-6-7 tandis que la seconde est composée des exons 1-2-4-5-6-7 (Geng et al., 2012; Snider et al., 2010). L'exon 3 est toujours exclu. Le signal de polyadénylation utilisé ici est localisé dans l'exon 7. Notre stratégie thérapeutique, ciblant exclusivement des éléments de l'exon 3, ne permet donc pas d'induire une diminution de l'expression de ces deux isoformes spécifiques aux cellules germinales et donc de potentiellement interférer avec les fonctions biologiques de la reproduction.

### **3.3.2. Quelle expression spatio-temporelle pour les différents variants ARNm *DUX4* ?**

Une autre question essentielle concerne l'expression spatio-temporelle de ces variants ARNm *DUX4*. Différentes études ont montré que *DUX4-s* est préférentiellement retrouvé dans les biopsies ou cultures de cellules musculaires de patients contrôles (myoblastes). Néanmoins, on retrouve également *DUX4-s* dans 30 à 40% des biopsies ou des cultures de cellules musculaires de patients FSHD. Au contraire, *DUX4-fl* est principalement détecté dans les biopsies et cultures cellulaires de patients FSHD (Jones et al., 2012; Snider et al., 2010; Tsumagari et al., 2011). Par ailleurs, des études menées sur des cellules iPS d'individus contrôles laissent à penser que *DUX4-fl* serait exprimé dès le développement fœtal, dans les tissus musculaires et non musculaires de patients FSHD et d'individus contrôles (Ferreboeuf et al., 2014a; Snider et al., 2010). Or, c'est l'isoforme *DUX4-s* qui prédomine par la suite dans les cellules musculaires et non musculaires des individus contrôles. Le mécanisme permettant ce passage d'expression de *DUX4-fl* vers *DUX4-s* et inversement n'est pas établi. Néanmoins, il est possible qu'il implique une modification de l'épissage alternatif avec une utilisation du site cryptique d'épissage localisé dans l'exon 1. Ce changement d'épissage pourrait dépendre du contexte épigénétique au locus 4q35. En effet, de nombreuses études ont mis en évidence

le rôle de la structure de la chromatine, des marques des histones ou encore de la méthylation de l'ADN dans le choix des introns ou exons à épisser (Brown et al., 2012; Luco et al., 2010, 2011; Zhou et al., 2014). Or la différence entre des iPS, des cellules musculaires FSHD et des cellules musculaires contrôles concerne l'état de méthylation de l'histone H3K9 et de l'ADN (Cabianca et al., 2012; Gaillard et al., 2014; de Greef et al., 2007, 2009, 2010; Hartweck et al., 2013; Jones et al., 2014, 2015, van Overveld et al., 2003, 2005; Sacconi et al., 2012; Zeng et al., 2009). Ainsi, la perte des marques épigénétiques répressives pourrait permettre un épissage produisant les isoformes *DUX4-fl*. A l'inverse, en présence de marques épigénétiques répressives, c'est majoritairement l'isoforme *DUX4-s* qui serait générée.

En ce qui concerne les deux isoformes *DUX4-fl1* (sans intron 1 et 2) et *DUX4-fl2* (sans intron 2), le mécanisme qui conduit à la production de l'une ou l'autre des isoformes semble être plus compliqué qu'une régulation épigénétique. Par ailleurs, bien que ces deux isoformes s'expriment simultanément dans les biopsies et cultures cellulaires de patients FSHD, aucune étude n'a encore permis de déterminer si ces deux isoformes étaient exprimées à partir du même noyau au sein d'une fibre musculaire. Ces deux isoformes, ainsi que celles associées à la lignée germinale, possèdent l'ORF *DUX4* complet. Il est ainsi fort probable que ces isoformes conduisent toutes à la production d'une protéine *DUX4* fonctionnelle. La différence de longueur de la région 3'UTR entre *DUX4-fl1* et *DUX4-fl2* (ainsi qu'entre les isoformes de la lignée germinale) pourrait être à l'origine d'une différence de stabilité ou d'efficacité de traduction entre ces isoformes (Di Giammartino et al., 2011).

### **3.3.3. A partir de quand inhiber l'expression de *DUX4* dans les cellules somatiques ?**

La stratégie thérapeutique développée au cours de ma thèse permet d'inhiber l'expression de *DUX4* et donc probablement de stopper le développement de la FSHD. Ainsi dans le cas d'un traitement de la FSHD, le plus judicieux est d'inhiber rapidement l'expression de *DUX4* dans les cellules somatiques avant que les effets liés à une toxicité de *DUX4*, comme l'atrophie musculaire, ne soient trop importants. L'évolution de la pathologie ainsi que sa sévérité varie d'un individu à un autre. Aucune indication n'existe sur le niveau ou la durée d'exposition à une expression de *DUX4* nécessaire à l'apparition des premiers symptômes.

Malgré tout, il a été montré que l'expression de *DUX4* précède de plusieurs mois ou années les premiers symptômes cliniques (Broucqsault et al., 2013; Ferreboeuf et al., 2014a). Dans les cas familiaux, un diagnostic génétique et une recherche d'expression de *DUX4* ou de marqueurs de la FSHD pourraient être envisagés parmi la descendance d'un patient FSHD afin de traiter la pathologie avant que le phénotype FSHD ne se développe. Mais dans 10 à 30% des cas, les patients FSHD sont des cas *de novo* sans antécédents familiaux. Pour ces patients FSHD, le diagnostic clinique et/ou génétique est généralement long à être établi et s'effectue souvent lorsque les symptômes associés à la pathologie sont déjà bien exprimés. Une amélioration du diagnostic de la FSHD ainsi que de la prise en charge des patients doivent être envisagées pour pouvoir fournir un traitement au patient avant que son intégrité physique ne soit trop affectée.

### **3.4. Combinaison de plusieurs stratégies de thérapie génique pour la FSHD**

#### **3.4.1. Faut-il combiner plusieurs stratégies thérapeutiques ciblant *DUX4* pour le traitement de la FSHD ?**

Au cours de ma thèse, nous nous sommes focalisés sur le développement d'une stratégie thérapeutique ciblant uniquement les éléments clés en 3' de l'ARNm. Chez les individus contrôles, l'ARNm *DUX4* n'est pas détecté. L'objectif des stratégies thérapeutiques pour la FSHD est donc d'inhiber totalement ou suffisamment l'expression de *DUX4* pour éliminer ces effets toxiques. En effet, il n'y a aucune indication sur l'existence potentielle d'un seuil acceptable d'expression de *DUX4*, mais on peut imaginer qu'en dessous d'un certain niveau d'expression, *DUX4* n'est pas toxique pour la cellule.

*In vitro*, nos oligonucléotides antisens permettent au mieux une diminution moyenne de l'ordre de 50% de l'ARNm *DUX4*. De ce fait, des protéines DUX4, en quantité non négligeable, sont encore produites dans les cellules traitées à partir de l'ARNm *DUX4* résiduel et induisent probablement une dérégulation de l'expression de nombreux gènes. Bien entendu, les effets toxiques de DUX4 sont quantitativement moins importants dans les cellules traitées mais ne sont probablement pas totalement supprimés. *In vitro*, une

augmentation de la quantité de l'oligonucléotide antisens pourrait, en théorie, permettre d'augmenter l'inhibition de *DUX4*. Cependant, nous avons observé, dans nos conditions expérimentales, au-delà de 50nM de PMO-PAS ou PMO-CS3 lipotransfectés, que l'efficacité de l'oligonucléotide antisens ne pouvait être améliorée (atteinte d'un plateau). Des modifications au niveau de la séquence de l'oligonucléotide antisens permettront probablement de diminuer la quantité d'ARNm *DUX4* résiduel. Malgré tout, la probabilité est faible de restreindre la quantité d'ARNm *DUX4* résiduel à une valeur nulle ou non détectable, identique à ce qui est retrouvé chez des individus contrôles, par l'utilisation uniquement de cette stratégie thérapeutique. De plus, *DUX4* faisant partie d'une boucle de rétro-contrôle positif par inhibition du NMD, il est facile d'imaginer que la persistance de l'expression de protéines *DUX4*, même en faible quantité, favorise la production d'autres protéines (Feng et al., 2015).

Dans le cadre de l'élaboration d'un traitement thérapeutique pour la FSHD, il est possible d'envisager inhiber totalement l'expression de *DUX4* dans les tissus somatiques en combinant plusieurs stratégies thérapeutiques. Par exemple, la stratégie thérapeutique développée par Himeda et al., consiste à diminuer la transcription de *DUX4* en utilisant un système dCas9-KRAB. L'efficacité de répression transcriptionnelle du système dCas9-KRAB n'est pas totale. Des ARNm *DUX4* sont transcrits. Cette stratégie pourrait être associée à notre stratégie thérapeutique ciblant les éléments clés en 3' du pré-ARNm *DUX4* (Himeda et al., 2016a). La combinaison de notre stratégie à celle de Himeda et al. permettrait alors de déstabiliser les quelques pré-ARNm *DUX4* transcrits en inhibant leur polyadénylation. Il en résulterait une réduction un peu plus importante de l'expression de *DUX4* en comparaison à l'utilisation seule de l'une ou l'autre de ces stratégies. D'autres stratégies thérapeutiques peuvent être combinées comme l'utilisation de miARN/siARN ou la surexpression de *DUX4-s*, l'objectif *in fine* étant toujours la même : inhiber totalement ou presque l'expression de *DUX4* dans les cellules somatiques, et notamment dans les fibres musculaires (Lim et al., 2015; Mitsuhashi et al., 2013a; Wallace et al., 2012).

### **3.4.2. Faut-il développer des stratégies thérapeutiques pour la FSHD ne ciblant que *DUX4* ?**

Bien que plusieurs loci distincts soient en cause chez les patients FSHD, les patients FSHD sont tous caractérisés par une expression de *DUX4*. De ce fait, la majorité des stratégies thérapeutiques en développement ont pour objectif d'éteindre l'expression de ce gène. Or l'expression de *DUX4* ne permet pas d'expliquer certaines caractéristiques de la FSHD, comme l'atteinte asymétrique (Jones et al., 2012). La FSHD n'est donc pas une pathologie monogénique mais multigénique. Quelques gènes candidats ont été proposés : *FRG1*, *FRG2*, *ANTI* et *FAT1*. Cependant, l'implication de la majorité de ces gènes dans la FSHD est sujette à controverse, à l'exception de *FAT1*. En effet, des souris homozygotes hypomorphiques *Fat1* récapitulent le phénotype musculaire et non musculaire de la FSHD, y compris l'asymétrie et la rétinopathie (Caruso et al., 2013). Par ailleurs, *FAT1* est plus faiblement exprimé chez des patients FSHD comparés à des individus contrôles (Caruso et al., 2013; Mariot et al., 2015). Ainsi, pour développer la FSHD, les individus doivent exprimer *DUX4* et/ou potentiellement présenter un niveau d'expression de *FAT1* en dessous de la moyenne établie dans la population (Mariot et al., 2015; Puppo et al., 2015). Des études complémentaires sont nécessaires pour confirmer cette hypothèse. Malgré tout, le développement d'une stratégie thérapeutique complétant l'expression de *FAT1* pourrait être envisagé pour la FSHD.

## **4. CONCLUSION GENERALE**



Mon projet de thèse consistait en l'élaboration d'une nouvelle stratégie thérapeutique en développant des oligonucléotides antisens ciblant les éléments clés en 3' du pré-ARNm *DUX4* afin d'inhiber sa polyadénylation, ce qui a pour conséquence sa déstabilisation et l'inhibition de la synthèse protéique de *DUX4*. En effet, le clivage et la polyadénylation des ARNm en 3', qui se déroulent dans le noyau, sont des mécanismes essentiels à la maturation de la majorité des ARNm eucaryotes. Cette maturation est généralement nécessaire à l'export nucléaire de l'ARNm, joue un rôle dans la stabilité de l'ARNm et dans l'efficacité de sa traduction. L'hypothèse de travail était que cibler les éléments clés impliqués dans la maturation de l'ARNm avec des oligonucléotides antisens empêcherait l'ajout d'une queue poly(A) et donc permettrait la suppression génique de *DUX4*. Les éléments clés de cette région 3'UTR sont une séquence de six nucléotides correspondant au signal de polyadénylation, un site de clivage positionné 20 à 30 nucléotides en aval du signal de polyadénylation et une séquence riche en G/U ou U (*Downstream sequence element* (DSE)) localisée à 30 nucléotides après le site de clivage.

Au cours de cette étude, nous avons pu montrer *in vitro*, et ce pour la première fois, que l'utilisation d'oligonucléotides antisens, ici des PMOs (*phosphorodiamidate morpholino oligomer*), ciblant les séquences clés impliquées dans l'ajout d'une queue poly(A), notamment les séquences contenant le signal de polyadénylation et le site de clivage, conduit à une sous-expression du gène ciblé et en particulier de *DUX4*. Cette stratégie thérapeutique pourrait être utilisée dans le cadre de la FSHD mais également dans le cas de maladies génétiques à gain de fonction. Nous avons par ailleurs commencé à tester l'efficacité *in vivo* de cette stratégie. Les premiers résultats obtenus dans un modèle de souris transgénique exprimant *DUX4* traité avec les séquences des AOs les plus prometteurs insérées dans un vecteur AAV sous promoteur U7 (AAV-U7-PAS) sont encourageants.

En conclusion, notre étude, montre la faisabilité pour la FSHD d'une approche thérapeutique ciblant les éléments clés en 3' d'un pré-ARNm qui est ici *DUX4*.

# **BIBLIOGRAPHIE**

- Afione, S.A., Conrad, C.K., Kearns, W.G., Chunduru, S., Adams, R., Reynolds, T.C., Guggino, W.B., Cutting, G.R., Carter, B.J., and Flotte, T.R. (1996). In vivo model of adeno-associated virus vector persistence and rescue. *J. Virol.* *70*, 3235–3241.
- Agrotis, A., and Ketteler, R. (2015). A new age in functional genomics using CRISPR/Cas9 in arrayed library screening. *Front. Genet.* *6*.
- Akache, B., Grimm, D., Pandey, K., Yant, S.R., Xu, H., and Kay, M.A. (2006). The 37/67-kilodalton laminin receptor is a receptor for adeno-associated virus serotypes 8, 2, 3, and 9. *J. Virol.* *80*, 9831–9836.
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). Chromosomal DNA and Its Packaging in the Chromatin Fiber.
- Allikian, M.J., Bhabha, G., Dospoy, P., Heydemann, A., Ryder, P., Earley, J.U., Wolf, M.J., Rockman, H.A., and McNally, E.M. (2007). Reduced life span with heart and muscle dysfunction in *Drosophila* sarcoglycan mutants. *Hum. Mol. Genet.* *16*, 2933–2943.
- Ames, D., Murphy, N., Helentjaris, T., Sun, N., and Chandler, V. (2008). Comparative analyses of human single- and multilocus tandem repeats. *Genetics* *179*, 1693–1704.
- Amitai, G., and Sorek, R. (2016). CRISPR-Cas adaptation: insights into the mechanism of action. *Nat. Rev. Microbiol.* *14*, 67–76.
- Andersen, G., Prahm, K.P., Dahlqvist, J.R., Citirak, G., and Vissing, J. (2015). Aerobic training and postexercise protein in facioscapulohumeral muscular dystrophy: RCT study. *Neurology* *85*, 396–403.
- André, F.M., Gehl, J., Sersa, G., Pr at, V., Hojman, P., Eriksen, J., Golzio, M., Cemazar, M., Pavselj, N., Rols, M.-P., et al. (2008). Efficiency of high- and low-voltage pulse combinations for gene electrotransfer in muscle, liver, tumor, and skin. *Hum. Gene Ther.* *19*, 1261–1271.
- Andreassi, C., and Riccio, A. (2009). To localize or not to localize: mRNA fate is in 3'UTR ends. *Trends Cell Biol.* *19*, 465–474.
- Andrulis, E.D., Neiman, A.M., Zappulla, D.C., and Sternglanz, R. (1998). Perinuclear localization of chromatin facilitates transcriptional silencing. *Nature* *394*, 592–595.
- Ansseau, E., Laoudj-Chenivesse, D., Marcowycz, A., Tassin, A., Vanderplanck, C., Sauvage, S., Barro, M., Mahieu, I., Leroy, A., Leclercq, I., et al. (2009). DUX4c is up-regulated in FSHD. It induces the MYF5 protein and human myoblast proliferation. *PloS One* *4*, e7482.
- Ansseau, E., Eidahl, J.O., Lancelot, C., Tassin, A., Matteotti, C., Yip, C., Liu, J., Leroy, B., Hubeau, C., Gerbaux, C., et al. (2016). Homologous Transcription Factors DUX4 and DUX4c Associate with Cyttoplasmic Proteins during Muscle Differentiation. *PLOS ONE* *11*, e0146893.
- Antes, T.J., Namciu, S.J., Fournier, R.E., and Levy-Wilson, B. (2001). The 5' boundary of the human apolipoprotein B chromatin domain in intestinal cells. *Biochemistry (Mosc.)* *40*, 6731–6742.
- Arahata, K., Ishihara, T., Fukunaga, H., Orimo, S., Lee, J.H., Goto, K., and Nonaka, I. (1995). Inflammatory response in facioscapulohumeral muscular dystrophy (FSHD): immunocytochemical and genetic analyses. *Muscle Nerve. Suppl.* *2*, S56-66.
- Aran (1850). Recherches sur une maladie non encore d crit du syst me musculaire (atrophie musculaire progressive). *Arch. G n rales M decine* *3*: 172.

- Arashiro, P., Eisenberg, I., Kho, A.T., Cerqueira, A.M.P., Canovas, M., Silva, H.C.A., Pavanello, R.C.M., Verjovski-Almeida, S., Kunkel, L.M., and Zatz, M. (2009). Transcriptional regulation differs in affected facioscapulohumeral muscular dystrophy patients compared to asymptomatic related carriers. *Proc. Natl. Acad. Sci. U. S. A.* *106*, 6220–6225.
- Argyros, O., Wong, S.P., Niceta, M., Waddington, S.N., Howe, S.J., Coutelle, C., Miller, A.D., and Harbottle, R.P. (2008). Persistent episomal transgene expression in liver following delivery of a scaffold/matrix attachment region containing non-viral vector. *Gene Ther.* *15*, 1593–1605.
- Arnould, S., Chames, P., Perez, C., Lacroix, E., Duclert, A., Epinat, J.-C., Stricher, F., Petit, A.-S., Patin, A., Guillier, S., et al. (2006). Engineering of large numbers of highly specific homing endonucleases that induce recombination on novel DNA targets. *J. Mol. Biol.* *355*, 443–458.
- Ashworth, J., Havranek, J.J., Duarte, C.M., Sussman, D., Monnat, R.J., Stoddard, B.L., and Baker, D. (2006). Computational redesign of endonuclease DNA binding and cleavage specificity. *Nature* *441*, 656–659.
- Attarian, S., Salort-Campana, E., Nguyen, K., Behin, A., and Andoni Urtizberea, J. (2012). Recommendations for the management of facioscapulohumeral muscular dystrophy in 2011. *Rev. Neurol. (Paris)* *168*, 910–918.
- Awerbuch, G.I., Nigro, M.A., and Wishnow, R. (1990). Beevor’s sign and facioscapulohumeral dystrophy. *Arch. Neurol.* *47*, 1208–1209.
- Ayuso, E., Mingozzi, F., and Bosch, F. (2010). Production, purification and characterization of adeno-associated vectors. *Curr. Gene Ther.* *10*, 423–436.
- Bakker, E., Wijmenga, C., Vossen, R.H., Padberg, G.W., Hewitt, J., van der Wielen, M., Rasmussen, K., and Frants, R.R. (1995). The FSHD-linked locus D4F104S1 (p13E-11) on 4q35 has a homologue on 10qter. *Muscle Nerve. Suppl.* *2*, S39-44.
- Baldwin, B.J., Talley, R.C., Johnson, C., and Nutter, D.O. (1973). Permanent paralysis of the atrium in a patient with facioscapulohumeral muscular dystrophy. *Am. J. Cardiol.* *31*, 649–653.
- Bankolé, L.-C., Millet, G.Y., Temesi, J., Bachasson, D., Ravelojaona, M., Wuyam, B., Verges, S., Ponsot, E., Antoine, J.-C., Kadi, F., et al. (2016). Safety and efficacy of a 6-month home-based exercise program in patients with facioscapulohumeral muscular dystrophy: A randomized controlled trial. *Medicine (Baltimore)* *95*, e4497.
- Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* *410*, 120–124.
- Barabino, S.M., and Keller, W. (1999). Last but not least: regulated poly(A) tail formation. *Cell* *99*, 9–11.
- Bareja, A., Holt, J.A., Luo, G., Chang, C., Lin, J., Hinken, A.C., Freudenberg, J.M., Kraus, W.E., Evans, W.J., and Billin, A.N. (2014). Human and mouse skeletal muscle stem cells: convergent and divergent mechanisms of myogenesis. *PLoS One* *9*, e90398.
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* *315*, 1709–1712.

- Barro, M., Carnac, G., Flavier, S., Mercier, J., Vassetzky, Y., and Laoudj-Chenivresse, D. (2010). Myoblasts from affected and non-affected FSHD muscles exhibit morphological differentiation defects. *J. Cell. Mol. Med.* *14*, 275–289.
- Bartlett, J.S., Wilcher, R., and Samulski, R.J. (2000). Infectious entry pathway of adeno-associated virus and adeno-associated virus vectors. *J. Virol.* *74*, 2777–2785.
- Bassett, A.R., Tibbit, C., Ponting, C.P., and Liu, J.-L. (2013). Highly efficient targeted mutagenesis of *Drosophila* with the CRISPR/Cas9 system. *Cell Rep.* *4*, 220–228.
- Baum, C., Kustikova, O., Modlich, U., Li, Z., and Fehse, B. (2006). Mutagenesis and oncogenesis by chromosomal insertion of gene transfer vectors. *Hum. Gene Ther.* *17*, 253–263.
- Baur, J.A., Zou, Y., Shay, J.W., and Wright, W.E. (2001). Telomere position effect in human cells. *Science* *292*, 2075–2077.
- Baxter, J., Merckenschlager, M., and Fisher, A.G. (2002). Nuclear organisation and gene expression. *Curr. Opin. Cell Biol.* *14*, 372–376.
- Beaudoing, E., Freier, S., Wyatt, J.R., Claverie, J.M., and Gautheret, D. (2000). Patterns of variant polyadenylation signal usage in human genes. *Genome Res.* *10*, 1001–1010.
- Beckmann, J.S., and Spencer, M. (2008). Calpain 3, the “gatekeeper” of proper sarcomere assembly, turnover and maintenance. *Neuromuscul. Disord. NMD* *18*, 913–921.
- Bedell, V.M., Wang, Y., Campbell, J.M., Poshusta, T.L., Starker, C.G., Krug Ii, R.G., Tan, W., Penheiter, S.G., Ma, A.C., Leung, A.Y.H., et al. (2012). In vivo genome editing using a high-efficiency TALEN system. *Nature* *491*, 114–118.
- Berli, R.R., Segal, D.J., Dreier, B., and Barbas, C.F. (1998). Toward controlling gene expression at will: specific regulation of the *erbB-2/HER-2* promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc. Natl. Acad. Sci. U. S. A.* *95*, 14628–14633.
- Berli, R.R., Dreier, B., and Barbas, C.F. (2000). Positive and negative regulation of endogenous genes by designed transcription factors. *Proc. Natl. Acad. Sci. U. S. A.* *97*, 1495–1500.
- Bell, A.C., and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature* *405*, 482–485.
- Bengtsson, N.E., Hall, J.K., Odom, G.L., Phelps, M.P., Andrus, C.R., Hawkins, R.D., Hauschka, S.D., Chamberlain, J.R., and Chamberlain, J.S. (2017). Muscle-specific CRISPR/Cas9 dystrophin gene editing ameliorates pathophysiology in a mouse model for Duchenne muscular dystrophy. *Nat. Commun.* *8*, 14454.
- Bengtsson, U., Altherr, M.R., Wasmuth, J.J., and Winokur, S.T. (1994). High resolution fluorescence in situ hybridization to linearly extended DNA visually maps a tandem repeat associated with facioscapulohumeral muscular dystrophy immediately adjacent to the telomere of 4q. *Hum. Mol. Genet.* *3*, 1801–1805.
- Bennett, F.C., and Harvey, K.F. (2006). Fat cadherin modulates organ size in *Drosophila* via the Salvador/Warts/Hippo signaling pathway. *Curr. Biol. CB* *16*, 2101–2110.
- Bennett, C.L., Brunkow, M.E., Ramsdell, F., O’Briant, K.C., Zhu, Q., Fuleihan, R.L., Shigeoka, A.O., Ochs, H.D., and Chance, P.F. (2001). A rare polyadenylation signal mutation of the *FOXP3* gene (AAUAAA-->AAUGAA) leads to the IPEX syndrome. *Immunogenetics* *53*, 435–439.

- Bensimon, A., Simon, A., Chiffaudel, A., Croquette, V., Heslot, F., and Bensimon, D. (1994). Alignment and sensitive detection of DNA by a moving interface. *Science* 265, 2096–2098.
- Bentzinger, C.F., Wang, Y.X., and Rudnicki, M.A. (2012). Building Muscle: Molecular Regulation of Myogenesis. *Cold Spring Harb. Perspect. Biol.* 4, a008342.
- Bergman, Y., and Cedar, H. (2013). DNA methylation dynamics in health and disease. *Nat. Struct. Mol. Biol.* 20, 274–281.
- Bernard, P., Maure, J.F., Partridge, J.F., Genier, S., Javerzat, J.P., and Allshire, R.C. (2001). Requirement of heterochromatin for cohesion at centromeres. *Science* 294, 2539–2542.
- Bernstein, B.E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D.K., Huebert, D.J., McMahon, S., Karlsson, E.K., Kulbokas, E.J., Gingeras, T.R., et al. (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* 120, 169–181.
- Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409, 363–366.
- Berto, E., Bozac, A., and Marconi, P. (2005). Development and application of replication-incompetent HSV-1-based vectors. *Gene Ther.* 12 Suppl 1, S98-102.
- Bessis, N., GarciaCozar, F.J., and Boissier, M.-C. (2004). Immune responses to gene therapy vectors: influence on vector function and effector mechanisms. *Gene Ther.* 11 Suppl 1, S10-17.
- Bessonov, S., Anokhina, M., Krasauskas, A., Golas, M.M., Sander, B., Will, C.L., Urlaub, H., Stark, H., and Lührmann, R. (2010). Characterization of purified human Bact spliceosomal complexes reveals compositional and morphological changes during spliceosome activation and first step catalysis. *RNA N. Y. N* 16, 2384–2403.
- Bestor, T.H. (2000). The DNA methyltransferases of mammals. *Hum. Mol. Genet.* 9, 2395–2402.
- Betts, C., Saleh, A.F., Arzumanov, A.A., Hammond, S.M., Godfrey, C., Coursindel, T., Gait, M.J., and Wood, M.J. (2012). Pip6-PMO, A New Generation of Peptide-oligonucleotide Conjugates With Improved Cardiac Exon Skipping Activity for DMD Treatment. *Mol. Ther. Nucleic Acids* 1, e38.
- Bibikova, M., Carroll, D., Segal, D.J., Trautman, J.K., Smith, J., Kim, Y.G., and Chandrasegaran, S. (2001). Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Mol. Cell. Biol.* 21, 289–297.
- Bickmore, W.A., Maarel, V.D., and M, S. (2003). Perturbations of chromatin structure in human genetic disease: recent advances. *Hum. Mol. Genet.* 12, R207–R213.
- Bish, L.T., Sleeper, M.M., Forbes, S.C., Morine, K.J., Reynolds, C., Singletary, G.E., Trafny, D., Pham, J., Bogan, J., Kornegay, J.N., et al. (2011). Long-term systemic myostatin inhibition via liver-targeted gene transfer in golden retriever muscular dystrophy. *Hum. Gene Ther.* 22, 1499–1509.
- Blaese, R.M., Culver, K.W., Miller, A.D., Carter, C.S., Fleisher, T., Clerici, M., Shearer, G., Chang, L., Chiang, Y., Tolstoshev, P., et al. (1995). T lymphocyte-directed gene therapy for ADA- SCID: initial trial results after 4 years. *Science* 270, 475–480.
- Blewitt, M.E., Gendrel, A.-V., Pang, Z., Sparrow, D.B., Whitelaw, N., Craig, J.M., Apedaile, A., Hilton, D.J., Dunwoodie, S.L., Brockdorff, N., et al. (2008). SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation. *Nat. Genet.* 40, 663–669.

