



UNIVERSITAT DE  
BARCELONA

# Unraveling potential disease modifiers of myotonic dystrophy type 1

Emma Agathe Koehorts

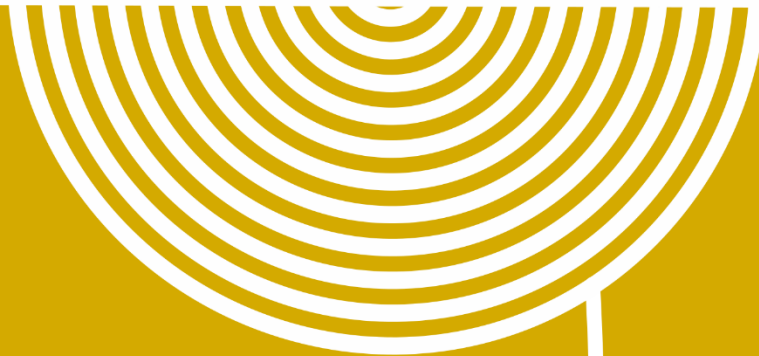
**ADVERTIMENT.** La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX ([www.tdx.cat](http://www.tdx.cat)) i a través del Dipòsit Digital de la UB ([diposit.ub.edu](http://diposit.ub.edu)) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

**ADVERTENCIA.** La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR ([www.tdx.cat](http://www.tdx.cat)) y a través del Repositorio Digital de la UB ([diposit.ub.edu](http://diposit.ub.edu)) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

**WARNING.** On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX ([www.tdx.cat](http://www.tdx.cat)) service and by the UB Digital Repository ([diposit.ub.edu](http://diposit.ub.edu)) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

# UNRAVELING POTENTIAL DISEASE MODIFIERS OF MYOTONIC DYSTROPHY TYPE 1

---



**Emma Koehorst**  
Doctoral Thesis







**Doctoral Thesis**

**Unraveling Potential Disease Modifiers of  
Myotonic Dystrophy Type 1**

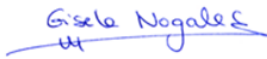
Emma Agathe Koehorst

GERMANS TRIAS I PUJOL RESEARCH INSTITUTE (IGTP)

Neuromuscular and Neuropediatric Research Group

PhD in Biomedicine at the Universitat de Barcelona

Dr. Gisela Nogales Gadea  
(Supervisor)



Emma Agathe Koehorst  
(Doctoral Student)



Dr. Mònica Suelves  
(Co-Supervisor)



Dr. Josep Saura Martí  
(Tutor)





voor mijn Papa en Mama

voor mijn Broertjes

voor Ons







# **ACKNOWLEDGMENTS**



## 'IT ALWAYS SEEMS IMPOSSIBLE, UNTIL IT IS DONE'

---

---

Six years ago, I could not have imagined sitting here, writing the acknowledgments for my doctoral thesis. I had reached a point in life where my disease had taken over and I, quite frankly, had given up on any kind of 'normal' life, let alone a PhD life. With the help of family and friends I found the perseverance, but more importantly a way to shape life to me, instead of the other way around, and here we are today. I am immensely grateful and proud and I could not have done this without any of you.

**Gisela**, thank you for seeing my strengths through my disability, and giving me a place to grow not only as a researcher, but also as a human. I started with you as a Master student and leaving as a well-rounded researcher, made possible by your support and guidance. You helped me to develop my strengths, identify and improve my weaknesses and challenged me to always reach for more, for better. Thank you for going on this amazing and crazy ride with me. I think it is safe to say, mission accomplished!

**Mònica**, thank you for stepping up when it was much needed, as another 'neuromuscular lab' baby joined the team. I am so grateful to have had you as my second supervisor. You gave me another perspective and view on research. Your endless drive and motivation helped me through the last part and your personality was a joy to be around. No phone call was too much, no question too difficult. You are an amazing person, with an incredible brain. Thank you for sharing it with me.

To all the neurologists of our group, **Alicia, Alba, Andrea, Giuseppe, Miriam, Guillem** and **Jaume**, thank you for teaching me the clinical aspects of Myotonic Dystrophy and many other neuromuscular disorders. You guys were always there whenever I had a question and were a major help in developing this doctoral thesis. I have learned so much during the weekly meetings and thank you for always adapting to this Dutch girl and her bad Spanish.

**Darren** and **Sarah**, it has been such a pleasure to have met you and to have worked with you. You made me feel incredibly welcome in your lab in Glasgow and it has been an experience I will not soon forget. Thank you for teaching me and for always lending a helping hand, even after the stay ended.

My family away from home, always and forever the **Buuumbaas**!!!! My little shits, I cannot put into words how much you mean to me, but I will try, haha! I am so grateful for our hour long coffee breaks, our 'lab outings' to Mallorca, the pizza parties, the drinks, the fun, the laughs. I could not have wished for better colleagues and friends. Although soon we will be in very different places on this earth, know that you guys will always have a very

special place in my heart and I am always there for you guys! **Alfonsina**, the girl who will always be younger than me (even though she is not, for the record). You made me feel so welcome the first time I stepped foot inside the lab. It was just us two in the beginning and you have taught me so much during that time. You were an amazing colleague, but more importantly, became an amazing friend. I have loved our 'work' trips together, from Bilbao, to Sweden and let's not forget Glasgow. Only you can convince me that sleeping in a car at -4°C is fun. Although it is safe to say I would never do it again, it is an experience I wouldn't want to miss for the world. Girl, or as I like to call you, my favorite little shit, it has been such a crazy and fun ride with you, filled with so much laughter, long talks and tears. We will always be connected, wherever we are. **Ian**, the technician from heaven. Being the only guy in the group, we did not always take it easy on you, with our oversharing. Luckily, after a while you fit right in and became the 'little brother' to us girls. It has been a pleasure working with you. Your open and kind personality, always willing to lend a hand or give a big bear hug when it was very much needed. Your 'frying eggs' portrayal still makes me laugh out loud and wherever I see a tutu, I think of you. Promise me you come and see the cows in the Netherlands, they are waiting for you! **Judit**, the girl that did not say a single word when we first met, but who has become one of my dearest friends. Thank you for confiding in me and for always being a listening ear. You always saw it when I had one of my pain flairs and I cannot tell you how much the little squeeze of your hand on my shoulder have meant to me. No words necessary to say 'I see you and I am here'. Thank you for the many coffees, and the patience while I was on one of my seemingly endless rants. I am so grateful to have had you by my side during this whole process, but especially at the end, when it was just you and me of the Buumbaas left. I could not have done it without you and I will miss seeing your face every day as your lab mate, but I will settle for a lifelong friend.

Thank you to my other colleagues, collaborators and the many a student that passed by the lab. **Renato**, although we did not get to spend a lot of time in the lab together, the short time that we did work together was a blast. Thank you for your helping hand in what seemed to be an endless stream of PCRs and sequencing. I also much enjoyed our coffee breaks, and the daily 'spilling the tea with Renato'. I am sure you are going to make a great PhD fellow and wishing you all the luck! **Julia**, thank you for your help with the final stretches in the methylation project, while my contract had already finished and not being able to do it myself. It has been a great help. **Marina**, my carpool buddy to work. Thank you for setting aside your fear of English speaking and welcoming me into your space. I have very much enjoyed our little car rides where we would talk about the day we had. Especially on tough days, I stepped out of the car a little lighter as when I stepped in. Thank you for that! **Shelly, Marta, Jorge, Adrián, Jaume, Aida, and Soheil**, each one of you has passed through our lab as interns for various times and projects and it has been a pleasure teaching you and also learning from you. **Rafa** and **Ana Pilar**, thank you for lending us a vector that was essential in this thesis project and for always answering my (many) questions. It has been a pleasure collaborating with you guys. **Jakub**, thank you for your help and guidance in the

world of microscopy. I have learned so much during our microscopy sessions and my images would not have turned out this way if it wasn't for you. Thank you to **Pilar** and the other members of the Genomics platform, **Anna, Irina** and **Xavi**, for their help in the sequencing analysis and their always friendly faces when I came downstairs with yet another big batch of samples. To **Nùria, Marc** and **Alexia**, thank you for your help with the Western Blots and cell culture. It was great to have such a nice team in the same lab, always up for a nice little chat. Thank you **Danni** for saving me many a time when I ran out of medium or RNA kits, and lending me some from your group. It was much appreciated. **Mar**, thank you for navigating through the world of bacteria with me, without you I would still be clueless, haha! A big thanks to **Natalia, David, Elisabeth, Quim, Harvey** and **Roser**. You guys are an essential part of the IGTP institute and I have much appreciated how you were always trying to find solutions when I found myself in a pinch. Thank you for your patience and your help. **Natalia**, in addition to being an amazing lab manager, our little coffee breaks and your infectious laugh were little highlights in my day. Thank you! Talking about essential beings, **Emmi, Maribel** and **Montse**, the three most amazing receptionists you can find. Your faces were the first thing I saw every morning and they always carried smiles. Although verbal communication was tough due to my horrible Spanish, your energies were always friendly and warm. Thank you for that!

To my friends scattered across this globe, I cannot put into words how much you guys mean to me. We have popped in and out of each other's lives, but every single one of you has helped me through this amazing and intense time. **Berend**, thank you for your timeless friendship and for being a friend when I couldn't. **JoJo**, we did not meet under the best of circumstances, but I am so grateful that you stepped into that apartment of crazy. Thank you for the brunches, the glasses of wine, the Toad-in-the-hole diner extravaganza, the ALL THOSE food truck festivals, the talks, the laughs, the tears, and for keeping me grounded when needed. Missing you loads. My Spanish class friends, who became so so so much more, **Carolyn** and **Deborah**. Mastering the Spanish language was not really in the cards for us, but boy did we have fun during those classes and long after. Schmuckies, you kept me sane during the times that this thesis and life just became too much. Thank you for all those glasses of wine, the salsa nights with really strong mojito's, the Sunday brunches, the sun-filled beach days, the belly-aching laughter and the long talks. You ladies have always been in my corner, rooting for me during this entire process and I am so grateful that I have met you. **Martha**, de leeftijdsloze yoga lerares, die maar niet ouder lijkt te worden. Dankje voor de yoga sessies, de praat sessies en voor je warme hart. Ik heb me altijd enorm 'gezien' gevoeld door je en je hebt me geholpen om langzaam mijn lijf als een veilige haven te gaan zien, niet als de vijand. Daar ben ik je zo dankbaar voor. **Carolien**, met je vrolijke rode krullenbos en je vriendelijke lach, altijd even checken hoe het hier in Barcelona gaat. Dankje daarvoor en natuurlijk bedankt voor de vele body pump lessen die je me toegestuurd hebt. Het was een fijn alternatief voor jouw energieke lessen en was hetgene waar ik alle frustratie en stress in kwijt kon. Super bedankt!

A special thanks to **Ramon Langen** and **Harry Gosker**. Ook al waren jullie officieel geen onderdeel van mijn promotie onderzoek, jullie hebben wel een belangrijke rol gespeeld in het beginnen ervan. Dankzij jullie hulp en positieve houding kon ik aan een deeltijd stage in het buitenland beginnen, iets wat ik daarvoor niet voor mogelijk had gehouden. Die stage heeft geleid tot dit promotie onderzoek en mijn tijd in Barcelona. Ik ben jullie daar erg dankbaar voor.

Last, but certainly not least, mijn **familie**. Mijn rotsen in de branding, zonder wie ik dit absoluut niet had gekund. Wat heb ik jullie meegenomen op een rollercoaster ride. In typische Emma stijl ging deze weg met veel hoogte en dieptepunten en jullie waren er voor elke stap. Dankje voor de eindeloze steun, het niet veroordelen en het luisterende oor wanneer het weer eens tegen zat. Mijn **mama**, zo'n geweldig voorbeeld van een mooie, sterke vrouw, die me geleerd heeft dat gevoeligheid en kwetsbaarheid je niet zwak maakt, integendeel. Dankje voor je eindeloze geduld, de lange, lange facetime gesprekken en de dikke knuffels wanneer ik eindelijk wel eens een keer in Nederland was. Ik hou zielsveel van je en ben je zo dankbaar dat je dit samen met mijn aangegaan bent, want ik weet dat het niet makkelijk was. Mijn **papa**, de reden waarom ik in het onderzoek gegaan ben, ook al heeft hij mij gewaarschuwd om het niet te doen, haha. Een man van weinig woorden, maar elk woord was altijd raak en zette me aan het denken. Dankje voor het veilige gevoel dat je me hebt gegeven, dat ik altijd een plek had wat ik thuis kon noemen en waar ik naar terug kon, mocht het nodig zijn. Die veilige thuishaven heeft me de kracht gegeven om er op uit te gaan en me te ontwikkelen tot de persoon die ik nu ben. Mijn **broertjes**, waar ik als grote zus vroeger de beschermeling was, begint dat langzaam om te draaien of misschien beter, gelijkwaardig te worden. Jullie hebben me in deze tijd altijd een geborgen gevoel gegeven, ook al waren jullie niet fysiek hier. Ik heb jullie zien opgroeien tot twee mooie mannen, en ik ben zo trots dat ik jullie zus ben en zo blij dat ik jullie naast me had tijdens dit proces. Ook al is bellen niet ons sterkste punt, het weten dat jullie er waren mocht ik jullie nodig hebben, was een ontzettend fijn gevoel. Onze broers en zus weekenden waar voor mij altijd een hoogtepunt in het jaar. **Dirk**, ik ben enorm dankbaar voor jouw positiviteit en jouw vermogen om zware situaties licht te maken. **Gijs**, waar Dirk dingen lichter kon maken, was jij de persoon die mijn struggles kon zien en me een veilige plek gaf waar ik ze kon laten. Zonder oordeel, alleen maar warmte, dankjewel daarvoor.



# **TABLE OF CONTENTS**





<b>GLOSSARY</b>	19
Abbreviations	21
Synonyms	24
<b>INTRODUCTION</b>	25
1. Myotonic Dystrophy Type 1: a Brief Summary	27
2. The Clinical Spectrum of Myotonic Dystrophy Type 1	27
2.1 Congenital	28
2.2 Childhood	28
2.3 Juvenile	28
2.4 Adult	29
2.5 Late-onset	29
3. Genetics of Myotonic Dystrophy Type 1	29
3.1 Somatic Instability	30
3.2 Anticipation	30
4. RNA Toxicity: Leading Pathomechanism	32
4.1 A Toxic RNA Gain-of-Function Disorder	32
4.2 Splicing Defects: a Dual Role for MBNL1 and CELF1	32
5. Disease Models for Myotonic Dystrophy type 1	35
6. Therapeutic Approaches in Myotonic Dystrophy Type 1	37
7. Disease Modifiers	39
7.1 Disease Modifiers at Genetic level	39
7.2 Disease Modifiers at Transcriptional level	43
7.3 Disease Modifiers at Protein level	46
7.4 Disease Modifiers at Epigenetic level	48
7.5 Disease Modifiers at non-coding RNA level	56
<b>OVERVIEW</b>	59
<b>OBJECTIVES</b>	63

Table of Contents

<b>RESULTS</b>	67
Director's Report Regarding Thesis by Papers	70
Chapter I	73
Chapter II	93
Chapter III	127
Chapter IV	175
<b>GENERAL DISCUSSION</b>	199
<b>CONCLUSIONS</b>	211
<b>BIBLIOGRAPHY</b>	215
<b>FUNDING</b>	245
<b>APPENDIX</b>	249
Appendix I. List of Other Publications	251



# **GLOSSARY**



---

---

## ABBREVIATIONS

5mC = 5-methylcytosine

A = adenine

APP = amyloid  $\beta$  precursor protein

ASOs = antisense oligonucleotides

ATP5MC2 = ATP synthase membrane subunit C locus 2

BIN1 = the bridging integrator 1/ amphispysin 2

BSP = bisulfite Sanger sequencing PCR

C = cytosine

C9ORF72 ALS/FTD = C9orf72 amyotrophic lateral sclerosis /frontotemporal dementia

CACNA1S = calcium voltage-gated channel subunit alpha 1 S

CAPN3 = calpain

Cas-8 = caspase 8

CDM1 = congenital myotonic dystrophy type 1

Celf = mice homologue for CELF1

CELF1/CUGBP1 = CUG-binding protein 1

CLCN1 = chloride voltage-gated channel 1

CNS = central nervous system

CpG = cytosine-guanine sites

CpGi = CpG island

CTCF = CCTC-binding factor

cTNT = cardiac troponin C

DM1 = myotonic dystrophy type 1

DM1-AS = DM1 anti-sense

DM2 = myotonic dystrophy type 2

DM-300 = mice carrying a fragment of the human *DMPK* locus containing 300 repeats

DMD = dystrophin

DMPK = dystrophia myotonica protein kinase

Dmpk = dystrophia myotonica protein kinase homolog in mice

## Glossary

DMSXL = mice carrying a fragment of the human *DMPK* locus containing 1000-1800 repeats

DMWD = dystrophia myotonica WD repeat-containing protein

DNMTs = DNA methyltransferases

DTNA = dystrobrevin alpha

ePAL = estimated progenitor allele length

FXPOI = fragile X-associated primary ovarian insufficiency

FXTAS = fragile X-associated tremor/ataxia syndrome

G = guanine

GSK3 $\beta$  = glycogen synthase kinase 3 beta

HD = Huntington's disease

hESCs = human embryonic stem cells

HSA<sup>LR</sup> = mice expressing human skeletal actin gene with long CTG repeats

HTTAS = Huntingtin antisense

iPSCs = induced pluripotent stem cells

IR = insulin receptor

IRES = internal ribosome entry site

LDB3 = LIM domain binding 3

MAPT = microtubule-associated protein tau

Mbml = mice homologue for MBNL1

MBNL = muscleblind-like

miRNA = microRNA

MSH3 = mutS homolog 3 gene

MYOM1 = myomesin 1

ALPK3 = alpha kinase 3

Myo-miR = muscle-specific microRNA

NCOR2 = nuclear receptor corepressor 2

NEB = nebulin

NFIX = nuclear factor I X

NMDAR1 = NMDA receptor 1

NTMR1 = myotubularin related protein 1

PKC = protein kinase C

PKM2 = pyruvate kinase M2

polyGln = polyglutamine

RAN = repeat-associated non ATG

RBFOX2 = RNA binding fox-1 homolog-2

RISC = RNA-induced silencing complex

RNAi = RNA interference

RyR1 = ryanodine receptor 1

SAM = S-adenyl methionine

SCA2 = spinocerebellar ataxia type 2

SCA31 = spinocerebellar ataxia type 31

SCA8 = spinocerebellar ataxia type 8

SCN5 = gene encoding for a  $\alpha$ -subunit of the cardiac voltage channel  $NaV1.5$

SERCA1 = sarcoplasmic/endoplasmic reticulum calcium ATPase 1

SERCA2 = ion channel encoded by ATP2A2

siRNA = small interfering RNA

SIX5 = SIX homeobox 5

SOS1 = SOS Ras/Rac Guanine Nucleotide Exchange Factor 1

SP-PCR = small pool polymerase chain reaction

T = thymine

TBP = TATA-box binding protein

TNNT3 = troponin T3

TP-PCR = triplet-primed polymerase chain reaction

TTN = titin

U = uracil

UTR = untranslated region



## SYNONYMS

---

---

Adult-onset DM1; classical DM1

CTGs; repeats; CTG repeats

Somatic mosaicism; somatic instability

CTG expansion size; CTG size; repeat size; expansion size; CTG repeat length

Variant repeat; interruptions



# **INTRODUCTION**

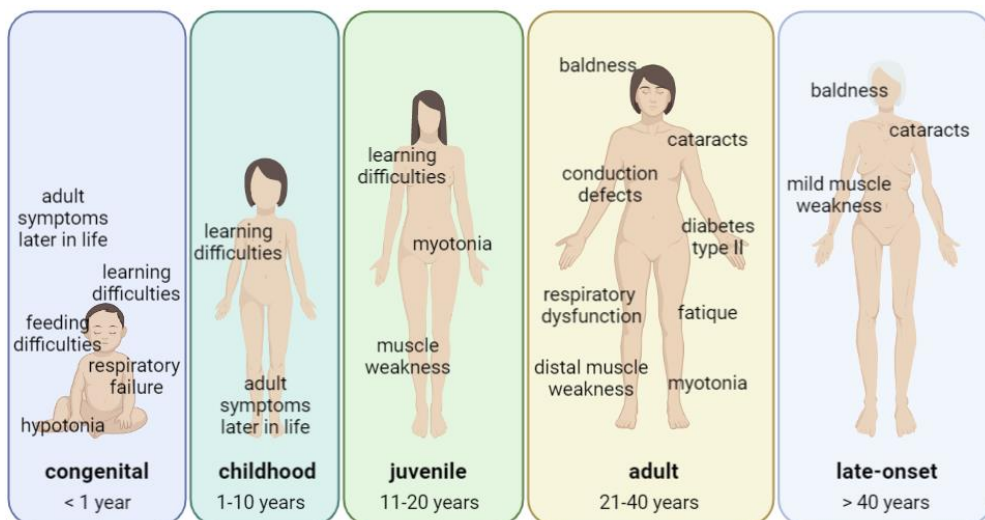


## 1. MYOTONIC DYSTROPHY TYPE 1: A BRIEF SUMMARY

Myotonic dystrophy type 1 (DM1; OMIM: 160900) is an incurable, autosomal dominant inherited muscular dystrophy with an overall prevalence of 1:8000 [1]. A more recent population-wide screening estimated a much higher prevalence, with a genetic prevalence at 4.8 per 10,000 individuals, making it one of the most common rare diseases [2]. DM1 is also known as Steinert's disease, named after its discoverer Hans Gustav Wilhelm Steinert, who first described its clinical characteristics in 1909 [3]. DM1 is viewed as one of the most variable manifestation of a monogenic disease, characterized by its wide variability in both symptomatology and age of onset. It is a progressive, multi-systemic disorder, that strongly impairs quality of life and reduces life expectancy [4,5].

## 2. THE CLINICAL SPECTRUM OF MYOTONIC DYSTROPHY TYPE 1

Due to the wide spectrum of clinical phenotype manifestation observed in DM1, in both symptomatology as well as age of onset, the disease has been categorized into five different clinical subtypes, namely congenital, childhood, juvenile, adult and late-onset DM1 [6]. The DM1 clinical triad is distal muscle weakness, myotonia and early-onset cataracts. However, this triad is not always present at diagnosis, especially in the younger subcategories. Each clinical subtype shows therefore a distinct clinical phenotype and its distinct challenges in disease management (Figure 1).



**Figure 1. Main clinical symptoms of the five DM1 subcategories.** Congenital, childhood, juvenile, adult and late-onset.

## 2.1 CONGENITAL

---

The most severe and early onset subcategory in DM1 is congenital DM1 (CDM1), with age of onset at birth or in the first year of life [7]. The incidence of CDM1 is estimated to up to 1 in 47,619 live births [8] and the mortality in neonatal period is around 30-40% [9]. CDM1 is not only the most severe form of DM1, it also presents with symptoms unique to this clinical category, which are not exhibited by other DM1 subcategories. These symptoms already start prenatally, where CDM1 is characterized by polyhydramnios, reduced fetal movement and delivery is often pre-term [10]. Neonatal manifestations include hypotonia, respiratory failure, feeding difficulties, failure to thrive and clubfoot deformities [8,11,12]. Due to severe facial weakness, affected infants have an inverted V-shaped or 'fish-shaped' upper lip. Respiratory failure is a common cause of death in these patients in the first year of life [13,14].

During childhood, congenital DM1 children show learning difficulties and delayed cognitive and motor milestones, as well as behavioral disorders [15,16]. Symptoms seen in the adult form of DM1, such as progressive myopathy, can develop from early adulthood on, but progression is usually slow [9].

## 2.2 CHILDHOOD

---

The second form of DM1 is childhood DM1, with age of onset between one and ten years of age. Diagnosis of childhood DM1 is complicated and it is often misdiagnosed due to the uncharacteristic symptom manifestation. The first signs are cognitive and learning abnormalities, rather than muscle impairment [17]. These abnormalities include internalizing disorders, borderline low intelligence and attention deficit disorders. The typical muscular signature, such as muscle myopathy and myotonia, can develop at various ages, but often do not develop until late adolescence [9,10]. There is however, a small subset of patients in which muscle wasting develops during the second decade of life, with rapid worsening and loss of ambulation. Early heart conduction abnormalities might be present, but often the severe cardio-respiratory complications arise after their thirties, similar to CDM1.

## 2.3 JUVENILE

---

Juvenile DM1 patients are an interesting clinical subgroup, as they are often placed either under the childhood or adult DM1-onset umbrella. Therefore, their clinical presentation is not as well understood. A certain overlap can be seen between this category and the other two, however, juvenile onset differs from childhood DM1 in their increased presence of myotonia and from adult-onset due to their more pronounced central nervous system involvement [6]. Juvenile DM1 patients faced greater employment challenges and needed more frequently specialized professional attention. In addition, their phenotype is in general more severe compared to adult-onset in terms of myotonia, dysphagia, muscle weakness

and facial dysmorphism. Age of onset for this category is between eleven and twenty years old.

## 2.4 ADULT

---

The adult-onset, also known as classical DM1, is the most prevalent DM1 phenotype and arises typically around the third or fourth decade of life. Core features are progressive muscle weakness, with preferential involvement of the cranial, trunk and distal limb muscles, myotonia and early-onset cataracts (<50 years) [14]. Ptosis and the involvement of facial muscles results in a characteristic myopathic facial appearance, often accompanied by premature balding [18]. Cardiac conduction defects are common and the leading cause of death in these patients [5,19,20].

Fatigue is one of the most impactful and debilitating DM1 symptoms and is caused by a combination of excessive daytime sleepiness, sleep apnea and respiratory failure [21]. Furthermore, gastrointestinal involvement with symptoms resembling irritable bowel syndrome, such as constipation, diarrhea, abdominal pain and fecal incontinence are commonly present [22,23]. Creatine kinase levels can be elevated and impairment of endocrine function is common, resulting in impaired insulin resistance, hypogonadism and thyroid dysfunction. Cognitive impairment in adult-DM1 varies widely, from no cognitive impairment to global intellectual impairment. Personality wise patients often are perceived as apathetic, with decreased emotional participation and a psychomotor delay [24].

## 2.5 LATE-ONSET

---

For late-onset DM1, symptoms manifest after the age of forty, and include low-grade muscle weakness, premature cataracts and alopecia. Due to the mild symptom display, this category often goes undiagnosed or misdiagnosed, until one of the patient's relatives receives a DM1 diagnosis with a more severe and earlier onset and the origin of the disease is investigated.

## 3. GENETICS OF MYOTONIC DYSTROPHY TYPE 1

---

---

The underlying genetic mutation for DM1 is a cytosine-thymine-guanine (CTG) expansion in the 3' untranslated region of the dystrophin myotonia protein kinase (*DMPK*) gene on chromosome 19q13.3, which can vary in length. Healthy individuals carry between 5 and 37 CTGs, whereas diseased individuals carry >50 repeats. When an individual carries between 38 and 50 repeats it is considered pre-mutational, the CTG repeat is less stable than in healthy individuals and there is a higher chance of expansion upon transmission [25–27]. Commonly, asymptomatic and/or late-onset DM1 individuals carry 50-80 CTG repeats and these relatively small expansions are called protomutations [28]. Over a hundred repeats is

associated with the adult DM1 onset, while the larger expansions (> 1000 repeats) are more likely to result in CDM1.

### 3.1 SOMATIC INSTABILITY

---

The CTG expansion is a highly unstable repeat, causing dynamic gene defects. This means that the CTG expansion continues to expand over time at varies speeds in different tissues. Ultimately, this gives rise to the phenomenon of 'somatic mosaicism', were different tissues in the same patient carry different CTG sizes and even cells of the same organ can have varying repeat sizes. The largest expansion sizes can be found in heart, skin and muscle [29,30]. The CTG expansion size is most often measured in blood, where the somatic mosaicism is highly biased toward expansions, contributing to the progressive nature of the symptomatology observed in DM1 [31–34]. However, overall the CTG expansion size is lower in blood and blood-derived cells compared to other tissues, such as skeletal muscle, skin and heart [29,30,35–37].

The degree of somatic instability is tissue-specific, age-dependent and has been correlated to the original size of the repeat, often referred to as the estimated progenitor allele length (ePAL) [32,34,38–40]. Although, there is a strong correlation between ePAL and the degree of somatic instability, not all variation in somatic instability can be accounted for by ePAL and age [31]. In addition, cases have been found where patients with the similar repeat lengths develop varies degrees of somatic instability, indicating individual-specific modifiers [31]. These modifiers are still poorly understood. It has been postulated that somatic instability is a highly heritable trait, implying a role for individual-specific trans-acting genetic modifiers [31]. The MutS homolog 3 gene (MSH3) has been proposed as one of those trans-acting modifiers. It has been shown that three polymorphisms in the MSH3 were associated with the variation in somatic instability [41]. Additionally, variant repeats have been shown to have a stabilizing effect on the CTG expansion and could therefore modify somatic instability [42–44].

### 3.2 ANTICIPATION

---

The instability of the CTG expansion is also present in germline, biased towards expansion, and leading to larger CTG expansions in successive generations. This is known as genetic anticipation, where the disease severity increases and/or age of onset decreases from one generation to the next [14]. Although in the majority of cases expansions are transmitted, the occasional contraction can also be observed. A large pedigree analysis revealed 6.4% of all transmissions results in a contraction, which was higher for paternal transmissions (10%) and lower for maternal transmission (3%) [45]. The intergenerational instability depends on the sex, the age and CTG expansion size of the transmitting parent [25,46,47]. Maternal transmission, especially when the affected mother carries between 80 and 250 repeats, results in larger expansions, often related to CDM1 offspring. In fact, CDM1

is almost exclusively maternally transmitted with an estimation of 87.5-91% [7], which is generally accompanied by CTG expansions of >1000 repeats [7,46–50]. The percentage of maternal transmission in other categories ranges from 30-58% in the other categories (Table 1). Interestingly, for paternal transmission to result in larger expansions, the CTG repeat of the affected father is often <80 repeats, suggesting higher instability in the shorter expansions in males [50,51]. These transmitted larger expansions rarely result in CDM1 offspring [8,52].

**Table 1. Overview of parental transmission of the five clinical subtypes observed in DM1.**

		Disease severity				
Age of onset	Prenatal Perinatal Neonatal	1-10 years	11-20 years	21-40 years	> 40 years	
Paternal transmission	9 - 12.5%	42 - 50%	68 - 72%	70%	70%	
Maternal transmission	87.5 - 91%	50 - 58%	28 - 32%	30%	30%	

The clinical subtypes consist of congenital, childhood, juvenile, adult and late-onset, which show a decrease in severity with increasing age of onset. Age of onset shows a direct correlation with paternal transmission and an inverse correlation with maternal transmission. *Table adapted from Lanni & Pearson (2019).*

The maternal bias for CDM1 offspring is poorly understood. Early hypotheses pointed towards the presence of maternal environmental or intrauterine factors, but none have been found so far [50,53,54]. Recently, a role for DNA methylation status around the CTG repeat has been postulated, which will be discussed in the DNA methylation section of this introduction [55]. Estimation of intergenerational instability has proven difficult due to the somatic instability present in patients. Its existence however, has been supported by the analysis of germ cells, human embryonic stem cells (hESCs) and embryos [51,56–59]. These studies have shown that intergenerational instability already occurs in the very early stages of life, with oocytes and spermatozoa showing changes in CTG repeat length [56,57]. Interestingly, 14.3% of spermatozoa showed contractions while no contractions were present in oocytes [57].



## 4. RNA TOXICITY: LEADING PATHOMECHANISM

---

The mutation underlying DM1 was found to be a CTG expansion located in the 3'-UTR of the *DMPK* gene in 1992, over eighty years after the first description of the disease [60]. At first, it was thought that the CTG expansion blocks *DMPK* mRNA or protein production, resulting in *DMPK* haploinsufficiency. The observation of decreased levels of *DMPK* mRNA and protein in DM1 muscle supported this notion [61]. However, *Dmpk*-knockout mice did not show the characteristic DM1 pathology, only mild myopathy and cardiac conduction defects in older animals, which might make it a contributing factor, but not the main pathological mechanism [62–64].

### 4.1 A TOXIC RNA GAIN-OF-FUNCTION DISORDER

---

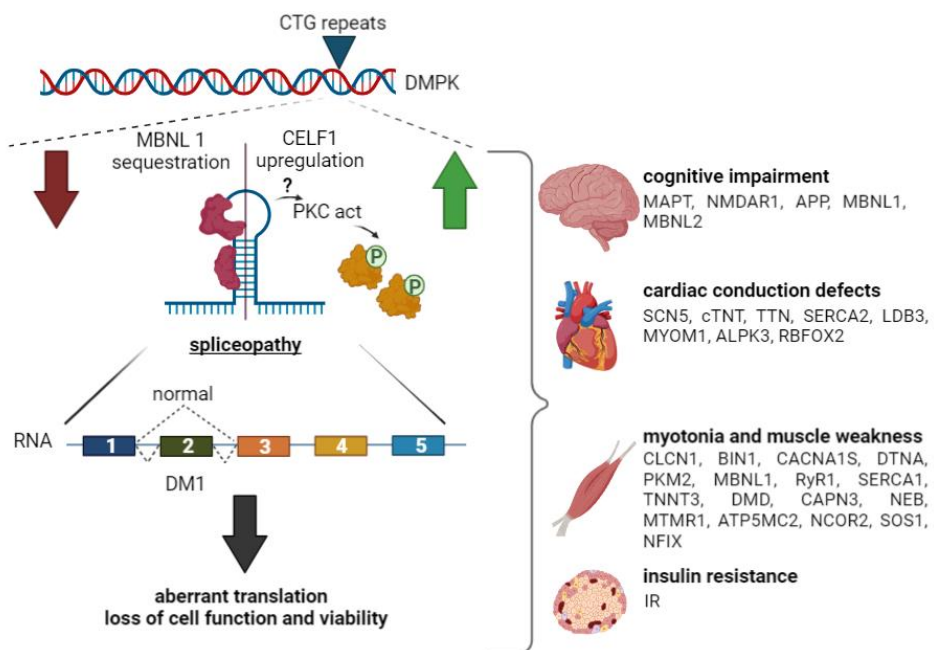
Evidence for an RNA gain-of-function mechanism was provided by the observation that although both wild type and mutant alleles were transcribed into mRNA, the mutant mRNA accumulated in the nucleus in discrete aggregates, so-called RNA foci [65]. Mouse models with the expanded repeat showed several DM1 features, including the presence of RNA foci, myotonia and a muscle histology similar to what was observed in DM1 [66–69]. Cytosine-uracil-guanine (CUG) expansions in mRNA with more than eleven repeats have been shown to form hairpin-like secondary structures, which are defined by Watson-Crick G-C base pairs, interrupted by U-U mismatches [70]. The RNA foci are located at the periphery of nuclear speckles, structures known to be enriched with small nuclear ribonucleoproteins, the spliceosome assembly factor SC35 and many other transcription and splicing regulating factors [71]. The hairpin-like secondary structures, formed by the RNA foci, were found to be able to dysregulate two important proteins, muscleblind-like 1 (MBNL1) and CUG-binding protein 1 (CELF1 or CUGBP1), which we now know stand at the base of the spliceopathy observed in DM1 (Figure 2) [72,73].

### 4.2 SPLICING DEFECTS: A DUAL ROLE FOR MBNL1 AND CELF1

---

Alternative splicing is a regulatory mechanism that modifies pre-mRNA constructs prior to translation and contributes to proteome complexity. This mechanism allows the productions of a diversity of mRNAs from a single gene by including and excluding exons from recently spliced RNA transcripts, which is tissue- and cell type-specific [74,75]. MBNL1 and CELF1 are two important regulators of alternative splicing and their dysregulation contributes vastly to the DM1 pathology, with over thirty genes described to be affected (Figure 2) [76], of which a subset is described in more detail below. MBNL1 is part of the muscleblind-like (MBNL) family, which consists of three different isoforms. MBNL1 is expressed most abundantly in skeletal muscle, but is present in most tissues. MBNL2 is almost exclusively expressed in the central nervous system and MBNL3 expression seems to be more restricted to muscle cell differentiation and regeneration [72,77]. All three isoforms

are sequestered by RNA foci and seem to play an important role in DM1 pathology [72,78,79]. MBNLs have structural similarities and exon 1, 2 and 4 encode for four zinc fingers domains, which are important for RNA binding and splicing activities [80]. Interestingly, these zinc fingers domains bind 5'-YGCY-3' motifs, which are abundant in CUG expanded RNAs [81,82]. Additionally, the MBNL1 binding site on the cardiac troponin C (cTNT) transcripts forms mismatched hairpin-like structures similar to the ones observed in the expanded repeat [83,84]. Taken together, this gives a possible explanation as to why MBNL is sequestered to RNA foci, as it seems they mimic the MBNL natural binding site. MBNL sequestration leads to a depletion of MBNL and deregulation of the alternative splicing of numerous genes, of which some can be directly linked to DM1 symptomatology [76].



**Figure 2. DM1's RNA gain-of-function overview.** The disease causing CTG expansion in the *DMPK* gene results in expanded mRNA transcripts, which form hairpin-like structure that sequester MBNL1 and hyperphosphorylation and stabilization of CELF1 through PKC activation. This results in decreased levels of MBNL1 and overexpression of CELF1, which alters the splicing of different transcripts, predominantly switching to embryonic isoforms. The incorrect splicing of several transcripts is connected to DM1 pathology. *Adapted from López-Martínez et al. (2020).* Abbreviations: MBNL1= muscleblind-like 1; CELF1= CUG-binding protein 1; PKC= protein kinase C; MAPT= microtubule-associated protein tau; NMDAR1= NMDA receptor 1; APP= amyloid  $\beta$  precursor protein; SCN5= gene encoding for a  $\alpha$ -subunit of the cardiac voltage channel  $\text{NaV1.5}$ ; cTNT= cardiac troponin C; SERCA2= ion channel encoded by ATP2A2; LDB3= LIM domain binding 3; TTN= titin; MYOM1= myomesin 1; ALPK3= alpha kinase 3; RBFOX2= RNA binding fox-1 homolog-2; CLCN1= chloride voltage-gated channel 1; BIN1= the bridging integrator 1; CACNA1S= Calcium voltage-gated channel subunit alpha1 S; DTNA= dystrobrevin alpha; PKM2= pyruvate kinase M2; RyR1= ryanodine receptor 1; SERCA1= sarcoplasmic/endoplasmic reticulum calcium ATPase 1; TNNT3= troponin T3; DMD= dystrophin; CAPN3= calpain 3; NEB= nebulin; NTMR1= Myotubularin Related Protein 1; ATP5MC2= ATP Synthase Membrane Subunit C Locus 2; NCOR2= Nuclear Receptor Corepressor 2; SOS1= SOS Ras/Rac Guanine Nucleotide Exchange Factor 1; NFIX= Nuclear Factor I X; IR= insulin receptor.

Out of the three isoforms, MBNL1 is the most contributing splicing regulator in DM1, due to its abundance in skeletal muscle and other tissues. Knockout mice of MBNL1 mimicked the molecular DM1 environment and these mice developed several DM1 features, such as myotonia and cataracts [78]. Entrapment of MBNL1 in nuclear RNA foci leads to a misregulation of alternative splicing of several transcripts. One of the best-described splicing alterations is the inclusion of exon 7a in the chloride voltage-gated channel 1 (CLCN1) transcript. This is the main chloride channel in skeletal muscle and missplicing results in one of the main features of DM1, namely myotonia [85,86]. Another example of a gene which alternative splicing is affected, is the bridging integrator 1 or amphispophysin 2 (BIN1), which encodes for a protein involved in tubular invaginations of membranes and is required for the biogenesis of muscle T tubules, structures essential for excitation-contraction coupling in skeletal muscle. Impairment due to the missplicing, resulting in the inactive form of BIN1, is coupled to muscle weakness in DM1 [87]. Intracellular calcium homeostasis plays a key role in muscle degeneration in DM1 and aberrant splicing of the calcium channel CaV1.1, encoded by the *CACNA1S* gene, is linked to contraction impairment and muscle weakness [88]. Insulin resistance in DM1 can be linked to the increased skipping of exon 11 in the insulin receptor (IR) [89,90]. In cardiac muscle, the missplicing of *SCN5* and cTNT is well studied. The *SCN5* gene encodes for an  $\alpha$ -subunit of the cardiac voltage channel NaV1.5. Missplicing results in lower conductance of the channel, which slows normal conduction and contributes to the cardiac conduction defects observed in DM1 [91]. The increased inclusion of exon 5 in cTNT might also contribute to cardiac conduction defects [89,92].

Where MBNL1 is predominantly involved in the missplicing events in skeletal muscle and heart, MBNL2 is predominantly present in the central nervous system and its depletion in knockout mice resulted in DM1-related abnormalities in the central nervous system [93]. The best-described splicing alteration due to the loss of MBNL2 is in the microtubule-associated protein tau (MAPT) found in DM1 frontal cortex samples. The abnormal expression of MAPT isoforms was associated with the presence of neurofibrillary tangles containing tau protein, suggesting a tauopathy-like degeneration of brain tissue [94,95].

CELF1 is the other main protein described in the spliceopathy found in DM1, but where MBNL1 is trapped by the nuclear RNA foci, CELF1 is not co-localizing with RNA foci. Although CELF1 can bind to the expanded repeats, it is not entrapped, but rather its expression is upregulated [96,97]. CELF1 expression is upregulated in skeletal muscle and cardiac muscle by hyperphosphorylation and stabilization through the inappropriate activation of protein kinase C (PKC) [73]. Involvement of CELF1 in DM1 pathology was supported by the observation that transgenic mice overexpressing CELF1 could reproduce splicing misregulation, as well as DM1 muscle features [98]. Although, upregulation of CELF1 has resulted in splicing defects uniquely attributed to CELF1 [99,100], interestingly, there is a big overlap with genes affected by MBNL1, but in an antagonistic manner. At the embryonic stage, MBNL1 is primarily localized in the cytoplasm, whereas CELF1 is mostly nuclear. During

development, MBNL1 nuclear levels increase, while nuclear CELF1 levels decrease, inducing an embryonic-to-adult transition of downstream splicing targets [101,102]. In DM1, due to the decrease in MBNL1 and the increase in CELF1, the embryonic state is mimicked, resulting in increased levels of embryonic isoforms in adult tissue and subsequent DM1 symptomatology. Examples of the antagonistic regulation are found for cTNT, IR and CLCN1 [85,86,90].

The main role for MBNL and CELF1 is alternative splicing regulation, but they also take part in other cellular processes, such as regulation of mRNA stability and decay and protein translation [103–106]. For example, cytoplasmic CELF1 is involved in translational regulation of proteins like Cyclin-dependent kinase inhibitor 1 and Myocyte Enhancer Factor 2A, which are involved in muscle differentiation [104,105].

In addition to the MBNL family and CELF1, other splicing regulators have been found to be affected. Increased levels of heterogeneous nuclear ribonucleoprotein H have been found in DM1 myoblasts and increased levels of Staufen1 has also been proposed to be involved in DM1 pathology [107,108]. Interestingly, the upregulation of Staufen1 might be a protective mechanism used by muscle fibers to reduce and/or delay the detrimental effects caused by MBNL1 depletion and CELF1 upregulation. Although several splicing defects can be linked to DM1 symptomatology, direct evidence of a cause-effect relationship is lacking and not the entire DM1 pathology can be explained. Moreover, the splicing changes observed in DM1 are not exclusive to the disease, as several neuromuscular disorders share these splicing changes [109,110]. Taken together, it is now clear that the DM1 pathomechanism is far more complex than originally thought.

---

## 5. DISEASE MODELS FOR MYOTONIC DYSTROPHY TYPE 1

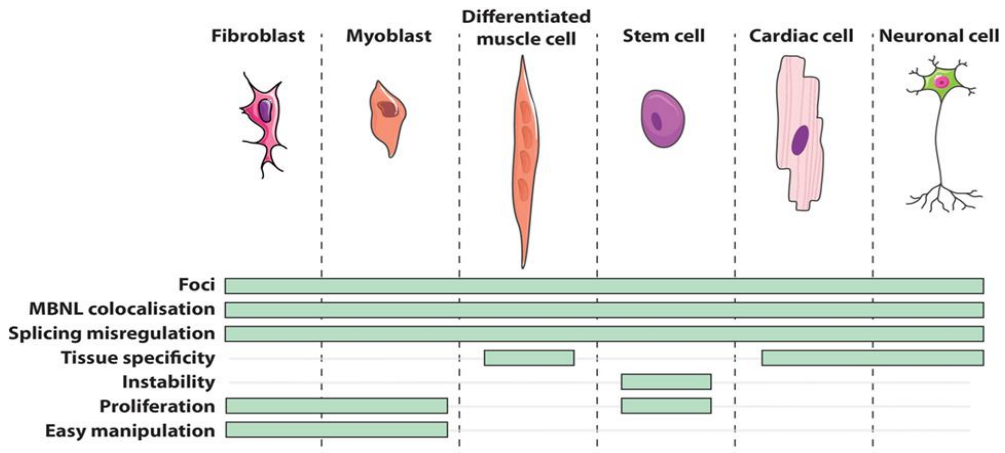
---

Disease models are quintessential in the dissection of DM1 pathology and molecular mechanisms. Over the past decades, several animal models, including mouse, fly, zebrafish, and worm have been developed to investigate DM1 pathology. For example, *Drosophila melanogaster* flies, which express a part of the *DMPK* gene with varying expanded repeat lengths using the Gal4/UAS gene expression system have been constructed. These flies form nuclear foci, present with cardiac defects, muscle wasting and eye degeneration [111,112]. Moreover, nearly 20 mouse models have contributed significant insights to DM1 pathology. Early mouse models were constructed by the inactivation of genes of the human DM1 locus, and included *DMPK* and *SIX5* homeobox 5 (*SIX5*) knockout mice [63,64,113]. These models failed to reproduce the complex DM1 pathology, as only mild myopathy and the development of cataracts was found. The hypothesis of an RNA-gain-of-function mechanism resulted in a second generation of transgenic mice, expressing the toxic RNA repeats. For example, the overexpression of toxic CUG repeats in skeletal muscle of HSA<sup>LR</sup> (human skeletal actin gene with long repeat length) mice [66]. Although mice containing the longest repeat of 250 CUG

developed a severe phenotype, including myotonia, muscle weakness and wasting was absent, limiting the usage of this mouse model. Two of the most extensively used animal models of this category are the DM-300 mice, carrying a large fragment of the human DM1 locus, containing 300 repeats and the DMSXL mice, carrying 1000-1800 CTG repeats. The latter are a result of the expansion bias upon intergenerational transmission of DM-300 mice. DM-300 mice exhibit RNA foci, myotonia, muscle weakness and insulin resistance, among others [67,114,115]. The derived DMSXL mice are characterized by a more pronounced disease phenotype, including splicing abnormalities in muscle, heart and central nervous system (CNS) and an earlier onset of symptoms [116]. This model is often used to resemble CDM1. Although these mice exhibit the multi-systemic nature of the disease, their overall phenotype is rather mild, limiting their usage. Another class of mouse models is the induction of splicing defects, through Mbnl inactivation [78,102,117] or Celf overexpression [98,105], resulting in a subset of classical DM1 symptoms.

Although animal models have significantly contributed to our knowledge of DM1 pathology, their development is an expensive undertaking and they have failed to completely reproduce the multi-systemic nature of the disease. An alternative is the use of cellular models, which can be divided into cell models expressing exogenous CTG repeats and DM1 tissue-derived cells. Several cellular models have been developed that exogenously express the toxic expanded repeat, most often by inserting repeats in the 3' UTR of a truncated DMPK gene, which are transiently or stably expressed in well-characterized human or murine cells, such as HeLa, HEK, or C2 cells [89,118,119]. These in vitro models show several DM1-associated features, such as RNA foci formation and splicing defects and have been used for therapeutic approaches and molecular mechanism identification. The in vitro models are, however, limited due to the absence of *DMPK* genomic context.

DM1 tissue-derived cells overcome this limitation and are a great tool for molecular mechanism explorations and treatment efficacy studies (Figure 3). These cells resemble the pool of CTG expansions present in DM1 patients and show RNA foci formation, and subsequent consequences leading to splicing defects and cellular dysfunction [71,72,120-123]. DM1 tissue-derived primary cell cultures are a great way to study the disease and the most often used cells are myoblasts and skin fibroblasts [121,122,124-126]. Both can be differentiated into myotubes, which are more closely related to the muscle-like environment. Skin fibroblast differentiation results rather in muscle-like cells, upon MyoD transfection, instead of real myotubes, and are used as an alternative for hard to harvest muscle biopsies [123,127,128]. Although primary cells resemble the variability observed within a patient with regards to CTG expansion size, this can also be seen as a limitation as it adds to an already complex mechanism. This is worsened by the effect of clinical subtype and age at sampling of the patient in question. Other limitations include low biopsy availability and the entering of replicative senescence after a certain amount of cell divisions.



**Figure 3. Features of DM1 cells in vitro.** Reproduced from Matloka et al. (2018).

Replicative senescence can be circumvented by the use of immortalized cell lines, such as lymphoblastoid cell lines, immortalized fibroblasts, trans-differentiated fibroblasts and myoblasts [129–131]. The immortalization process can however potentially alter cellular behavior. A solution for the limited biopsy material might be the use of DM1 human pluripotent stem cells, which consist of both hESCs [132–135] and induced pluripotent stem cells (iPSCs) [136,137], which have the potential to differentiate into a wide spectrum of cell types, such as myotubes and neuronal cells. The ethical issues surrounding the use of hESCs have resulted in the development of iPSCs, and iPSCs have already been created from DM1 primary skin fibroblasts [136,138–140]. More recently iPSCs were made from DM1 lymphoblastoid cell lines and directly from blood, which would be an excellent non-invasive way to obtain iPSCs [140,141].

## 6. THERAPEUTIC APPROACHES IN MYOTONIC DYSTROPHY TYPE 1

To date, no treatment to cure or halt disease progression exists for DM1 and management is based on symptomatic treatment. Nevertheless, great advances have been made in the field of therapeutics and several therapeutic targets have been discovered. These drugs are either targeting the root of the disease or are targeted towards ceasing specific DM1 symptoms.

One such group of potential therapies are small molecule-drug therapies, which are mainly already approved drugs in other illnesses and are potentially repurposed for usage in DM1 symptomatology control. For example, Tideglusib is a marine-derived glycogen synthase kinase 3 beta (GSK3 $\beta$ ) inhibitor, initially developed to treat Alzheimer's disease [142]. In

DM1 skeletal muscle, GSK3 $\beta$  levels are elevated and its correction might prevent myopathy in DM1 [143]. Tideglusib was found to significantly correct GSK3 $\beta$  levels and CELF1 levels and reduces expression of toxic expanded repeats in CDM1 myoblasts [144]. The drug has currently gone through a phase II trial and CDM1 and childhood DM1 patients reported improved CNS and clinical neuromuscular symptoms [145]. One of the most common therapeutic paths taken is the inhibition of the interaction between expanded repeats and MBNL1. Currently, the most advanced therapies are MYD-0124 and ERX-963, which can bind the hairpin-like structure, resulting in decreased foci formation and missplicing in DM1 cells and animal models [146,147]. More recently, a 'decoy' approach has been developed, where a modified MBNL1 protein functions as a decoy protein due to its strong affinity to the expanded repeats. This modified MBNL1 protein displaces endogenous MBNL1, restoring MBNL1 levels, subsequently improving missplicing [148]. Drugs targeting specific symptoms include metformin and resveratrol for insulin resistance [149,150], mexiletine to relieve myotonia [151], and thiamine to improve DM1 patient muscular strength [152].

In DM1, the main focus has been on RNA-based therapies due to the toxic RNA-gain-of-function and include RNA interference (RNAi) and antisense oligonucleotides (ASOs). RNAi based therapies mimic the structure of endogenously present small interfering (siRNAs) and microRNAs (miRNAs). siRNAs and miRNAs are small single-stranded RNAs inside a protein complex that can bind mRNAs, signaling their repression or degradation [153,154]. Chemically synthesized small RNAs that can bind to disease-causing mRNAs might repress their expression and alleviate symptoms. Several of these small RNAs have been engineered to bind to the expanded repeats in toxic *DMPK* transcripts. They have been found to reduce foci formation and improve splicing alterations [155,156]. ASOs also block gene expression, but through other mechanisms, such as steric blocking, splicing blocking, 5' cap blocking and activation of the RNase-H machinery in the nucleus [154,157]. Steric blocking ASOs bind to the start codon of mRNAs, blocking the binding of ribosomes. The most studied class of ASOs in DM1 are the ones that activate the RNase-H machinery, directly degrading the toxic expanded transcripts, often referred to as gapmers. Both RNAi and ASOs effectiveness is hindered by bad cellular uptake and off-target effects. In DM1 cells, the cellular membrane has equal permeability to healthy subjects [158], and overcoming this barrier has been one of the major challenges in DM1 therapy development. To improve RNAi bio distribution their delivery by adeno-associated virus vector has been evaluated, with promising results [159]. With regards to ASOs, to improve cellular uptake and binding affinity, several structurally different ASOs have been synthesized, including the use of a phosphorothioate backbone in conjunction with sugar modifications, the use of locked 2'-nucleic acids, and 2',4'-bridged nucleic acids [160,161]. Both steric blockers [162–166] and gapmers [118,167–169] have been tested in DM1 cells and animal models, with mixed results. For example, a 25 base-pair ASO restored MBNL1 loss in DM1 mouse models [162]. The advantage of RNA blocking is that wild-type *DMPK* mRNA is protected. Nevertheless, most advancement has been made with gapmers. A gapmer with RNase-H degradation activity developed by IONIS

pharmaceuticals has entered clinical trials (<https://clinicaltrials.gov/ct2/show/NCT02312011>), and although murine models and DM1 cells have previously shown good results [167,168], the trial was discontinued due to inadequate therapeutic benefits, probably caused by poor penetrance of the skeletal muscle (<https://strongly.mda.org/ionis-reports-setback-dmpkrx-program-myotonic-dystrophy/>). A more recent development on gene therapy level is the use of CRISPR-Cas9 (clustered regularly interspaced short palindromic repeats-associated protein 9) genetic scissors, which directly eliminates the expanded repeat at DNA level, and therefore eradicating the entire DM1 pathology [170,171].

Another class of therapeutic targets are miRNAs, as a global miRNA deregulation in DM1 is present, which can be linked to the clinical phenotype (discussed in detail in the miRNA section of this introduction). This opens up the possibility of them also being therapeutic targets. Several studies have been conducted, primarily focusing on miRNAs that are associated with the two important splicing regulators MBNL1 and CELF1. MBNL1 expression levels can be increased by either using artificial miRNA mimics of for example miR-1, with the goal to 'overexpress' miR-1, an upregulator of MBNL1 expression or by using anti-miRs (miR-30-5-p, miR-23b, miR-218) to inhibit miRNAs that negatively regulate MBNL1 expression [172–174]. For CELF1, mi-206 and miR-23a/b have been shown to inhibit CELF1 expression, therefore the use of artificial miRNA mimics to overexpress miR-206 and miR-23a/b have been proposed to inhibit CELF1 expression [175,176]. In both cases, MBNL1 upregulation and CELF1 downregulation would potentially reverse spliceopathy and alleviate related symptoms.

---

## 7. DISEASE MODIFIERS

---

There is still a vast part unknown about the origin of the wide variability in DM1 clinical phenotype and the disease is far more complex than originally thought. In the past few decades, several emerging disease modifiers have been discovered and further research is needed to fully understand their contribution to the molecular pathomechanisms and more importantly to the phenotypic variability observed in DM1 patients. Both to be able to manage the disease more accurately and to find new targets for therapies.