- Block, G.J., Petek, L.M., Narayanan, D., Amell, A.M., Moore, J.M., Rabaia, N.A., Tyler, A., van der Maarel, S.M., Tawil, R., Filippova, G.N., et al. (2012). Asymmetric bidirectional transcription from the FSHD-causing D4Z4 array modulates DUX4 production. *PLoS One* 7, e35532.
- Block, G.J., Narayanan, D., Amell, A.M., Petek, L.M., Davidson, K.C., Bird, T.D., Tawil, R., Moon, R.T., and Miller, D.G. (2013). Wnt/ $\beta$ -catenin signaling suppresses DUX4 expression and prevents apoptosis of FSHD muscle cells. *Hum. Mol. Genet.* 22, 4661–4672.
- Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., and Bonas, U. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326, 1509–1512.
- Bockstael, O., Foust, K.D., Kaspar, B., and Tenenbaum, L. (2011). Recombinant AAV delivery to the central nervous system. *Methods Mol. Biol. Clifton NJ* 807, 159–177.
- Bode, J., Benham, C., Knopp, A., and Mielke, C. (2000). Transcriptional augmentation: modulation of gene expression by scaffold/matrix-attached regions (S/MAR elements). *Crit. Rev. Eukaryot. Gene Expr.* 10, 73–90.
- Bodega, B., Ramirez, G.D.C., Grasser, F., Cheli, S., Brunelli, S., Mora, M., Meneveri, R., Marozzi, A., Mueller, S., Battaglioli, E., et al. (2009). Remodeling of the chromatin structure of the facioscapulohumeral muscular dystrophy (FSHD) locus and upregulation of FSHD-related gene 1 (FRG1) expression during human myogenic differentiation. *BMC Biol.* 7, 41.
- Bodine, S.C., and Baehr, L.M. (2014). Skeletal muscle atrophy and the E3 ubiquitin ligases MuRF1 and MAFbx/atrogen-1. *Am. J. Physiol. Endocrinol. Metab.* 307, E469–484.
- Bodine, S.C., Latres, E., Baumhueter, S., Lai, V.K., Nunez, L., Clarke, B.A., Poueymirou, W.T., Panaro, F.J., Na, E., Dharmarajan, K., et al. (2001). Identification of ubiquitin ligases required for skeletal muscle atrophy. *Science* 294, 1704–1708.
- Bogdanove, A.J., and Voytas, D.F. (2011). TAL effectors: customizable proteins for DNA targeting. *Science* 333, 1843–1846.
- Bonner, M., and Kmiec, E.B. (2009). DNA breakage associated with targeted gene alteration directed by DNA oligonucleotides. *Mutat. Res.* 669, 85–94.
- van den Boogaard, M.L., Jfl Lemmers, R., Camaño, P., van der Vliet, P.J., Voermans, N., van Engelen, B.G., Lopez de Munain, A., Tapscott, S.J., van der Stoep, N., Tawil, R., et al. (2015). Double SMCHD1 variants in FSHD2: the synergistic effect of two SMCHD1 variants on D4Z4 hypomethylation and disease penetrance in FSHD2. *Eur. J. Hum. Genet. EJHG.*
- van den Boogaard, M.L., Lemmers, R.J.L.F., Balog, J., Wohlgemuth, M., Auranen, M., Mitsuhashi, S., van der Vliet, P.J., Straasheijm, K.R., van den Akker, R.F.P., Kriek, M., et al. (2016). Mutations in DNMT3B Modify Epigenetic Repression of the D4Z4 Repeat and the Penetrance of Facioscapulohumeral Dystrophy. *Am. J. Hum. Genet.* 98, 1020–1029.
- Booth, H.A.F., and Holland, P.W.H. (2007). Annotation, nomenclature and evolution of four novel homeobox genes expressed in the human germ line. *Gene* 387, 7–14.
- Bordignon, C., Notarangelo, L.D., Nobili, N., Ferrari, G., Casorati, G., Panina, P., Mazzolari, E., Maggioni, D., Rossi, C., Servida, P., et al. (1995). Gene therapy in peripheral blood lymphocytes and bone marrow for ADA- immunodeficient patients. *Science* 270, 470–475.
- Bortolanza, S., Nonis, A., Sanvito, F., Maciotta, S., Sitia, G., Wei, J., Torrente, Y., Di Serio, C., Chamberlain, J.R., and Gabellini, D. (2011). AAV6-mediated systemic shRNA delivery reverses

disease in a mouse model of facioscapulohumeral muscular dystrophy. *Mol. Ther. J. Am. Soc. Gene Ther.* *19*, 2055–2064.

Bosnakovski, D., Xu, Z., Gang, E.J., Galindo, C.L., Liu, M., Simsek, T., Garner, H.R., Agha-Mohammadi, S., Tassin, A., Coppée, F., et al. (2008). An isogenetic myoblast expression screen identifies DUX4-mediated FSHD-associated molecular pathologies. *EMBO J.* *27*, 2766–2779.

Bosnakovski, D., Choi, S.H., Strasser, J.M., Toso, E.A., Walters, M.A., and Kyba, M. (2014). High-throughput screening identifies inhibitors of DUX4-induced myoblast toxicity. *Skelet. Muscle* *4*, 4.

Bouard, D., Alazard-Dany, N., and Cosset, F.-L. (2009). Viral vectors: from virology to transgene expression. *Br. J. Pharmacol.* *157*, 153–165.

Brand, M.D., Pakay, J.L., Ocloo, A., Kokoszka, J., Wallace, D.C., Brookes, P.S., and Cornwall, E.J. (2005). The basal proton conductance of mitochondria depends on adenine nucleotide translocase content. *Biochem. J.* *392*, 353–362.

Braun, T., and Gautel, M. (2011). Transcriptional mechanisms regulating skeletal muscle differentiation, growth and homeostasis. *Nat. Rev. Mol. Cell Biol.* *12*, 349–361.

Brideau, N.J., Coker, H., Gendrel, A.-V., Siebert, C.A., Bezstarosti, K., Demmers, J., Poot, R.A., Nesterova, T.B., and Brockdorff, N. (2015). Independent Mechanisms Target SMCHD1 to Trimethylated Histone H3 Lysine 9-Modified Chromatin and the Inactive X Chromosome. *Mol. Cell Biol.* *35*, 4053–4068.

Brockstedt, D.G., Podsakoff, G.M., Fong, L., Kurtzman, G., Mueller-Ruchholtz, W., and Engleman, E.G. (1999). Induction of immunity to antigens expressed by recombinant adeno-associated virus depends on the route of administration. *Clin. Immunol. Orlando Fla* *92*, 67–75.

Broucqsault, N., Morere, J., Gaillard, M.-C., Dumonceaux, J., Torrents, J., Salort-Campana, E., Maues De Paula, A., Bartoli, M., Fernandez, C., Chesnais, A.L., et al. (2013). Dysregulation of 4q35- and muscle-specific genes in fetuses with a short D4Z4 array linked to facio-scapulo-humeral dystrophy. *Hum. Mol. Genet.* *22*, 4206–4214.

Brouwer, J.R., Willemsen, R., and Oostra, B.A. (2009). Microsatellite repeat instability and neurological disease. *BioEssays News Rev. Mol. Cell. Dev. Biol.* *31*, 71–83.

Brouwer, O.F., Padberg, G.W., Ruys, C.J., Brand, R., de Laat, J.A., and Grote, J.J. (1991). Hearing loss in facioscapulohumeral muscular dystrophy. *Neurology* *41*, 1878–1881.

Brower, J.V., Rodic, N., Seki, T., Jorgensen, M., Fliess, N., Yachnis, A.T., McCarrey, J.R., Oh, S.P., and Terada, N. (2007). Evolutionarily conserved mammalian adenine nucleotide translocase 4 is essential for spermatogenesis. *J. Biol. Chem.* *282*, 29658–29666.

Brown, S.J., Stoilov, P., and Xing, Y. (2012). Chromatin and epigenetic regulation of pre-mRNA processing. *Hum. Mol. Genet.* *21*, R90-96.

Brun, C., Suter, D., Pauli, C., Dunant, P., Lochmüller, H., Burgunder, J.M., Schümperli, D., and Weis, J. (2003). U7 snRNAs induce correction of mutated dystrophin pre-mRNA by exon skipping. *Cell. Mol. Life Sci. CMLS* *60*, 557–566.

Buckingham, M. (2006). Myogenic progenitor cells and skeletal myogenesis in vertebrates. *Curr. Opin. Genet. Dev.* *16*, 525–532.



- Buj-Bello, A., Furling, D., Tronchère, H., Laporte, J., Lerouge, T., Butler-Browne, G.S., and Mandel, J.-L. (2002). Muscle-specific alternative splicing of myotubularin-related 1 gene is impaired in DM1 muscle cells. *Hum. Mol. Genet.* *11*, 2297–2307.
- Büller, H.R., Gailani, D., and Weitz, J.I. (2015). Factor XI antisense oligonucleotide for venous thrombosis. *N. Engl. J. Med.* *372*, 1672.
- Buller, R.M., Janik, J.E., Sebring, E.D., and Rose, J.A. (1981). Herpes simplex virus types 1 and 2 completely help adenovirus-associated virus replication. *J. Virol.* *40*, 241–247.
- Büning, H., Perabo, L., Coutelle, O., Quadt-Humme, S., and Hallek, M. (2008). Recent developments in adeno-associated virus vector technology. *J. Gene Med.* *10*, 717–733.
- Buongiorno-Nardelli, M., Micheli, G., Carri, M.T., and Marilley, M. (1982). A relationship between replicon size and supercoiled loop domains in the eukaryotic genome. *Nature* *298*, 100–102.
- Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., et al. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* *55*, 611–622.
- Cabianca, D.S., Casa, V., Bodega, B., Xynos, A., Ginelli, E., Tanaka, Y., and Gabellini, D. (2012). A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in FSHD muscular dystrophy. *Cell* *149*, 819–831.
- Cacurri, S., Piazza, N., Deidda, G., Vigneti, E., Galluzzi, G., Colantoni, L., Merico, B., Ricci, E., and Felicetti, L. (1998). Sequence homology between 4qter and 10qter loci facilitates the instability of subtelomeric KpnI repeat units implicated in facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.* *63*, 181–190.
- Cantone, I., and Fisher, A.G. (2013). Epigenetic programming and reprogramming during development. *Nat. Struct. Mol. Biol.* *20*, 282–289.
- Canudas, S., Houghtaling, B.R., Bhanot, M., Sasa, G., Savage, S.A., Bertuch, A.A., and Smith, S. (2011). A role for heterochromatin protein 1 $\gamma$  at human telomeres. *Genes Dev.* *25*, 1807–1819.
- Carroll, D. (2008). Zinc-finger Nucleases as Gene Therapy Agents. *Gene Ther.* *15*, 1463–1468.
- Caruso, N., Herberth, B., Bartoli, M., Puppo, F., Dumonceaux, J., Zimmermann, A., Denadai, S., Lebossé, M., Roche, S., Geng, L., et al. (2013). Deregulation of the protocadherin gene FAT1 alters muscle shapes: implications for the pathogenesis of facioscapulohumeral dystrophy. *PLoS Genet.* *9*, e1003550.
- Cathomen, T., and Joung, J.K. (2008). Zinc-finger nucleases: the next generation emerges. *Mol. Ther. J. Am. Soc. Gene Ther.* *16*, 1200–1207.
- Cavazzana-Calvo, M., Hacein-Bey, S., de Saint Basile, G., Gross, F., Yvon, E., Nusbaum, P., Selz, F., Hue, C., Certain, S., Casanova, J.L., et al. (2000). Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science* *288*, 669–672.
- Cedar, H., and Bergman, Y. (2012). Programming of DNA methylation patterns. *Annu. Rev. Biochem.* *81*, 97–117.
- Celegato, B., Capitanio, D., Pescatori, M., Romualdi, C., Pacchioni, B., Cagnin, S., Viganò, A., Colantoni, L., Begum, S., Ricci, E., et al. (2006). Parallel protein and transcript profiles of FSHD patient muscles correlate to the D4Z4 arrangement and reveal a common impairment of slow to fast fibre differentiation and a general deregulation of MyoD-dependent genes. *Proteomics* *6*, 5303–5321.

- Cerritelli, S.M., and Crouch, R.J. (2009). Ribonuclease H: the enzymes in eukaryotes. *FEBS J.* 276, 1494–1505.
- Chan, Y.B., Miguel-Aliaga, I., Franks, C., Thomas, N., Trülzsch, B., Sattelle, D.B., Davies, K.E., and van den Heuvel, M. (2003). Neuromuscular defects in a *Drosophila* survival motor neuron gene mutant. *Hum. Mol. Genet.* 12, 1367–1376.
- Chang, N., Sun, C., Gao, L., Zhu, D., Xu, X., Zhu, X., Xiong, J.-W., and Xi, J.J. (2013). Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. *Cell Res.* 23, 465–472.
- Chang, Y.-F., Imam, J.S., and Wilkinson, M.F. (2007). The nonsense-mediated decay RNA surveillance pathway. *Annu. Rev. Biochem.* 76, 51–74.
- Chartier, A., Benoit, B., and Simonelig, M. (2006). A *Drosophila* model of oculopharyngeal muscular dystrophy reveals intrinsic toxicity of PABPN1. *EMBO J.* 25, 2253–2262.
- Chaudhuri, T., Mukherjea, M., Sachdev, S., Randall, J.D., and Sarkar, S. (2005). Role of the fetal and alpha/beta exons in the function of fast skeletal troponin T isoforms: correlation with altered Ca<sup>2+</sup> regulation associated with development. *J. Mol. Biol.* 352, 58–71.
- Cheli, S., François, S., Bodega, B., Ferrari, F., Tenedini, E., Roncaglia, E., Ferrari, S., Ginelli, E., and Meneveri, R. (2011). Expression profiling of FSHD-1 and FSHD-2 cells during myogenic differentiation evidences common and distinctive gene dysregulation patterns. *PloS One* 6, e20966.
- Chen, J.C., King, O.D., Zhang, Y., Clayton, N.P., Spencer, C., Wentworth, B.M., Emerson, C.P., and Wagner, K.R. (2016a). Morpholino-mediated Knockdown of DUX4 Toward Facioscapulohumeral Muscular Dystrophy Therapeutics. *Mol. Ther.* 24, 1405–1411.
- Chen, K., Czabotar, P.E., Blewitt, M.E., and Murphy, J.M. (2016b). The hinge domain of the epigenetic repressor Smc4d1 adopts an unconventional homodimeric configuration. *Biochem. J.* 473, 733–742.
- Chen, S.C., Frett, E., Marx, J., Bosnakovski, D., Reed, X., Kyba, M., and Kennedy, B.K. (2011). Decreased proliferation kinetics of mouse myoblasts overexpressing FRG1. *PloS One* 6, e19780.
- Chen, T., Ueda, Y., Dodge, J.E., Wang, Z., and Li, E. (2003). Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Mol. Cell. Biol.* 23, 5594–5605.
- Chen, T., Tsujimoto, N., and Li, E. (2004). The PWWP domain of Dnmt3a and Dnmt3b is required for directing DNA methylation to the major satellite repeats at pericentric heterochromatin. *Mol. Cell. Biol.* 24, 9048–9058.
- Cheng, X., and Blumenthal, R.M. (2008). Mammalian DNA methyltransferases: a structural perspective. *Struct. Lond. Engl.* 1993 16, 341–350.
- Chevalier, B.S., and Stoddard, B.L. (2001). Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Res.* 29, 3757–3774.
- Chirmule, N., Propert, K., Magosin, S., Qian, Y., Qian, R., and Wilson, J. (1999). Immune responses to adenovirus and adeno-associated virus in humans. *Gene Ther.* 6, 1574–1583.
- Cho, E., Feng, Y., Rauskolb, C., Maitra, S., Fehon, R., and Irvine, K.D. (2006). Delineation of a Fat tumor suppressor pathway. *Nat. Genet.* 38, 1142–1150.

- Christian, M., Cermak, T., Doyle, E.L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A.J., and Voytas, D.F. (2010). Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* 186, 757–761.
- Ciani, L., Patel, A., Allen, N.D., and French-Constant, C. (2003). Mice lacking the giant protocadherin mFAT1 exhibit renal slit junction abnormalities and a partially penetrant cyclopia and anophthalmia phenotype. *Mol. Cell. Biol.* 23, 3575–3582.
- Ciosk, R., Shirayama, M., Shevchenko, A., Tanaka, T., Toth, A., Shevchenko, A., and Nasmyth, K. (2000). Cohesin's binding to chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins. *Mol. Cell* 5, 243–254.
- Clapp, J., Mitchell, L.M., Bolland, D.J., Fantes, J., Corcoran, A.E., Scotting, P.J., Armour, J.A.L., and Hewitt, J.E. (2007). Evolutionary conservation of a coding function for D4Z4, the tandem DNA repeat mutated in facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.* 81, 264–279.
- Cockell, M., and Gasser, S.M. (1999). Nuclear compartments and gene regulation. *Curr. Opin. Genet. Dev.* 9, 199–205.
- Cockerill, P.N., and Garrard, W.T. (1986). Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites. *Cell* 44, 273–282.
- Coker, H., and Brockdorff, N. (2014). SMCHD1 accumulates at DNA damage sites and facilitates the repair of DNA double-strand breaks. *J. Cell Sci.* 127, 1869–1874.
- Cole, N.J., Tanaka, M., Prescott, A., and Tickle, C. (2003). Expression of limb initiation genes and clues to the morphological diversification of threespine stickleback. *Curr. Biol. CB* 13, R951-952.
- Collins, C.A., and Partridge, T.A. (2005). Self-renewal of the adult skeletal muscle satellite cell. *Cell Cycle Georget. Tex* 4, 1338–1341.
- Collins, K., and Mitchell, J.R. (2002). Telomerase in the human organism. *Oncogene* 21, 564–579.
- Colson, S.S., Benchortane, M., Tanant, V., Faghan, J.-P., Fournier-Mehouas, M., Benaim, C., Desnuelle, C., and Sacconi, S. (2010). Neuromuscular electrical stimulation training: a safe and effective treatment for facioscapulohumeral muscular dystrophy patients. *Arch. Phys. Med. Rehabil.* 91, 697–702.
- Cornu, T.I., and Cathomen, T. (2010). Quantification of zinc finger nuclease-associated toxicity. *Methods Mol. Biol. Clifton NJ* 649, 237–245.
- Cox, B., Hadjantonakis, A.K., Collins, J.E., and Magee, A.I. (2000). Cloning and expression throughout mouse development of mfat1, a homologue of the *Drosophila* tumour suppressor gene fat. *Dev. Dyn. Off. Publ. Am. Assoc. Anat.* 217, 233–240.
- Cox, D.B.T., Platt, R.J., and Zhang, F. (2015). Therapeutic Genome Editing: Prospects and Challenges. *Nat. Med.* 21, 121–131.
- Cremer, T., and Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat. Rev. Genet.* 2, 292–301.
- Cruveilhier (1852). Sur la paralysie musculaire, progressive, atrophique. *Bull. L'académie Natl. Médecine* 17ième Année 490 (séance du 29 mars 1853).

- Crystal, R.G. (1999). In vivo and ex vivo gene therapy strategies to treat tumors using adenovirus gene transfer vectors. *Cancer Chemother. Pharmacol.* *43 Suppl*, S90-99.
- Csink, A.K., and Henikoff, S. (1996). Genetic modification of heterochromatic association and nuclear organization in *Drosophila*. *Nature* *381*, 529–531.
- Cullen, B.R. (2005). Induction of stable RNA interference in mammalian cells. *Gene Ther.* *13*, 503–508.
- Cup, E.H., Pieterse, A.J., Ten Broek-Pastoor, J.M., Munneke, M., van Engelen, B.G., Hendricks, H.T., van der Wilt, G.J., and Oostendorp, R.A. (2007). Exercise therapy and other types of physical therapy for patients with neuromuscular diseases: a systematic review. *Arch. Phys. Med. Rehabil.* *88*, 1452–1464.
- Czerny, T., and Busslinger, M. (1995). DNA-binding and transactivation properties of Pax-6: three amino acids in the paired domain are responsible for the different sequence recognition of Pax-6 and BSAP (Pax-5). *Mol. Cell. Biol.* *15*, 2858–2871.
- Dahlberg, J.E., and Lund, E. (1987). Structure and expression of U-snRNA genes. *Mol. Biol. Rep.* *12*, 139–143.
- Damianov, A., and Black, D.L. (2010). Autoregulation of Fox protein expression to produce dominant negative splicing factors. *RNA N. Y. N* *16*, 405–416.
- Danckwardt, S., Hentze, M.W., and Kulozik, A.E. (2008). 3' end mRNA processing: molecular mechanisms and implications for health and disease. *EMBO J.* *27*, 482–498.
- Dandapat, A., Bosnakovski, D., Hartweck, L.M., Arpke, R.W., Baltgalvis, K.A., Vang, D., Baik, J., Darabi, R., Perlingeiro, R.C.R., Hamra, F.K., et al. (2014). Dominant Lethal Pathologies in Male Mice Engineered to Contain an X-Linked DUX4 Transgene. *Cell Rep.* *8*, 1484–1496.
- Das, S., and Chadwick, B.P. (2016). Influence of Repressive Histone and DNA Methylation upon D4Z4 Transcription in Non-Myogenic Cells. *PLoS ONE* *11*.
- Davis, C.G. (1990). The many faces of epidermal growth factor repeats. *New Biol.* *2*, 410–419.
- Daya, S., and Berns, K.I. (2008). Gene Therapy Using Adeno-Associated Virus Vectors. *Clin. Microbiol. Rev.* *21*, 583–593.
- De Angelis, F.G., Sthandier, O., Berarducci, B., Toso, S., Galluzzi, G., Ricci, E., Cossu, G., and Bozzoni, I. (2002). Chimeric snRNA molecules carrying antisense sequences against the splice junctions of exon 51 of the dystrophin pre-mRNA induce exon skipping and restoration of a dystrophin synthesis in Delta 48-50 DMD cells. *Proc. Natl. Acad. Sci. U. S. A.* *99*, 9456–9461.
- Deak, K.L., Lemmers, R.J.L.F., Stajich, J.M., Klooster, R., Tawil, R., Frants, R.R., Speer, M.C., van der Maarel, S.M., and Gilbert, J.R. (2007). Genotype-phenotype study in an FSHD family with a proximal deletion encompassing p13E-11 and D4Z4. *Neurology* *68*, 578–582.
- Dean, D.A. (2000). Peptide nucleic acids: versatile tools for gene therapy strategies. *Adv. Drug Deliv. Rev.* *44*, 81–95.
- Decary, S., Hamida, C.B., Mouly, V., Barbet, J.P., Hentati, F., and Butler-Browne, G.S. (2000). Shorter telomeres in dystrophic muscle consistent with extensive regeneration in young children. *Neuromuscul. Disord. NMD* *10*, 113–120.

Deidda, G., Cacurri, S., Grisanti, P., Vigneti, E., Piazzo, N., and Felicetti, L. (1995). Physical mapping evidence for a duplicated region on chromosome 10qter showing high homology with the facioscapulohumeral muscular dystrophy locus on chromosome 4qter. *Eur. J. Hum. Genet. EJHG* 3, 155–167.

Deidda, G., Cacurri, S., Piazzo, N., and Felicetti, L. (1996). Direct detection of 4q35 rearrangements implicated in facioscapulohumeral muscular dystrophy (FSHD). *J. Med. Genet.* 33, 361–365.

Déjardin, J., and Kingston, R.E. (2009). Purification of proteins associated with specific genomic Loci. *Cell* 136, 175–186.

Dejerine, and Landouzy (1884). De la myopathie atrophique progressive (myopathie héréditaire débutant, dans l'enfance, par la face, sans altération du système nerveux). *Comptes Rendus Hebd. Séances L'académie Sci. Tome* 98, p.53.

Dellavalle, A., Sampaolesi, M., Tonlorenzi, R., Tagliafico, E., Sacchetti, B., Perani, L., Innocenzi, A., Galvez, B.G., Messina, G., Morosetti, R., et al. (2007). Pericytes of human skeletal muscle are myogenic precursors distinct from satellite cells. *Nat. Cell Biol.* 9, 255–267.

Deng, D., Yan, C., Pan, X., Mahfouz, M., Wang, J., Zhu, J.-K., Shi, Y., and Yan, N. (2012). Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science* 335, 720–723.

Denti, M.A., Rosa, A., D'Antona, G., Sthandier, O., De Angelis, F.G., Nicoletti, C., Allocca, M., Pansarasa, O., Parente, V., Musarò, A., et al. (2006). Chimeric adeno-associated virus/antisense U1 small nuclear RNA effectively rescues dystrophin synthesis and muscle function by local treatment of mdx mice. *Hum. Gene Ther.* 17, 565–574.

Desjarlais, J.R., and Berg, J.M. (1993). Use of a zinc-finger consensus sequence framework and specificity rules to design specific DNA binding proteins. *Proc. Natl. Acad. Sci. U. S. A.* 90, 2256–2260.

van Deutekom, J.C., Wijmenga, C., van Tienhoven, E.A., Gruter, A.M., Hewitt, J.E., Padberg, G.W., van Ommen, G.J., Hofker, M.H., and Frants, R.R. (1993). FSHD associated DNA rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit. *Hum. Mol. Genet.* 2, 2037–2042.

van Deutekom, J.C., Bakker, E., Lemmers, R.J., van der Wielen, M.J., Bik, E., Hofker, M.H., Padberg, G.W., and Frants, R.R. (1996a). Evidence for subtelomeric exchange of 3.3 kb tandemly repeated units between chromosomes 4q35 and 10q26: implications for genetic counselling and etiology of FSHD1. *Hum. Mol. Genet.* 5, 1997–2003.

van Deutekom, J.C., Lemmers, R.J., Grewal, P.K., van Geel, M., Romberg, S., Dauwerse, H.G., Wright, T.J., Padberg, G.W., Hofker, M.H., Hewitt, J.E., et al. (1996b). Identification of the first gene (FRG1) from the FSHD region on human chromosome 4q35. *Hum. Mol. Genet.* 5, 581–590.

DeVincenzo, J., Cehelsky, J.E., Alvarez, R., Elbashir, S., Harborth, J., Toudjarska, I., Nechev, L., Murugaiah, V., Van Vliet, A., Vaishnav, A.K., et al. (2008). Evaluation of the safety, tolerability and pharmacokinetics of ALN-RSV01, a novel RNAi antiviral therapeutic directed against respiratory syncytial virus (RSV). *Antiviral Res.* 77, 225–231.

DeVos, S.L., and Miller, T.M. (2013). Antisense oligonucleotides: treating neurodegeneration at the level of RNA. *Neurother. J. Am. Soc. Exp. Neurother.* 10, 486–497.

Di Giammartino, D.C., Nishida, K., and Manley, J.L. (2011). Mechanisms and consequences of alternative polyadenylation. *Mol. Cell* 43, 853–866.

- Dias, N., and Stein, C.A. (2002). Antisense Oligonucleotides: Basic Concepts and Mechanisms. *Mol. Cancer Ther.* *1*, 347–355.
- Dillon, N. (2004). Heterochromatin structure and function. *Biol. Cell* *96*, 631–637.
- DiMattia, M.A., Nam, H.-J., Van Vliet, K., Mitchell, M., Bennett, A., Gurda, B.L., McKenna, R., Olson, N.H., Sinkovits, R.S., Potter, M., et al. (2012). Structural insight into the unique properties of adeno-associated virus serotype 9. *J. Virol.* *86*, 6947–6958.
- Dimitri, P., Caizzi, R., Giordano, E., Carmela Accardo, M., Lattanzi, G., and Biamonti, G. (2009). Constitutive heterochromatin: a surprising variety of expressed sequences. *Chromosoma* *118*, 419–435.
- Ding, H., Beckers, M.C., Plaisance, S., Marynen, P., Collen, D., and Belayew, A. (1998). Characterization of a double homeodomain protein (DUX1) encoded by a cDNA homologous to 3.3 kb dispersed repeated elements. *Hum. Mol. Genet.* *7*, 1681–1694.
- Ding, W., Zhang, L., Yan, Z., and Engelhardt, J.F. (2005). Intracellular trafficking of adeno-associated viral vectors. *Gene Ther.* *12*, 873–880.
- Dixit, M., Anseau, E., Tassin, A., Winokur, S., Shi, R., Qian, H., Sauvage, S., Mattéotti, C., van Acker, A.M., Leo, O., et al. (2007). DUX4, a candidate gene of facioscapulohumeral muscular dystrophy, encodes a transcriptional activator of PITX1. *Proc. Natl. Acad. Sci. U. S. A.* *104*, 18157–18162.
- Dmitriev, P., Petrov, A., Anseau, E., Stankevics, L., Charron, S., Kim, E., Bos, T.J., Robert, T., Turki, A., Coppée, F., et al. (2011). The Krüppel-like factor 15 as a molecular link between myogenic factors and a chromosome 4q transcriptional enhancer implicated in facioscapulohumeral dystrophy. *J. Biol. Chem.* *286*, 44620–44631.
- Doerner, A., Pauschinger, M., Badorff, A., Noutsias, M., Giessen, S., Schulze, K., Bilger, J., Rauch, U., and Schultheiss, H.P. (1997). Tissue-specific transcription pattern of the adenine nucleotide translocase isoforms in humans. *FEBS Lett.* *414*, 258–262.
- Dolce, V., Scarcia, P., Iacopetta, D., and Palmieri, F. (2005). A fourth ADP/ATP carrier isoform in man: identification, bacterial expression, functional characterization and tissue distribution. *FEBS Lett.* *579*, 633–637.
- Dominski, Z., and Marzluff, W.F. (1999). Formation of the 3' end of histone mRNA. *Gene* *239*, 1–14.
- Dominski, Z., and Marzluff, W.F. (2007). Formation of the 3' end of histone mRNA: getting closer to the end. *Gene* *396*, 373–390.
- Dominski, Z., Zheng, L.X., Sanchez, R., and Marzluff, W.F. (1999). Stem-loop binding protein facilitates 3'-end formation by stabilizing U7 snRNP binding to histone pre-mRNA. *Mol. Cell. Biol.* *19*, 3561–3570.
- Dominski, Z., Erkmann, J.A., Yang, X., Sánchez, R., and Marzluff, W.F. (2002). A novel zinc finger protein is associated with U7 snRNP and interacts with the stem-loop binding protein in the histone pre-mRNP to stimulate 3'-end processing. *Genes Dev.* *16*, 58–71.
- Dominski, Z., Yang, X., and Marzluff, W.F. (2005). The polyadenylation factor CPSF-73 is involved in histone-pre-mRNA processing. *Cell* *123*, 37–48.