### 7.1 DISEASE MODIFIERS AT GENETIC LEVEL

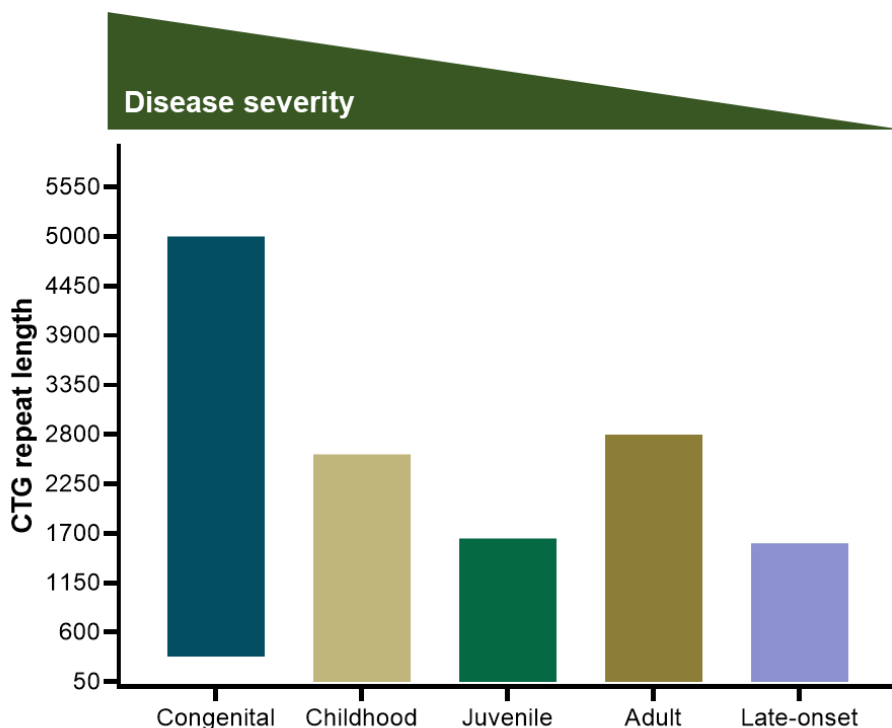
---

As DM1 is a monogenic disorder, caused by a CTG expanded repeat in the *DMPK* gene, it was one of the first places to look to account for the phenotypic variability observed in DM1. The CTG expansion can vary in length, somatically as well as generationally, potentially explaining part of the variability found in DM1. Additionally, in more recent years, it was found that the CTG expanded repeat can carry variant repeats in a subset of patients, which can also be a potential genetic modifier of the disease.



### 7.1.1 CTG EXPANSION SIZE CORRELATES TO DISEASE PHENOTYPE

The CTG expansion can vary in size, increases over time in a tissue-specific manner and larger expansions can be passed down to subsequent generations. The inheritance of a larger expansion is thought to account for the different clinical subtypes observed in DM1 and the general consensus is that the CTG expansion size in blood correlates to the disease severity and age of onset, where bigger CTG expansions relate to a lower age of onset and a more severe disease manifestation (Figure 4). CDM1, the most severe form, is often associated with CTG repeats surpassing a thousand repeats, whereas adult DM1 is associated with  $\geq 100$  repeats, but not surpassing 1000 repeats. The mildest form, late-onset, often does not show expansions over 150 CTG repeats [14]. However, quite an extensive overlap in the repeat size ranges has been observed in the different clinical subtypes (Figure 4) [6,55,177]. Furthermore, the somatic mosaicism present in the tissues of DM1 patients makes it challenging to draw genotype-phenotype correlations and gives these correlations little predictive value. The correlations to age of onset and disease severity were only present or at least stronger below a certain threshold, either  $\leq 400$  CTG repeats [178] or  $\leq 250$  CTGs [179].



**Figure 4. Estimated CTG expansion sizes across the clinical subtypes of DM1.** The highest CTG expansion sizes are observed with the congenital cases, with a downward gradient with increasing age of onset. However, sizes are variable and much overlap exists. *Adapted from Lanni & Pearson (2019).*

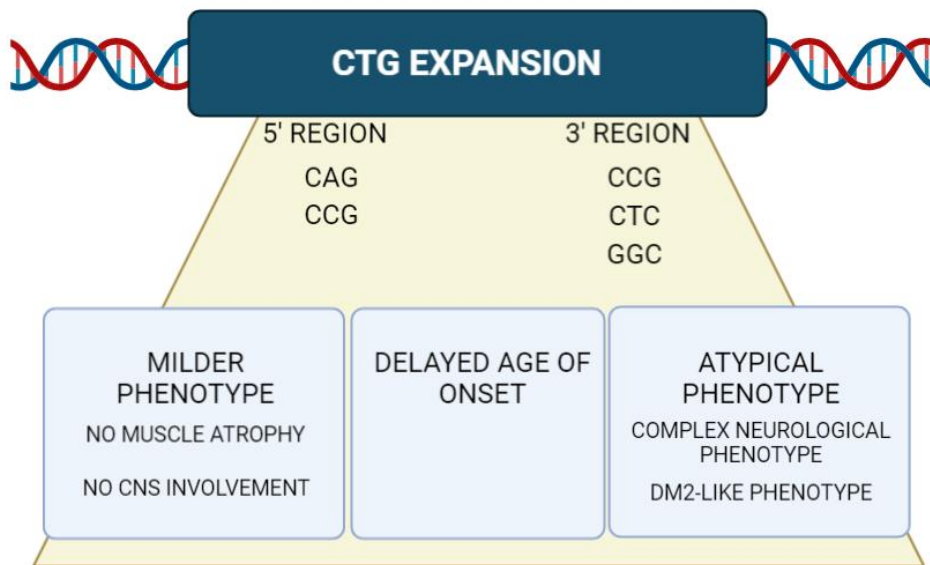
Traditionally, molecular diagnosis of DM1 was done by a Southern Blot analysis of restriction-enzyme digested genomic blood. This set-up showed the expanded repeat usually as big smears, instead of discrete bands and the mid-part of the smear was taken as the CTG expansion size [60,61,180]. The problem is that this method does not account for the confounding effect of the somatic instability and the age of sampling. To eliminate the confounding effect of somatic mosaicism, a new method was developed where small amount of input DNA is used in small pool polymerase chain reaction (SP-PCR), resulting in discrete bands and the ability to more accurately predict CTG expansion size [181]. Furthermore, it gives the opportunity to establish the ePAL, which is thought to be the allele transmitted from the parent. This allows to correct for somatic instability and age at sampling, resulting in more accurate genotype-phenotype correlations, based on an individual-specific variation level [31–34,41]. These studies revealed that not only the CTG expansion size, but also the degree of somatic instability could be linked to disease severity, adding an additional genetic modifier.

In addition to the correlation of the CTG expansion size to overall disease severity, attempts have been made to link the CTG expansion size to individual DM1 symptoms. For example, the CTG length measured in blood associates with cardiac involvement, where higher expansions were correlated to sudden death, conduction defects, left ventricular dysfunction and supraventricular arrhythmias [182,183]. Furthermore, it has been shown to relate to the muscular profile of DM1, such as muscle weakness, and myotonia [33,184,185], to cognitive deficits [186], and to respiratory function [33,187]. Genotype-phenotype correlations are hindered by the degree of somatic instability and the age at sampling, as previously mentioned. Additionally, the method of CTG expansion sizing can cause discrepancies. A study by our group has shown that the usage of different sizing methods commonly used in the DM1 field, resulted in differential CTG sizes for the same patient, further complicating genotype-phenotype correlations [188].

#### 7.1.2 VARIANT REPEATS: DISEASE STABILIZERS

Where it was long thought that the CTG expansion was an uninterrupted sequence, in the last two decades increasing evidence was found for the existence of variant repeats within the CTG expansion. To date, the known pathological variant expansions contain either unstable CCG, CTC, GGC or cytosine-adenine-guanine (CAG) sequence interruptions at the 3' end [26,43,189–191] or less frequently at the 5' end [190,192] (Figure 5). The most common variant repeat found is the CCG variant, which can be found as single repeats, in CCGCTG hexamers or as small or large (CCG)*n* arrays. Variant repeats are estimated to be present in 3-5% of the population [26,43,189–191], but this could be an underestimation due to the limitations of current techniques, as they are limited to the outer regions of the CTG expansion. Variant repeats are hypothesized to originate from rare base pair substitutions, which then spread during DNA metabolism processes [42]. This is supported by the finding of de novo variant repeats [26,44].

Variant repeats have been linked to disease phenotype, but results are controversial (Figure 5). The most observed finding is the delay in age of onset in interrupted versus non-interrupted DM1 patients [26,189,190]. Large cohort screenings have found that interrupted patients showed a delay in age of onset of 13.2 years [34] or seven years [193] compared to pure repeat DM1 patients. Moreover, to date no congenital or childhood forms have been found carrying the variant repeats [43,189,192]. Although it is often reported to associate with a delay in age of onset, there are exceptions where a variant repeat resulted in an earlier age of onset than expected [26].



**Figure 5. Variant repeats scheme.** Variant repeats found in DM1 at the 5' end and the 3' end region of the CTG expansion, which are associated with several clinical attributes, including milder phenotype, delayed age of onset and atypical phenotypes. Abbreviations: CNS= central nervous system; DM2= myotonic dystrophy type 2.

In addition to the delay in age of onset, a milder phenotype is often mentioned [26,44,189,190,193,194], where variant repeat carrying DM1 patients have shown less severe muscle weakness, less myotonia and better respiratory function. Moreover, it was shown that variant repeat containing sequences were associated with a significant decrease in the severity of progressive symptoms [33]. Moreover, CNS involvement in variant repeat carrying DM1 patients seems limited and they are found to do better in terms of cognitive and behavioral performance [34,193]. A possible explanation to the observed milder phenotype is the potentially stabilizing effect of the variant repeats in both somatic and germline cells. Variant repeat carrying individuals have shown a more stable intergenerational transmission and a lesser degree of somatic instability [26,43,192,195]. Interestingly, this is independent of the size, location and pattern of the interruptions. Even

a single interruption can result in a more stable CTG expansion [192,195]. The stabilizing effect in germ line cells has resulted in the transmission of either similar or shorter CTG expansions in offspring [26,43,189,196], independent of parental transmission. More precisely, interrupted DM1 patients vs. non-interrupted patients showed, upon transmission, a higher prevalence of stable or contracted repeats (68.4 vs 6.4%) and a lower frequency of expansion (31.6% vs. 93.6%) [42]. Contradictory to the milder phenotype, either no differences or a worsening of symptoms compared to pure repeat carrying DM1 patients has also been found [189,191]. Of note, Cumming and collaborators have found variant repeats to interfere with genetic diagnosis of DM1, where an individual with mild symptoms and dominant family history showed normal CTG repeat sizes after triplet-primed polymerase chain reaction (TP-PCR). However, further analysis revealed an expanded CTG repeat, carrying variant repeats [197].

Another observation regarding associations to clinical phenotype is the finding of both atypical and extremely complex phenotypes. For example, two independent studies have found a clinical phenotype more resembling the DM2 phenotype rather than DM1, with proximal weakness, calf hypertrophy and the absence of myotonia [26,191]. An extremely complex neurological phenotype was observed in a Dutch family, where the variant carrying DM1 family was associated with Charcot-Marie-Tooth disease [43]. These vast differences in clinical presentation and age of onset may be explained by the difference in variant repeat patterns, as almost each individual shows a unique pattern, with a wide variety in type, size and location of these variant repeats, which at least in part, can explain some of the intervariability found.

## 7.2 DISEASE MODIFIERS AT TRANSCRIPTIONAL LEVEL

---

The expanded mRNA transcripts have been shown to stand at the base of the disease by an RNA gain-of-function mechanism due to accumulation of expanded mRNA transcripts, leading to a spliceopathy. Other transcriptional mechanisms have been proposed to contribute to DM1 pathology, including altered gene expression profiles and bidirectional transcription.

### 7.2.1 ALTERED GENE EXPRESSION LEVELS SURPASSING THE *DMPK* GENE

As mentioned above, the expanded mRNA transcripts are trapped inside the nucleus, which prevents protein translation, but additionally it has been found that the mRNA *DMPK* levels themselves are downregulated [61,198,199]. However, findings are not unanimous [200]. The decrease in *DMPK* mRNA levels might contribute to further depletion of the *DMPK* protein in DM1 patients. The *DMPK* gene resides in a crowded part of chromosome 19, with several genes overlapping, raising the question whether the expanded repeat in the 3'UTR might have an effect on expression levels of neighboring genes as well; further worsening DM1 pathology. Indeed, the downstream adjacent and partly overlapping *SIX5* gene has

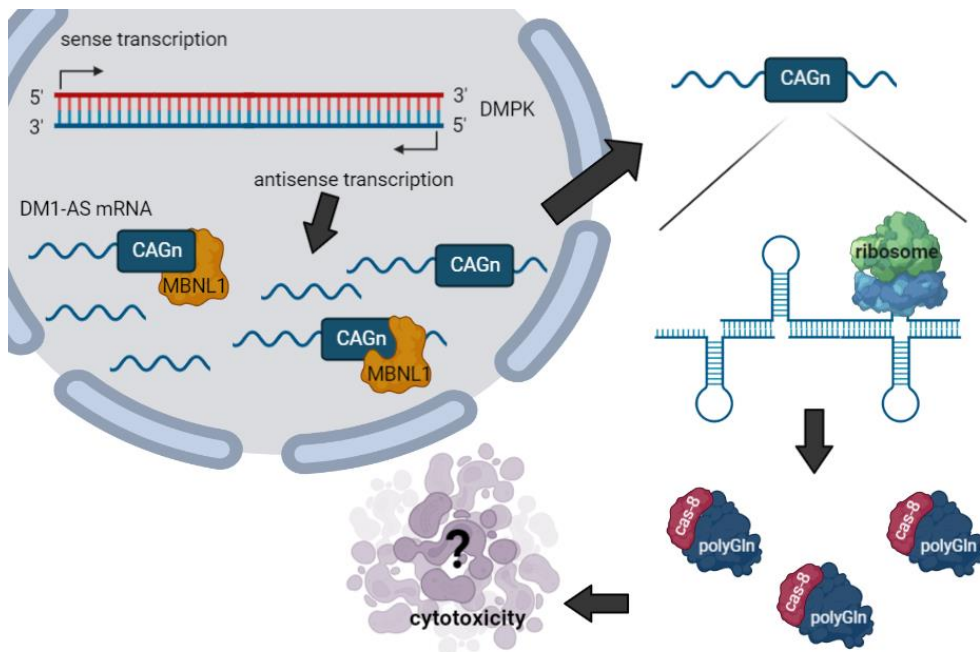
been shown to have lower mRNA expression levels, whereas the upstream Dystrophia myotonica WD repeat-containing protein (*DMWD*) gene seems to not be altered [201,202]. *SIX5* downregulation might contribute to DM1 pathology as *SIX5* knockout mice have been found to develop cataracts [113]. Furthermore, the expanded mRNA transcripts have been found to have other roles besides alternative splicing misregulation through MBNL1 and CELF1. For example, expanded CUG repeats can be processed into small RNAs, that can activate an RNA interference pathway to silence specific targets and increase toxicity [203].

### 7.2.2 BIDIRECTIONAL TRANSCRIPTION: DOUBLE THE TROUBLE

Another emerging transcriptional disease modifier is the mechanism of bidirectional transcription, where transcription occurs in both the sense and antisense direction at the gene locus. Antisense transcripts have been found to play a role in several microsatellite expansion disorders, such as Huntington's disease (HD), spinocerebellar ataxia type 8 (SCA8) and also in both Myotonic Dystrophies [204]. Of note, the mechanism of bidirectional transcription is not disease inducing per se, as wild-type alleles also produce antisense transcripts, and are emerging as important regulators of gene expression [205–207]. They have been found to regulate gene expression and genome integrity by interfering with sense transcription, and modulation of histone modifications and DNA methylation [208,209]. Over 70% of transcripts in humans are bidirectionally transcribed [210]. Antisense transcripts are generally low in abundance and preferentially retained in the nucleus [211,212]. The process of bidirectional transcription in itself is not pathogenic, but deregulation in its expression or in terms of microsatellite expansion disorders, the inclusion of the expansion, might be. Indeed, both overexpression of certain antisense transcripts and the presence of the expanded repeat have been linked to disease. For example in HD, Huntingtin antisense (HTTAS) is alternatively spliced into HTTAS-v1 and HTTAS-v2, of which the former has been observed to be downregulated in human HD frontal cortex, and in in vitro models overexpression of HTTAS-v1 resulted in reduced levels of Huntingtin, suggesting a role for HTTAS-v1 in Huntingtin regulation and HD progression [213]. Moreover, in spinocerebellar ataxia type 2 (SCA2), bidirectional transcription of the gene was present in both SCA2 and controls, but only the presence of the expanded antisense transcript led to cell death of primary mouse cortical neurons and formed RNA foci in SCA2 human cerebella, most likely co-localizing with MBNL1 [214].

DM1 anti-sense (DM1-AS) transcription was first described by Cho and collaborators in 2005 [215]. An antisense transcript, emanating from the adjacent *SIX5* regulatory region, was reported. This transcript was converted into 21 nucleotides siRNAs, which were proposed to have a regulatory role in heterochromatin formation. More recent studies, however, show that DM1-AS transcription extends across the CAG repeat, since DM1-AS RNA was detected, originating from downstream of the repeat, in the 3' region of DM1 tissue [216]. Additionally, these antisense mRNA transcripts are found to form antisense RNA foci, indicating that they indeed carry the expanded repeat. These antisense RNA foci do not co-localize with sense

RNA foci, but were found to co-localize with MBNL1 in DM1 heart samples [216]. Co-localization with MBNL1 might indicate that they potentially contribute to the DM1 pathology in a similar fashion as sense RNA foci (Figure 6). Similar results were found by Michel and collaborators in human fetal samples, indicating a similar role for DM1-AS transcription in CDM1 [217]. The inclusion of the expanded repeat was confirmed by Gudde and collaborators in 2017, although they found DM1-AS transcripts to be low in abundance and with varying lengths, both including and excluding the CAG repeat [218].



**Figure 6. Bidirectional transcription and repeat associated non-ATG (RAN) translation in DM1.** Antisense transcription of the *DMPK* gene (DM1-AS) leads to a pool of mRNAs with and without the expanded repeat (CAGn from the 5'-to-3' direction of the antisense strand). The expanded DM1-AS can both sequester MBNL1 and travel outside of the nucleus, where ribosomes attach to the hairpin like structures, resulting in the production of long stretches of polyGln RAN protein, co-localizing with caspase-8, potentially contributing to cytotoxicity. Abbreviations: CAGn= expanded CAG repeat; MBNL1= muscleblind-like 1; DMPK= dystrophia myotonia protein kinase gene; polyGln= polyglutamine RAN protein; cas-8= caspase 8.

Little is known about the contribution to molecular mechanism or clinical phenotype of these antisense transcripts. Evidence from other CAG expansion studies shows that the presence of the expanded CAG repeats can induce cytotoxicity in animal models and in a similar fashion to what is observed with CTG repeats, enhance cytotoxicity through a triple repeat-derived siRNA mechanism [219,220]. Moreover, CAG antisense transcripts can be translated through repeat-associated non ATG (RAN) translation into potentially toxic peptides, further discussed in the next section [221,222].

### 7.3 DISEASE MODIFIERS AT PROTEIN LEVEL

---

At protein level, the main focus has been on proteins sequestered or upregulated by the toxic expanded repeats, involved in the DM1 spliceopathy. Especially, it was thought that the expanded mRNA transcripts are retained in the nucleus, preventing protein translation. This of course does have effects on the *DMPK* protein levels itself, which can attribute to DM1 pathology to some extent, but the discovery of RAN translation in microsatellite expansion disorders, including DM1, has put into focus a disease modifier on protein level outside of the scope of the spliceopathy.

#### 7.3.1 THE *DMPK* PROTEIN

The CTG expansion resides in the 3' end UTR of the *DMPK* gene. This gene encodes for a serine/threonine protein kinase and six major isoforms are found in both humans and mice, of which an 80 kD protein is predominantly found in skeletal and heart muscle [223,224]. Its full functional and regulatory properties are still not fully understood, but the *DMPK* protein has been found to be involved in skeletal muscle integrity, cardiac muscle atrioventricular conduction, ion-channel gating and cell metabolism [225]. Both decreased *DMPK* mRNA transcript levels and the retention of mRNA transcripts in the nucleus have been found in DM1 [199,201]. Subsequently, decreased levels of *DMPK* protein in skeletal and cardiac muscles are also observed [61,223]. In skeletal muscle from DM1, this decrease was found to be about 50% and did not correlate to CTG expansion length [226].

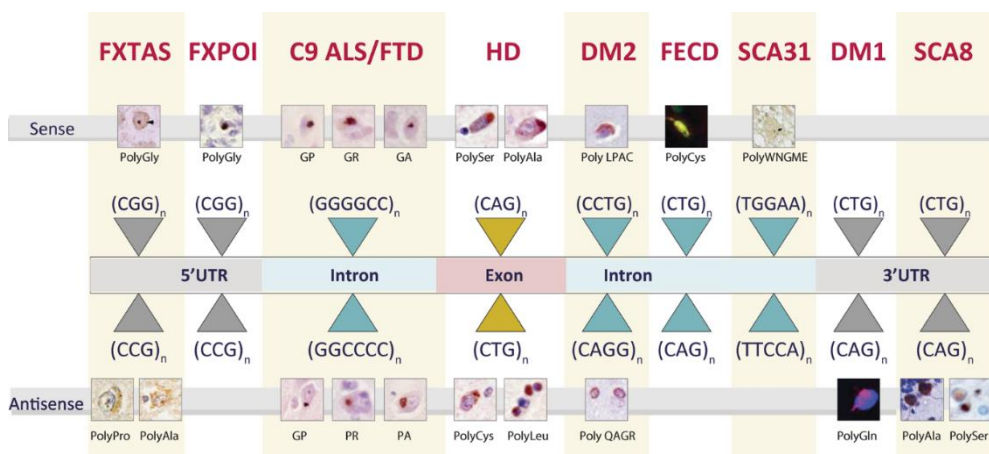
The decrease in *DMPK* protein may be an additional contributor to DM1 pathology. Indeed, *DMPK* transgenic and knock out mice show a mild late-onset myopathy and smaller head and neck muscle fibers, similar to the comprised muscle function of these regions in DM1 patients [63,64]. Additionally, they have been found to have abnormal sodium channel gating in skeletal muscle [227], a mechanism previously linked to myotonia in DM1 patients, and DM1-like calcium homeostasis [228]. Moreover, cardiac conduction defects are found similar to those observed in DM1 [62,63,229] and mice present with muscle insulin resistance [230].

#### 7.3.2 RAN TRANSLATION: ADDITIONAL CYTOTOXICITY

For RNA-gain-of-function disorders, such as DM1, the main focus has always been on the accumulation of foci and the entrapment of RNA-binding-proteins. It was thought that the expanded mRNA transcripts were unable to travel outside of the nucleus and therefore protein translation from the expanded repeat was absent [65]. However, the discovery of RAN translation in 2011 by Zu and collaborators has changed the view of the pathology of RNA-gain-of-function disorders and microsatellite expansion disorders in general [221]. RAN translation goes against the historical view that eukaryotic translation initiates at an AUG start codon. Canonical translation initiation consists of the step-wise assembly of 80S ribosomes at start codons of mRNA. It is a highly complex process, in which at least nine eukaryotic initiation factors are involved [231]. The mRNA secondary and tertiary structures

contribute significantly to the dynamics and regulation of translation initiation. These structures in a 5' UTR region can influence translation initiation both negatively and positively, depending on their position. Highly structured regions upstream of an AUG start codon can inhibit initiation, while downstream secondary structures can facilitate initiation at imperfect start codons [232–235]. Furthermore, several atypical translation mechanisms exist, which are modulated by mRNA secondary structures. For example, several viral RNAs are translated via an internal ribosome entry site (IRES)-mediated pathway. IRESs are complex RNA structures that can directly recruit ribosomal subunits and initiation factors [236]. In the case of RAN translation, proteins are formed from the microsatellite expansions without the presence of an AUG start codon, probably through the use of the mRNA secondary structures in these expanded transcripts. It can occur in all three reading frames and from both the sense and antisense strand of the expanded transcripts, which can result in up to six proteins from a single expansion mutation [237].

Since its discovery in 2011 in DM1 and SCA8, RAN translation has been reported in seven other microsatellite expansion disorders (Figure 7): C9orf72 amyotrophic lateral sclerosis /frontotemporal dementia (C9ORF72 ALS/FTD) [238–240], fragile X-associated tremor/ataxia syndrome (FXTAS) [241], Fragile X-associated primary ovarian insufficiency (FXPOI) [242], HD [222,243], spinocerebellar ataxia 31 (SCA31) [244], Fuchs' endothelial corneal dystrophy (FECD) [245] and myotonic dystrophy type 2 (DM2) [246]. Expanded repeats in a variety of RNA contexts have resulted in RAN translation, including within 3' and 5' UTR regions, protein-coding open reading frames, or introns and non-coding RNAs.



**Figure 7. RAN proteins identified in nine microsatellite expansion disorders.** Both originating from the sense and antisense strand. Repeats are given in the 5' to 3' direction of the individual strands. Abbreviations: FXTAS= fragile X-associated tremor/ataxia syndrome; FXPOI= Fragile X-associated primary ovarian insufficiency; C9 ALS/FTD= C9orf72 amyotrophic lateral sclerosis /frontotemporal dementia; HD= Huntington's disease; DM1= Myotonic Dystrophy type 2; FECD= Fuchs' endothelial corneal dystrophy; SCA31= spinocerebellar ataxia 31; DM1=Myotonic Dystrophy type 1; SCA8= spinocerebellar ataxia type 8. *Reproduced from Banez-Coronel & Ranum (2019).*



The mechanism initiating RAN translation and the effect of the RAN translated proteins is still poorly understood. Studies on FXTAS, a neurodegenerative disorder, caused by CGG repeats in the 5' UTR of the fragile X mental retardation 1 gene, found RAN translation initiation in at least two reading frames, resulting in polyalanine and polyglycine proteins. Data on RAN translation in FXTAS thus far points to a scanning mechanism, using a near-AUG codon for initiation just upstream of the repeat, thus relying on a canonical translation mechanism [241,247,248]. Moreover, it has been postulated that RAN translation is initiated through the IRES-mediated pathway, but proof-of-concept is missing [249]. Although, the role of specific RAN translated proteins in individual microsatellite expansion disorders is not yet clear, there is growing evidence that the peptides are toxic and contribute to pathology. For example, studies have found that RAN proteins are toxic, independent of the RNA-gain-of-function mechanism, illustrated by codon-replacement strategies [221,222,246]. Additionally, in HD, the RAN translated proteins were found primarily in regions of the brain showing neuronal loss, cell death and microglial activation [222]. Mice expressing the polyglycine RAN protein found in FXTAS showed behavioral deficits, which were not observed in FXTAS mice only expressing the expanded repeat [250]. In C9ORF72 ALS/FTD in vitro and in vivo experiments, overexpressing RAN proteins has been reported to be toxic, with impaired dendritic branching, endoplasmic reticulum stress and apoptosis in neuronal cells [251,252].

In DM1, Zu and collaborators discovered a novel polyglutamine (polyGln) RAN protein expressed from the antisense CAG expansion transcript of the *DMPK* gene [221]. Nuclear polyGln RAN protein aggregates were found at a low frequency in a DM1 patient's myoblasts and skeletal muscle (n= 1) and at a higher frequency in leukocytes from peripheral blood (n= 1) [221]. The nuclear aggregates co-localized with caspase-8, an early indicator of polyGln-induced apoptosis. This suggests that RAN proteins may be an additional mechanism of cytotoxicity in DM1 cells (Figure 6). The contribution of RAN translation to DM1 pathology has not been further studied since its first report in 2011 and much remains unknown regarding the presence of RAN translation and its contribution to DM1 pathology.

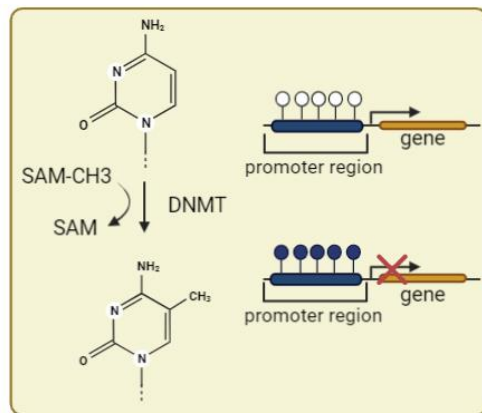
## 7.4 DISEASE MODIFIERS AT EPIGENETIC LEVEL

---

Epigenetics is the study of heritable changes in gene expression and function that are not attributed to alterations of the DNA sequence. The word epigenetics is of Greek origin and literally means over and above (epi) the genome. Epigenetic modifications includes DNA methylation and histone modifications [253]. Deregulation of gene expression by epigenetic mechanisms have been shown to play a role in pathogenesis of cancer, autoimmune diseases, neuromuscular disorders and expansion disorders, including DM1 [254–260].

#### 7.4.1 ABERRANT DNA METHYLATION PATTERNS IN THE *DMPK* LOCUS

DNA methylation is a heritable, yet reversible epigenetic modification and is essential for silencing retroviral elements, regulating tissue-specific gene expression, genomic imprinting, and X chromosome inactivation [261,262]. DNA methylation occurs through a family of DNA methyltransferases (DNMTs) that transfer a methyl group from S-adenyl methionine to the fifth carbon of a cytosine to form 5-methylcytosine (5mC) (Figure 8). 5mC only accounts for ~1% of the nucleic acids in the human genome [263]. There are three main DNMTs at play in DNA methylation. DNMT3a and DNMT3b are able to produce de novo DNA methylation, whereas DNMT1 is active during DNA replication to preserve the DNA methylation pattern from the parental DNA [262,264]. All three DNMTs are active during development. The current model shows two big waves of demethylation and remethylation during development, after which DNMT levels are reduced, resulting in rather stable DNA methylation patterns in post mitotic cells [265–267]. DNA methylation occurs most often on so-called CpG sites, where a cytosine is positioned upstream of a guanine. CpG sites are overall underrepresented in the genome, and the vast majority are heavily methylated, except the ones located in CpG islands (CpGis) [268].



**Figure 8. The process of DNA methylation.** Methylated cytosines are formed through DNMTs, which transfer a methyl group from SAM onto a cytosine, resulting in repression of gene expression due to methylated CpG islands, located in promoters. Every dot in the promoter region represents a CpG dinucleotide, where a filled in dot means hypermethylation. The collection of dots indicates a CpG island. Abbreviations: SAM= S-adenyl methionine; CH3= methyl group; DNMT= DNA methyl transferases.

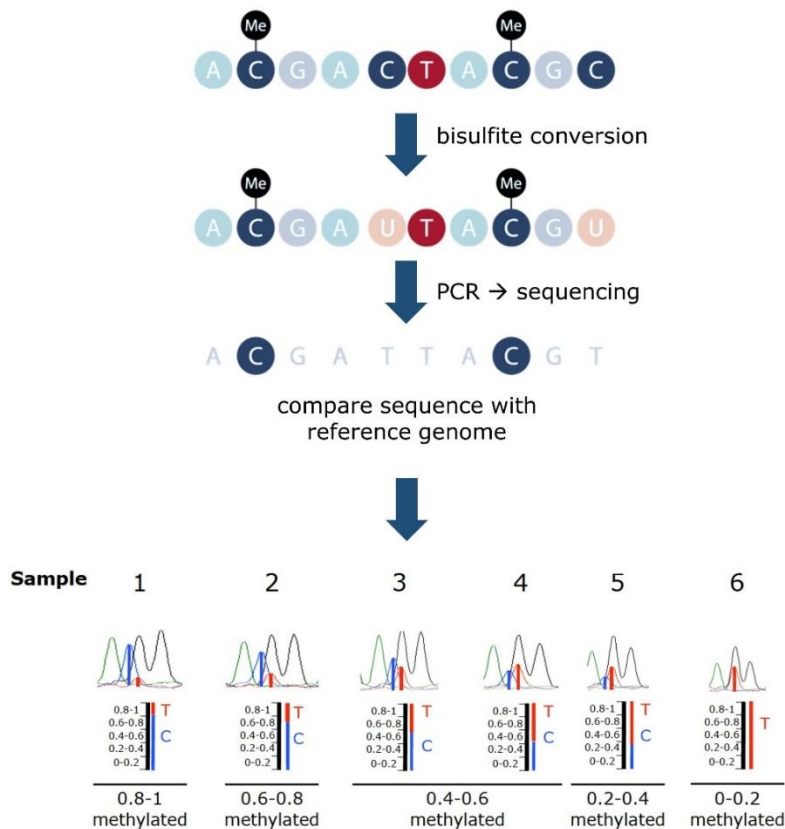
CpG islands are defined as short regions over 200 bp, with more than 50% CG content and a ratio greater than 0.6 of observed number of CG dinucleotides relative to the expected one. They are generally found at promoters of housekeeping and developmental regulatory genes, but can be also found in exons, introns and regulatory regions [269], whereas tissue-specific genes contain archetypically CpG-poor promoters [270]. Interestingly, the CpGis are mainly untouched by the methylation waves during development, and remain largely unmethylated throughout the course of development and afterwards in post mitotic cells [271]. CpG islands are highly conserved regions throughout evolution, indicating their important functional role [272]. CpG islands enhance the accessibility of DNA and promote transcription factor binding [273]. The methylation of CpG islands results in stable silencing of gene expression (Figure 8) [269,270]. DNA methylation in itself can reduce

gene expression by impairing the binding of transcription factors, but it can also favor the binding of a class of proteins with high affinity for 5mC, which inhibits transcription by recruiting chromatin remodelers associated with gene repression. This class of proteins consists of for example the methyl-CpG-binding domain proteins and the ubiquitin-like, containing PHD and RING finger domain proteins [274–277]. Although CpG islands are largely unmethylated, DNA methylation patterns have been found to be disrupted in disease-state and CpG island hypermethylation has been linked to several diseases [278].

Distinct methods exist to address DNA methylation, all requiring a pre-treatment of the genomic DNA to be able to distinguish between methylated and unmethylated cytosines. Several methylation dependent pre-treatments have been established, including endonuclease digestion, affinity enrichment and bisulfite conversion. These pre-treatments can then be followed by an array of different analysis techniques, either genome-wide approaches, such as DNA methylation microarrays and next-generation sequencing, or single nucleotide resolution approaches, including bisulfite Sanger sequencing PCR (BSP), pyrosequencing and methylation-specific PCR. The first pre-treatment, endonuclease digestion, is based on the knowledge that each sequence-specific restriction enzyme has an accompanying DNA methyl transferase to protect endogenous DNA from the restriction defense system by methylating bases in the recognition site. Some restriction enzymes are inhibited by 5mC at CpG sites, which means that cutting patterns by these enzymes can provide information on DNA methylation patterns [279]. The most utilized restriction enzymes inhibited by 5mC are HpaII and SmaI. Mostly due to the existence of isoschizomers (MspI for HpaII) or neoschizomers (XmaI for SmaI) that are not inhibited by methylation, allowing easy comparisons [280,281]. Affinity enrichment of methylated regions using 5mC-specific antibodies or methyl-binding proteins are particularly used in whole-genome profiling through array hybridization and more recently next-generation sequencing [282–284]. Although this approach allows for rapid and efficient whole-genome analysis, it does not yield information on individual CpG dinucleotides.

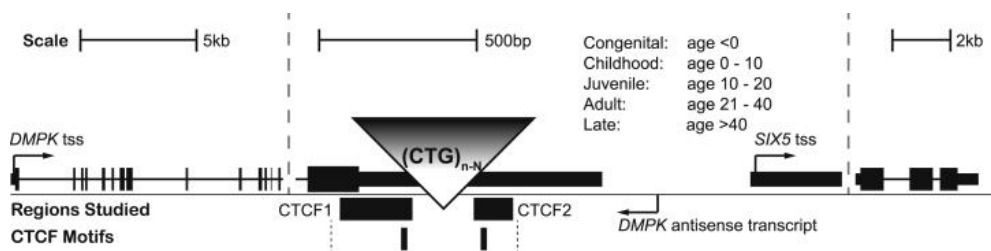
One of the most used pre-treatments is bisulfite conversion, where denaturation of genomic DNA with sodium bisulfite results in the conversion of unmethylated cytosines into thymines, whereas methylated cytosines remain unchanged, allowing the investigation of DNA methylation on genetic level (Figure 9) [285]. At the single nucleotide levels, the most utilized techniques are BSP, and pyrosequencing. BSP has been the gold standard for many years, because it provides a qualitative, quantitative and efficient approach to identify 5-methylcytosine at single base-pair resolution. In BSP, the bisulfite treatment is followed by (nested) PCR and subsequent Sanger sequencing [286]. The degree of methylation in the cell pool is determined by the height difference between the cytosine and thymine nucleotides at the CpG site, as often the methylation is not absolute due to the mix of DNA molecules originating from the mixture of cells and a range is given from 0-100% (Figure 9). However, pyrosequencing has become more popular in recent years. Pyrosequencing is a sequencing-

by-synthesis technology that relies on the luminometric detection of pyrophosphate release upon nucleotide incorporation through a cascade consisting of four enzymes. It results in precise quantification of methylation at each single CpG analyzed [287,288]. Both have advantages and disadvantages in their usage. BSP is quite labor-intensive, due to the high number of samples and the need for nested-PCR to overcome unspecific amplification, whereas pyrosequencing requires specialized equipment and is a fairly expensive technique. Moreover, it has been shown that BSP might be more sensitive to strong hypermethylation of DNA [289,290].



**Figure 9. Bisulfite conversion principle.** Denaturation of genomic DNA with sodium bisulphite results in the conversion of unmethylated Cs into Ts, whereas methylated Cs remain unchanged. After PCR and Sanger sequencing, comparison with the reference genome can establish which CpG sites are methylated. The degree of methylation is often not absolute due to the DNA pool and is most often presented as a range of methylation between 0% to 100%, depending on the difference between the two peaks observed.

DNA methylation profiles gained interested in DM1 due to the location of the expanded repeat, which happens to reside in a 3.5 kb CpG island (CpGi 374). Additionally, the CTG repeat is flanked by two CTCF-binding factor (CTCF) binding sites (Figure 10). CTCF is an important transcription factor and can act as a transcriptional activator, repressor and insulator, and mediates chromatin looping [291]. Early studies suggest that the two CTCF binding sites together with the expanded repeat establish an insulator element between the *DMPK* promoter and the *SIX5* enhancer [292]. The most important binding site for CTCF seems to be the upstream CTCF1 binding site, as binding of CTCF to the downstream CTCF2 binding site is controversial [135,292,293]. Hypermethylation of CTCF1 has been shown to result in the loss of binding, potentially affecting chromatin dynamics [215,292]. Several studies have found aberrant DNA methylation profiles in these two regions flanking the repeat (upstream CTCF1 and downstream CTCF2) in DM1, but results are controversial with variable degrees of hypermethylation observed at both sites in different clinical subtypes of DM1 (Table 2) [32,55,300,135,215,294–299].



**Figure 10. The DM1 locus.** The DM1 locus, associated genes, and mapped functional regions are schematically shown. The CTG repeat is located in the 3' UTR and *SIX5* promoter, of which part of the DNA sequence is shown. *Reproduced from Barbé et al. (2017).*

Overall, the vast majority of hypermethylation around the expanded repeat was observed in CDM1 cases, predominantly at the CTCF1 site, suggesting an important role for DNA methylation in the younger, more severe clinical phenotype. Interestingly, DNA methylation status has also been linked to parental transmission, where hypermethylation was almost exclusively observed upon maternal transmission [55,298]. Together, this has led to the idea of the existence of a CpG methylation based parent-of-origin effect, that can explain the maternal bias for CDM1 offspring [55]. The proposed hypothesis is that hypermethylation around the repeat results in decreased levels of *SIX5* expression, which is detrimental for spermatogonia as their survival relies on *SIX5* protein levels, and may in turn prevent the transmission of CDM1. This protective mechanism is not present in females, as oogonia do not rely on *SIX5* expression, hence the maternal bias for CDM1 transmission. Prenatal screening for CDM1 has relied primarily on the CTG expansion size and while larger expansions are often related to CDM1 offspring, this relation is not absolute and Barbé and

collaborators found that methylation status might be a better indicator for CDM1 offspring and a good prenatal screening biomarker [55].

In addition to the relation with CDM1 and maternal transmission, CTCF1 hypermethylation has been associated with greater expansion sizes, independent of clinical subtype [296–298,300]. However, this association is controversial and not always found [55,299]. DNA methylation profiles have also been studied in relation to variant repeats and the presence of variant repeats in the CTG expansions are linked to a differential DNA methylation pattern. Hypermethylation was predominantly observed in the downstream region of the CTG repeat, as opposed to the upstream region in non-interrupted DM1 patients [294,298,301].

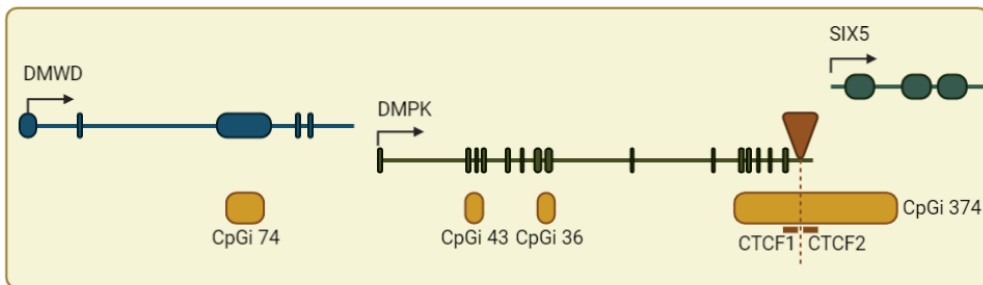
**Table 2. Summary of studies on DNA methylation in the region surrounding the CTG expansion.**

Disease Form (tissue of origin)	Sample Size	Genomic Context	Analysis outcome	Ref
CDM1; Adult (dura mater, skeletal muscle, skin biopsies and white blood cells)	30 DM1	Upstream region of (CTG) <sub>n</sub> repeats	Hypermethylation in intron 12 at restriction sites of SacII and HhaI in CDM1 patients	[300]
DM1 fetuses; DM1 adults; transgenic DM1 mice (several tissues)	13 DM1 vs. 3 CTRs	Upstream (CTCF1) and downstream (CTCF2) regions	Hypermethylation of CTCF1 in DM1 individuals. In DM1 mice methylation pattern was present at both sites	[302]
Childhood; Juvenile CDM1; DM1 "atypical" (whole blood)	66 DM1 (9 VRs) vs. 30 CTRs	Upstream (CTCF1) and downstream (CTCF2) regions	Hypermethylation of CTCF1 in CDM1 and childhood patients, significantly associated with MT. DM1 patients with VRs show a distinctive DNAm pattern	[298]
DM1- hESCs (hESCs)	14 DM1	DNA sequence spanning from exon 11 to the (CTG) <sub>n</sub>	Marked increase in DNAm levels of the expanded allele	[135]
Late-onset; Adult; Juvenile; Childhood (whole blood)	92 DM1 vs. 10 CTRs	Upstream (CTCF1) and downstream (CTCF2) regions	DNAm levels of both CTCF sites higher in CDM1 than in non-CDM1 patients	[55]
Adult; DM1 "atypical" (whole blood)	90 DM1 (8 VRs)	Upstream (CTCF1) and downstream (CTCF2) regions	CTCF1 region DNAm levels correlated with CTG repeat length, and the presence of a VRs was associated with higher DNAm levels	[301]
Adult; DM1 "atypical" (whole blood)	115 DM1 (12 VRs)	Downstream region (no CTCF binding sites) of the (CTG) <sub>n</sub>	Patients with VRs alleles had distinctive DNAm and cognitive profile	[294]
Non-CDM1 DM1 (whole blood)	68 DM1 vs. 73 CTRs	Upstream and downstream regions (no CTCF binding sites) of the (CTG) <sub>n</sub>	Hypermethylation of both upstream and downstream regions	[296]

Abbreviations: CDM1= congenital myotonic dystrophy type 1; VRs= variant repeats; MT= maternal transmission; DM1= myotonic dystrophy type 1; CTRs= controls; DNAm= DNA methylation. Adapted from Visconti et al. (2021)

The effect of aberrant DNA methylation profiles on clinical phenotype is poorly understood. It seems to correlate with age of onset of the disease, where earlier onsets have a higher prevalence of hypermethylation [55,297]. Data on the contribution to specific aspects of the clinical phenotype are scarce. Interestingly, although the incidence of aberrant DNA methylation profiles in adult DM1 is low, it has been linked to respiratory and muscular profiles and a decline in cognitive function in adult DM1 [294,301]. This might suggest that DNA methylation status can function as a prognostic marker in DM1.

Although blood has been the tissue of origin in most studies, the DNA methylation status has been investigated in other tissues. Two studies to date have looked at DNA methylation status in tissues other than blood. In fetal samples, high levels of DNA methylation were found for all tissues studied, including muscle, liver, kidney, heart, pancreas, brain and skin. This methylation profile was highly polarized, as only the CTCF1 region showed hypermethylation, whereas CTCF2 showed sparse methylation in only one of the three fetuses analyzed. Adult DM1 heart, liver and cortex showed high-to-moderate methylation levels in CTCF1, whereas cerebellum, kidney and skeletal muscle showed low-to-no methylation levels. CTCF2 was completely devoid of methylation in all adult DM1 tissues [302]. Similarly, Hildonen and collaborators showed higher levels of CTCF2, but not CTCF1, methylation in adult DM1 muscle compared to control muscle and DM1 blood samples, but differences were less than 5% in most samples [296].



**Figure 11. The CpG islands in the DMPK locus.** CpGi 74 resides in the DMWD gene, while the other three reside in the DMPK gene. CpGi 374 is split into two regions, encompassing the CTG expansion and referring to the two CTCF binding sites in these regions. DMWD= myotonic dystrophy WD containing protein; DMPK= dystrophia myotonica protein kinase; SIX5= SIX homeobox 5; CpGi=

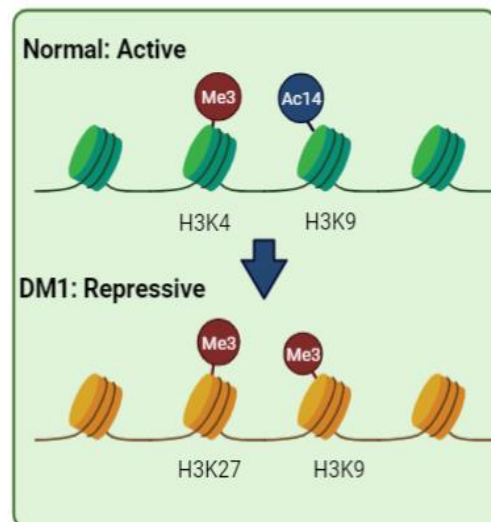
The DMPK gene harbors two more CpG islands (CpGi 36 and 43), and an additional one is found in the neighboring *DMWD* gene (CpGi 74), henceforth referred to in its entirety as the *DMPK* locus (Figure 11). Although DNA methylation is a fairly stable epigenetic mechanism, it is known that it can spread at low levels during an individual's lifetime and might influence neighboring CpG islands [303]. In addition, in the only one publication that has analyzed epigenetics at the entire *DMPK* gene and neighboring genes in control tissues, it has been shown that CpGi 74 and 36 are fully methylated in blood, while CpGi 43 is non-

methylated [293]. Conversely, for CpGi 74 and 43 this profile is reversed in muscle and muscle-derived cells, with complete hypomethylation of CpGi 74 and hypermethylation of CpGi 43. Interestingly, CpGi 43 contains a proposed alternative *DMPK* promoter, suggesting that the myogenic hypermethylation could suppress the usage of this promoter and drive the usage of the strong/canonical *DMPK* promoter [293]. Regarding DM1 samples, nothing has been published addressing the epigenetic behavior of these neighboring CpG islands in the *DMPK* locus.

#### 7.4.2 HISTONE MODIFICATIONS: A REPRESSIVE SWITCH

Histone modifications are important regulators that control chromatin structure and gene transcription. Histones can be post-translationally modified at their N-terminal tails through a large number of different histone post-translational modifications, including acetylation, phosphorylation and methylation. These modifications define chromatin structure by the recruitment of proteins and complexes with enzymatic activities, which in turn regulates transcription, as the gene will be more or less accessible for transcription [304].

Little is known to date on the chromatin structure of the *DMPK* locus and the contribution of histone modifications. The presence of the CTG expansion has been shown to alter regional chromatin structure, where it increases the efficiency of DNA nucleosome assembly, resulting in the creation of more stable nucleosomes and hence repression of transcription [305–307]. Indeed, the presence of the CTG expansion in DM1 has been associated with heterochromatinization of the *DMPK* locus [295,308]. More specifically, a change in local chromatin structure, about 500 base pairs downstream of the expansion, has been reported in DM1 muscle and skin fibroblasts, suggesting a switch to a more transcriptional repressive heterochromatin [308]. This repressive chromatin structure might explain the reduced expression of *DMPK* and *SIX5* observed in DM1 [61,198,199,201,202]. Upon expansion of the CTG repeat, heterochromatin spreading was observed and active H3K4me3 was replaced by a repressive H3K9me3 mark [292]. Moreover, the less active heterochromatin environment surrounding the expanded repeat is characterized by a decrease in H3K9/14Ac (a marker for active chromatin) at the CTCF binding sites and an enrichment



**Figure 12. Histone modifications observed in DM1.** The active chromatin state surrounding the CpG island 374 is replaced by a repressive chromatin state.



of H3K27me and H3K9me3, both markers for repressed heterochromatin [295]. Taken together, this suggests a chromatin remodeling at the *DMPK* locus in DM1 towards a more repressive heterochromatin state (Figure 12). However, a more recent study by Sorek and collaborators have found an opposite histone modification profile, indicative of an active chromatin state [309].

## 7.5 DISEASE MODIFIERS AT NON-CODING RNA LEVEL

---

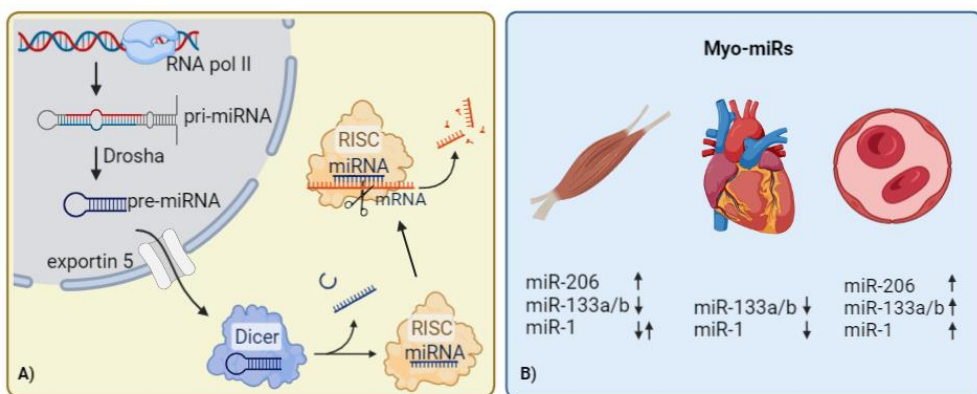
The genomes of eukaryotes are mostly comprised of non-protein coding DNA. A portion of this non-protein coding DNA is transcribed into RNA, commonly referred to as non-coding RNA (ncRNA) and in the past simply referred to as 'junk' RNA. However, in the last decades a subset of these ncRNAs have been found to have important biological roles in development, homeostasis and also in disease. ncRNAs are comprised of an array of different ncRNAs, such as long non-coding RNAs, small nucleolar RNAs and the most well studied ncRNAs, the miRNAs [310]. Little is known about the role of ncRNAs in DM1, with the exception of miRNAs, which will be further discussed below.

### 7.5.1 GLOBAL miRNA DEREGULATION IN MYOTONIC DYSTROPHY TYPE 1

miRNAs are small, single-stranded RNAs about twenty-two nucleotides long, that regulate gene expression levels by either inhibiting translation or promoting degradation of their target mRNAs [311]. It is estimated that the human genome encodes for over a thousand miRNAs, which can target dozens of mRNAs, and every individual mRNA can be targeted by several miRNAs [312]. The majority of miRNAs result from RNA polymerase II transcription, which yields long primary miRNAs transcripts. Primary miRNA transcripts are trimmed in the nucleus by the RNase Drosha, yielding premature hair-looped miRNAs of  $\pm 70$  nucleotides. These pre-miRNAs are transported to the cytoplasm where they are further processed by the RNase Dicer, resulting in mature miRNA. Mature miRNAs are then incorporated into the RNA-induced silencing complex, where the miRNA strand anneals to the 3' UTRs of target mRNAs, leading to the degradation or translation inhibition of mRNAs and subsequent protein repression (Figure 13A). miRNAs are mostly known for their gene regulation properties. However, they also play an important role in intercellular signaling and can therefore be found abundantly in bodily fluids, including blood and urine [313]. miRNA expression profiles show tissue-specific expression patterns and are shown to be different in diseased state.

In DM1, several studies have found changes in miRNA expression levels in blood, skeletal muscle and heart. The focus has been mainly on the four muscle-specific miRNAs (myo-miRs), which are expressed in both skeletal and cardiac muscle. miR-1, miR-133a/b, and miR-206 are highly enriched in skeletal muscle, whereas cardiac muscle predominantly shows expression of miR-1 and miR-133a/b. In skeletal muscle, miR-133a and miR-133b have been found to be downregulated [314,315], whereas miR-206 is upregulated [314,316]

and mixed results have been found for miR-1 (Figure 13B) [314,315,317,318]. Non-myomiRs were found to be either up or down regulated in DM1 muscle compared to controls [315,317,318] and these aberrant expression profiles were linked to changes in muscle development, alternative splicing, atrophy, transcription and free radical removal, among others [317,318]. In cardiac muscle, possibly due to the scarcity of tissue, only two studies to date have studied the miRNA expression levels. The myo-miRs abundant in cardiac muscle were found to be downregulated (Figure 13B), in addition to a variety of other miRNAs, twenty-two in total [319,320]. This downregulation was associated with arrhythmias, conduction defects and fibrosis.



**Figure 13. miRNAs in DM1.** A) Production of miRNAs by transcription through RNA polymerase II, resulting in pri-miRNAs, which are further trimmed by Drosha and transported to the cytoplasm where they are further processed by Dicer, and forms a complex with RISC, resulting in mRNA degradation. B) the most studied miRNAs for DM1 are the muscle-specific myo-miRs, a summary of their up or down regulations (indicated by arrows) are given for heart, blood and skeletal muscle. Abbreviations: RNA pol II= RNA polymerase II; pri-miRNA= primary microRNA; RISC= RNA-induced silencing complex; myo-miRs= muscle specific microRNAs.

Where skeletal and cardiac muscle show predominantly a downregulation pattern in miRNAs, the opposite is found in blood (Figure 13B), suggesting that blood does not mirror the changes happening in tissues. Nevertheless, the upregulation in blood might be caused by the leakage of certain miRNAs into the bloodstream due to muscle damage, which means they might still be used as biomarkers [321]. An array of miRNAs are found to be upregulated, including the four myo-miRs [315,322–326]. These aberrant expression profiles have been linked to muscle weakness [322,324] and progressive muscle wasting [323,325]. Interestingly, Pegoraro and collaborators studied miRNA expression levels after DM1 patients underwent a six week exercise rehabilitation program and found that the four myo-miRs were significantly decreased in parallel with improved muscle function [326].





# OVERVIEW



Myotonic Dystrophy type 1 is an incurable, progressive, multi-systemic disorder with a CTG expansion in the *DMPK* gene at the base of its pathology. The expanded mRNA transcripts emanating from the gene are entrapped inside the nucleus, deregulating splicing factors, and subsequently creating a spliceopathy that can account for several of the DM1 symptoms observed. However, the complex multi-systemic nature and wide variability in symptom manifestation cannot be completely accounted for and the identification of other disease modifiers is key to understanding the complete DM1 pathology, for both better disease management, as well as the identification of new therapeutic targets. Therefore, the main objective of this doctoral thesis was to investigate potential disease modifiers and study their link to DM1 clinical phenotype. The focus was on five main potential disease modifiers on a genetic (variant repeats), transcriptional (antisense transcription), protein (RAN translation), epigenetic (DNA methylation) and non-coding RNAs (miRNAs) level. This dissertation is subdivided into four chapters, outlining the findings on the disease modifiers.

Where it was long thought that the CTG expansion was an uninterrupted sequence, in the last two decades increasing evidence was found for the existence of variant repeats within the CTG expansion. To date, the known pathological variant expansions contain either unstable CCG, CTC, GGC or CAG sequence interruptions at the 3' end, and less frequently at the 5' end. Variant repeats are estimated to be present in 3-5% of the population and have been linked to disease phenotype, but results are controversial. Due to the low prevalence, only a few families to date have been identified and much remains unclear about the effect of variant repeats in the expanded allele on DM1 clinical phenotype. Improving our knowledge on the influence of variant repeats in DM1 pathology is essential for disease management, therapy development and genetic counseling (**CHAPTER I**).

Another emerging potential disease modifier is the mechanism of bidirectional transcription, where transcription occurs in both the sense and antisense direction at the gene locus. Antisense transcription in itself is not directly related to disease, as it is a common mechanism in humans, but the incorporation of the expanded repeat may worsen disease pathology. Expanded antisense transcripts have been found in several microsatellite expansion disorders, and have been associated with additionally toxicity. In DM1, its presence has been established, although its expression levels are controversial, and more importantly, its contribution to DM1 pathology is unknown (**CHAPTER II**). Moreover, the antisense strand has been found to give rise to polyglutamine proteins through the phenomenon of RAN translation. RAN translation has been proposed to add cytotoxicity in DM1 and could potentially be another disease modifier. However, to date only one study has been conducted in an extremely limited sample collection and much remains unknown about the contribution of RAN translation to DM1 pathology (**CHAPTER II**).

DNA methylation is a heritable, yet reversible epigenetic modification and is essential for silencing retroviral elements, regulating tissue-specific gene expression, genomic imprinting, and X chromosome inactivation. DNA methylation profiles gained

interested in DM1 due to the location of the expanded repeat, which happens to reside in a 3.5 kb CpG island. Additionally, the CTG repeat is flanked by two CTCF binding sites. Several studies have found aberrant DNA methylation profiles in these two regions flanking the repeat (upstream CTCF1 and downstream CTCF2) in DM1, but results are controversial with variable degrees of hypermethylation observed. More importantly, little is known about its contribution to the clinical phenotype. Nevertheless, it seems to be associated with the most severe form of DM1, highlighting the importance of further elucidating the contribution of epigenetics to DM1 pathology. This CpG island is not the only one present in the *DMPK* gene, as it harbors two additional islands and another resides in the adjacent *DMWD* gene. Whether these CpG islands have an altered DNA methylation status in DM1 is unknown. Moreover, whether a tissue-specific epigenetic landscape exist in DM1 is also a question that remains to be answered (**CHAPTER III**).

miRNAs, a class of small non-coding regulatory RNA molecules have been proposed as an excellent non-invasive biomarker for disease progression due to their wide abundance in extracellular body fluids, including blood serum and plasma. The original idea was to conduct a miRNA expression study by using the digital droplet PCR in both DM1 serum and muscle simultaneously, focusing on the four myo-miRs. Although we were in the process of setting up the technique, we were unfortunately forced to stop this line of research due to the COVID pandemic. Instead, it was decided to use that time to write a comprehensive review on the disease modifying and biomarker potential of miRNAs (**CHAPTER IV**) and upon return focus in the other research lines.



# **OBJECTIVES**





## **CHAPTER I**

*OBJECTIVE:* Study the presence of variant repeats in a Spanish cohort, consisting of 49 DM1 patients and identify potential links to the clinical phenotype.

## **CHAPTER II**

*OBJECTIVE:* Determine the presence of antisense transcription (the origin of the RAN-translated polyGln protein) and RAN-translated polyGln protein in DM1 primary cell cultures, e.g. myoblasts, skin fibroblasts and lymphoblastoids, and their potential role in DM1 pathology.

## **CHAPTER III**

*OBJECTIVE:* Analyze the CpG DNA methylation patterns of four CpG islands in the *DMPK* locus in three different tissues (blood, muscle, skin) of six subtypes of DM1 and its relationship to CTG expansion length and clinical phenotype and to establish whether DNA methylation is preserved in patient-derived cells.

## **CHAPTER IV**

*OBJECTIVE:* Elucidate the disease modifying effect and biomarker potential of miRNAs in DM1 through an exhaustive literature study.





## **RESULTS**



The results section is divided into four chapters, corresponding to the four articles that this doctoral thesis is comprised of:

**CHAPTER I.** A DM1 family with interruptions associated with atypical symptoms and late onset but not with a milder phenotype.

**CHAPTER II.** Characterization of RAN Translation and Antisense Transcription in Primary Cell Cultures of Patients with Myotonic Dystrophy Type 1.

**CHAPTER III.** An integrative analysis of DNA methylation pattern in Myotonic Dystrophy type 1 samples reveals a distinct DNA methylation profile between tissues and a novel muscle-associated epigenetic dysregulation.

**CHAPTER IV.** The Biomarker Potential of miRNAs in Myotonic Dystrophy Type I



## DIRECTOR'S REPORT REGARDING THESIS BY PAPERS

---

---

**CHAPTER I.** A DM1 family with interruptions associated with atypical symptoms and late onset but not with a milder phenotype

**Hum Mutat. 2020 Feb;41(2):420-431. doi: 10.1002/humu.23932**

Available from: <https://onlinelibrary.wiley.com/doi/10.1002/humu.23932>

*JOURNAL:* Human Mutation *QUARTILE:* Q1 *IMPACT FACTOR:* 4.878

*CONTRIBUTION STUDENT:* For this scientific article, the doctoral student shares first authorship with another PhD Student of the lab due to their equal contribution to the work. The doctoral student has single handedly screened the entire population for variant repeats through TP-PCR, has identified the family presented and was involved in the set-up of the sequencing to identify the found variant repeats and later in their interpretation. She was also in charge of this part of the manuscript and has written most of the first draft of the introduction and discussion and was heavily involved in the corrections after review, leading to the submitted manuscript.

**CHAPTER II.** Characterization of RAN Translation and Antisense Transcription in Primary Cell Cultures of Patients with Myotonic Dystrophy Type 1

**J Clin Med. 2021 Nov 25;10(23):5520. doi: 10.3390/jcm10235520**

Available from: <https://www.mdpi.com/2077-0383/10/23/5520>

*JOURNAL:* Journal of Clinical Medicine *QUARTILE:* Q1 *IMPACT FACTOR:* 5.583

*CONTRIBUTION STUDENT:* For this scientific article the doctoral student has first authorship and was the main force behind the development of this article in close partnership with her supervisors. She has done the majority of bench work, analysis and interpretation and wrote the first draft of the manuscript.

**CHAPTER III.** An integrative analysis of DNA methylation pattern in Myotonic Dystrophy type 1 samples reveals a distinct DNA methylation profile between tissues and a novel muscle-associated epigenetic dysregulation

### **Under Revision**

*CONTRIBUTION STUDENT:* For this scientific article the doctoral student has first authorship and was the main force behind the development of this article in close partnership with her supervisors. She has done the majority of bench work, analysis and interpretation and wrote the first draft of the manuscript.

**CHAPTER IV.** The Biomarker Potential of miRNAs in Myotonic Dystrophy Type I

**J Clin Med. 2020 Dec 4;9(12):3939. doi: 10.3390/jcm9123939**

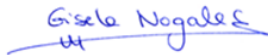
Available from: <https://www.mdpi.com/2077-0383/9/12/3939>

*JOURNAL:* Journal of Clinical Medicine *QUARTILE:* Q1 *IMPACT FACTOR:* 5.583

*CONTRIBUTION STUDENT:* For this scientific review, the doctoral student has first authorship and did the extensive literature study and subsequent first draft of the review, which was edited in collaboration with the other authors.

We hereby declare that the articles presented in this doctoral thesis are a direct result of the work from the doctoral student Emma Agathe Koehorst.

Dr. Gisela Nogales Gadea



Dr. Mònica Suelves









# CHAPTER I



## A DM1 family with interruptions associated with atypical symptoms and late onset but not with a milder phenotype

Alfonsina Ballester-Lopez<sup>1, 2,\*</sup>, **Emma Koehorst**<sup>1,\*</sup>, Miriam Almendrote<sup>1, 3</sup>, Alicia Martínez-Piñeiro<sup>1, 3</sup>, Giuseppe Lucente<sup>1, 3</sup>, Ian Linares-Pardo<sup>1</sup>, Judit Núñez-Manchón<sup>1</sup>, Nicolau Guanyabens<sup>3</sup>, Antoni Cano<sup>4</sup>, Alejandro Lucia<sup>5, 6</sup>, Gayle Overend<sup>7</sup>, Sarah A Cumming<sup>7</sup>, Darren G Monckton<sup>7</sup>, Teresa Casadevall<sup>8</sup>, Irina Isern<sup>9</sup>, Josep Sánchez-Ojanguren<sup>9</sup>, Albert Planas<sup>10</sup>, Agustí Rodríguez-Palmero<sup>1, 11</sup>, Laura Monlleó-Neila<sup>1, 11</sup>, Guillem Pintos-Morell<sup>1, 2, 12</sup>, Alba Ramos-Fransi<sup>1, 3</sup>, Jaume Coll-Cantí<sup>1, 2, 3</sup>, Gisela Nogales-Gadea<sup>1, 2</sup>

### AFFILIATIONS.

- <sup>1</sup> Neuromuscular and Neuropediatric Research Group, Institut d'Investigació en Ciències de la Salut Germans Trias i Pujol, Campus Can Ruti, Universitat Autònoma de Barcelona, Badalona, Spain.
- <sup>2</sup> Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III, Madrid, Spain.
- <sup>3</sup> Neuromuscular Pathology Unit, Neurology Service, Neuroscience Department, Hospital Universitari Germans Trias i Pujol, Barcelona, Spain.
- <sup>4</sup> Neurology Unit, Neuroscience Department, Hospital de Mataró, Barcelona, Spain.
- <sup>5</sup> Universidad Europea (Faculty of Sport Sciences), Madrid, Spain.
- <sup>6</sup> Instituto de Investigación Hospital 12 de Octubre (i+12), Madrid, Spain.
- <sup>7</sup> Institute of Molecular, Cell and Systems Biology, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, UK.
- <sup>8</sup> Neurology Service, Hospital Comarcal Sant Jaume de Calella, Barcelona, Spain.
- <sup>9</sup> Unitat de Neurologia, Hospital de l'Esperit Sant, Barcelona, Spain.
- <sup>10</sup> Servei de medicina interna, Secció de neurologia, Hospital Municipal de Badalona, Barcelona, Spain.
- <sup>11</sup> Neuropediatric Unit, Pediatric Service, Hospital Universitari Germans Trias i Pujol, Barcelona, Spain.
- <sup>12</sup> Division of Rare Diseases, University Hospital Vall d'Hebron, Barcelona, Spain.
- \* Equal contribution

**Hum Mutat. 2020 Feb;41(2):420-431. doi: 10.1002/humu.23932**

Available from: <https://onlinelibrary.wiley.com/doi/10.1002/humu.23932>





## SUMMARY OF THE RESULTS

---

---

To further elucidate the relation of variant repeats to the DM1 clinical phenotype, a Spanish cohort consisting of 49 DM1 patients (36 DM1 families) was screened for variant repeats by TP-PCR of both the 3' and 5' end region. In TP-PCR, interrupted alleles can be identified by gaps in the pattern of contiguous peaks, detectable by capillary electrophoresis. Five out of 49 DM1 patients (~10% at individual level, 3% at family level) presented with these gaps at the 3' end and were all part of the same family. P1, P2 and P3 were sisters, and P4 was the son of P2 and P5 the daughter of P3.

Their clinical phenotype revealed a delay in age of onset for the three sisters, with an age of onset >50 years. However, after initial onset, symptoms developed quite rapidly and included cardiac conduction defects and respiratory dysfunction, with variable severity. Interestingly, P1 and P2 presented with several traits atypical for DM1, such as severe axial weakness with dropped-head and moderate facial weakness, but without ptosis and no temporal atrophy. P4 (aged 35 years) was asymptomatic at the time of inclusion and P5 showed first symptoms at the age of 27 and at the time of examination revealed mild neck flexor and facial weakness and myotonia, but no limb weakness or other DM1 symptoms.

AciI digestion was performed to investigate the presence of CCG and/or GGC variant repeats. A downward shift of the smear in the gel confirmed the presence of one or both of these variant repeats in the five variant carrying DM1 patients. Since TP-PCR is limited to the outer regions of the CTG expansion, the digestion was additionally performed on the entire cohort to discard potential variant repeats further inside the CTG expansion and no new variant repeat carrying individual were found.

Sanger sequencing with specific primers targeting the potential variant repeats revealed the presence of a complex pattern of CCG interruptions in all five DM1 patients. Interruption patterns differed between the family members, except P2 and her son P4 whom showed identical interruption patterns. CCG variant repeats were present as single repeats, CCGCTG hexamers and short (CCG)<sub>n</sub> arrays. During the sequencing process, several bands were purified from the same patient to discard a potential influence of somatic instability on the differences observed between patients. The different bands belonging to the same patients showed identical patterns, discarding an effect of somatic instability on variant repeat pattern.

SP-PCR provided information on the ePAL (P1, 319 CTGs; P2, 241 CTGs; P4, 222 CTGs; P5, 547 CTGs) and somatic instability (P1, 319–900 CTGs; P2, 241–651 CTGs; P4, 222–332 CTGs; P5, 547–897 CTGs) of four out of the five interrupted patients. These revealed a contraction in the CTG expansion size from patient P2 to P4 (i.e. from mother to son) and an expansion from patient P3 to P5 (i.e. from mother to daughter). This expansion was linked to anticipation, with an earlier age of onset for P5 compared to her mother P2.

Altogether, a family carrying CCG variant repeats was identified in our Spanish cohort. Both a contraction and an expansion in the expanded allele upon generational transmission were associated with the presence of variant repeats and our clinical findings contribute to the observation that variant repeat carrying individuals present with atypical clinical features, hampering DM1 diagnosis, accompanied by a delayed age of onset and a previously unreported aging-related severe disease manifestation.

Received: 21 August 2018 | Revised: 18 September 2019 | Accepted: 6 October 2019

DOI: 10.1002/humu.23932

## RESEARCH ARTICLE

Human Mutation  WILEY

## A DM1 family with interruptions associated with atypical symptoms and late onset but not with a milder phenotype

Alfonsina Ballester-Lopez<sup>1,2\*</sup> | Emma Koehorst<sup>1\*</sup> | Miriam Almendrote<sup>1,3</sup> |  
 Alicia Martínez-Piñero<sup>1,3</sup> | Giuseppe Lucente<sup>1,3</sup> | Ian Linares-Pardo<sup>1</sup> |  
 Judit Núñez-Manchón<sup>1</sup> | Nicolau Guanyabens<sup>3</sup> | Antoni Cano<sup>4</sup> | Alejandro Lucia<sup>5,6</sup> |  
 Gayle Overend<sup>7</sup> | Sarah A. Cumming<sup>7</sup> | Darren G. Monckton<sup>7</sup> |  
 Teresa Casadevall<sup>8</sup> | Irina Isern<sup>9</sup> | Josep Sánchez-Ojanguren<sup>9</sup> | Albert Planas<sup>10</sup> |  
 Agustí Rodríguez-Palmero<sup>1,11</sup> | Laura Monlleó-Neila<sup>1,11</sup> | Guillem Pintos-Morell<sup>1,2,12</sup> |  
 Alba Ramos-Fransi<sup>1,3</sup> | Jaume Coll-Canti<sup>1,2,3</sup> | Gisela Nogales-Gadea<sup>1,2</sup>

<sup>1</sup>Neuromuscular and Neuropediatric Research Group, Institut d'Investigació en Ciències de la Salut Germans Trias i Pujol, Campus Can Ruti, Universitat Autònoma de Barcelona, Badalona, Spain

<sup>2</sup>Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III, Madrid, Spain

<sup>3</sup>Neuromuscular Pathology Unit, Neurology Service, Neuroscience Department, Hospital Universitari Germans Trias i Pujol, Barcelona, Spain

<sup>4</sup>Neurology Unit, Neuroscience Department, Hospital de Mataró, Barcelona, Spain

<sup>5</sup>Universidad Europea (Faculty of Sport Sciences), Madrid, Spain

<sup>6</sup>Instituto de Investigación Hospital 12 de Octubre (i+12), Madrid, Spain

<sup>7</sup>Institute of Molecular, Cell and Systems Biology, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, UK

<sup>8</sup>Neurology Service, Hospital Comarcal Sant Jaume de Calella, Barcelona, Spain

<sup>9</sup>Unitat de Neurologia, Hospital de l'Esperit Sant, Barcelona, Spain

<sup>10</sup>Servei de medicina interna, Secció de neurologia, Hospital Municipal de Badalona, Barcelona, Spain

<sup>11</sup>Neuropediatric Unit, Pediatric Service, Hospital Universitari Germans Trias i Pujol, Barcelona, Spain

<sup>12</sup>Division of Rare Diseases, University Hospital Vall d'Hebron, Barcelona, Spain

## Correspondence

Gisela Nogales-Gadea, Grup de Recerca en Malalties Neuromusculars i Neuropediatrics, Institut d'Investigació en Ciències de la Salut Germans Trias i Pujol, Ctra. de Can Ruti, Camí de les Escoles, s/n 08916 Badalona (Barcelona), Spain.  
 Email: gnogales@igtp.cat

## Funding information

Agència de Gestió d'Ajuts Universitaris i de Recerca, Grant/Award Number: FI\_B 01090; AFM-Téléthon, Grant/Award Number: #21108; "la Caixa" Foundation, Grant/Award Number: LCF/BQ/IN18/11660019; Instituto de Salud Carlos III, Grant/Award Numbers: CM16/00016, CP14/00032, CPII19/00021, PI15/00558, PI15/01756, PI18/00713

## Abstract

Carriage of interruptions in CTG repeats of the myotonic dystrophy protein kinase gene has been associated with a broad spectrum of myotonic dystrophy type 1 (DM1) phenotypes, mostly mild. However, the data available on interrupted DM1 patients and their phenotype are scarce. We studied 49 Spanish DM1 patients, whose clinical phenotype was evaluated in depth. Blood DNA was obtained and analyzed through triplet-primed polymerase chain reaction (PCR), long PCR-Southern blot, small pool PCR, *Acil* digestion, and sequencing. Five patients of our registry (10%), belonging to the same family, carried CCG interruptions at the 3'-end of the CTG expansion. Some of them presented atypical traits such as very late onset of symptoms (> 50 years) and a severe axial and proximal weakness requiring walking assistance. They also showed classic DM1 symptoms including cardiac and respiratory dysfunction, which were severe in some of them. Sizes and interrupted allele patterns were determined,

\*Alfonsina Ballester-Lopez and Emma Koehorst contributed equally to this work.



and we found a contraction and an expansion in two intergenerational transmissions. Our study contributes to the observation that DM1 patients carrying interruptions present with atypical clinical features that can make DM1 diagnosis difficult, with a later than expected age of onset and a previously unreported aging-related severe disease manifestation.

#### KEYWORDS

atypical symptoms, interruptions, late onset, myotonic dystrophy type 1, severe phenotype, Steinert disease, variant repeats

## 1 | INTRODUCTION

Myotonic dystrophy type 1 (DM1, Steinert disease; MIM# 160900) is a multisystemic disorder with an overall estimated prevalence of 1:8000 (Harper PS, 2001), being the most common form of inherited muscular dystrophy in adults. DM1 patients show wide phenotypic heterogeneity, not only in the age of onset but also in severity and type of clinical manifestation. DM1 patients can be broadly divided into five subtypes based mainly on their age of onset: congenital (<1 month), childhood (1 month–10 years), juvenile (10–20 years), adulthood/classic (20–40 years), or late onset (>40 years; De Antonio et al., 2016). Classic DM1 symptoms include muscle weakness, myotonia, respiratory failure, cardiac conduction defects, cataracts, and endocrine disturbances. The younger subtypes, congenital and childhood onset, are characterized primarily by cognitive and learning abnormalities (Douniol et al., 2012; Meola & Cardani, 2015).

DM1 is an autosomal dominant disorder caused by a CTG expansion in the 3' untranslated region of the myotonic dystrophy protein kinase (*DMPK*) gene. Unaffected individuals carry 5–35 CTG repeats whereas individuals carrying between 35 and 50 repeats are usually asymptomatic. Yet in the latter, *DMPK* alleles have a higher mutation rate and are labeled as "pre-mutational alleles" (Imbert, Kretz, Johnson, & Mandel, 1993). The length of the CTG expansion varies widely between patients, ranging from 50 to thousands of CTGs and has been associated with age of symptom onset and severity (Groh et al., 2011; Logigian et al., 2004). A CTG repeat size  $\leq 150$  CTGs,  $\leq 1,000$  CTGs and  $>1000$  CTGs is common to late onset, adulthood/classic and congenital DM1, respectively (Meola & Cardani, 2015). However, a high individual variability exists among DM1 patients of the same subtype and thus caution is needed when using CTG expansion length to predict disease progression. For instance, congenital cases have been found with CTG repeat lengths clearly below 1,000 CTG repeats (Tsilfidis, MacKenzie, Mettler, Barceló, & Korneluk, 1992) and late onset DM1 cases have been reported with over 1,000 CTG repeats (Clark, Petty, & Strong, 1998). Another feature of the disease that makes it difficult to infer potential genotype/phenotype correlations is the presence of "somatic mosaicism." Indeed, the CTG expansion is highly unstable in both germline and somatic cells, and this instability persists through the lifetime of the patient. Thus, the CTG repeat size of a given patient represents the mean value for different CTG repeat

sizes, which in turn can vary depending on the age at which the patient is studied. These potential confounders for sizing CTG repeat makes it difficult to find genotype-phenotype correlations for DM1. In this respect, estimating the inherited allele length has proven to be a more accurate predictor of potential genotype-phenotype correlations in this disease. (Higham, Morales, Cobbold, Haydon, & Monckton, 2012; Morales et al., 2012).

Because the aforementioned fact that CTG expansion instability is also present in germline cells, new alleles with different CTG repeat sizes are constantly generated and children may inherit CTG repeat sizes considerably longer than those found in the transmitting parent. This leads to the so-called "anticipation" phenomenon, which occurs in DM1 and in other triplet disorders, and is characterized by the fact that the disease may develop earlier in life in each successive generation (Harper, Harley, Reardon, & Shaw, 1992). In DM1, the sex of the transmitting parent plays an important role in anticipation, although both paternal and maternal transmission have been described. The paternal allele seems more unstable and leads more frequently to higher expansions in offspring, especially with CTG expansions below 100 repeats (Ashizawa et al., 1994). However, very large expansions causing congenital DM1 are transmitted almost exclusively by affected mothers (Harley et al., 1993), with few exceptions reported (Di Costanzo et al., 2009; Zeesman, Carson, & Whelan, 2002). In contrast, the congenital form is frequently observed after transmission from mothers who are carriers of more than 500 CTG repeats. Contractions of the CTG expansion upon transmission have also been reported, with a higher estimated prevalence in paternal transmission compared with maternal transmissions (6.7 vs. 19.5%; López de Munain et al., 1996).

In most cases, the CTG expansion in expanded *DMPK* alleles is an uninterrupted sequence. However, in the last decade, pathological variant expansions containing unstable CCG, CTC, GGC, and CAG sequence interruptions at the 3' and 5' ends of the *DMPK* allele have been reported, with a prevalence of 3–5% among DM1 (Botta et al., 2017; Braida et al., 2010; Cumming et al., 2018; Musova et al., 2009; Pešović et al., 2017; Santoro et al., 2013; Tomé et al., 2018). In addition, intergenerational transmissions typically lead to smaller CTG expansions when compared with noninterrupted DM1 families, suggesting a stabilizing effect of the expansion on germline transmission (Botta et al., 2017; Pešović et al., 2017; Tomé et al., 2018). These findings might also explain why no congenital cases have been described in maternal transmission of interrupted alleles.

Most of the phenotype consequences of interruptions remain poorly understood and vary considerably between studies, ranging from a complex neurological phenotype to a later age of onset (Botta et al., 2017; Braida et al., 2010; Cumming et al., 2018; Santoro et al., 2015; Musova et al., 2009; Pešović et al., 2017). There is an urgent need to determine the phenotypes that associate with the subset of DM1 patients presenting with interruptions. This information is required for patient management, genetic counseling and future clinical trials. In the literature, only a few families and some isolated cases have been described, and their reported clinical data are scarce. In the present study, we have analyzed a large cohort of Spanish DM1 patients belonging to several families. Our aim was to identify DM1 patients carrying variant repeats and to perform an in-depth analysis of their clinical phenotypes. This might help to gain insight into the modifying effect that these repeat interruptions could have in DM1 diagnosis, clinical manifestation, and patient follow-up.

## 2 | MATERIALS AND METHODS

### 2.1 | Editorial policies and ethical considerations

This study was approved by the ethics committee of the University Hospital *Germans Trias i Pujol* (ref. PI-15-129) and was performed in accordance with the Declaration of Helsinki for Human Research. Written informed consent was obtained from all the participants.

### 2.2 | Participants

Forty-nine DM1 patients belonging to 36 different families who were evaluated in our center during the 2015–2018 period participated in this study. Clinical and genetic information was collected and stored in a secure registry. Their clinical phenotype was evaluated by the neurologists of our team. Muscle strength was assessed using the manual Medical Research Council (MRC) scale. The most recent ophthalmological, cardiological and respiratory examinations carried out by the corresponding specialists were reviewed, as well as blood analyses, electrocardiograms, echocardiograms, and functional respiratory and swallowing tests. Functional status and disability were assessed using the Muscular Impairment Rating Scale (MIRS), the modified Rankin Scale (mRS), and the Rasch-Built Myotonic Dystrophy type 1 activity and participation scale (DM1-Activ).

### 2.3 | DNA extraction and bidirectional triplet-primed polymerase chain reaction (TP-PCR)

Total genomic DNA was extracted from peripheral blood samples, as previously described (Miller, Dykes, & Polesky, 1988). To assess the size and the presence of interruptions in the expanded allele, all DM1 blood DNA samples were analyzed by bidirectional TP-PCR. TP-PCR was performed with primers DM1for-FAM, DM1-CAG-rev, and P3 at the 5'-end of the CTG expansion, or DM1rev-FAM, DM1-CTG-for, and P3 at the 3'-end of the CTG expansion, as previously described by Radvansky, Ficek, Minarik, Palffy, and Kadasi (2011). Both

TP-PCRs (5' and 3') were performed with 100 ng of genomic DNA, 10× PCR Buffer containing 15 mM of MgCl<sub>2</sub>, 10 nM of dNTP mixture, 0.5 U of TaKaRa DNA polymerase (TaKaRa, Kusatsu, Japan), 3% dimethyl sulfoxide (DMSO), and 0.2 μM of each primer. PCR amplification conditions were the same for both TP-PCRs: initial denaturation at 94°C for 5 min, followed by 34 cycles at 94°C for 1 min, 65°C for 1 min, and 72°C for 2 min and a final extension step at 72°C for 7 min. Correct amplification was assessed on a 2% agarose gel. PCR products were separated on an ABI PRISM 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA, EEUU) and data were analyzed with PeakScanner Software v1.0 (Applied Biosystems).

### 2.4 | *Acil* digestion and Southern blot

We used a digestion with *Acil* and Southern blot-long PCR strategy to determine the presence of interruptions of the CCG/CGG type. DNA (100 ng) was amplified using the primers MDY1D-F GCTCGAA GGGTCCCTGTAGCCG and DM1-rev GTGCGTGGAGGATGGAA. The conditions of the long PCR were as follows: initial denaturation at 94°C for 4 min, followed by 35 cycles of denaturation at 94°C for 30 s and annealing-extension at 65°C for 7 min. The final extension was performed at 65°C for 10 min. Fifty microliters of long PCR products were divided into two parts, one digested with *Acil* and the other not digested. An aliquot (10 μl) of each sample was resolved in an agarose gel and the products were detected by Southern blot hybridization. A DIG-labeled LNA probe (5'-gcAgCagcAgCagCagcAgca-3', with lower and upper-case letters representing an unmodified and an LNA nucleotide, respectively) was used to detect the expansions through chemiluminescence.

### 2.5 | Sequencing

To determine the pattern of the interruptions we first amplified the DNA using primers GC1\_CC, GC1\_CCG, P2-rev, and P3, as described elsewhere (Pešović et al., 2017). Products were resolved in a 3% agarose gel and purified using QIAquick gel extraction kit (Werfen, Barcelona, Spain). Purified products were sequenced with BigDye™ Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). Sequences were analyzed with Chromas version 2.6.2. The *DMPK* gene reference sequence used was NG\_009784.1.

### 2.6 | Small pool PCR and *Acil* digestion

To estimate the length of the expanded progenitor allele (ePAL), small-pool PCR (SP-PCR) was carried out using flanking primers DM-C and DM-DR as previously described (Gomes-Pereira, Bidichandani, & Monckton, 2004; Monckton, Wong, Ashizawa, & Caskey, 1995). PCR was performed using Custom PCR Master Mix (Thermo Fisher Scientific, Waltham, MA, EEUU) supplemented with 69 mM 2-mercaptoethanol, and Taq polymerase (Sigma-Aldrich, Gillingham, UK) at 1 unit per 10 μl. All reactions were supplemented with 5% DMSO and the annealing temperature was 63.5°C. DNA fragments were resolved by

electrophoresis on a 1% agarose gel, and Southern blot hybridized as described (Gomes-Pereira et al., 2004; Monckton et al., 1995). Autoradiographic images were scanned and ePAL estimated from the lower boundary by comparison against the molecular weight ladder, using CLIQS 1D gel analysis software (TotalLab, Newcastle upon Tyne, UK). To analyze again the presence of CCG or CGG variant repeats, an additional step was added to the SP-PCR protocol. PCR products were purified using the QIAquick (Qiagen, Venlo, the Netherlands) PCR purification kit and split into two aliquots, one of which was digested with *Acil*. They were then resolved and blotted as before.

### 3 | RESULTS

#### 3.1 | Clinical phenotypes

Five of 49 DM1 patients (~10%) were found to have interruptions in the 3'-end of the CTG expansion. They belonged to the same family (Figure 1). Patient P1, P2, and P3 are sisters who paternally inherited the disease, and patient P4 is the son of P2 (Figure 1) whereas patient P5 is the daughter of P3. A summary of their clinical characteristics is shown in Table 1.

Patient P1 is the oldest of the siblings, and currently the most severely affected of all five patients. The first symptom she reported was a weakness at the age of 52. Subsequently, she developed a generalized weakness, which interfered with her ability to cope with daily life activities. We studied her when she was 72-year-old and the most striking feature upon clinical examination was a severe axial weakness with dropped-head. The patient also presented with mild weakness in the upper and lower limbs, with only little myotonia. Another remarkable fact was that she had moderate facial weakness, but almost no ptosis and no temporal atrophy. She also presented with bilateral cataracts, dysphagia for liquids and frontal baldness. She had a heart pacemaker implanted since the age of 71 and used nocturnal noninvasive mechanical ventilation, and had no cognitive impairment.

Patient P2: Symptoms started at the age of 50 with mild fatigue and myotonia. At the moment of inclusion in the study (aged 62), the clinical examination revealed the only mild weakness of the neck flexor muscles, with mild handgrip myotonia and minimum ptosis. Complementary explorations showed a first-degree atrioventricular block and low values of maximum inspiratory and expiratory pressure (38% and 22% of normal, respectively), and of both forced vital capacity (81%) and expiratory volume in 1 s (96%). The patient presented with bilateral cataracts and severe baldness. No limb weakness, dysphagia or cognitive impairment was found.

Patient P3: The first sign reported was handgrip myotonia in her fifties. At the moment of examination (age 60) she also had a severe axial weakness with mild proximal limb weakness and moderate distal weakness. Like her oldest sister (Patient P1), she had a moderate facial weakness with no ptosis or temporal atrophy. She also presented with bilateral cataracts, frontal baldness, and dysphagia for liquids. Cardiological studies revealed a first-degree

atrioventricular block. No respiratory or cognitive involvement was found.

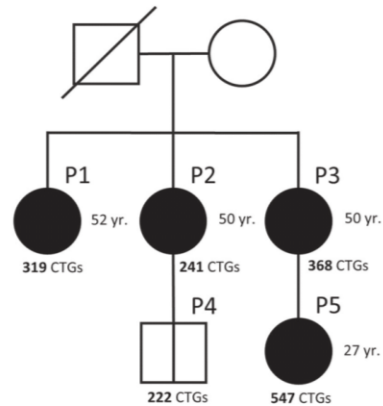
Patient P4: This male patient (aged 35 years) carrying an interrupted allele was asymptomatic upon clinical examination and had no detectable myotonia or cardiac alterations.

Patient P5: She was diagnosed at the age of 25 based on the family history, although clinical manifestation did not start until 2 years later, starting with handgrip myotonia. At the moment of assessment (age 32) she presented with mild neck flexor and facial weakness and handgrip and percussion myotonia, without limb weakness. She has a first-degree atrioventricular block and cataracts, but no respiratory impairment.

#### 3.2 | Molecular analysis of interruptions

Interrupted alleles were first detected as gaps in the pattern of contiguous peaks detectable by capillary electrophoresis by 3' TP-PCR (Figure 2). Patients P2 and P4 showed a similar interruption pattern, whereas P1, P3, and P5 showed different interruption patterns (Figure 2). No alterations were found with 5' TP-PCR (Figure S1).

We performed an *Acil* digestion of PCR products to test for the presence of CCG or GGC variant repeats in patients P1–P5. In all these patients, the results showed a downward shift of the smear in the gel of the digested product compared with the nondigested product (Figure 3). This indicated that *Acil* had cleaved the PCR product and the interruption was likely either a CCG or a GGC triplet. In addition, since the bidirectional TP-PCR performed in the entire cohort was limited to the outer regions of the CTG expansion, we performed an *Acil* digestion in the entire DM1 cohort to search for possible undetected CCG or GGC interruptions in the middle

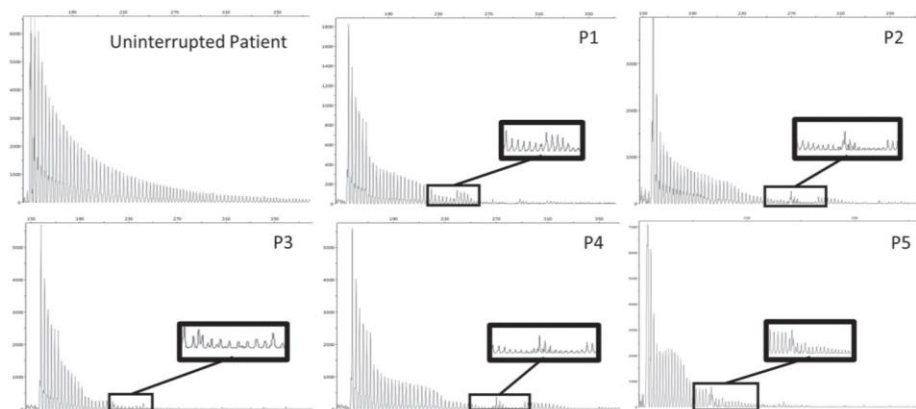


**FIGURE 1** Pedigree of the interrupted patients in our cohort. CTG, number of repeats in CTG; P1, patient P1; P2, patient P2; P3, patient P3; P4, patient P4; P5, patient P5. The father of P1, P2, and P3 died (sudden cardiac death)

**TABLE 1** Clinical characteristics of the interrupted cases

	P1	P2	P3	P4	P5
Sex	Female	Female	Female	Male	Female
Age of onset (years)	52	50	50	Asymptomatic	27
Age of assessment (years)	72	62	60	35	32
Cardiopathy	Pacemaker	1st degree AV-block	1st degree AV-block	None	1st degree AV-block
Respiratory disturbance	Yes, nocturnal NMV	Alteration MIP and MEP	None	None	None
Dysphagia	Liquids	No	Liquids	None	None
Cognitive impairment	None	None	None	None	None
Cataracts	Yes	Yes	Yes	None	Yes
Metabolic disturbance	None	None	Hypothyroidism	None	None
Myotonia	Yes	Yes	Yes	None	Yes
Polyneuropathy	None	None	None	None	None
CK level	Normal	Normal	213 U/L	Normal	ND
<i>Limb weakness</i>					
Facial ptosis	Yes	Mild	Yes	None	Mild
Flexor/extensor neck	1 (Dropped head)	4	2	5	5
Axial weakness	Severe	None	Severe	None	None
Upper limb proximal (MRC)	4	5	4	5	5
Upper limb distal (MRC)	4	5	3	5	5
Lower limb proximal (MRC)	4	5	4	5	5
Lower limb distal (MRC)	4	5	3	5	5
6 MWT (m)	250	436	240	658	800
MIRS	4	2	4	1	2
mRS	3	1	3	0	2
DM1-Activ	23	37	23	40	39

Abbreviations: 6 MWT, six-minute walking test; AV, atrioventricular; CK, creatine kinase; DM1-Activ, Rasch-built myotonic dystrophy type 1 activity and participation scale; MEP, maximum expiratory pressure; MIP, maximum inspiratory pressure; MIRS, Muscular Impairment Rating Scale; mRS, modified Rankin Scale; MRC, Medical Research Council; ND, not determined; NMV, noninvasive mechanical ventilation.



**FIGURE 2** Peak scan results of triplet-primed polymerase chain reaction (TP-PCR) of the 3'-end, obtained with DNA extracted from blood. Interruptions indicated by black box. P1, patient P1; P2, patient P2; P3, patient P3; P4, patient P4; P5, patient P5

region of the CTG expansion. No additional CCG or GGC interruptions were found in our 44 remaining DM1 patients.

Sequencing revealed the presence of several CCG interruptions in the CTG expansion of our five patients carrying interrupted alleles (Figure 4). The pattern of CCG interruptions was identical in the mother (patient P2) and son (P4), but different between all the other family members (patients P1, P3, and P5). In P1, we found some isolated CCG repeats scattered across the expansion. Patients P2 and P4 showed a complex CCG pattern, with one pair of CCGs together with other isolated CCG repeats. Patient P3 had a few CCGCTG hexamers, but inside of a more complex pattern including CCG interruptions in other positions. Patient P5 (the daughter of P3) showed a pattern similar to that of her mother with respect to the hexamers, but with some extra CCGs located in different positions, generating three consecutive CCG repeats. During the sequencing process, we purified different bands from the same patient to assess the influence of somatic instability in the interruption pattern (Figure 4b). In the different bands analyzed, the same pattern was observed in each patient.

SP-PCR (Figure 5) provided information on the repeat size of the ePAL for some of the patients: P1, 319 CTGs; P2, 241 CTGs; P4, 222 CTGs; P5, 547 CTGs. The expansion range due to the instability of the repeat was also determined: P1, 319–900 CTGs; P2, 241–651 CTGs; P4, 222–332 CTGs; and P5 = 547 to 897 CTGs. For patient P3, the expanded allele did not amplify well under these conditions, so it was not possible to determine the ePAL or expansion range. This may be due to the specific pattern of variant repeats present. However, in the SP-PCR we could amplify 368 CTGs, which was the only sizing of the expanded allele that we could make. Additionally, the type of interruptions was analyzed through *Acil* digestion in SP-PCR experiments, which again showed that the interruptions were of the CCG type (data not shown). By comparing the range of these bands, we could determine a contraction in the repeat size from patient P2 to P4 (i.e., from mother to son) but expansion from patient P3 to P5 (i.e., from mother to daughter). This expansion was also linked to anticipation, with an early age of onset for P5 when compared with her mother (P3).

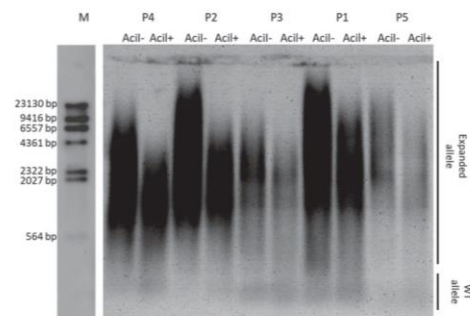
#### 4 | DISCUSSION

The effect of variant repeat patterns on the DM1 clinical phenotype is still unclear. On one hand, this genetic alteration has been shown to be associated (albeit in one family only) with a complex co-segregated neurological phenotype, including an intermediate Charcot-Marie-Tooth neuropathy, early hearing loss and encephalopathic attacks (Braidia et al., 2010). In contrast, variant repeats have been associated with a milder or atypical phenotype, including a later age of onset (Cumming et al., 2018; Musova et al., 2009; Pešović et al., 2017), a DM2-like muscle phenotype (Pešović et al., 2017), as well as with an absence of muscular dystrophy (Musova et al., 2009) or central nervous system symptoms (Santoro, Masciullo, Silvestri, Novelli, & Botta, 2017). These reports have led to a tendency to believe that patients with interrupted alleles have some

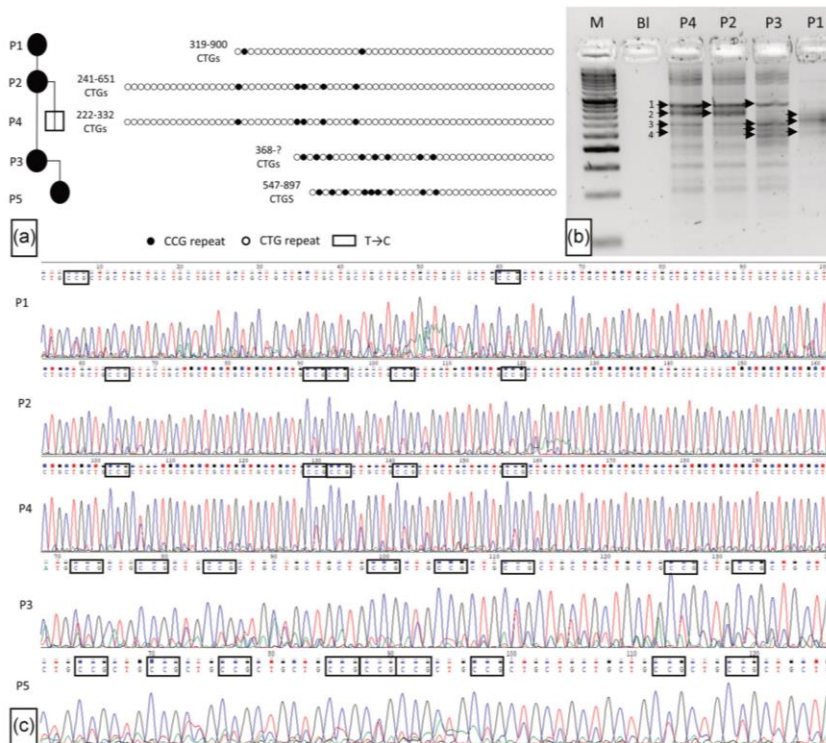
atypical symptoms, but overall a milder phenotype than their age-matched DM1 noninterrupted peers with a similar repeat length. In this respect, we had the unique possibility to study a family containing interrupted cases of whom three were aged above 60 years. In this regard, although our data were obtained in a small number of patients within the same family, our results support the occurrence of atypical DM1 features and late age of onset, but not of a milder phenotype in patients carrying interruptions.

Despite the fact that several of the classical symptoms of DM1 could be found in the three sisters—such as myotonia, cataracts, and cardiopathy—some peculiarities need to be highlighted. An atypical trait was the distribution pattern of muscle weakness in two of the sisters. Indeed, besides the distal limb weakness commonly found in DM1 patients, these two sisters presented with proximal limb weakness and severe axial involvement. One of them also had a dropped-head, which resembled limb-girdle muscle dystrophy and severely affected her ability to perform activities of daily living. Another atypical trait of these patients is that they did not have the typical myopathic face expected in DM1 patients, despite the presence of moderate facial weakness. Although the interrupted cases showed several classic DM1 symptoms, the presentation of atypical symptoms could interfere with (and thus delay) the diagnosis.

Based on the algorithm published by Morales et al (Morales et al., 2012), the ePAL of patients P1, P2 and the expanded allele size of P3 should be theoretically associated with an age of onset around 30 s whereas in our patients symptoms did not actually start until they were in their 50 s. In this regard, it should be first noted that it is very difficult to assess the age of onset in DM1 patients. The definition of age of onset refers to the age at which an individual starts to develop one or more clinical features or symptoms of a disease. In actual clinical practice, this depends on the capacity of the patient to report such symptoms or to remember the time when they started, and also



**FIGURE 3** Southern blot of long polymerase chain reaction (PCR) products from patients carrying variant repeats. For each patient, we show two conditions: digestion with (+) and without (-) the enzyme *Acil* (recognizing the pattern CCGC). bp, base pairs; M, molecular weight marker; P1, patient P1; P2, patient P2; P3, patient P3; P4, patient P4; P5, patient P5; WT, wild type

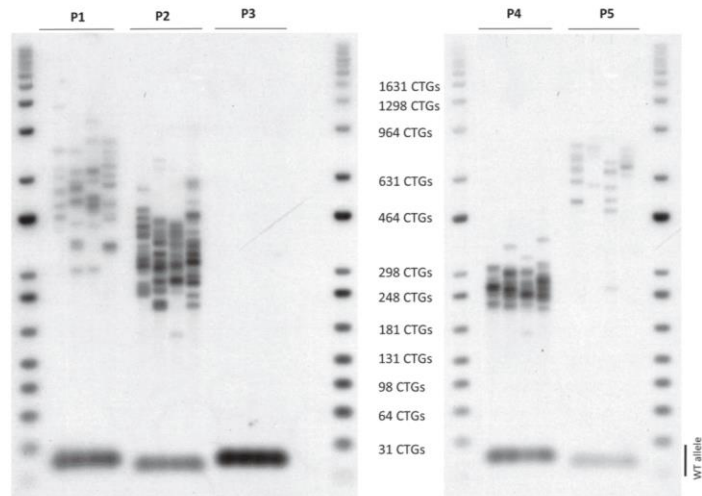


**FIGURE 4** Sequencing the interrupted alleles. (a) Schematic structure of *DMPK* expanded alleles of the interrupted DM1 family in our study. CTG repeats are shown in white, CCG repeats in black. Indicated size ranges were estimated by Southern blot analysis. (b) Cutting and purifying strategy for several bands of polymerase chain reaction product (indicated by arrows), from each interrupted patient, which are affected by somatic instability. (c) Sequences showing the CCG interruptions are marked by black rectangles. BI, PCR reaction with no DNA; M, molecular weight marker; P1, patient P1; P2, patient P2; P3, patient P3; P4, patient P4; P5, patient P5. The *DMPK* gene reference sequence used was NG\_009784.1

on the ability of the physician to recognize them. Thus, the reported age of onset can be quite variable, depending on which symptoms are searched for by the physician and on the patient's own reports. Our patients P1–P3 (the three sisters) reported their first symptoms in their 50s (being myotonia, and difficulty to walk, the first abnormalities that made them suspect they had a major medical condition). Patient P4 was a 35-year-old and was still asymptomatic. Based on his ePAL length (222 CTGs), he should show a classic DM1 phenotype (Morales et al., 2012), but no signs could be detected upon neurological examination. This late onset of symptoms has been previously reported in interrupted DM1 families and seems to be a fingerprint for most of the cases (Botta et al., 2017; Cumming et al., 2018; Musova et al., 2009; Pešović et al., 2017).

In the family we studied, anticipation was observed in one of the two intergenerational transmissions that we assessed, since in the other intergenerational transmission one of the patients (P4) was still

asymptomatic. In patient P5 (whose first symptom was myotonia, at the age of 27), we found a bigger size of the expansion and an earlier age of onset than her progenitor. Although this anticipation in interrupted families has been previously reported (Pešović et al., 2017), after reviewing all the published families (Table 2), we assessed anticipation in every single reported family. In the rest of intergenerational transmissions reported, and in the case of our patients P2 and P4, anticipation could not be assessed since patients in the next generation are still asymptomatic. The explanation for these findings is not apparent, since anticipation is not expected in these families; indeed, interruptions are thought to be related to a stabilization or even contraction of the pathological expansion (Braidă et al., 2010; Cumming et al., 2018; Musova et al., 2009; Pešović et al., 2017; Tomé et al., 2018). However, anticipation was found in our studied intergenerational transmission, with this finding being also reported in other interrupted DM1 patients based on reported age of onset (Table 2). In our family, no



**FIGURE 5** Small pool polymerase chain reaction from patients carrying variant repeats. For each patient, several lines show the normal and the expanded alleles. CTGs, number of repeats; P1, patient P1; P2, patient P2; P3, patient P3; P4, patient P4; P5, patient P5; WT allele, wild type allele

congenital, childhood or juvenile cases of DM1 were observed. Among the interrupted families reported in the literature (Table 2), at least three juvenile DM1 cases (age < 18 years) have been described (Braidà et al., 2010; Pešović et al., 2017), but no congenital or childhood case. Thus, the absence of infantile DM1 seems also to be a distinctive trait for interrupted expansions.

The prevalence of interrupted alleles among our patients was ~10%, and 3% among the studied DM1 families. This is in overall agreement with previous studies in which the prevalence in families ranged from 3% to 5% (Botta et al., 2017; Braidà et al., 2010; Musova et al., 2009; Pešović et al., 2017). The type of interruption present in our cohort was CCG, which is currently the most frequently reported variant repeat. However, the difficulties we experienced in characterizing the pattern of interruptions in our family members must be emphasized, with such difficulties mainly due to a technical limitation of TP-PCR and sequencing, which have a limited ability to detect interruptions deeper inside the expansion. In addition, characterization is affected by the PCR slippage and by somatic mosaicism (implying more noise in readouts and thus a higher difficulty to identify the interrupted pattern). Our sequences showed in some cases double peaks of C and T at the same position, and we decided to consider only those interruptions where the C peaks were above T in the electropherograms, which might have resulted in the loss of CCG interruptions in our patients' sequences. We sequenced several amplified bands coming from the same TP-PCR to determine whether somatic mosaicism was also affecting the pattern of CCG interruptions, but the same patterns were found in all the sequences. Despite the aforementioned limitations, we determined the interrupted

pattern in all the studied family members. We observed a substantial change in the interruption pattern in every transmission, and the number and position of the CCGs changed in every generation, except for one intergenerational transmission. Thus, interruption patterns can be conserved or vary upon transmission. Both situations have been previously described in interrupted families (Musova et al., 2009; Pešović et al., 2017; Tomé et al., 2018). Our CCG interruptions were found in blocks of two or three (in hexamers of CCGCTG that were repeated two or three times), and also as isolated cases. Due to technical limitations, we cannot be certain that other interruptions are not present deeper in the CTG expansion. TP-PCR and sequencing allowed us to study the flanking regions of the CTG expansion, but the middle part remained undetected. We detected the contraction of the expansion between patients P2 and P4, but expansion between patients P3 and P5. Previous studies suggest that CTG expansion containing variant repeat patterns display more frequently stable, or even contracted, *DMPK* alleles instead of further expanded *DMPK* alleles (Cumming et al., 2018; Musova et al., 2009; Pešović et al., 2017; Tomé et al., 2018). However, some studies have also found the expansion of the interrupted alleles from one generation to the other (Braidà et al., 2010; Cumming et al., 2018; Pešović et al., 2017). Perhaps these expansions are less frequent than in pure CTG expansions transmission, but they do occur. Therefore, caution is needed with genetic counseling with regard to prospective parents with DM1.

Our study contributes to the observation that DM1 patients carrying interruptions may have atypical symptoms that can make the diagnosis of DM1 difficult, with a later age of onset and a previously

TABLE 2 Analysis of the literature reported interrupted families

Article	Cases	Patient code	Age S	Age O	Relationship	Anticipation	Atypical findings in the examination	CTG repeats number	Type of interruption		
Musova et al. (2009)	3 families	Family A	A-1	0	Fetus of A2			230	CTC and CCG		
			(Fetus)								
			A-2	31	Daughter of A4	NPS			300	CTC and CCG	
			A-3	23	Daughter of A4	NPS			400–500	CTC and CCG	
			A-4	54	Brother of A5				600–800	CTC and CCG	
			A-5	53	Sister of A4				450–650	CTC and CCG	
		Family B	A-6	29	Daughter of A5	NPS			600–750	CTC and CCG	
			A-7	31	Son of A5	NPS			270	CTC and CCG	
			B-1	50	Father of B2				450	CTC and CCG	
		Family E	B-2	25	Son of B2	NPS			400	CCG	
			E-1	20	1 year <sup>a</sup>	Son of E2			43	CCG	
			E-2	56	Father of E1	A			43	CCG	
Braida et al. (2010)	1 family	Family 1	III-9	55	Cousin of III-17 and III-16		Charcot-Marie-Tooth disease, acute encephalopathy, and early hearing loss	229 <sup>b</sup>	CCG and GGC		
			III-16	57	Sister of III-17		Charcot-Marie-Tooth disease, acute encephalopathy, and early hearing loss	170 <sup>b</sup>	CCG and GGC		
			IV-19	37	Son of III-16	A	Charcot-Marie-Tooth disease, acute encephalopathy, and early hearing loss	213 <sup>b</sup>	CCG and GGC		
			IV-20	34	Son of III-16	A	Charcot-Marie-Tooth disease	213 <sup>b</sup>	CCG and GGC		
			III-17	61	Sister of III-16		Charcot-Marie-Tooth disease, acute encephalopathy, and early hearing loss	179 <sup>b</sup>	CCG and GGC		
			IV-21	30	Son of III-17	A	Charcot-Marie-Tooth disease	220 <sup>b</sup>	CCG and GGC		
			IV-22	28	Daughter of III-17	A	Charcot-Marie-Tooth disease	225 <sup>b</sup>	CCG and GGC		
Botta et al., 2017	3 families	Family A	A1	66	Father of A2		Absence of myotonia and cataracts	1,000–1,400	CCG		
			A2	39	Daughter of A1	A	Absence of muscle weakness	475–640	CCG		
			A3	0	Fetus of A2			500	CCG		
		Family B	B1	55	Mother of B2			740–930	CCG		
			B2	28	Daughter of B1	NPS		450–550	CCG		
			C1	58	Mother of C2			140	CCG		
		Family C	C2	40	Daughter of C1	A		121	CCG		
			C3	0	Fetus of C2			113	CCG		
		Pešović et al. (2017)	3 families	Family DF1	DF1-1	57	Mother of DF1-2 and DF1-3		Absence of: percussion myotonia, ptosis, cataracts, and muscle wasting. Presence of calf hypertrophy, suggesting DM2	520–1,250	CCG
					DF1-2	37	Son of DF1-1	A		370–730	CCG

(Continues)



TABLE 2 (Continued)

Article	Cases	Patient code	Age S	Age O	Relationship	Anticipation	Atypical findings in the examination	CTG repeats number	Type of interruption
Cumming et al. (2018)	Family DF2	DF1-3	30	15#	Son of DF1-1	A	Calf hypertrophy	450-970	CCG
		DF2-1	45	40	Father of DF2-2		Similar involvement of both proximal and distal muscles and winging scapulae in the right side.	320-600	CCG
	Family DF5	DF2-2	14	12#	Daughter of DF-1	A	Mild ptosis and mild percussion myotonia	200-240	CCG
		DF5-2	27	22	Sister of DF5-3		Normal strength of the sternocleidomastoid muscle and very mild myotonia	250-350	CTC
	DF5-3	22	21	Sister of DF5-2			300-620	Noninterrupted	
Our study	Family 1	14	25.5	-	Daughter of 165	NPS	Absence of muscle weakness, myotonia, and cataracts	381 <sup>ePAL</sup>	CCG
		57	20.5	5	Son of 165	A		357 <sup>ePAL</sup>	Noninterrupted
	Family 2	165	59	28	Brother of 83			383 <sup>ePAL</sup>	Noninterrupted
		83	46	38	Brother of 165			105 <sup>ePAL</sup>	Noninterrupted
		182	35.5	-	Brother of 184	NPS	Absence of muscle weakness, mild masseter myotonia and peripheral membrane irritability on EMG.	293 <sup>ePAL</sup>	CCG
	Family 3	184	28	20	Brother of 182			288 <sup>ePAL</sup>	Noninterrupted
		206	70	60	Father of 182 and 184			90 <sup>ePAL</sup>	Noninterrupted
		242	65	ND	Sister of 206			80 <sup>ePAL</sup>	Noninterrupted
		15	39	-	Daughter of 234	NPS	No clinical apparent weakness or myotonia and no cataracts	303 <sup>ePAL</sup>	CCG
		54	40	35	Brother of 15			146 <sup>ePAL</sup>	Noninterrupted
234	ND	ND	Father of 15 and 54			496 <sup>ePAL</sup>	Noninterrupted		
Family 1	Patient 1	72	52	Sister of Patient 2 and 3		Severe axial weakness with dropped-head, Mild weakness in upper and lower limb muscles, with only little myotonia.	319 <sup>ePAL</sup>	CCG	
		62	50	Sister of Patient 1 and 3		Moderate facial weakness, almost no ptosis, and no temporal atrophy	241 <sup>ePAL</sup>	CCG	
	Patient 3	60	50	Sister of Patient 1 and 2		Mild weakness of the neck flexor muscles, no limb weakness, and minimum ptosis	368	CCG	
		35	32	Son of Patient 3	NPS	Severe axial weakness, mild proximal limb weakness, and moderate distal weakness. Moderate facial weakness with no ptosis or temporary atrophy	222 <sup>ePAL</sup>	CCG	
	Patient 5	27	27	Daughter of Patient 2	A	Mild neck flexor and facial weakness, but no limb weakness.	547 <sup>ePAL</sup>	CCG	

Abbreviations: -, asymptomatic; A, anticipation; Age S, age of Sampling; Age O, age at Onset; ePAL, estimated progenitor allele; N/A, not applicable; ND, no data; NPS, not possible to establish yet.