- Dorobek, M., Szmidt-Sałkowska, E., Rowińska-Marcińska, K., Gawęł, M., and Hausmanowa-Petrusewicz, I. (2013). Relationships between clinical data and quantitative EMG findings in facioscapulohumeral muscular dystrophy. *Neurol. Neurochir. Pol.* 47, 8–17.
- Douar, A.-M., Poulard, K., Stockholm, D., and Danos, O. (2001). Intracellular Trafficking of Adeno-Associated Virus Vectors: Routing to the Late Endosomal Compartment and Proteasome Degradation. *J. Virol.* 75, 1824–1833.
- Dreier, B., Beerli, R.R., Segal, D.J., Flippin, J.D., and Barbas, C.F. (2001). Development of zinc finger domains for recognition of the 5'-ANN-3' family of DNA sequences and their use in the construction of artificial transcription factors. *J. Biol. Chem.* 276, 29466–29478.
- Duan, D., Li, Q., Kao, A.W., Yue, Y., Pessin, J.E., and Engelhardt, J.F. (1999). Dynamin is required for recombinant adeno-associated virus type 2 infection. *J. Virol.* 73, 10371–10376.
- Duchenne, G. (1853). Etude comparée des lésions anatomiques de l'atrophie musculaire graisseuse progressive et de la paralysie générale. *Mém. Lu À Société Méd.-Chir. Paris Dans Séances* 11 Mars 8 Avril 1853.
- Duchenne, G. (1858). De l'ataxie locomotrice progressive. *Arch. Générales Médecine* 5 sér., 12: 641-652; 13: 36-62, 158-1881, 417-451.
- Duchenne, G. (1868). De la paralysie musculaire pseudo-hypertrophique ou paralysie myo-sclérotique. *Arch. Générales Médecine* 11 : 5, 178, 305, 421, 552.
- Duffy, J.B. (2002). GAL4 system in *Drosophila*: a fly geneticist's Swiss army knife. *Genes. N. Y. N* 2000 34, 1–15.
- Dunne, J., Hanby, A.M., Poulson, R., Jones, T.A., Sheer, D., Chin, W.G., Da, S.M., Zhao, Q., Beverley, P.C., and Owen, M.J. (1995). Molecular cloning and tissue expression of FAT, the human homologue of the *Drosophila* fat gene that is located on chromosome 4q34-q35 and encodes a putative adhesion molecule. *Genomics* 30, 207–223.
- Dutta, R., and Inouye, M. (2000). GHKL, an emergent ATPase/kinase superfamily. *Trends Biochem. Sci.* 25, 24–28.
- Eberl, H.C., Spruijt, C.G., Kelstrup, C.D., Vermeulen, M., and Mann, M. (2013). A map of general and specialized chromatin readers in mouse tissues generated by label-free interaction proteomics. *Mol. Cell* 49, 368–378.
- Eberle, A.B., Stalder, L., Mathys, H., Orozco, R.Z., and Mühlemann, O. (2008). Posttranscriptional gene regulation by spatial rearrangement of the 3' untranslated region. *PLoS Biol.* 6, e92.
- Eckstein, F. (2000). Phosphorothioate oligodeoxynucleotides: what is their origin and what is unique about them? *Antisense Nucleic Acid Drug Dev.* 10, 117–121.
- Edwards, R.A., and Bryan, J. (1995). Fascins, a family of actin bundling proteins. *Cell Motil. Cytoskeleton* 32, 1–9.
- Egholm, M., Buchardt, O., Christensen, L., Behrens, C., Freier, S.M., Driver, D.A., Berg, R.H., Kim, S.K., Norden, B., and Nielsen, P.E. (1993). PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules. *Nature* 365, 566–568.
- Ehrlich, M., and Lacey, M. (2012). Deciphering transcription dysregulation in FSH muscular dystrophy. *J. Hum. Genet.* 57, 477–484.

- Ehrlich, M., Jackson, K., Tsumagari, K., Camaño, P., and Lemmers, R.J.F.L. (2007). Hybridization analysis of D4Z4 repeat arrays linked to FSHD. *Chromosoma* *116*, 107–116.
- Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* *411*, 494–498.
- Elkon, R., Ugalde, A.P., and Agami, R. (2013). Alternative cleavage and polyadenylation: extent, regulation and function. *Nat. Rev. Genet.* *14*, 496–506.
- Emery, A.E.H. (2002). The muscular dystrophies. *Lancet Lond. Engl.* *359*, 687–695.
- Escors, D., and Breckpot, K. (2010). Lentiviral vectors in gene therapy: their current status and future potential. *Arch. Immunol. Ther. Exp. (Warsz.)* *58*, 107–119.
- Farrar, G.J., Millington-Ward, S., Chadderton, N., Humphries, P., and Kenna, P.F. (2012). Gene-based therapies for dominantly inherited retinopathies. *Gene Ther.* *19*, 137–144.
- Faustmann, P.M., Farahati, J., Rupilius, B., Dux, R., Koch, M.C., and Reiners, C. (1996). Cardiac involvement in facio-scapulo-humeral muscular dystrophy: a family study using Thallium-201 single-photon-emission-computed tomography. *J. Neurol. Sci.* *144*, 59–63.
- Feinberg, A.P. (2007). Phenotypic plasticity and the epigenetics of human disease. *Nature* *447*, 433–440.
- Feng, Q., Snider, L., Jagannathan, S., Tawil, R., van der Maarel, S.M., Tapscott, S.J., and Bradley, R.K. (2015). A feedback loop between nonsense-mediated decay and the retrogene DUX4 in facioscapulohumeral muscular dystrophy. *eLife* *4*.
- Ferrari, F.K., Samulski, T., Shenk, T., and Samulski, R.J. (1996). Second-strand synthesis is a rate-limiting step for efficient transduction by recombinant adeno-associated virus vectors. *J. Virol.* *70*, 3227–3234.
- Ferreboeuf, M., Mariot, V., Bessières, B., Vasiljevic, A., Attié-Bitach, T., Collardeau, S., Morere, J., Roche, S., Magdinier, F., Robin-Ducellier, J., et al. (2014a). DUX4 and DUX4 downstream target genes are expressed in fetal FSHD muscles. *Hum. Mol. Genet.* *23*, 171–181.
- Ferreboeuf, M., Mariot, V., Furling, D., Butler-Browne, G., Mouly, V., and Dumonceaux, J. (2014b). Nuclear protein spreading: implication for pathophysiology of neuromuscular diseases. *Hum. Mol. Genet.* *23*, 4125–4133.
- Ferrer, A., Foster, H., Wells, K.E., Dickson, G., and Wells, D.J. (2004). Long-term expression of full-length human dystrophin in transgenic mdx mice expressing internally deleted human dystrophins. *Gene Ther.* *11*, 884–893.
- Ferri, G., Huichalaf, C.H., Caccia, R., and Gabellini, D. (2015). Direct interplay between two candidate genes in FSHD muscular dystrophy. *Hum. Mol. Genet.* *24*, 1256–1266.
- Filipowicz, W., Bhattacharyya, S.N., and Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat. Rev. Genet.* *9*, 102–114.
- Filippova, G.N. (2008). Genetics and epigenetics of the multifunctional protein CTCF. *Curr. Top. Dev. Biol.* *80*, 337–360.
- Finlan, L.E., Sproul, D., Thomson, I., Boyle, S., Kerr, E., Perry, P., Ylstra, B., Chubb, J.R., and Bickmore, W.A. (2008). Recruitment to the nuclear periphery can alter expression of genes in human cells. *PLoS Genet.* *4*, e1000039.



- Finsterer, J., Stöllberger, C., and Meng, G. (2005). Cardiac involvement in facioscapulohumeral muscular dystrophy. *Cardiology* *103*, 81–83.
- Fiorini, A., Gouveia, F. de S., and Fernandez, M.A. (2006). Scaffold/Matrix Attachment Regions and intrinsic DNA curvature. *Biochem. Biokhimiia* *71*, 481–488.
- Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* *391*, 806–811.
- Fitzsimons, R.B., Gurwin, E.B., and Bird, A.C. (1987). Retinal vascular abnormalities in facioscapulohumeral muscular dystrophy. A general association with genetic and therapeutic implications. *Brain J. Neurol.* *110* ( Pt 3), 631–648.
- Flower, D.R. (1996). The lipocalin protein family: structure and function. *Biochem. J.* *318* ( Pt 1), 1–14.
- Foletta, V.C., White, L.J., Larsen, A.E., Léger, B., and Russell, A.P. (2011). The role and regulation of MAFbx/atrogin-1 and MuRF1 in skeletal muscle atrophy. *Pflugers Arch.* *461*, 325–335.
- Fox, D.K., Ebert, S.M., Bongers, K.S., Dyle, M.C., Bullard, S.A., Dierdorff, J.M., Kunkel, S.D., and Adams, C.M. (2014). p53 and ATF4 mediate distinct and additive pathways to skeletal muscle atrophy during limb immobilization. *Am. J. Physiol. Endocrinol. Metab.* *307*, E245-261.
- Francastel, C., Schübeler, D., Martin, D.I., and Groudine, M. (2000). Nuclear compartmentalization and gene activity. *Nat. Rev. Mol. Cell Biol.* *1*, 137–143.
- Frontera, W.R., and Ochala, J. (2015). Skeletal muscle: a brief review of structure and function. *Calcif. Tissue Int.* *96*, 183–195.
- Fukumura, K., Kato, A., Jin, Y., Ideue, T., Hirose, T., Kataoka, N., Fujiwara, T., Sakamoto, H., and Inoue, K. (2007). Tissue-specific splicing regulator Fox-1 induces exon skipping by interfering E complex formation on the downstream intron of human F1gamma gene. *Nucleic Acids Res.* *35*, 5303–5311.
- Funakoshi, M., Goto, K., and Arahata, K. (1998). Epilepsy and mental retardation in a subset of early onset 4q35-facioscapulohumeral muscular dystrophy. *Neurology* *50*, 1791–1794.
- Furger, A., Schaller, A., and Schümperli, D. (1998). Functional importance of conserved nucleotides at the histone RNA 3' processing site. *RNA N. Y. N* *4*, 246–256.
- Gabellini, D., Green, M.R., and Tupler, R. (2002). Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. *Cell* *110*, 339–348.
- Gabellini, D., D'Antona, G., Moggio, M., PELLE, A., Zecca, C., Adami, R., Angeletti, B., Ciscato, P., Pellegrino, M.A., Bottinelli, R., et al. (2006). Facioscapulohumeral muscular dystrophy in mice overexpressing FRG1. *Nature* *439*, 973–977.
- Gabriëls, J., Beckers, M.C., Ding, H., De Vriese, A., Plaisance, S., van der Maarel, S.M., Padberg, G.W., Frants, R.R., Hewitt, J.E., Collen, D., et al. (1999). Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. *Gene* *236*, 25–32.
- Gaillard, M.-C., Roche, S., Dion, C., Tasmadjian, A., Bouget, G., Salort-Campana, E., Vovan, C., Chaix, C., Broucqsaault, N., Morere, J., et al. (2014). Differential DNA methylation of the D4Z4 repeat in patients with FSHD and asymptomatic carriers. *Neurology* *83*, 733–742.

- Galetta, F., Franzoni, F., Sposito, R., Plantinga, Y., Femia, F.R., Galluzzi, F., Rocchi, A., Santoro, G., and Siciliano, G. (2005). Subclinical cardiac involvement in patients with facioscapulohumeral muscular dystrophy. *Neuromuscul. Disord. NMD* 15, 403–408.
- Galińska-Rakoczy, A., Engel, P., Xu, C., Jung, H., Craig, R., Tobacman, L.S., and Lehman, W. (2008). Structural basis for the regulation of muscle contraction by troponin and tropomyosin. *J. Mol. Biol.* 379, 929–935.
- Galli, G., Hofstetter, H., Stunnenberg, H.G., and Birnstiel, M.L. (1983). Biochemical complementation with RNA in the *Xenopus* oocyte: a small RNA is required for the generation of 3' histone mRNA termini. *Cell* 34, 823–828.
- Galluzzi, G., Deidda, G., Cacurri, S., Colantoni, L., Piazza, N., Vigneti, E., Ricci, E., Servidei, S., Merico, B., Pachì, A., et al. (1999). Molecular analysis of 4q35 rearrangements in facioscapulohumeral muscular dystrophy (FSHD): application to family studies for a correct genetic advice and a reliable prenatal diagnosis of the disease. *Neuromuscul. Disord. NMD* 9, 190–198.
- Gan, H.H., and Gunsalus, K.C. (2015). Assembly and analysis of eukaryotic Argonaute–RNA complexes in microRNA-target recognition. *Nucleic Acids Res.* 43, 9613–9625.
- Gao, G., Wang, Q., Calcedo, R., Mays, L., Bell, P., Wang, L., Vandenberghe, L.H., Grant, R., Sanmiguel, J., Furth, E.E., et al. (2009). Adeno-associated virus-mediated gene transfer to nonhuman primate liver can elicit destructive transgene-specific T cell responses. *Hum. Gene Ther.* 20, 930–942.
- Gao, G., Zhong, L., and Danos, O. (2011). Exploiting natural diversity of AAV for the design of vectors with novel properties. *Methods Mol. Biol. Clifton NJ* 807, 93–118.
- García, L., Urbiola, K., Düzgüneş, N., and Tros de Ilarduya, C. (2012). Lipopolyplexes as nanomedicines for therapeutic gene delivery. *Methods Enzymol.* 509, 327–338.
- Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* 109, E2579–2586.
- Gasser, S.M., and Laemmli, U.K. (1986). Cohabitation of scaffold binding regions with upstream/enhancer elements of three developmentally regulated genes of *D. melanogaster*. *Cell* 46, 521–530.
- Gaszner, M., and Felsenfeld, G. (2006). Insulators: exploiting transcriptional and epigenetic mechanisms. *Nat. Rev. Genet.* 7, 703–713.
- Gatchel, J.R., and Zoghbi, H.Y. (2005). Diseases of unstable repeat expansion: mechanisms and common principles. *Nat. Rev. Genet.* 6, 743–755.
- van Geel, M., Dickson, M.C., Beck, A.F., Bolland, D.J., Frants, R.R., van der Maarel, S.M., de Jong, P.J., and Hewitt, J.E. (2002). Genomic analysis of human chromosome 10q and 4q telomeres suggests a common origin. *Genomics* 79, 210–217.
- Gehl, J. (2014). Gene electrotransfer in clinical trials. *Methods Mol. Biol. Clifton NJ* 1121, 241–246.
- Geiman, T.M., and Muegge, K. (2010). DNA methylation in early development. *Mol. Reprod. Dev.* 77, 105–113.
- Gendrel, A.-V., Tang, Y.A., Suzuki, M., Godwin, J., Nesterova, T.B., Grealley, J.M., Heard, E., and Brockdorff, N. (2013). Epigenetic functions of *smc4d1* repress gene clusters on the inactive X chromosome and on autosomes. *Mol. Cell. Biol.* 33, 3150–3165.

- Geng, L.N., Tyler, A.E., and Tapscott, S.J. (2011). Immunodetection of human double homeobox 4. *Hybrid*. 2005 *30*, 125–130.
- Geng, L.N., Yao, Z., Snider, L., Fong, A.P., Cech, J.N., Young, J.M., van der Maarel, S.M., Ruzzo, W.L., Gentleman, R.C., Tawil, R., et al. (2012). DUX4 activates germline genes, retroelements, and immune mediators: implications for facioscapulohumeral dystrophy. *Dev. Cell* *22*, 38–51.
- Genovese, P., Schirotti, G., Escobar, G., Di Tomaso, T., Firrito, C., Calabria, A., Moi, D., Mazzieri, R., Bonini, C., Holmes, M.C., et al. (2014). Targeted genome editing in human repopulating haematopoietic stem cells. *Nature* *510*, 235–240.
- Giannini, S., Faldini, C., Pagkrati, S., Grandi, G., Digennaro, V., Luciani, D., and Merlini, L. (2007). Fixation of Winged Scapula in Facioscapulohumeral Muscular Dystrophy. *Clin. Med. Res.* *5*, 155–162.
- Gilbert, J.R., Stajich, J.M., Speer, M.C., Vance, J.M., Stewart, C.S., Yamaoka, L.H., Samson, F., Fardeau, M., Potter, T.G., and Roses, A.D. (1992). Linkage studies in facioscapulohumeral muscular dystrophy (FSHD). *Am. J. Hum. Genet.* *51*, 424–427.
- Gilbert, J.R., Stajich, J.M., Wall, S., Carter, S.C., Qiu, H., Vance, J.M., Stewart, C.S., Speer, M.C., Pufky, J., and Yamaoka, L.H. (1993). Evidence for heterogeneity in facioscapulohumeral muscular dystrophy (FSHD). *Am. J. Hum. Genet.* *53*, 401–408.
- Gillis, J.-M. (2002). Multivariate evaluation of the functional recovery obtained by the overexpression of utrophin in skeletal muscles of the mdx mouse. *Neuromuscul. Disord. NMD* *12 Suppl 1*, S90-94.
- Ginn, S.L., Alexander, I.E., Edelstein, M.L., Abedi, M.R., and Wixon, J. (2013). Gene therapy clinical trials worldwide to 2012 - an update. *J. Gene Med.* *15*, 65–77.
- Girod, A., Wobus, C.E., Zádori, Z., Ried, M., Leike, K., Tijssen, P., Kleinschmidt, J.A., and Hallek, M. (2002). The VP1 capsid protein of adeno-associated virus type 2 is carrying a phospholipase A2 domain required for virus infectivity. *J. Gen. Virol.* *83*, 973–978.
- Goll, M.G., and Bestor, T.H. (2005). Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.* *74*, 481–514.
- Gombert, W.M., Farris, S.D., Rubio, E.D., Morey-Rosler, K.M., Schubach, W.H., and Krumm, A. (2003). The c-myc Insulator Element and Matrix Attachment Regions Define the c-myc Chromosomal Domain. *Mol. Cell. Biol.* *23*, 9338–9348.
- Gomes, M.D., Lecker, S.H., Jagoe, R.T., Navon, A., and Goldberg, A.L. (2001). Atrogin-1, a muscle-specific F-box protein highly expressed during muscle atrophy. *Proc. Natl. Acad. Sci. U. S. A.* *98*, 14440–14445.
- Gordon, A.M., Homsher, E., and Regnier, M. (2000). Regulation of contraction in striated muscle. *Physiol. Rev.* *80*, 853–924.
- Gordon, C.T., Xue, S., Yigit, G., Filali, H., Chen, K., Rosin, N., Yoshiura, K.-I., Oufadem, M., Beck, T.J., McGowan, R., et al. (2017). De novo mutations in SMCHD1 cause Bosma arhinia microphthalmia syndrome and abrogate nasal development. *Nat. Genet.* *49*, 249–255.
- Gorgonie, B., Andrews, S., Schaller, A., Schümperli, D., Gray, N.K., and Müller, B. (2005). The stem-loop binding protein stimulates histone translation at an early step in the initiation pathway. *RNA* *11*, 1030–1042.

- Gorman, L., Suter, D., Emerick, V., Schümperli, D., and Kole, R. (1998). Stable alteration of pre-mRNA splicing patterns by modified U7 small nuclear RNAs. *Proc. Natl. Acad. Sci. U. S. A.* *95*, 4929–4934.
- Gottschling, D.E., Aparicio, O.M., Billington, B.L., and Zakian, V.A. (1990). Position effect at *S. cerevisiae* telomeres: reversible repression of Pol II transcription. *Cell* *63*, 751–762.
- Goudy, K.S., Wang, B., and Tisch, R. (2008). Gene gun-mediated DNA vaccination enhances antigen-specific immunotherapy at a late preclinical stage of type 1 diabetes in nonobese diabetic mice. *Clin. Immunol. Orlando Fla* *129*, 49–57.
- Goyenvalle, A., Vulin, A., Fougères, F., Leturcq, F., Kaplan, J.-C., Garcia, L., and Danos, O. (2004). Rescue of dystrophic muscle through U7 snRNA-mediated exon skipping. *Science* *306*, 1796–1799.
- Goyenvalle, A., Babbs, A., van Ommen, G.-J.B., Garcia, L., and Davies, K.E. (2009). Enhanced Exon-skipping Induced by U7 snRNA Carrying a Splicing Silencer Sequence: Promising Tool for DMD Therapy. *Mol. Ther. J. Am. Soc. Gene Ther.* *17*, 1234–1240.
- Goyenvalle, A., Wright, J., Babbs, A., Wilkins, V., Garcia, L., and Davies, K.E. (2012). Engineering Multiple U7snRNA Constructs to Induce Single and Multiexon-skipping for Duchenne Muscular Dystrophy. *Mol. Ther.* *20*, 1212–1221.
- Gratz, S.J., Cummings, A.M., Nguyen, J.N., Hamm, D.C., Donohue, L.K., Harrison, M.M., Wildonger, J., and O'Connor-Giles, K.M. (2013). Genome Engineering of *Drosophila* with the CRISPR RNA-Guided Cas9 Nuclease. *Genetics* *194*, 1029–1035.
- de Greef, J.C., Wohlgemuth, M., Chan, O.A., Hansson, K.B., Smeets, D., Frants, R.R., Weemaes, C.M., Padberg, G.W., and van der Maarel, S.M. (2007). Hypomethylation is restricted to the D4Z4 repeat array in phenotypic FSHD. *Neurology* *69*, 1018–1026.
- de Greef, J.C., Lemmers, R.J.L.F., van Engelen, B.G.M., Sacconi, S., Venance, S.L., Frants, R.R., Tawil, R., and van der Maarel, S.M. (2009). Common epigenetic changes of D4Z4 in contraction-dependent and contraction-independent FSHD. *Hum. Mutat.* *30*, 1449–1459.
- de Greef, J.C., Lemmers, R.J.L.F., Camaño, P., Day, J.W., Sacconi, S., Dunand, M., van Engelen, B.G.M., Kiuru-Enari, S., Padberg, G.W., Rosa, A.L., et al. (2010). Clinical features of facioscapulohumeral muscular dystrophy 2. *Neurology* *75*, 1548–1554.
- Gregorevic, P., Blankinship, M.J., Allen, J.M., Crawford, R.W., Meuse, L., Miller, D.G., Russell, D.W., and Chamberlain, J.S. (2004). Systemic delivery of genes to striated muscles using adeno-associated viral vectors. *Nat. Med.* *10*, 828–834.
- Gregory, G.D., Vakoc, C.R., Rozovskaia, T., Zheng, X., Patel, S., Nakamura, T., Canaani, E., and Blobel, G.A. (2007). Mammalian ASH1L is a histone methyltransferase that occupies the transcribed region of active genes. *Mol. Cell. Biol.* *27*, 8466–8479.
- Greider, C.W. (1991). Telomerase is processive. *Mol. Cell. Biol.* *11*, 4572–4580.
- Greider, C.W., and Blackburn, E.H. (1985). Identification of a specific telomere terminal transferase activity in *Tetrahymena* extracts. *Cell* *43*, 405–413.
- Grewal, P.K., van Geel, M., Frants, R.R., de Jong, P., and Hewitt, J.E. (1999). Recent amplification of the human FRG1 gene during primate evolution. *Gene* *227*, 79–88.

- Grieger, J.C., and Samulski, R.J. (2012). Adeno-associated virus vectorology, manufacturing, and clinical applications. *Methods Enzymol.* *507*, 229–254.
- Griggs, R.C., Tawil, R., McDermott, M., Forrester, J., Figlewicz, D., and Weiffenbach, B. (1995). Monozygotic twins with facioscapulohumeral dystrophy (FSHD): implications for genotype/phenotype correlation. *Muscle Nerve. Suppl.* S50-55.
- Grimm, C., Stefanovic, B., and Schümperli, D. (1993). The low abundance of U7 snRNA is partly determined by its Sm binding site. *EMBO J.* *12*, 1229–1238.
- Grimm, D., Streetz, K.L., Jopling, C.L., Storm, T.A., Pandey, K., Davis, C.R., Marion, P., Salazar, F., and Kay, M.A. (2006). Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature* *441*, 537–541.
- Grosso, S., Mostardini, R., Di Bartolo, R.M., Balestri, P., and Verrotti, A. (2011). Epilepsy, speech delay, and mental retardation in facioscapulohumeral muscular dystrophy. *Eur. J. Paediatr. Neurol. EJPN Off. J. Eur. Paediatr. Neurol. Soc.* *15*, 456–460.
- Gruber, S., Haering, C.H., and Nasmyth, K. (2003). Chromosomal cohesin forms a ring. *Cell* *112*, 765–777.
- Grüter, P., Tabernero, C., von Kobbe, C., Schmitt, C., Saavedra, C., Bachi, A., Wilm, M., Felber, B.K., and Izaurralde, E. (1998). TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus. *Mol. Cell* *1*, 649–659.
- Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W., et al. (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* *453*, 948–951.
- Guttridge, D.C., Mayo, M.W., Madrid, L.V., Wang, C.Y., and Baldwin, A.S. (2000). NF-kappaB-induced loss of MyoD messenger RNA: possible role in muscle decay and cachexia. *Science* *289*, 2363–2366.
- Ha, M., and Kim, V.N. (2014). Regulation of microRNA biogenesis. *Nat. Rev. Mol. Cell Biol.* *15*, 509–524.
- Hacein-Bey-Abina, S., Le Deist, F., Carlier, F., Bouneaud, C., Hue, C., De Villartay, J.-P., Thrasher, A.J., Wulffraat, N., Sorensen, R., Dupuis-Girod, S., et al. (2002). Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. *N. Engl. J. Med.* *346*, 1185–1193.
- Haidet, A.M., Rizo, L., Handy, C., Umapathi, P., Eagle, A., Shilling, C., Boue, D., Martin, P.T., Sahenk, Z., Mendell, J.R., et al. (2008). Long-term enhancement of skeletal muscle mass and strength by single gene administration of myostatin inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 4318–4322.
- Halbert, C.L., Standaert, T.A., Wilson, C.B., and Miller, A.D. (1998). Successful readministration of adeno-associated virus vectors to the mouse lung requires transient immunosuppression during the initial exposure. *J. Virol.* *72*, 9795–9805.
- Hall-Pogar, T., Zhang, H., Tian, B., and Lutz, C.S. (2005). Alternative polyadenylation of cyclooxygenase-2. *Nucleic Acids Res.* *33*, 2565–2579.
- Hanakahi, L.A., Dempsey, L.A., Li, M.-J., and Maizels, N. (1997). Nucleolin is one component of the B cell-specific transcription factor and switch region binding protein, LR1. *Proc. Natl. Acad. Sci. U. S. A.* *94*, 3605–3610.

- Hanel, M.L., Wuebbles, R.D., and Jones, P.L. (2009). Muscular dystrophy candidate gene FRG1 is critical for muscle development. *Dev. Dyn. Off. Publ. Am. Assoc. Anat.* 238, 1502–1512.
- Hanel, M.L., Sun, C.-Y.J., Jones, T.I., Long, S.W., Zanotti, S., Milner, D., and Jones, P.L. (2011). Facioscapulohumeral muscular dystrophy (FSHD) region gene 1 (FRG1) is a dynamic nuclear and sarcomeric protein. *Differ. Res. Biol. Divers.* 81, 107–118.
- Haraguchi, Y., Chung, A.B., Torroni, A., Stepien, G., Shoffner, J.M., Wasmuth, J.J., Costigan, D.A., Polak, M., Altherr, M.R., and Winokur, S.T. (1993). Genetic mapping of human heart-skeletal muscle adenine nucleotide translocator and its relationship to the facioscapulohumeral muscular dystrophy locus. *Genomics* 16, 479–485.
- Hark, A.T., Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M., and Tilghman, S.M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* 405, 486–489.
- Harley, C.B. (1991). Telomere loss: mitotic clock or genetic time bomb? *Mutat. Res.* 256, 271–282.
- Harley, C.B., Futcher, A.B., and Greider, C.W. (1990). Telomeres shorten during ageing of human fibroblasts. *Nature* 345, 458–460.
- de Haro, M., Al-Ramahi, I., De Gouyon, B., Ukani, L., Rosa, A., Faustino, N.A., Ashizawa, T., Cooper, T.A., and Botas, J. (2006). MBNL1 and CUGBP1 modify expanded CUG-induced toxicity in a Drosophila model of myotonic dystrophy type 1. *Hum. Mol. Genet.* 15, 2138–2145.
- Harold, C. (2006). Atlas of genetic diagnosis and counseling. 375–380.
- Hartweck, L.M., Anderson, L.J., Lemmers, R.J., Dandapat, A., Toso, E.A., Dalton, J.C., Tawil, R., Day, J.W., van der Maarel, S.M., and Kyba, M. (2013). A focal domain of extreme demethylation within D4Z4 in FSHD2. *Neurology* 80, 392–399.
- Harumoto, T., Ito, M., Shimada, Y., Kobayashi, T.J., Ueda, H.R., Lu, B., and Uemura, T. (2010). Atypical cadherins Dachous and Fat control dynamics of noncentrosomal microtubules in planar cell polarity. *Dev. Cell* 19, 389–401.
- Hastie, N.D., Dempster, M., Dunlop, M.G., Thompson, A.M., Green, D.K., and Allshire, R.C. (1990). Telomere reduction in human colorectal carcinoma and with ageing. *Nature* 346, 866–868.
- Hauck, B., Zhao, W., High, K., and Xiao, W. (2004). Intracellular viral processing, not single-stranded DNA accumulation, is crucial for recombinant adeno-associated virus transduction. *J. Virol.* 78, 13678–13686.
- Hayakawa, T., Haraguchi, T., Masumoto, H., and Hiraoka, Y. (2003). Cell cycle behavior of human HP1 subtypes: distinct molecular domains of HP1 are required for their centromeric localization during interphase and metaphase. *J. Cell Sci.* 116, 3327–3338.
- Heitz, E. (1928). Das Heterochromatin der Moose. *I. Jahrb Wiss Bot* 69, 762–818.
- Hermonat, P.L., and Muzyczka, N. (1984). Use of adeno-associated virus as a mammalian DNA cloning vector: transduction of neomycin resistance into mammalian tissue culture cells. *Proc. Natl. Acad. Sci. U. S. A.* 81, 6466–6470.
- Herrero, M.J., Sabater, L., Guenechea, G., Sendra, L., Montilla, A.I., Abargues, R., Navarro, V., and Aliño, S.F. (2012). DNA delivery to “ex vivo” human liver segments. *Gene Ther.* 19, 504–512.

- Herweijer, H., and Wolff, J.A. (2003). Progress and prospects: naked DNA gene transfer and therapy. *Gene Ther.* *10*, 453–458.
- Hewitt, J.E., Lyle, R., Clark, L.N., Valleley, E.M., Wright, T.J., Wijmenga, C., van Deutekom, J.C., Francis, F., Sharpe, P.T., and Hofker, M. (1994). Analysis of the tandem repeat locus D4Z4 associated with facioscapulohumeral muscular dystrophy. *Hum. Mol. Genet.* *3*, 1287–1295.
- Hikida, R.S. (2011). Aging changes in satellite cells and their functions. *Curr. Aging Sci.* *4*, 279–297.
- Himeda, C.L., Debarnot, C., Homma, S., Beermann, M.L., Miller, J.B., Jones, P.L., and Jones, T.I. (2014). Myogenic Enhancers Regulate Expression of the Facioscapulohumeral Muscular Dystrophy-Associated DUX4 Gene. *Mol. Cell. Biol.* *34*, 1942–1955.
- Himeda, C.L., Jones, T.I., and Jones, P.L. (2015). Facioscapulohumeral muscular dystrophy as a model for epigenetic regulation and disease. *Antioxid. Redox Signal.* *22*, 1463–1482.
- Himeda, C.L., Jones, T.I., and Jones, P.L. (2016a). CRISPR/dCas9-mediated Transcriptional Inhibition Ameliorates the Epigenetic Dysregulation at D4Z4 and Represses DUX4-fl in FSH Muscular Dystrophy. *Mol. Ther. J. Am. Soc. Gene Ther.* *24*, 527–535.
- Himeda, C.L., Jones, T.I., and Jones, P.L. (2016b). Scalpel or Straitjacket: CRISPR/Cas9 Approaches for Muscular Dystrophies. *Trends Pharmacol. Sci.* *37*, 249–251.
- Hirano, T. (2006). At the heart of the chromosome: SMC proteins in action. *Nat. Rev. Mol. Cell Biol.* *7*, 311–322.
- Hoeng, J.C., Höng, J.C., Ivanov, N.V., Hodor, P., Xia, M., Wei, N., Blevins, R., Gerhold, D., Borodovsky, M., and Liu, Y. (2004). Identification of new human cadherin genes using a combination of protein motif search and gene finding methods. *J. Mol. Biol.* *337*, 307–317.
- Hoggan, M.D., Blacklow, N.R., and Rowe, W.P. (1966). Studies of small DNA viruses found in various adenovirus preparations: physical, biological, and immunological characteristics. *Proc. Natl. Acad. Sci. U. S. A.* *55*, 1467–1474.
- Holland, P.W., Booth, H.A.F., and Bruford, E.A. (2007). Classification and nomenclature of all human homeobox genes. *BMC Biol.* *5*, 47.
- Holmes, N.G. (1994). Microsatellite markers and the analysis of genetic disease. *Br. Vet. J.* *150*, 411–421.
- Horike, S., Cai, S., Miyano, M., Cheng, J.-F., and Kohwi-Shigematsu, T. (2005). Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. *Nat. Genet.* *37*, 31–40.
- Hou, R., and Sibinga, N.E.S. (2009). Atrophin Proteins Interact with the Fat1 Cadherin and Regulate Migration and Orientation in Vascular Smooth Muscle Cells. *J. Biol. Chem.* *284*, 6955–6965.
- Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell* *157*, 1262–1278.
- Hughes, B.P. (1971). Creatine Phosphokinase in Facioscapulohumeral Muscular Dystrophy. *Br. Med. J.* *3*, 464–465.
- Huichalaf, C., Micheloni, S., Ferri, G., Caccia, R., and Gabellini, D. (2014). DNA methylation analysis of the macrosatellite repeat associated with FSHD muscular dystrophy at single nucleotide level. *PloS One* *9*, e115278.

- Hutvagner, G., and Zamore, P.D. (2002). RNAi: nature abhors a double-strand. *Curr. Opin. Genet. Dev.* *12*, 225–232.
- Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.-R.J., and Joung, J.K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* *31*, 227–229.
- Illingworth, R.S., and Bird, A.P. (2009). CpG islands – “A rough guide.” *FEBS Lett.* *583*, 1713–1720.
- Infante, T., Mancini, F.P., Lanza, A., Soricelli, A., de Nigris, F., and Napoli, C. (2015). Polycomb YY1 is a critical interface between epigenetic code and miRNA machinery after exposure to hypoxia in malignancy. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* *1853*, 975–986.
- Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.* *169*, 5429–5433.
- Jabbari, K., and Bernardi, G. (2004). Cytosine methylation and CpG, TpG (CpA) and TpA frequencies. *Gene* *333*, 143–149.
- Jackson, D.A. (2003). The principles of nuclear structure. *Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* *11*, 387–401.
- Jacqueline, M. (2011). The identification of Landouzy-Dejerine disease: an investigative history. *Neuromuscul. Disord.* *NMD 21*, 291–297.
- Janssen, B.H., Voet, N.B.M., Nabuurs, C.I., Kan, H.E., de Rooy, J.W.J., Geurts, A.C., Padberg, G.W., van Engelen, B.G.M., and Heerschap, A. (2014a). Distinct disease phases in muscles of facioscapulohumeral dystrophy patients identified by MR detected fat infiltration. *PloS One* *9*, e85416.
- Janssen, B.H., Pillen, S., Voet, N.B.M., Heerschap, A., van Engelen, B.G.M., and van Alfen, N. (2014b). Quantitative muscle ultrasound versus quantitative MRI in facioscapulohumeral dystrophy. *Muscle Nerve*.
- Jenal, M., Elkon, R., Loayza-Puch, F., van Haaften, G., Kühn, U., Menzies, F.M., Oude Vrielink, J.A.F., Bos, A.J., Drost, J., Rooijers, K., et al. (2012). The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. *Cell* *149*, 538–553.
- Jensen, K.K., Orum, H., Nielsen, P.E., and Nordén, B. (1997). Kinetics for hybridization of peptide nucleic acids (PNA) with DNA and RNA studied with the BIAcore technique. *Biochemistry (Mosc.)* *36*, 5072–5077.
- Jensen, M.P., Hoffman, A.J., Stoelb, B.L., Abresch, R.T., Carter, G.T., and McDonald, C.M. (2008). Chronic pain in persons with myotonic dystrophy and facioscapulohumeral dystrophy. *Arch. Phys. Med. Rehabil.* *89*, 320–328.
- Jiang, C., and Pugh, B.F. (2009). Nucleosome positioning and gene regulation: advances through genomics. *Nat. Rev. Genet.* *10*, 161–172.
- Jiang, G., Yang, F., van Overveld, P.G.M., Vedanarayanan, V., van der Maarel, S., and Ehrlich, M. (2003). Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q. *Hum. Mol. Genet.* *12*, 2909–2921.



- Jiao, S., Williams, P., Berg, R.K., Hodgeman, B.A., Liu, L., Repetto, G., and Wolff, J.A. (1992). Direct gene transfer into nonhuman primate myofibers in vivo. *Hum. Gene Ther.* 3, 21–33.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821.
- Johnson, J.S., and Samulski, R.J. (2009). Enhancement of adeno-associated virus infection by mobilizing capsids into and out of the nucleolus. *J. Virol.* 83, 2632–2644.
- Jones, P.L., Veenstra, G.J., Wade, P.A., Vermaak, D., Kass, S.U., Landsberger, N., Strouboulis, J., and Wolffe, A.P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat. Genet.* 19, 187–191.
- Jones, T.I., Chen, J.C.J., Rahimov, F., Homma, S., Arashiro, P., Beermann, M.L., King, O.D., Miller, J.B., Kunkel, L.M., Emerson, C.P., Jr, et al. (2012). Facioscapulohumeral muscular dystrophy family studies of DUX4 expression: evidence for disease modifiers and a quantitative model of pathogenesis. *Hum. Mol. Genet.* 21, 4419–4430.
- Jones, T.I., Yan, C., Sapp, P.C., McKenna-Yasek, D., Kang, P.B., Quinn, C., Salameh, J.S., King, O.D., and Jones, P.L. (2014). Identifying diagnostic DNA methylation profiles for facioscapulohumeral muscular dystrophy in blood and saliva using bisulfite sequencing. *Clin. Epigenetics* 6, 23.
- Jones, T.I., King, O.D., Himeda, C.L., Homma, S., Chen, J.C.J., Beermann, M.L., Yan, C., Emerson, C.P., Miller, J.B., Wagner, K.R., et al. (2015). Individual epigenetic status of the pathogenic D4Z4 macrosatellite correlates with disease in facioscapulohumeral muscular dystrophy. *Clin. Epigenetics* 7, 37.
- Jones, T.I., Parilla, M., and Jones, P.L. (2016). Transgenic *Drosophila* for Investigating DUX4 and FRG1, Two Genes Associated with Facioscapulohumeral Muscular Dystrophy (FSHD). *PloS One* 11, e0150938.
- Joung, J.K., and Sander, J.D. (2013). TALENs: a widely applicable technology for targeted genome editing. *Nat. Rev. Mol. Cell Biol.* 14, 49–55.
- Junquera, E., and Aicart, E. (2014). Cationic lipids as transfecting agents of DNA in gene therapy. *Curr. Top. Med. Chem.* 14, 649–663.
- Jurica, M.S., Licklider, L.J., Gygi, S.R., Grigorieff, N., and Moore, M.J. (2002). Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis. *RNA N. Y.* N 8, 426–439.
- Kaczmarek, A., Schneider, S., Wirth, B., and Riessland, M. (2015). Investigational therapies for the treatment of spinal muscular atrophy. *Expert Opin. Investig. Drugs* 24, 867–881.
- Kalkman, J., Schillings, M., van der Werf, S.P., Padberg, G., Zwarts, M., van Engelen, B.G.M., and Bleijenberg, G. (2005). Experienced fatigue in facioscapulohumeral dystrophy, myotonic dystrophy, and HMSN-I. *J. Neurol. Neurosurg. Psychiatry* 76, 1406–1409.
- Kalkman, J.S., Schillings, M.L., Zwarts, M.J., van Engelen, B.G.M., and Bleijenberg, G. (2007). The development of a model of fatigue in neuromuscular disorders: a longitudinal study. *J. Psychosom. Res.* 62, 571–579.
- Kalkman, J.S., Zwarts, M.J., Schillings, M.L., van Engelen, B.G.M., and Bleijenberg, G. (2008). Different types of fatigue in patients with facioscapulohumeral dystrophy, myotonic dystrophy and

- HMSN-I. Experienced fatigue and physiological fatigue. *Neurol. Sci. Off. J. Ital. Neurol. Soc. Ital. Soc. Clin. Neurophysiol.* 29 *Suppl* 2, S238-240.
- Kalverda, B., Röling, M.D., and Fornerod, M. (2008). Chromatin organization in relation to the nuclear periphery. *FEBS Lett.* 582, 2017–2022.
- Kanduri, C., Pant, V., Loukinov, D., Pugacheva, E., Qi, C.F., Wolffe, A., Ohlsson, R., and Lobanenkov, V.V. (2000). Functional association of CTCF with the insulator upstream of the H19 gene is parent of origin-specific and methylation-sensitive. *Curr. Biol. CB* 10, 853–856.
- Kaneda, Y. (2001). Gene therapy: a battle against biological barriers. *Curr. Mol. Med.* 1, 493–499.
- Käs, E., and Chasin, L.A. (1987). Anchorage of the Chinese hamster dihydrofolate reductase gene to the nuclear scaffold occurs in an intragenic region. *J. Mol. Biol.* 198, 677–692.
- Kashiwakura, Y., Tamayose, K., Iwabuchi, K., Hirai, Y., Shimada, T., Matsumoto, K., Nakamura, T., Watanabe, M., Oshimi, K., and Daida, H. (2005). Hepatocyte growth factor receptor is a coreceptor for adeno-associated virus type 2 infection. *J. Virol.* 79, 609–614.
- Katoh, Y., and Katoh, M. (2006). Comparative integromics on FAT1, FAT2, FAT3 and FAT4. *Int. J. Mol. Med.* 18, 523–528.
- Kawabata, K., Takakura, Y., and Hashida, M. (1995). The fate of plasmid DNA after intravenous injection in mice: involvement of scavenger receptors in its hepatic uptake. *Pharm. Res.* 12, 825–830.
- Kawamura-Saito, M., Yamazaki, Y., Kaneko, K., Kawaguchi, N., Kanda, H., Mukai, H., Gotoh, T., Motoi, T., Fukayama, M., Aburatani, H., et al. (2006). Fusion between CIC and DUX4 up-regulates PEA3 family genes in Ewing-like sarcomas with t(4;19)(q35;q13) translocation. *Hum. Mol. Genet.* 15, 2125–2137.
- Kearns, N.A., Pham, H., Tabak, B., Genga, R.M., Silverstein, N.J., Garber, M., and Maehr, R. (2015). Functional annotation of native enhancers with a Cas9-histone demethylase fusion. *Nat. Methods* 12, 401–403.
- Keaton, M.A., Taylor, C.M., Layer, R.M., and Dutta, A. (2011). Nuclear Scaffold Attachment Sites within ENCODE Regions Associate with Actively Transcribed Genes. *PLoS ONE* 6.
- Keller, W., Bienroth, S., Lang, K.M., and Christofori, G. (1991). Cleavage and polyadenylation factor CPF specifically interacts with the pre-mRNA 3' processing signal AAUAAA. *EMBO J.* 10, 4241–4249.
- Kemler, R. (1993). From cadherins to catenins: cytoplasmic protein interactions and regulation of cell adhesion. *Trends Genet. TIG* 9, 317–321.
- van Kernebeek, C.R., Seijger, C.G.W., Goselink, R.J.M., Heijdra, Y.F., and van Engelen, B.G.M. (2017). Respiratory pattern in a FSHD pediatric population. *Respir. Med.*
- Kervestin, S., and Jacobson, A. (2012). NMD: a multifaceted response to premature translational termination. *Nat. Rev. Mol. Cell Biol.* 13, 700–712.
- Kerwitz, Y., Kühn, U., Lilie, H., Knoth, A., Scheuermann, T., Friedrich, H., Schwarz, E., and Wahle, E. (2003). Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA. *EMBO J.* 22, 3705–3714.
- Kim, J.S., and Pabo, C.O. (1998). Getting a handhold on DNA: design of poly-zinc finger proteins with femtomolar dissociation constants. *Proc. Natl. Acad. Sci. U. S. A.* 95, 2812–2817.

- Kim, J.K., Samaranyake, M., and Pradhan, S. (2009a). Epigenetic mechanisms in mammals. *Cell. Mol. Life Sci. CMLS* 66, 596–612.
- Kim, V.N., Han, J., and Siomi, M.C. (2009b). Biogenesis of small RNAs in animals. *Nat. Rev. Mol. Cell Biol.* 10, 126–139.
- Kim, Y.G., Cha, J., and Chandrasegaran, S. (1996). Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. U. S. A.* 93, 1156–1160.
- Kislauskis, E.H., and Singer, R.H. (1992). Determinants of mRNA localization. *Curr. Opin. Cell Biol.* 4, 975–978.
- Kissel, J.T., McDermott, M.P., Natarajan, R., Mendell, J.R., Pandya, S., King, W.M., Griggs, R.C., and Tawil, R. (1998). Pilot trial of albuterol in facioscapulohumeral muscular dystrophy. FSH-DY Group. *Neurology* 50, 1402–1406.
- Kissel, J.T., McDermott, M.P., Mendell, J.R., King, W.M., Pandya, S., Griggs, R.C., Tawil, R., and FSH-DY Group (2001). Randomized, double-blind, placebo-controlled trial of albuterol in facioscapulohumeral dystrophy. *Neurology* 57, 1434–1440.
- Kisseljova, N.P., Dmitriev, P., Katargin, A., Kim, E., Ezerina, D., Markozashvili, D., Malysheva, D., Planche, E., Lemmers, R.J.L.F., van der Maarel, S.M., et al. (2014). DNA polymorphism and epigenetic marks modulate the affinity of a scaffold/matrix attachment region to the nuclear matrix. *Eur. J. Hum. Genet. EJHG*.
- Klingenberg, M. (2008). The ADP and ATP transport in mitochondria and its carrier. *Biochim. Biophys. Acta* 1778, 1978–2021.
- Klooster, R., Straasheijm, K., Shah, B., Sowden, J., Frants, R., Thornton, C., Tawil, R., and van der Maarel, S. (2009). Comprehensive expression analysis of FSHD candidate genes at the mRNA and protein level. *Eur. J. Hum. Genet. EJHG* 17, 1615–1624.
- Klose, R.J., Sarraf, S.A., Schmiedeberg, L., McDermott, S.M., Stancheva, I., and Bird, A.P. (2005). DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG. *Mol. Cell* 19, 667–678.
- Koenig, M., Beggs, A.H., Moyer, M., Scherpf, S., Heindrich, K., Bettecken, T., Meng, G., Müller, C.R., Lindlöf, M., Kaariainen, H., et al. (1989). The molecular basis for Duchenne versus Becker muscular dystrophy: correlation of severity with type of deletion. *Am. J. Hum. Genet.* 45, 498–506.
- Koering, C.E., Pollice, A., Zibella, M.P., Bauwens, S., Puisieux, A., Brunori, M., Brun, C., Martins, L., Sabatier, L., Pulitzer, J.F., et al. (2002). Human telomeric position effect is determined by chromosomal context and telomeric chromatin integrity. *EMBO Rep.* 3, 1055–1061.
- Koetsier, C. (1997). Pain related to FSH dystrophy: an underestimated problem? Results of an inquiry in the Netherlands, Baarn VSN. *FSH Watch Publ. FSH Soc. USA* 23–24.
- Kohwi-Shigematsu, T., and Kohwi, Y. (1990). Torsional stress stabilizes extended base unpairing in suppressor sites flanking immunoglobulin heavy chain enhancer. *Biochemistry (Mosc.)* 29, 9551–9560.
- Kole, R., Krainer, A.R., and Altman, S. (2012). RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nat. Rev. Drug Discov.* 11, 125–140.
- Kolev, N.G., and Steitz, J.A. (2006). In vivo assembly of functional U7 snRNP requires RNA backbone flexibility within the Sm-binding site. *Nat. Struct. Mol. Biol.* 13, 347–353.

van Koningsbruggen, S., Dirks, R.W., Mommaas, A.M., Onderwater, J.J., Deidda, G., Padberg, G.W., Frants, R.R., and van der Maarel, S.M. (2004). FRG1P is localised in the nucleolus, Cajal bodies, and speckles. *J. Med. Genet.* *41*, e46.

van Koningsbruggen, S., Straasheijm, K.R., Sterrenburg, E., de Graaf, N., Dauwerse, H.G., Frants, R.R., and van der Maarel, S.M. (2007). FRG1P-mediated aggregation of proteins involved in pre-mRNA processing. *Chromosoma* *116*, 53–64.

van der Kooi, E.L., Vogels, O.J.M., van Asseldonk, R.J.G.P., Lindeman, E., Hendriks, J.C.M., Wohlgenuth, M., van der Maarel, S.M., and Padberg, G.W. (2004). Strength training and albuterol in facioscapulohumeral muscular dystrophy. *Neurology* *63*, 702–708.

van der Kooi, E.L., Kalkman, J.S., Lindeman, E., Hendriks, J.C.M., van Engelen, B.G.M., Bleijenberg, G., and Padberg, G.W. (2007). Effects of training and albuterol on pain and fatigue in facioscapulohumeral muscular dystrophy. *J. Neurol.* *254*, 931–940.

Kornberg, R.D., and Lorch, Y. (1999). Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell* *98*, 285–294.

Kotin, R.M., Siniscalco, M., Samulski, R.J., Zhu, X.D., Hunter, L., Laughlin, C.A., McLaughlin, S., Muzyczka, N., Rocchi, M., and Berns, K.I. (1990). Site-specific integration by adeno-associated virus. *Proc. Natl. Acad. Sci. U. S. A.* *87*, 2211–2215.

Kowaljow, V., Marcowycz, A., Anseau, E., Conde, C.B., Sauvage, S., Mattéotti, C., Arias, C., Corona, E.D., Nuñez, N.G., Leo, O., et al. (2007). The DUX4 gene at the FSHD1A locus encodes a pro-apoptotic protein. *Neuromuscul. Disord. NMD* *17*, 611–623.

Krol, J., Loedige, I., and Filipowicz, W. (2010). The widespread regulation of microRNA biogenesis, function and decay. *Nat. Rev. Genet.* *11*, 597–610.

Krom, Y.D., Dumonceaux, J., Mamchaoui, K., den Hamer, B., Mariot, V., Negroni, E., Geng, L.N., Martin, N., Tawil, R., Tapscott, S.J., et al. (2012). Generation of isogenic D4Z4 contracted and noncontracted immortal muscle cell clones from a mosaic patient: a cellular model for FSHD. *Am. J. Pathol.* *181*, 1387–1401.

Krom, Y.D., Thijssen, P.E., Young, J.M., den Hamer, B., Balog, J., Yao, Z., Maves, L., Snider, L., Knopp, P., Zammit, P.S., et al. (2013). Intrinsic epigenetic regulation of the D4Z4 macrosatellite repeat in a transgenic mouse model for FSHD. *PLoS Genet.* *9*, e1003415.

Kuang, S., Kuroda, K., Le Grand, F., and Rudnicki, M.A. (2007). Asymmetric Self-Renewal and Commitment of Satellite Stem Cells in Muscle. *Cell* *129*, 999–1010.

Kühn, U., Gündel, M., Knoth, A., Kerwitz, Y., Rüdell, S., and Wahle, E. (2009). Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor. *J. Biol. Chem.* *284*, 22803–22814.

Kureishy, N., Sapountzi, V., Prag, S., Anilkumar, N., and Adams, J.C. (2002). Fascins, and their roles in cell structure and function. *BioEssays News Rev. Mol. Cell. Dev. Biol.* *24*, 350–361.

Kuroyanagi, H. (2009). Fox-1 family of RNA-binding proteins. *Cell. Mol. Life Sci. CMLS* *66*, 3895–3907.

Kurz, A., Lampel, S., Nickolenko, J.E., Bradl, J., Benner, A., Zirbel, R.M., Cremer, T., and Lichter, P. (1996). Active and inactive genes localize preferentially in the periphery of chromosome territories. *J. Cell Biol.* *135*, 1195–1205.

- Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* *410*, 116–120.
- Laforêt, P., de Toma, C., Eymard, B., Becane, H.M., Jeanpierre, M., Fardeau, M., and Duboc, D. (1998). Cardiac involvement in genetically confirmed facioscapulohumeral muscular dystrophy. *Neurology* *51*, 1454–1456.
- Lagirand-Cantaloube, J., Cornille, K., Csibi, A., Batonnet-Pichon, S., Leibovitch, M.P., and Leibovitch, S.A. (2009). Inhibition of Atrogin-1/MAFbx Mediated MyoD Proteolysis Prevents Skeletal Muscle Atrophy In Vivo. *PLOS ONE* *4*, e4973.
- Lanctôt, C., Moreau, A., Chamberland, M., Tremblay, M.L., and Drouin, J. (1999). Hindlimb patterning and mandible development require the Ptx1 gene. *Dev. Camb. Engl.* *126*, 1805–1810.
- Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. *Nature* *409*, 860–921.
- Landouzy, and Dejerine (1885). De la myopathie atrophique progressive. *Rev. Médicale Fr.* *5*:81.
- de Lange, T. (2004). T-loops and the origin of telomeres. *Nat. Rev. Mol. Cell Biol.* *5*, 323–329.
- Laoudj-Chenivesse, D., Carnac, G., Bisbal, C., Hugon, G., Bouillot, S., Desnuelle, C., Vassetzky, Y., and Fernandez, A. (2005). Increased levels of adenine nucleotide translocator 1 protein and response to oxidative stress are early events in facioscapulohumeral muscular dystrophy muscle. *J. Mol. Med. Berl. Ger.* *83*, 216–224.
- Larsen, F., Gundersen, G., Lopez, R., and Prydz, H. (1992). CpG islands as gene markers in the human genome. *Genomics* *13*, 1095–1107.
- Larsen, M., Rost, S., El Hajj, N., Ferbert, A., Deschauer, M., Walter, M.C., Schoser, B., Tacik, P., Kress, W., and Müller, C.R. (2015). Diagnostic approach for FSHD revisited: SMCHD1 mutations cause FSHD2 and act as modifiers of disease severity in FSHD1. *Eur. J. Hum. Genet. EJHG* *23*, 808–816.
- Lebofsky, R., and Bensimon, A. (2003). Single DNA molecule analysis: applications of molecular combing. *Brief. Funct. Genomic. Proteomic.* *1*, 385–396.
- Lecker, S.H., Solomon, V., Mitch, W.E., and Goldberg, A.L. (1999). Muscle protein breakdown and the critical role of the ubiquitin-proteasome pathway in normal and disease states. *J. Nutr.* *129*, 227S–237S.
- Lee, G.D., Chen, V.M., Barnes, A.C., Goldman, D.R., and Duker, J.S. (2014). Retinal telangiectasis detected during a vision screening examination in a child with hearing loss led to the diagnosis of facioscapulohumeral muscular dystrophy. *J. AAPOS Off. Publ. Am. Assoc. Pediatr. Ophthalmol. Strabismus Am. Assoc. Pediatr. Ophthalmol. Strabismus* *18*, 303–305.
- Lee, J.H., Goto, K., Sahashi, K., Nonaka, I., Matsuda, C., and Arahata, K. (1995). Cloning and mapping of a very short (10-kb) EcoRI fragment associated with facioscapulohumeral muscular dystrophy (FSHD). *Muscle Nerve. Suppl.* *2*, S27-31.
- Lee, M.S., Gippert, G.P., Soman, K.V., Case, D.A., and Wright, P.E. (1989). Three-dimensional solution structure of a single zinc finger DNA-binding domain. *Science* *245*, 635–637.
- Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* *75*, 843–854.

- Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Rådmark, O., Kim, S., et al. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature* *425*, 415–419.
- Lee, Y., Kim, M., Han, J., Yeom, K.-H., Lee, S., Baek, S.H., and Kim, V.N. (2004). MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* *23*, 4051–4060.
- Lehming, N., Le Saux, A., Schüller, J., and Ptashne, M. (1998). Chromatin components as part of a putative transcriptional repressing complex. *Proc. Natl. Acad. Sci. U. S. A.* *95*, 7322–7326.
- Lehnertz, B., Ueda, Y., Derijck, A.A.H.A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A.H.F.M. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr. Biol. CB* *13*, 1192–1200.
- Leidenroth, A., and Hewitt, J.E. (2010). A family history of DUX4: phylogenetic analysis of DUXA, B, C and Duxbl reveals the ancestral DUX gene. *BMC Evol. Biol.* *10*, 364.
- Leidenroth, A., Clapp, J., Mitchell, L.M., Coneyworth, D., Dearden, F.L., Iannuzzi, L., and Hewitt, J.E. (2012). Evolution of DUX gene macrosatellites in placental mammals. *Chromosoma* *121*, 489–497.
- Lek, A., Rahimov, F., Jones, P.L., and Kunkel, L.M. (2015). Emerging preclinical animal models for FSHD. *Trends Mol. Med.* *21*, 295–306.
- Lemmers, R.J.L.F. (2017). Analyzing Copy Number Variation Using Pulsed-Field Gel Electrophoresis: Providing a Genetic Diagnosis for FSHD1. *Methods Mol. Biol. Clifton NJ* *1492*, 107–125.
- Lemmers, R.J., van der Maarel, S.M., van Deutekom, J.C., van der Wielen, M.J., Deidda, G., Dauwerse, H.G., Hewitt, J., Hofker, M., Bakker, E., Padberg, G.W., et al. (1998). Inter- and intrachromosomal sub-telomeric rearrangements on 4q35: implications for facioscapulohumeral muscular dystrophy (FSHD) aetiology and diagnosis. *Hum. Mol. Genet.* *7*, 1207–1214.
- Lemmers, R.J.F.L., Wohlgemuth, M., Frants, R.R., Padberg, G.W., Morava, E., and van der Maarel, S.M. (2004a). Contractions of D4Z4 on 4qB subtelomeres do not cause facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.* *75*, 1124–1130.
- Lemmers, R.J.L.F., de Kievit, P., Sandkuijl, L., Padberg, G.W., van Ommen, G.-J.B., Frants, R.R., and van der Maarel, S.M. (2002). Facioscapulohumeral muscular dystrophy is uniquely associated with one of the two variants of the 4q subtelomere. *Nat. Genet.* *32*, 235–236.
- Lemmers, R.J.L.F., Osborn, M., Haaf, T., Rogers, M., Frants, R.R., Padberg, G.W., Cooper, D.N., van der Maarel, S.M., and Upadhyaya, M. (2003). D4F104S1 deletion in facioscapulohumeral muscular dystrophy: phenotype, size, and detection. *Neurology* *61*, 178–183.
- Lemmers, R.J.L.F., Van Overveld, P.G.M., Sandkuijl, L.A., Vrieling, H., Padberg, G.W., Frants, R.R., and van der Maarel, S.M. (2004b). Mechanism and timing of mitotic rearrangements in the subtelomeric D4Z4 repeat involved in facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.* *75*, 44–53.
- Lemmers, R.J.L.F., van der Wielen, M.J.R., Bakker, E., Padberg, G.W., Frants, R.R., and van der Maarel, S.M. (2004c). Somatic mosaicism in FSHD often goes undetected. *Ann. Neurol.* *55*, 845–850.
- Lemmers, R.J.L.F., Wohlgemuth, M., van der Gaag, K.J., van der Vliet, P.J., van Teijlingen, C.M.M., de Knijff, P., Padberg, G.W., Frants, R.R., and van der Maarel, S.M. (2007). Specific sequence

variations within the 4q35 region are associated with facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.* 81, 884–894.

Lemmers, R.J.L.F., van der Vliet, P.J., Klooster, R., Sacconi, S., Camaño, P., Dauwerse, J.G., Snider, L., Straasheijm, K.R., van Ommen, G.J., Padberg, G.W., et al. (2010a). A unifying genetic model for facioscapulohumeral muscular dystrophy. *Science* 329, 1650–1653.

Lemmers, R.J.L.F., van der Vliet, P.J., van der Gaag, K.J., Zuniga, S., Frants, R.R., de Knijff, P., and van der Maarel, S.M. (2010b). Worldwide population analysis of the 4q and 10q subtelomeres identifies only four discrete interchromosomal sequence transfers in human evolution. *Am. J. Hum. Genet.* 86, 364–377.

Lemmers, R.J.L.F., Tawil, R., Petek, L.M., Balog, J., Block, G.J., Santen, G.W.E., Amell, A.M., van der Vliet, P.J., Almomani, R., Straasheijm, K.R., et al. (2012a). Digenic inheritance of an SMCHD1 mutation and an FSHD-permissive D4Z4 allele causes facioscapulohumeral muscular dystrophy type 2. *Nat. Genet.* 44, 1370–1374.

Lemmers, R.J.L.F., O’Shea, S., Padberg, G.W., Lunt, P.W., and van der Maarel, S.M. (2012b). Best practice guidelines on genetic diagnostics of Facioscapulohumeral muscular dystrophy: workshop 9th June 2010, LUMC, Leiden, The Netherlands. *Neuromuscul. Disord.* NMD 22, 463–470.

Lemmers, R.J.L.F., Goeman, J.J., van der Vliet, P.J., van Nieuwenhuizen, M.P., Balog, J., Vos-Versteeg, M., Camano, P., Ramos Arroyo, M.A., Jerico, I., Rogers, M.T., et al. (2015a). Inter-individual differences in CpG methylation at D4Z4 correlate with clinical variability in FSHD1 and FSHD2. *Hum. Mol. Genet.* 24, 659–669.

Lemmers, R.J.L.F., van den Boogaard, M.L., van der Vliet, P.J., Donlin-Smith, C.M., Nations, S.P., Ruivenkamp, C.A.L., Heard, P., Bakker, B., Tapscott, S., Cody, J.D., et al. (2015b). Hemizygoty for SMCHD1 in Facioscapulohumeral Muscular Dystrophy Type 2: Consequences for 18p Deletion Syndrome. *Hum. Mutat.*

LeRoy, G., Weston, J.T., Zee, B.M., Young, N.L., Plazas-Mayorca, M.D., and Garcia, B.A. (2009). Heterochromatin Protein 1 Is Extensively Decorated with Histone Code-like Post-translational Modifications. *Mol. Cell. Proteomics MCP* 8, 2432–2442.

Leung, D.G., Carrino, J.A., Wagner, K.R., and Jacobs, M.A. (2015). Whole-body magnetic resonance imaging evaluation of facioscapulohumeral muscular dystrophy. *Muscle Nerve.*

Levy, S., Sutton, G., Ng, P.C., Feuk, L., Halpern, A.L., Walenz, B.P., Axelrod, N., Huang, J., Kirkness, E.F., Denisov, G., et al. (2007). The diploid genome sequence of an individual human. *PLoS Biol.* 5, e254.

Li, E. (2002). Chromatin modification and epigenetic reprogramming in mammalian development. *Nat. Rev. Genet.* 3, 662–673.

Li, S. (2004). Electroporation gene therapy: new developments in vivo and in vitro. *Curr. Gene Ther.* 4, 309–316.

Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69, 915–926.

Li, E., Beard, C., and Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. *Nature* 366, 362–365.

- Li, H., Haurigot, V., Doyon, Y., Li, T., Wong, S.Y., Bhagwat, A.S., Malani, N., Anguela, X.M., Sharma, R., Ivanciu, L., et al. (2011a). In vivo genome editing restores haemostasis in a mouse model of haemophilia. *Nature* 475, 217–221.
- Li, K., Warner, C.K., Hodge, J.A., Minoshima, S., Kudoh, J., Fukuyama, R., Maekawa, M., Shimizu, Y., Shimizu, N., and Wallace, D.C. (1989). A human muscle adenine nucleotide translocator gene has four exons, is located on chromosome 4, and is differentially expressed. *J. Biol. Chem.* 264, 13998–14004.
- Li, T., Huang, S., Zhao, X., Wright, D.A., Carpenter, S., Spalding, M.H., Weeks, D.P., and Yang, B. (2011b). Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. *Nucleic Acids Res.* 39, 6315–6325.
- Li, Y., Danzer, J.R., Alvarez, P., Belmont, A.S., and Wallrath, L.L. (2003). Effects of tethering HP1 to euchromatic regions of the *Drosophila* genome. *Dev. Camb. Engl.* 130, 1817–1824.
- Li, Y.-C., Korol, A.B., Fahima, T., Beiles, A., and Nevo, E. (2002). Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Mol. Ecol.* 11, 2453–2465.
- Li, Y.S., Davidson, E., Reid, C.N., and McHale, A.P. (2009). Optimising ultrasound-mediated gene transfer (sonoporation) in vitro and prolonged expression of a transgene in vivo: potential applications for gene therapy of cancer. *Cancer Lett.* 273, 62–69.
- Liang, G., Chan, M.F., Tomigahara, Y., Tsai, Y.C., Gonzales, F.A., Li, E., Laird, P.W., and Jones, P.A. (2002). Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements. *Mol. Cell. Biol.* 22, 480–491.
- Lieber, M.R., Ma, Y., Pannicke, U., and Schwarz, K. (2003). Mechanism and regulation of human non-homologous DNA end-joining. *Nat. Rev. Mol. Cell Biol.* 4, 712–720.
- Liebich, I., Bode, J., Reuter, I., and Wingender, E. (2002). Evaluation of sequence motifs found in scaffold/matrix-attached regions (S/MARs). *Nucleic Acids Res.* 30, 3433–3442.
- Lim, J.-W., Snider, L., Yao, Z., Tawil, R., Van Der Maarel, S.M., Rigo, F., Bennett, C.F., Filippova, G.N., and Tapscott, S.J. (2015). DICER/AGO-dependent epigenetic silencing of D4Z4 repeats enhanced by exogenous siRNA suggests mechanisms and therapies for FSHD. *Hum. Mol. Genet.* 24, 4817–4828.
- Liu, D.X., and Lobie, P.E. (2007). Transcriptional activation of p53 by Pitx1. *Cell Death Differ.* 14, 1893–1907.
- Liu, Y., and Chen, X.J. (2013). Adenine Nucleotide Translocase, Mitochondrial Stress, and Degenerative Cell Death. *Oxid. Med. Cell. Longev.* 2013.
- Liu, Q., Segal, D.J., Ghiara, J.B., and Barbas, C.F. (1997). Design of polydactyl zinc-finger proteins for unique addressing within complex genomes. *Proc. Natl. Acad. Sci. U. S. A.* 94, 5525–5530.
- Liu, Q., Jones, T.I., Tang, V.W., Briehner, W.M., and Jones, P.L. (2010). Facioscapulohumeral muscular dystrophy region gene-1 (FRG-1) is an actin-bundling protein associated with muscle-attachment sites. *J. Cell Sci.* 123, 1116–1123.
- Liu, S., Asparuhova, M., Brondani, V., Ziekau, I., Klimkait, T., and Schümperli, D. (2004). Inhibition of HIV-1 multiplication by antisense U7 snRNAs and siRNAs targeting cyclophilin A. *Nucleic Acids Res.* 32, 3752–3759.



- Lockhart-Clarke (1867). Case of musculaire atrophie with disease of the spinal cord, and medulla oblongata. *Medico-Chir. Trans.*
- Logan, M., and Tabin, C.J. (1999). Role of Pitx1 upstream of Tbx4 in specification of hindlimb identity. *Science* 283, 1736–1739.
- Lomvardas, S., Barnea, G., Pisapia, D.J., Mendelsohn, M., Kirkland, J., and Axel, R. (2006). Interchromosomal interactions and olfactory receptor choice. *Cell* 126, 403–413.
- Lonez, C., Vandenbranden, M., and Ruyschaert, J.-M. (2008). Cationic liposomal lipids: from gene carriers to cell signaling. *Prog. Lipid Res.* 47, 340–347.
- Long, C., McAnally, J.R., Shelton, J.M., Mireault, A.A., Bassel-Duby, R., and Olson, E.N. (2014). Prevention of muscular dystrophy in mice by CRISPR/Cas9-mediated editing of germline DNA. *Science* 345, 1184–1188.
- Long, C., Amoasii, L., Mireault, A.A., McAnally, J.R., Li, H., Sanchez-Ortiz, E., Bhattacharyya, S., Shelton, J.M., Bassel-Duby, R., and Olson, E.N. (2016). Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science* 351, 400–403.
- Lou, Z., Wei, J., Riethman, H., Baur, J.A., Voglauer, R., Shay, J.W., and Wright, W.E. (2009). Telomere length regulates ISG15 expression in human cells. *Aging* 1, 608–621.
- Luco, R.F., Pan, Q., Tominaga, K., Blencowe, B.J., Pereira-Smith, O.M., and Misteli, T. (2010). Regulation of alternative splicing by histone modifications. *Science* 327, 996–1000.
- Luco, R.F., Allo, M., Schor, I.E., Kornblihtt, A.R., and Misteli, T. (2011). Epigenetics in alternative pre-mRNA splicing. *Cell* 144, 16–26.
- Lucotte, G., Berriche, S., and Fardeau, M. (1989). Linkage analysis of French families with facioscapulohumeral muscular dystrophy. *J. Med. Genet.* 26, 485–486.
- Lukasik, S.M., Cierpicki, T., Borloz, M., Grembecka, J., Everett, A., and Bushweller, J.H. (2006). High resolution structure of the HDGF PWWP domain: a potential DNA binding domain. *Protein Sci. Publ. Protein Soc.* 15, 314–323.
- Lunt, P.W., and Harper, P.S. (1991). Genetic counselling in facioscapulohumeral muscular dystrophy. *J. Med. Genet.* 28, 655–664.
- Lunt, P.W., Noades, J.G., Upadhyaya, M., Sarfarazi, M., and Harper, P.S. (1988). Evidence against location of the gene for facioscapulohumeral muscular dystrophy on the distal long arm of chromosome 14. *J. Neurol. Sci.* 88, 287–292.
- Lunt, P.W., Compston, D.A., and Harper, P.S. (1989). Estimation of age dependent penetrance in facioscapulohumeral muscular dystrophy by minimising ascertainment bias. *J. Med. Genet.* 26, 755–760.
- Lunt, P.W., Jardine, P.E., Koch, M., Maynard, J., Osborn, M., Williams, M., Harper, P.S., and Upadhyaya, M. (1995a). Phenotypic-genotypic correlation will assist genetic counseling in 4q35-facioscapulohumeral muscular dystrophy. *Muscle Nerve. Suppl.* 2, S103-109.
- Lunt, P.W., Jardine, P.E., Koch, M.C., Maynard, J., Osborn, M., Williams, M., Harper, P.S., and Upadhyaya, M. (1995b). Correlation between fragment size at D4F104S1 and age at onset or at wheelchair use, with a possible generational effect, accounts for much phenotypic variation in 4q35-facioscapulohumeral muscular dystrophy (FSHD). *Hum. Mol. Genet.* 4, 951–958.

- Lutz, C.S., and Moreira, A. (2011). Alternative mRNA polyadenylation in eukaryotes: an effective regulator of gene expression. *Wiley Interdiscip. Rev. RNA* 2, 23–31.
- Lutz, K.L., Holte, L., Kliethermes, S.A., Stephan, C., and Mathews, K.D. (2013). Clinical and genetic features of hearing loss in facioscapulohumeral muscular dystrophy. *Neurology* 81, 1374–1377.
- Lyle, R., Wright, T.J., Clark, L.N., and Hewitt, J.E. (1995). The FSHD-associated repeat, D4Z4, is a member of a dispersed family of homeobox-containing repeats, subsets of which are clustered on the short arms of the acrocentric chromosomes. *Genomics* 28, 389–397.
- Ma, W., Lin, Y., Xuan, W., Iversen, P.L., Smith, L.J., and Benchimol, S. (2012). Inhibition of p53 expression by peptide-conjugated phosphorodiamidate morpholino oligomers sensitizes human cancer cells to chemotherapeutic drugs. *Oncogene* 31, 1024–1033.
- van der Maarel, S.M., Deidda, G., Lemmers, R.J., van Overveld, P.G., van der Wielen, M., Hewitt, J.E., Sandkuijl, L., Bakker, B., van Ommen, G.J., Padberg, G.W., et al. (2000). De novo facioscapulohumeral muscular dystrophy: frequent somatic mosaicism, sex-dependent phenotype, and the role of mitotic transchromosomal repeat interaction between chromosomes 4 and 10. *Am. J. Hum. Genet.* 66, 26–35.
- Macaione, V., Aguenouz, M., Rodolico, C., Mazzeo, A., Patti, A., Cannistraci, E., Colantone, L., Di Giorgio, R.M., De Luca, G., and Vita, G. (2007). RAGE-NF-kappaB pathway activation in response to oxidative stress in facioscapulohumeral muscular dystrophy. *Acta Neurol. Scand.* 115, 115–121.
- MacDonald, C.C., Wilusz, J., and Shenk, T. (1994). The 64-kilodalton subunit of the CstF polyadenylation factor binds to pre-mRNAs downstream of the cleavage site and influences cleavage site location. *Mol. Cell. Biol.* 14, 6647–6654.
- Mahfouz, M.M., Li, L., Shamimuzzaman, M., Wibowo, A., Fang, X., and Zhu, J.-K. (2011). De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. *Proc. Natl. Acad. Sci. U. S. A.* 108, 2623–2628.
- Mahoney, P.A., Weber, U., Onofrechuk, P., Biessmann, H., Bryant, P.J., and Goodman, C.S. (1991). The fat tumor suppressor gene in *Drosophila* encodes a novel member of the cadherin gene superfamily. *Cell* 67, 853–868.
- Mak, A.N.-S., Bradley, P., Cernadas, R.A., Bogdanove, A.J., and Stoddard, B.L. (2012). The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science* 335, 716–719.
- Makarov, E.M., Makarova, O.V., Urlaub, H., Gentzel, M., Will, C.L., Wilm, M., and Lührmann, R. (2002). Small nuclear ribonucleoprotein remodeling during catalytic activation of the spliceosome. *Science* 298, 2205–2208.
- Malhas, A., Lee, C.F., Sanders, R., Saunders, N.J., and Vaux, D.J. (2007). Defects in lamin B1 expression or processing affect interphase chromosome position and gene expression. *J. Cell Biol.* 176, 593–603.
- Mamchaoui, K., Trollet, C., Bigot, A., Negroni, E., Chaouch, S., Wolff, A., Kandalla, P.K., Marie, S., Di Santo, J., St Guily, J.L., et al. (2011). Immortalized pathological human myoblasts: towards a universal tool for the study of neuromuscular disorders. *Skelet. Muscle* 1, 34.
- Mandegar, M.A., Huebsch, N., Frolov, E.B., Shin, E., Truong, A., Olvera, M.P., Chan, A.H., Miyaoka, Y., Holmes, K., Spencer, C.I., et al. (2016). CRISPR Interference Efficiently Induces Specific and Reversible Gene Silencing in Human iPSCs. *Cell Stem Cell* 18, 541–553.

- Mandel, C.R., Bai, Y., and Tong, L. (2008). Protein factors in pre-mRNA 3'-end processing. *Cell. Mol. Life Sci. CMLS* 65, 1099–1122.
- Mangus, D.A., Evans, M.C., and Jacobson, A. (2003). Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol.* 4, 223.
- Manning, W.C., Zhou, S., Bland, M.P., Escobedo, J.A., and Dwarki, V. (1998). Transient immunosuppression allows transgene expression following readministration of adeno-associated viral vectors. *Hum. Gene Ther.* 9, 477–485.
- Manservigi, R., Argnani, R., and Marconi, P. (2010). HSV Recombinant Vectors for Gene Therapy. *Open Virol. J.* 4, 123–156.
- Marcil, A., Dumontier, É., Chamberland, M., Camper, S.A., and Drouin, J. (2003). Pitx1 and Pitx2 are required for development of hindlimb buds. *Development* 130, 45–55.
- Margueron, R., and Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. *Nature* 469, 343–349.
- Mariot, V., Roche, S., Hourdé, C., Portilho, D., Sacconi, S., Puppo, F., Duguez, S., Rameau, P., Caruso, N., Delezoide, A.-L., et al. (2015). Correlation between low FAT1 expression and early affected muscle in FSHD. *Ann. Neurol.*
- Marshall, W.F. (2003). Gene expression and nuclear architecture during development and differentiation. *Mech. Dev.* 120, 1217–1230.
- Marsollier, A.-C., Ciszewski, L., Mariot, V., Popplewell, L., Voit, T., Dickson, G., and Dumonceaux, J. (2016). Antisense targeting of 3' end elements involved in DUX4 mRNA processing is an efficient therapeutic strategy for facioscapulohumeral dystrophy: a new gene-silencing approach. *Hum. Mol. Genet.* 25, 1468–1478.
- Martineau, L., Horan, M.A., Rothwell, N.J., and Little, R.A. (1992). Salbutamol, a beta 2-adrenoceptor agonist, increases skeletal muscle strength in young men. *Clin. Sci. Lond. Engl.* 1979 83, 615–621.
- Masny, P.S., Bengtsson, U., Chung, S.-A., Martin, J.H., van Engelen, B., van der Maarel, S.M., and Winokur, S.T. (2004). Localization of 4q35.2 to the nuclear periphery: is FSHD a nuclear envelope disease? *Hum. Mol. Genet.* 13, 1857–1871.
- Masny, P.S., Chan, O.Y.A., de Greef, J.C., Bengtsson, U., Ehrlich, M., Tawil, R., Lock, L.F., Hewitt, J.E., Stocksdale, J., Martin, J.H., et al. (2010). Analysis of allele-specific RNA transcription in FSHD by RNA-DNA FISH in single myonuclei. *Eur. J. Hum. Genet. EJHG* 18, 448–456.
- Mathews, K.D., Mills, K.A., Bosch, E.P., Ionasescu, V.V., Wiles, K.R., Buetow, K.H., and Murray, J.C. (1992). Linkage localization of facioscapulohumeral muscular dystrophy (FSHD) in 4q35. *Am. J. Hum. Genet.* 51, 428–431.
- Mátrai, J., Chuah, M.K.L., and VandenDriessche, T. (2010). Preclinical and clinical progress in hemophilia gene therapy. *Curr. Opin. Hematol.* 17, 387–392.
- Matsumura, T., Goto, K., Yamanaka, G., Lee, J.H., Zhang, C., Hayashi, Y.K., and Arahata, K. (2002). Chromosome 4q;10q translocations; comparison with different ethnic populations and FSHD patients. *BMC Neurol.* 2, 7.
- Mattout-Drubezki, A., and Gruenbaum, Y. (2003). Dynamic interactions of nuclear lamina proteins with chromatin and transcriptional machinery. *Cell. Mol. Life Sci. CMLS* 60, 2053–2063.

- Mauro, A. (1961). Satellite cell of skeletal muscle fibers. *J. Biophys. Biochem. Cytol.* 9, 493–495.
- Mays, L.E., Vandenberghe, L.H., Xiao, R., Bell, P., Nam, H.-J., Agbandje-McKenna, M., and Wilson, J.M. (2009). Adeno-associated virus capsid structure drives CD4-dependent CD8+ T cell response to vector encoded proteins. *J. Immunol. Baltim. Md 1950* 182, 6051–6060.
- McBride, J.L., Boudreau, R.L., Harper, S.Q., Staber, P.D., Monteys, A.M., Martins, I., Gilmore, B.L., Burstein, H., Peluso, R.W., Polisky, B., et al. (2008). Artificial miRNAs mitigate shRNA-mediated toxicity in the brain: implications for the therapeutic development of RNAi. *Proc. Natl. Acad. Sci. U. S. A.* 105, 5868–5873.
- McCready, S.J., Akrigg, A., and Cook, P.R. (1979). Electron-microscopy of intact nuclear DNA from human cells. *J. Cell Sci.* 39, 53–62.
- McKee, A.E., Minet, E., Stern, C., Riahi, S., Stiles, C.D., and Silver, P.A. (2005). A genome-wide in situ hybridization map of RNA-binding proteins reveals anatomically restricted expression in the developing mouse brain. *BMC Dev. Biol.* 5, 14.
- Medina-Kauwe, L.K., Xie, J., and Hamm-Alvarez, S. (2005). Intracellular trafficking of nonviral vectors. *Gene Ther.* 12, 1734–1751.
- Meijer, H.A., Smith, E.M., and Bushell, M. (2014). Regulation of miRNA strand selection: follow the leader? *Biochem. Soc. Trans.* 42, 1135–1140.
- Meister, G. (2013). Argonaute proteins: functional insights and emerging roles. *Nat. Rev. Genet.* 14, 447–459.
- Melters, D.P., Nye, J., Zhao, H., and Dalal, Y. (2015). Chromatin Dynamics in Vivo: A Game of Musical Chairs. *Genes* 6, 751–776.
- Mendell, J., Goemans, N., Rodino-Klapac, L., Sahenk, Z., Lowes, L., Alfano, L., Berry, K., Peterson, E., Lewis, S., Shontz, K., et al. (2016). Eteplirsen, a Phosphorodiamidate Morpholino Oligomer (PMO) for Duchenne Muscular Dystrophy (DMD): Clinical Update and Longitudinal Comparison to External Controls on Six-Minute Walk Test (6MWT) (S28.001). *Neurology* 86, S28.001.
- Mendell, J.R., Sahenk, Z., Malik, V., Gomez, A.M., Flanigan, K.M., Lowes, L.P., Alfano, L.N., Berry, K., Meadows, E., Lewis, S., et al. (2015). A phase 1/2a follistatin gene therapy trial for becker muscular dystrophy. *Mol. Ther. J. Am. Soc. Gene Ther.* 23, 192–201.
- Meneveri, R., Agresti, A., Marozzi, A., Saccone, S., Rocchi, M., Archidiacono, N., Corneo, G., Valle, G.D., and Ginelli, E. (1993). Molecular organization and chromosomal location of human GC-rich heterochromatic blocks. *Gene* 123, 227–234.
- Meyne, J., Ratliff, R.L., and Moyzis, R.K. (1989). Conservation of the human telomere sequence (TTAGGG)<sub>n</sub> among vertebrates. *Proc. Natl. Acad. Sci. U. S. A.* 86, 7049–7053.
- Midoux, P., Breuzard, G., Gomez, J.P., and Pichon, C. (2008). Polymer-based gene delivery: a current review on the uptake and intracellular trafficking of polyplexes. *Curr. Gene Ther.* 8, 335–352.
- Miller, N. (2012). Glybera and the future of gene therapy in the European Union. *Nat. Rev. Drug Discov.* 11, 419.
- Miller, J., McLachlan, A.D., and Klug, A. (1985). Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J.* 4, 1609–1614.

- Miller, J.C., Tan, S., Qiao, G., Barlow, K.A., Wang, J., Xia, D.F., Meng, X., Paschon, D.E., Leung, E., Hinkley, S.J., et al. (2011). A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* *29*, 143–148.
- Millevoi, S., and Vagner, S. (2010). Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation. *Nucleic Acids Res.* *38*, 2757–2774.
- Mills, K.A., Buetow, K.H., Xu, Y., Ritty, T.M., Mathews, K.D., Bodrug, S.E., Wijmenga, C., Balazs, I., and Murray, J.C. (1992). Genetic and physical mapping on chromosome 4 narrows the localization of the gene for facioscapulohumeral muscular dystrophy (FSHD). *Am. J. Hum. Genet.* *51*, 432–439.
- Minc, E., Courvalin, J.C., and Buendia, B. (2000). HP1gamma associates with euchromatin and heterochromatin in mammalian nuclei and chromosomes. *Cytogenet. Cell Genet.* *90*, 279–284.
- Mingozzi, F., and High, K.A. (2011). Therapeutic in vivo gene transfer for genetic disease using AAV: progress and challenges. *Nat. Rev. Genet.* *12*, 341–355.
- Mingozzi, F., and High, K.A. (2013). Immune responses to AAV vectors: overcoming barriers to successful gene therapy. *Blood* *122*, 23–36.
- Mingozzi, F., Maus, M.V., Hui, D.J., Sabatino, D.E., Murphy, S.L., Rasko, J.E.J., Ragni, M.V., Manno, C.S., Sommer, J., Jiang, H., et al. (2007). CD8(+) T-cell responses to adeno-associated virus capsid in humans. *Nat. Med.* *13*, 419–422.
- Mirkovitch, J., Mirault, M.E., and Laemmli, U.K. (1984). Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. *Cell* *39*, 223–232.
- Mitropant, C., Porensky, P., Zhou, H., Price, L., Muntoni, F., Fletcher, S., Wilton, S.D., and Burghes, A.H.M. (2013). Improved antisense oligonucleotide design to suppress aberrant SMN2 gene transcript processing: towards a treatment for spinal muscular atrophy. *PLoS One* *8*, e62114.
- Mitsuhashi, H., Mitsuhashi, S., Lynn-Jones, T., Kawahara, G., and Kunkel, L.M. (2013a). Expression of DUX4 in zebrafish development recapitulates facioscapulohumeral muscular dystrophy. *Hum. Mol. Genet.* *22*, 568–577.
- Mitsuhashi, S., Boyden, S.E., Estrella, E.A., Jones, T.I., Rahimov, F., Yu, T.W., Darras, B.T., Amato, A.A., Folkert, R.D., Jones, P.L., et al. (2013b). Exome sequencing identifies a novel SMCHD1 mutation in facioscapulohumeral muscular dystrophy 2. *Neuromuscul. Disord.* *23*, 975–980.
- Mitsui, K., Nakajima, D., Ohara, O., and Nakayama, M. (2002). Mammalian fat3: a large protein that contains multiple cadherin and EGF-like motifs. *Biochem. Biophys. Res. Commun.* *290*, 1260–1266.
- Miura, K., Kumagai, T., Matsumoto, A., Iriyama, E., Watanabe, K., Goto, K., and Arahata, K. (1998). Two cases of chromosome 4q35-linked early onset facioscapulohumeral muscular dystrophy with mental retardation and epilepsy. *Neuropediatrics* *29*, 239–241.
- Moeller, M.J., Soofi, A., Braun, G.S., Li, X., Watzl, C., Kriz, W., and Holzman, L.B. (2004). Protocadherin FAT1 binds Ena/VASP proteins and is necessary for actin dynamics and cell polarization. *EMBO J.* *23*, 3769–3779.
- Monaco, A.P., Bertelson, C.J., Liechti-Gallati, S., Moser, H., and Kunkel, L.M. (1988). An explanation for the phenotypic differences between patients bearing partial deletions of the DMD locus. *Genomics* *2*, 90–95.