\*Authors clarify in the paper that he had isolated symptoms, cannot be considered childhood DM1.

unreported aging-related severe disease manifestation. Indeed, some of our older patients needed mechanical ventilation and a pacemaker, and besides their cardiorespiratory problems, they had muscle weakness with subsequent impairment in daily life activities and walking ability. Despite the small sample size of our study sample, our results challenge the notion that interrupted patients who remain asymptomatic until their late 30s or 40s are not at risk for having a severe phenotype later in life. Indeed, our patients developed a classical DM1 phenotype after their 50s. These patients require clinical follow-up and genetic counseling similar to noninterrupted DM1 patients. In the family we studied, we found some characteristics that add to the current body of knowledge regarding interrupted families: a later age of onset, variation of CCG repeat pattern between intergenerational transmission, anticipation due to the earlier age of onset of symptoms in next generation and no cases of congenital or childhood onset of DM1. In addition, we have found other previously undescribed characteristics, such as a predominant axial weakness. However, the small number of interrupted patients present in the DM1 population makes it hard to perform genotype-phenotype correlations and there is still much uncertainty. Studies with larger DM1 cohorts, preferably with DM1 families, are needed to unravel the phenotypic consequences of variant repeat patterns and to study their effect on intergenerational transmissions of the DMPK expanded allele.

#### ACKNOWLEDGMENTS

We gratefully acknowledge to other researchers in the Myotonic dystrophy type I for their insightful advices for sequencing interruptions and their discussions regarding clinical data. The research of Gisela Nogales-Gadea and Alejandro Lucia is funded by Instituto de Salud Carlos III (grant numbers PI15/01756, PI15/00558, and PI18/00713) and co-financed by Fondos FEDER. Gisela Nogales-Gadea is supported by a Miguel Servet research contract (ISCIII CD14/00032, ISCIII CPII19/00021 and FEDER) and by a Trampoline Grant #21108 from AMF Telethon. Alfonsina Ballester-Lopez is funded by an FI Agaur fellowship ref. F1\_B 01090. Emma Koehorst is funded by the "La Caixa" Foundation (ID 100010434), fellowship code LCF/BQ/IN18/11660019, cofunded by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement n°713673. Ian Linares-Pardo is funded by CP14/00032. Judit Núñez-Manchón is funded by AFM Telethon Trampoline Grant #21108. Giuseppe Lucente is supported by a Rio Hortega contract (ISCIII CM16/00016 and FEDER). Darren Monckton, Gayle Overend, and Sarah Cumming received funding from the Myotonic Dystrophy Support Group (UK). The funding bodies had no role in the design of the study and collection, analysis, and interpretation of data.

#### CONFLICTS OF INTEREST

G. P.-M reports personal honoraria from Shire, and Sanofi-Genzyme, outside the submitted work. D. G. M. has been a scientific consultant and/or received honoraria or stock options from Biogen Idec, AMO Pharma, Charles River, Vertex Pharmaceuticals, Triplet Therapeutics,

LoQus23, BridgeBio, Small Molecule RNA and Lion Therapeutics and he also had a research contract with AMO Pharma. The remaining co-authors declare no conflicts of interest.

#### ORCID

Alfonsina Ballester-Lopez  <http://orcid.org/0000-0002-8922-1664>  
 Emma Koehorst  <http://orcid.org/0000-0002-8830-7941>  
 Miriam Almendrote  <http://orcid.org/0000-0003-0212-275X>  
 Alicia Martínez-Piñero  <http://orcid.org/0000-0003-1988-606X>  
 Giuseppe Lucente  <http://orcid.org/0000-0003-1120-9136>  
 Ian Linares-Pardo  <http://orcid.org/0000-0002-5725-1201>  
 Judit Núñez-Manchón  <http://orcid.org/0000-0001-5154-1454>  
 Nicolau Guanyabens  <http://orcid.org/0000-0002-2074-2744>  
 Alejandro Lucia  <http://orcid.org/0000-0002-3025-2060>  
 Gayle Overend  <http://orcid.org/0000-0003-4033-1747>  
 Sarah A. Cumming  <http://orcid.org/0000-0002-0201-3660>  
 Darren G. Monckton  <http://orcid.org/0000-0002-8298-8264>  
 Agustí Rodríguez-Palmero  <http://orcid.org/0000-0002-4141-5515>  
 Guillem Pintos-Morell  <http://orcid.org/0000-0002-9347-2386>  
 Alba Ramos-Fransi  <http://orcid.org/0000-0002-4114-4575>  
 Jaume Coll-Cantí  <http://orcid.org/0000-0001-7128-1186>  
 Gisela Nogales-Gadea  <http://orcid.org/0000-0002-7414-212X>

#### REFERENCES

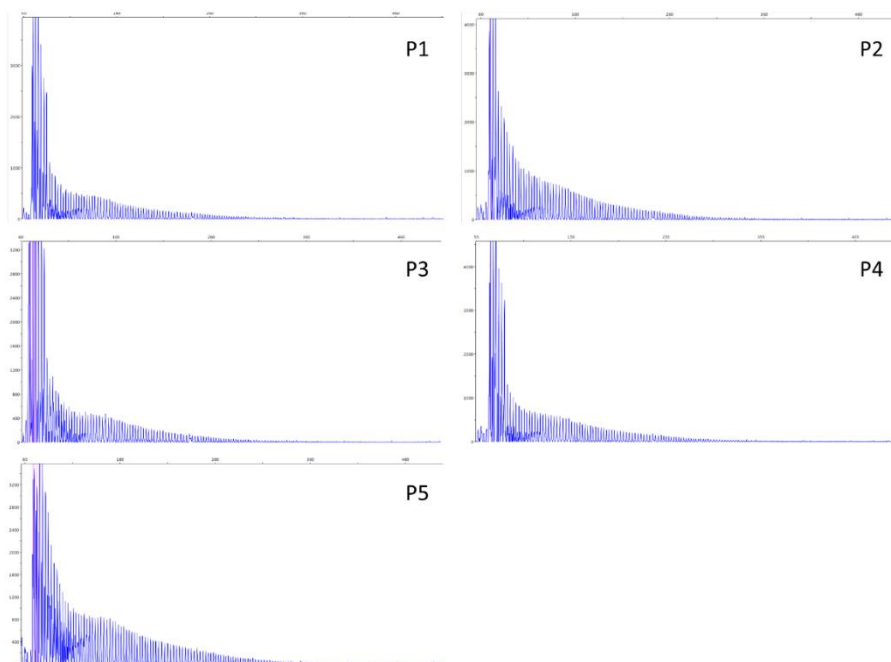
- Ashizawa, T., Anvret, M., Baiget, M., Barceló, J. M., Brunner, H., Cobo, A. M., & Harley, H. (1994). Characteristics of intergenerational contractions of the CTG repeat in myotonic dystrophy. *American Journal of Human Genetics*, 54(3), 414–423.
- Botta, A., Rossi, G., Marcarello, M., Fontana, L., D'Apice, M. R., Brancati, F., & Novelli, G. (2017). Identification and characterization of 5' CCG interruptions in complex DMPK expanded alleles. *European Journal of Human Genetics*, 25(2), 257–261. <https://doi.org/10.1038/ejhg.2016.148>
- Braida, C., Stefanatos, R. K. A., Adam, B., Mahajan, N., Smeets, H. J. M., Niel, F., & Monckton, D. G. (2010). Variant CCG and GGC repeats within the CTG expansion dramatically modify mutational dynamics and likely contribute toward unusual symptoms in some myotonic dystrophy type 1 patients. *Human Molecular Genetics*, 19(8), 1399–1412. <https://doi.org/10.1093/hmg/ddq015>
- Clark, C., Petty, R. K., & Strong, A. M. (1998). Late presentation of myotonic dystrophy. *Clinical and Experimental Dermatology*, 23(1), 47–48.
- Cumming, S. A., Hamilton, M. J., Robb, Y., Gregory, H., McWilliam, C., Cooper, A., & Monckton, D. G. (2018). De novo repeat interruptions are associated with reduced somatic instability and mild or absent clinical features in myotonic dystrophy type 1. *European Journal of Human Genetics*, 26(11), 1635–1647. <https://doi.org/10.1038/s41431-018-0156-9>
- De Antonio, M., Dogan, C., Hamroun, D., Mati, M., Zerrouki, S., & Eymard, B., French Myotonic Dystrophy Clinical Network. (2016). Unravelling the myotonic dystrophy type 1 clinical spectrum: A systematic registry-based study with implications for disease classification. *Revue Neurologique*, 172(10), 572–580. <https://doi.org/10.1016/j.neuro.2016.08.003>
- Di Costanzo, A., de Cristofaro, M., Di Iorio, G., Daniele, A., Bonavita, S., & Tedeschi, G. (2009). Paternally inherited case of congenital DM1: Brain MRI and review of literature. *Brain & Development*, 31(1), 79–82. <https://doi.org/10.1016/j.braindev.2008.04.008>

- Douniol, M., Jacquette, A., Cohen, D., Bodeau, N., Rachidi, L., Angeard, N., & Guilé, J.-M. (2012). Psychiatric and cognitive phenotype of childhood myotonic dystrophy type 1. *Developmental Medicine and Child Neurology*, 54(10), 905–911. <https://doi.org/10.1111/j.1469-8749.2012.04379.x>
- Gomes-Pereira, M., Bidichandani, S. I., & Monckton, D. G. (2004). Analysis of unstable triplet repeats using small-pool polymerase chain reaction. *Methods in Molecular Biology*, 277, 61–76. <https://doi.org/10.1385/1-59259-804-8:061>
- Groh, W. J., Groh, M. R., Shen, C., Monckton, D. G., Bodkin, C. L., & Pascuzzi, R. M. (2011). Survival and CTG repeat expansion in adults with myotonic dystrophy type 1. *Muscle & Nerve*, 43(5), 648–651. <https://doi.org/10.1002/mus.21934>
- Harley, H. G., Rundle, S. A., MacMillan, J. C., Myring, J., Brook, J. D., Crow, S., & Harper, P. S. (1993). Size of the unstable CTG repeat sequence in relation to phenotype and parental transmission in myotonic dystrophy. *American Journal of Human Genetics*, 52(6), 1164–1174.
- Harper, P. S., Harley, H. G., Reardon, W., & Shaw, D. J. (1992). Anticipation in myotonic dystrophy: New light on an old problem. *American Journal of Human Genetics*, 51(1), 10–16.
- Harper PS (2001). *Major problems in neurology: Myotonic dystrophy* (3rd ed.). London: WB Saunders.
- Higham, C. F., Morales, F., Cobbold, C. A., Haydon, D. T., & Monckton, D. G. (2012). High levels of somatic DNA diversity at the myotonic dystrophy type 1 locus are driven by ultra-frequent expansion and contraction mutations. *Human Molecular Genetics*, 21(11), 2450–2463. <https://doi.org/10.1093/hmg/dds059>
- Imbert, G., Kretz, C., Johnson, K., & Mandel, J. L. (1993). Origin of the expansion mutation in myotonic dystrophy. *Nature Genetics*, 4(1), 72–76. <https://doi.org/10.1038/ng0593-72>
- Logigian, E. L., Moxley, R. T., Blood, C. L., Barbieri, C. A., Martens, W. B., Wiegner, A. W., & Moxley, R. T. (2004). Leukocyte CTG repeat length correlates with severity of myotonia in myotonic dystrophy type 1. *Neurology*, 62(7), 1081–1089.
- López de Munain, A., Cobo, A. M., Sáenz, A., Blanco, A., Poza, J. J., Martorell, L., & Baiget, M. (1996). Frequency of intergenerational contractions of the CTG repeats in myotonic dystrophy. *Genetic Epidemiology*, 13(5), 483–487. [https://doi.org/10.1002/\(SICI\)1098-2272\(1996\)13:5<483::AID-GEPI4>3.0.CO;2-3](https://doi.org/10.1002/(SICI)1098-2272(1996)13:5<483::AID-GEPI4>3.0.CO;2-3)
- Meola, G., & Cardani, R. (2015). Myotonic dystrophies: An update on clinical aspects, genetic, pathology, and molecular pathomechanisms. *Biochimica et Biophysica Acta—Molecular Basis of Disease*, 1852(4), 594–606. <https://doi.org/10.1016/j.bbdis.2014.05.019>
- Miller, S. A., Dykes, D. D., & Polesky, H. F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research*, 16(3), 1215.
- Monckton, D. G., Wong, L. J., Ashizawa, T., & Caskey, C. T. (1995). Somatic mosaicism, germline expansions, germline reversions and intergenerational reductions in myotonic dystrophy males: Small pool PCR analyses. *Human Molecular Genetics*, 4(1), 1–8. <https://doi.org/10.1093/hmg/4.1.1>
- Morales, F., Couto, J. M., Higham, C. F., Hogg, G., Cuenca, P., Braida, C., & Monckton, D. G. (2012). Somatic instability of the expanded CTG triplet repeat in myotonic dystrophy type 1 is a heritable quantitative trait and modifier of disease severity. *Human Molecular Genetics*, 21(16), 3558–3567. <https://doi.org/10.1093/hmg/dds185>
- Musova, Z., Mazanec, R., Krepelova, A., Ehler, E., Vales, J., Jaklova, R., & Sedlacek, Z. (2009). Highly unstable sequence interruptions of the CTG repeat in the myotonic dystrophy gene. *American Journal of Medical Genetics, Part A*, 149(7), 1365–1369. <https://doi.org/10.1002/ajmg.a.32987>
- Pešović, J., Perić, S., Brkušanić, M., Brajušković, G., Rakočević-Stojanović, V., & Savić-Pavičević, D. (2017). Molecular genetic and clinical characterization of myotonic dystrophy type 1 patients carrying variant repeats within DMPK expansions. *Neurogenetics*, 18(4), 207–218. <https://doi.org/10.1007/s10048-017-0523-7>
- Radvansky, J., Fieck, A., Minarik, G., Palffy, R., & Kadasi, L. (2011). Effect of unexpected sequence interruptions to conventional PCR and repeat primed PCR in myotonic dystrophy type 1 testing. *Diagnostic Molecular Pathology: The American Journal of Surgical Pathology, Part B*, 20(1), 48–51. <https://doi.org/10.1097/PDM.0b013e3181efe290>
- Santoro, M., Masciullo, M., Silvestri, G., Novelli, G., & Botta, A. (2017). Myotonic dystrophy type 1: Role of CCG, CTC and CGG interruptions within DMPK alleles in the pathogenesis and molecular diagnosis. *Clinical Genetics*, 92(4), 355–364. <https://doi.org/10.1111/cge.12954>
- Santoro, M., Fontana, L., Masciullo, M., Bianchi, M. L. E., Rossi, S., Leoncini, E., & Silvestri, G. (2015). Expansion size and presence of CCG/CTC/CGG sequence interruptions in the expanded CTG array are independently associated to hypermethylation at the DMPK locus in myotonic dystrophy type 1 (DM1). *Biochimica et Biophysica Acta*, 1852(12), 2645–2652. <https://doi.org/10.1016/j.bbdis.2015.09.007>
- Santoro, M., Masciullo, M., Pietrobbono, R., Conte, G., Modoni, A., Bianchi, M. L. E., & Silvestri, G. (2013). Molecular, clinical, and muscle studies in myotonic dystrophy type 1 (DM1) associated with novel variant CCG expansions. *Journal of Neurology*, 260(5), 1245–1257. <https://doi.org/10.1007/s00415-012-6779-9>
- Tomé, S., Dandelot, E., Dogan, C., Bertrand, A., Geneviève, D., Péréon, Y., & Gourdon, G. (2018). Unusual association of a unique CAG interruption in 5' of DM1 CTG repeats with intergenerational contractions and low somatic mosaicism. *Human Mutation*, 39(7), 970–982. <https://doi.org/10.1002/humu.23531>
- Tsilfidis, C., MacKenzie, A. E., Mettler, G., Barceló, J., & Korneluk, R. G. (1992). Correlation between CTG trinucleotide repeat length and frequency of severe congenital myotonic dystrophy. *Nature Genetics*, 1(3), 192–195. <https://doi.org/10.1038/ng0692-192>
- Zeesman, S., Carson, N., & Whelan, D. T. (2002). Paternal transmission of the congenital form of myotonic dystrophy type 1: A new case and review of the literature. *American Journal of Medical Genetics*, 107(3), 222–226.

#### SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Ballester-Lopez A, Koehorst E, Almendrote M, et al. A DM1 family with interruptions associated with atypical symptoms and late onset but not with a milder phenotype. *Human Mutation*. 2020;41:420–431. <https://doi.org/10.1002/humu.23932>

**Supporting Information**

Supp. Figure S1: 5' TP-PCR analysis for the patients of this study.





## **CHAPTER II**



## Characterization of RAN Translation and Antisense Transcription in Primary Cell Cultures of Patients with Myotonic Dystrophy Type 1

**Emma Koehorst**<sup>1</sup>, Judit Núñez-Manchón<sup>1</sup>, Alfonsina Ballester-López<sup>1, 2</sup>, Miriam Almendrote<sup>1, 3</sup>, Giuseppe Lucente<sup>1, 3</sup>, Andrea Arbex<sup>1, 3</sup>, Jakub Chojnacki<sup>4</sup>, Rafael P Vázquez-Manrique<sup>2, 5, 6</sup>, Ana Pilar Gómez-Escribano<sup>2, 5, 6</sup>, Guillem Pintos-Morell<sup>1, 7</sup>, Jaume Coll-Cantí<sup>3</sup>, Alba Ramos-Fransi<sup>1, 3</sup>, Alicia Martínez-Piñeiro<sup>1, 3</sup>, Mònica Suelves<sup>1</sup>, Gisela Nogales-Gadea<sup>1, 2</sup>

### AFFILIATIONS.

- <sup>1</sup> Neuromuscular and Neuropediatric Research Group, Institut d'Investigació en Ciències de la Salut Germans Trias i Pujol (IGTP), Campus Can Ruti, Universitat Autònoma de Barcelona, 08916 Badalona, Spain.
- <sup>2</sup> Centre for Biomedical Network Research on Rare Diseases (CIBERER), Instituto de Salud Carlos III, 28029 Madrid, Spain.
- <sup>3</sup> Neuromuscular Pathology Unit, Neurology Service, Neuroscience Department, Hospital Universitari Germans Trias i Pujol, 08916 Badalona, Spain.
- <sup>4</sup> IrsiCaixa AIDS Research Institute, 08916 Badalona, Spain.
- <sup>5</sup> Laboratory of Molecular, Cellular and Genomic Biomedicine, Instituto de Investigación Sanitaria La Fe, 46026 Valencia, Spain.
- <sup>6</sup> Joint Unit for Rare Diseases IIS La Fe-CIPF, 46012 Valencia, Spain.
- <sup>7</sup> Reference Unit for Hereditary Metabolic Disorders (MetabERN), Vall d'Hebron University Hospital, 08035 Barcelona, Spain.

**J Clin Med. 2021 Nov 25;10(23):5520. doi: 10.3390/jcm10235520**

Available from: <https://www.mdpi.com/2077-0383/10/23/5520>





---

---

## SUMMARY OF THE RESULTS

---

---

To study the influence of DM1-AS transcription and RAN translation on DM1 pathology, we used three primary cell cultures, namely myoblasts, skin fibroblasts and lymphoblastoids from ten DM1 patients in total. The studied DM1 cohort consisted of eight females and two males with an age of onset ranging from 15 to 50 years. All patients presented with clinical myotonia, but mild muscle impairment was only observed in two patients. The muscular impairment rating scale revealed minimal signs of muscular impairment in three of the patients, while five patients showed distal weakness and two patients had mild-moderate proximal weakness. Cardiac problems occurred in all patients, except P3 and P10. Five patients needed nocturnal mechanical ventilation, whereas three patients only showed mild changes in the respiratory function test. The average repeat size in blood was 387 CTGs (range 130–619 CTGs).

The RAN-translated polyGln has been described to originate from the antisense strand of the *DMPK* gene. We therefore decided to first validate the presence of DM1-AS transcripts in our three patient-derived primary cell culture by using three DM1-AS specific primer combinations for reverse transcription PCR (LK1/anti-1B, LK1/anti-N3, LK2/anti-N3), normalized by three different housekeeping genes (GAPDH,  $\beta$ 2-MG and PSMC4) and validated by Sanger sequencing. Overall, a lower expression was found in DM1 patients compared to controls with all three primer combinations, which reached significance for LK2/anti-N3 and LK1/anti-1B in myoblasts and LK1/anti-N3 and LK1/anti-1B in lymphoblastoids. Of note, the LK1/anti-N3 combination includes the CTG expansion, but seems to favor the wild-type allele, as for the DM1 patients only one patient showed an extra band, which could be the expanded allele, based on the 130 CTG repeat this patient carried.

For RAN translation to occur, DM1-AS transcripts need to be able to reach the cytoplasm. Subcellular fractionation of the DM1 cells revealed that, in both patients and controls, the DM1-AS transcripts of two different regions were present in the cytoplasm. In addition, we performed fluorescence in situ hybridization using a Cy3-labeled (CTG)<sub>10</sub> probe to detect antisense RNA foci and a Cy3-labeled (CAG)<sub>10</sub> to detect sense RNA foci in DM1 cells. We showed the presence of antisense and sense RNA foci in all DM1 cells, with less abundance of antisense foci compared to sense foci. Notably, for myoblasts and skin fibroblasts, these antisense RNA foci were also present in the cytoplasm (12.5% of myoblasts and 8.75% of skin fibroblasts). As expected, no sense or antisense RNA foci were found in controls. Importantly, antisense RNA foci were detected in cytoplasm, indicating that RAN translation is theoretically possible.

To study the presence of RAN translation we used three different antibodies for protein blots and immunofluorescence, two antibodies detect polyGln and are commercially available (1C2 and #1874), and one custom antibody,  $\alpha$ -DM1, directed against the C-terminus of a predicted glutamine frame of DM1 in the CAG direction. Positive and negative

controls were included to ensure the commercial antibodies were able to recognize longer stretches of polyGln. Protein blots with 1C2 and #1874 showed a 42 kD protein in both controls and patients in all primary cell cultures, which was also present in both our negative and positive controls. This protein was later identified as the TATA-box binding protein (TBP). Lymphoblastoid primary cell cultures showed polyGln containing proteins in the upper regions, which were not visible in the controls. However, these proteins were not recognized by the  $\alpha$ -DM1 antibody, which showed several bands in the 37 to 75 kD region in both patients and controls, and these higher located polyGln-containing proteins could therefore not be validated as a polyGln RAN protein.




Although the protein blots did not reveal the polyGln RAN-translated protein, we opted to use a second approach to validate our findings, using immunofluorescence. No differences were observed between patients and controls in all three primary cell cultures. A wide range of concentrations from 1:200 to 1:20,000 was used, but did not alter the original findings. Both anti-polyGln antibodies, 1C2 and #1874, showed infrequent staining of the nucleus and an intense aggregate around the nucleus in myoblasts and skin fibroblasts in all cells of both patients and controls.  $\alpha$ -DM1 displayed a more intense staining around the nucleus, roughly at the same place as the aggregates found with 1C2 and #1874, illustrated by double staining.

Due to the unexpected result of finding a positive staining in both DM1 patients and controls, we decided to further study the origin of this positive staining. A double immunostaining with TGN-38, a marker for the Golgi apparatus, showed an exact match to the structure we found with the 1C2 antibody in DM1 cells.

Taken together, DM1 patients had lower levels of DM1-AS transcripts compared to controls, which presented as a heterogeneous pool with and without the inclusion of the expanded repeat. RAN translation was not present in patient-derived DM1 cells, or in such low quantities that current techniques are unable to detect its presence.

## Article

# Characterization of RAN Translation and Antisense Transcription in Primary Cell Cultures of Patients with Myotonic Dystrophy Type 1

Emma Koehorst<sup>1</sup>, Judit Núñez-Manchón<sup>1</sup>, Alfonsina Ballester-López<sup>1,2</sup>, Miriam Almendrote<sup>1,3</sup>, Giuseppe Lucente<sup>1,3</sup> , Andrea Arbex<sup>1,3</sup>, Jakub Chojnacki<sup>4</sup>, Rafael P. Vázquez-Manrique<sup>2,5,6</sup> , Ana Pilar Gómez-Escribano<sup>2,5,6</sup> , Guillem Pintos-Morell<sup>1,7</sup>, Jaume Coll-Cantí<sup>1,3</sup>, Alba Ramos-Fransi<sup>1,3</sup>, Alicia Martínez-Piñero<sup>1,3,†</sup>, Mònica Suelves<sup>1,†</sup> and Gisela Nogales-Gadea<sup>1,2,\*,†</sup>



**Citation:** Koehorst, E.; Núñez-Manchón, J.; Ballester-López, A.; Almendrote, M.; Lucente, G.; Arbex, A.; Chojnacki, J.; Vázquez-Manrique, R.P.; Gómez-Escribano, A.P.; Pintos-Morell, G.; et al. Characterization of RAN Translation and Antisense Transcription in Primary Cell Cultures of Patients with Myotonic Dystrophy Type 1. *J. Clin. Med.* **2021**, *10*, 5520. <https://doi.org/10.3390/jcm10235520>

Academic Editor: Sabrina Ravaglia

Received: 30 October 2021

Accepted: 19 November 2021

Published: 25 November 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

- <sup>1</sup> Neuromuscular and Neuropediatric Research Group, Institut d'Investigació en Ciències de la Salut Germans Trias i Pujol (IGTP), Campus Can Ruti, Universitat Autònoma de Barcelona, 08916 Badalona, Spain; ekoehorst@igtp.cat (E.K.); jnunez@igtp.cat (J.N.-M.); aballester@igtp.cat (A.B.-L.); malmendrote.germanstrias@gencat.cat (M.A.); glucente@igtp.cat (G.L.); aarbex@igtp.cat (A.A.); gpintos@igtp.cat (G.P.-M.); jcoll@igtp.cat (J.C.-C.); aramosf@igtp.cat (A.R.-F.); amartinezp.germanstrias@gencat.cat (A.M.-P.); msuelves@igtp.cat (M.S.)
  - <sup>2</sup> Centre for Biomedical Network Research on Rare Diseases (CIBERER), Instituto de Salud Carlos III, 28029 Madrid, Spain; rafael\_vazquez@isslfe.es (R.P.V.-M.); ana\_pilar\_gomez@iislafe.es (A.P.G.-E.)
  - <sup>3</sup> Neuromuscular Pathology Unit, Neurology Service, Neuroscience Department, Hospital Universitari Germans Trias i Pujol, 08916 Badalona, Spain
  - <sup>4</sup> IrsiCaixa AIDS Research Institute, 08916 Badalona, Spain; jchojnacki@irsicaixa.es
  - <sup>5</sup> Laboratory of Molecular, Cellular and Genomic Biomedicine, Instituto de Investigación Sanitaria La Fe, 46026 Valencia, Spain
  - <sup>6</sup> Joint Unit for Rare Diseases IIS La Fe-CIPE, 46012 Valencia, Spain
  - <sup>7</sup> Reference Unit for Hereditary Metabolic Disorders (MetabERN), Vall d'Hebron University Hospital, 08035 Barcelona, Spain
- \* Correspondence: gnogales@igtp.cat; Tel.: +34-930330530  
 † Equal contribution.

**Abstract:** Myotonic Dystrophy type 1 (DM1) is a muscular dystrophy with a multi-systemic nature. It was one of the first diseases in which repeat associated non-ATG (RAN) translation was described in 2011, but has not been further explored since. In order to enhance our knowledge of RAN translation in DM1, we decided to study the presence of DM1 antisense (DM1-AS) transcripts (the origin of the polyglutamine (polyGln) RAN protein) using RT-PCR and FISH, and that of RAN translation via immunoblotting and immunofluorescence in distinct DM1 primary cell cultures, e.g., myoblasts, skin fibroblasts and lymphoblastoids, from ten patients. DM1-AS transcripts were found in all DM1 cells, with a lower expression in patients compared to controls. Antisense RNA foci were found in the nuclei and cytoplasm of a subset of DM1 cells. The polyGln RAN protein was undetectable in all three cell types with both approaches. Immunoblots revealed a 42 kD polyGln containing protein, which was most likely the TATA-box-binding protein. Immunofluorescence revealed a cytoplasmic aggregate, which co-localized with the Golgi apparatus. Taken together, DM1-AS transcript levels were lower in patients compared to controls and a small portion of the transcripts included the expanded repeat. However, RAN translation was not present in patient-derived DM1 cells, or was in undetectable quantities for the available methods.

**Keywords:** RAN translation; antisense transcription; myotonic dystrophies; primary cell cultures; phenotypic modulators

## 1. Introduction

Myotonic Dystrophy type 1 (DM1) is an autosomal dominant inherited muscular dystrophy with a multi-systemic nature. Patients display a wide variety of symptoms,

including muscle weakness, myotonia, respiratory failure, cardiac conduction defects, cataracts, and endocrine disturbances. In addition, the age of onset varies greatly, from birth up to >70 years. The cause of the variability in clinical manifestation is poorly understood. DM1 is viewed as an RNA gain of function disorder, wherein a CTG expansion in the 3' untranslated region of the myotonic dystrophy protein kinase (*DMPK*) gene causes the accumulation of expanded transcripts as intranuclear RNA foci, which sequester a number of splicing factors, resulting in loss of function and downstream deregulation of the alternative splicing of several genes [1]. In recent years, the view of DM1 as solely an RNA gain-of-function disorder has changed. Several new discoveries, such as antisense transcription and repeat associated non-ATG (RAN) translation, have added to the complexity of this disease. DM1 antisense (DM1-AS) transcription was first described by Cho and collaborators in 2005 [2]. They reported an antisense transcript, emanating from the adjacent *SIX5* regulatory region, which was converted into 21 nt siRNAs. These siRNAs were proposed to have a regulatory role in heterochromatin formation. Huguet and collaborators, however, showed that DM1-AS transcription extends across the CAG repeat, as they could detect DM1-AS RNA after the repeat in the 3' region of DM1 tissue and they also showed the presence of antisense RNA foci, which did not co-localize with sense RNA foci, in adult mouse models [3]. Similar results were found by Michel and collaborators in human fetal samples [4]. The inclusion of the expanded repeat was confirmed by Gudde and collaborators in 2017, although they found DM1-AS transcripts to be low in abundance and with varying lengths, both including and excluding the CAG repeat [5]. In addition, the DM1-AS strand was found by Zu and collaborators to give rise to RAN-translated polyglutamine stretches [6]. RAN translation is a typical phenomenon seen in repeat expansion disorders, in which peptides are produced from all frames without ATG start codon recognition. Zu and collaborators discovered a novel polyglutamine (polyGln) RAN protein expressed from the antisense CAG expansion transcript of the *DMPK* gene [6]. Nuclear polyGln RAN protein aggregates were found at a low frequency in a DM1 patient's myoblasts and skeletal muscle ( $n = 1$ ) and at a higher frequency in leukocytes from peripheral blood ( $n = 1$ ) [6]. The nuclear aggregates co-localized with caspase-8, an early indicator of polyGln-induced apoptosis. This suggests that RAN proteins may be an additional mechanism of cytotoxicity in DM1 cells.

Since its first discovery in 2011 by Zu and collaborators, RAN translation has been extensively studied in multiple expansion disorders and great advances have been made [7]. However, the contribution of RAN translation to DM1 pathology has not been further studied since its first report in 2011. Much remains unknown regarding the presence of RAN translation and its mechanism in DM1. Is it equally present across patients, and is its distribution across tissues similar? To what extent does it contribute to the pathology of the disease? In order to further enhance our knowledge of RAN translation in DM1, we decided to study the presence of RAN translation in DM1 primary cell cultures—myoblasts, skin fibroblasts and lymphoblastoids—derived from ten DM1 patients, with a heterogeneous display of subtypes. The RAN-translated polyGln has been described to originate from the antisense strand of the *DMPK* gene. We therefore validated the presence of DM1-AS transcription in our three patient-derived cellular models and lower expression levels were found in patients compared to controls. Additionally, we found that the DM1-AS transcripts were found in both the nucleus and the cytoplasm, of which at least a portion contained the expanded repeat, as shown by the presence of antisense RNA foci in patients. However, the polyGln RAN protein was not present in patient-derived DM1 cells, or was present in such low quantities that it was below the detection limit of the currently available techniques.

## 2. Materials and Methods

### 2.1. Samples

This study was approved by the Ethics Committee of the University Hospital Germans Trias i Pujol and was performed in accordance with the Declaration of Helsinki for

Human Research. Written informed consent was obtained from all participants. The study included ten patients with DM1 and thirteen controls with no previous family history of neuromuscular disorders (recruited from the traumatology department, in whom surgery was needed). DM1 diagnosis was confirmed or discarded via triplet-primed PCR in all the study participants. Clinical information of DM1 patients was obtained from medical records and updated at the last visit. We obtained three different samples from eight patients and eleven controls: blood, muscle biopsy, and skin biopsy. The other two patients and two controls only provided a blood sample. All samples were simultaneously obtained from patients with confirmed juvenile, adult or late-onset DM1. The muscle biopsy was obtained from biceps brachii ( $n = 7$ ) or vastus lateralis ( $n = 1$ ) of DM1 patients and from intrinsic forearm or hand muscles of eleven non-DM1 patients. Skin biopsy was obtained with a 0.5 cm skin punch.

### 2.2. Small Pool PCR for Sizing the CTG Repeat

Total genomic DNA was extracted from peripheral blood samples, as previously described [8]. To estimate the length of the expanded mode allele, small-pool PCR (SP-PCR) was carried out with small amounts of input DNA (300 pg), using flanking primers DM-C and DM-DR, as previously described [9,10]. PCR was performed using a Custom PCR Master Mix (Thermo Fisher Scientific, Waltham, MA, USA) supplemented with 69 mM 2-mercaptoethanol, and Taq polymerase (Sigma-Aldrich, Gillingham, UK) at 1 unit per 10  $\mu$ L. All reactions were supplemented with 5% DMSO and the annealing temperature was 63.5 °C. DNA fragments were resolved by electrophoresis on a 1% agarose gel, and Southern blot hybridized as described in references [9,10]. Autoradiographic images were scanned and the CTG size of the mode was estimated through comparison against the molecular weight ladder using GelAnalyzer 19.1 software ([www.gelanalyzer.com](http://www.gelanalyzer.com), by Istvan Lazar Jr. and Istvan Lazar Sr.).

### 2.3. Cell Culture

Myoblasts were isolated from the biopsied tissue by CD56 magnetic separation. Myoblasts were grown until 60–70% confluency on 0.1% gelatin-coated coverslips in six-well tissue-culture plates in a proliferation medium containing DMEM supplemented with 10% FBS, 22% M-199, 1  $\times$  PSF, insulin 1.74  $\mu$ M, L-glutamine 2 mM, FGF 1.39 nM and EGF 0.135 mM. Skin fibroblasts were isolated from biopsied tissue using the explant method. Skin fibroblasts were grown until 70% confluency on 0.5% poly-D-lysine coated coverslips in a six-well tissue culture plate in a proliferation medium containing DMEM, supplemented with 10% FBS and 1  $\times$  PSF. Peripheral blood mononuclear cells were isolated from blood collected in heparin tubes according to the Ficoll gradient and immortalized using Epstein-Barr virus. Lymphoblastoids were grown until 70% confluency on 0.1% poly-D-lysine coated coverslips in a six-well tissue culture plate in a proliferation medium containing RPMI, supplemented with 10% FBS, 1  $\times$  PSF and 2 mM L-glutamine. HEK293 cells were cultured until 80% confluency on 0.5% poly-D-lysine coated coverslips in a six-well tissue culture plate in MEM, supplemented with 5% horse serum, 5% FBS, 1% Glutamine and 1% PSF. In addition to the coverslips, cell pellets from each of the cell types were extracted and stored for later use in RNA and protein studies.

### 2.4. cDNA Construct Huntingtin and HEK293 Transfection

The Q17 vector was generously provided to us by R.P.V.-M.. In short, the construct is part of the cDNA of human huntingtin (585 amino acids), containing 17 CAGs, and was cloned into a pcDNA3.1-Gateway. The human huntingtin with the 17 CAGs was in frame with mCherry, a fluorescent. DNA transfections were performed when HEK293 cells reached 80% confluency using Lipofectamine 2000 reagent (Thermo Fisher Scientific, Waltham, MA, USA), according to the manufacturer's instructions. DNA:lipofectamine ratio was 1:2 and incubation lasted 48 h.

### 2.5. RNA Isolation and Subcellular Fractionation

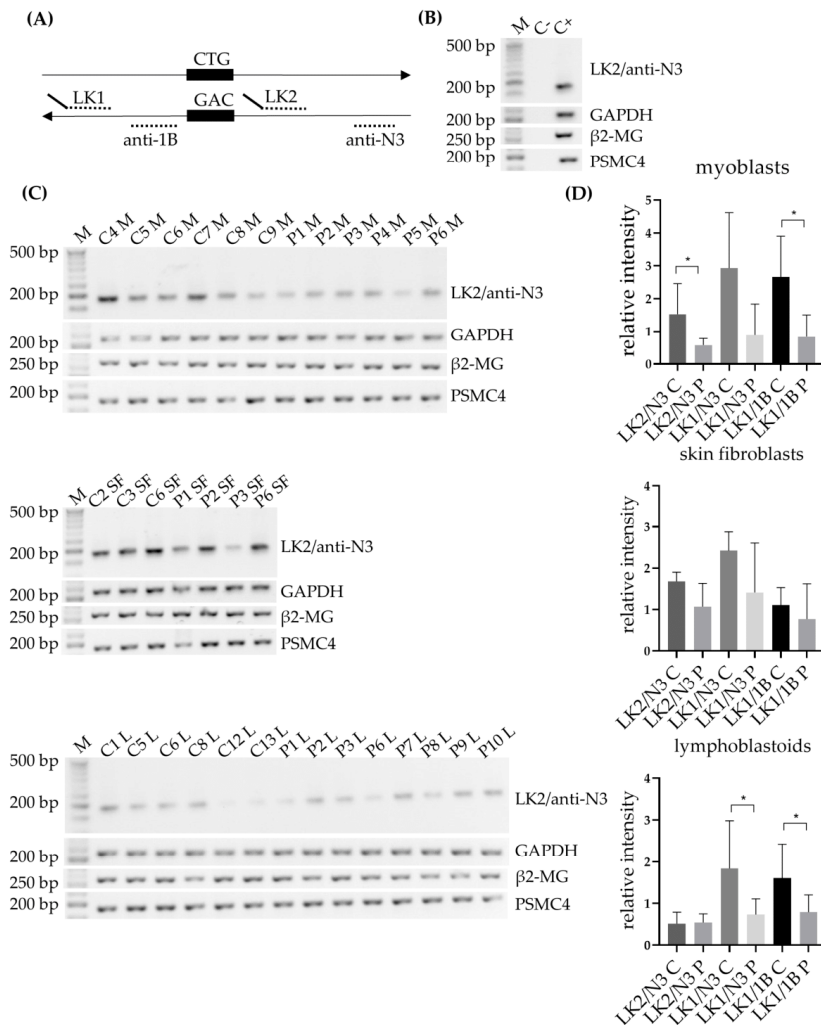
Total RNA from cultured myoblast, skin fibroblast and lymphoblastoid cell lines was isolated using TRIzol reagent (Thermo Fisher Scientific, Waltham, MA, USA) or the RNeasy plus mini kit (Qiagen, Hilden, Germany), according to the manufacturer's instructions. Total RNA from a healthy human heart was used as a positive control in all RT-PCR analyses (AM7966, Thermo Fisher Scientific, Waltham, MA, USA).

For RNA isolation from the subcellular fractions, skin fibroblasts were grown to 80% confluency, collected through trypsinization and pelleted through centrifugation at 2000 rpm for 5 min at 4 °C. Cell pellets were washed twice with ice-cold PBS. Cells were resuspended in ice-cold cell disruption buffer (10 mM KCl, 1.5 mM MgCl<sub>2</sub>, 20 mM Tris-Cl (pH 7.5), 1 mM DTT) and incubated on ice for 10 min [11]. Samples were homogenized using a tissue disruptor for 30 s and then Triton X-100 was added to a final concentration of 0.1%. The lysate was centrifuged at 2000 rpm for 5 min at 4 °C, after which the supernatant (cytoplasmic fraction) was removed. Both cytoplasmic and nuclear fraction underwent RNA isolation by the RNeasy plus mini kit (Qiagen, Hilden, Germany), according to the manufacturer's instructions.

### 2.6. RT-PCR Analysis for DM1-AS Transcript Detection

For detection of the DM1-AS transcripts, a similar strategy to Zu et al. was followed [6], with some minor changes. In brief, an equivalent of 1 µg total RNA was subjected to cDNA synthesis using SuperScript IV reverse transcriptase (Thermo Fisher Scientific, Waltham, MA, USA) at 55 °C, using random hexamers (50 µM, N8080127, Thermo Fisher Scientific, Waltham, MA, USA). The subsequent PCR was carried out using DM1-AS specific primers LK1, together with anti-N3 or anti-1B, and LK2, together with anti-N3 (Figure 1A; primer details and PCR conditions can be found in Table S1). Both LK1 and LK2 contained a linker sequence, for which a specific primer was created and used in some of the PCR reactions. For endogenous controls GAPDH, β2-microglobulin (β2-MG) and PSMC4 a similar approach was followed (primer details and PCR condition found in Table S1).

To analyze nuclear and cytoplasmic RNA fractions, 150 ng RNA was used for the RT reaction for both fractions. RNA was reverse transcribed using SuperScript IV reverse transcriptase (Thermo Fisher Scientific, Waltham, MA, USA) at 55 °C, using random hexamers (50 µM, N8080127, Thermo Fisher Scientific, Waltham, MA, USA). For RT-PCR analyses, the same approach was used as described above with primer combinations LK1/anti-N3 and LK2/anti-N3. As a nuclear marker, pre-mRNA DMPK was used and GAPDH as an endogenous control (primers and PCR conditions in Table S1).



**Figure 1.** DM1 antisense (DM1-AS) transcripts in DM1 cells. **(A)** Schematic diagram indicating the location of the primers used for the DM1-AS amplification. Both DM1-AS specific primers (LK1 and LK2) have a linker sequence attached, indicated by the tail. Either a linker primer or the strand-specific primers were used for the RT-PCR reactions. A total of three primer combinations were used, LK1 with primers anti-1B/anti-N3 and LK2 with primer anti-N3. **(B)** Positive and negative controls from the RT-PCR reactions; C- = no DNA in the RT-reaction; C+ = control heart RNA as input, the tissue used in the original paper. **(C)** Strand-specific RT-PCRs of the DM1-AS with the DM1-AS specific primers LK2 and anti-N3 in all three primary cell cultures. GAPDH, β2-MG and PSMC4 were used as an endogenous controls. **(D)** Expression of DM1-AS determined by measuring intensity with ImageJ, normalized against the mean of the endogenous controls for all three primer combinations in all three primary cell cultures. Gels of the other two primer combinations can be found in Figure S1. Error bars indicate standard deviations. \* = *p*-value below 0.05. Abbreviations: M = marker; C = control; P = DM1 patient; M = myoblasts; SF = skin fibroblasts; L = lymphoblastoids.



### 2.7. Immunoblotting

The cell pellets collected from the cell cultures were lysed in a RIPA buffer, supplemented with a cOmplete™ Protease Inhibitor Cocktail (Roche, Basel, Switzerland) and homogenized with a tissue disruptor. After centrifugation, the supernatant was collected and the protein concentration determined with a DC™ Protein Assay Kit II (Bio-Rad Laboratories, Hercules, CA, USA). 20–70 µg of protein was separated on a 3–8% gradient Tris-Acetate gel or an 8% acrylamide gel. Gel proteins were transferred by the iBlot2 system (Thermo Fisher Scientific, Waltham, MA, USA) for Tris-Acetate gels to a nitrocellulose membrane and by wet transfer to a PVDF membrane for the acrylamide gels (Merck, Darmstadt, Germany). Membranes were blocked for one hour in Intercept (TBS) blocking buffer (LI-COR, Lincoln, NE, USA). Immunoblotting was performed with α-DM1 antibody (1:1000, kindly provided by Laura Ranum), 1C2 (clone 5TF1-1C2, 1:250-1:1000, Merck, Darmstadt, Germany), #1874 (clone 3B5H10, Merck, Darmstadt, Germany, 1:1000), TATA-box-binding protein (TBP) (1:250-1:1000, Abcam, Cambridge, UK) or α-tubulin (1:5000, Merck, Darmstadt, Germany) overnight at 4 °C. Appropriate secondary antibodies, anti-rabbit 1:8000 conjugated with IRDye 680RD and anti-mouse 1:8000 conjugated with IRDye 800CW or IRDye 680RD (Thermo Fisher Scientific, Waltham, MA, USA), were used. Band pattern was revealed with an Odyssey Imager (LI-QOR, Lincoln, NE, USA). For favoring the detection of bigger proteins, one of the membranes was stripped with a mild stripping buffer (20 mM glycine, 3 mM SDS, 0.1% Tween 20 in deionized water, with a pH of 2.2), after which the membrane was cut below the 50kD marker and re-blocked and probed with #1874.

### 2.8. Fluorescence In Situ Hybridization (FISH) and Immunofluorescence

Myoblasts, skin fibroblasts and lymphoblastoids grown on coverslips were fixed in 4% paraformaldehyde for 30 min and permeabilized with 0.3% Triton X-100 at 4 °C for 5 min (0.1% Triton X-100 for lymphoblastoids).

For the FISH, coverslips were incubated with 30% formamide in 2× SSC buffer for 30 min at room temperature, followed by an overnight incubation at 37 °C in darkness with a hybridization buffer, containing 0.01 µM Cy3-labelled (CAG)<sub>10</sub> or (CTG)<sub>10</sub> probe, 30% formamide, 1% dextran sulfate, 0.02% BSA and 2 mM vanadyl in 2× SSC buffer. The following day, the coverslips were washed with 30% formamide in 2× SSC buffer at 45 °C, 1× SSC buffer at 37 °C and 1× PBS at room temperature, and mounted with ProLong Gold Anti-Fade Mountant with DAPI (Thermo Fisher Scientific, Waltham, MA, USA).

For immunofluorescence, three approaches were followed: a simple immunofluorescence for the detection of RAN-translated polyglutamine protein in all three primary cell cultures, a single-step double immunofluorescence to compare one of the commercial antibodies with the custom antibody used and a two-step double immunofluorescence to assess the co-localization of the found protein with the Golgi apparatus.

For the simple immunofluorescence, coverslips with myoblasts, skin fibroblasts or lymphoblastoids were blocked for 1 h in 5% filtered NGS and incubated overnight at 4 °C with either α-DM1 (1:200, kindly provided by Laura Ranum), 1C2 (Merck, Darmstadt, Germany, ref MAB1574, 1:1000) or #1874 (Merck, Darmstadt, Germany, ref P1874, 1:1000). Next, cells were washed and incubated with goat anti-rabbit conjugated with alexa fluor 488 (Thermo Fisher Scientific, Waltham, MA, USA) for α-DM1 and goat anti-mouse conjugated with alexa fluor 488 (Thermo Fisher Scientific, Waltham, MA, USA) for 1C2 and #1874 for 1 h at room temperature, in darkness. After another round of washes in 1× PBS supplemented with 0.025% Tween-20, coverslips were mounted with ProLong Gold or Diamond Anti-Fade Mountant with DAPI (Thermo Fisher Scientific, Waltham, MA, USA).

For the single-step double immunofluorescence, coverslips were blocked for 1 h in 5% filtered NGS and incubated overnight at 4 °C with two antibodies in the same mix, α-DM1 (1:500 for myoblasts and skin fibroblasts and 1:5000 for lymphoblastoids, kindly provided by Laura Ranum) and 1C2 (1:500, Merck, Darmstadt, Germany). Next, cells were washed and incubated with goat anti-rabbit conjugated with alexa fluor 488 (Thermo Fisher Scientific, Waltham, MA, USA) and goat anti-mouse conjugated with alexa fluor 594

(Thermo Fisher Scientific, Waltham, MA, USA). After another round of washes in  $1 \times$  PBS, coverslips were mounted with ProLong Gold Anti-Fade Mountant with DAPI (Thermo Fisher Scientific, Waltham, MA, USA).

For the two-step immunofluorescence, coverslips, containing either skin fibroblasts or myoblasts, were blocked for 30 min with the horse blocking solution, containing 5% normal horse serum, 10% normal human serum and 0.02% bovine serum albumin in  $1 \times$  TBS. Subsequently, the coverslips were incubated overnight at  $4^\circ\text{C}$  with 1C2 (1:200, Merck, Darmstadt, Germany). The following day, the cells were washed three times for 5 min with  $1 \times$  TBS and incubated at room temperature for 1 h with biotin-labeled horse anti-mouse (1:2000, Thermo Fisher Scientific, Waltham, MA, USA). After another round of washes, the cells were incubated at room temperature for 1 h with streptavidin conjugated with alexa fluor 594 (1:2000, Thermo Fisher Scientific, Waltham, MA, USA). Next, the cells were washed and blocked with the goat blocking solution, containing 5% normal goat serum, 10% normal human serum, 10% Goat anti-horse IgG (Thermo Fisher Scientific, Waltham, MA, USA) and 0.02% bovine serum albumin in  $1 \times$  TBS. Subsequently, the coverslips were incubated overnight at  $4^\circ\text{C}$  with TGN-38 (1:200, Merck, Darmstadt, Germany). The following day, the cells were washed three times for 5 min with  $1 \times$  TBS and incubated at room temperature for 1 h with goat anti-mouse conjugated with alexa fluor 488 (1:500, Thermo Fisher Scientific, Waltham, MA, USA). After another round of washes, the cells were mounted with ProLong Gold Anti-Fade Mountant with DAPI (Thermo Fisher Scientific, Waltham, MA, USA).

### 2.9. Image and Statistical Analysis

To assess DM1-AS expression and protein quantity, intensity measurements were taken using ImageJ software, and normalized against the endogenous controls. Differences in transcript expression were calculated using the Mann–Whitney’s non-parametric U test with Graphpad Prism 9.1.2 software; significance level was set at 0.05.

## 3. Results

### 3.1. The DM1 Clinical Phenotype of the Studied Cohort

Our study population consisted of ten DM1 patients, of which the clinical characteristics have been, in part, described previously (Table 1, [12,13]). Eight out of ten DM1 patients provided skin and muscle biopsies, whereas from the other two only lymphoblastoid cell lines were available. The studied DM1 cohort consisted of eight females and two males with an age of onset ranging from 15 to 50 years. Seven individuals were unrelated and three were sisters (P3, P4 and P8). All patients presented with clinical myotonia, but mild muscle impairment was only observed in two patients, reflected by a Biceps MRC of four. Performance in the six-min walking distance test averaged 377 m (range 251–519 m). The muscular impairment rating scale (MIRS) revealed minimal signs of muscular impairment in three of the patients, while five patients showed distal weakness and two patients had mild-moderate proximal weakness. Cardiac problems occurred in all patients, except P3 and P10. Six DM1 patients showed minor ECG alterations, one (P2) a structural cardiopathy (valvulopathy) and P5 had a pacemaker. Five patients needed nocturnal mechanical ventilation, whereas three patients only showed mild changes in the respiratory function test and two patients (P7 and P10) showed no altered respiration. The majority of patients were independent in daily life activities (score of 0–2 on the modified Rankin (mRS) scale), and two patients (P5 and P9) had a moderate to moderately severe disability (scores of 3 and 4, respectively). The average repeat size in blood was 387 CTGs (range 130–619 CTGs) and one patient presented with CCG variant repeats (P7, previously reported [14]).

Table 1. Clinical characteristics of Myotonic Dystrophy type 1 (DM1) patients.

Patient	Sex	Age of Onset (y)	Age at Sampling (y)	Biceps MRC	Myotonia (s)	6-MWD (m)	MIRS	mRS	Cataracts	Cardiopathy	Spirometry	Repeat Size (CTGs)
P1	F	15*	36	4	0.52	348	4	2	no	LAFB	Altered PFT	445
P2	M	48	54	5	0.67	251	3	2	yes	Valvulopathy	NMV	381
P3	F	36	41	5	0.73	368	2	1	yes	none	NMV	338
P4	F	42	46	5	0.98	338	3	1	yes	1st-degree AV block	NMV	246
P5	F	27	40	4	NP	NP	4	4	yes	Pacemaker	Altered PFT	374
P6	M	36	41	5	0.96	519	3	2	no	LAFB	Altered PFT	130
P7	F	50	62	5	NP	436	2	1	yes	1st-degree AV block	none	561
P8	F	35	38	5	NP	NP	3	2	no	1st-degree AV block	NMV	619
P9	F	30	51	5	NP	NP	4	3	yes	1st-degree AV block	NMV	NP
P10	F	18	34	5	NP	NP	3	2	yes	none	none	NP

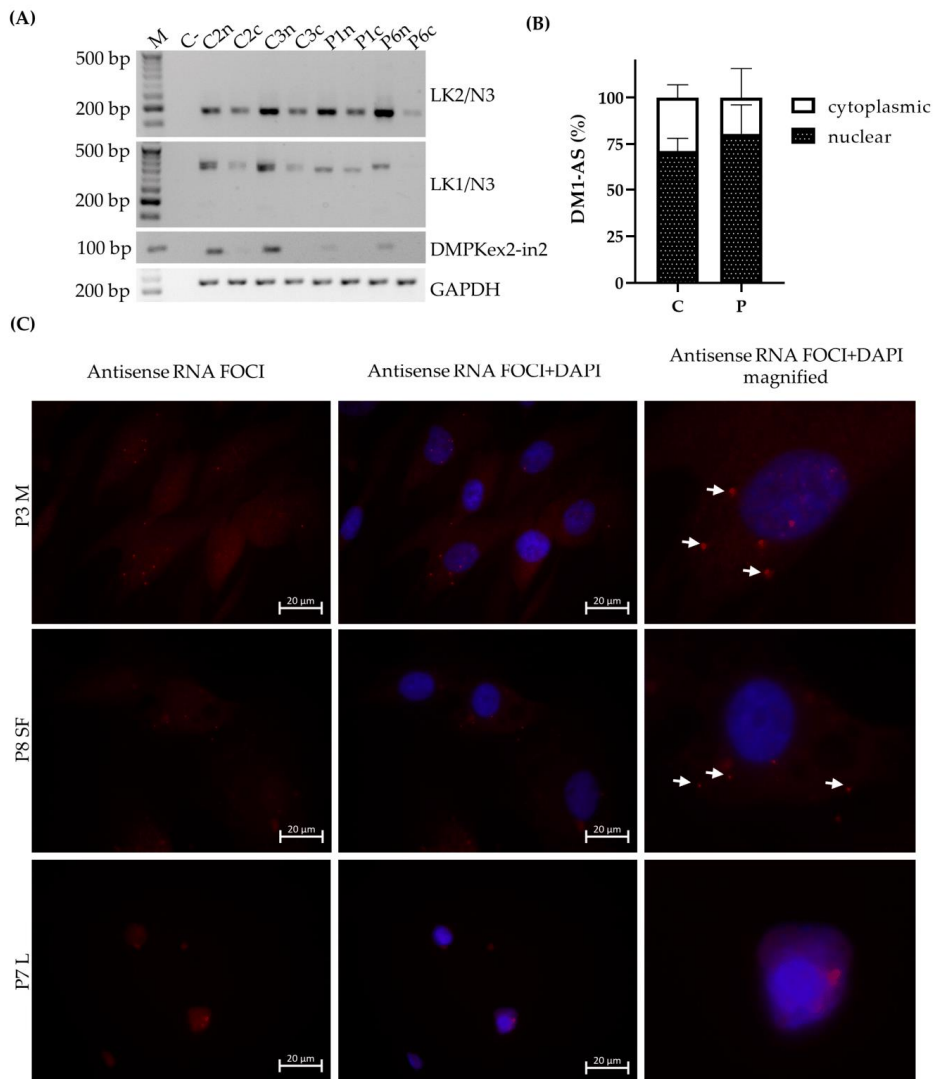
F = female; M = male; MRC = Medical Research Council; NP = not performed; 6MWD = 6-min walking test; AV = atrioventricular; LAFB = left anterior fascicular block; mRS = modified Rankin Scale; MIRS = Muscular Impairment Rating Scale; NMV = nocturnal mechanical ventilation; PFT = pulmonary function test; \* Although the exact age of disease onset was impossible to determine, based on the symptoms displayed at first visit (age 36), which commonly appear with an early onset (including oval pallor and temporal atrophy), disease onset was considered to have been during adolescence.