- Monia, B.P., Lesnik, E.A., Gonzalez, C., Lima, W.F., McGee, D., Guinosso, C.J., Kawasaki, A.M., Cook, P.D., and Freier, S.M. (1993). Evaluation of 2'-modified oligonucleotides containing 2'-deoxy gaps as antisense inhibitors of gene expression. *J. Biol. Chem.* 268, 14514–14522.
- Moore, C.B., Guthrie, E.H., Huang, M.T.-H., and Taxman, D.J. (2010). Short Hairpin RNA (shRNA): Design, Delivery, and Assessment of Gene Knockdown. *Methods Mol. Biol.* Clifton NJ 629, 141–158.
- Moore, L.D., Le, T., and Fan, G. (2013). DNA Methylation and Its Basic Function. *Neuropsychopharmacology* 38, 23–38.
- Moroziwicz, D., and Kaufman, H.L. (2005). Gene therapy with poxvirus vectors. *Curr. Opin. Mol. Ther.* 7, 317–325.
- Moscou, M.J., and Bogdanove, A.J. (2009). A simple cipher governs DNA recognition by TAL effectors. *Science* 326, 1501.
- Mould, A.W., Pang, Z., Pakusch, M., Tonks, I.D., Stark, M., Carrie, D., Mukhopadhyay, P., Seidel, A., Ellis, J.J., Deakin, J., et al. (2013). Smchd1 regulates a subset of autosomal genes subject to monoallelic expression in addition to being critical for X inactivation. *Epigenetics Chromatin* 6, 19.
- Mountain, A. (2000). Gene therapy: the first decade. *Trends Biotechnol.* 18, 119–128.
- Moyzis, R.K., Buckingham, J.M., Cram, L.S., Dani, M., Deaven, L.L., Jones, M.D., Meyne, J., Ratliff, R.L., and Wu, J.R. (1988). A highly conserved repetitive DNA sequence, (TTAGGG)<sub>n</sub>, present at the telomeres of human chromosomes. *Proc. Natl. Acad. Sci. U. S. A.* 85, 6622–6626.
- Mul, K., Lassche, S., Voermans, N.C., Padberg, G.W., Horlings, C.G., and van Engelen, B.G. (2016). What's in a name? The clinical features of facioscapulohumeral muscular dystrophy. *Pract. Neurol.* 16, 201–207.
- Muntoni, F., Torelli, S., and Ferlini, A. (2003). Dystrophin and mutations: one gene, several proteins, multiple phenotypes. *Lancet Neurol.* 2, 731–740.
- Murlidharan, G., Samulski, R.J., and Asokan, A. (2014). Biology of adeno-associated viral vectors in the central nervous system. *Front. Mol. Neurosci.* 7, 76.
- Mussolino, C., Morbitzer, R., Lütge, F., Dannemann, N., Lahaye, T., and Cathomen, T. (2011). A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.* 39, 9283–9293.
- Muzyczka, N. (1992). Use of adeno-associated virus as a general transduction vector for mammalian cells. *Curr. Top. Microbiol. Immunol.* 158, 97–129.
- Nagy, E., and Maquat, L.E. (1998). A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. *Trends Biochem. Sci.* 23, 198–199.
- Nakamura, Y., Koyama, K., and Matsushima, M. (1998). VNTR (variable number of tandem repeat) sequences as transcriptional, translational, or functional regulators. *J. Hum. Genet.* 43, 149–152.
- Nakayama, M., Nakajima, D., Yoshimura, R., Endo, Y., and Ohara, O. (2002). MEGF1/fat2 proteins containing extraordinarily large extracellular domains are localized to thin parallel fibers of cerebellar granule cells. *Mol. Cell. Neurosci.* 20, 563–578.
- Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N., and Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* 393, 386–389.

- Napoli, C., Lemieux, C., and Jorgensen, R. (1990). Introduction of a Chimeric Chalcone Synthase Gene into *Petunia* Results in Reversible Co-Suppression of Homologous Genes in trans. *Plant Cell* 2, 279–289.
- Nathwani, A.C., Tuddenham, E.G.D., Rangarajan, S., Rosales, C., McIntosh, J., Linch, D.C., Chowdary, P., Riddell, A., Pie, A.J., Harrington, C., et al. (2011). Adenovirus-associated virus vector-mediated gene transfer in hemophilia B. *N. Engl. J. Med.* 365, 2357–2365.
- Nayerossadat, N., Maedeh, T., and Ali, P.A. (2012). Viral and nonviral delivery systems for gene delivery. *Adv. Biomed. Res.* 1.
- Neguembor, M.V., and Gabellini, D. (2010). In junk we trust: repetitive DNA, epigenetics and facioscapulohumeral muscular dystrophy. *Epigenomics* 2, 271–287.
- Nelson, C.E., Hakim, C.H., Ousterout, D.G., Thakore, P.I., Moreb, E.A., Rivera, R.M.C., Madhavan, S., Pan, X., Ran, F.A., Yan, W.X., et al. (2016). In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science* 351, 403–407.
- Newman, C.M., Lawrie, A., Brisken, A.F., and Cumberland, D.C. (2001). Ultrasound gene therapy: on the road from concept to reality. *Echocardiogr. Mt. Kisco* 18, 339–347.
- Nguyen, K., Walrafen, P., Bernard, R., Attarian, S., Chaix, C., Vovan, C., Renard, E., Dufrane, N., Pouget, J., Vannier, A., et al. (2011). Molecular combing reveals allelic combinations in facioscapulohumeral dystrophy. *Ann. Neurol.* 70, 627–633.
- Nicholson, P., Yepiskoposyan, H., Metze, S., Zamudio Orozco, R., Kleinschmidt, N., and Mühlemann, O. (2010). Nonsense-mediated mRNA decay in human cells: mechanistic insights, functions beyond quality control and the double-life of NMD factors. *Cell. Mol. Life Sci. CMLS* 67, 677–700.
- Nielsen, P.E., Egholm, M., Berg, R.H., and Buchardt, O. (1991). Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* 254, 1497–1500.
- Niemeyer, G.P., Herzog, R.W., Mount, J., Arruda, V.R., Tillson, D.M., Hathcock, J., van Ginkel, F.W., High, K.A., and Lothrop, C.D. (2009). Long-term correction of inhibitor-prone hemophilia B dogs treated with liver-directed AAV2-mediated factor IX gene therapy. *Blood* 113, 797–806.
- Nikolov, M., Stützer, A., Mosch, K., Krasauskas, A., Soeroes, S., Stark, H., Urlaub, H., and Fischle, W. (2011). Chromatin affinity purification and quantitative mass spectrometry defining the interactome of histone modification patterns. *Mol. Cell. Proteomics MCP* 10, M110.005371.
- Noh, K.-M., Allis, C.D., and Li, H. (2016). Reading between the Lines: “ADD”-ing Histone and DNA Methylation Marks toward a New Epigenetic “Sum.” *ACS Chem. Biol.* 11, 554–563.
- Noland, C.L., Ma, E., and Doudna, J.A. (2011). siRNA Repositioning for Guide Strand Selection by Human Dicer Complexes. *Mol. Cell* 43, 110–121.
- Nollet, F., Kools, P., and van Roy, F. (2000). Phylogenetic analysis of the cadherin superfamily allows identification of six major subfamilies besides several solitary members. *J. Mol. Biol.* 299, 551–572.
- Nonaka, N., Kitajima, T., Yokobayashi, S., Xiao, G., Yamamoto, M., Grewal, S.I.S., and Watanabe, Y. (2002). Recruitment of cohesin to heterochromatic regions by Swi6/HP1 in fission yeast. *Nat. Cell Biol.* 4, 89–93.
- Nonnenmacher, M., and Weber, T. (2012). Intracellular transport of recombinant adeno-associated virus vectors. *Gene Ther.* 19, 649–658.

- Nowrouzi, A., Penaud-Budloo, M., Kaepfel, C., Appelt, U., Le Guiner, C., Moullier, P., von Kalle, C., Snyder, R.O., and Schmidt, M. (2012). Integration frequency and intermolecular recombination of rAAV vectors in non-human primate skeletal muscle and liver. *Mol. Ther. J. Am. Soc. Gene Ther.* *20*, 1177–1186.
- Nozawa, R.-S., Nagao, K., Igami, K.-T., Shibata, S., Shirai, N., Nozaki, N., Sado, T., Kimura, H., and Obuse, C. (2013). Human inactive X chromosome is compacted through a PRC2-independent SMCHD1-HBiX1 pathway. *Nat. Struct. Mol. Biol.* *20*, 566–573.
- Nunes, N.M., Li, W., Tian, B., and Furger, A. (2010). A functional human Poly(A) site requires only a potent DSE and an A-rich upstream sequence. *EMBO J.* *29*, 1523–1536.
- O'Brien, J.A., and Lummis, S.C.R. (2006). Biolistic transfection of neuronal cultures using a hand-held gene gun. *Nat. Protoc.* *1*, 977–981.
- Odermatt, P., Trüb, J., Furrer, L., Fricker, R., Marti, A., and Schümperli, D. (2016). Somatic Therapy of a Mouse SMA Model with a U7 snRNA Gene Correcting SMN2 Splicing. *Mol. Ther. J. Am. Soc. Gene Ther.* *24*, 1797–1805.
- Okamura, K., Liu, N., and Lai, E.C. (2009). Distinct mechanisms for microRNA strand selection by *Drosophila Argonautes*. *Mol. Cell* *36*, 431–444.
- Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* *99*, 247–257.
- Olsen, D.B., Ørngreen, M.C., and Vissing, J. (2005). Aerobic training improves exercise performance in facioscapulohumeral muscular dystrophy. *Neurology* *64*, 1064–1066.
- Osborne, R.J., Welle, S., Venance, S.L., Thornton, C.A., and Tawil, R. (2007). Expression profile of FSHD supports a link between retinal vasculopathy and muscular dystrophy. *Neurology* *68*, 569–577.
- Ostlund, C., Garcia-Carrasquillo, R.M., Belayew, A., and Worman, H.J. (2005). Intracellular trafficking and dynamics of double homeodomain proteins. *Biochemistry (Mosc.)* *44*, 2378–2384.
- Ottaviani, A., Gilson, E., and Magdinier, F. (2008). Telomeric position effect: from the yeast paradigm to human pathologies? *Biochimie* *90*, 93–107.
- Ottaviani, A., Rival-Gervier, S., Boussouar, A., Foerster, A.M., Rondier, D., Sacconi, S., Desnuelle, C., Gilson, E., and Magdinier, F. (2009a). The D4Z4 macrosatellite repeat acts as a CTCF and A-type lamins-dependent insulator in facio-scapulo-humeral dystrophy. *PLoS Genet.* *5*, e1000394.
- Ottaviani, A., Schluth-Bolard, C., Rival-Gervier, S., Boussouar, A., Rondier, D., Foerster, A.M., Morere, J., Bauwens, S., Gazzo, S., Callet-Bauchu, E., et al. (2009b). Identification of a perinuclear positioning element in human subtelomeres that requires A-type lamins and CTCF. *EMBO J.* *28*, 2428–2436.
- Ottaviani, A., Schluth-Bolard, C., Gilson, E., and Magdinier, F. (2010). D4Z4 as a prototype of CTCF and lamins-dependent insulator in human cells. *Nucl. Austin Tex* *1*, 30–36.
- Oudet, P., Gross-Bellard, M., and Chambon, P. (1975). Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell* *4*, 281–300.
- van Overveld, P.G., Lemmers, R.J., Deidda, G., Sandkuijl, L., Padberg, G.W., Frants, R.R., and van der Maarel, S.M. (2000). Interchromosomal repeat array interactions between chromosomes 4 and 10: a model for subtelomeric plasticity. *Hum. Mol. Genet.* *9*, 2879–2884.



- van Overveld, P.G.M., Lemmers, R.J.F.L., Sandkuijl, L.A., Enthoven, L., Winokur, S.T., Bakels, F., Padberg, G.W., van Ommen, G.-J.B., Frants, R.R., and van der Maarel, S.M. (2003). Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy. *Nat. Genet.* 35, 315–317.
- van Overveld, P.G.M., Enthoven, L., Ricci, E., Rossi, M., Felicetti, L., Jeanpierre, M., Winokur, S.T., Frants, R.R., Padberg, G.W., and van der Maarel, S.M. (2005). Variable hypomethylation of D4Z4 in facioscapulohumeral muscular dystrophy. *Ann. Neurol.* 58, 569–576.
- Ozaki, T., and Nakagawara, A. (2011). Role of p53 in Cell Death and Human Cancers. *Cancers* 3, 994–1013.
- Padberg, G. (1982). Facioscapulohumeral disease. MD Thesis Univ. Leiden.
- Padberg, G.W., Lunt, P.W., Koch, M., and Fardeau, M. (1991). Diagnostic criteria for facioscapulohumeral muscular dystrophy. *Neuromuscul. Disord.* NMD 1, 231–234.
- Padberg, G.W., Frants, R.R., Brouwer, O.F., Wijmenga, C., Bakker, E., and Sandkuijl, L.A. (1995a). Facioscapulohumeral muscular dystrophy in the Dutch population. *Muscle Nerve. Suppl.* 2, S81-84.
- Padberg, G.W., Brouwer, O.F., de Keizer, R.J., Dijkman, G., Wijmenga, C., Grote, J.J., and Frants, R.R. (1995b). On the significance of retinal vascular disease and hearing loss in facioscapulohumeral muscular dystrophy. *Muscle Nerve. Suppl.* 2, S73-80.
- Pandey, S.N., Cabotage, J., Shi, R., Dixit, M., Sutherland, M., Liu, J., Muger, S., Harper, S.Q., Nagaraju, K., and Chen, Y.-W. (2012). Conditional over-expression of PITX1 causes skeletal muscle dystrophy in mice. *Biol. Open* 1, 629–639.
- Pandey, S.N., Lee, Y.-C., Yokota, T., and Chen, Y.-W. (2014). Morpholino treatment improves muscle function and pathology of Pitx1 transgenic mice. *Mol. Ther. J. Am. Soc. Gene Ther.* 22, 390–396.
- Pandya, S., King, W.M., and Tawil, R. (2008). Facioscapulohumeral dystrophy. *Phys. Ther.* 88, 105–113.
- Pantoja, M., and Ruohola-Baker, H. (2013). Drosophila as a starting point for developing therapeutics for the rare disease Duchenne Muscular Dystrophy. *Rare Dis.* 1.
- Parsi, K.M., Hennessy, E., Kearns, N., and Maehr, R. (2017). Using an Inducible CRISPR-dCas9-KRAB Effector System to Dissect Transcriptional Regulation in Human Embryonic Stem Cells. *Methods Mol. Biol. Clifton NJ* 1507, 221–233.
- Pasquinelli, A.E. (2012). MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship. *Nat. Rev. Genet.* 13, 271–282.
- Passerieux, E., Hayot, M., Jaussent, A., Carnac, G., Gouzi, F., Pillard, F., Picot, M.-C., Böcker, K., Hugon, G., Pincemail, J., et al. (2015). Effects of vitamin C, vitamin E, zinc gluconate, and selenomethionine supplementation on muscle function and oxidative stress biomarkers in patients with facioscapulohumeral dystrophy: a double-blind randomized controlled clinical trial. *Free Radic. Biol. Med.* 81, 158–169.
- Patel, S.B., and Bellini, M. (2008). The assembly of a spliceosomal small nuclear ribonucleoprotein particle. *Nucleic Acids Res.* 36, 6482–6493.
- Pattanayak, V., Ramirez, C.L., Joung, J.K., and Liu, D.R. (2011). Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nat. Methods* 8, 765–770.

- Paule, M.R., and White, R.J. (2000). Survey and summary: transcription by RNA polymerases I and III. *Nucleic Acids Res.* 28, 1283–1298.
- Pauleikhoff, D., Bornfeld, N., Bird, A.C., and Wessing, A. (1992). Severe visual loss associated with retinal telangiectasis and facioscapulohumeral muscular dystrophy. *Graefes Arch. Clin. Exp. Ophthalmol. Albrecht Von Graefes Arch. Klin. Exp. Ophthalmol.* 230, 362–365.
- Pavletich, N.P., and Pabo, C.O. (1991). Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science* 252, 809–817.
- Payan, C.A., Hogrel, J.Y., Hammouda, E.H., Lacomblez, L., Ollivier, G., Doppler, V., Eymard, B., Attarian, S., Pouget, J., Desnuelle, C., et al. (2009). Periodic salbutamol in facioscapulohumeral muscular dystrophy: a randomized controlled trial. *Arch. Phys. Med. Rehabil.* 90, 1094–1101.
- Pedram, M., Sprung, C.N., Gao, Q., Lo, A.W.I., Reynolds, G.E., and Murnane, J.P. (2006). Telomere position effect and silencing of transgenes near telomeres in the mouse. *Mol. Cell. Biol.* 26, 1865–1878.
- Pelisek, J., Gaedtke, L., DeRouche, J., Walker, G.F., Nikol, S., and Wagner, E. (2006). Optimized lipopolyplex formulations for gene transfer to human colon carcinoma cells under in vitro conditions. *J. Gene Med.* 8, 186–197.
- Peng, Z. (2005). Current status of gendicine in China: recombinant human Ad-p53 agent for treatment of cancers. *Hum. Gene Ther.* 16, 1016–1027.
- Perry, S.V. (1998). Troponin T: genetics, properties and function. *J. Muscle Res. Cell Motil.* 19, 575–602.
- Peters, A.H.F.M., Kubicek, S., Mechtler, K., O’Sullivan, R.J., Derijck, A.A.H.A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tachibana, M., Shinkai, Y., et al. (2003). Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Mol. Cell* 12, 1577–1589.
- Petrov, A., Pirozhkova, I., Carnac, G., Laoudj, D., Lipinski, M., and Vassetzky, Y.S. (2006). Chromatin loop domain organization within the 4q35 locus in facioscapulohumeral dystrophy patients versus normal human myoblasts. *Proc. Natl. Acad. Sci. U. S. A.* 103, 6982–6987.
- Petrov, A., Allinne, J., Pirozhkova, I., Laoudj, D., Lipinski, M., and Vassetzky, Y.S. (2008). A nuclear matrix attachment site in the 4q35 locus has an enhancer-blocking activity in vivo: implications for the facio-scapulo-humeral dystrophy. *Genome Res.* 18, 39–45.
- Pichon, C., Billiet, L., and Midoux, P. (2010). Chemical vectors for gene delivery: uptake and intracellular trafficking. *Curr. Opin. Biotechnol.* 21, 640–645.
- Pillai, R.S., Will, C.L., Lührmann, R., Schümperli, D., and Müller, B. (2001). Purified U7 snRNPs lack the Sm proteins D1 and D2 but contain Lsm10, a new 14 kDa Sm D1-like protein. *EMBO J.* 20, 5470–5479.
- Pillai, R.S., Grimmmer, M., Meister, G., Will, C.L., Lührmann, R., Fischer, U., and Schümperli, D. (2003). Unique Sm core structure of U7 snRNPs: assembly by a specialized SMN complex and the role of a new component, Lsm11, in histone RNA processing. *Genes Dev.* 17, 2321–2333.
- Pistoni, M., Shiue, L., Cline, M.S., Bortolanza, S., Neguembor, M.V., Xynos, A., Ares, M., Jr, and Gabellini, D. (2013). Rbfox1 downregulation and altered calpain 3 splicing by FRG1 in a mouse model of Facioscapulohumeral muscular dystrophy (FSHD). *PLoS Genet.* 9, e1003186.

- Ponassi, M., Jacques, T.S., Ciani, L., and French Constant, C. (1999). Expression of the rat homologue of the Drosophila fat tumour suppressor gene. *Mech. Dev.* 80, 207–212.
- Pongracz, K., and Gryaznov, S.M. (1998). alpha-Oligodeoxyribonucleotide N3'-->P5' phosphoramidates: synthesis and duplex formation. *Nucleic Acids Res.* 26, 1099–1106.
- Porensky, P.N., and Burghes, A.H.M. (2013). Antisense oligonucleotides for the treatment of spinal muscular atrophy. *Hum. Gene Ther.* 24, 489–498.
- Portela, A., and Esteller, M. (2010). Epigenetic modifications and human disease. *Nat. Biotechnol.* 28, 1057–1068.
- Proudfoot, N.J. (2011). Ending the message: poly(A) signals then and now. *Genes Dev.* 25, 1770–1782.
- Provost, E., and Rimm, D.L. (1999). Controversies at the cytoplasmic face of the cadherin-based adhesion complex. *Curr. Opin. Cell Biol.* 11, 567–572.
- Puppo, F., Dionnet, E., Gaillard, M.-C., Gaildrat, P., Castro, C., Vovan, C., Bertaux, K., Bernard, R., Attarian, S., Goto, K., et al. (2015). Identification of variants in the 4q35 gene FAT1 in patients with a facioscapulohumeral dystrophy-like phenotype. *Hum. Mutat.* 36, 443–453.
- Qiao, C., Li, J., Jiang, J., Zhu, X., Wang, B., Li, J., and Xiao, X. (2008). Myostatin propeptide gene delivery by adeno-associated virus serotype 8 vectors enhances muscle growth and ameliorates dystrophic phenotypes in mdx mice. *Hum. Gene Ther.* 19, 241–254.
- Qiao, C., Koo, T., Li, J., Xiao, X., and Dickson, J.G. (2011). Gene therapy in skeletal muscle mediated by adeno-associated virus vectors. *Methods Mol. Biol. Clifton NJ* 807, 119–140.
- Qing, K., Mah, C., Hansen, J., Zhou, S., Dwarki, V., and Srivastava, A. (1999). Human fibroblast growth factor receptor 1 is a co-receptor for infection by adeno-associated virus 2. *Nat. Med.* 5, 71–77.
- Qiu, C., Sawada, K., Zhang, X., and Cheng, X. (2002). The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. *Nat. Struct. Biol.* 9, 217–224.
- Quilez, J., Guilmatre, A., Garg, P., Highnam, G., Gymrek, M., Erlich, Y., Joshi, R.S., Mittelman, D., and Sharp, A.J. (2016). Polymorphic tandem repeats within gene promoters act as modifiers of gene expression and DNA methylation in humans. *Nucleic Acids Res.* 44, 3750–3762.
- Rao, D.D., Vorhies, J.S., Senzer, N., and Nemunaitis, J. (2009). siRNA vs. shRNA: Similarities and differences. *Adv. Drug Deliv. Rev.* 61, 746–759.
- Rappsilber, J., Ryder, U., Lamond, A.I., and Mann, M. (2002). Large-scale proteomic analysis of the human spliceosome. *Genome Res.* 12, 1231–1245.
- Raz, V., Buijze, H., Raz, Y., Verwey, N., Anvar, S.Y., Aartsma-Rus, A., and van der Maarel, S.M. (2014). A novel feed-forward loop between ARIH2 E3-ligase and PABPN1 regulates aging-associated muscle degeneration. *Am. J. Pathol.* 184, 1119–1131.
- Razin, S.V. (2001). The nuclear matrix and chromosomal DNA loops: is there any correlation between partitioning of the genome into loops and functional domains? *Cell. Mol. Biol. Lett.* 6, 59–69.
- Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., et al. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406, 593–599.

- Reddy, K.L., Zullo, J.M., Bertolino, E., and Singh, H. (2008). Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature* *452*, 243–247.
- Relaix, F., and Zammit, P.S. (2012). Satellite cells are essential for skeletal muscle regeneration: the cell on the edge returns centre stage. *Dev. Camb. Engl.* *139*, 2845–2856.
- Relaix, F., Montarras, D., Zaffran, S., Gayraud-Morel, B., Rocancourt, D., Tajbakhsh, S., Mansouri, A., Cumano, A., and Buckingham, M. (2006). Pax3 and Pax7 have distinct and overlapping functions in adult muscle progenitor cells. *J. Cell Biol.* *172*, 91–102.
- Reynolds, B.C., Lemmers, R.J.L.F., Tolmie, J., Howatson, A.G., and Hughes, D.A. (2010). Focal segmental glomerulosclerosis, Coats'-like retinopathy, sensorineural deafness and chromosome 4 duplication: a new association. *Pediatr. Nephrol. Berl. Ger.* *25*, 1551–1554.
- Rezaee, M., Oskuee, R.K., Nassirli, H., and Malaekheh-Nikouei, B. (2016). Progress in the development of lipopolyplexes as efficient non-viral gene delivery systems. *J. Control. Release Off. J. Control. Release Soc.* *236*, 1–14.
- Ricci, E., Galluzzi, G., Deidda, G., Cacurri, S., Colantoni, L., Merico, B., Piazzo, N., Servidei, S., Vigneti, E., Pasceri, V., et al. (1999). Progress in the molecular diagnosis of facioscapulohumeral muscular dystrophy and correlation between the number of KpnI repeats at the 4q35 locus and clinical phenotype. *Ann. Neurol.* *45*, 751–757.
- Ricci, G., Scionti, I., Sera, F., Govi, M., D'Amico, R., Frambolli, I., Mele, F., Filosto, M., Vercelli, L., Ruggiero, L., et al. (2013). Large scale genotype-phenotype analyses indicate that novel prognostic tools are required for families with facioscapulohumeral muscular dystrophy. *Brain J. Neurol.* *136*, 3408–3417.
- Rice, J.C., Briggs, S.D., Ueberheide, B., Barber, C.M., Shabanowitz, J., Hunt, D.F., Shinkai, Y., and Allis, C.D. (2003). Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. *Mol. Cell* *12*, 1591–1598.
- Rickard, A.M., Petek, L.M., and Miller, D.G. (2015). Endogenous DUX4 expression in FSHD myotubes is sufficient to cause cell death and disrupts RNA splicing and cell migration pathways. *Hum. Mol. Genet.* *24*, 5901–5914.
- Ridgway, P., and Almouzni, G. (2001). Chromatin assembly and organization. *J. Cell Sci.* *114*, 2711–2712.
- Rijkers, T., Deidda, G., van Koningsbruggen, S., van Geel, M., Lemmers, R.J.L.F., van Deutekom, J.C.T., Figlewicz, D., Hewitt, J.E., Padberg, G.W., Frants, R.R., et al. (2004). FRG2, an FSHD candidate gene, is transcriptionally upregulated in differentiating primary myoblast cultures of FSHD patients. *J. Med. Genet.* *41*, 826–836.
- Ringrose, L., and Paro, R. (2004). Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu. Rev. Genet.* *38*, 413–443.
- Rivera, V.M., Gao, G., Grant, R.L., Schnell, M.A., Zoltick, P.W., Rozamus, L.W., Clackson, T., and Wilson, J.M. (2005). Long-term pharmacologically regulated expression of erythropoietin in primates following AAV-mediated gene transfer. *Blood* *105*, 1424–1430.
- Robertson, K.D. (2005). DNA methylation and human disease. *Nat. Rev. Genet.* *6*, 597–610.
- Robertson, K.D., and Wolffe, A.P. (2000). DNA methylation in health and disease. *Nat. Rev. Genet.* *1*, 11–19.

- Robin, J.D., Ludlow, A.T., Batten, K., Gaillard, M.-C., Stadler, G., Magdinier, F., Wright, W.E., and Shay, J.W. (2015). SORBS2 transcription is activated by telomere position effect-over long distance upon telomere shortening in muscle cells from patients with facioscapulohumeral dystrophy. *Genome Res.* 25, 1781–1790.
- Rodino-Klapac, L.R., Montgomery, C.L., Bremer, W.G., Shontz, K.M., Malik, V., Davis, N., Sprinkle, S., Campbell, K.J., Sahenk, Z., Clark, K.R., et al. (2010). Persistent expression of FLAG-tagged micro dystrophin in nonhuman primates following intramuscular and vascular delivery. *Mol. Ther. J. Am. Soc. Gene Ther.* 18, 109–117.
- Rodino-Klapac, L.R., Janssen, P.M.L., Shontz, K.M., Canan, B., Montgomery, C.L., Griffin, D., Heller, K., Schmelzer, L., Handy, C., Clark, K.R., et al. (2013). Micro-dystrophin and follistatin co-delivery restores muscle function in aged DMD model. *Hum. Mol. Genet.* 22, 4929–4937.
- Rodriguez, A., Griffiths-Jones, S., Ashurst, J.L., and Bradley, A. (2004). Identification of mammalian microRNA host genes and transcription units. *Genome Res.* 14, 1902–1910.
- Rosazza, C., Meglic, S.H., Zumbusch, A., Rols, M.-P., and Miklavcic, D. (2016). Gene Electrotransfer: A Mechanistic Perspective. *Curr. Gene Ther.* 16, 98–129.
- Rossi, M., Ricci, E., Colantoni, L., Galluzzi, G., Frusciante, R., Tonali, P.A., and Felicetti, L. (2007). The Facioscapulohumeral muscular dystrophy region on 4qter and the homologous locus on 10qter evolved independently under different evolutionary pressure. *BMC Med. Genet.* 8, 8.
- Sacconi, S., Camaño, P., de Greef, J.C., Lemmers, R.J.L.F., Salviati, L., Boileau, P., Lopez de Munain Arregui, A., van der Maarel, S.M., and Desnuelle, C. (2012). Patients with a phenotype consistent with facioscapulohumeral muscular dystrophy display genetic and epigenetic heterogeneity. *J. Med. Genet.* 49, 41–46.
- Sacconi, S., Lemmers, R.J.L.F., Balog, J., van der Vliet, P.J., Lahaut, P., van Nieuwenhuizen, M.P., Straasheijm, K.R., Debipersad, R.D., Vos-Versteeg, M., Salviati, L., et al. (2013). The FSHD2 gene SMCHD1 is a modifier of disease severity in families affected by FSHD1. *Am. J. Hum. Genet.* 93, 744–751.
- Sachs, A. (1990). The role of poly(A) in the translation and stability of mRNA. *Curr. Opin. Cell Biol.* 2, 1092–1098.
- Saito, Y., Miyashita, S., Yokoyama, A., Komaki, H., Seki, A., Maegaki, Y., and Ohno, K. (2007). Facioscapulohumeral muscular dystrophy with severe mental retardation and epilepsy. *Brain Dev.* 29, 231–233.
- Saksouk, N., Simboeck, E., and Déjardin, J. (2015). Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin* 8, 3.
- Sánchez, R., and Marzluff, W.F. (2002). The stem-loop binding protein is required for efficient translation of histone mRNA in vivo and in vitro. *Mol. Cell. Biol.* 22, 7093–7104.
- Sancisi, V., Germinario, E., Esposito, A., Morini, E., Peron, S., Moggio, M., Tomelleri, G., Danieli-Betto, D., and Tupler, R. (2014). Altered Tnnt3 characterizes selective weakness of fast fibers in mice overexpressing FSHD region gene 1 (FRG1). *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 306, R124–137.
- Sandri, M., El Meslemani, A.H., Sandri, C., Schjerling, P., Vissing, K., Andersen, J.L., Rossini, K., Carraro, U., and Angelini, C. (2001). Caspase 3 expression correlates with skeletal muscle apoptosis in Duchenne and facioscapulo human muscular dystrophy. A potential target for pharmacological treatment? *J. Neuropathol. Exp. Neurol.* 60, 302–312.

- Sands, M.S. (2011). AAV-mediated liver-directed gene therapy. *Methods Mol. Biol.* Clifton NJ 807, 141–157.
- Sanjana, N.E., Cong, L., Zhou, Y., Cunniff, M.M., Feng, G., and Zhang, F. (2012). A transcription activator-like effector toolbox for genome engineering. *Nat. Protoc.* 7, 171–192.
- Sanlioglu, S., Benson, P.K., Yang, J., Atkinson, E.M., Reynolds, T., and Engelhardt, J.F. (2000). Endocytosis and nuclear trafficking of adeno-associated virus type 2 are controlled by rac1 and phosphatidylinositol-3 kinase activation. *J. Virol.* 74, 9184–9196.
- Santiago, Y., Chan, E., Liu, P.-Q., Orlando, S., Zhang, L., Urnov, F.D., Holmes, M.C., Guschin, D., Waite, A., Miller, J.C., et al. (2008). Targeted gene knockout in mammalian cells by using engineered zinc-finger nucleases. *Proc. Natl. Acad. Sci. U. S. A.* 105, 5809–5814.
- Sarfarazi, M., Upadhyaya, M., Padberg, G., Pericak-Vance, M., Siddique, T., Lucotte, G., and Lunt, P. (1989). An exclusion map for facioscapulohumeral (Landouzy-Déjérine) disease. *J. Med. Genet.* 26, 481–484.
- Sarfarazi, M., Wijmenga, C., Upadhyaya, M., Weiffenbach, B., Hyser, C., Mathews, K., Murray, J., Gilbert, J., Pericak-Vance, M., and Lunt, P. (1992). Regional mapping of facioscapulohumeral muscular dystrophy gene on 4q35: combined analysis of an international consortium. *Am. J. Hum. Genet.* 51, 396–403.
- Sarraf, S.A., and Stancheva, I. (2004). Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. *Mol. Cell* 15, 595–605.
- Saxonov, S., Berg, P., and Brutlag, D.L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc. Natl. Acad. Sci. U. S. A.* 103, 1412–1417.
- Schlehofer, J.R., Ehrbar, M., and zur Hausen, H. (1986). Vaccinia virus, herpes simplex virus, and carcinogens induce DNA amplification in a human cell line and support replication of a helpervirus dependent parvovirus. *Virology* 152, 110–117.
- Schnepp, B.C., Clark, K.R., Klemanski, D.L., Pacak, C.A., and Johnson, P.R. (2003). Genetic Fate of Recombinant Adeno-Associated Virus Vector Genomes in Muscle. *J. Virol.* 77, 3495–3504.
- Schnepp, B.C., Jensen, R.L., Chen, C.-L., Johnson, P.R., and Clark, K.R. (2005). Characterization of adeno-associated virus genomes isolated from human tissues. *J. Virol.* 79, 14793–14803.
- Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B., and Cavalli, G. (2007). Genome Regulation by Polycomb and Trithorax Proteins. *Cell* 128, 735–745.
- Schulz, M., Schneider, S., Lottspeich, F., Renkawitz, R., and Eggert, M. (2001). Identification of nucleolin as a glucocorticoid receptor interacting protein. *Biochem. Biophys. Res. Commun.* 280, 476–480.
- Schümperli, D., and Pillai, R.S. (2004). The special Sm core structure of the U7 snRNP: far-reaching significance of a small nuclear ribonucleoprotein. *Cell. Mol. Life Sci. CMLS* 61, 2560–2570.
- Scionti, I., Greco, F., Ricci, G., Govi, M., Arashiro, P., Vercelli, L., Berardinelli, A., Angelini, C., Antonini, G., Cao, M., et al. (2012). Large-scale population analysis challenges the current criteria for the molecular diagnosis of facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.* 90, 628–635.

- Scott, M.P., Tamkun, J.W., and Hartzell III, G.W. (1989). The structure and function of the homeodomain. *Biochim. Biophys. Acta BBA - Rev. Cancer* 989, 25–48.
- Shapiro, L., Fannon, A.M., Kwong, P.D., Thompson, A., Lehmann, M.S., Grübel, G., Legrand, J.F., Als-Nielsen, J., Colman, D.R., and Hendrickson, W.A. (1995). Structural basis of cell-cell adhesion by cadherins. *Nature* 374, 327–337.
- Shapiro, M.D., Marks, M.E., Peichel, C.L., Blackman, B.K., Nereng, K.S., Jónsson, B., Schluter, D., and Kingsley, D.M. (2004). Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature* 428, 717–723.
- Sharer, J.D. (2005). The adenine nucleotide translocase type 1 (ANT1): a new factor in mitochondrial disease. *IUBMB Life* 57, 607–614.
- Sharma, S., and Namrata, S. (2004). Association of schizophrenia and mental retardation with facio-scapulohumeral muscular dystrophy. *Neurol. India* 52, 104–105.
- Sharma, V., Harafuji, N., Belayew, A., and Chen, Y.-W. (2013). DUX4 differentially regulates transcriptomes of human rhabdomyosarcoma and mouse C2C12 cells. *PloS One* 8, e64691.
- Sharp, A.J., Stathaki, E., Migliavacca, E., Brahmachary, M., Montgomery, S.B., Dupre, Y., and Antonarakis, S.E. (2011). DNA methylation profiles of human active and inactive X chromosomes. *Genome Res.* 21, 1592–1600.
- Shaw, N.D., Brand, H., Kupchinsky, Z.A., Bengani, H., Plummer, L., Jones, T.I., Erdin, S., Williamson, K.A., Rainger, J., Stortchevoi, A., et al. (2017). SMCHD1 mutations associated with a rare muscular dystrophy can also cause isolated arhinia and Bosma arhinia microphthalmia syndrome. *Nat. Genet.* 49, 238–248.
- Shay, J.W. (2003). Telomerase therapeutics: telomeres recognized as a DNA damage signal: commentary re: K. Kraemer et al., antisense-mediated hTERT inhibition specifically reduces the growth of human bladder cancer cells. *Clin. Cancer Res.*, 9: 3794–3800, 2003. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 9, 3521–3525.
- Shay, J.W., and Wright, W.E. (2011). Role of telomeres and telomerase in cancer. *Semin. Cancer Biol.* 21, 349–353.
- Shen, B., Zhang, J., Wu, H., Wang, J., Ma, K., Li, Z., Zhang, X., Zhang, P., and Huang, X. (2013). Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. *Cell Res.* 23, 720–723.
- Shi, Y., and Manley, J.L. (2015). The end of the message: multiple protein–RNA interactions define the mRNA polyadenylation site. *Genes Dev.* 29, 889–897.
- Shields, C.L., Zahler, J., Falk, N., Furuta, M., Eagle, R.C., Espinosa, L.E.B., Fischer, P.R., and Shields, J.A. (2007). Neovascular glaucoma from advanced Coats disease as the initial manifestation of facioscapulohumeral dystrophy in a 2-year-old child. *Arch. Ophthalmol. Chic. Ill* 1960 125, 840–842.
- Shopland, L.S., Johnson, C.V., Byron, M., McNeil, J., and Lawrence, J.B. (2003). Clustering of multiple specific genes and gene-rich R-bands around SC-35 domains: evidence for local euchromatic neighborhoods. *J. Cell Biol.* 162, 981–990.
- Silva, G., Poirot, L., Galetto, R., Smith, J., Montoya, G., Duchateau, P., and Pâques, F. (2011). Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy. *Curr. Gene Ther.* 11, 11–27.

- Simons, M., and Mlodzik, M. (2008). Planar cell polarity signaling: from fly development to human disease. *Annu. Rev. Genet.* *42*, 517–540.
- Skouloudaki, K., Puetz, M., Simons, M., Courbard, J.-R., Boehlke, C., Hartleben, B., Engel, C., Moeller, M.J., Englert, C., Bollig, F., et al. (2009). Scribble participates in Hippo signaling and is required for normal zebrafish pronephros development. *Proc. Natl. Acad. Sci. U. S. A.* *106*, 8579–8584.
- Smith, Z.D., and Meissner, A. (2013). DNA methylation: roles in mammalian development. *Nat. Rev. Genet.* *14*, 204–220.
- Smith, J., Bibikova, M., Whitby, F.G., Reddy, A.R., Chandrasegaran, S., and Carroll, D. (2000). Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. *Nucleic Acids Res.* *28*, 3361–3369.
- Smith, J., Grizot, S., Arnould, S., Duclert, A., Epinat, J.-C., Chames, P., Prieto, J., Redondo, P., Blanco, F.J., Bravo, J., et al. (2006). A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. *Nucleic Acids Res.* *34*, e149.
- Smith, T.G., Van Hateren, N., Tickle, C., and Wilson, S.A. (2007). The expression of Fat-1 cadherin during chick limb development. *Int. J. Dev. Biol.* *51*, 173–176.
- Snider, L., Asawachaicharn, A., Tyler, A.E., Geng, L.N., Petek, L.M., Maves, L., Miller, D.G., Lemmers, R.J.L.F., Winokur, S.T., Tawil, R., et al. (2009). RNA transcripts, miRNA-sized fragments and proteins produced from D4Z4 units: new candidates for the pathophysiology of facioscapulohumeral dystrophy. *Hum. Mol. Genet.* *18*, 2414–2430.
- Snider, L., Geng, L.N., Lemmers, R.J.L.F., Kyba, M., Ware, C.B., Nelson, A.M., Tawil, R., Filippova, G.N., van der Maarel, S.M., Tapscott, S.J., et al. (2010). Facioscapulohumeral dystrophy: incomplete suppression of a retrotransposed gene. *PLoS Genet.* *6*, e1001181.
- Soleimani, V.D., Punch, V.G., Kawabe, Y., Jones, A.E., Palidwor, G.A., Porter, C.J., Cross, J.W., Carvajal, J.J., Kockx, C.E.M., van IJcken, W.F.J., et al. (2012). Transcriptional dominance of Pax7 in adult myogenesis is due to high-affinity recognition of homeodomain motifs. *Dev. Cell* *22*, 1208–1220.
- Song, E., Lee, S.-K., Wang, J., Ince, N., Ouyang, N., Min, J., Chen, J., Shankar, P., and Lieberman, J. (2003). RNA interference targeting Fas protects mice from fulminant hepatitis. *Nat. Med.* *9*, 347–351.
- Song, J., Teplova, M., Ishibe-Murakami, S., and Patel, D.J. (2012). Structure-based mechanistic insights into DNMT1-mediated maintenance DNA methylation. *Science* *335*, 709–712.
- Sonntag, F., Schmidt, K., and Kleinschmidt, J.A. (2010). A viral assembly factor promotes AAV2 capsid formation in the nucleolus. *Proc. Natl. Acad. Sci. U. S. A.* *107*, 10220–10225.
- Sonoda, E., Hohegger, H., Saberi, A., Taniguchi, Y., and Takeda, S. (2006). Differential usage of non-homologous end-joining and homologous recombination in double strand break repair. *DNA Repair* *5*, 1021–1029.
- Sood, P., Krek, A., Zavolan, M., Macino, G., and Rajewsky, N. (2006). Cell-type-specific signatures of microRNAs on target mRNA expression. *Proc. Natl. Acad. Sci. U. S. A.* *103*, 2746–2751.
- Sopko, R., and McNeill, H. (2009). The skinny on Fat: an enormous cadherin that regulates cell adhesion, tissue growth, and planar cell polarity. *Curr. Opin. Cell Biol.* *21*, 717–723.
- Spector, D.L. (2001). Nuclear domains. *J. Cell Sci.* *114*, 2891–2893.



- Spilianakis, C.G., Lalioti, M.D., Town, T., Lee, G.R., and Flavell, R.A. (2005). Interchromosomal associations between alternatively expressed loci. *Nature* *435*, 637–645.
- Spurlock, G., Jim, H.-P., and Upadhyaya, M. (2010). Confirmation that the specific SSLP microsatellite allele 4qA161 segregates with fascioscapulohumeral muscular dystrophy (FSHD) in a cohort of multiplex and simplex FSHD families. *Muscle Nerve* *42*, 820–821.
- Srinivas, S., Watanabe, T., Lin, C.-S., William, C.M., Tanabe, Y., Jessell, T.M., and Costantini, F. (2001). Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Dev. Biol.* *1*, 4.
- Stadler, G., Chen, J.C., Wagner, K., Robin, J.D., Shay, J.W., Emerson, C.P., Jr, and Wright, W.E. (2011). Establishment of clonal myogenic cell lines from severely affected dystrophic muscles - CDK4 maintains the myogenic population. *Skelet. Muscle* *1*, 12.
- Stadler, G., Rahimov, F., King, O.D., Chen, J.C.J., Robin, J.D., Wagner, K.R., Shay, J.W., Emerson, C.P., Jr, and Wright, W.E. (2013). Telomere position effect regulates DUX4 in human facioscapulohumeral muscular dystrophy. *Nat. Struct. Mol. Biol.* *20*, 671–678.
- Stahnke, S., Lux, K., Uhrig, S., Kreppel, F., Hösel, M., Coutelle, O., Ogris, M., Hallek, M., and Büning, H. (2011). Intrinsic phospholipase A2 activity of adeno-associated virus is involved in endosomal escape of incoming particles. *Virology* *409*, 77–83.
- Statland, J.M., Sacconi, S., Farmakidis, C., Donlin-Smith, C.M., Chung, M., and Tawil, R. (2013). Coats syndrome in facioscapulohumeral dystrophy type 1: frequency and D4Z4 contraction size. *Neurology* *80*, 1247–1250.
- Stein, C.A., Subasinghe, C., Shinozuka, K., and Cohen, J.S. (1988). Physicochemical properties of phosphorothioate oligodeoxynucleotides. *Nucleic Acids Res.* *16*, 3209–3221.
- Stepien, G., Torroni, A., Chung, A.B., Hodge, J.A., and Wallace, D.C. (1992). Differential expression of adenine nucleotide translocator isoforms in mammalian tissues and during muscle cell differentiation. *J. Biol. Chem.* *267*, 14592–14597.
- Stevenson, W.G., Perloff, J.K., Weiss, J.N., and Anderson, T.L. (1990). Facioscapulohumeral muscular dystrophy: evidence for selective, genetic electrophysiologic cardiac involvement. *J. Am. Coll. Cardiol.* *15*, 292–299.
- Stieger, K., Cronin, T., Bennett, J., and Rolling, F. (2011). Adeno-associated virus mediated gene therapy for retinal degenerative diseases. *Methods Mol. Biol. Clifton NJ* *807*, 179–218.
- Stoddard, B.L. (2005). Homing endonuclease structure and function. *Q. Rev. Biophys.* *38*, 49–95.
- Strätling, W.H., and Yu, F. (1999). Origin and roles of nuclear matrix proteins. Specific functions of the MAR-binding protein MeCP2/ARBP. *Crit. Rev. Eukaryot. Gene Expr.* *9*, 311–318.
- Strub, K., Galli, G., Busslinger, M., and Birnstiel, M.L. (1984). The cDNA sequences of the sea urchin U7 small nuclear RNA suggest specific contacts between histone mRNA precursor and U7 RNA during RNA processing. *EMBO J.* *3*, 2801–2807.
- Stübgen, J.P. (2007). Facioscapulohumeral muscular dystrophy. A quantitative electromyographic study. *Electromyogr. Clin. Neurophysiol.* *47*, 175–182.
- Sugisaki, H., and Kanazawa, S. (1981). New restriction endonucleases from *Flavobacterium okeanokoites* (FokI) and *Micrococcus luteus* (MluI). *Gene* *16*, 73–78.

- Suhonen, J., Ray, J., Blömer, U., Gage, F.H., and Kaspar, B. (2006). Ex vivo and in vivo gene delivery to the brain. *Curr. Protoc. Hum. Genet. Chapter 13*, Unit 13.3.
- Summerford, C., and Samulski, R.J. (1998). Membrane-associated heparan sulfate proteoglycan is a receptor for adeno-associated virus type 2 virions. *J. Virol.* 72, 1438–1445.
- Sun, X., and Zhang, N. (2010). Cationic polymer optimization for efficient gene delivery. *Mini Rev. Med. Chem.* 10, 108–125.
- Sun, C.-Y.J., van Koningsbruggen, S., Long, S.W., Straasheijm, K., Klooster, R., Jones, T.I., Bellini, M., Levesque, L., Briehar, W.M., van der Maarel, S.M., et al. (2011). Facioscapulohumeral muscular dystrophy region gene 1 is a dynamic RNA-associated and actin-bundling protein. *J. Mol. Biol.* 411, 397–416.
- Sun, H.B., Shen, J., and Yokota, H. (2000). Size-dependent positioning of human chromosomes in interphase nuclei. *Biophys. J.* 79, 184–190.
- Szeto, D.P., Rodriguez-Esteban, C., Ryan, A.K., O’Connell, S.M., Liu, F., Kioussi, C., Gleiberman, A.S., Izpisua-Belmonte, J.C., and Rosenfeld, M.G. (1999). Role of the Bicoid-related homeodomain factor Pitx1 in specifying hindlimb morphogenesis and pituitary development. *Genes Dev.* 13, 484–494.
- Tajbakhsh, S., Rocancourt, D., Cossu, G., and Buckingham, M. (1997). Redefining the genetic hierarchies controlling skeletal myogenesis: Pax-3 and Myf-5 act upstream of MyoD. *Cell* 89, 127–138.
- Takeichi, M. (1995). Morphogenetic roles of classic cadherins. *Curr. Opin. Cell Biol.* 7, 619–627.
- Tam, R., Smith, K.P., and Lawrence, J.B. (2004). The 4q subtelomere harboring the FSHD locus is specifically anchored with peripheral heterochromatin unlike most human telomeres. *J. Cell Biol.* 167, 269–279.
- Tan, M., Luo, H., Lee, S., Jin, F., Yang, J.S., Montellier, E., Buchou, T., Cheng, Z., Rousseaux, S., Rajagopal, N., et al. (2011). Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* 146, 1016–1028.
- Tanabe, H., Habermann, F.A., Solovei, I., Cremer, M., and Cremer, T. (2002). Non-random radial arrangements of interphase chromosome territories: evolutionary considerations and functional implications. *Mutat. Res.* 504, 37–45.
- Tanaka, M., Hale, L.A., Amores, A., Yan, Y.-L., Cresko, W.A., Suzuki, T., and Postlethwait, J.H. (2005). Developmental genetic basis for the evolution of pelvic fin loss in the pufferfish *Takifugu rubripes*. *Dev. Biol.* 281, 227–239.
- Tang, M., Li, Y., Zhang, X., Deng, T., Zhou, Z., Ma, W., and Songyang, Z. (2014). Structural Maintenance of Chromosomes Flexible Hinge Domain Containing 1 (SMCHD1) Promotes Non-homologous End Joining and Inhibits Homologous Recombination Repair upon DNA Damage. *J. Biol. Chem.* 289, 34024–34032.
- Tang, Z.Z., Zheng, S., Nikolic, J., and Black, D.L. (2009). Developmental control of CaV1.2 L-type calcium channel splicing by Fox proteins. *Mol. Cell. Biol.* 29, 4757–4765.
- Tanguay, R.L., and Gallie, D.R. (1996). Translational efficiency is regulated by the length of the 3’ untranslated region. *Mol. Cell. Biol.* 16, 146–156.

- Tanoue, T., and Takeichi, M. (2004). Mammalian Fat1 cadherin regulates actin dynamics and cell-cell contact. *J. Cell Biol.* *165*, 517–528.
- Tasca, G., Monforte, M., Iannaccone, E., Laschena, F., Ottaviani, P., Leoncini, E., Boccia, S., Galluzzi, G., Pelliccioni, M., Masciullo, M., et al. (2014). Upper girdle imaging in facioscapulohumeral muscular dystrophy. *PloS One* *9*, e100292.
- Tassin, A., Laoudj-Chenivesse, D., Vanderplanck, C., Barro, M., Charron, S., Ansseau, E., Chen, Y.-W., Mercier, J., Coppée, F., and Belayew, A. (2013). DUX4 expression in FSHD muscle cells: how could such a rare protein cause a myopathy? *J. Cell. Mol. Med.* *17*, 76–89.
- Tawil, R. (2004). Facioscapulohumeral muscular dystrophy. *Curr. Neurol. Neurosci. Rep.* *4*, 51–54.
- Tawil, R. (2008). Facioscapulohumeral muscular dystrophy. *Neurother. J. Am. Soc. Exp. Neurother.* *5*, 601–606.
- Tawil, R., Forrester, J., Griggs, R.C., Mendell, J., Kissel, J., McDermott, M., King, W., Weiffenbach, B., and Figlewicz, D. (1996). Evidence for anticipation and association of deletion size with severity in facioscapulohumeral muscular dystrophy. The FSH-DY Group. *Ann. Neurol.* *39*, 744–748.
- Tawil, R., Kissel, J.T., Heatwole, C., Pandya, S., Gronseth, G., Benatar, M., Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology, and Practice Issues Review Panel of the American Association of Neuromuscular & Electrodiagnostic Medicine (2015). Evidence-based guideline summary: Evaluation, diagnosis, and management of facioscapulohumeral muscular dystrophy: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology and the Practice Issues Review Panel of the American Association of Neuromuscular & Electrodiagnostic Medicine. *Neurology* *85*, 357–364.
- Taylor, M.V. (2002). Muscle Differentiation: How Two Cells Become One. *Curr. Biol.* *12*, R224–R228.
- Tebas, P., Stein, D., Tang, W.W., Frank, I., Wang, S.Q., Lee, G., Spratt, S.K., Surosky, R.T., Giedlin, M.A., Nichol, G., et al. (2014). Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N. Engl. J. Med.* *370*, 901–910.
- Tepass, U., Truong, K., Godt, D., Ikura, M., and Peifer, M. (2000). Cadherins in embryonic and neural morphogenesis. *Nat. Rev. Mol. Cell Biol.* *1*, 91–100.
- Teplova, M., Minasov, G., Tereshko, V., Inamati, G.B., Cook, P.D., Manoharan, M., and Egli, M. (1999). Crystal structure and improved antisense properties of 2'-O-(2-methoxyethyl)-RNA. *Nat. Struct. Biol.* *6*, 535–539.
- Thakore, P.I., D'Ippolito, A.M., Song, L., Safi, A., Shivakumar, N.K., Kabadi, A.M., Reddy, T.E., Crawford, G.E., and Gersbach, C.A. (2015). Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat. Methods* *12*, 1143–1149.
- Thierry, A., and Dujon, B. (1992). Nested chromosomal fragmentation in yeast using the meganuclease I-Sce I: a new method for physical mapping of eukaryotic genomes. *Nucleic Acids Res.* *20*, 5625–5631.
- Thijssen, P.E., Balog, J., Yao, Z., Pham, T.P., Tawil, R., Tapscott, S.J., and Van der Maarel, S.M. (2014). DUX4 promotes transcription of FRG2 by directly activating its promoter in facioscapulohumeral muscular dystrophy. *Skelet. Muscle* *4*, 19.

- Thomas, J.O., and Travers, A.A. (2001). HMG1 and 2, and related “architectural” DNA-binding proteins. *Trends Biochem. Sci.* 26, 167–174.
- Thomas, M.J., and Seto, E. (1999). Unlocking the mechanisms of transcription factor YY1: are chromatin modifying enzymes the key? *Gene* 236, 197–208.
- Thomas, N.S.T., Wiseman, K., Spurlock, G., MacDonald, M., Ustek, D., and Upadhyaya, M. (2007). A large patient study confirming that facioscapulohumeral muscular dystrophy (FSHD) disease expression is almost exclusively associated with an FSHD locus located on a 4qA-defined 4qter subtelomere. *J. Med. Genet.* 44, 215–218.
- Tian, B., Hu, J., Zhang, H., and Lutz, C.S. (2005). A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res.* 33, 201–212.
- Tinsley, J., Deconinck, N., Fisher, R., Kahn, D., Phelps, S., Gillis, J.M., and Davies, K. (1998). Expression of full-length utrophin prevents muscular dystrophy in mdx mice. *Nat. Med.* 4, 1441–1444.
- Tintignac, L.A., Lagirand, J., Batonnet, S., Sirri, V., Leibovitch, M.P., and Leibovitch, S.A. (2005). Degradation of MyoD mediated by the SCF (MAFbx) ubiquitin ligase. *J. Biol. Chem.* 280, 2847–2856.
- Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol. Cell* 10, 1453–1465.
- Tonini, M.M.O., Passos-Bueno, M.R., Cerqueira, A., Matioli, S.R., Pavanello, R., and Zatz, M. (2004). Asymptomatic carriers and gender differences in facioscapulohumeral muscular dystrophy (FSHD). *Neuromuscul. Disord. NMD* 14, 33–38.
- Tonini, M.M.O., Lemmers, R.J.L.F., Pavanello, R.C.M., Cerqueira, A.M.P., Frants, R.R., van der Maarel, S.M., and Zatz, M. (2006). Equal proportions of affected cells in muscle and blood of a mosaic carrier of facioscapulohumeral muscular dystrophy. *Hum. Genet.* 119, 23–28.
- Tratschin, J.D., West, M.H., Sandbank, T., and Carter, B.J. (1984). A human parvovirus, adeno-associated virus, as a eucaryotic vector: transient expression and encapsidation of the procaryotic gene for chloramphenicol acetyltransferase. *Mol. Cell. Biol.* 4, 2072–2081.
- Trevisan, C.P., Pastorello, E., Armani, M., Angelini, C., Nante, G., Tomelleri, G., Tonin, P., Mongini, T., Palmucci, L., Galluzzi, G., et al. (2006). Facioscapulohumeral muscular dystrophy and occurrence of heart arrhythmia. *Eur. Neurol.* 56, 1–5.
- Trevisan, C.P., Pastorello, E., Tomelleri, G., Vercelli, L., Bruno, C., Scapolan, S., Siciliano, G., and Comacchio, F. (2008). Facioscapulohumeral muscular dystrophy: hearing loss and other atypical features of patients with large 4q35 deletions. *Eur. J. Neurol. Off. J. Eur. Fed. Neurol. Soc.* 15, 1353–1358.
- Trojer, P., and Reinberg, D. (2007). Facultative heterochromatin: is there a distinctive molecular signature? *Mol. Cell* 28, 1–13.
- Tros de Ilarduya, C., Sun, Y., and Düzgüneş, N. (2010). Gene delivery by lipoplexes and polyplexes. *Eur. J. Pharm. Sci. Off. J. Eur. Fed. Pharm. Sci.* 40, 159–170.
- Tsien, F., Sun, B., Hopkins, N.E., Vedanarayanan, V., Figlewicz, D., Winokur, S., and Ehrlich, M. (2001). Methylation of the FSHD syndrome-linked subtelomeric repeat in normal and FSHD cell cultures and tissues. *Mol. Genet. Metab.* 74, 322–331.