### 3.2. Lower Expression of DM1-AS Transcripts Found in DM1 Patients Compared to Controls

The RAN-translated polyGln has been described to originate from the antisense strand of the *DMPK* gene. We therefore decided to first validate the presence of DM1-AS transcripts in our three patient-derived primary cell cultures (myoblasts, skin fibroblasts and lymphoblastoids) by using three DM1-AS specific primer combinations, previously described ([6], Figure 1A). The original set-up published by Cho and collaborators and used in the original paper on DM1 RAN translation consisted of using DM1-AS specific primers with a linker sequence attached for strand-specific priming, and a primer complementary to that linker sequence for the subsequent PCRs [2]. However, in our cohort, this setup resulted in an inability to find the DM1-AS transcripts, including our positive control consisting of RNA of a heart control (data not shown), the tissue of origin used in the original paper [6]. When using the DM1-AS specific primers for the subsequent PCRs instead of the linker primer, we detected extreme variability in both DM1-AS transcripts and endogenous controls. We therefore opted for the use of random hexamers for cDNA synthesis and the use of the DM1-AS specific primers for the subsequent PCR. This allowed us to show DM1-AS transcription in all patients and controls with all three primer combinations and with a homogenous housekeeping gene distribution (Figures 1B,C and S1). PCR products were validated by Sanger sequencing. Overall, a lower expression was found in DM1 patients compared to controls with all three primer combinations, which reached significance for LK2/anti-N3 and LK1/anti-1B in myoblasts and LK1/anti-N3 and LK1/anti-1B in lymphoblastoids (Figure 1D). Expression was normalized against each housekeeping gene individually and against the mean of the three endogenous controls. Similar results were obtained with each approach, and the latter was used for the normalization shown in Figure 1. Of note, the LK1/anti-N3 combination, which included the CTG expansion, showed two distinct bands in some of the controls, indicating heterogeneity in their wild-type alleles (Figure S1). This primer combination seemed to favor the wild-type allele, as for the DM1 patients only one patient showed an extra band, which could be the expanded allele, based on the 130 CTG repeat this patient carried (Figure S1A,B).

### 3.3. DM1-AS Transcripts Are Present in the Cytoplasm of DM1 Cells

The results shown above indicate the presence of the DM1-AS transcript in both patients and controls, but for RAN translation to occur, these transcripts need to be able to reach the cytoplasm. To study this, we decided on a dual approach. The first approach was subcellular fractionation of the DM1 cells prior to RNA isolation. We verified the absence of nuclear pre-mRNA *DMPK* in the cytoplasmic fraction to ensure that no nuclear contamination could alter our results. RT-PCRs revealed that, in both patients and controls, the DM1-AS transcripts of two different regions were present in the cytoplasm (Figure 2A). The signal was higher in the nucleus compared to the cytoplasm, and this was slightly more apparent in patients (Figure 2B). In addition, we performed FISH using a Cy3-labeled (CTG)<sub>10</sub> probe to detect antisense RNA foci and a Cy3-labeled (CAG)<sub>10</sub> to detect sense RNA foci in DM1 cells. We showed the presence of antisense and sense RNA foci in all DM1 cells (Figures 2C and S2A). Notably, for myoblasts and skin fibroblasts, these antisense RNA foci were also present in the cytoplasm (12.5% of myoblasts and 8.75% of skin fibroblasts). However, this percentage was highly dependent on the patient, as some patients showed only 5% of cells with cytoplasmic antisense RNA foci, whereas others showed 30% (Tables S3 and S5). This variability could not be correlated to, for example, the CTG repeat size. As expected, no sense or antisense RNA foci were found in controls (Figure S2B). Overall, the presence of antisense RNA foci was shown in DM1 cells, although they were less abundant compared to sense RNA foci (Tables S3–S7). Importantly, antisense RNA foci were detected in cytoplasm, indicating that RAN translation is theoretically possible.

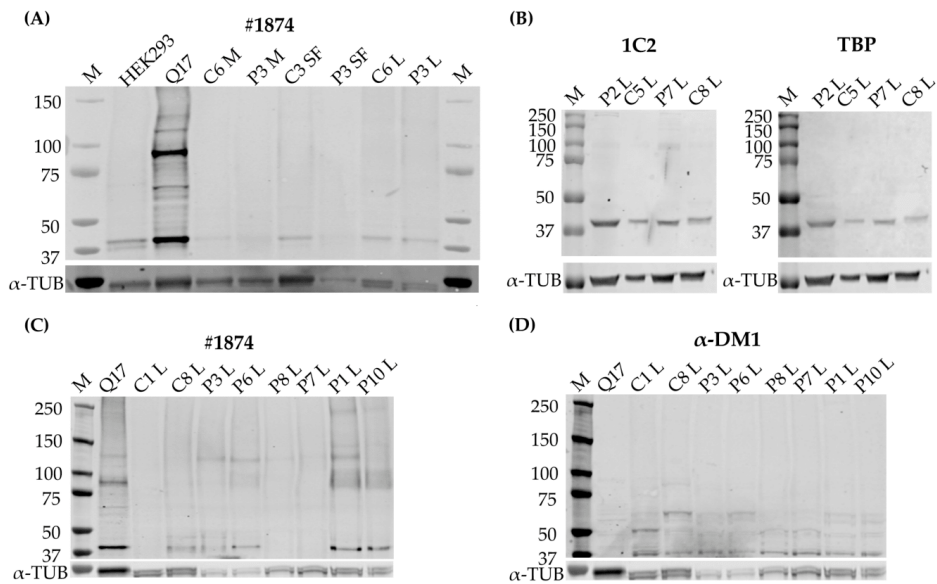


**Figure 2.** DM1-AS presence in the cytoplasm. (A) DM1-AS specific RT-PCR for subcellular fractionated skin fibroblasts with two primer combinations, LK2/anti-N3 and LK1/anti-N3. (B) Quantification of DM1-AS signals; bars represent mean +SD. (C) Presence of antisense RNA foci in all three cell types of DM1 patients. Cy3-labeled (CTG)<sub>10</sub> probe (RED) showing antisense RNA foci; nuclei indicated by DAPI (blue) and white arrows indicate cytoplasmic antisense RNA foci. Abbreviations: C = control; P = DM1 patient; n = nuclear; c = cytoplasmic; DMPKex2-in 2 = pre-mRNA DMPK as a nuclear marker; GAPDH = endogenous control; M = myoblasts; SF = skin fibroblasts; L = lymphoblastoids.

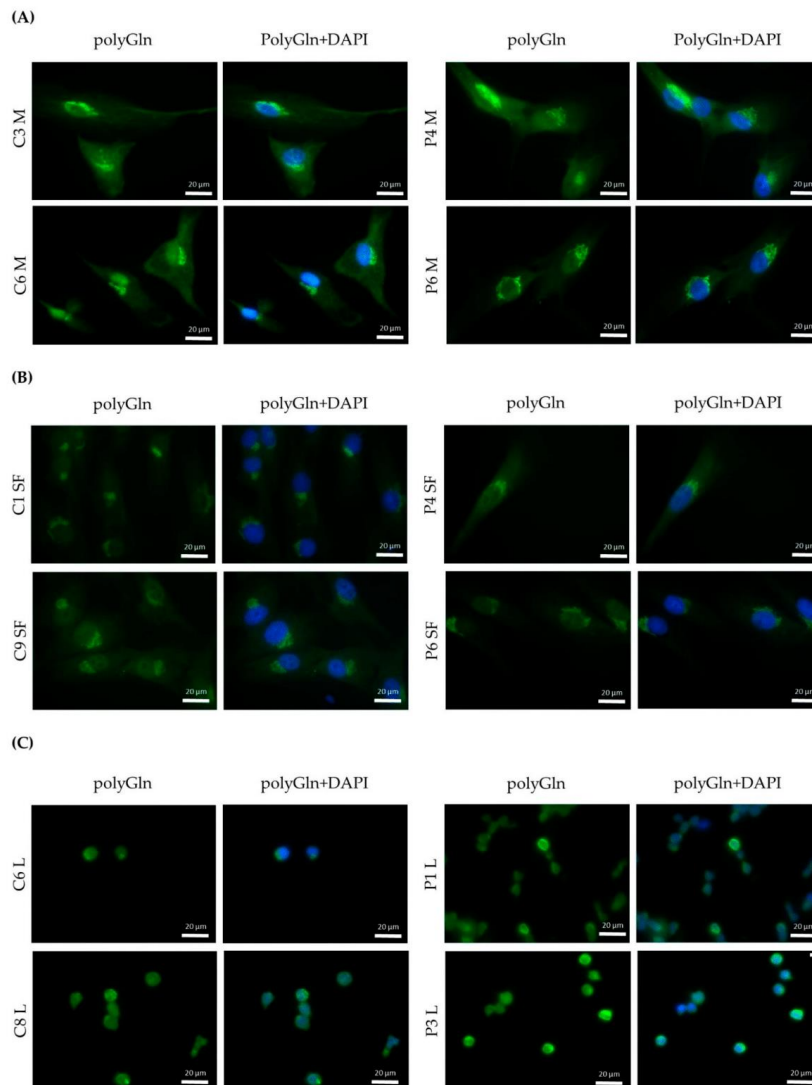
### 3.4. RAN Protein Was Undetectable in All Three Primary Cell Cultures

After validation that the origin of the RAN-translated polyGln was detectable in our primary cell cultures, we moved on to detection of the protein itself. For this, we used three different antibodies, of which two detect polyGln and are commercially available (1C2 and #1874), and one custom antibody,  $\alpha$ -DM1, directed against the C-terminus of a predicted glutamine frame of DM1 in the CAG direction (generously provided by Laura Ranum). To make sure our commercial antibodies were able to recognize longer stretches of polyGln, we added two positive controls to our experiments: firstly a huntingtin vector containing 17 glutamines (Q17), in frame with mCherry, transfected into HEK293 cells, with an expected size of approximately 85 kD and secondly a Huntington's patient lymphoblastoid cell line (Huntingtin expected size: 340–350 kD). The commercially available anti-polyGln antibody #1874 was able to detect the polyGln-containing proteins in the two positive controls, and 1C2 was able to detect the Q17 vector as well (Figure S3A,B), indicating that our approach was able to show polyGln-containing proteins as large as 350 kD. In addition, HEK293 untransfected cells were added as a negative control and only showed a 42 kD protein (Figures 3A and S3B). Protein blots with 1C2 and #1874 in most cases showed a 42 kD protein in both controls and patients in all primary cell cultures, which was also present in both our negative and positive controls (Figures 3A and S3B). This protein was later identified as the TATA-box binding protein (TBP) (Figure 3B). Cutting the membrane above the 42 kD band to favor binding to other proteins did not alter our results (Figure S3C).  $\alpha$ -DM1 revealed several different sizes of proteins in all three cell types, ranging from 26 kD to 150 kD, but these were visible in both patients and controls and with no significant differences in intensity (Figure S3D). In addition, the striking observation here was the different size of the polyGln-containing protein that both antibodies bound. Although small, the double staining method clearly showed a different band and the antibodies therefore did not bind the same protein (Figure S3E). The only cellular model to show a band that might correspond to the RAN protein were the lymphoblastoid cell lines, as four out of six patients showed polyGln containing proteins in the upper regions, which were not visible in the controls (Figure 3C). However, these proteins were not recognized by the  $\alpha$ -DM1 antibody, which showed several bands in the 37 to 75 kD region in both patients and controls (Figure 3D), and these higher located polyGln-containing proteins could therefore not be validated as a polyGln RAN protein.

Although the protein blots did not reveal the polyGln RAN-translated protein, we opted to use a second approach to validate our findings, using immunocytochemistry. By use of the Q17-transfected HEK293 cells we validated the use of our commercial anti-polyGln antibodies in immunocytochemistry (Figure S3F). No differences were observed between patients and controls in all three primary cell cultures (Figure 4). A wide range of concentrations from 1:200 to 1:20,000 was used, but did not alter the original findings (data not shown). Both anti-polyGln antibodies, 1C2 and #1874, showed infrequent staining of the nucleus and an intense aggregate around the nucleus in myoblasts and skin fibroblasts in all cells of both patients and controls (Figures 4A,B, S4 and S5). In lymphoblastoids, due to the limited amount of cytoplasm, it did not show as an aggregate, but rather as an overall intense staining (Figures 4C, S4 and S5). This staining was present in approximately half of the cells, but still visible in both DM1 patients and controls. With the  $\alpha$ -DM1, we saw an overall staining of the cytoplasm, with a more intense staining around the nucleus in both DM1 patients and controls and again infrequent staining of the nucleus (Figure S6). In myoblasts, in addition to the intense staining around the nucleus, small bright dots were observed in approximately half of the cells. Again, these were seen in both DM1 patients and controls (Figure S6). This more intense staining was roughly at the same place as the aggregates found with 1C2 and #1874, as illustrated by the double staining (Figure S7).



**Figure 3.** Quantitative analysis of repeat associated non-ATG (RAN) translated polyglutamine (polyGln) protein. (A) A representative immunoblot with #1874 1:1000, including all three primary cell cultures, our positive control, a huntingtin vector containing 17 polyglutamine stretches (Q17), and a negative control, consisting of untransfected HEK293 cells (C:  $n = 4$ , P:  $n = 4$ ). (B) Immunoblot showing co-localization of the 42 kD protein found with #1874 and 1C2 with TATA-box-binding protein (TBP) in lymphoblastoids. The antibody shown here is 1C2. (C) Immunoblot of lymphoblastoid cell lines, two controls and six patients with #1874 showing several higher bands in certain DM1 patients. (D) Immunoblot with the custom antibody  $\alpha$ -DM1, showing several non-specific bands, but none that overlap with the bands found with #1874. Abbreviations: M = marker; C = control; P = DM1 patient; HEK293 = untransfected HEK293 cells; Q17 = Q17 huntingtin vector; M = myoblasts; SF = skin fibroblasts; L = lymphoblastoids. #1874 = commercial antibody recognizing polyGln;  $\alpha$ -DM1 = custom antibody against the predicted C-terminus of the polyGln RAN protein;  $\alpha$ -TUB =  $\alpha$ -tubulin as endogenous control.

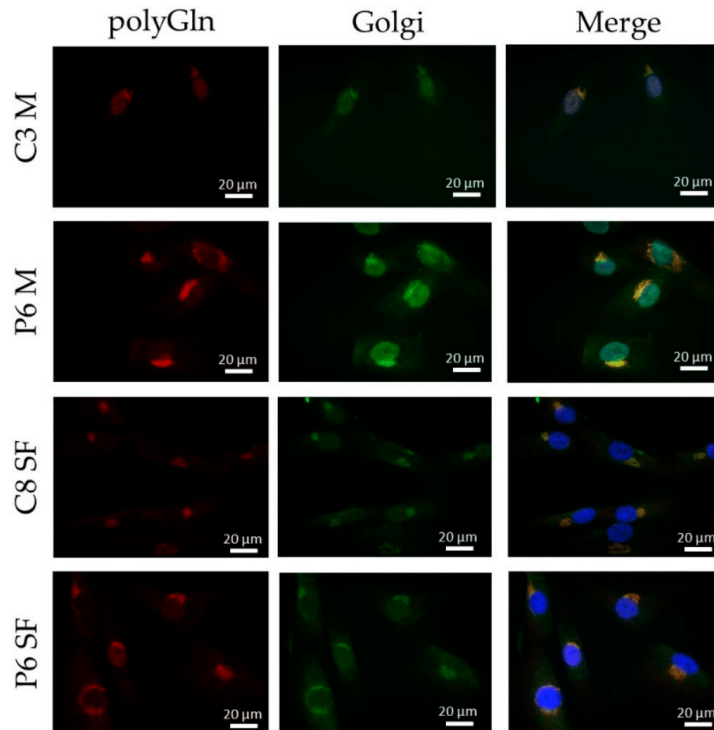


**Figure 4.** Qualitative analysis of polyGln RAN proteins in all DM1 cells. (A) Immunofluorescence polyGln staining with #1874 (alexa fluor-488, green) of human control and DM1 myoblasts. Nuclei indicated by DAPI (blue) (C:  $n = 5$ , P:  $n = 6$ ) (B) Immunofluorescence polyGln staining with #1874 (alexa fluor-488, green) of human control and DM1 skin fibroblasts. Nuclei indicated by DAPI (blue) (C:  $n = 5$ , P:  $n = 8$ ). (C) Immunofluorescence polyGln staining with #1874 (alexa fluor-488, green) of human control and DM1 lymphoblastoids (C:  $n = 4$ , P:  $n = 5$ ). Nuclei indicated by DAPI (blue). Abbreviations: C = control; P = DM1 patient; polyGln = polyglutamine; M = myoblasts; SF = skin fibroblasts; L = lymphoblastoids.



### 3.5. The Polyglutamine Containing Protein Found Resides in the Golgi Apparatus

Due to the unexpected result of finding a positive staining in both DM1 patients and controls, we decided to further study the origin of this positive staining. The distinct structure found with the 1C2 and #1874 antibody resembled the structure of an organelle and we therefore decided to investigate this hypothesis. A double immunostaining with TGN-38, a marker for the Golgi apparatus, showed an exact match to the structure we found with the 1C2 antibody in DM1 cells (Figure 5).



**Figure 5.** Determination of the origin of the polyGln containing proteins found via the immunofluorescence. Immunofluorescence showing co-localization of the polyGln aggregate found in myoblasts and skin fibroblasts with the commercial antibodies (1C2, alexa fluor-594, red) with the Golgi apparatus (TGN-38, alexa fluor-488 in green). Nuclei stained with DAPI (blue). Abbreviations: C = control; P = DM1 patient; M = myoblasts; SF = skin fibroblasts; polyGln = polyglutamine; Golgi = Golgi apparatus.

## 4. Discussion

We studied the presence of antisense transcription and polyGln RAN protein in three primary cell cultures of patients with DM1, namely myoblast, skin fibroblast and lymphoblastoid cell lines, in order to further elucidate its contribution to DM1 pathology.

The presence of antisense transcription, the origin of RAN-translated polyGln, was validated in our three primary cell cultures with three different primer combinations, and lower levels of expression were observed in DM1 patients compared to controls, which reached significance for LK2/anti-N3 and LK1/anti-1B in myoblasts and LK1/anti-N3 and LK1/anti-1B in lymphoblastoids. Of note, the LK1/anti-N3 primer combination, which

encompasses the repeat region, revealed that it primarily detected the smaller transcripts, most likely corresponding to the wild-type allele. Only P6 showed an additional band in myoblasts and skin fibroblasts, which, based on the CTG repeat size, could correspond to the expanded repeat. P6 had the smallest expanded repeat (130 repeats) in our cohort and it could be that the other patients, carrying much larger expanded repeats, could not be detected by this method. This could be a potential explanation for the lower levels of expression seen in patients compared to controls. However, the lower levels were also observed with the other two primer combinations that did not encompass the expanded repeat, making it highly unlikely that the lower levels seen were solely due to the binding of the wild-type allele. DM1-AS expression has only been studied by a handful of other groups. Zu and collaborators showed its presence in a heart sample of a DM1 patient and a healthy control; however, the expression patterns were hard to interpret, since an endogenous control was lacking [6]. A study by Gudde and collaborators showed a slightly higher expression in muscle biopsies of DM1 patients when studying RNA-seq data from the myotonic dystrophy deep sequencing data repository [5]. They did, however, mention that globally no obvious differences in read density were observed between DM1 patients and controls. However, when stratified based on inferred MBNL concentration, the most severely affected patients showed a three-fold increase in DM1-AS expression compared to controls, which was in vast contrast to the lower expression levels found in our cohort [5]. Another study, performed by Brouwer and collaborators, showed that in mouse models with increasing CTG repeat length, the DM1-AS transcription levels remained unchanged [15]. Unfortunately, disease severity in Gudde and collaborators' report was based on the inferred MBNL concentration of DM1 patients, and this was not available for our patients, which meant we could not do a similar stratification. We did, however, have extensive knowledge on the clinical phenotype of our DM1 cohort and had patients from three different clinical subtypes included in our studies, namely juvenile, adult and late-onset. Upon revision, a correlation between expression levels and clinical phenotype could not be found, based on for example, age of onset, muscle involvement (muscle weakness, myotonia) or CTG repeat size, the latter in agreement with the report of Brouwer and collaborators [15]. The sample size of our cohort was rather small for such comparisons or stratification, hindering the analysis. To determine whether DM1-AS transcript expression is linked to disease severity, a bigger cohort is needed.

The presence of DM1-AS transcripts in DM1 cells does not necessarily mean that these transcripts can reach the cytoplasm and be RAN-translated. To further elucidate the localization of these transcripts, cellular fractionation was performed and revealed the presence of DM1-AS transcripts in the cytoplasm of both patients and controls, with a higher percentage in the nuclear fraction. The presence of cytoplasmic DM1-AS transcripts was previously reported by Gudde and collaborators, as they showed the presence of DM1-AS transcripts in the cytoplasmic fraction of myoblasts [5]. However, in both the fractionated and unfracionated DM1-AS pool, it was unclear whether the transcripts possessed the expanded repeat. The LK1/anti-N3 combination already hinted that not all of the DM1-AS transcripts had the expanded repeat. This was also shown by Gudde and collaborators, who found a heterogeneous pool of DM1-AS transcript sizes, with and without the expanded repeat [5]. To determine whether the DM1-AS transcripts in DM1 cells included the expanded repeat, we performed a FISH to detect antisense RNA foci and found that antisense RNA foci were present in both the nucleus and cytoplasm of DM1 cells, indicating that these DM1-AS transcripts contained the expanded repeat and RAN translation was therefore, hypothetically, possible. However, the number of antisense RNA foci compared to sense RNA foci was much lower and they were not present in all cells. In addition, the cytoplasmic antisense RNA foci were even rarer, with only approximately 10% of myoblasts and skin fibroblasts containing cytoplasmic RNA foci, indicating that the presence of DM1-AS transcripts with the expanded repeat in the cytoplasm was quite low. Previous reports on antisense RNA foci in DM1 have also shown that the amount of antisense RNA foci in the nucleus was less compared to sense RNA

foci [3,4]. The polyGln RAN protein was undetectable in all three of our primary cell cultures via two different approaches. Western blots revealed a 42 kD polyGln-containing protein with the two commercial anti-polyGln antibodies, which was most likely TBP. In fact, the original immunogen for the 1C2 antibody was the general transcription factor TBP, which contains a 38-Gln stretch and therefore matches our results. It was shown, however, that although TBP will always show up on Western blots in both patients and controls, the antibody favored the binding of longer stretches of polyGln, such as were present in Huntington's disease and cerebellar ataxia type 1 and 3 [16]. Accordingly, a certain subset of lymphoblastoids did show a band that might correspond to the polyGln RAN protein, but the custom  $\alpha$ -DM1 antibody showed a range of non-specific bands in both patients and controls and we were therefore unable to determine the origin of this protein with certainty. In addition, it is difficult to know the exact size of the polyGln RAN protein produced by the DM1-AS, as the disease is prone to somatic mosaicism. This means that cells of the same tissue can carry different CTG expansion sizes and it is therefore also possible to have a range of potential sizes for the protein originating from these transcripts [17,18]. However, Zu and collaborators showed a band just below 60 kD in a patient carrying 85 CTG-CAGs [6]. Our patients carried expansions much larger than that, and when estimating the molecular weight based on the CTG expansion size, it was possible to have polyGln RAN proteins in the range we found within the lymphoblastoid cell lines. This will remain, however, hypothetical, as it seems we do not have a proper functional custom DM1 polyGln RAN antibody and no positive control available to test its functionality.

Immunofluorescence revealed a cytoplasmic aggregate surrounding the nucleus in myoblasts and skin fibroblasts with both commercial anti-polyGln antibodies, which was found to be co-localized with the Golgi apparatus. Since the aggregate was visible in both patients and controls and no apparent differences were seen, this might indicate the detection of another endogenous polyGln-containing protein. For example, ataxin-2, the product of the spinocerebellar ataxia type 2 gene, contains 22 glutamines and resides in the Golgi apparatus [19]. In addition, our immunofluorescence did show staining of the nucleus at high antibody concentrations, which might be due to binding of the transcription factor TBP, also detected by the immunoblots (42 kD band). Taken together, this would mean that both commercial anti-polyGln antibodies bind to several endogenous polyGln-containing proteins, especially at higher antibody concentrations. However, no apparent differences were found between patients and controls across a wide range of concentrations, and the use of  $\alpha$ -DM1 antibody did not reveal these similar aggregates. This is in vast contrast to the results previously reported by Zu and collaborators, as they found nuclear polyGln RAN protein aggregates at low frequencies in a DM1 patient's myoblasts and skeletal muscle ( $n = 1$ ) and at higher frequencies in leukocytes from peripheral blood ( $n = 1$ ) [6]. The 1C2 antibody was used in their experiments to validate the specificity of their custom  $\alpha$ -DM1 antibody. Although one of the cell types we used was the same as theirs, i.e., myoblasts, neither antibody was able to find the polyGln RAN protein in our myoblast cell lines, nor in the other two primary cell cultures. In fact, although both types of antibodies showed a protein of approximately 42 kD, our simultaneous staining showed that it was not the same protein, indicating that the antibodies were not able to recognize the same proteins. This was surprising, as the commercial antibody was used to validate the custom antibody in the paper of Zu and collaborators [6]. Our DM1-AS results suggested that the presence of DM1-AS transcripts containing the expanded repeat in the cytoplasm of DM1 cells is quite a rare occurrence. This highly affects the chance of producing polyGln RAN proteins. In addition, polyGln-containing proteins are very common in healthy subjects. Taking these two notions together, it might be plausible that with current techniques, sensitivity is too low to detect such low quantities of the polyGln RAN protein, which in addition is hindered by the presence of other polyGln containing proteins.

Although we were unable to detect polyGln RAN proteins in our DM1 cells, much progress has been made in other repeat expansion disorders displaying RAN translation, which could help in the field of DM1. Nine expansion disorders have been added since the first discovery of RAN translation in SCA8 and DM1: C9orf72 amyotrophic lateral sclerosis/frontotemporal dementia [20–22], fragile X tremor/ataxia syndrome [23], Huntington’s disease (HD) [24,25], spinocerebellar ataxia 3 and 31 [26,27], Fuchs’ endothelial corneal dystrophy [28] and myotonic dystrophy type 2 (DM2) [29]. Of these, SCA8, SCA3 and HD are the three repeat expansion disorders in which the RAN proteins originate from a CAG expansion, and can therefore result in polyGln RAN proteins. Interestingly, *in vivo*, none of these diseases show polyGln RAN proteins, but instead produce poly-alanine, and for HD additionally poly-serine RAN proteins. It might be interesting to include custom antibodies for the two additional homo-polymeric protein possibilities with regard to DM1. Although the name suggests a close relationship between DM1 and DM2, the underlying expansion in DM2 is a CCTG expansion and therefore results in complex poly-LPAC (sense) or poly-QAGR (antisense) RAN proteins, and is thus not hindered by the presence of endogenous polyGln proteins. The study was performed in autopsy brains, a tissue not yet studied for DM1, which might also be worth exploring.

In conclusion, DM1-AS transcript levels were lower in patients compared to controls and were present in both the nucleus and the cytoplasm of DM1 cells. Only a small portion of the DM1-AS transcripts contained the expanded repeat, substantially lowering the possibility of RAN translation in DM1. The polyGln RAN protein was not present in patient-derived DM1 cells, or was present in such low quantities that it is below the detection limit of the currently available techniques.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/jcm10235520/s1>. Table S1: Primer sequences. Figure S1: DM1-AS transcripts in DM1 cells of the additional two primer combinations studied. Figure S2: RNA foci overview. Table S2: Sense RNA foci details in DM1 myoblasts. Table S3: Antisense RNA foci details in DM1 myoblasts. Table S4: Sense RNA foci details in DM1 skin fibroblasts. Table S5: Antisense RNA foci details in DM1 skin fibroblasts. Table S6: Sense RNA foci details in DM1 lymphoblastoids. Table S7: Antisense RNA foci details in DM1 lymphoblastoids. Figure S3: Quantitative analysis of the polyGln RAN protein and antibody validation. Figure S4: Qualitative analysis of polyGln RAN proteins with the #1874 antibody of additional samples. Figure S5: Qualitative analysis of polyGln RAN proteins with the 1C2 antibody. Figure S6: Qualitative analysis of polyGln RAN proteins with the  $\alpha$ -DM1 antibody. Figure S7: Double immunofluorescence with the  $\alpha$ -DM1 antibody and 1C2 antibody.

**Author Contributions:** Conceptualization, G.N.-G. and E.K.; methodology, E.K., M.S., G.N.-G. and J.C.; validation, M.S. and G.N.-G.; formal analysis, E.K.; investigation, E.K., J.N.-M., A.B.-L., J.C., A.P.G.-E., M.A., G.L., A.A., G.P.-M., J.C.-C., A.R.-F. and A.M.-P.; resources, G.N.-G. and R.P.V.-M.; data curation, G.N.-G.; writing—original draft preparation, E.K.; writing—review and editing, G.N.-G., M.S., A.M.-P., J.N.-M., A.B.-L., M.A., G.L., A.A., J.C., R.P.V.-M., A.P.G.-E., G.P.-M., J.C.-C. and A.R.-F.; visualization, E.K. and J.C.; supervision, M.S. and G.N.-G.; project administration, E.K., M.S. and G.N.-G.; funding acquisition, G.N.-G., M.S. and A.M.-P. All authors have read and agreed to the published version of the manuscript.

**Funding:** The research of G. Nogales-Gadea and A. Ramos-Fransi is funded by Instituto de Salud Carlos III (grant numbers PI15/01756 and PI18/00713) and co-financed by Fondos FEDER. G. Nogales-Gadea is supported by a Miguel Servet research contract (ISCIII CD14/00032, ISCIII CPII19/00021, and FEDER) and by a Trampoline Grant #21108 from AFM Telethon. E. Koehorst is funded by the La Caixa Foundation (ID 100010434), fellowship code LCF/BQ/IN18/11660019, cofunded by the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no. 713673. The research of M. Suelves is funded by Ministerio de Ciencia e Innovación (grant number PID2020-118730RB-I00) and co-financed by Fondos FEDER. J. Núñez-Manchón is funded by Instituto de Salud Carlos III I-PFIS fellowship (grant number IFI20/00022). G. Lucente is supported by a Rio Hortega contract (ISCIII CM16/00016 and FEDER). J. Chojnacki is supported by European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no. 793830. The work of A.P. Gómez-Escribano and R.P. Vázquez-

Manrique is funded by the ISCIII (CPII16/00004, PI17/00011 and PI20/00114) and the Fundació Ramón Areces (CIVP19S8119). This work was supported by the CERCA program / Government of Catalonia. The funding bodies had no role in the design of the study and the collection, analysis, and interpretation of data.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of Hospital Universitario Germans Trias i Pujol, protocol code PI15/01756, date of approval 27 November 2015.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Acknowledgments:** The authors wish to thank Laura Ranum for generously providing the custom  $\alpha$ -DM1 antibody and Darren Monckton and Sarah Cumming for their help with the SP-PCRs. The authors also wish to thank Manuel Rodríguez-Allue and Daniel del Toro-Ruiz for their helpful comments and suggestions and the DM1 patients and the Huntington's patient for providing the samples needed to perform this study. We thank the IGTP core facilities for their contribution to this publication.

**Conflicts of Interest:** G. Pintos-Morell reports personal honoraria from Shire-Takeda, Amicus, and Sanofi-Genzyme, outside the submitted work.

## References

- Meola, G.; Cardani, R. Myotonic dystrophies: An update on clinical aspects, genetic, pathology, and molecular pathomechanisms. *Biochim. Biophys. Acta Mol. Basis Dis.* **2015**, *1852*, 594–606. [[CrossRef](#)]
- Cho, D.H.; Thienes, C.P.; Mahoney, S.E.; Analau, E.; Filippova, G.N.; Tapscott, S.J. Antisense transcription and heterochromatin at the DM1 CTG repeats are constrained by CTCF. *Mol. Cell* **2005**, *20*, 483–489. [[CrossRef](#)]
- Huguet, A.; Medja, F.; Nicole, A.; Vignaud, A.; Guiraud-Dogan, C.; Ferry, A.; Decostre, V.; Hogrel, J.Y.; Metzger, F.; Hoeflich, A.; et al. Molecular, Physiological, and Motor Performance Defects in DMSXL Mice Carrying >1000 CTG Repeats from the Human DM1 Locus. *PLoS Genet.* **2012**, *8*, e1003043. [[CrossRef](#)]
- Michel, L.; Huguet-Lachon, A.; Gourdon, G. Sense and antisense DMPK RNA foci accumulate in DM1 tissues during development. *PLoS ONE* **2015**, *10*, e1003043. [[CrossRef](#)]
- Gudde, A.E.E.G.; van Heeringen, S.J.; de Oude, A.I.; van Kessel, I.D.G.; Estabrook, J.; Wang, E.T.; Wieringa, B.; Wansink, D.G. Antisense transcription of the myotonic dystrophy locus yields low-abundant RNAs with and without (CAG)<sub>n</sub> repeat. *RNA Biol.* **2017**, *14*, 1374–1388. [[CrossRef](#)]
- Zu, T.; Gibbens, B.; Doty, N.S.; Gomes-Pereira, M.; Huguet, A.; Stone, M.D.; Margolis, J.; Peterson, M.; Markowski, T.W.; Ingram, M.A.C.; et al. Non-ATG-initiated translation directed by microsatellite expansions. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 260–265. [[CrossRef](#)]
- Nguyen, L.; Cleary, J.D.; Ranum, L.P.W. Repeat-Associated Non-ATG Translation: Molecular Mechanisms and Contribution to Neurological Disease. *Annu. Rev. Neurosci.* **2019**, *42*, 227–247. [[CrossRef](#)]
- Miller, S.A.; Dykes, D.D.; Polesky, H.F. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* **1988**, *16*, 1215. [[CrossRef](#)] [[PubMed](#)]
- Gomes-Pereira, M.; Bidichandani, S.I.; Monckton, D.G. Analysis of Unstable Triplet Repeats Using Small-Pool Polymerase Chain Reaction. In *Trinucleotide Repeat Protocols*; Humana Press: Totowa, NJ, USA, 2004; Volume 277, pp. 61–76.
- Monckton, D.G.; Wong, L.J.; Ashizawa, T.; Caskey, C.T. Somatic mosaicism, germline expansions, germline reversions and intergenerational reductions in myotonic dystrophy males: Small pool PCR analyses. *Hum. Mol. Genet.* **1995**, *4*, 1–8. [[CrossRef](#)]
- Rio, D.C.; Ares, M.; Hannon, G.J.; Nilsen, T.W. Preparation of cytoplasmic and nuclear RNA from tissue culture cells. *Cold Spring Harb. Protoc.* **2010**, *2010*, pdb.prot5441. [[CrossRef](#)] [[PubMed](#)]
- Ballester-Lopez, A.; Koehorst, E.; Linares-Pardo, I.; Núñez-Manchón, J.; Almendrote, M.; Lucente, G.; Arbex, A.; Alonso, C.P.; Lucia, A.; Monckton, D.G.; et al. Preliminary findings on ctg expansion determination in different tissues from patients with myotonic dystrophy type 1. *Genes* **2020**, *11*, 1321. [[CrossRef](#)]
- Ballester-Lopez, A.; Núñez-Manchón, J.; Koehorst, E.; Linares-Pardo, I.; Almendrote, M.; Lucente, G.; Guanyabens, N.; Lopez-Osias, M.; Suárez-Mesa, A.; Hanick, S.A.; et al. Three-dimensional imaging in myotonic dystrophy type 1. *Neurol. Genet.* **2020**, *6*, e484. [[CrossRef](#)]
- Ballester-Lopez, A.; Koehorst, E.; Almendrote, M.; Martínez-Piñero, A.; Lucente, G.; Linares-Pardo, I.; Núñez-Manchón, J.; Guanyabens, N.; Cano, A.; Lucia, A.; et al. A DM1 family with interruptions associated with atypical symptoms and late onset but not with a milder phenotype. *Hum. Mutat.* **2020**, *41*, 420–431. [[CrossRef](#)]

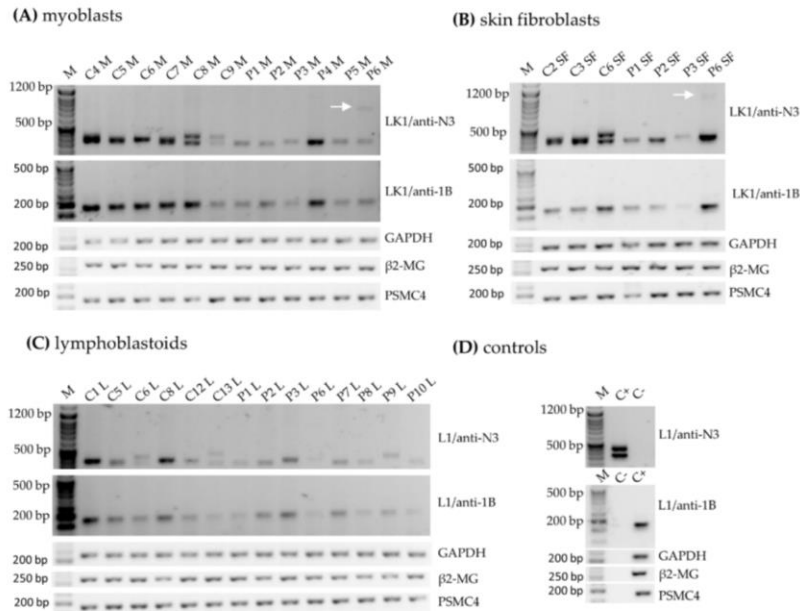
15. Brouwer, J.R.; Huguet, A.; Nicole, A.; Munnich, A.; Gourdon, G. Transcriptionally repressive chromatin remodelling and CpG methylation in the presence of expanded CTG-repeats at the DM1 locus. *J. Nucleic Acids* **2013**, *2013*, 567435. [[CrossRef](#)] [[PubMed](#)]
16. Trottier, Y.; Lutz, Y.; Stevanin, G.; Imbert, G.; Devys, D.; Cancel, G.; Saudou, F.; Weber, C.; David, G.; Tora, L.; et al. Polyglutamine expansion as a pathological epitope in Huntington's disease and four dominant cerebellar ataxias. *Nature* **1995**, *378*, 403–406. [[CrossRef](#)]
17. Morales, F.; Couto, J.M.; Higham, C.F.; Hogg, G.; Cuenca, P.; Braida, C.; Wilson, R.H.; Adam, B.; del Valle, G.; Brian, R.; et al. Somatic instability of the expanded CTG triplet repeat in myotonic dystrophy type 1 is a heritable quantitative trait and modifier of disease severity. *Hum. Mol. Genet.* **2012**, *21*, 3558–3567. [[CrossRef](#)] [[PubMed](#)]
18. Morales, F.; Vásquez, M.; Corrales, E.; Vindas-Smith, R.; Santamaría-Ulloa, C.; Zhang, B.; Sirito, M.; Estecio, M.; Krahe, R.; Monckton, D. Longitudinal increases in somatic mosaicism of the expanded CTG repeat in myotonic dystrophy type 1 are associated with variation in age-at-onset. *Hum. Mol. Genet.* **2020**, *29*, 2496–2507. [[CrossRef](#)]
19. Huynh, D.P.; Yang, H.T.; Vakharia, H.; Nguyen, D.; Pulst, S.M. Expansion of the polyQ repeat in ataxin-2 alters its Golgi localization, disrupts the Golgi complex and causes cell death. *Hum. Mol. Genet.* **2003**, *12*, 1485–1496. [[CrossRef](#)]
20. Ash, P.E.A.; Bieniek, K.F.; Gendron, T.F.; Caulfield, T.; Lin, W.-L.; DeJesus-Hernandez, M.; van Blitterswijk, M.M.; Jansen-West, K.; Paul, J.W.; Rademakers, R.; et al. Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron* **2013**, *77*, 639–646. [[CrossRef](#)] [[PubMed](#)]
21. Zu, T.; Liu, Y.; Bañez-Coronel, M.; Reid, T.; Pletnikova, O.; Lewis, J.; Miller, T.M.; Harms, M.B.; Falchook, A.E.; Subramony, S.H.; et al. RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E4968–E4977. [[CrossRef](#)]
22. Mori, K.; Weng, S.M.; Arzberger, T.; May, S.; Rentzsch, K.; Kremmer, E.; Schmid, B.; Kretzschmar, H.A.; Cruts, M.; Van Broeckhoven, C.; et al. The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTL/ALS. *Science* **2013**, *339*, 1335–1338. [[CrossRef](#)] [[PubMed](#)]
23. Todd, P.K.; Oh, S.Y.; Krans, A.; He, F.; Sellier, C.; Frazer, M.; Renoux, A.J.; Chen, K.; Scaglione, K.M.; Basrur, V.; et al. CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome. *Neuron* **2013**, *78*, 440–455. [[CrossRef](#)]
24. Banez-Coronel, M.; Ayhan, F.; Tarabochia, A.D.; Zu, T.; Perez, B.A.; Tusi, S.K.; Pletnikova, O.; Borchelt, D.R.; Ross, C.A.; Margolis, R.L.; et al. RAN Translation in Huntington Disease. *Neuron* **2015**, *88*, 667–677. [[CrossRef](#)]
25. Davies, J.E.; Rubinsztein, D. Polyalanine and polyserine frameshift products in Huntington's disease. *J. Med. Genet.* **2006**, *43*, 893–896. [[CrossRef](#)]
26. Ishiguro, T.; Sato, N.; Ueyama, M.; Fujikake, N.; Sellier, C.; Kanegami, A.; Tokuda, E.; Zamiri, B.; Gall-Duncan, T.; Mirceta, M.; et al. Regulatory Role of RNA Chaperone TDP-43 for RNA Misfolding and Repeat-Associated Translation in SCA31. *Neuron* **2017**, *94*, 108–124.e7. [[CrossRef](#)]
27. Toulouse, A.; Au-Yeung, F.; Gaspar, C.; Roussel, J.; Dion, P.; Rouleau, G.A. Ribosomal frameshifting on MJD-1 transcripts with long CAG tracts. *Hum. Mol. Genet.* **2005**, *14*, 2649–2660. [[CrossRef](#)] [[PubMed](#)]
28. Soragni, E.; Petrosyan, L.; Rinkoski, T.A.; Wieben, E.; Baratz, K.; Fautsch, M.; Gottesfeld, J. Repeat-Associated Non-ATG (RAN) Translation in Fuchs' Endothelial Corneal Dystrophy. *Investig. Ophthalmol. Vis. Sci.* **2018**, *59*, 1888–1896. [[CrossRef](#)]
29. Zu, T.; Cleary, J.D.; Liu, Y.; Bañez-Coronel, M.; Bubenik, J.L.; Ayhan, F.; Ashizawa, T.; Xia, G.; Clark, H.B.; Yachnis, A.T.; et al. RAN Translation Regulated by Muscleblind Proteins in Myotonic Dystrophy Type 2. *Neuron* **2017**, *95*, 1292–1305.e5. [[CrossRef](#)] [[PubMed](#)]

## Supplementary Materials

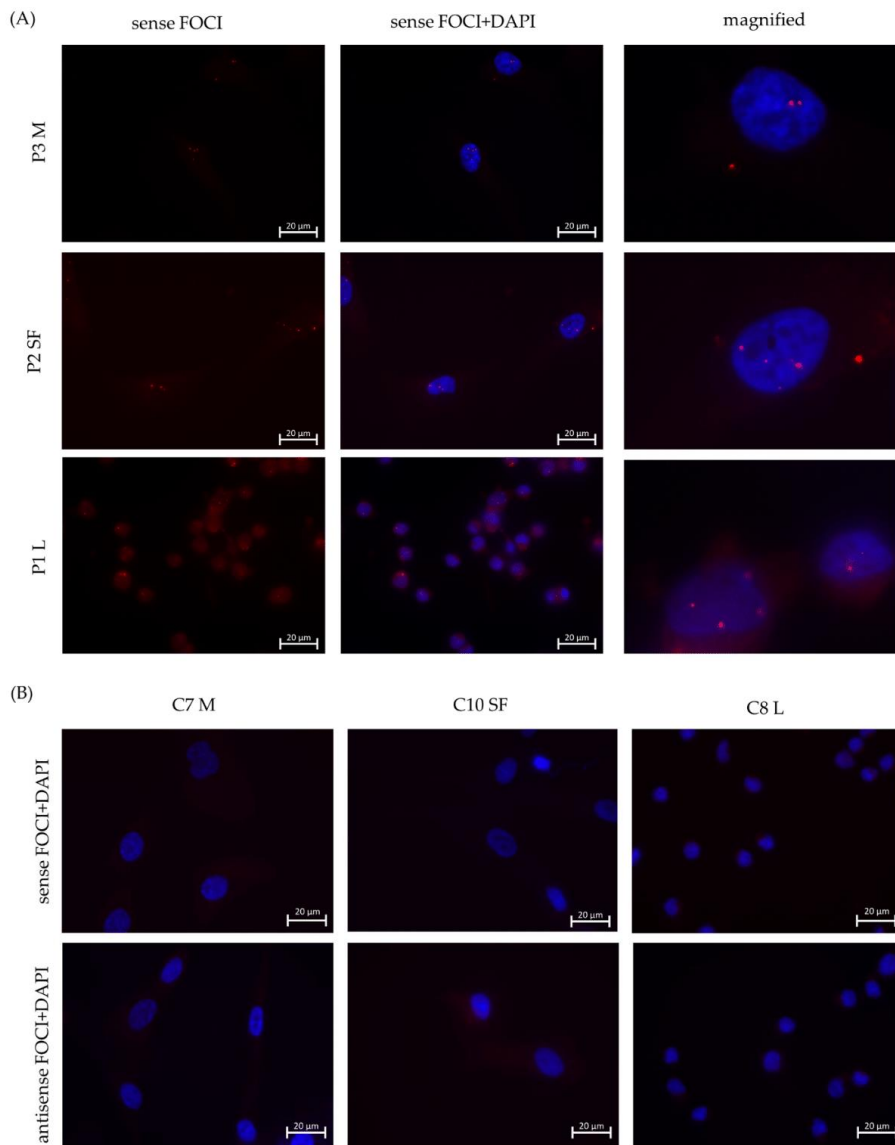
Table S1. Primer sequences.

Name primer	Sequence	Tm	Cycles	Expected size
LK1	5'-CGCCTGCCAGTTCACAACCGCTCCGAGCGT-3'			
LK2	5'-GACCATTTCITTTCTTTCCGGCCAGGCTGAGGC-3'			
linker	5'-CGACTGGAGCACGAGGACACTGA-3'			
Anti-N3	5'-GAGCAGGGCGTCATGCACAAG-3'	63	30	LK1: 349 bp LK2: 162 bp
Anti-1B	5'-GCAGCATTCCCGCTACAAGACCCTTC-3'	67	30	LK1: 150 bp
GAPDH-fw	5'-GAAGGTGAAGGTCGGAGTC-3'			
GAPDH-rev	5'-GAAGATGGTGTATGGGATTTC-3'	58	30	226 bp
$\beta$ 2-MG-fw	5'-CCAGCAGAGAATGGAAAGTC-3'			
$\beta$ 2-MG-rev	5'-GATGCTGCTTACATGTCTCG-3'	60	40	269 bp
PSMC4-fw	5'-TGTGGCAAAGGCGGTGGCA-3'			
PSMC4-rev	5'-TCTCTTGGTGGCGATGGCAT-3'	60	40	182 bp
DMPKex2-in2-fw	5'-GAGGGACGACTTCGAGATTCTGAA-3'			
DMPKex2-in2-rev	5'-CACCACGAGTCAAGTCAGGC-3'	67	40	92 bp

fw= forward; rev= reverse; bp= base pairs; Tm= melting temperature.



**Figure S1.** DM1-AS transcripts in DM1 cells of the additional two primer combinations studied. (A) DM1-AS specific primers LK1 and anti-N3/anti-1B for myoblasts (A), skin fibroblasts (B) and lymphoblastoids (C). RT-PCR controls are depicted in (D). C- = no DNA in RT-reaction and C+ = RNA of a control heart sample, the tissue used in original paper. Abbreviations: C= control; P= DM1 patient; bp= base pairs; M= marker. White arrows indicate the potential expanded repeat.



**Figure S2.** RNA foci overview. (A) sense RNA foci in all three cell types of DM1 patients. Cy3-labeled (CAG)10 probe (RED) showing sense RNA foci, nuclei indicated by DAPI (blue). Note that cytoplasmic sense RNA foci are present in both myoblasts and skin fibroblasts. (B) No presence of sense or antisense RNA foci in all three cell types of controls. Cy3-labeled (CAG)10 for sense RNA foci and Cy3-labeled (CTG)10 probe for antisense RNA foci (RED), nuclei indicated by DAPI (blue). P= DM1 patient; C= control; M= myoblasts; SF= skin fibroblasts; L= lymphoblastoids.



**Table S2.** Sense RNA foci details in DM1 myoblasts.

Myoblasts	Average RNA foci / cell		Maximum RNA foci / cell		Minimum RNA foci / cell		% Cells with RNA foci	
	nuclear	cytoplasm	nuclear	cytoplasm	nuclear	cytoplasm	nuclear	cytoplasm
P1	2.94 ± 1.70	0.00 ± 0.00	4.00	0.00	1.00	0.00	100.00	0.00
P3	4.30 ± 2.30	0.10 ± 0.31	11.00	1.00	1.00	0.00	100.00	10.00
P4	3.50 ± 1.40	0.50 ± 0.69	7.00	2.00	2.00	0.00	100.00	40.00
P5	3.00 ± 1.86	0.20 ± 0.42	8.00	1.00	1.00	0.00	100.00	20.00
Mean	3.44 ± 1.89	0.27 ± 0.52	7.50 ± 2.89	1.00 ± 0.82	1.25 ± 0.50	0.00 ± 0.00	100.00 ± 0.00	17.50 ± 17.08

Average/mean ± SD.

**Table S3.** Antisense RNA foci details in DM1 myoblasts.

Myoblasts	Average RNA foci / cell		Maximum RNA foci / cell		Minimum RNA foci / cell		% Cells with RNA foci	
	nuclear	cytoplasm	nuclear	cytoplasm	nuclear	cytoplasm	nuclear	cytoplasm
P1	0.30 ± 0.57	0.05 ± 0.22	2.00	1.00	0.00	0.00	25.00	5.00
P3	1.10 ± 0.97	0.70 ± 1.30	3.00	4.00	0.00	0.00	70.00	30.00
P4	0.25 ± 0.72	0.15 ± 0.49	1.00	1.00	0.00	0.00	15.00	10.00
P5	0.35 ± 0.67	0.10 ± 0.45	2.00	2.00	0.00	0.00	25.00	5.00
Mean	0.50 ± 0.81	0.25 ± 0.77	2.00 ± 0.82	2.00 ± 1.41	0.00 ± 0.00	0.00 ± 0.00	33.75 ± 24.62	12.50 ± 11.90

Average/mean ± SD.

**Table S4.** Sense RNA foci details in DM1 skin fibroblasts.

Skin fibroblasts	Average RNA foci / cell		Maximum RNA foci / cell		Minimum RNA foci / cell		% Cells with RNA foci	
	nuclear	cytoplasm	nuclear	cytoplasm	nuclear	cytoplasm	nuclear	cytoplasm
P2	3.55 ± 1.61	0.40 ± 0.82	8.00	2.00	1.00	0.00	100.00	25.00
P4	3.60 ± 1.43	0.10 ± 0.31	6.00	1.00	1.00	0.00	100.00	10.00
P7	3.45 ± 1.19	0.00 ± 0.00	6.00	0.00	0.00	0.00	100.00	0.00
P8	3.85 ± 1.63	0.35 ± 0.81	8.00	3.00	2.00	0.00	100.00	20.00
Mean	3.61 ± 1.45	0.21 ± 0.61	7.00 ± 1.17	1.50 ± 1.29	1.00 ± 0.82	0.00 ± 0.00	100.00 ± 0.00	13.75 ± 11.09

Average/mean ± SD.

**Table S5.** Antisense RNA foci details in skin fibroblasts.

Skin fibroblasts	Average RNA foci / cell		Maximum RNA foci / cell		Minimum RNA foci / cell		% Cells with RNA foci	
	nuclear	cytoplasm	nuclear	cytoplasm	nuclear	cytoplasm	nuclear	cytoplasm
P2	0.65 ± 0.93	0.05 ± 0.22	4.00	1.00	0.00	0.00	50.00	5.00
P4	0.20 ± 0.41	0.15 ± 0.49	1.00	2.00	0.00	0.00	20.00	10.00
P7	0.20 ± 0.41	0.10 ± 0.31	1.00	1.00	0.00	0.00	20.00	10.00
P8	0.20 ± 0.52	0.10 ± 0.31	2.00	1.00	0.00	0.00	15.00	10.00
Mean	0.31 ± 0.63	0.10 ± 0.34	2.00 ± 1.41	1.25 ± 0.50	0.00 ± 0.00	0.00 ± 0.00	26.25 ± 16.01	8.75 ± 2.50

Average/mean ± SD.

**Table S6.** Sense RNA foci details in lymphoblastoids.

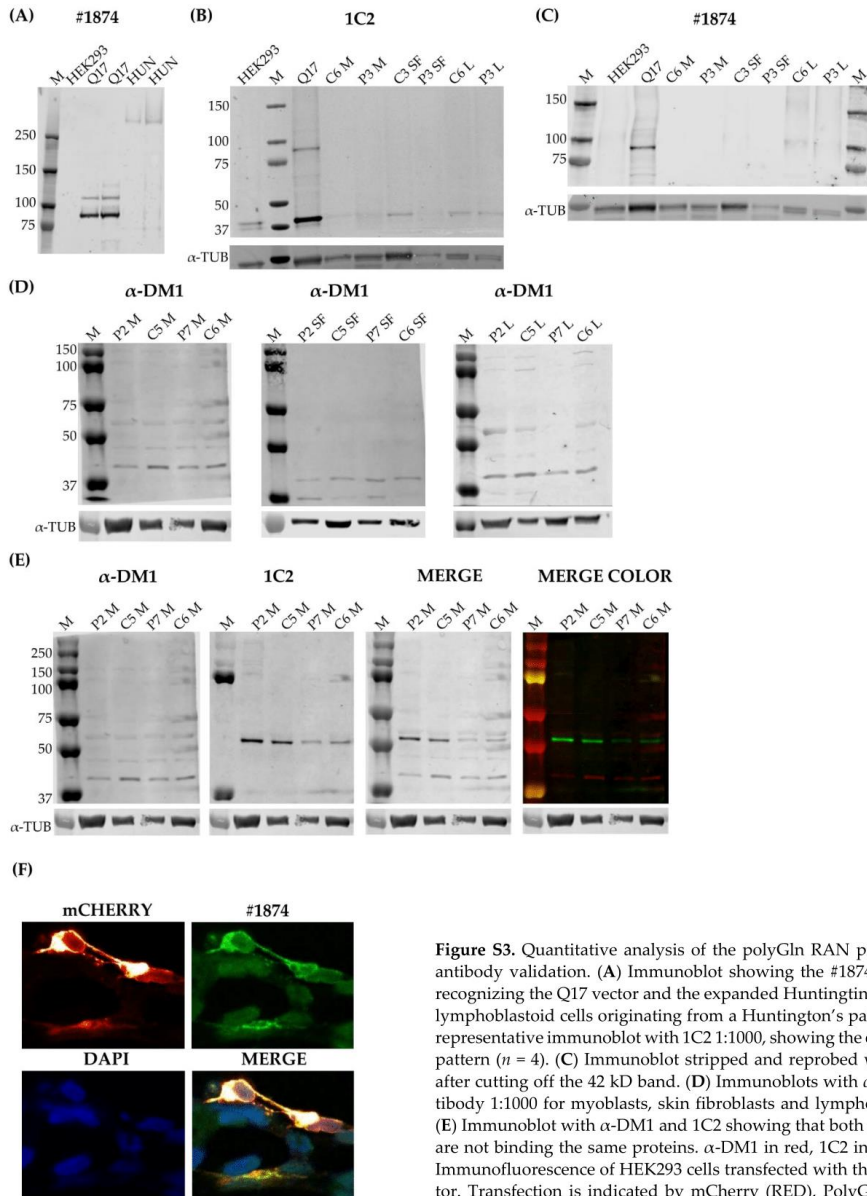
Lymphoblastoids	Average RNA foci / cell		Maximum RNA foci / cell		Minimum RNA foci / cell		% Cells with RNA foci	
	nuclear	cytoplasm	nuclear	cytoplasm	nuclear	cytoplasm	nuclear	cytoplasm
P1	0.85 ± 1.09	N.D.	4.00	N.D.	0.00	N.D.	50.00	N.D.
P2	0.81 ± 1.11	N.D.	3.00	N.D.	0.00	N.D.	44.00	N.D.
P7	1.06 ± 1.30	N.D.	5.00	N.D.	0.00	N.D.	58.80	N.D.
P8	1.29 ± 1.38	N.D.	5.00	N.D.	0.00	N.D.	60.00	N.D.
Mean	1.01 ± 1.22	N.D.	4.25 ± 0.96	N.D.	0.00 ± 0.00	N.D.	53.20 ± 7.58	N.D.

Average/mean ± SD. Cytoplasm too small to make an accurate distinction, therefore cytoplasmic RNA foci were not counted, N.D. = not determined.

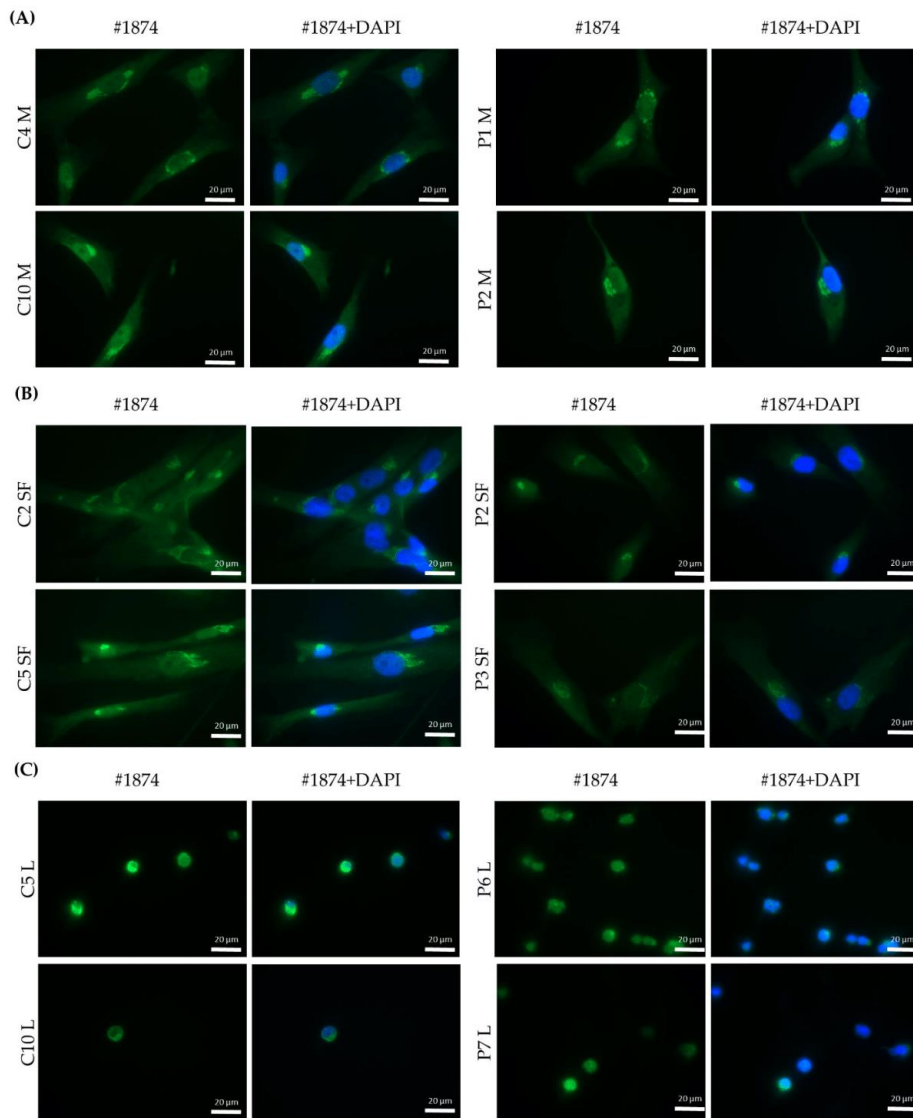
Table S7. Antisense RNA foci details in lymphoblastoids.

Lymphoblastoids	Average RNA foci / cell		Maximum RNA foci / cell		Minimum RNA foci / cell		% Cells with RNA foci	
	nuclear	cytoplasm	nuclear	cytoplasm	nuclear	cytoplasm	nuclear	cytoplasm
P1	0.05 ± 0.22	N.D.	1.00	N.D.	0.00	N.D.	5.00	N.D.
P2	0.15 ± 0.37	N.D.	1.00	N.D.	0.00	N.D.	15.00	N.D.
P7	0.00 ± 0.00	N.D.	0.00	N.D.	0.00	N.D.	0.00	N.D.
P8	0.00 ± 0.00	N.D.	0.00	N.D.	0.00	N.D.	0.00	N.D.
Mean	0.05 ± 0.22	N.D.	0.50 ± 0.58	N.D.	0.00 ± 0.00	N.D.	5.00 ± 7.07	N.D.

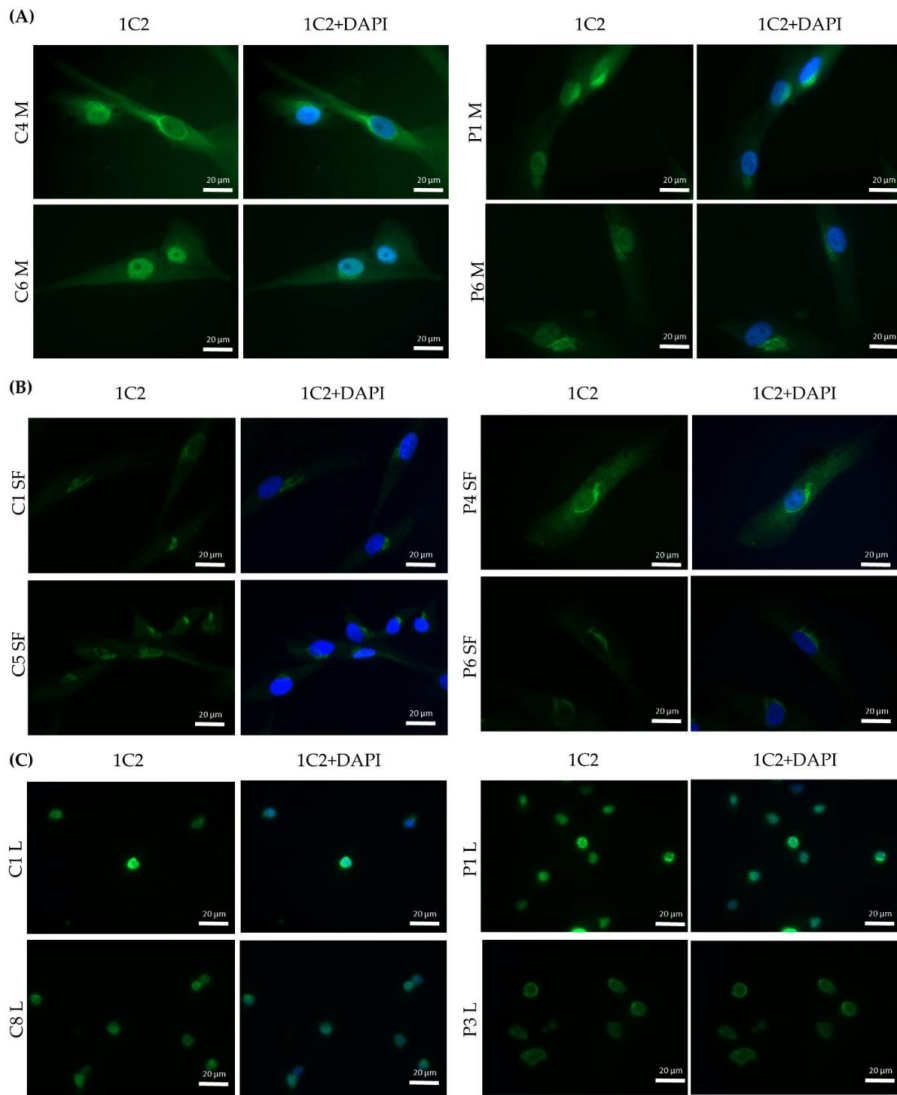
Average/mean ± SD. Cytoplasm too small to make an accurate distinction, therefore cytoplasmic RNA foci were not counted, N.D. = not determined.



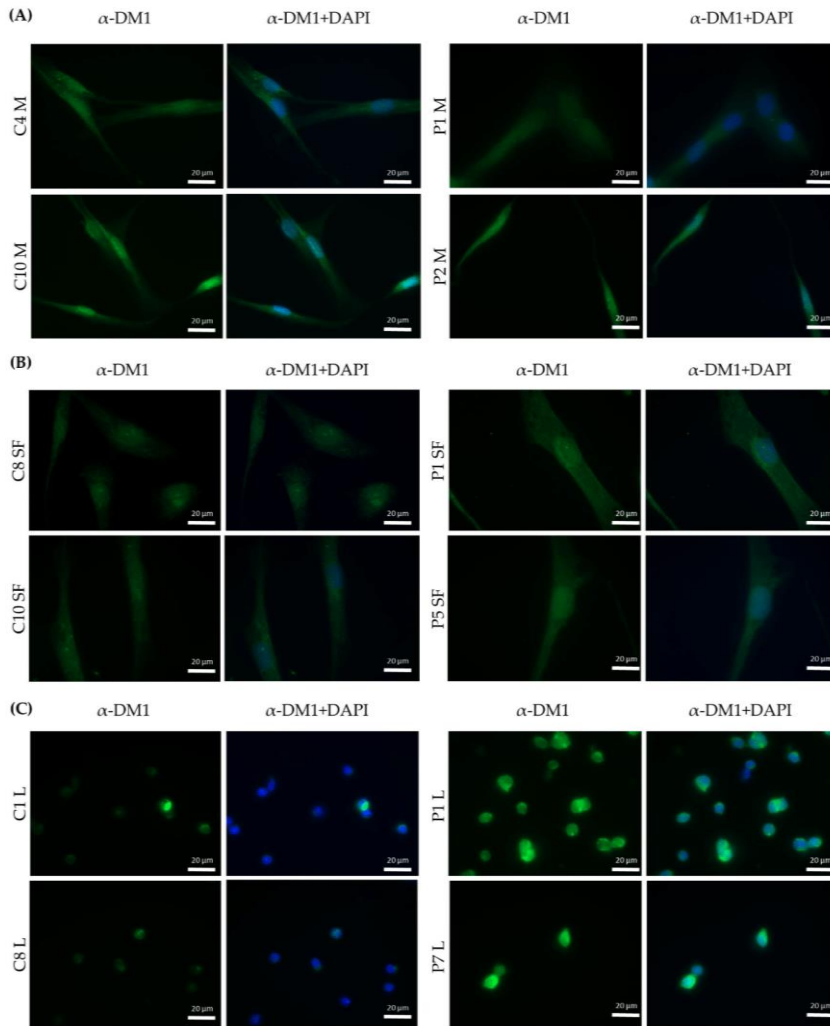
**Figure S3.** Quantitative analysis of the polyGln RAN protein and antibody validation. (A) Immunoblot showing the #1874 antibody recognizing the Q17 vector and the expanded Huntingtin protein in lymphoblastoid cells originating from a Huntington's patient. (B) a representative immunoblot with 1C2 1:1000, showing the exact same pattern ( $n = 4$ ). (C) Immunoblot stripped and reprobed with #1874 after cutting off the 42 kD band. (D) Immunoblots with  $\alpha$ -DM1 antibody 1:1000 for myoblasts, skin fibroblasts and lymphoblastoids. (E) Immunoblot with  $\alpha$ -DM1 and 1C2 showing that both antibodies are not binding the same proteins.  $\alpha$ -DM1 in red, 1C2 in green. (F) Immunofluorescence of HEK293 cells transfected with the Q17 vector. Transfection is indicated by mCherry (RED), PolyGln is indicated with #1874 (GREEN) and nuclei by DAPI (BLUE). Abbreviations: C = control; P = DM1 patient; M = myoblasts; SF = skin fibroblasts; L = lymphoblastoids; Q17 = Q17 vector; HUN = lymphoblastoids of a Huntington's patient.



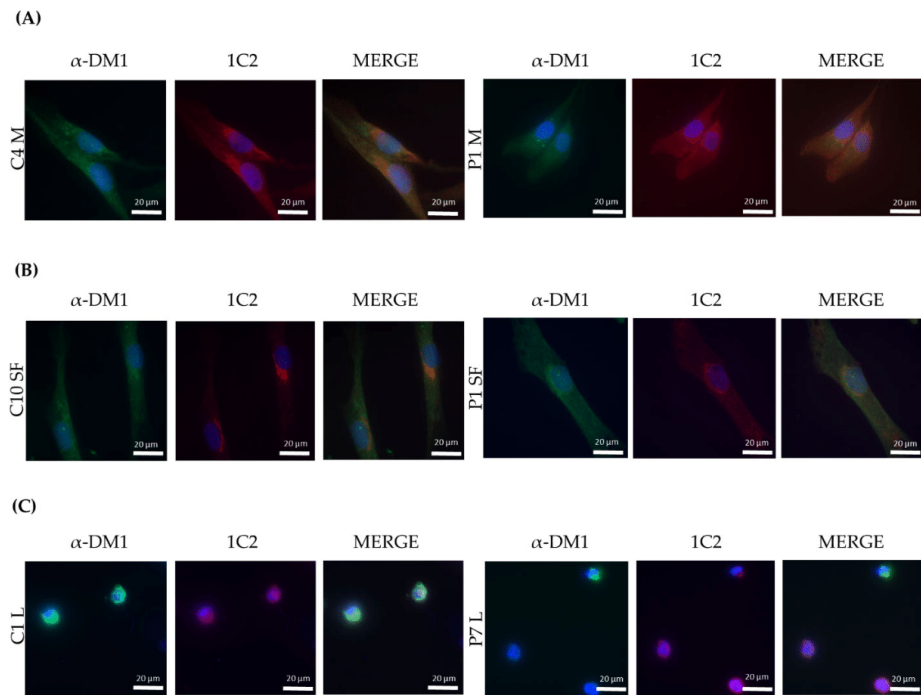
**Figure S4.** Qualitative analysis of polyGln RAN proteins with the #1874 antibody of additional samples. (A) Immunofluorescence polyGln staining with #1874 (alexa fluor-488, green) of human control and DM1 myoblasts (C:  $n = 5$ , P:  $n = 6$ ). Nuclei indicated by DAPI (blue) (B) Immunofluorescence polyGln staining with #1874 (alexa fluor-488, green) of human control and DM1 skin fibroblasts (C:  $n = 8$ , P:  $n = 5$ ). (C) Immunofluorescence polyGln staining with #1874 (alexa fluor-488, green) of human control and DM1 lymphoblastoids (C:  $n = 4$ , P:  $n = 5$ ). Nuclei indicated by DAPI (blue). Abbreviations: C = control; P = DM1 patient; M = myoblasts; SF = skin fibroblasts; L = lymphoblastoids.



**Figure S5.** Qualitative analysis of polyGln RAN proteins with the 1C2 antibody. (A) Immunofluorescence polyGln staining with 1C2 (alexa fluor-488, green) of human control and DM1 myoblasts (C: n=6, P: n=6). Nuclei indicated by DAPI (blue) (B) Immunofluorescence polyGln staining with 1C2 (alexa fluor-488, green) of human control and DM1 skin fibroblasts (C: n=8, P: n=8). (C) Immunofluorescence polyGln staining with 1C2 (alexa fluor-488, green) of human control and DM1 lymphoblastoids (C: n=4, P: n=6). Nuclei indicated by DAPI (blue). Abbreviations: C= control; P= DM1 patient; M= myoblasts; SF= skin fibroblasts; L= lymphoblastoids.



**Figure S6.** Qualitative analysis of polyGln RAN proteins with the  $\alpha$ -DM1 antibody. (A) Immunofluorescence polyGln staining with  $\alpha$ -DM1 (alexa fluor-488, green) of human control and DM1 myoblasts (C:  $n = 3$ , P:  $n = 4$ ). Nuclei indicated by DAPI (blue) (B) Immunofluorescence polyGln staining with  $\alpha$ -DM1 (alexa fluor-488, green) of human control and DM1 skin fibroblasts (C:  $n = 4$ , P:  $n = 4$ ). (C) Immunofluorescence polyGln staining with  $\alpha$ -DM1 (alexa fluor-488, green) of human control and DM1 lymphoblastoids (C:  $n = 4$ , P:  $n = 4$ ). Nuclei indicated by DAPI (blue). Abbreviations: C = control; P = DM1 patient; M = myoblasts; SF = skin fibroblasts; L = lymphoblastoids.



**Figure S7.** Double immunofluorescence with the  $\alpha$ -DM1 antibody and 1C2 antibody. (A) Immunofluorescence polyGln staining with  $\alpha$ -DM1 (alexa fluor-488, green) and 1C2 antibody (alexa fluor-594) of human control and DM1 myoblasts. Nuclei indicated by DAPI (blue) (B) Immunofluorescence polyGln staining with  $\alpha$ -DM1 (alexa fluor-488, green) and 1C2 antibody (alexa fluor-594) of human control and DM1 skin fibroblasts (C) Immunofluorescence polyGln staining with  $\alpha$ -DM1 (alexa fluor-488, green) and 1C2 antibody (alexa fluor-594) of human control and DM1 lymphoblastoids. Nuclei indicated by DAPI (blue). Abbreviations: C = control; P = DM1 patient; M = myoblasts; SF = skin fibroblasts; L = lymphoblastoids.



## **CHAPTER III**





## **An integrative analysis of DNA methylation pattern in myotonic dystrophy type 1 samples reveals a distinct DNA methylation profile between tissues and a novel muscle-associated epigenetic dysregulation**

**Emma Koehorst**<sup>1</sup>, Renato Odria<sup>1</sup>, Júlia Capó<sup>1</sup>, Judit Núñez-Manchón<sup>1</sup>, Andrea Arbex<sup>1,2</sup>, Miriam Almendrote<sup>1,2</sup>, Ian Linares-Pardo<sup>1</sup>, Daniel Natera-de Benito<sup>3</sup>, Verónica Saez<sup>3</sup>, Andrés Nascimento<sup>3</sup>, Carlos Ortez<sup>3</sup>, Miguel Ángel Rubio<sup>4</sup>, Jordi Díaz-Manera<sup>5,6</sup>, Jorge Alonso-Pérez<sup>5</sup>, Giuseppe Lucente<sup>1,2</sup>, Agustín Rodríguez-Palmero<sup>1,7</sup>, Alba Ramos-Fransi<sup>1,2</sup>, Alicia Martínez-Piñeiro<sup>1,2</sup>, Gisela Nogales-Gadea<sup>1,\*</sup>, Mònica Suelves<sup>1,\*,\$</sup>

### **AFFILIATIONS.**

1. Neuromuscular and Neuropediatric Research Group, Institut d'Investigació en Ciències de la Salut Germans Trias i Pujol (IGTP), Campus Can Ruti, Universitat Autònoma de Barcelona, 08916 Badalona, Spain.
  2. Neuromuscular Pathology Unit, Neurology Service, Neuroscience Department, Hospital Universitari Germans Trias i Pujol, 08916 Badalona, Spain.
  3. Neuromuscular Unit, Neuropediatric Department, Institut de Recerca Pediàtrica Hospital Sant Joan de Déu, Barcelona, Spain
  4. Neuromuscular Unit, Department of Neurology, Hospital del Mar, Barcelona, Spain.
  5. Neuromuscular Diseases Unit, Department of Neurology, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain.
  6. John Walton Muscular Dystrophy Research Centre, Newcastle University and Newcastle Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK
  7. Pediatric Neurology Unit, Department of Pediatrics. Hospital Universitari Germans Trias i Pujol, Autonomous University from Barcelona, Spain
- \* Equal contribution

**UNDER REVIEW**





---

---

## SUMMARY OF THE RESULTS

---

---

To investigate the contribution of epigenetics to the complexity of DM1, we compared the DNA methylation profiles across the four CpG islands residing in the *DMPK* locus in distinct DM1 tissues and tissue-derived cells across the different clinical phenotypes. For this study, DM1 patients from six different subcategories were included: congenital (n=6), childhood (n=6) and juvenile (n=23), which will also be referred to commonly as the developmental cases, whereas adult (n=20), late onset (n=7) and asymptomatic (n=2) will be referred to as the non-developmental cases. Age of onset ranged from just after birth until sixty-seven years, with a mean of seven years for childhood, 15 years for juvenile, 31 years for adult and 50 years for late onset. The presence of seven families was identified in this cohort. For 59 patients the CTG expansion could be sized, ranging from 115 to 1011 CTG repeats. Detailed information on their clinical phenotypes was obtained by neurologists and recorded in a clinical database.

The four CpGis were divided into five distinct regions: CpGi 74, CpGi 43, CpGi 36, CTCF1 and CTCF2. The latter two reside in the same CpG island and refer to the two regions that contain a CTCF binding site and also encompass the CTG expansion. In blood, CpGi 74, CpGi 43 and CpGi 36 showed no differences between patients and controls, with hypermethylation in CpGi 74 and CpGi 36 and a completely unmethylated region in CpGi 43. For the CTCF1 region upstream of the CTG repeat, 25 CpG sites were studied and increased levels of methylation were observed exclusively in the developmental cases, with 100% of the congenital cases, 50% of the childhood and 13% of juvenile cases. No changes in the non-developmental cases and the controls were found, with the exception of one adult case (P50). The two CpGs (CpG 18 and 19) that reside inside the CTCF1 binding site were both hypermethylated in the methylated cases. Fifty percent of the CTCF1 hypermethylated cases also showed hypermethylation in the CTCF2 region, of which CpG 5 resided in the CTCF2 binding site and was hypermethylated in the found cases.

Revision of the clinical phenotype of the developmental cases revealed a more severe muscular, cardiac and cognitive manifestation of the disease in the methylated childhood cases. No differences could be observed in the juvenile cases, potentially due to the low number of methylation cases in this subtype. No comparisons could be made in the congenital cases, as all were hypermethylated.

A positive association was found between two measures of CTG expansion size and methylation status, which suggests that the larger the CTG expansion, the more likely the occurrence of hypermethylation at the CTCF1 region. Additionally, we found an increased maternal transmission in CTCF1 methylated cases, but notably there were patients that were methylated and paternally transmitted and patients that were maternally transmitted and unmethylated, indicating hypermethylation is not exclusively maternally transmitted.

Our study cohort included several families, giving us the opportunity to study inheritance of the DNA methylation profiles. We found the DNA methylation status of our patients to be not inheritable, as several unmethylated mothers gave birth to methylated cases. Interestingly, we observed that the offspring of methylated mothers carried contractions of the CTG expansion, while the offspring of unmethylated mothers carried expansions of the repeat.

Next, we addressed whether tissue-specific epigenetics at the *DMPK* locus exist in DM1. For this, we acquired of a subset of patients (juvenile, adult and late-onset origin) a muscle and skin biopsy. For skin biopsies, a similar DNA methylation pattern to blood was found in all five CpG regions, and for muscle, this also was true for CpGi 74, CpGi 36 and CTCF2. For CpGi 43, the muscle showed an overall hypermethylation in both patients and controls. However, the hypermethylation levels observed in DM1 muscle were much lower compared to control muscle; reductions were in some cases as high as 70%. For CTCF1, six out of seven muscle biopsies of the same patients showed hypermethylation. This hypermethylation was highest in the youngest biopsy, belonging to a juvenile case. Of these biopsies, we have the CTG expansion size available in muscle. We could, however, not link the CTG expansion size to the degree of methylation.

Of the blood samples and biopsies, cells were isolated to assess whether cellular models reflect accurately the origin tissue in terms of DNA methylation status. Lymphoblastoids preserved the DNA methylation pattern observed in all five regions in all clinical subtypes, with a few exceptions. P7 showed hypermethylation for CTCF1 in blood, but not in lymphoblastoid cells and in CTCF2, two patients gained low-grade hypermethylation in lymphoblastoids (~10% methylation), that was not present in blood. For skin fibroblasts, CpGi 74 and CpGi 36 and CTCF2 preserved their DNA methylation pattern, but CpGi 43 showed gained hypermethylation in the majority of both controls and patients (~10% methylation), with no differences between the two groups. CTCF1 in skin fibroblasts showed low-grade hypermethylation (average of around 10%) in DM1 patients, but no methylation in controls. For myoblasts and myotubes, similar patterns were observed as described in muscle for all five regions.

Altogether, our results showed a CTCF1 DNA hypermethylation gradient in blood of the developmental cases and CTCF1 hypermethylation correlated to disease severity and CTG expansion size. Hypermethylated cases showed a higher chance of maternal transmission and CTCF1 hypermethylation in the parent was associated with a contraction of the CTG expansion upon generational transmission. Notably, DM1 patient-derived cells mostly preserved the DNA methylation profiles observed in tissues. Finally, our results showed a DM1 muscle-specific epigenetic landscape, with a loss of methylation at CpGi 43, a region containing an alternative *DMPK* promoter, accompanied by a hypermethylation of the CTCF1 region in muscle and muscle-derived cells.

## An integrative analysis of DNA methylation pattern in myotonic dystrophy type 1 samples reveals a distinct DNA methylation profile between tissues and a novel muscle-associated epigenetic dysregulation

Emma Koehorst<sup>1</sup>, Renato Odria<sup>1</sup>, Júlia Capó<sup>1</sup>, Judit Núñez-Manchón<sup>1</sup>, Andrea Arbex<sup>1,2</sup>, Miriam Almendrote<sup>1,2</sup>, Ian Linares-Pardo<sup>1</sup>, Daniel Natera-de Benito<sup>3</sup>, Verónica Saez<sup>3</sup>, Andrés Nascimento<sup>3</sup>, Carlos Ortiz<sup>3</sup>, Miguel Ángel Rubio<sup>4</sup>, Jordi Díaz-Manera<sup>5,6</sup>, Jorge Alonso-Pérez<sup>5</sup>, Giuseppe Lucente<sup>1,2</sup>, Agustín Rodríguez-Palmero<sup>1,7</sup>, Alba Ramos-Fransi<sup>1,2</sup>, Alicia Martínez-Piñeiro<sup>1,2</sup>, Gisela Nogales-Gadea<sup>1,\*</sup>, Mònica Suelves<sup>1,\*</sup>

### AFFILIATIONS.

1. Neuromuscular and Neuropediatric Research Group, Institut d'Investigació en Ciències de la Salut Germans Trias i Pujol (IGTP), Campus Can Ruti, Universitat Autònoma de Barcelona, 08916 Badalona, Spain.
  2. Neuromuscular Pathology Unit, Neurology Service, Neuroscience Department, Hospital Universitari Germans Trias i Pujol, 08916 Badalona, Spain.
  3. Neuromuscular Unit, Neuropediatric Department, Institut de Recerca Pediàtrica Hospital Sant Joan de Déu, Barcelona, Spain
  4. Neuromuscular Unit, Department of Neurology, Hospital del Mar, Barcelona, Spain.
  5. Neuromuscular Diseases Unit, Department of Neurology, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain.
  6. John Walton Muscular Dystrophy Research Centre, Newcastle University and Newcastle Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK
  7. Pediatric Neurology Unit, Department of Pediatrics, Hospital Universitari Germans Trias i Pujol, Autonomous University from Barcelona, Spain
- \* Equal contribution

<sup>§</sup> Correspondence: Mònica Suelves: msuelves@igtp.cat

### ABSTRACT

Myotonic dystrophy type 1 (DM1) is a progressive, non-treatable, multi-systemic disorder. To investigate the contribution of epigenetics to the complexity of DM1, we compared DNA methylation profiles of four annotated CpG islands (CpGis) in the *DMPK* locus and neighbouring genes; in distinct DM1 tissues and derived cells, representing six DM1 subtypes, by bisulphite sequencing. In blood, we found no differences in CpGi 74, 43 and 36 in DNA methylation profile. In contrast, a CTCF1 DNA hypermethylation gradient was found with 100% hypermethylation in congenital, 50% in childhood and 13% in juvenile cases. CTCF1 hypermethylation correlated to disease severity and CTG expansion size. Notably, fifty percent of CTCF1 hypermethylated cases showed hypermethylation in the CTCF2 regions as well. Additionally, hypermethylation was associated with maternal transmission. Interestingly, the evaluation of seven families showed that unmethylated mothers passed on an expansion of the CTG repeat, whereas the methylated mothers transmitted a contraction. The analysis of patient-derived cells showed that DNA methylation profiles were highly preserved, validating their use as faithful DM1 cellular models. Importantly, the comparison of DNA methylation levels of distinct DM1 tissues revealed a novel muscle-specific epigenetic signature with hypermethylation of the CTCF1 region accompanied by demethylation of CpGi 43, a region containing an alternative *DMPK* promoter, which may decrease the canonical promoter activity. Altogether, our results showed a distinct DNA methylation profile across DM1 tissues and uncovered a dual epigenetic signature in DM1 muscle samples. Our results highlight the contribution of epigenetic changes at the *DMPK* locus to DM1 pathology.

**KEYWORDS:** Myotonic dystrophy, CpG islands, DNA methylation, Epigenetics, Phenotype severity, DM1 biopsies and cellular models

**CURRENTLY UNDER REVISION**

## 1. INTRODUCTION

Myotonic dystrophy type 1 (DM1) is an autosomal dominant inherited, multi-systemic disorder, with predominant muscle involvement and an estimated incidence of 1:8000<sup>1</sup>. DM1 has been recognized as one of the muscle dystrophies with the more variable phenotype, as it affects several tissues and systems, and because it has varied manifestations. It can be classified into five different clinical subtypes; which are based on the age of onset and they range from fetal to late-adult onset<sup>2</sup>.

The most severe form is congenital DM1 (CDM1), with an age of onset in the first year of life and an almost exclusive maternal transmission<sup>3</sup>. Neonatal manifestations include hypotonia, respiratory failure, feeding difficulties, failure to thrive and clubfoot deformities. Respiratory failure is a common cause of death in these patients in the first year of life<sup>4,5</sup>. When age of onset is between one to ten years, it is called childhood DM1. This form of DM1 often has a delay in diagnosis due to the uncharacteristic symptom manifestation. The first signs are cognitive and learning abnormalities, rather than muscle impairment<sup>6</sup>. The typical muscular signature, such as muscle weakness and myotonia, often does not develop until late adolescence<sup>7,8</sup>. The third clinical subset is the juvenile DM1, with an age of onset between eleven and twenty years. This group is often placed under the childhood or adult umbrella, but differs from childhood DM1 in their increased presence of myotonia and from adult-onset due to their more pronounced central nervous system involvement<sup>2</sup>. The most prevalent DM1 clinical subtype is the adult onset DM1 and is categorized by an age of onset between twenty and forty years. Core symptoms are progressive muscle weakness, myotonia and early-onset cataracts<sup>5,9–11</sup>. Additionally, cardiac conduction defects are common and are the leading cause of death<sup>12</sup>. The last clinical category is late-onset DM1, with an age of onset after forty and mild symptom presentation, such as low-grade muscle weakness, premature cataracts and alopecia. In addition to the five clinical categories, there is another special set of 'DM1 patients', the asymptomatic or paucisymptomatic DM1 category, characterized by the absence or just minor symptoms

across an individual's life span. Due to the wide variability in age of onset and symptomatology, and the presence of anticipation (development of the disease earlier in life with successive generations), genetic counselling of relatives is a common clinical practice, which leads to the identification of asymptomatic patients.

The causes of the clinical variability observed in DM1 are poorly understood. The main pathological mechanism of the disease has been shown to be RNA toxicity, caused by a CTG expansion in the 3' untranslated region of the dystrophin myotonia protein kinase (*DMPK*) gene. The CTG expansion results in the accumulation of expanded transcripts as intranuclear RNA foci, which in turn sequester a number of splicing factors, resulting in loss of function and downstream deregulation of the alternative splicing of several genes. CTG expansion size has been correlated to disease severity and age of onset, with the greater repeat sizes corresponding to a more severe disease and earlier onset<sup>13,14</sup>. This correlation is however not absolute, as several studies have shown CDM1 patients with small expanded repeats and late-onset patients with over a thousand repeats<sup>15–17</sup>. Alternative splicing defects and CTG expansion size cannot account for the entire symptomatology seen in DM1 patients. Additional potential mechanisms include bidirectional transcription<sup>18,19</sup>, repeat-associated non-ATG translation (although results are controversial)<sup>19,20</sup>, and epigenetic changes, such as distinct DNA methylation profiles<sup>21,22</sup>.

Epigenetics is defined as heritable changes that do not affect the DNA sequence itself but influence gene expression and it includes DNA methylation, histone modifications, and non-coding RNAs<sup>25</sup>. DNA methylation is the most widely studied epigenetic mark, which is essential for mammalian development, crucial for the establishment and maintenance of cell identity, and it affects gene expression by regulating promoters and distal regulatory elements, such as enhancers and insulators<sup>26–28</sup>. DNA methylation occurs most often on a cytosine, leading a guanine, which are referred to as CpG dinucleotides. They are globally underrepresented in the genome, except in CpG islands. CpG islands are CpG-dense regions largely

resistant to DNA methylation<sup>29,30</sup>. CpG islands are generally found at promoters of housekeeping and developmental genes, and are represented in 70% of promoters, but can be also found in exons, introns and regulatory regions<sup>31</sup>. Importantly, approximately 6% of CpG islands become methylated in a tissue-specific manner during early development or in differentiated tissues, highlighting DNA methylation as an important epigenetic mechanism in the establishment and maintenance of cellular identity<sup>32</sup>

The *DMPK* gene and neighboring genes (henceforth referred to as the *DMPK* locus) contain several CpG islands (CpGis). CpGi 374 has gained the most attention because this 3.5 kb island contains the expanded repeat. In addition, the CTG repeat is flanked by two CTCF-binding factor (CTCF) binding sites, named CTCF1 and CTCF2. Early studies suggested that the two CTCF binding sites together with the expanded repeat establish an insulator element between the *DMPK* promoter and the six homeobox 5 (*SIX5*) enhancer, affecting chromatin dynamics<sup>33</sup>. Several studies have found aberrant DNA methylation profiles in CTCF1 and CTCF2 regions in DM1, but the results are controversial<sup>33–43</sup>. CDM1 seems to be the clinical phenotype where hypermethylation plays the most important role, and CTCF1 hypermethylation has been associated with greater expansion sizes<sup>39,41–43</sup>, although this association is controversial and not always found<sup>34,37</sup>. The effect on clinical phenotype of aberrant DNA methylation profiles in the CTCF1 region is largely unknown. Nevertheless, it has been linked to respiratory and muscular profiles and a decline in cognitive function in adults<sup>38,44</sup>.

The *DMPK* locus harbors three more CpG islands, one in the neighboring myotonic dystrophy WD repeat containing (*DMWD*) gene, CpGi 74, and two in the *DMPK* gene, namely CpGi 43 and 36; however, nothing is known about the epigenetic state of these CpG islands in DM1. To date, only one publication has looked at the entire *DMPK* locus and this was solely done in control tissues and cell cultures<sup>45</sup>. Therefore, the main goal of this study was to elucidate the DNA methylation profiles across the four CpG islands residing in the *DMPK* locus in distinct DM1 tissues and tissue-derived cells across the different clinical phenotypes. Our results showed

a CTCF1 DNA hypermethylation gradient in blood of the developmental cases and CTCF1 hypermethylation correlated to disease severity and CTG expansion size. Hypermethylated cases showed a higher chance of maternal transmission and CTCF1 hypermethylation in the parent was associated with a contraction of the CTG expansion upon generational transmission. Notably, DM1 patient-derived cells preserved the DNA methylation profiles observed in tissues. Finally, our results showed a DM1 muscle-specific epigenetic landscape, with a loss of methylation at CpGi 43, a region containing an alternative *DMPK* promoter, accompanied by a hypermethylation of the CTCF1 region in muscle and muscle-derived cells. Altogether, our results offer novel insights into the epigenetic changes in DM1 pathology.

## 2. MATERIAL AND METHODS

### 2.1. Patient registry

This study was approved by the Ethics Committee of the University Hospital Germans Trias i Pujol and was performed in accordance with the Declaration of Helsinki for Human Research. Written informed consent was obtained for all participants. The study included 65 DM1 patients and 8 controls with no previous family history of neuromuscular disorders (recruited from the traumatology department in whom surgery was needed). DM1 diagnosis was confirmed or discarded with triplet primed-PCR in all the study participants. Clinical information of DM1 patients was obtained from the medical records and updated in the last visit by neurologists. Patients were subdivided into five different categories based on age of onset: congenital= first year of life, childhood= 1 to 10 years, juvenile= 11-20 years, adult= 21 to 40 years, late-onset= >40 years. Additionally, a group of asymptomatic patients was added. For three patients the exact year of age of onset could not be determined, but based on the clinical information all three were classified as adults. Clinical information included family history; details on the last ophthalmological, cardiological and respiratory examination by the corresponding specialists, including the electrocardiograms, echocardiograms



and spirometry tests performed in the last year. A full neurological work-up was performed by neurologists, including the presence of myotonia, ptosis, axial and facial weakness; muscle strength with the manual Medical Research Council (MRC) scale; and muscle impairment by the Muscular Impairment Rating scale (MIRS). In addition, the presence of cataracts; alopecia; intestinal problems and sleep disturbances were catalogued and the functional status and degree of disability were evaluated using the DM1-Activ questionnaire and modified Rankin Scale (mRS), respectively.

### 2.2. Tissue and cell culture

A total of three different samples from patients and controls were obtained: blood, muscle biopsy, and skin biopsy. Blood was obtained from all patients and lymphoblastoids were isolated when possible. From a subset of patients and controls, an additional muscle (Biceps Brachialis or Vastus Lateralis) and skin biopsy was taken, of which muscle and skin-derived cell cultures were obtained. Of note, to increase the number of biopsies, a subset of biopsies for which no blood was available were included. They included three extra patients and two controls. All samples were obtained at the same time. Blood was collected in EDTA and heparin tubes. The EDTA tubes were frozen at  $-20^{\circ}\text{C}$  before DNA extraction and the heparin tubes were used for peripheral blood mononuclear cell (PBMCs) isolation using a standard ficoll procedure. PBMCs were subsequently incubated with anti-human CD3 antibody to suppress T cells, and immortalized with Epstein Bar virus. Lymphoblastoids were further cultured with B95-8 medium (80% RPMI, 10% Fetal Bovine Serum, PSF 1x and L-Glutamine 200 nM).

The muscle biopsy was obtained from the left biceps brachialis in all individuals except for one patient (P68, Vastus Lateralis muscle). Muscle biopsies were frozen immediately and stored at  $-80^{\circ}\text{C}$  before DNA extraction. Skin biopsies were obtained with a 0.5 cm skin punch. For cell isolation, muscle and skin biopsies were cleaned and fragmented to explants, which were placed in plates with human serum and gelatin 1.5% (1:2), and cultured with DMEM supplemented with 5% of fetal bovine serum and PSF 1x, at  $37^{\circ}\text{C}$  for two to three

days. Derived cells from the tissue explants were further cultured with DMEM supplemented with 15% of fetal bovine serum and PSF 1x in the case of the skin fibroblasts, and supplemented additionally with 22% M-199, insulin 1.74  $\mu\text{M}$ , L-glutamine 2 mM, FGF 1.39 nM and EGF 0.135 mM in the case of the muscle cells. Myoblasts were purified through CD56 magnetic beads according to manufacturer's instructions (Miltenyi Biotec, Bergisch Gladbach, Germany), and they were further grown on pre-coated surfaces with 0.1% gelatin until 60-70% of confluence was achieved. To differentiate a subset of myoblasts into myotubes, after 24hr incubation in the above mentioned proliferation medium, the medium was changed to differentiation medium containing DMEM, supplemented with 2% horse serum, 1% L-Glutamine and 1x PSF. Differentiation medium was changed every 24 hours during 7 days to achieve differentiation. Skin Fibroblasts, isolated from biopsied tissue using the explant method, were grown until 70% confluence in a proliferation medium containing DMEM, supplemented with 10% FBS and 1x PSF.

### 2.3. DNA isolation

Genomic DNA was isolated from peripheral blood by the use of the QIAamp DNA mini kit (Qiagen, Hilden, Germany), the PureLink genomic DNA mini kit (Thermo Fisher Scientific, Waltham, MA, USA) or a simple salting procedure, as previously described<sup>46</sup>. Genomic DNA was extracted from muscle and skin tissue by homogenization in 100 mM Tris-HCl, pH 7.8, and 5 mM EDTA until these tissues were disaggregated. Thereafter tissues were digested in 20 mg/mL proteinase K and 10% SDS for 16 h at  $37^{\circ}\text{C}$ , and treated with 5.5 M NaCl, phenol and chloroform isoamyl (1:24) before DNA precipitation with isopropanol. Genomic DNA from lymphoblastoids, myoblasts, myotubes and skin fibroblasts was extracted by using the QIAamp DNA mini kit (Qiagen, Hilden, Germany), according to manufacturer's protocol.

### 2.4. CTG expansion size analysis

To estimate the length of the expanded progenitor allele, a specific long PCR with digested

DNA, followed by a Southern Blot was carried out. First, 250ng DNA was digested with EcoR I (New England Biolabs, Ipswich, MA, USA), according to manufacturer's protocol. 750pg of digested DNA was used in the subsequent PCR, with the Expand Long Template PCR System (Roche, Basel, Switzerland) and primers DM-D and DM-DR (Table 1), according to manufacturer's guidelines, supplemented with 2% DMSO. The following thermocycler conditions were used: initial 3 min at 96 °C, followed by 28 cycles of 15s 96°C, 45s 63.5 °C, 5 min 68°C, and a final extension step of 1 min 63.5°C and 7 min 68°C. DNA fragments were resolved by electrophoresis on a 1% agarose gel. The gel was run for an initial 10 min at 200V, followed by ±19 hours at 27V and blotted onto a positively charged nylon membrane (Roche, Basel, Switzerland). The membrane was hybridized with a digoxigenin labelled seven CAG LNA probe overnight and detected by chemiluminescence using the anti-Dig-CDP-Star system (Roche, Basel, Switzerland). Two CTG expansion sizes were estimated through comparison against the molecular weight ladder using GelAnalyzer 19.1 software (www.gelanalyzer.com, by Istvan Lazar Jr. and Istvan Lazar Sr.). The CTG size of the progenitor (ePAL), which is the lowest range of band thought to originate from the transmitting parent, and the mode allele, which shows the densest collection of CTG sizes, thought to be the most representative size for the patient at that specific time.

### 2.5. Bisulfite treatment and Sanger sequencing of four CpG islands

The methylation status of four annotated CpG islands (CpGi), divided into 5 individual areas, in the DMPK locus was studied by using bisulfite treatment and subsequent Sanger sequencing (Figure 1A). For CpGi 74, CpGi 43 and CpGi 36, nineteen, fifteen and seventeen CpGs were studied respectively, (Figure S1A-C for detailed location). In the CpGi 374, two separate regions were studied, namely CTCF1 and CTCF2, which surround the CTG expansion and hold each a CTCF binding site. For CTCF1 25 CpGs were studied and for CTCF2 11 CpGs (Figure S1D-E for detailed information).

200-400 ng of DNA was bisulfite treated using the EZ DNA Methylation Gold kit (Zymo Research, Irvine, CA, USA), according to

manufacturer's guidelines. Bisulfite-treated DNA was amplified using nested and hemi-nested PCR for the CTCF1 and CTCF2 region, located in the CpGi 374 surrounding the CTG expansion, previously described<sup>34</sup>, with some minor modifications. Here we use the TaKaRa Taq DNA polymerase (TaKaRa Bio Inc., Kioto, Japan) on a Mastercycler nexus x2 thermocycler, primer combinations and thermocycler settings are listed in Table S1. Additionally, three regions further upstream of the CTG repeat were analyzed, namely the CpGi 36, CpGi 43 and CpGi 74 regions in a similar fashion, using different primer combinations and thermocycler settings (Table S1).

Amplicons were purified using illustra™ ExoProStar 1-Step (Merck, Darmstadt, Germany), according to manufacturer's protocol. This was followed by sequencing using the BigDye Terminator v3.1 cycle sequencing kit (Thermo Fisher Scientific, Waltham, MA, USA), following manufacturer's guidelines. Afterwards, amplicons were run on an ABI Prism 3130 Genetic Analyzer (Applied Biosystems, Waltham, MA, USA) and analyzed using Chromas software version 2.6.6 or FinchTV software version 1.5.0, as detailed in Carrió et al, 2016<sup>47</sup>. Sodium bisulfite sequencing data were represented with the Methylation Plotter web tool<sup>48</sup>. A sample was considered hypermethylated when more than one CpG showed ≥ 10% methylation.

### Statistical analysis

Dichotomous variable CTCF1 methylation status, the occurrence of abnormal methylation upstream of the repeat, was modelled as dependent variable using a logistic regression model, against the independent variable ePAL and modal allele. The program used was SPSS 28.0.0.0 and significance level was set at 0.05.

## 3. RESULTS

### 3.1. A study cohort encompassing all clinical subtypes of DM1

For this study, DM1 patients from six different subcategories were included. The first five categories are the different established clinical phenotypes:

congenital, childhood, juvenile, adult and late-onset. The sixth category is a special subset of patients, which are known to carry the CTG expansion, but are as of yet asymptomatic. Congenital (n=6), childhood (n=6) and juvenile (n=23) will also be referred to commonly as the developmental cases, whereas adult (n=22), late onset (n=6) and asymptomatic (n=2) will be referred to as the non-developmental cases. Age of onset ranged from first year of life until 67 years, with a mean of seven years for childhood, 15 years for juvenile, 31 years for adult and 52 years for late-onset. The presence of seven families was identified in this cohort. In 59 out of 65 patients, the CTG expansion could be sized, ranging from 115 to 1011 CTG repeats. Detailed information on the clinical phenotypes can be found in Table 1.

### 3.2. DNA methylation profiles across the *DMPK* locus in blood

This study aimed to elucidate the DNA methylation profiles across the four CpG islands residing in the *DMPK* locus in distinct DM1 patient samples and derived primary cell cultures. The four CpGs were divided into five distinct regions: CpGi 74, CpGi 43, CpGi 36, CTCF1 and CTCF2. The latter two reside in the same CpG island and refer to the two regions that contain a CTCF binding site and also encompass the CTG expansion (Figure 1A). In blood samples, for the first three CpGs, no differences in DNA methylation levels across the six phenotypes and the controls were observed. CpGi 74 and CpGi 36 showed hypermethylation (90-100%) across the 19 and 17 CpGs studied, respectively (Figure 1B and Tables S2 and S4). CpGi 43 showed no methylation across the 15 CpG sites studied (Figure 1B and Table S3). For the CTCF1 region upstream of the CTG repeat, 25 CpG sites were studied and increased levels of methylation were observed almost exclusively in the developmental cases, with 100% of the congenital cases, 50% of the childhood cases and 13% of juvenile cases (Figure 1C and Table S5). No increased methylation levels were found in the non-developmental cases and the controls, except for one adult case out of 30 (P50). The two CpGs (CpG 18 and 19) that reside inside the proposed CTCF1 binding site were both hypermethylated in the methylated cases. Regarding the CTCF2 region found downstream of the CTG repeat, eleven CTG

sites were studied and fifty percent of the CTCF1 hypermethylated cases also showed hypermethylation in the CTCF2 region (Figure 1C and Table S6). No CTCF2 hypermethylation was found in cases that were not hypermethylated in CTCF1. CpG 5 resides in the CTCF2 binding site and it was hypermethylated in the found hypermethylated cases. Interestingly, for one of the congenital cases (P1), we obtained another blood sample from a five-year follow-up. The hypermethylation pattern that this patient showed in the CTCF1 region was preserved after five years, whereas CTCF2 remained unmethylated (Table S5 and S6, annotated as P1 and P1-2).

### 3.3. Aberrant DNA methylation profiles of CTCF1 associated with higher disease severity in childhood cases

Since the methylation profiles showed exclusive hypermethylation in the CTCF1 region of the developmental cases, we reviewed their clinical phenotypes in-depth to see whether this aberrant methylation profile is associated with a differential disease severity. The congenital cases all showed hypermethylation and no clinical phenotype distinction based on DNA methylation status can be made (detailed clinical information in Table S7). Our focus was therefore on the childhood and juvenile cases. In the group of childhood-onset DM1, three out of six patients showed hypermethylation (detailed clinical information in Table 2). One case with hypermethylation was female; all the other childhood cases were male. Age of onset was on average 6.83 years old, but the age at sampling was a few decades later with a mean of 41.83 years.

At the time of revision, all patients showed muscle weakness and myotonia, but the muscular symptoms seemed to be more significant in the methylated group. Two out of three of the hypermethylated patients experienced cramps and myalgia, while just one non-methylated patient suffered this symptom. All six patients had the characteristic facial dysmorphism and ptosis. The ptosis seemed to be more pronounced in the methylated patients: two of the patients had a severe ptosis covering the pupil and the last one had a

**Table 1.** Clinical characteristics cohort

Clinical subtype	n Patients	Age of onset (years)	Age at sampling (years)	Inheritance maternal	Gender (male)	ePAL (CTGs)	Myotonia	Biceps MRC scale	MIRS	Cardiac involvement	NVM	Cataracts	mRS
<i>Congenital</i>	6	At birth	12.83 ± 5.43	(6/6)	(4/6)	610 (222-1011)	(2/6)	4.25 (3-5)	3.80 (3-5)	(1/6)	(2/6)	(0/6)	3.60 (2-5)
<i>Childhood</i>	6	6.83 ± 2.99	41.83 ± 12.04	(2/6)	(5/6)	549 (296-796)	(6/6)	4.17 (3-5)	3.33 (2-4)	(4/6)	(4/6)	(3/6)	3.33 (2-4)
<i>Juvenile</i>	23	15.05 ± 2.50	27.09 ± 12.54	(7/23)	(11/22)	317 (189-642)	(23/23)	4.87 (4-5)	2.30 (1-4)	(6/23)	(6/23)	(3/23)	1.48 (1-3)
<i>Adult</i>	22	31.50 ± 4.50	49.00 ± 9.33	(4/13)	(6/22)	290 (115-628)	(21/21)	4.86 (4-5)	2.57 (1-4)	(7/21)	(7/21)	(8/21)	1.62 (0-4)
<i>Late Onset</i>	6	52.17 ± 7.99	61.33 ± 10.23	(0/4)	(5/6)	332 (131-911)	(3/5)	5.00 (5-5)	2.40 (1-4)	(5/5)	(1/5)	(5/5)	2.00 (0-4)
<i>Asymptomatic</i>	2	N/A	30.00 ± 22.63	(0/2)	(2/2)	238 <sup>a</sup>	(0/2)	5.00 (5-5)	1 (1-1)	(0/2)	(0/2)	(0/2)	0 (0-0)

<sup>a</sup> CTG size of only one of the two patients available. Age of onset and sampling is given as mean ±SD; MRC scale, MIRS scale, mRS scale and ePAL are given as mean (range); Abbreviations: MRC = Medical Research Council; MIRS = Muscular Impairment Rating Scale; NVM= nocturnal mechanical ventilation; mRS = modified Rankin Scale; ePAL= estimated progenitor CTG size. N/A = not applicable.

**Table 2.** Clinical characteristics of childhood cases

Patient	Methyl CTGF1/2	Age at onset (years)	Age at sampling (years)	Gender	ePAL (CTGs)	Myotonia	Facial weakness	Axial weakness	Limb weakness	MIRS	Cardiac involvement	NVM	CNS involvement	Hyper-somnolence	Cataracts	mRS	DM1-ACTIV
P7	yes/no	6	56	male	756	yes	moderate	moderate	severe distal	4	pacemaker	yes	learning disability	yes	yes	4	9
P8	no/no	7	43	male	296	yes	mild	mild	mild distal	3	no	yes	no	no	no	2	25
P9	no/no	2	20	male	476	yes	moderate	mild	mild distal	2	1°AV block	no	moderate learning disability	no	no	4	22
P10	yes/no	6	40	female	324	yes	severe	moderate	mild distal	3	no	no	severe cognitive delay	yes	yes	3	14
P11	no/no	10	44	male	644	yes	mild	moderate	mild proximal + distal	4	1°AV block	yes	no	no	yes	3	38
P12	yes/no	10	48	male	796	yes	moderate	moderate	severe proximal + distal	4	pacemaker	yes	severe cognitive delay	yes	no	4	14

Abbreviations: Methyl CTGF1/2= hypermethylated at the CTGF1 and CTGF2 site; ePAL= estimated progenitor allele size; AV = atrioventricular; MIRS = Muscular Impairment Rating Scale mRS = modified Rankin Scale; NVM = nocturnal mechanical ventilation; CNS= central nervous system

moderate ptosis covering part of the pupil. In the non-methylated group, just one patient had a moderate ptosis, while the other two had a mild ptosis. Facial weakness was also present in all childhood-onset-DM1 patients, it being severe in 1/3 and moderate in 2/3 of the methylated patients; whereas the non-methylated group showed only mild (two patients) and moderate facial weakness (one patient).

Dysarthria was present in all patients, but to a higher degree in methylated patients, where it ranged from moderate to severe, compared to mild to moderate in non-methylated patients. Axial and limb weakness was also more pronounced in the methylated versus the non-methylated group. In the hypermethylated cases, 2/3 had a severe limb weakness (MRC scale of 1-2): one of the patients had a proximal and distal weakness, requiring a wheelchair and the other patient showed a proximal and distal weakness pattern needing just a cane for walking. The last one of the methylated childhood cases had a mild distal weakness (MRC scale 3-4). In the non-methylated group, all of the patients showed a mild weakness (MRC of 3-4), two of the patients with a distal pattern and one patient with proximal/distal weakness.

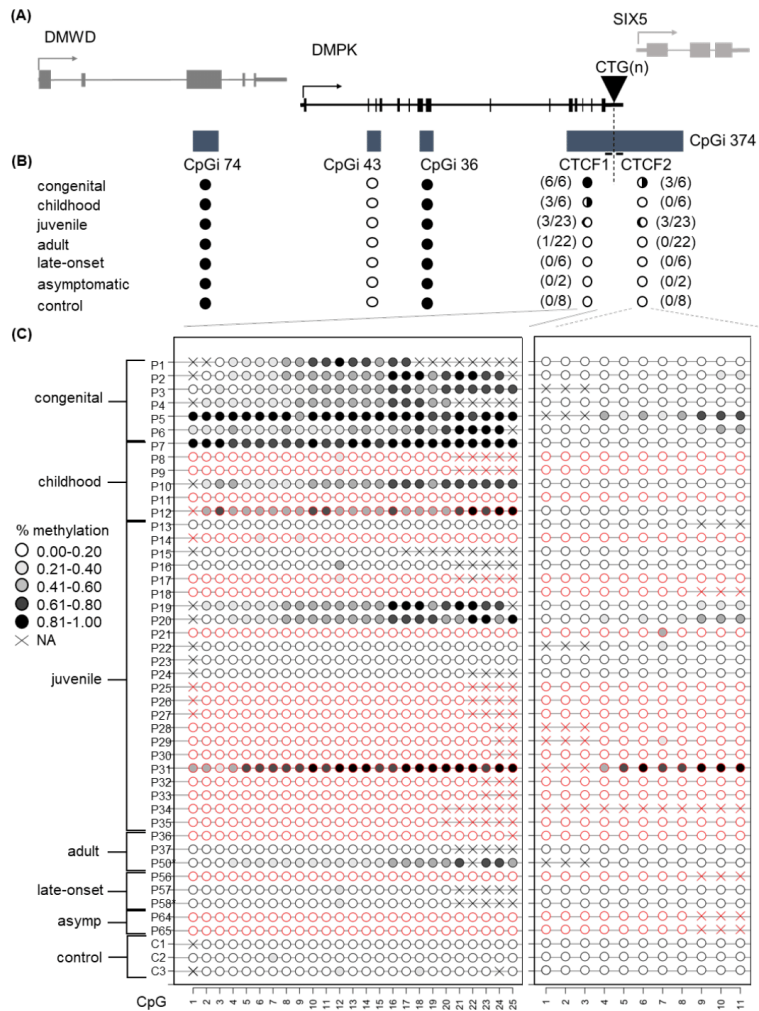
Cardiac manifestations were more severe in the methylated group, as 2/3 patients had a pacemaker, whereas in the case of the non-methylated group only mild changes in electrocardiogram were seen, such as a mild first grade AV block in 2/3 cases. Two of the methylated patients and one of the non-methylated patients used ventilatory support. Cognitive manifestations were noticed in all methylated childhood DM1 patients: two of the patients had severe cognitive delay, while moderate learning difficulty was observed in the third patient. In the non-methylated group, just one patient had learning difficulties, while the others showed no mental disabilities. All methylated patients experienced hypersomnolence and none of the non-methylated patients suffered this symptom. Intestinal rhythm dysfunction was found exclusively in the methylated patients, while alopecia was found in the non-methylated patients only. At the time of revision, all patients had some degree of dependence in the mRS. In the methylated group 2/3 needed help in the basic activities of daily life but did not require

continuous supervision (mRS 4), while the other one required help for instrumental activities (mRS 3). The average score in the DM1-ACTIV scale was 12.33. In the non-methylated group, we found one patient with mRS 4, another one with mRS 3 and the last one had a milder dependence (mRS 2) and the average score in the DM1-ACTIV scale was 28.33, which means they were better at performing daily and social activities. Considering all this data as a whole, there seemed to be a more severe muscular, cardiac and cognitive manifestation of the disease in the methylated childhood cases.

Upon revision of the juvenile cases, no such differences in muscular, cardiac and cognitive manifestation could be found between the methylated (n=3) and non-methylated group (n=20). Although, this might be due to the low number of cases, and analysis of a larger cohort of this DM1 subcategory would be needed to address the impact of CTCF1 hypermethylation on this clinical subtype (detailed information in table S8).