- Tsumagari, K., Chang, S.-C., Lacey, M., Baribault, C., Chittur, S.V., Sowden, J., Tawil, R., Crawford, G.E., and Ehrlich, M. (2011). Gene expression during normal and FSHD myogenesis. *BMC Med. Genomics* 4, 67.
- Tupler, R., Berardinelli, A., Barbierato, L., Frants, R., Hewitt, J.E., Lanzi, G., Maraschio, P., and Tiepolo, L. (1996a). Monosomy of distal 4q does not cause facioscapulohumeral muscular dystrophy. *J. Med. Genet.* 33, 366–370.
- Tupler, R., Berardinelli, A., Barbierato, L., Frants, R., Hewitt, J.E., Lanzi, G., Maraschio, P., and Tiepolo, L. (1996b). Monosomy of distal 4q does not cause facioscapulohumeral muscular dystrophy. *J. Med. Genet.* 33, 366–370.
- Tupler, R., Barbierato, L., Memmi, M., Sewry, C.A., De Grandis, D., Maraschio, P., Tiepolo, L., and Ferlini, A. (1998). Identical de novo mutation at the D4F104S1 locus in monozygotic male twins affected by facioscapulohumeral muscular dystrophy (FSHD) with different clinical expression. *J. Med. Genet.* 35, 778–783.
- Turki, A., Hayot, M., Carnac, G., Pillard, F., Passerieux, E., Bommart, S., Raynaud de Mauverger, E., Hugon, G., Pincemail, J., Pietri, S., et al. (2012). Functional muscle impairment in facioscapulohumeral muscular dystrophy is correlated with oxidative stress and mitochondrial dysfunction. *Free Radic. Biol. Med.* 53, 1068–1079.
- Tyler, F.H., and Stephens, F.E. (1950). Studies in disorders of muscle. II. Clinical manifestations and inheritance of facioscapulohumeral dystrophy in a large family. *Ann. Intern. Med.*
- Tyner, S.D., Venkatachalam, S., Choi, J., Jones, S., Ghebraniou, N., Igelmann, H., Lu, X., Soron, G., Cooper, B., Brayton, C., et al. (2002). p53 mutant mice that display early ageing-associated phenotypes. *Nature* 415, 45–53.
- Uitte de Willige, S., Rietveld, I.M., De Visser, M.C.H., Vos, H.L., and Bertina, R.M. (2007). Polymorphism 10034C>T is located in a region regulating polyadenylation of FGG transcripts and influences the fibrinogen gamma'/gammaA mRNA ratio. *J. Thromb. Haemost. JTH* 5, 1243–1249.
- Upadhyaya, M., Lunt, P.W., Sarfarazi, M., Broadhead, W., Daniels, J., Owen, M., and Harper, P.S. (1990). DNA marker applicable to presymptomatic and prenatal diagnosis of facioscapulohumeral disease. *Lancet Lond. Engl.* 336, 1320–1321.
- Upadhyaya, M., Lunt, P.W., Sarfarazi, M., Broadhead, W., Daniels, J., Owen, M., and Harper, P.S. (1991). A closely linked DNA marker for facioscapulohumeral disease on chromosome 4q. *J. Med. Genet.* 28, 665–671.
- Urnov, F.D., Miller, J.C., Lee, Y.-L., Beausejour, C.M., Rock, J.M., Augustus, S., Jamieson, A.C., Porteus, M.H., Gregory, P.D., and Holmes, M.C. (2005). Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* 435, 646–651.
- Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S., and Gregory, P.D. (2010). Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* 11, 636–646.
- Urtizberea, J.A. (1997). Results of a French survey of individuals with FSHD. *FSH Watch Publ. FSH Soc. USA* 37–40.
- Vacek, M., Sazani, P., and Kole, R. (2003). Antisense-mediated redirection of mRNA splicing. *Cell. Mol. Life Sci. CMLS* 60, 825–833.

- Vakoc, C.R., Mandat, S.A., Olenchok, B.A., and Blobel, G.A. (2005). Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. *Mol. Cell* 19, 381–391.
- Vanamee, E.S., Santagata, S., and Aggarwal, A.K. (2001). FokI requires two specific DNA sites for cleavage. *J. Mol. Biol.* 309, 69–78.
- Vanderplanck, C., Anseau, E., Charron, S., Stricwant, N., Tassin, A., Laoudj-Chenivesse, D., Wilton, S.D., Coppée, F., and Belayew, A. (2011). The FSHD atrophic myotube phenotype is caused by DUX4 expression. *PloS One* 6, e26820.
- Vasale, J., Jocson, M., Boyar, F., Liaquat, K., Hoffman, C., Sulcova, V., Tsao, D., Hensley, K., Liu, Y., Chan, P., et al. (2015). Molecular Combing compared to Southern Blot in Measuring D4Z4 Contractions in FSHD (P2.027). *Neurology* 84, P2.027.
- Vassetzky, Y., Lemaitre, J.M., and Méchali, M. (2000a). Specification of chromatin domains and regulation of replication and transcription during development. *Crit. Rev. Eukaryot. Gene Expr.* 10, 31–38.
- Vassetzky, Y.S., Bogdanova, A.N., and Razin, S.V. (2000b). Analysis of the chicken DNA fragments that contain structural sites of attachment to the nuclear matrix: DNA-matrix interactions and replication. *J. Cell. Biochem.* 79, 1–14.
- Vergnaud, G., and Denoeud, F. (2000). Minisatellites: mutability and genome architecture. *Genome Res.* 10, 899–907.
- Vetrini, F., and Ng, P. (2010). Gene Therapy with Helper-Dependent Adenoviral Vectors: Current Advances and Future Perspectives. *Viruses* 2, 1886–1917.
- Vickers, T.A., Wyatt, J.R., Burckin, T., Bennett, C.F., and Freier, S.M. (2001). Fully modified 2' MOE oligonucleotides redirect polyadenylation. *Nucleic Acids Res.* 29, 1293–1299.
- Voet, N., Bleijenberg, G., Hendriks, J., de Groot, I., Padberg, G., van Engelen, B., and Geurts, A. (2014). Both aerobic exercise and cognitive-behavioral therapy reduce chronic fatigue in FSHD: an RCT. *Neurology* 83, 1914–1922.
- Voet, N.B.M., van der Kooi, E.L., Riphagen, I.I., Lindeman, E., van Engelen, B.G.M., and Geurts, A.C.H. (2013). Strength training and aerobic exercise training for muscle disease. *Cochrane Database Syst. Rev.* 7, CD003907.
- Wada, T., Hara, M., Taneda, T., Qingfu, C., Takata, R., Moro, K., Takeda, K., Kishimoto, T., and Handa, H. (2012). Antisense morpholino targeting just upstream from a poly(A) tail junction of maternal mRNA removes the tail and inhibits translation. *Nucleic Acids Res.* 40, e173.
- Wah, D.A., Hirsch, J.A., Dorner, L.F., Schildkraut, I., and Aggarwal, A.K. (1997). Structure of the multimodular endonuclease FokI bound to DNA. *Nature* 388, 97–100.
- Wahlestedt, C., Salmi, P., Good, L., Kela, J., Johnsson, T., Hökfelt, T., Broberger, C., Porreca, F., Lai, J., Ren, K., et al. (2000). Potent and nontoxic antisense oligonucleotides containing locked nucleic acids. *Proc. Natl. Acad. Sci. U. S. A.* 97, 5633–5638.
- Wallace, L.M., Garwick, S.E., Mei, W., Belayew, A., Coppee, F., Ladner, K.J., Guttridge, D., Yang, J., and Harper, S.Q. (2011a). DUX4, a candidate gene for facioscapulohumeral muscular dystrophy, causes p53-dependent myopathy in vivo. *Ann. Neurol.* 69, 540–552.

- Wallace, L.M., Garwick-Coppens, S.E., Tupler, R., and Harper, S.Q. (2011b). RNA interference improves myopathic phenotypes in mice over-expressing FSHD region gene 1 (FRG1). *Mol. Ther. J. Am. Soc. Gene Ther.* *19*, 2048–2054.
- Wallace, L.M., Liu, J., Domire, J.S., Garwick-Coppens, S.E., Guckes, S.M., Mendell, J.R., Flanigan, K.M., and Harper, S.Q. (2012). RNA interference inhibits DUX4-induced muscle toxicity in vivo: implications for a targeted FSHD therapy. *Mol. Ther. J. Am. Soc. Gene Ther.* *20*, 1417–1423.
- Walther, T.N., Wittop Koning, T.H., Schümperli, D., and Müller, B. (1998). A 5'-3' exonuclease activity involved in forming the 3' products of histone pre-mRNA processing in vitro. *RNA N. Y. N* *4*, 1034–1046.
- Walton, E.L., Francastel, C., and Velasco, G. (2011). Maintenance of DNA methylation: Dnmt3b joins the dance. *Epigenetics* *6*, 1373–1377.
- Wang, Y.X., and Rudnicki, M.A. (2012). Satellite cells, the engines of muscle repair. *Nat. Rev. Mol. Cell Biol.* *13*, 127–133.
- Wang, B., Li, J., and Xiao, X. (2000). Adeno-associated virus vector carrying human minidystrophin genes effectively ameliorates muscular dystrophy in mdx mouse model. *Proc. Natl. Acad. Sci. U. S. A.* *97*, 13714–13719.
- Wang, D., Zhong, L., Nahid, M.A., and Gao, G. (2014). The potential of adeno-associated viral vectors for gene delivery to muscle tissue. *Expert Opin. Drug Deliv.* *11*, 345–364.
- Wang, H., Maurano, M.T., Qu, H., Varley, K.E., Gertz, J., Pauli, F., Lee, K., Canfield, T., Weaver, M., Sandstrom, R., et al. (2012). Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Res.* *22*, 1680–1688.
- Wang, L., Figueredo, J., Calcedo, R., Lin, J., and Wilson, J.M. (2007). Cross-presentation of adeno-associated virus serotype 2 capsids activates cytotoxic T cells but does not render hepatocytes effective cytolytic targets. *Hum. Gene Ther.* *18*, 185–194.
- Wang, Z., Zhu, T., Qiao, C., Zhou, L., Wang, B., Zhang, J., Chen, C., Li, J., and Xiao, X. (2005). Adeno-associated virus serotype 8 efficiently delivers genes to muscle and heart. *Nat. Biotechnol.* *23*, 321–328.
- Watts, J.K., and Corey, D.R. (2012). Gene silencing by siRNAs and antisense oligonucleotides in the laboratory and the clinic. *J. Pathol.* *226*, 365–379.
- Wei, J.-X., Yang, J., Sun, J.-F., Jia, L.-T., Zhang, Y., Zhang, H.-Z., Li, X., Meng, Y.-L., Yao, L.-B., and Yang, A.-G. (2009). Both Strands of siRNA Have Potential to Guide Posttranscriptional Gene Silencing in Mammalian Cells. *PLoS ONE* *4*.
- Weiffenbach, B., Bagley, R., Falls, K., Hyser, C., Storvick, D., Jacobsen, S.J., Schultz, P., Mendell, J., Willems van Dijk, K., and Milner, E.C. (1992). Linkage analyses of five chromosome 4 markers localizes the facioscapulohumeral muscular dystrophy (FSHD) gene to distal 4q35. *Am. J. Hum. Genet.* *51*, 416–423.
- Wengel, J. (1999). Synthesis of 3'-C- and 4'-C-Branched Oligodeoxynucleotides and the Development of Locked Nucleic Acid (LNA). *Acc. Chem. Res.* *32*, 301–310.
- West, A.G., Gaszner, M., and Felsenfeld, G. (2002). Insulators: many functions, many mechanisms. *Genes Dev.* *16*, 271–288.

- Wiedenheft, B., Sternberg, S.H., and Doudna, J.A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 482, 331–338.
- Wiestner, A., Tehrani, M., Chiorazzi, M., Wright, G., Gibellini, F., Nakayama, K., Liu, H., Rosenwald, A., Muller-Hermelink, H.K., Ott, G., et al. (2007). Point mutations and genomic deletions in CCND1 create stable truncated cyclin D1 mRNAs that are associated with increased proliferation rate and shorter survival. *Blood* 109, 4599–4606.
- Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell* 75, 855–862.
- Wijmenga, C., Frants, R.R., Brouwer, O.F., van der Klift, H.M., Meera Khan, P., and Padberg, G.W. (1990a). Facioscapulohumeral muscular dystrophy gene in Dutch families is not linked to markers for familial adenomatous polyposis on the long arm of chromosome 5. *J. Neurol. Sci.* 95, 225–229.
- Wijmenga, C., Frants, R.R., Brouwer, O.F., Moerer, P., Weber, J.L., and Padberg, G.W. (1990b). Location of facioscapulohumeral muscular dystrophy gene on chromosome 4. *Lancet* 336, 651–653.
- Wijmenga, C., Padberg, G.W., Moerer, P., Wiegant, J., Liem, L., Brouwer, O.F., Milner, E.C., Weber, J.L., van Ommen, G.B., and Sandkuyf, L.A. (1991). Mapping of facioscapulohumeral muscular dystrophy gene to chromosome 4q35-qter by multipoint linkage analysis and in situ hybridization. *Genomics* 9, 570–575.
- Wijmenga, C., Hewitt, J.E., Sandkuyf, L.A., Clark, L.N., Wright, T.J., Dauwerse, H.G., Gruter, A.M., Hofker, M.H., Moerer, P., and Williamson, R. (1992). Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. *Nat. Genet.* 2, 26–30.
- Wijmenga, C., Winokur, S.T., Padberg, G.W., Skraastad, M.I., Altherr, M.R., Wasmuth, J.J., Murray, J.C., Hofker, M.H., and Frants, R.R. (1993). The human skeletal muscle adenine nucleotide translocator gene maps to chromosome 4q35 in the region of the facioscapulohumeral muscular dystrophy locus. *Hum. Genet.* 92, 198–203.
- Williams, A.S., Ingledue, T.C., Kay, B.K., and Marzluff, W.F. (1994). Changes in the stem-loop at the 3' terminus of histone mRNA affects its nucleocytoplasmic transport and cytoplasmic regulation. *Nucleic Acids Res.* 22, 4660–4666.
- Wilson, J.M. (2005). Gendicine: the first commercial gene therapy product. *Hum. Gene Ther.* 16, 1014–1015.
- Wilson, A.S., Power, B.E., and Molloy, P.L. (2007). DNA hypomethylation and human diseases. *Biochim. Biophys. Acta* 1775, 138–162.
- Wilson, D., Sheng, G., Lecuit, T., Dostatni, N., and Desplan, C. (1993). Cooperative dimerization of paired class homeo domains on DNA. *Genes Dev.* 7, 2120–2134.
- Winokur, S.T., Bengtsson, U., Feddersen, J., Mathews, K.D., Weiffenbach, B., Bailey, H., Markovich, R.P., Murray, J.C., Wasmuth, J.J., and Altherr, M.R. (1994). The DNA rearrangement associated with facioscapulohumeral muscular dystrophy involves a heterochromatin-associated repetitive element: implications for a role of chromatin structure in the pathogenesis of the disease. *Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* 2, 225–234.
- Winokur, S.T., Bengtsson, U., Vargas, J.C., Wasmuth, J.J., and Altherr, M.R. (1997). The evolutionary distribution and structural organization of the homeobox-containing repeat D4Z4 indicates a functional role for the ancestral copy in the FSHD region. *Hum. Mol. Genet.* 6, 502.



- Winokur, S.T., Chen, Y.-W., Masny, P.S., Martin, J.H., Ehmsen, J.T., Tapscott, S.J., van der Maarel, S.M., Hayashi, Y., and Flanigan, K.M. (2003a). Expression profiling of FSHD muscle supports a defect in specific stages of myogenic differentiation. *Hum. Mol. Genet.* *12*, 2895–2907.
- Winokur, S.T., Barrett, K., Martin, J.H., Forrester, J.R., Simon, M., Tawil, R., Chung, S.-A., Masny, P.S., and Figlewicz, D.A. (2003b). Facioscapulohumeral muscular dystrophy (FSHD) myoblasts demonstrate increased susceptibility to oxidative stress. *Neuromuscul. Disord. NMD* *13*, 322–333.
- Winston, J., Duerden, L., Mort, M., Frayling, I.M., Rogers, M.T., and Upadhyaya, M. (2014). Identification of two novel SMCHD1 sequence variants in families with FSHD-like muscular dystrophy. *Eur. J. Hum. Genet. EJHG*.
- Wirth, T., Parker, N., and Ylä-Herttuala, S. (2013). History of gene therapy. *Gene* *525*, 162–169.
- Wohlgemuth, M., van der Kooi, E.L., van Kesteren, R.G., van der Maarel, S.M., and Padberg, G.W. (2004). Ventilatory support in facioscapulohumeral muscular dystrophy. *Neurology* *63*, 176–178.
- Wokke, J.H.J., van Doorn, P.A., Hoogendijk, J.E., and De Visser, M. (2013). *Neuromuscular Disease : A case-based approach*. Camb. Univ. Press.
- Wolfe, S.A., Greisman, H.A., Ramm, E.I., and Pabo, C.O. (1999). Analysis of zinc fingers optimized via phage display: evaluating the utility of a recognition code. *J. Mol. Biol.* *285*, 1917–1934.
- Wolfe, S.A., Nekludova, L., and Pabo, C.O. (2000). DNA recognition by Cys2His2 zinc finger proteins. *Annu. Rev. Biophys. Biomol. Struct.* *29*, 183–212.
- Wright, T.J., Wijmenga, C., Clark, L.N., Frants, R.R., Williamson, R., and Hewitt, J.E. (1993a). Fine mapping of the FSHD gene region orientates the rearranged fragment detected by the probe p13E-11. *Hum. Mol. Genet.* *2*, 1673–1678.
- Wright, T.J., Wijmenga, C., Clark, L.N., Frants, R.R., Williamson, R., and Hewitt, J.E. (1993b). Fine mapping of the FSHD gene region orientates the rearranged fragment detected by the probe p13E-11. *Hum. Mol. Genet.* *2*, 1673–1678.
- Wu, S.-L., Li, G.-Z., Chou, C.-Y., Tsai, M.-S., Chen, Y.-P., Li, C.-J., Liou, G.-G., Chang, W.-W., Chen, S.-L., and Wang, S.-H. (2014). Double homeobox gene, *Duxbl*, promotes myoblast proliferation and abolishes myoblast differentiation by blocking MyoD transactivation. *Cell Tissue Res.*
- Wu, Z., Asokan, A., and Samulski, R.J. (2006). Adeno-associated virus serotypes: vector toolkit for human gene therapy. *Mol. Ther. J. Am. Soc. Gene Ther.* *14*, 316–327.
- Wu, Z.-Y., Wang, Z.-Q., Murong, S.-X., and Wang, N. (2004). FSHD in Chinese population: characteristics of translocation and genotype-phenotype correlation. *Neurology* *63*, 581–583.
- Wuebbles, R.D., Hanel, M.L., and Jones, P.L. (2009). FSHD region gene 1 (FRG1) is crucial for angiogenesis linking FRG1 to facioscapulohumeral muscular dystrophy-associated vasculopathy. *Dis. Model. Mech.* *2*, 267–274.
- Wuebbles, R.D., Long, S.W., Hanel, M.L., and Jones, P.L. (2010). Testing the effects of FSHD candidate gene expression in vertebrate muscle development. *Int. J. Clin. Exp. Pathol.* *3*, 386–400.
- Xiao, A., Wang, Z., Hu, Y., Wu, Y., Luo, Z., Yang, Z., Zu, Y., Li, W., Huang, P., Tong, X., et al. (2013). Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. *Nucleic Acids Res.* *41*, e141.

- Xie, Q., Bu, W., Bhatia, S., Hare, J., Somasundaram, T., Azzi, A., and Chapman, M.S. (2002). The atomic structure of adeno-associated virus (AAV-2), a vector for human gene therapy. *Proc. Natl. Acad. Sci. U. S. A.* *99*, 10405–10410.
- Yamaguchi, T., Miki, Y., and Yoshida, K. (2010). The c-Abl tyrosine kinase stabilizes Pitx1 in the apoptotic response to DNA damage. *Apoptosis Int. J. Program. Cell Death* *15*, 927–935.
- Yamazaki, Y., Nango, M., Matsuura, M., Hasegawa, Y., Hasegawa, M., and Oku, N. (2000). Polycation liposomes, a novel nonviral gene transfer system, constructed from cetylated polyethylenimine. *Gene Ther.* *7*, 1148–1155.
- Yang, Y.T., and McElligott, M.A. (1989). Multiple actions of beta-adrenergic agonists on skeletal muscle and adipose tissue. *Biochem. J.* *261*, 1–10.
- Yang, L., Li, J., and Xiao, X. (2011). Directed evolution of adeno-associated virus (AAV) as vector for muscle gene therapy. *Methods Mol. Biol. Clifton NJ* *709*, 127–139.
- Yao, Z., Snider, L., Balog, J., Lemmers, R.J.L.F., Van Der Maarel, S.M., Tawil, R., and Tapscott, S.J. (2014). DUX4-induced gene expression is the major molecular signature in FSHD skeletal muscle. *Hum. Mol. Genet.*
- Yeom, K.-H., Lee, Y., Han, J., Suh, M.R., and Kim, V.N. (2006). Characterization of DGCR8/Pasha, the essential cofactor for Drosha in primary miRNA processing. *Nucleic Acids Res.* *34*, 4622–4629.
- Yew, N.S., and Scheule, R.K. (2005). Toxicity of cationic lipid-DNA complexes. *Adv. Genet.* *53*, 189–214.
- Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev.* *17*, 3011–3016.
- Yi, Y., Jong Noh, M., and Hee Lee, K. (2011). Current Advances in Retroviral Gene Therapy. *Curr. Gene Ther.* *11*, 218–228.
- Yin, H., Kanasty, R.L., Eltoukhy, A.A., Vegas, A.J., Dorkin, J.R., and Anderson, D.G. (2014). Non-viral vectors for gene-based therapy. *Nat. Rev. Genet.* *15*, 541–555.
- Ying, G.G., Proost, P., van Damme, J., Bruschi, M., Introna, M., and Golay, J. (2000). Nucleolin, a novel partner for the Myb transcription factor family that regulates their activity. *J. Biol. Chem.* *275*, 4152–4158.
- Ylä-Herttuala, S. (2012). Endgame: glybera finally recommended for approval as the first gene therapy drug in the European union. *Mol. Ther. J. Am. Soc. Gene Ther.* *20*, 1831–1832.
- Yu, M., Sha, H., Gao, Y., Zeng, H., Zhu, M., and Gao, X. (2006). Alternative 3' UTR polyadenylation of Bzw1 transcripts display differential translation efficiency and tissue-specific expression. *Biochem. Biophys. Res. Commun.* *345*, 479–485.
- Yu, Z., Ren, M., Wang, Z., Zhang, B., Rong, Y.S., Jiao, R., and Gao, G. (2013). Highly Efficient Genome Modifications Mediated by CRISPR/Cas9 in *Drosophila*. *Genetics* *195*, 289–291.
- Yusufzai, T.M., Tagami, H., Nakatani, Y., and Felsenfeld, G. (2004). CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. *Mol. Cell* *13*, 291–298.
- Zacchigna, S., Zentilin, L., and Giacca, M. (2014). Adeno-Associated Virus Vectors as Therapeutic and Investigational Tools in the Cardiovascular System. *Circ. Res.* *114*, 1827–1846.

- Zaiss, A.-K., Liu, Q., Bowen, G.P., Wong, N.C.W., Bartlett, J.S., and Muruve, D.A. (2002). Differential Activation of Innate Immune Responses by Adenovirus and Adeno-Associated Virus Vectors. *J. Virol.* 76, 4580–4590.
- Zakian, V.A. (2012). Telomeres: The beginnings and ends of eukaryotic chromosomes. *Exp. Cell Res.* 318, 1456–1460.
- Zambrowicz, B.P., Imamoto, A., Fiering, S., Herzenberg, L.A., Kerr, W.G., and Soriano, P. (1997). Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. *Proc. Natl. Acad. Sci. U. S. A.* 94, 3789–3794.
- Zamecnik, P.C., and Stephenson, M.L. (1978). Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide. *Proc. Natl. Acad. Sci. U. S. A.* 75, 280–284.
- Zatz, M., Marie, S.K., Passos-Bueno, M.R., Vainzof, M., Campiotto, S., Cerqueira, A., Wijmenga, C., Padberg, G., and Frants, R. (1995). High proportion of new mutations and possible anticipation in Brazilian facioscapulohumeral muscular dystrophy families. *Am. J. Hum. Genet.* 56, 99–105.
- Zatz, M., Marie, S.K., Cerqueira, A., Vainzof, M., Pavanello, R.C., and Passos-Bueno, M.R. (1998). The facioscapulohumeral muscular dystrophy (FSHD1) gene affects males more severely and more frequently than females. *Am. J. Med. Genet.* 77, 155–161.
- Zeng, W., de Greef, J.C., Chen, Y.-Y., Chien, R., Kong, X., Gregson, H.C., Winokur, S.T., Pyle, A., Robertson, K.D., Schmieging, J.A., et al. (2009). Specific loss of histone H3 lysine 9 trimethylation and HP1gamma/cohesin binding at D4Z4 repeats is associated with facioscapulohumeral dystrophy (FSHD). *PLoS Genet.* 5, e1000559.
- Zeng, W., Chen, Y.-Y., Newkirk, D.A., Wu, B., Balog, J., Kong, X., Ball, A.R., Jr, Zanotti, S., Tawil, R., Hashimoto, N., et al. (2014). Genetic and epigenetic characteristics of FSHD-associated 4q and 10q D4Z4 that are distinct from non-4q/10q D4Z4 homologs. *Hum. Mutat.*
- Zeng, Y., Zhang, C., and Su, Q. (2001). [Gene diagnosis of facioscapulohumeral muscular dystrophy]. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi Zhonghua Yixue Yichuanxue Zazhi Chin. J. Med. Genet.* 18, 213–215.
- Zhang, G., Gao, X., Song, Y., Vollmer, R., Stolz, D., Gasiorowski, J., Dean, D., and Liu, D. (2004). Hydroporation as the mechanism of hydrodynamic delivery. *Gene Ther.* 11, 675–682.
- Zhang, X.-X., McIntosh, T.J., and Grinstaff, M.W. (2012). Functional lipids and lipoplexes for improved gene delivery. *Biochimie* 94, 42–58.
- Zhang, X.Y., Loflin, P.T., Gehrke, C.W., Andrews, P.A., and Ehrlich, M. (1987). Hypermethylation of human DNA sequences in embryonal carcinoma cells and somatic tissues but not in sperm. *Nucleic Acids Res.* 15, 9429–9449.
- Zhang, Y., Jurkowska, R., Soeroes, S., Rajavelu, A., Dhayalan, A., Bock, I., Rathert, P., Brandt, O., Reinhardt, R., Fischle, W., et al. (2010). Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Res.* 38, 4246–4253.
- Zhang, Y., King, O.D., Rahimov, F., Jones, T.I., Ward, C.W., Kerr, J.P., Liu, N., Emerson, C.P., Kunkel, L.M., Partridge, T.A., et al. (2014). Human skeletal muscle xenograft as a new preclinical model for muscle disorders. *Hum. Mol. Genet.* 23, 3180–3188.
- Zhang, Y., Lee, J.K., Toso, E.A., Lee, J.S., Choi, S.H., Slattery, M., Aihara, H., and Kyba, M. (2016). DNA-binding sequence specificity of DUX4. *Skelet. Muscle* 6, 8.

- Zhao, W., and Manley, J.L. (1996). Complex alternative RNA processing generates an unexpected diversity of poly(A) polymerase isoforms. *Mol. Cell. Biol.* *16*, 2378–2386.
- Zhao, J., Hyman, L., and Moore, C. (1999). Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol. Mol. Biol. Rev. MMBR* *63*, 405–445.
- Zhou, H., Janghra, N., Mitrpant, C., Dickinson, R.L., Anthony, K., Price, L., Eperon, I.C., Wilton, S.D., Morgan, J., and Muntoni, F. (2013). A novel morpholino oligomer targeting ISS-N1 improves rescue of severe spinal muscular atrophy transgenic mice. *Hum. Gene Ther.* *24*, 331–342.
- Zhou, H.-L., Luo, G., Wise, J.A., and Lou, H. (2014). Regulation of alternative splicing by local histone modifications: potential roles for RNA-guided mechanisms. *Nucleic Acids Res.* *42*, 701–713.
- Zhu, C.-H., Mouly, V., Cooper, R.N., Mamchaoui, K., Bigot, A., Shay, J.W., Di Santo, J.P., Butler-Browne, G.S., and Wright, W.E. (2007). Cellular senescence in human myoblasts is overcome by human telomerase reverse transcriptase and cyclin-dependent kinase 4: consequences in aging muscle and therapeutic strategies for muscular dystrophies. *Aging Cell* *6*, 515–523.
- Zincarelli, C., Soltys, S., Rengo, G., and Rabinowitz, J.E. (2008). Analysis of AAV serotypes 1-9 mediated gene expression and tropism in mice after systemic injection. *Mol. Ther. J. Am. Soc. Gene Ther.* *16*, 1073–1080.
- Zouvelou, V., Manta, P., Kalfakis, N., Evdokimidis, I., and Vassilopoulos, D. (2009). Asymptomatic elevation of serum creatine kinase leading to the diagnosis of 4q35 facioscapulohumeral muscular dystrophy. *J. Clin. Neurosci. Off. J. Neurosurg. Soc. Australas.* *16*, 1218–1219.