#### *3.4. A higher chance of hypermethylation in CTCF1 with increasing CTG expansion size*

As mentioned before, hypermethylation of the CTCF1 region is almost exclusively found in the developmental subtypes. These subtypes are associated with a higher disease severity and overall greater CTG expansion sizes. To see whether methylation status in DM1 patients is associated with the CTG expansion size, we performed a logistic regression on the entire cohort using the ePAL and modal allele. This revealed a positive association between ePAL and methylation status of the CTCF1 region in DM1 patients (Table S9, model 1,  $p=0.004$ ) and between the modal allele and the methylations status (Table S9, model 2,  $p=0.001$ ), which suggests that the larger the ePAL/modal, the more likely hypermethylation at the CTCF1 regions is going to occur.



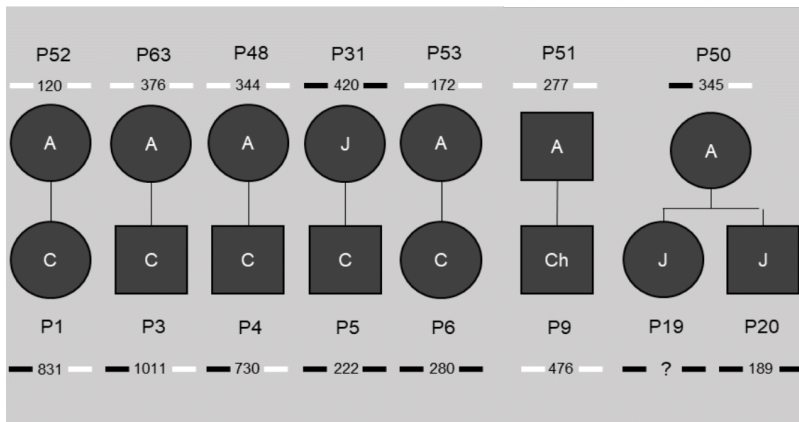
**Figure 1. DNA methylation profile at the DMPK locus in DM1 blood samples representing all clinical subtypes.** A) Schematic representation of the four CpGis residing in the DMPK locus and neighboring genes. CpGi 374 is divided into the CTCF1 and CTCF2 region, harboring the CTCF binding sites and encompassing the CTG expansion. B) Summary of the methylation profiles of the five CpG regions across the DMPK locus in the studied clinical subtypes, in which black indicates the degree of hypermethylation. Congenital n= 6, Childhood n= 6, Juvenile n= 23, Adult n= 22, Late onset n= 6, asymptomatic n= 2, Controls n= 8. C) Detailed DNA methylation profiles of the clinical subtypes in the CTCF1 and CTCF 2 region. Each circle represents a CpG dinucleotide. The colour gradient represents the level of methylation indicated in the legend assessed by sodium bisulphite sequencing. Red indicates paternal inheritance. Black indicates maternal inheritance. \* unknown inheritance. For the non-developmental cases, only a representative subset of three samples has been displayed. Detailed methylation profiles of all patients in all categories can be found in the supplemental tables. Abbreviations: DMWD= Dystrophia myotonica WD repeat-containing gene, DMPK= myotonic dystrophy protein kinase gene, SIX5= six homeobox 5 gene, CTG(n)= the CTG expansion, CpGi= CpG island, Asymp= asymptomatic, C= control, P= DM1 patient

### 3.5. Increased, but not exclusive, maternal transmission in CTCF1 hypermethylated cases

It has been reported that the hypermethylation observed in DM1 cases is associated with maternal transmission. We therefore decided to evaluate, where possible, the transmission in this DM1 cohort (Figure 1C). We found that all congenital cases, which were all hypermethylated, were maternally transmitted. In both childhood and juvenile cases, two out of the three hypermethylated cases in each category were maternally transmitted. For the childhood subcategory as a whole, a total of two maternal transmissions were registered and one paternal; meaning that all maternally transmitted cases reside in the methylated group. However, for the juvenile subcategory, a total of seven patients were maternally transmitted, of which only two reside in the methylated group. Taken together, we could corroborate a higher chance of maternal transmission when hypermethylated, but notably there were patients that were methylated and paternally transmitted and patients that were maternally transmitted and unmethylated.

### 3.6. Methylation status is not inheritable and associated with the transmission of CTG repeat contractions

This study cohort included seven families, giving us the opportunity to study the inheritance of the differential DNA methylation profiles shown in CTCF1 and CTCF2 (pedigrees in Figure 2). Of five out of six congenital cases, the mother was included in the study cohort as well. Only one of the mothers showed hypermethylation at both the CTCF1 and CTCF2 region, whereas the other four were unmethylated. This hypermethylated mother belonged to the juvenile subtype. For the childhood subset, only one family could be studied, where both the childhood case and the adult-onset father were unmethylated. For the juvenile subset, one family consisting of two siblings with both juvenile onsets could be studied. These two siblings were both methylated on the CTCF1 and CTCF2 regions and interestingly the mother was the only adult hypermethylated case in our cohort, showing only CTCF1 hypermethylation. Of note, when reviewing the CTG expansion sizes, the unmethylated mothers have passed on an expansion of the CTG repeat, whereas the methylated mothers have transmitted a contraction of the CTG repeat.



**Figure 2. Pedigrees of the known families in our study cohort.** The number beneath the patient identification code indicates the ePAL (estimated progenitor allele size), with the bars next to it indicated methylation status of CTCF1 (left) and CTCF2 (right). Black indicates hypermethylation, white no methylation. A= adult, J= juvenile, Ch= childhood, C= congenital. ? = unknown CTG expansion size.

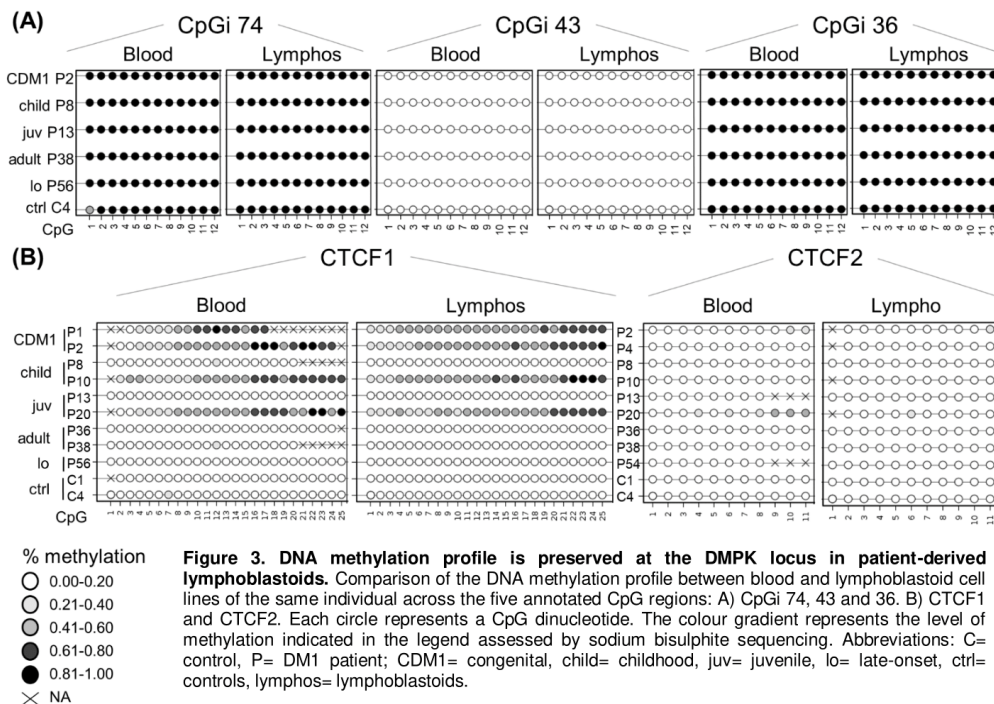
### 3.7. DNA methylation profiles are preserved in blood-derived cells

To see whether lymphoblastoids preserve the epigenetic landscape and can be used as a faithful DM1 cellular model, we decided to study whether the DNA methylation profiles are similar between blood and the blood-derived lymphoblastoids (Figure 3). CpGi 74, CpGi 43 and CpGi 36 showed no differences between blood and lymphoblastoids across all clinical subtypes (Figure 3A and Tables S10-12). Both CpGi 74 and CpGi 36 remained completely hypermethylated, whereas CpGi 43 was totally unmethylated. For CTCF1, the pattern observed in blood, a gradient of hypermethylation in the developmental cases, was preserved in all the studied lymphoblastoids (Figure 3B), except for P7 which showed hypermethylation in blood, but not in lymphoblastoid cells in CTCF1 (Table S13). Regarding CTCF2, the patients that

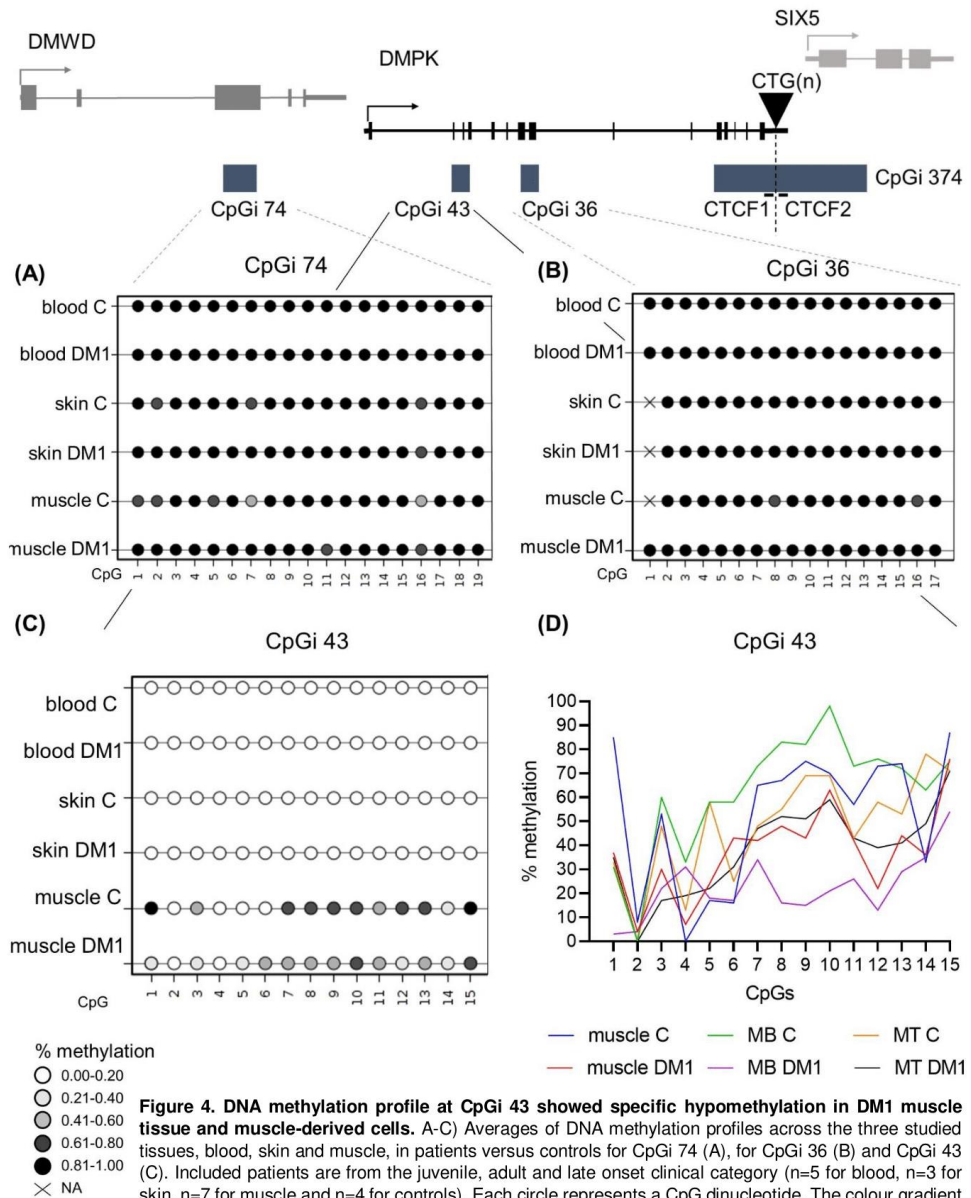
displayed hypermethylation of the CTCF2 region in blood, preserved their methylation status in the studied lymphoblastoids (Figure 3B and Table S14). However, two patients that were not hypermethylated in blood for CTCF2 (P4 and P10) showed low-grade hypermethylation in lymphoblastoids (~10% methylation) (Table S14).

### 3.8. DM1 is associated with hypomethylation of CpGi 43 in muscle tissue and muscle-derived cells

Next, we wanted to study the DNA methylation profiles of the five regions in tissues other than blood to address whether tissue-specific epigenetics at the *DMPK* locus exist in DM1. For this, we acquired a muscle and skin biopsy from a subset of patients (juvenile, adult and late-onset origin). From these biopsies, cells were isolated to assess whether cellular models accurately reflect the origin tissue in terms of DNA methylation status. In CpGi 74







**Figure 4. DNA methylation profile at CpGi 43 showed specific hypomethylation in DM1 muscle tissue and muscle-derived cells.** A-C) Averages of DNA methylation profiles across the three studied tissues, blood, skin and muscle, in patients versus controls for CpGi 74 (A), for CpGi 36 (B) and CpGi 43 (C). Included patients are from the juvenile, adult and late onset clinical category (n=5 for blood, n=3 for skin, n=7 for muscle and n=4 for controls). Each circle represents a CpG dinucleotide. The colour gradient represents the level of methylation indicated in the legend assessed by sodium bisulphite sequencing. D) Representation of the DNA methylation profiles of muscle and muscle-derived cells (myoblasts and myotubes) in CpGi 43 assessed by bisulphite sequencing.

and CpG 36, DNA hypermethylation was found in skin and muscle of both DM1 patients and controls, similar to what was found in blood (Figure 4A-B). Additionally, the tissue-derived cells of both muscle and skin preserved the observed hypermethylation (Tables S15-18). Regarding CpG 43, skin samples showed no methylation in the DM1 patients and controls, similar to the observations in blood (Figure 4C, Table S19). However, skin-derived fibroblasts, from both controls and patients, showed a slight hypermethylation with no differences between the two groups (Table S19). Interestingly, a distinct DNA methylation pattern was observed in muscle tissue, with an overall hypermethylation in control samples, which was much lower in DM1 muscles, showing reductions as high as 70% (Figure 4C-D and Table S20). This muscle-specific DNA methylation profile was preserved in both control- and patient-derived myoblasts and myotubes (Figure 4D and Table S20). Interestingly, DM1 myoblasts or myogenic precursor cells showed the biggest decrease of DNA methylation compared to controls.

### 3.9. DM1 is associated with hypermethylation in CTCF1 in muscle tissue and muscle derived cells

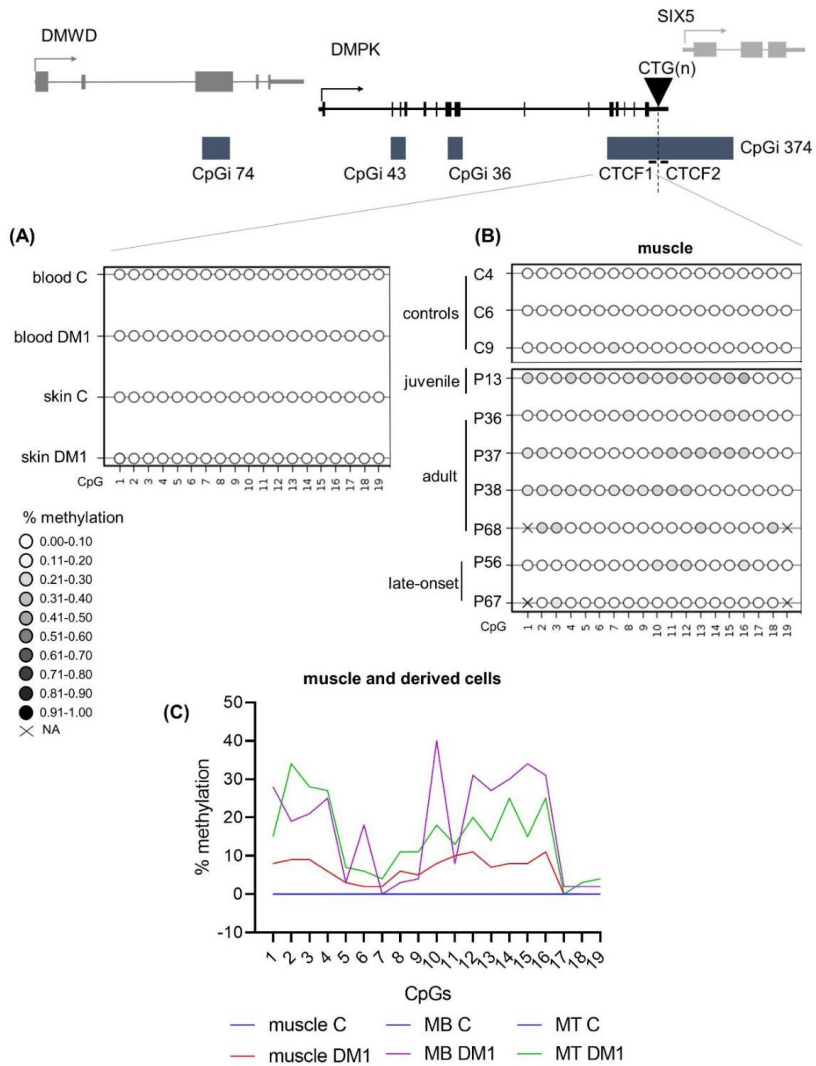
The CTCF1 region showed similar DNA methylation profiles in skin compared to blood (Figure 5A and Table S21). However, skin fibroblasts showed low-grade hypermethylation (average of around 10%) in DM1 patients, but none in controls (Table S21). Due to difficulties with the sequencing analysis, only in one patient could both the skin and the skin-derived fibroblasts be analyzed. Not to mention that the other analyzed DM1 skin fibroblasts were not derived from the analyzed skin biopsies; therefore, we cannot rule out the possibility that these samples already show a partial methylation, especially in the case of the juvenile sample. Interestingly, six out of seven muscle biopsies showed hypermethylation (Figure 5B and Table S22); with the highest average found in the youngest biopsy, belonging to a juvenile case. Of these biopsies, we have the CTG expansion size available in muscle and blood (Table S23). However, CTG expansion size could not be linked to the degree of methylation. Surprisingly, the two CpGs residing in the CTCF1 binding site (CpG 18 and 19) were not methylated in these biopsies, with the exception of

CpG 18 in the biopsy of P68. Total unmethylation in all CpGs was observed in the control biopsies. This methylation profile was maintained in the muscle-derived cells, although DM1 patient-derived myoblasts and myotubes showed, when compared to the muscle, slightly higher levels of methylation (Figure 5C and Table S22). This may be due to the purity of the myogenic cultures versus the analysis of whole muscle tissue, containing over 10 distinct cell types. Control muscle cells showed no methylation in any CpG (Figure 5C and Table S22). Of note, the differences in myogenic precursor cells between patients and controls were the highest (Figure 5C). The last analyzed region, CTCF2, showed no methylation in the tissues and tissue-derived cells studied from the patients of which we have the biopsies (Table S24 and S25), suggesting that muscle-specific hypermethylation only happens in the CTCF1 region.

## 4. DISCUSSION

The overall goal of this study was to investigate the contribution of epigenetics to DM1 pathology, by analyzing for the first time the DNA methylation profiles across the four CpG islands residing in the *DMPK* locus in several tissues and tissue-derived cells in all clinical subtypes of DM1. Our results showed a distinct DNA methylation profile across DM1 tissues and uncovered a novel and dual epigenetic signature in DM1 muscle samples, involving a gain of DNA methylation in the flanking region of the CTG expansion accompanied by specific DNA demethylation in the *DMPK* gene body (Figure 6).

In blood samples, we found a hypermethylation across all clinical subtypes for CpG 74 and 36, while CpG 43 remained completely unmethylated, with no differences observed between patients and controls. Previously, it was reported that in leukocytes derived from control individuals, CpG 36 and 74 (located in the *DMPK* and *DMWD* gene bodies, respectively) were highly methylated, meanwhile CpG 43 (located in *DMPK* gene body and overlapping with a proposed alternative promoter) was unmethylated<sup>45</sup>. Our results showed that in DM1 blood samples, the DNA methylation profile of CpG

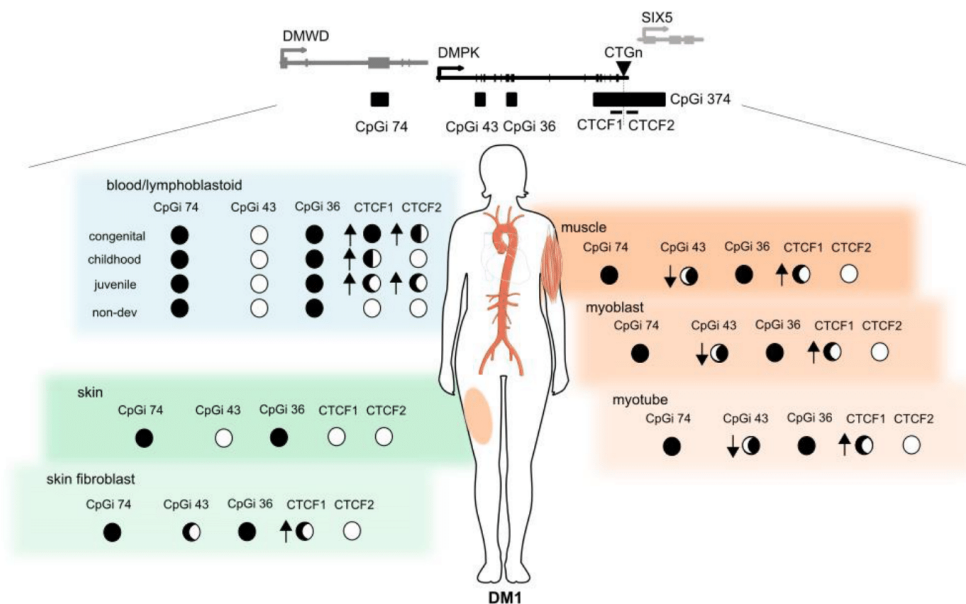


**Figure 5. DNA methylation profiles at the CTCF1 regions were tissue-specific and increased in DM1 muscle samples.** A) Averages of DNA methylation profiles across blood and skin in patients versus controls for CTCF1. Included patients are from the juvenile, adult and late onset clinical category (n=5 for blood, n=3 for skin and n=4 for controls). B) Overview of the muscle biopsy DNA methylation profiles of DM1 patients and controls. Each circle represents a CpG dinucleotide. The colour gradient represents the level of methylation indicated in the legend assessed by sodium bisulphite sequencing. C) Representation of the DNA methylation profiles of muscle and muscle-derived cells (myoblasts and myotubes) in CTCF1. For CTCF1 all control muscle cells are at zero and have been given the same colour to aid visualization.

36, 43 and 74 does not change in the distinct DM1 subtypes. Conversely, the CTCF1 and CTCF2 regions did show a change in DNA methylation levels in DM1 compared to controls. Developmental cases showed an upward gradient of hypermethylation with increasing severity of the disease and decreasing age of onset. No methylation was observed in the non-developmental cases, except for a single adult case. Additionally, fifty percent of the CTCF1 hypermethylated cases also showed hypermethylation in the CTCF2 region, but interestingly, CTCF2 methylation without CTCF1 hypermethylation was not observed in any case. That may indicate that the beginning of the aberrant DNA methylation is not random and it spreads beyond the CTG repeat only in certain cases/conditions. The mechanism behind the hypermethylation observed in the CTCF1 region and why the mechanism seems to be biased towards developmental cases are still

relevant questions that need addressing in the DM1 field.

The current studies on DNA methylation profiles in the two CTCF binding regions are controversial. Several studies have shown a similar trend in the levels of CTCF1 methylation as we observed in our developmental cases<sup>34,37,41,42</sup>. However, their observations were not as clear of a trend as ours were. For example, Barbé and collaborators have seen the majority of CTCF1 hypermethylation in CDM1 (95% of cases) and to a lesser degree in childhood DM1, with only two out of seven childhood cases and no juvenile cases<sup>34</sup>. Santoro and collaborators reported a very similar gradient in congenital and childhood cases, where all congenital cases and about fifty percent of the childhood cases were methylated<sup>42</sup>. Unfortunately, in this study no distinction between adult and juvenile was made, so it is unclear to which subtype the six



**Figure 6. Summary of the epigenetic landscape of the *DMPK* locus in DM1.** An overview is given of the methylation status across the *DMPK* locus in muscle, skin and blood and their derived-cells in DM1 patients. The circle underneath each CpG island represents its methylation status, ranging from fully methylated (black) to completely unmethylated (white) in DM1 patients. The arrows indicate the changes compared to controls, meaning an upward arrow means increased methylation in DM1 compared to controls. This figure was partially made using Servier Medical Art (smart.servier.com).

methylated cases in this hybrid group belonged to. The CTCF2 region is showing similar controversy, with several studies showing a similar trend as we have observed. For example, Morales and collaborators showed CTCF2 methylation in the developmental cases, but in less patients compared to CTCF1<sup>41</sup> and Barbé and collaborators found CTCF2 hypermethylation primarily in CDM1 patients and only in one out of six non-DM1 patients<sup>34</sup>. However, other studies have shown no methylation of this region in any clinical subtype, including Santoro and collaborators, which showed a similar methylation gradient to ours in developmental cases, but found the CTCF2 region to be completely unmethylated<sup>42</sup>. Several studies have focused solely on adult cases, and interestingly have found hypermethylation among these non-developmental cases, contradictory to our findings<sup>38,39,49</sup>. For example, Hildonen and collaborators found more than a 10 percent increase in methylation in CTCF2 in 9/68 studied DM1 patients<sup>39</sup> and Legaré and collaborators found a 3-5% increase in methylation at a few CpG sites downstream of the CTG repeat in adult cases ( $\geq 18$  years)<sup>49</sup>. These differences found between the published works could be due to the different techniques used to assess DNA methylation levels, the different criteria used to decide what is considered methylated and the difficulties to establish the age of onset in DM1. To better identify the DNA methylation differences in DM1 samples, a general consensus in DM1 DNA methylation studies would be needed. In our cohort we could see whether the methylation status changes over time, since we had the five-year follow-up of one of the congenital cases, where we found that the methylation status was stable, which is in concordance with previous studies<sup>38-40</sup>.

Few studies have addressed the link between clinical phenotype and DNA methylation profiles. Due to the extensive clinical data obtained from this cohort, we had the opportunity to assess whether DNA methylation status was associated with clinical phenotype. Methylated childhood cases showed a more severe muscular, cardiac and cognitive manifestations of the disease. This type of phenotypic association in this particular disease subtype has not been made previously and the few available studies on clinical phenotype correlations are from adults. Légaré and collaborators have

shown that methylation status is linked to muscular and respiratory profiles in adults<sup>44</sup> and Breton found a correlation between hypermethylation and a decline in cognitive function<sup>38</sup>. Santoro and collaborators are the only ones that studied the association between DNA methylation and MIRS scale across the clinical subtypes and could not find an association<sup>42</sup>. However, it has been stipulated that hypermethylation seems to be associated with the more severe forms of the disease, as hypermethylation is found overwhelmingly in CDM1 cases<sup>34,41-43</sup>. Our study supports this notion and adds the novel finding that it is also linked to the more severe cases of childhood DM1. Caution must be taken, however, as our sample size was quite low and further studies are needed to confirm this association.

Disease severity and age of onset have been previously linked to the CTG expansion size, where the more severe disease forms and earlier age of onset are correlating with the greater CTG expansion sizes<sup>13,14</sup>. Since we have observed a gradient in hypermethylation in our developmental cases, similar to the CTG expansion size gradient, we decided to analyse whether methylation status is also associated to the CTG expansion size or whether these two modifiers work independently. We found a positive association between two CTG size predictors and methylation status of CTCF1. This suggests that the larger the CTG expansion size, the higher the likelihood of CTCF1 hypermethylation. This association has been found previously by several studies, both for ePAL<sup>41</sup> and modal allele<sup>39,42,44</sup>. However, not all studies were able to find such association between CTG expansion size and the methylation status<sup>34,37</sup>. This discrepancy might be due to the age at which the CTG expansion size was analysed, the used CTG expansion measure (ePAL vs. modal) or the sizing technique used. For example, somatic instability might be another disease modifier, which increases over time. Using the modal allele might make establishing correlations more challenging<sup>50,51</sup>. Moreover, it has been shown that different techniques to size the CTG expansion result in different sizes, hindering correlations<sup>52</sup>. Our study cohort included several families, giving us the opportunity to study inheritance of the DNA methylation profiles. We found the DNA methylation status of our patients to be not inheritable, as several

unmethylated mothers gave birth to methylated cases. This is in accordance with previous studies<sup>34,41</sup>. Interestingly, we observed that the offspring of methylated mothers carried contractions of the CTG expansion, while the offspring of unmethylated mothers carried expansions of the repeat. This suggests that although the bigger CTG expansion sizes are associated with hypermethylation, when a methylated parent passes on the methylation status, it coincides with the transference of a smaller CTG expansion. Some authors have evaluated the effect of DNA methylation on the stability of the CTG expansion repeat<sup>53</sup>. When using bacterial and primate cellular models of 83 to 100 CTG repeat expansions, DNA methylation was found to be associated with a stabilization of the repeat size. However, caution must be taken with these observations, as the sample size in their study and our study was low and further studies are needed to further elucidate this observation.

CDM1 cases have been linked to maternal transmission, as about ninety percent of CDM1 cases are maternally transmitted, while non-developmental DM1 is about thirty percent maternally transmitted<sup>3</sup>. This brought up the question whether methylation status can also be linked to maternal transmission and indeed we found an increased maternal transmission rate in CTCF1 hypermethylated cases. However, this was not absolute. We have found several cases of hypermethylation that were paternally transmitted and not all maternal transmission resulted in hypermethylation. Barbé and collaborators have suggested the presence of a parent-of-origin effect, where DNA methylation may account for the maternal bias for CDM1 transmission, the larger maternal CTG expansions, age of onset and clinical phenotype<sup>34</sup>. The hypothesis is based on the potential reduced survival of spermatozoa due to the hypermethylation of the CTCF1 region, disrupting the insulator element and decreasing levels of *SIX5*, which is essential for spermatozoa survival. This was supported by Morales and collaborators and they additionally showed that independent of clinical subtype, the parental transmission and methylation status were associated with almost exclusive maternal transmission<sup>41</sup>. Although we do see a similar trend, as all our CDM1 cases are both maternally transmitted and hypermethylated, parental

inheritance is not a good predictor for methylation status or vice versa for the other clinical phenotypes. This is strengthened by the observation made by Morales and collaborators, where a large family showed several paternally transmitted methylated cases<sup>41</sup>. It therefore may be a good diagnostic indicator during prenatal screening, but less efficient as a general disease marker. Of note, a recent study in DM1 spermatozoa found that methylation was not affecting sperm viability and these spermatozoa were compatible with “*in vitro*” fertilization<sup>54</sup>. These findings go against the hypothesis that reduced survival is associated with methylated spermatozoa, preventing the transmission of CDM1<sup>3</sup>, and therefore, other explanations for this maternal bias should be explored.

All the above observations were made in blood and blood-derived cells. Next, we investigated whether tissue-specific epigenetics exists at the *DMPK* locus in DM1. We found no differences between blood and skin/skin-derived cells for the majority of the CpG regions studied. Exceptions were CpGi 43 in which skin fibroblasts showed slight hypermethylation in both DM1 patients and controls, and CTCF1, which showed a small gain of methylation exclusively in the DM1 skin fibroblasts. These results are in agreement with the findings of Buckley and collaborators<sup>45</sup>, where they showed a similar DNA methylation pattern for all regions, except CTCF2, which was not reported in control skin and skin fibroblast<sup>45</sup>. Interestingly, our results uncovered a novel and specific DNA methylation signature in DM1 muscles and muscle-derived cells for CpGi 43 and the CTCF1 region. CpGi 43 showed an almost complete absence of methylation in blood, skin and their derived cells, but a clear hypermethylation of non-affected muscle and muscle-derived cells. This is in line with previous findings by Buckley and collaborators in control samples<sup>45</sup>. However, in DM1 patients our results showed much lower methylation levels compared to muscle and muscle-derived cells from unaffected individuals. Furthermore, CpGi 43 demethylation in DM1 muscles and myogenic cells was accompanied by hypermethylation in the CTCF1, but not CTCF2, region suggesting that this specific epigenetic landscape could alter gene expression in muscle samples.

Buckley and collaborators extensively studied the epigenetics in the *DMPK* locus in several control tissues and cell types. Interestingly, they reported the existence of a *DMPK* alternative promoter (also referred to as the downstream promoter, which overlaps with CpGi 43), as well as cell type-dependent differences in promoter usage according to epigenetic features<sup>45</sup>. The use of the *DMPK* canonical promoter was associated with high levels of *DMPK* expression (in muscle cells and to a lower extent in skin fibroblasts and osteoclasts), meanwhile the use of the alternative promoter was associated with low levels of expression (in leukocytes and embryonic stem cells, among others). Importantly, cell type-specific DNA methylation patterns, together with transcription factor binding (MyoD, CTCF, ZNF143) would regulate the use of these promoters, modulating *DMPK* expression. It is well known that although *DMPK* is expressed in several tissues, skeletal muscle and heart display higher expression levels, and are the most affected tissues in DM1 pathology. The data presented by Buckley and collaborators showed a predominant use of the canonical upstream promoter in skeletal muscle and myogenic cells from control individuals, by RNA-CAGE (cap analysis gene expression) data, supported by complete hypomethylation of this promoter and hypermethylation of the alternative downstream promoter<sup>45</sup>. Additionally, they reported that MyoD and CTCF binding and the open chromatin marks H3K4me3 and H3K27ac were hardly detected in the alternative promoter region, further indicating that the use of this alternative promoter was silenced in skeletal muscle and myogenic cells. Conversely, in blood, a predominant usage of the alternative *DMPK* promoter (in CpGi 43) was suggested, by showing its hypomethylation together with hypermethylation of the canonical promoter, strong binding of CTCF and high levels of H3K4me3 in control leukocytes<sup>45</sup>. Interestingly, our study of DM1 muscle samples revealed a novel epigenetic change by specific demethylation of this alternative promoter located at CpGi 43, in skeletal muscle tissue and muscle-derived cells. This could potentially alter chromatin conformation and result in a shift of the promoter usage from the strongest/canonical one to the weak/alternative promoter, decreasing *DMPK* expression levels in DM1 myogenic samples.

Additionally, this DM1 tissue-dependent demethylation was accompanied by a gain of methylation in the CTCF1, but not the CTCF2 region. Previous studies showed that CTCF, a transcription factor that can function as an insulator, binds strongly to CTCF1, but not CTCF2, in a methylation-dependent manner<sup>33</sup>. Importantly, the hypermethylation of the CTCF regions could inhibit CTCF-binding and disrupt the insulator element formed by the CTG expansion and the two CTCF-binding sites, affecting *DMPK* and *SIX5* expression. The loss of the insulator activity by DNA methylation would allow the interaction of the *SIX5* enhancer with the *DMPK* promoter, increasing *DMPK* expression, meanwhile reducing *SIX5* expression<sup>33</sup>. Notably, our results showed that in blood samples and derived-leukocytes, the CpGs located inside the CTCF1 and CTCF2 binding sites were hypermethylated in all hypermethylated cases (almost exclusively developmental cases), meanwhile in muscle samples remained unmethylated (almost exclusively non-developmental cases). This may imply that although there is a disease-specific gain of methylation for the CTCF1 region in muscle, the CTCF binding site is not disrupted, allowing CTCF binding. However, this hypermethylation found might affect other chromatin interactions, in turn affecting gene expression. The analysis of publicly available ChIP-seq data of histone post translational modifications (H3K4me3, H3K4me1 and H3K27ac), done by Buckley and collaborators at *DMPK* and neighboring genes, showed that CTCF regions and intragenic regions of *DMWD* and radial spoke head 6 homolog A genes (*RSPH6A*), located next to *DMPK* gene, displays enhancer chromatin features in control muscle cells<sup>45</sup>. This is interesting since *DMPK* lies in the middle of a chromosomal domain with three genes preferentially expressed in testis, indicating that its expression, mainly in skeletal muscle and heart, has to be tightly regulated. Additionally, Brouwer and collaborators showed an increase of the H3K9me3 chromatin repressive mark, together with gain of DNA methylation, in the CTCF1 region (and to a lesser extent in CTCF2) in DM1 heart mice, which correlates with decreased *DMPK* expression<sup>36</sup>. To address whether DNA methylation changes in the CTCF1 region in DM1 skeletal muscles may alter chromatin interactions between *DMPK* promoter and these

potential myogenic enhancers, further experiments are needed.

Finally, this study addressed for first time the DNA methylation status of patients-derived DM1 cells. The availability of several DM1 tissues and their corresponding tissue-derived cells give us the opportunity to determine that most cellular models maintained the DNA methylation state observed in the original tissue. However, in some cases we observed a slightly higher increase in methylation levels in cells (e.g., some DM1 skin fibroblasts and myoblast/myotubes) versus the corresponding tissues. This can be explained by the observation that cellular models, especially immortalized cell lines or primary cell cultures that have been in culture for a substantial amount of time, can increase DNA methylation levels<sup>25,55,56</sup>, and/or because of the purity of cell cultures compared to tissues containing distinct cell types. Overall, our results showed that the DM1 patient-derived cells preserve the genetic and epigenetic features, which make them excellent models to study DM1 pathology.

In conclusion, our results showed a distinct DNA methylation profile across DM1 tissues and uncovered a novel, dual epigenetic signature involving a gain of DNA methylation in the flanking region of the CTG expansion, accompanied by specific DNA demethylation in the *DMPK* gene body of DM1 muscle samples, which highlighted the contribution of epigenetic changes to DM1 pathology.

#### SUPPLEMENTAL DATA

Supplemental data include one figure and 25 tables

#### DECLARATION OF INTEREST

The authors declare that they have no competing interest.

#### ACKNOWLEDGMENTS

The authors wish to thank the DM1 patients for providing the samples needed to perform this study.

We also thank the IGTP core facilities for their contribution to this publication.

This work was supported by Instituto de Salud Carlos III (grant number PI18/00713 to G. Nogales-Gadea and A. Ramos-Fransi), Trampoline Grant (#21108) from AFM-Telethon to G. Nogales-Gadea, Ministerio de Ciencia e Innovación (grant number PID2020-118730RB-I00) and Grant Project (#23557) from AFM-Telethon to M. Suelves, and co-financed by Fondos FEDER. E. Koehorst is funded by the La Caixa Foundation (ID 100010434), fellowship code LCF/BQ/IN18/11660019, cofunded by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no. 713673. J. Núñez-Manchón is funded by Instituto de Salud Carlos III I-PFIS fellowship (grant number IFI20/00022). G. Nogales-Gadea is supported by a Miguel Servet research contract (ISCIII CPII19/00021, and FEDER). This work was supported by the CERCA program/ Government of Catalonia. The funding bodies had no role in the design of the study and the collection, analysis, and interpretation of data.

#### REFERENCES

1. Harper PS (2001). Major Problems in Neurology: Myotonic Dystrophy (London, UK: WB Saunders).
2. De Antonio, M., Dogan, C., Hamroun, D., Mati, M., Zerrouki, S., Eymard, B., Katsahian, S., Bassez, G., and French Myotonic Dystrophy Clinical Network (2016). Unravelling the myotonic dystrophy type 1 clinical spectrum: A systematic registry-based study with implications for disease classification. *Rev. Neurol. (Paris)*. 172, 572–580.
3. Lanni, S., and Pearson, C.E. (2019). Molecular genetics of congenital myotonic dystrophy. *Neurobiol. Dis.* 132.
4. Campbell, C., Sherlock, R., Jacob, P., and Blayney, M. (2004). Congenital myotonic dystrophy: assisted ventilation duration and outcome. *Pediatrics* 113, 811–816.
5. Meola, G., and Cardani, R. (2015). Myotonic dystrophies: An update on clinical aspects, genetic, pathology, and molecular pathomechanisms. *Biochim. Biophys. Acta - Mol. Basis Dis.* 1852, 594–606.
6. Douniol, M., Jacqueline, A., Cohen, D., Bodeau, N., Rachidi, L., Angeard, N., Cuisset, J.-M., Vallée, L., Eymard, B., Plaza, M., et al. (2012). Psychiatric



- and cognitive phenotype of childhood myotonic dystrophy type 1. *Dev. Med. Child Neurol.* *54*, 905–911.
7. Enchev, B., and Bassez, G. (2013). Congenital and infantile myotonic dystrophy. *Handb. Clin. Neurol.* *113*, 1387–1393.
  8. Ho, G., Carey, K.A., Cardamone, M., and Farrar, M.A. (2019). Myotonic dystrophy type 1: Clinical manifestations in children and adolescents. *Arch. Dis. Child.* *104*, 48–52.
  9. Mathieu, J., Allard, P., Potvin, L., Prévost, C., and Bégin, P. (1999). A 10-year study of mortality in a cohort of patients with myotonic dystrophy. *Neurology* *52*, 1658–1662.
  10. de Die-Smulders, C.E., Höweler, C.J., Thijs, C., Mirandolle, J.F., Anten, H.B., Smeets, H.J., Chandler, K.E., and Geraedts, J.P. (1998). Age and causes of death in adult-onset myotonic dystrophy. *Brain* *155*, 1557–1563.
  11. Chen, H., and Chen, H. (2017). Myotonic Dystrophy Type 1. In *Atlas of Genetic Diagnosis and Counseling*, pp. 1999–2011.
  12. Groh, W.J., Groh, M.R., Saha, C., Kincaid, J.C., Simmons, Z., Ciafaloni, E., Pourmand, R., Otten, R.F., Bhakta, D., Nair, G. V., et al. (2008). Electrocardiographic abnormalities and sudden death in myotonic dystrophy type 1. *N. Engl. J. Med.* *358*, 2688–2697.
  13. Groh, W.J., Groh, M.R., Shen, C., Monckton, D.G., Bodkin, C.L., and Pascuzzi, R.M. (2011). Survival and CTG repeat expansion in adults with myotonic dystrophy type 1. *Muscle Nerve* *43*, 648–651.
  14. Logigian, E.L., Moxley, R.T., Blood, C.L., Barbieri, C.A., Martens, W.B., Wiegner, A.W., Thornton, C.A., and Moxley, R.T. (2004). Leukocyte CTG repeat length correlates with severity of myotonia in myotonic dystrophy type 1. *Neurology* *62*, 1081–1089.
  15. Yum, K., Wang, E.T., and Kalsotra, A. (2017). Myotonic Dystrophy: Disease Repeat Range, Penetrance, Age of Onset, and Relationship Between Repeat Size and Phenotypes. *Curr. Opin. Genet. Dev.* *44*, 30.
  16. Novelli, G., Gennarelli, M., Menegazzo, E., Angelini, C., and Dallapiccola, B. (1995). Discordant clinical outcome in myotonic dystrophy relatives showing (CTG)<sub>n</sub> > 700 repeats. *Neuromuscul. Disord.* *5*, 157–159.
  17. Campbell, C., Levin, S., Siu, V.M., Venance, S., and Jacob, P. (2013). Congenital myotonic dystrophy: Canadian population-based surveillance study. *J. Pediatr.* *163*.
  18. Gudde, A.E.E.G., van Heeringen, S.J., de Oude, A.I., van Kessel, I.D.G., Estabrook, J., Wang, E.T., Wieringa, B., and Wansink, D.G. (2017). Antisense transcription of the myotonic dystrophy locus yields low-abundant RNAs with and without (CAG)<sub>n</sub> repeat. *RNA Biol.* *14*, 1374–1388.
  19. Koehorst, E., Núñez-manchón, J., Ballester-lópez, A., Almendrote, M., Lucente, G., Arbex, A., Chojnacki, J., Vázquez-manrique, R.P., Gómez-escribano, A.P., Pintos-morell, G., et al. (2021). Characterization of RAN Translation and Antisense Transcription in Primary Cell Cultures of Patients with Myotonic Dystrophy Type 1. *J. Clin. Med.* *10*.
  20. Zu, T., Gibbens, B., Doty, N.S., Gomes-Pereira, M., Huguet, A., Stone, M.D., Margolis, J., Peterson, M., Markowski, T.W., Ingram, M.A.C., et al. (2011). Non-ATG-initiated translation directed by microsatellite expansions. *Proc. Natl. Acad. Sci. U. S. A.* *108*, 260–265.
  21. Visconti, V.V., Centofanti, F., Fittipaldi, S., Macri, E., Novelli, G., and Botta, A. (2021). Epigenetics of Myotonic Dystrophies: A Minireview. *Int. J. Mol. Sci.* *22*, 12594.
  22. Coppède, F. (2020). Epigenetics of neuromuscular disorders. *Epigenomics* *12*, 2125–2139.
  23. Castel, A.L., Overby, S.J., and Artero, R. (2019). MicroRNA-based therapeutic perspectives in myotonic dystrophy. *Int. J. Mol. Sci.* *20*.
  24. Koehorst, E., Ballester-lopez, A., Arechavala-gomez, V., Martínez-piñeiro, A., and Nogales-gadea, G. (2020). The Biomarker Potential of miRNAs in Myotonic Dystrophy Type I. *J. Clin. Med.* *9*, 1–18.
  25. Berger, S.L., Kouzarides, T., Shiekhattar, R., and Shilatifard, A. (2009). An operational definition of epigenetics. *Genes Dev.* *23*, 781.
  26. Jones, P.A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* *13*, 484–492.
  27. Carrió, E., Díez-Villanueva, A., Lois, S., Mallona, I., Cases, I., Forn, M., Peinado, M.A., and Suelves, M. (2015). Deconstruction of DNA methylation patterns during myogenesis reveals specific epigenetic events in the establishment of the skeletal muscle lineage. *Stem Cells* *33*, 2025–2036.
  28. Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* *447*, 425–432.
  29. Gardiner-Garden, M., and Frommer, M. (1987). CpG islands in vertebrate genomes. *J. Mol. Biol.* *196*, 261–282.
  30. Bird, A.P. (1986). CpG-rich islands and the function of DNA methylation. *Nature* *321*, 209–213.
  31. Deaton, A.M., and Bird, A. (2011). CpG islands and the regulation of transcription. *Genes Dev.* *25*, 1010–1022.
  32. Straussman, R., Nejman, D., Roberts, D., Steinfeld, I., Blum, B., Benvenisty, N., Simon, I.,

- Yakhini, Z., and Cedar, H. (2009). Developmental programming of CpG island methylation profiles in the human genome. *Nat. Struct. Mol. Biol.* *16*, 564–571.
33. Filippova, G.N., Thienes, C.P., Penn, B.H., Cho, D.H., Hu, Y.J., Moore, J.M., Klesert, T.R., Lobanenkov, V. V., and Tapscott, S.J. (2001). CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. *Nat. Genet.* *28*, 335–343.
34. Barbé, L., Lanni, S., López-Castel, A., Franck, S., Spits, C., Keymolen, K., Seneca, S., Tomé, S., Miron, I., Letourneau, J., et al. (2017). CpG Methylation, a Parent-of-Origin Effect for Maternal-Biased Transmission of Congenital Myotonic Dystrophy. *Am. J. Hum. Genet.* *100*, 488–505.
35. Yanovsky-Dagan, S., Avitzour, M., Altarescu, G., Renbaum, P., Eldar-Geva, T., Schonberger, O., Mitrani-Rosenbaum, S., Levy-Lahad, E., Birnbaum, R.Y., Gepstein, L., et al. (2015). Uncovering the Role of Hypermethylation by CTG Expansion in Myotonic Dystrophy Type 1 Using Mutant Human Embryonic Stem Cells. *Stem Cell Reports* *5*, 221–231.
36. Brouwer, J.R., Huguet, A., Nicole, A., Munnich, A., and Gourdon, G. (2013). Transcriptionally repressive chromatin remodelling and CpG methylation in the presence of expanded CTG-repeats at the DM1 locus. *J. Nucleic Acids* *2013*.
37. Spits, C., Seneca, S., Hilven, P., Liebaers, I., and Sermon, K. (2010). Methylation of the CpG sites in the myotonic dystrophy locus does not correlate with CTG expansion size or with the congenital form of the disease. *J. Med. Genet.* *47*, 700–703.
38. Breton, É., Légaré, C., Overend, G., Guay, S.P., Monckton, D., Mathieu, J., Gagnon, C., Richer, L., Gallais, B., and Bouchard, L. (2020). DNA methylation at the DMPK gene locus is associated with cognitive functions in myotonic dystrophy type 1. *Epigenomics* *12*, 2051–2064.
39. Hildonen, M., Knak, K.L., Dunø, M., Vissing, J., and Tümer, Z. (2020). Stable Longitudinal Methylation Levels at the CpG Sites Flanking the CTG Repeat of DMPK in Patients with Myotonic Dystrophy Type 1. *Genes (Basel)* *11*, 1–13.
40. Morales, F., Vásquez, M., Corrales, E., Vindas-Smith, R., Santamaría-Ulloa, C., Zhang, B., Sirito, M., Estecio, M.R., Krahe, R., and Monckton, D.G. (2020). Longitudinal increases in somatic mosaicism of the expanded CTG repeat in myotonic dystrophy type 1 are associated with variation in age-at-onset. *Hum. Mol. Genet.* *29*, 2496–2507.
41. Morales, F., Corrales, E., Zhang, B., Vásquez, M., Santamaría-Ulloa, C., Quesada, H., Sirito, M., Estecio, M.R., Monckton, D.G., and Krahe, R. (2021). Myotonic dystrophy type 1 (DM1) clinical sub-types and CTCF site methylation status flanking the CTG expansion are mutant allele length-dependent. *Hum. Mol. Genet.*
42. Santoro, M., Fontana, L., Masciullo, M., Bianchi, M.L.E., Rossi, S., Leoncini, E., Novelli, G., Botta, A., and Silvestri, G. (2015). Expansion size and presence of CCG/CTC/CGG sequence interruptions in the expanded CTG array are independently associated to hypermethylation at the DMPK locus in myotonic dystrophy type 1 (DM1). *Biochim. Biophys. Acta - Mol. Basis Dis.* *1852*, 2645–2652.
43. Steinbach, P., Gläser, D., Vogel, W., Wolf, M., and Schwemmler, S. (1998). The DMPK gene of severely affected myotonic dystrophy patients is hypermethylated proximal to the largely expanded CTG repeat. *Am. J. Hum. Genet.* *62*, 278–285.
44. Légaré, C., Overend, G., Guay, S.P., Monckton, D.G., Mathieu, J., Gagnon, C., and Bouchard, L. (2019). DMPK gene DNA methylation levels are associated with muscular and respiratory profiles in DM1. *Neurol. Genet.* *5*.
45. Buckley, L., Lacey, M., and Ehrlich, M. (2016). Epigenetics of the myotonic dystrophy-associated DMPK gene neighborhood. *Epigenomics* *8*, 13–31.
46. Miller, S.A., Dykes, D.D., and Polesky, H.F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* *16*, 1215.
47. Carrió, E., Magli, A., Muñoz, M., Peinado, M.A., Perlingeiro, R., and Suelves, M. (2016). Muscle cell identity requires Pax7-mediated lineage-specific DNA demethylation. *BMC Biol.* *14*.
48. Mallona, I., Díez-Villanueva, A., and Peinado, M.A. (2014). Methylation plotter: a web tool for dynamic visualization of DNA methylation data. *Source Code Biol. Med.* *9*.
49. Légaré, C., Overend, G., Guay, S.-P., Monckton, D.G., Mathieu, J., Gagnon, C., and Bouchard, L. (2019). DMPK gene DNA methylation levels are associated with muscular and respiratory profiles in DM1. *Neurol. Genet.* *5*, e338.
50. Morales, F., Couto, J.M., Higham, C.F., Hogg, G., Cuenca, P., Braida, C., Wilson, R.H., Adam, B., del Valle, G., Brian, R., et al. (2012). Somatic instability of the expanded CTG triplet repeat in myotonic dystrophy type 1 is a heritable quantitative trait and modifier of disease severity. *Hum. Mol. Genet.* *21*, 3558–3567.
51. Cumming, S.A., Jimenez-Moreno, C., Okkersen, K., Wenninger, S., Daidj, F., Hogarth, F., Littleford, R., Gorman, G., Bassez, G., Schoser, B., et al. (2019). Genetic determinants of disease severity in the myotonic dystrophy type 1 OPTIMISTIC cohort. *Neurology* *93*, E995–E1009.
52. Ballester-Lopez, A., Linares-Pardo, I., Koehorst, E., Núñez-Manchón, J., Pintos-Morell, G., Coll-

- Cantí, J., Almendrote, M., Lucente, G., Arbex, A., Magaña, J.J., et al. (2020). The need for establishing a universal CTG sizing method in myotonic dystrophy type 1. *Genes (Basel)*. *11*, 1–9.
53. Nichol, K., and Pearson, C.E. (2002). CpG methylation modifies the genetic stability of cloned repeat sequences. *Genome Res.* *12*, 1246–1256.
54. Yanovsky-Dagan, S., Cohen, E., Megalli, P., Altarescu, G., Schonberger, O., Eldar-Geva, T., Epsztejn-Litman, S., and Eiges, R. (2021). DMPK hypermethylation in sperm cells of myotonic dystrophy type 1 patients. *Eur. J. Hum. Genet.*
55. Antequera, F., Boyes, J., and Bird, A. (1990). High levels of de novo methylation and altered chromatin structure at CpG islands in cell lines. *Cell* *62*, 503–514.
56. Jones, P.A., Wolkowicz, M.J., Rideout, W.M., Gonzales, F.A., Marziasz, C.M., Coetzee, G.A., and Tapscott, S.J. (1990). De novo methylation of the MyoD1 CpG island during the establishment of immortal cell lines. *Proc. Natl. Acad. Sci. U. S. A.* *87*, 6117–6121.



**Table S1.** Primer combinations and thermocycler settings

<b>Primer Name</b>	<b>Sequence (5'-3')</b>	<b>T<sub>m</sub></b>	<b>Cycles</b>
<i>Long PCR</i>			
DM-C	AACGGGGCTCGAAGGGTCCT	63.5	28
DM-DR	CAGGCCTGCAGTTTGCCCATC		
<i>Methylation</i>			
CpGi 74-F-1	GGTAAGTAATGGAGTTAGTTT	57	40
CpGi 74-R-1	ACTTCTCTATCTATACTACCA		
CpGi 74-F-2	GTGTAGGGGTTAAAGGTTATAG	55	40
CpGi 74-R-2	CTATCTATACTACCAAAAACAA		
CpGi 43-F-1	GATAGTGAGATAGAGTGGAG	55	40
CpGi 43-R-1	AACATTTTTTAACCCCAAAAC		
CpGi 43-F-2	AGTTTAGTTTTAGGTTTTAAGGT	55	40
CpGi 43-R-2	CTTATCTCCAATACCCTTTCTAA		
CpGi 36-F-1	GGGAATGAGTGATTTAGGATTT	57	40
CpGi 36-R-1	CTCTCCCTCTAAACAAAACACCT		
CpGi 36-F-2	AGTAGGGGTTATAGGTATTTATTT	57	40
CpGi 36-R-2	TCTAAACAAAACACCTCTCTCTAC		
CTCF1-F-1	TGTYGTYGTTTTGGGTTGTATTG	57	40
CTCF1-R-1	TTCCYGACTACAAAACCCCTTYG		
CTCF1-F-2	GTTGTATTGGGTTGGTGGTTTA	57	40
CTCF1-R-2	CTACAAAACCCCTTYGAACCC		
CTCF2-F-1	TTYGGTTAGGTTGAGGTTT	57	40
CTCF2-R	TTAACAAAACAAATTTCCC		
CTCF2-F-2	TAAATTGTAGGTTTGGGAAG	57	40

1 and 2 in the primer name correspond to PCR 1 and 2 in the nested PCR. T<sub>m</sub>= melting temperature

**Table S2.** DNA methylation analysis of CpG 74 in blood samples across the clinical DM1 subtypes.

	CpG →	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19				
<b>Congenital</b>	P1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1			
	P2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
	P3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
	P5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
<b>Childhood</b>	P8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
<b>Juvenile</b>	P16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P23	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P31	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
<b>Adult</b>	P36	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P37	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P38	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P40	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P41	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
<b>Late-onset</b>	P50	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P51	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P52	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P53	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P54	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
<b>Asymptomatic Control</b>	P55	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P56	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P57	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P58	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P59	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
<b>Asymptomatic Control</b>	P60	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P61	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P65	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	C1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	C2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
<b>Asymptomatic Control</b>	C3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	C4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.75

Range 0 to 1 indicates 0 to 100% methylation.

**Table S3.** DNA methylation analysis of CpGi 43 in blood across the DM1 clinical subtypes

	CpG →	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
<i>Congenital</i>	P1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.2
	P2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.2
	P3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5
<i>Childhood</i>	P6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.2
	P7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3
	P10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Juvenile</i>	P11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P18	0	0	0	0	0	0	0	0	0	0.2	0	0.1	0.2	0.4	0.5
<i>Adult</i>	P19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.4
	P20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1
	P41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5
<i>Late-onset</i>	P51	0	0	0	0	0	0	0	0.2	0	0.25	0	0	0	0	0.5
	P52	0	0	0	0	0	0	0	0	0	0	0	0.2	0.3	0.45	
	P53	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.15
	P54	0	0	0	0	0	0	0	0	0	0	0	0	0	0.2	0.3
	P55	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P57	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.4
	P58	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P59	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P60	0	0	0	0	0	0	0	0	0	0.2	0	0	0.2	0.3	0.5
<i>Asymptomatic</i>	P61	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P64	0	0	0	0	0	0	0	0	0	0	0.2	0	0.3	0.4	0.6
<i>Control</i>	P65	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1
	C1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	C2	0	0	0	0	0	0	0	0	0	0	0	0	0	0.2	0.2
	C3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Range 0 to 1 indicates 0 to 100% methylation. NA = not available

**Table S4.** DNA methylation analysis of CpGi 36 in blood across the DM1 clinical subtypes

CpG →	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
<i>Congenital</i>	1	1	1	1	1	1	1	1	1	1	0.9	1	1	1	1	0.9	0.9	0.9
P1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P3	1	1	1	1	1	0.5	1	1	1	1	1	1	1	1	1	0.9	0.9	1
P4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.8	0.9	1	1
P16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P23	1	1	1	1	1	0.75	1	1	1	1	1	1	1	1	0.9	1	1	1
P31	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P36	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P37	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9	1	1	1
P38	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P40	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P41	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9	1	1	1
P50	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P51	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P52	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P53	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P54	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P55	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P56	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P57	1	1	1	1	1	0.75	1	1	1	1	1	1	1	1	1	1	1	1
P58	1	1	1	1	1	1	0.8	1	1	1	1	1	1	1	0.9	0.9	1	1
P59	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P60	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P61	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
P64	1	1	1	1	1	1	1	0.9	1	1	1	1	1	1	0.9	0.9	1	1
P65	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9	0.9	0.9	1	1
<i>Control</i>	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9	1	1	1
C1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
C2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
C3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
C4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9	0.9	1	1

Range 0 to 1 indicates 0 to 100% methylation.



Table S5. DNA methylation analysis of CTCF1 in blood across the DM1 clinical subtypes

CpG →	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	
<b>Congenital</b>	NA	NA	0.15	0.25	0.25	0.25	0.25	0.5	0.5	0.6	0.65	1	0.65	0.65	0.5	0.75	0.75	NA	NA	NA	NA	NA	NA	NA	NA	NA
P1-2	0.25	0.25	0.25	0.4	0.25	0.25	0.25	0.4	0.25	0.5	0.5	0.4	0.4	0.4	0.5	0.4	0.65	0.6	0.5	0.6	0.9	0.9	0.9	0.9	0.9	NA
P2	NA	0.15	0.15	0.25	0.25	0.25	0.25	0.4	0.4	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.8	1	0.5	0.75	0.9	0.9	0.75	0.75	0.75	NA
P3	0	0.15	0.25	0.15	0.15	0.25	0.25	0.25	0.35	0.4	0.4	0.4	0.4	0.4	0.4	0.55	0.6	0.5	0.5	0.5	0.75	0.75	0.75	0.75	0.75	0.75
P4	NA	0.25	0.25	0.25	0.25	0.25	0.25	0.3	0.4	0.5	0.5	0.5	0.4	0.4	0.4	0.75	0.75	0.75	0.5	0.5	NA	NA	NA	NA	NA	NA
P5	1	1	1	1	1	1	1	1	0.45	1	1	1	1	1	1	1	1	1	0.6	0.75	1	0.75	1	1	1	NA
P6	0.25	0.25	0.4	0.25	0.25	0.25	0.25	0.4	0.25	0.25	0.25	0.25	0.4	0.4	0.25	0.5	0.5	0.6	0.5	0.6	0.8	0.9	1	1	1	NA
P7	1	1	1	0.6	0.75	0.75	0.75	0.75	0.75	0.8	0.75	0.85	0.85	0.85	0.75	0.85	0.85	0.85	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9
P8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	NA	NA	NA	NA	NA	NA
P9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	NA	NA	NA	NA	NA	NA
P10	NA	0.25	0.5	0.5	0.25	0.25	0.25	0.25	0.25	0.4	0.5	0.45	0.45	0.5	0.4	0.75	0.75	0.75	0.5	0.75	0.75	0.75	0.75	0.75	0.75	0.75
P11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P12	NA	0.5	0.75	0.5	0.5	0.5	0.5	0.5	0.5	0.6	0.75	0.55	0.5	0.5	0.5	0.75	0.5	0.55	0.4	0.5	0.75	0.8	0.75	0.8	0.8	0.8
P13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P14	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	NA	NA	NA	NA	NA	NA
P15	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	NA	NA	NA	NA	NA	NA
P16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P19	NA	0.25	0.25	0.25	0.25	0.25	0.25	0.4	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.8	0.9	0.9	0.5	0.6	0.9	1	0.75	0.75	0.9	0.9
P20	NA	0.15	0.15	0.25	0.25	0.25	0.25	0.4	0.4	0.5	0.5	0.5	0.5	0.5	0.5	0.75	0.75	0.75	0.5	0.75	0.5	0.4	0.9	0.9	0.5	0.9
P21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P22	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P23	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P24	NA	0.05	0.05	0.05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P25	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P26	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P27	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P28	0.05	0.05	0.05	0.05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P29	0.05	0.05	0.05	0.05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P30	0.4	0.4	0.23	0.45	0.6	0.75	0.75	0.75	0.7	0.8	0.75	0.8	0.8	0.8	0.8	0.75	0.75	0.85	0.85	0.85	1	1	0.9	0.75	1	1
P31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P42	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P43	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P46	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P47	NA	0	0	0	0	0	0	0.1	0.25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P48	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P49	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P50	0.1	0.1	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.25	0.25	0.25	0.25	0.3	0.25	0.4	0.5	0.5	0.5	0.5	0.6	0.6	0.6	0.6	0.6	0.6
P51	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P52	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P53	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Continues on the next page



**Table S6.** DNA methylation analysis of CTCF2 in blood across the DM1 clinical subtypes

	CpG →	1	2	3	4	5	6	7	8	9	10	11
<b>Congenital</b>	P1	0	0	0	0	0	0	0.1	0	0	0	0
	P1-2	0	0	0	0	0	0.05	0	0.05	0.1	0.1	0
	P2	0	0	0	0.1	0.1	0.15	0.1	0.1	0.15	0.2	0.25
	P3	NA	NA	NA	0.05	0.05	0.1	0.05	0	0	0	0.1
	P4	0	0	0	0	0	0	0	0.05	0.05	0.05	0.05
	P5	NA	NA	NA	0.4	0.25	0.4	0.3	0.4	0.6	0.75	0.75
	P6	0	0	0	0.05	0.05	0.1	0.1	0.15	0.25	0.45	0.45
<b>Childhood</b>	P7	0	0	0	0	0	0	0	0	0	0	0
	P8	0	0	0	0	0	0	0	0	0	0	0
	P9	0	0	0	0	0	0	0	0	0	0	0
	P10	0	0	0	0	0	0	0	0.05	0.05	0.05	0
	P11	0	0	0	0	0	0	0	0	0	0	0
	P12	0	0	0	0	0	0	0	0	0	0	0
	P13	0	0	0	0	0	0	0	0	NA	NA	NA
	P14	0	0	0	0.05	0	0.05	0.05	0	0.05	0.05	0
	P15	0	0	0	0	0	0	0	0	0	0	0
	P16	0	0	0	0	0	0	0	0	0	0	0
<b>Juvenile</b>	P17	0	0	0	0	0	0	0.1	0	0	0	0
	P18	0	0	0	0	0	0	0	0	NA	NA	NA
	P19	0.05	0.05	0	0.15	0.1	0.15	0.1	0.1	0.2	0.25	0.25
	P20	0.05	0.05	0	0.2	0.15	0.2	0.1	0.25	0.45	0.5	0.5
	P21	0	0	0	0	0	0	0.5	0	0	0	0
	P22	NA	NA	NA	0	0	0	0.2	0	0	0	0
	P23	0	0	0	0	0	0.05	0.05	0.05	0.1	0.1	0.1
	P24	0	0	0	0	0	0	0	0	0	0	0
	P25	0	0	0	0	0	0	0	0	0.05	0.05	0.05
	P26	0	0	0	0	0	0	0	0	0	0	0
	P27	0	0	0	0.05	0.05	0	0	0	0	0	0
	P28	NA	NA	NA	0	0	0	0	0	0	0	0
	P29	NA	NA	NA	0	0	0	0.2	0	0	0	0
	P30	0	0	0	0.05	0	0	0	0	0	0	0
	P31	NA	NA	NA	0.5	0.6	1	0.75	0.75	1	1	1
	P32	0	0	0	0	0	0	0	0	0	0	0
	P33	0	0	0	0	0	0	0	0	0	0	0
	P35	0	0	0	0	0	0	0	0	0	0	0
	P37	0	0	0	0	0	0	0	0	0	0	0
	P38	0	0	0	0	0	0	0	0	0	0	0
	P39	0	0	0	0	0	0	0	0	0	0	0.05
	P40	0	0	0	0	0	0	0	0	0	0	0
	P41	0	0	0	0	0	0	0	0	0	0	0
	P42	0	0	0	0	0	0	0	0	0	0	0
	P43	0	0	0	0	0	0	0	0	0	0	0
	P44	NA	NA	NA	0	0	0	0.5	0	0	0	0
	P45	0	0	0	0	0	0	0	0	0	0	0
	P48	0	0	0	0	0	0	0	0	0	0	0
P49	0	0	0	0	0	0	0	0	0	0	0	
P50	NA	NA	NA	0	0	0	0	0	0	0	0	
P51	NA	NA	NA	0	0	0	0	0	0	0	0	
P52	0	0	0	0	0	0	0	0	0	0	0	
P53	0	0	0	0	0	0	0	0	0	0	0	
P54	0	0	0	0	0	0	0	0.1	0	0	0	
P55	0	0	0	0	0	0	0	0	0	NA	NA	NA
P56	0	0	0	0	0	0	0	0	0	NA	NA	NA
P57	0	0	0	0	0	0	0	0	0	0	0	0
P58	0	0	0	0	0	0	0	0	0	0	0	0
P59	0	0	0	0	0	0	0	0	0	0	0	0
P60	0	0	0	0	0	0	0	0.05	0	0	0	0
P63	0	0	0	0	0	0	0	0.1	0	0	0	0
P64	0	0	0	0	0	0	0	0	0	NA	NA	NA
P65	0	0	0	0	0	0	0	0	0	NA	NA	NA
<b>Control</b>	Ctrl 1	0	0	0	0	0	0	0	0	0	0	0
	Ctrl 2	0	0	0	0	0	0	0	0	0	0	0
	Ctrl 3	0	0	0	0	0	0	0	0	0	0	0
	Ctrl 4	0	0	0	0	0	0	0	0	0	0	0
	Ctrl 5	0	0	0	0	0	0	0	0	0	0	0
	Ctrl 6	0	0	0	0	0	0	0	0	0	0	0
	Ctrl 7	NA	0	0	0	0	0	0	0	0	0	0

Range 0 to 1 indicates 0 to 100% methylation. NA = not available. P1-2= five year follow-up P1

Table S7. Clinical characteristics of congenital cases

Patient	Methyl CTCF1/2	Age of onset (years)	Age at sampling (years)	Gender	ePAL (CTGs)	Myotonia	Facial weakness	Axial weakness	Limb weakness	MIRS	Cardiac involvement	NVM	Hyper-somnolence	Cataracts	mRS	DM1-ACTIV
P1	yes/no	birth	10	female	831	no	mild	moderate	moderate proximal	4	no	no	no	no	3	32
P2	yes/yes	birth	18	male	587	no	severe	mild	severe proximal	5	LAFB	no	U	no	5	1
P3	yes/no	birth	2	male	1011	no	moderate	moderate	NA	4	no	no	no	no	NA	NA
P4	yes/no	birth	16	male	730	yes	moderate	mild	moderate distal	3	no	yes	no	no	3	38
P5	yes/yes	birth	16	male	222	yes	unknown	none	moderate distal	3	no	no	no	no	2	40
P6	yes/yes	birth	15	female	280	no	mild	moderate	moderate proximal	4	no	no	yes	no	3	30

Methyl CTCF1/2= hypermethylated at the CTCF1 and CTCF2 site; ePAL= estimated progenitor allele size; NA = not applicable; U= unknown; AV = atrioventricular; LAFB = left anterior fascicular block; mRS = modified Rankin Scale; MIRS = Muscular Impairment Rating Scale; NVM = nocturnal mechanical ventilation.

**Table S8.** Clinical characteristics of juvenile cases

Patient	Methyl CTGF1/2	Age of onset (years)	Age at sampling (years)	Gender	ePAL (CTGs)	Myotonia	Facial weakness	Axial weakness	Limb weakness	MIRS	Cardiac involvement	NVM	Hyper-somnolence	Cataracts	mRS	DM1-ACTIV
P13	no/no	15	39	female	286	yes	mild	none	mild/moderate proximal + distal	4	LAFB	no	yes	no	2	U
P14	no/no	12	22	male	NA	yes	none	none	none	2	NA	no	no	no	1	40
P15	no/no	18	54	female	392	yes	moderate	U	moderate proximal + distal	4	1° AV block	no	no	yes	3	13
P16	no/no	11	12	male	517	yes	U	U	none	2	no	no	U	no	1	36
P17	no/no	16	19	female	273	yes	mild	U	mild distal	1	no	no	no	no	1	40
P18	no/no	11	16	female	222	yes	none	none	mild distal	2	no	no	no	no	1	40
P19	yes/yes	15	26	female	U	yes	none	none	none	2	no	no	yes	no	1	39
P20	yes/yes	15	24	male	189	yes	none	none	none	2	no	no	no	no	3	37
P21	no/no	12	40	female	428	yes	mild	moderate	mild distal	3	no	no	yes	no	2	18
P22	no/no	19	37	female	251	yes	mild	none	mild distal	2	no	no	no	no	1	29
P23	no/no	12	14	female	642	yes	mild	none	none	2	no	no	yes	no	1	40
P24	no/no	14	16	female	229	yes	mild	none	none	2	no	no	no	no	1	40
P25	no/no	15	15	male	235	yes	none	none	none	1	1°AV block	no	no	no	0	40
P26	no/no	15	16	male	323	yes	mild	none	mild/moderate proximal + distal	2	no	no	no	no	1	40
P27	no/no	16	37	female	330	yes	mild	mild	moderate proximal + distal	4	1°AV block	no	yes	yes	2	29
P28	no/no	17	45	male	236	yes	mild	U	mild/moderate proximal + distal	3	1°AV block + LAFB	yes	no	no	3	24
P29	no/no	18	26	female	254	yes	none	none	none	1	no	no	no	no	1	40
P30	no/no	18	24	male	274	yes	mild	none	mild distal	2	no	no	yes	no	1	33
P31	yes/yes	16	53	female	420	yes	mild	none	moderate proximal + distal	3	no	no	no	yes	2	27
P32	no/no	15	29	male	237	yes	moderate	none	none	2	pacemaker	yes	no	no	3	19
P33	no/no	19	24	male	228	yes	mild	none	mild distal	2	no	no	yes	no	1	37
P34	no/no	13	20	female	368	yes	mild	none	mild distal	2	no	no	yes	no	1	39
P35	no/no	14	15	male	318	yes	mild	none	none	2	no	no	yes	no	1	39

Methyl CTGF1/2= hypermethylated at the CTGF1 and CTGF2 site; ePAL= estimated progenitor allele size; U= unknown; AV = atrioventricular; LAFB = left anterior fascicular block; mRS = modified Rankin Scale; MIRS = Muscular Impairment Rating Scale; NVM = nocturnal mechanical ventilation

**Table S9.** Logistic regression model ePAL/modal vs. Methylation status

<i>Model</i>	<i>Variables</i>	<i>OR</i>	<i>CI</i>	<i>p-value</i>	<i>N</i>
1	ePAL versus MetS	1.005	1.002-1.008	0.004	53
2	Modal versus MetS	1.004	1.002-1.007	0.001	46

Significance is set at  $p \leq 0.05$ . ePAL= estimated progenitor allele size; MetS= methylation status

**Table S10.** DNA methylation analysis of CpGi 74 in lymphoblastoids

	<i>CpG</i> →	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
<i>Congenital</i>	P1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
<i>Childhood</i>	P4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
<i>Juvenile</i>	P10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
<i>Adult</i>	P20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P36	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
<i>Late-onset</i>	P38	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P56	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P56	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
<i>Control</i>	C1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	C4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	C5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	C6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Range 0 to 1 indicates 0 to 100% methylation.

**Table S11.** DNA methylation analysis of CpGi 43 in lymphoblastoids

	<i>CpG</i> →	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
<i>Congenital</i>	P1	0	0	0	0	0	0	0	0	0	0	0	0	0	0.35	0.3
	P2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.25
<i>Childhood</i>	P4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.15
	P7	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0.2
	P8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3
<i>Juvenile</i>	P10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P11	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.6
	P13	0	0	0	0	0	0	0	0	0	0	0	0	0	0.2	0.2
<i>Adult</i>	P20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Late-onset</i>	P38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	P56	0	0	0	0	0.2	0	0	0	0	0	0	0	0	0.4	0.3
	P56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Control</i>	C1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	C4	0	0	0	0	0	0	0	0	0	0	0	0	0	0.45	0.45
	C5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	C6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Range 0 to 1 indicates 0 to 100% methylation.

**Table S12.** DNA methylation analysis of CpGi 36 in lymphoblastoids

CpG →	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
<b>Congenital</b>																		
P2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.8	1	1	
P4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9	0.9	1	
<b>Childhood</b>																		
P7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
P8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
P9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
P10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9	0.9	1	
P13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9	0.9	1	
<b>Juvenile</b>																		
<b>Adult</b>																		
P36	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9	1	
P38	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9	0.9	1	
<b>Late-onset</b>																		
<b>Control</b>																		
C4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9	0.8	1	
C5	1	1	1	1	1	1	1	1	1	0.8	1	1	1	1	1	0.9	0.9	1
C6	1	1	1	1	1	1	1	0.8	1	1	1	1	1	1	0.8	0.8	1	

Range 0 to 1 indicates 0 to 100% methylation.

**Table S13.** DNA methylation analysis of CTCF1 in lymphoblastoids

CpG →	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	
<b>Congenital</b>																										
P1	0.25	0.25	0.25	0.4	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.45	0.5	0.5	0.6	0.4	0.6	0.75	0.75	0.75	0.75	0.75
P2	0.2	0.2	0.3	0.3	0.45	0.4	0.4	0.4	0.45	0.45	0.5	0.4	0.4	0.5	0.4	0.6	0.55	0.55	0.45	0.6	0.75	0.75	0.75	0.75	0.75	0.8
P4	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.15	0.1	0.1	0.1	0.1	0.1	0.2	0.25	0.25	0.2	0.3	0.25	0.25
P7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P10	0.25	0.25	0.25	0.5	0.4	0.4	0.45	0.5	0.5	0.5	0.5	0.5	0.5	0.65	0.5	0.75	0.5	0.55	0.45	0.55	0.75	0.8	0.8	0.8	0.8	0.75
P11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P20	0.2	0.2	0.2	0.4	0.3	0.3	0.3	0.4	0.4	0.35	0.4	0.4	0.4	0.4	0.45	0.5	0.5	0.55	0.5	0.75	0.75	0.75	0.75	0.7	0.7	0.7
P36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Range 0 to 1 indicates 0 to 100% methylation.

**Table S14.** DNA methylation analysis of CTCF2 in lymphoblastoids

CpG →	1	2	3	4	5	6	7	8	9	10	11
<b>Congenital</b>											
P2	NA	0	0	0.1	0.1	0.1	0.1	0.15	0.1	0.15	0.2
P4	NA	0	0	0.1	0.1	0.1	0.1	0.1	0.1	0.15	0.15
P7	NA	0	0	0	0.15	0	0	0.05	0.05	0.05	0.05
P8	0	0	0	0	0	0	0	0	0	0	0
P9	NA	0	0	0	0	0	0	0.05	0.05	0.05	0.05
P10	NA	0	0.05	0.1	0.1	0.1	0.1	0.1	0.15	0.15	0.1
P11	NA	0	0	0	0.05	0	0	0	0.5	0.1	0.1
P13	0	0	0	0	0	0	0	0	0	0	0
P20	NA	0	0	0.1	0.2	0.05	0.1	0.1	0.1	0.15	0.15
P36	0	0	0	0	0	0	0	0	0	0	0
P38	0	0	0	0	0	0	0	0	0	0	0
P56	0	0	0	0	0	0	0	0	0	0	0
C1	0	0	0	0	0	0	0	0	0	0	0
C4	0	0	0	0	0	0	0	0	0	0	0
C5	0	0	0	0	0	0	0	0	0	0	0
C6	0	0	0	0	0	0	0	0	0	0	0

Range 0 to 1 indicates 0 to 100% methylation. NA = not available



**Table S15.** DNA methylation analysis of CpGi 74 in skin and skin-derived cells

Skin	CpG →																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Adult	P37	1	0.85	1	1	0.7	1	0.75	1	1	1	1	1	1	1	0.7	1	1	1
	P68	1	0.75	1	1	1	1	1	1	1	1	1	1	1	1	0.5	1	1	1
	P69	0.75	1	1	1	1	1	0.95	1	1	1	1	1	1	1	0.7	1	1	1
Control	C2	1	0.75	1	1	0.9	0.75	0.9	1	1	1	1	1	1	0.6	1	0.8	0.8	0.8
	C10	1	0.8	1	1	1	0.75	0.8	0.8	1	1	1	1	1	1	0.75	1	0.95	0.95
Skin fibroblasts	P13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.5	1	1	1
	P36	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	P38	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Control	C2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	C6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.75	1	1	1
	C9	1	1	1	1	1	1	1	0.75	1	1	0.75	1	1	1	0.75	1	1	1

Range 0 to 1 indicates 0 to 100% methylation.

**Table S16.** DNA methylation analysis of CpGi 74 in muscle and muscle-derived cells

Muscle	CpG →																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Juvenile	P13	1	0.75	1	1	0.75	1	0.65	1	0.85	0.9	0.9	1	1	1	1	0.5	1	1
	P37	1	0.75	1	1	0.73	0.9	0.55	0.9	1	1	1	1	1	1	0.85	1	1	1
	P68	0	0.5	1	0.6	0.75	0.75	0.4	0.83	0.83	0.83	0.9	1	1	1	1	0.6	1	1
Late-onset	P56	0.68	0.75	1	0.95	0.73	0.9	0.55	0.85	1	0.9	1	1	1	1	0.75	1	1	1
	P67	0.4	0.75	0.75	1	0.95	1	0.75	1	1	1	1	1	1	0.9	0.2	1	1	1
Control	C6	1	1	1	1	1	1	1	0.85	1	1	0.5	1	1	1	1	0.5	1	0.7
	C9	1	0.67	1	1	1	1	0.6	0.9	1	1	1	1	1	1	0.8	1	0.9	0.9
Myoblasts	P36	1	1	1	NA	NA	NA	NA	1	1	NA	NA	1	1	1	1	1	1	1
	P37	1	1	1	1	1	1	0.75	0.75	1	0.9	0.8	0.9	1	1	0.87	0.8	1	1
	P38	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Late-onset	P56	1	1	1	NA	NA	1	0.95	1	1	1	1	0.9	0.9	1	1	1	0.9	0.8
	C2	1	0.75	1	NA	NA	1	0.8	1	1	1	1	1	1	1	0.8	1	1	1
Control	C3	NA	0.73	1	NA	NA	NA	NA	NA	1	1	1	1	1	1	1	1	1	1
	C9	1	1	1	NA	NA	NA	0.9	1	1	1	1	1	1	1	1	1	1	1
	C10	1	0.75	1	NA	NA	NA	0.8	1	0.8	0.9	1	1	0.9	0.9	1	1	0.87	0.87
Myotubes	P13	NA	0.4	1	NA	NA	1	0.75	0.65	1	1	1	1	1	1	0.5	1	1	1
	P36	1	1	1	1	0.9	1	0.8	1	1	1	1	1	1	1	1	1	1	1
	P37	1	0.5	1	1	0.75	1	0.75	0.75	0.75	1	1	1	1	1	1	0.6	1	1
Late-onset	P38	1	1	1	NA	NA	NA	0.5	0.9	1	1	1	1	1	1	0.9	1	NA	1
	P56	1	1	1	NA	NA	1	1	1	1	1	1	1	1	1	1	1	1	1
Control	C2	1	1	1	NA	NA	1	0.75	1	0.9	1	1	1	1	1	1	1	1	1
	C4	1	0.9	1	NA	NA	1	0.9	0.8	1	1	1	1	1	1	0.75	1	1	1
	C6	NA	1	1	1	1	1	1	0.9	1	0.95	0.95	1	1	0.9	0.75	0.8	0.85	1
C9	NA	1	1	NA	NA	NA	0.75	0.75	1	1	1	1	1	0.9	0.9	0.7	1	0.8	0.9

Range 0 to 1 indicates 0 to 100% methylation. NA = not available

**Table S17.** DNA methylation analysis of CpGi 36 in skin and skin-derived cells

Skin	CpG →	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
<b>Adult</b>	P37	NA	1	1	1	1	1	1	0.75	1	1	1	1	1	0.9	0.9	0.9	0.75
	P68	NA	1	1	1	1	1	1	0.8	1	1	1	1	1	1	1	1	1
	P69	NA	1	1	1	0.9	1	0.9	1	1	1	1	1	1	1	0.9	0.9	1
	C2	NA	1	1	1	0.8	1	0.8	1	1	1	1	1	1	1	1	1	0.85
<b>Control</b>	C10	NA	1	1	1	0.8	1	0.8	1	1	1	1	1	1	1	0.75	0.75	1
	P13	1	1	1	1	0.55	1	1	1	1	0.8	1	1	1	1	1	1	1
	P36	1	1	1	1	1	1	0.9	1	1	1	1	1	1	1	0.75	1	1
	P37	1	1	1	1	1	1	1	0.9	1	1	1	1	1	1	1	1	0.8
<b>Juvenile</b>	P38	1	1	1	1	1	1	1	0.8	0.9	1	1	1	1	1	0.8	0.9	0.8
	C2	0.7	1	1	1	1	1	1	0.9	1	1	1	1	1	1	1	0.9	0.9
	C5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	C6	1	1	1	1	1	1	1	1	0.9	1	1	1	1	1	1	1	1
	C9	NA	1	0.7	0.7	1	1	1	1	1	1	1	1	1	1	1	1	1

Range 0 to 1 indicates 0 to 100% methylation. NA = not available

**Table S18.** DNA methylation analysis of CpG1 36 in muscle and muscle-derived cells

	CpG →	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17		
<b>Muscle</b>	Juvenile	1	1	1	1	0.8	1	0.9	0.75	1	1	0.9	0.9	0.9	0.55	0.75	0.7	0.75		
	Adult	1	1	1	1	0.9	0.9	0.9	0.75	0.9	0.9	0.9	1	1	0.75	0.85	0.85	1		
	P36	1	1	1	1	1	1	0.9	0.8	1	1	1	1	1	0.9	1	1	0.9	0.8	
	P37	1	1	1	1	1	1	1	0.9	1	1	1	1	1	1	1	1	1	1	
	P38	1	1	1	1	1	1	1	0.8	0.8	1	1	1	1	1	0.9	0.9	0.8	0.9	
	P68	NA	NA	1	1	1	1	1	1	1	1	1	0.8	1	1	0.8	NA	NA	NA	NA
	P56	1	1	1	1	1	1	1	1	0.8	1	1	1	1	1	1	0.8	0.9	0.9	
	P67	NA	NA	1	0.8	1	1	1	1	0.8	1	1	1	1	1	1	0.8	0.9	0.9	
	C4	NA	0.5	1	1	1	1	1	1	0.5	1	1	1	1	1	1	1	0.75	0.63	
C6	NA	1	1	1	1	0.85	1	1	0.75	1	1	1	1	1	1	0.75	0.75	0.9		
C9	NA	1	1	1	1	1	0.8	1	0.9	1	1	1	1	0.9	1	0.75	0.5	0.9		
<b>Myoblasts</b>	Juvenile	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
	Adult	1	0.7	1	0.8	1	1	1	0.9	0.6	1	1	0.75	1	1	0.55	0.8	0.8		
	P36	NA	NA	1	1	1	1	1	1	0.95	0.95	0.9	0.9	0.9	0.75	0.7	0.8	0.75		
	P37	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9	0.85	0.9		
	P38	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
	P56	1	1	1	1	1	1	1	0.8	1	1	1	1	1	1	1	1	1	0.9	
<b>Late-onset Controls</b>	C3	0.7	1	1	1	1	1	1	0.7	1	1	1	1	1	1	1	1	1	0.9	
	C3	0.7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9	
	C9	1	0.7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.9	
	C10	1	0.8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
<b>Myotubes</b>	Juvenile	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
	Adult	1	1	1	1	1	1	1	1	1	1	0.8	1	1	1	0.6	0.9	0.8		
	P36	1	1	1	1	1	1	1	0.7	1	1	1	1	1	0.7	0.7	0.8	0.9		
	P37	1	0.5	0.75	0.75	1	0.55	1	0.55	1	1	0.8	1	1	1	0.7	1	1		
	P38	1	0.6	1	1	1	1	1	0.75	1	1	1	1	1	0.8	1	0.9	1		
	P56	1	1	1	1	1	1	1	0.7	1	1	1	1	1	1	1	1	0.6		
	C2	1	1	1	1	0.9	0.8	1	0.7	1	1	1	1	0.9	1	1	1	0.91		
	C4	1	0.6	1	1	1	1	1	1	0.9	1	1	1	1	0.8	0.9	1	0.9		
	C6	1	0.7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.6	
C9	1	0.7	1	1	1	1	0.7	1	0.9	1	1	1	1	1	1	1	0.8	0.8		

Range 0 to 1 indicates 0 to 100% methylation. NA = not available

Table S19. DNA methylation analysis of CpG: 43 in skin and skin-derived cells

Skin	CpG →														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
<b>Adult</b>															
P37	0	0	0	0	0	0.15	0	0.05	0.05	NA	NA	0	0	0	0
P68	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P69	0	0	0	0	0	0.7	0	0	0	0	0	0	0	0	0
<b>Control</b>															
C2	0	0	0	0	0.05	0.1	0.05	0	0.05	0	0	0	0.1	0.1	0.1
C10	0	0	0	0.05	0	0.1	0.05	0.05	0.05	0.1	0.1	0.1	0.1	0	0.1
<b>Skin fibroblasts</b>															
<b>Juvenile</b>															
P13	0	0	0	0	0	0	0	0	0.5	0.4	0	0	NA	NA	0.4
P36	0.2	0.25	0.5	0	0.25	0	0	0	0.4	0.5	0	0	0.5	0.05	0.8
<b>Adult</b>															
P37	0	0	0.05	0	0	0.05	0.05	0.1	0.1	0.05	0.1	0.1	0.1	NA	NA
P38	0	0	0	0	0	0	0	0	0.5	NA	0	0	0	0	0
<b>Control</b>															
C2	0	0	0.1	0	0.1	0	0.1	0.1	0.1	0.1	0.1	0.1	0.05	0.1	0.3
C5	1	0	0	0	0	0	0.6	0	0.75	0	0	0	0	0	0
C6	0	0	0.4	0	0.3	0	0	NA	0.4	0.4	0.2	NA	0	0.5	0.75
C9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Range 0 to 1 indicates 0 to 100% methylation. NA = not available

**Table S20.** DNA methylation analysis of CpG<sub>i</sub> 43 in muscle and muscle-derived cells

	CpG →	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
<i>Muscle</i> <i>Juvenile</i> <i>Adult</i>	P13	0.1	0	0.35	0	0.2	0.2	0.4	0.4	0.55	0.25	0	0.25	0.25	0.4	0.75
	P36	0	0	0.4	0	0.4	1	0.75	0.5	0.5	1	0.6	0	1	0.5	1
	P37	0.25	0	0.5	0	0.4	0.4	0.5	0.5	0.75	0.75	0.5	0	0.75	0.75	0.8
	P38	1	0	0.4	0.35	0.4	0.4	0.8	0.45	0.5	0.5	0.5	0.25	0	0.75	1
<i>Late-onset</i>	P68	0	0	0	0	0	0.15	0	0	0	0	0	0.25	0.35	0	0.5
	P56	0.2	0	0.25	0.25	0.25	0.4	0.4	0.4	0.55	0.5	0.5	0.4	0.4	0.4	0.65
	P67	0.35	0.25	0.25	0	0.05	0.35	0.35	1	0.75	1	0.45	0.45	0.75	0.1	0.8
<i>Control</i>	C6	0.85	0	0.25	0	0.2	0.2	0.75	0.4	0.75	0.55	0.45	0.45	0.45	0	0.85
	C9	0.85	0.25	0.2	0	0.2	0.1	0.55	1	0.6	1	0.65	0.25	0.65	0.25	1
<i>Myoblasts</i> <i>Adult</i>	P36	0	0.2	0	0.5	0.5	0.5	0.5	0.25	0.25	0.25	0.15	0.15	0.75	0.75	0.75
	P37	0.15	0	0	0.6	0	0	0	0	0	0	0.5	0	0	0.5	1
<i>Late-onset</i> <i>Control</i>	P38	0	0	0.25	0.15	0.25	0.15	0.45	0.4	0.5	0.6	0.4	0.5	0.4	0.5	0.75
	P56	0	0	0.85	NA	NA	0.2	0.75	NA	0	0.2	NA	0	NA	0	0.2
<i>Myotubes</i> <i>Juvenile</i> <i>Adult</i>	C2	0.25	0	0.6	0.25	0.65	0.65	0.78	0.85	0.9	1	0.5	0.75	0.85	0.6	0.9
	C3	0.35	0	0.55	0.4	0.55	NA	0.65	0.6	0.6	0.9	0.9	0.6	0.65	0.5	0.6
	C9	0.25	0	0.6	0.4	0.5	NA	0.75	0.85	0.85	1	0.75	0.95	0.75	0.65	0.75
<i>Late-onset</i> <i>Control</i>	C10	0.4	0	0.67	0.25	0.67	0.5	0.75	1	1	1	0.75	0.75	0.6	0.75	0.75
	P13	NA	NA	0	0	0	0	0.5	0.2	0.35	0.35	0.4	0.1	0.25	0.25	0.75
<i>Myotubes</i> <i>Juvenile</i> <i>Adult</i>	P36	0.4	NA	0.45	0.1	0.45	0.5	0.6	0.55	0.65	0.8	0.27	0.25	0.75	0.45	0.8
	P37	0.6	0	0.25	0.20	0.5	NA	0.78	1	0.8	1	0.70	0.85	0.35	0.65	0.9
	P38	0.05	0	0.15	0.1	0.15	0.2	0.25	0.25	0.5	0.3	0.2	0.25	0.6	0.5	0.75
<i>Late-onset</i> <i>Control</i>	P56	NA	0	0	0.55	0	0.55	0.2	0.6	0.25	0.5	0.6	0.5	0.1	0.6	0.35
	C2	0.3	0	0.45	0.1	0.75	NA	0.6	0.6	0.75	0.75	0.55	0.7	0.75	0.6	0.8
<i>Control</i>	C4	0.5	0	0.7	0	0.75	0.5	0.8	0.75	1	1	0.75	0.8	0.5	0.75	0.75
	C6	0	0	0	0	0.25	0	0	0	0	0	0	0	0	0	0.3
	C9	0.5	0	0.75	0.4	0.55	NA	NA	0.85	1	1	NA	0.75	0.85	1	1

Range 0 to 1 indicates 0 to 100% methylation. NA = not available

Table S21. DNA methylation analysis of CTCF1 in skin and skin-derived cells

CpG →	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	
<b>Skin</b>																										
<b>Adult</b>																										
P36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P68	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<b>Control</b>																										
C2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<b>Skin fibros</b>																										
<b>Juvenile</b>																										
P13	NA	0.15	0.25	0.1	0.1	0.1	0.05	0.1	0.1	0.75	0.25	0.1	0.25	0.3	0.5	0.2	0.1	0.1	0.1	0.1	0	0.05	0.15	0.1	0.15	
P36	NA	0.1	0.1	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.1	0.1	0.15	0.1	0.05	0	0	0	0	0	0	0	0	0	
<b>Adult</b>																										
P37	NA	NA	0	0	0.1	0.1	0	0	0.15	0.2	0.4	0.1	0.15	0.1	0.1	0.1	0.1	0.1	0.25	0.25	0.05	0.05	0	0	0	
P38	0	0	0	0	0	0	0	0	0	0	0.4	0.35	0.4	0	0	0	0	0	0	0	0	0	0	0	0	
<b>Control</b>																										
C2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
C5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
C6	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
C9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

Range 0 to 1 indicates 0 to 100% methylation. NA = not available. Skin fibros= skin fibroblasts

Table S22. DNA methylation analysis of CTCF1 in muscle and muscle-derived cells

CpG →	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	
<b>Muscle</b>																										
<b>Juvenile</b>																										
<b>Adult</b>																										
P13	0.2	0.1	0.1	0.2	0.1	0.1	0.1	0	0.15	0.2	0.15	0.2	0.25	0.1	0.25	0.25	0.4	0.6	0.2	0.1	0.1	0.1	0.15	0.1	0.1	
P36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0.15	0	0	0	0	0	0	0	0	0	
P37	0.1	0.1	0.05	0.1	0	0.05	0.05	0.05	0.05	0.1	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0	0	0	0	0	
P38	NA	0.2	0.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
<b>Late-onset</b>																										
P36	0	0.05	0.05	0	0	0	0	0	0	0.1	0.1	0.1	0	0	0	0.1	0	0	0	0	0	0	0	0	0	
P37	NA	0.05	0.15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
<b>Control</b>																										
C4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
C6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
C9	0	0	0	0	0	0	0	0.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
<b>Myoblasts</b>																										
<b>Juvenile</b>																										
<b>Adult</b>																										
P13	NA	0.25	0.4	0.2	0.1	0.1	0	0.15	0.15	0.75	0.25	0.25	0.25	0.4	0.6	0.2	0.1	0.1	0.1	0.1	0	0.1	0.15	0.1	0.1	
P36	NA	NA	NA	0.8	0	0	0	0	0.05	0.05	0.05	0.1	1	1	0.75	0	0	0	0	0	0	0	0	0	0	
P37	0.15	0.15	0.4	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.85	
P38	NA	0.25	0.05	0.25	0.05	0	0	0	0	0.2	0.1	0.2	0.1	0.1	0.35	0.2	0	0	0	0	0	0	0	0	0	
<b>Late-onset</b>																										
<b>Control</b>																										
C3	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
C9	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
C10	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
<b>Myotubes</b>																										
<b>Juvenile</b>																										
<b>Adult</b>																										
P13	NA	0.2	0.25	0.2	0.1	0.1	0.1	0.15	0.15	0.4	0.25	0.2	0.25	0.25	0.2	0.4	0	0.1	0.2	0	0	0	0.1	0.1	0.1	
P36	NA	0.15	0.1	0.4	0.1	0.1	0.05	0.2	0.2	0.25	0.2	0.5	0.25	0.25	0.25	0.65	0	0.05	0	0	0.05	0.05	0.1	0.15	0.1	
P37	NA	1	1	0.4	0.1	0.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.4	
P38	0.15	0.15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
<b>Late-onset</b>																										
<b>Control</b>																										
P34	NA	0.2	0.05	0.6	0.05	0	0.05	0.05	0.05	0.05	0.05	0.1	0.1	0.6	0.15	0	0	0	0	0	0	0	0	0	0	
C2	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
C4	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
C9	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
C6	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

Range 0 to 1 indicates 0 to 100% methylation. NA = not available

**Table S23.** CTG expansion sizes in blood and muscle of the patients providing a muscle biopsy

	<i>Blood progenitor</i>	<i>Blood modal</i>	<i>Muscle progenitor</i>	<i>Muscle modal</i>
<i>Juvenile P13</i>	286	445	691	736
<i>Adult P36</i>	180	338	208	335
<i>Adult P37</i>	199	374	469	788
<i>Adult P38</i>	115	130	135	207
<i>Adult P68</i>	322	619	351	758
<i>Late-onset P56</i>	300	381	NA	NA
<i>Late-onset P67</i>	108	246	128	512

NA = not available

**Table S24.** DNA methylation analysis of CTCF2 in skin and skin-derived cells

	<i>CpG→</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>	<i>10</i>	<i>11</i>	
<i>Skin</i>	<i>Adult</i>	<i>P37</i>	NA	NA	NA	0	0	0	0	0	0	0	0
		<i>P68</i>	NA	NA	NA	0	0	0	0	0	0	0	0
		<i>P69</i>	NA	NA	NA	0	0	0	0	0	0	0	0
	<i>Control</i>	<i>C2</i>	NA	NA	NA	0	0	0	0	0	0	0	0
		<i>C10</i>	NA	NA	NA	0	0	0	0	0	0	0	0
<i>Skin fibroblasts</i>	<i>Juvenile</i>	<i>P13</i>	NA	0	0	0	0	0	0	0	0	0	0
		<i>P36</i>	NA	0	0	0	0	0	0	0	0	0	0
		<i>P38</i>	0	0	0	0	0	0	0	0	0	0	0
	<i>Control</i>	<i>C2</i>	NA	0	0	0	0	0	0	0	0	0	0
		<i>C5</i>	0	0	0	0	0	0	0	0	0	0	0
		<i>C6</i>	0	0	0	0	0	0	0	0	0	0	0
		<i>C9</i>	0	0	0	0	0	0	0	0	0	0	0
		<i>C2</i>	0	0	0	0	0	0	0	0	0	0	0
		<i>C9</i>	0	0	0	0	0	0	0	0	0	0	0

Range 0 to 1 indicates 0 to 100% methylation. NA = not available

**Table S25.** DNA methylation analysis of CTCF2 in muscle and muscle-derived cells

	<i>CpG→</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>	<i>10</i>	<i>11</i>		
<i>Muscle</i>	<i>Juvenile</i>	<i>P13</i>	NA	NA	NA	0	0	0	0	0	0	0	0	
		<i>P36</i>	NA	NA	NA	0	0	0	0	0	0	0	0	
		<i>P37</i>	NA	NA	NA	0	0	0	0	0	0	0	0	
	<i>Adult</i>	<i>P38</i>	NA	NA	NA	0	0	0	0	0	0	0	0	
		<i>P68</i>	NA	NA	NA	0	0	0	0	0	0	0	0	
		<i>P56</i>	NA	NA	NA	0	0	0	0	0	0	0	0	
		<i>P67</i>	NA	NA	NA	0	0	0	0	0	0	0	0	
		<i>Late-onset</i>	<i>C4</i>	NA	NA	NA	0	0	0	0	0	0	0	0
			<i>C6</i>	NA	NA	NA	0	0	0	0	0	0	0	0
	<i>C9</i>		NA	NA	NA	0	0	0	0	0	0	0	0	
	<i>C10</i>		NA	NA	NA	0	0	0	0	0	0	0	0	
	<i>Myoblasts</i>	<i>Juvenile</i>	<i>P13</i>	NA	0	0	0	0	0	0	0	0	0	0
			<i>P36</i>	NA	0	0	0	0	0	0	0	0	0	0
<i>P37</i>			NA	0	0	0	0	0	0	0	0	0	0	
<i>Adult</i>		<i>P38</i>	NA	0	0	0	0	0	0	0	0	0	0	
		<i>P56</i>	NA	0	0	0	0	0	0	0	0	0	0	
		<i>C2</i>	0	0	0	0	0	0	0	0	0	0	0	
		<i>C3</i>	0	0	0	0	0	0	0	0	0	0	0	
<i>Late-onset</i>		<i>C2</i>	0	0	0	0	0	0	0	0	0	0	0	
		<i>C9</i>	0	0	0	0	0	0	0	0	0	0	0	
		<i>C10</i>	0	0	0	0	0	0	0	0	0	0	0	
<i>Myotubes</i>	<i>Juvenile</i>	<i>P13</i>	NA	0	0	0	0	0	0	0	0	0	0	
		<i>P36</i>	NA	0	0	0	0	0	0	0	0	0	0	
		<i>P37</i>	NA	0	0	0	0	0	0	0	0	0	0	
	<i>Adult</i>	<i>P38</i>	NA	0	0	0	0	0	0	0	0	0	0	
		<i>P56</i>	NA	0	0	0	0	0	0	0	0	0	0	
		<i>C2</i>	0	0	0	0	0	0	0	0	0	0	0	
		<i>C4</i>	0	0	0	0	0	0	0	0	0	0	0	
	<i>Late-onset</i>	<i>C2</i>	0	0	0	0	0	0	0	0	0	0	0	
		<i>C9</i>	0	0	0	0	0	0	0	0	0	0	0	
		<i>C6</i>	0	0	0	0	0	0	0	0	0	0	0	

Range 0 to 1 indicates 0 to 100% methylation. NA = not available



## **CHAPTER IV**





## The Biomarker Potential of miRNAs in Myotonic Dystrophy Type I

**Emma Koehorst**<sup>1</sup>, Alfonsina Ballester-Lopez<sup>1, 2</sup>, Virginia Arechavala-Gomez<sup>3, 4</sup>, Alicia Martínez-Piñeiro<sup>1, 5</sup>, Gisela Nogales-Gadea<sup>1, 2</sup>

### AFFILIATIONS.

- <sup>1</sup> Neuromuscular and Neuropediatric Research Group, Institut d'Investigació en Ciències de la Salut Germans Trias i Pujol, Campus Can Ruti, Universitat Autònoma de Barcelona, 08916 Badalona, Spain.
- <sup>2</sup> Centre for Biomedical Network Research on Rare Diseases (CIBERER), Instituto de Salud Carlos III, 28029 Madrid, Spain.
- <sup>3</sup> Neuromuscular Disorders Group, Biocruces Bizkaia Health Research Institute, 48903 Barakaldo, Spain.
- <sup>4</sup> Ikerbasque, Basque Foundation for Science, 48009 Bilbao, Spain.
- <sup>5</sup> Neuromuscular Pathology Unit, Neurology Service, Neuroscience Department, Hospital Universitari Germans Trias i Pujol, 08916 Badalona, Spain.

**J Clin Med. 2020 Dec 4;9(12):3939. doi: 10.3390/jcm9123939**

Available from: <https://www.mdpi.com/2077-0383/9/12/3939>



---

---

## SUMMARY OF THE RESULTS

---

---

Here we review the current knowledge on the potential of miRNAs as a biomarker for disease development in DM1. Due to the multi-systemic nature of DM1 and the different tissues involved, the advances will be discussed per tissue, and we will touch on the therapeutic potential of miRNAs in DM1.

myo-miRs are specifically expressed in skeletal muscle and cardiac muscle. They consist of miR-1, miR-133a/b, and miR-206, which are highly enriched in skeletal muscle, whereas miR-1 and miR-133a/b are also highly enriched in cardiac muscle. Skeletal muscle miRNA deregulation was first described by Gambardella and collaborators in 2010. Only miR-206 was found to be overexpressed in biopsies of DM1 patients and was localized in centralized nuclei and nuclear clumps in DM1 muscle sections, but predicted targets, such as Utrophin were unchanged. Perbellini and collaborators conducted a similar study, but found miR-1 and miR-335 to be upregulated and miR-29b/c and miR-33 to be downregulated. Predicted targets of miR-1 and miR-29, involved in muscle development, arrhythmia, splicing, and atrophy, were significantly upregulated in DM1 patients, indicating functional relevance and a possible contribution to DM1 pathology. The paper by Fernandez-Costa and collaborators used a *Drosophila* model, in which CTG expansions were introduced. They found 20 miRNAs to be differentially expressed. Only three of the deregulated miRNAs were preserved in humans, miR-1, miR-7, and miR-10, and their downregulation was also found in humans, and at least seven of their target genes were upregulated. In addition, they found that Muscleblind, the homolog of human MBNL1, is necessary for the regulation of miR-1 and miR-7 in *Drosophila* flies.

Information on miRNA patterns in the heart of DM1 patients is scarce, most likely due to the difficulty of sample collection. Rau and collaborators showed in 2010 that miR-1 is downregulated in DM1 patients compared to controls. Loss of miR-1 was due to MBNL1 and downstream targets of miR-1, which are responsible for intracardiomyocyte conductance, are upregulated in heart samples of DM1 patients. The second study, performed by Kalsotra and collaborators, screened a DM1 mouse model for over 500 miRNAs and identified 54 differentially expressed miRNAs, which were validated in DM1 heart tissue. Twenty were significantly downregulated in human DM1 heart tissue, including two myo-miRs (miR-1 and miR-133a). Pathway analysis revealed the loss of function of the myocyte enhancer factor-2 transcriptional network.

Although the above studies discussing muscle and heart, have given us valuable information regarding the involvement of miRNAs in DM1 pathology, due to the invasive nature of sample collection, their biomarker potential is limited. The field, therefore, shifted its focus on circulating miRNAs for biomarkers in DM1. The first study on circulating miRNAs in DM1 was performed by Perfetti and collaborators in 2014. In 24 plasma DM1 samples, 14 miRNAs were significantly different expressed out of a panel of 381 miRNAs. Validation in a

bigger cohort revealed nine differently expressed miRNAs. Pooling the nine miRNAs into a DM1-miRNA score resulted in an accurate discrimination between DM1 patients and controls. Interestingly, the DM1-miRNA score showed a negative correlation with global muscle strength. In a later study, these findings were validated in a bigger cohort (n=103). Koutsoulidou and collaborators performed a similar study in the sera of DM1 patients and found all four myo-miRs to be significantly increased in blood sera of twenty-three DM1 patients, and myo-miRs were distinguishable between patients and controls with Receiver Operating Characteristic curves. No correlation between the miRNA expression levels and DM1 severity were found. However, they found that the four miRNAs were increased in progressive DM1 patients compared to stable patients, and that the miRNAs could discriminate between progressive and stable DM1 patients. These findings were also validated in a bigger cohort (n=63). In contrast, Fernandez-Costa and collaborators reported no changes in serum miRNA profiles of DM1 patients.



Several correlations between miRNA expression levels and DM1 clinical phenotype have been made, and a recent paper has shown how miRNA expression levels can be a biomarker for rehabilitation. Interestingly, to date only one paper has compared serum and muscle miRNAs levels simultaneously. They found the levels of circulating miR-133a, miR-29b, and miR-33a to be increased in DM1 patients. However, in muscle tissue two myo-miRs (miR-1 and miR-133a) and miR29c were found to be downregulated, which indicates that circulating myo-miRs are not reflective of the situation in the muscle.

The search for a therapeutic, targeting miRNAs in DM1, has focused mainly on miRNAs associated with the two splice factors involved in DM1 pathology, namely MBNL1 and CELF1. Four different miRNAs (miR-1, miR-30-5p, miR-23b, and miR-218) associated with the splice factor MBNL1 have been proposed as therapeutic targets. For example, a study showed that introducing a miR-30-5p mimic into C2C12 muscle cells downregulated the expression of MBNL and deregulated downstream targets of MBNL, including Trim55 and IR. Cerro-Herreros and collaborators showed that miR-23b and miR-218 anti-miR treatment in DM1 myoblasts and DM1 mice increased MBNL1 levels, and in the latter, this was linked to improvements in spliceopathy profile, histopathological signs, and functional myotonia without toxicity. The therapeutic effect was dose-dependent and has long-lasting effects. For the splice factor CELF1, the focus has been primarily on miR-206, which may modulate CELF1 overexpression. The introduction of miR-206 mimics into cells overexpressing CELF1 inhibited CELF1 expression and improved the myoblast fusion index and myotube area.

Taken together, in the past two decades, extensive research has been conducted in the miRNA expression profiles of DM1 patients and their biomarker potential. Several expression profiles have been found to be able to distinguish between DM1 patients and healthy subjects, and even a link has been made between progressive and non-progressive muscle wasting in DM1 patients. Moreover, the deregulated miRNA expression profiles in DM1 have led to the identification of novel therapeutic targets, with promising results.

Review

# The Biomarker Potential of miRNAs in Myotonic Dystrophy Type I

Emma Koehorst<sup>1</sup>, Alfonsina Ballester-Lopez<sup>1,2</sup>, Virginia Arechavala-Gomez<sup>3,4</sup> ,  
 Alicia Martínez-Piñero<sup>1,5</sup> and Gisela Nogales-Gadea<sup>1,2,\*</sup> 

<sup>1</sup> Neuromuscular and Neuropediatric Research Group, Institut d'Investigació en Ciències de la Salut Germans Trias i Pujol, Campus Can Ruti, Universitat Autònoma de Barcelona, 08916 Badalona, Spain; ekoehorst@igtp.cat (E.K.); aballester@igtp.cat (A.B.-L.); amartinezp.germanstrias@gencat.cat (A.M.-P.)

<sup>2</sup> Centre for Biomedical Network Research on Rare Diseases (CIBERER), Instituto de Salud Carlos III, 28029 Madrid, Spain

<sup>3</sup> Neuromuscular Disorders Group, Biocruces Bizkaia Health Research Institute, 48903 Barakaldo, Spain; virginia.arechavalagomez@osakidetza.eus

<sup>4</sup> Ikerbasque, Basque Foundation for Science, 48009 Bilbao, Spain

<sup>5</sup> Neuromuscular Pathology Unit, Neurology Service, Neuroscience Department, Hospital Universitari Germans Trias i Pujol, 08916 Badalona, Spain

\* Correspondence: gnogales@igtp.cat; Tel.: +34-93-4978684

Received: 21 October 2020; Accepted: 1 December 2020; Published: 4 December 2020



**Abstract:** MicroRNAs (miRNAs) are mostly known for their gene regulation properties, but they also play an important role in intercellular signaling. This means that they can be found in bodily fluids, giving them excellent biomarker potential. Myotonic Dystrophy type I (DM1) is the most frequent autosomal dominant muscle dystrophy in adults, with an estimated prevalence of 1:8000. DM1 symptoms include muscle weakness, myotonia, respiratory failure, cardiac conduction defects, cataracts, and endocrine disturbances. Patients display heterogeneity in both age of onset and disease manifestation. No treatment or cure currently exists for DM1, which shows the necessity for a biomarker that can predict disease progression, providing the opportunity to implement preventative measures before symptoms arise. In the past two decades, extensive research has been conducted in the miRNA expression profiles of DM1 patients and their biomarker potential. Here we review the current state of the field with a tissue-specific focus, given the multi-systemic nature of DM1 and the intracellular signaling role of miRNAs.

**Keywords:** myotonic dystrophies; miRNAs; biomarkers; therapeutics

## 1. Introduction

MicroRNAs (miRNAs) are small, single-stranded RNAs about 22 nucleotides long that regulate gene expression by either inhibiting translation or promoting degradation of their target mRNAs [1]. It is estimated that the human genome encodes over a thousand miRNAs, which can target dozens of mRNAs, and every individual mRNA can be targeted by several miRNAs [2]. miRNAs are estimated to regulate about one-third of human protein-coding genes [3]. The majority of miRNAs result from RNA polymerase II transcription, which yields long primary miRNA (pri-miRNA) transcripts. Pri-miRNAs are trimmed in the nucleus by the RNase Drosha, yielding premature hair-looped miRNA of  $\pm 70$  nucleotides. These pre-mRNAs are transported to the cytoplasm where they are further processed by the RNase Dicer, resulting in mature miRNA. Mature miRNAs are then incorporated into the RNA-induced silencing complex (RISC), where the miRNA strand anneals to the 3' untranslated regions (UTRs) of target mRNAs, leading to the degradation or translation inhibition of mRNAs and subsequent protein repression [3]. miRNAs are mostly known for their gene regulation properties.

However, they also play an important role in intercellular signaling and can therefore be found abundantly in bodily fluids, including blood and urine [4]. There are several ways in which miRNAs can reach the bloodstream from the tissues, including lipid or lipoprotein complexes and extracellular vesicles (EVs), but they have also been found as free-floating complexes with Argonaute proteins [5]. This makes them excellent biomarker candidates, since one of the most important properties of a good biomarker is easy accessibility. Furthermore, their detection is easily achieved with low-cost techniques, such as quantitative reverse transcription PCR. In addition, miRNA expression profiles are different in the diseased state and in their tissue-specific expression patterns, further strengthening their biomarker potential, and extensive research has been conducted on this discovery. The most studied field is oncology, where miRNAs encapsulated in exosomes are suggested to function as biomarkers in the diagnosis and prognosis of several cancers, including breast cancer, prostate cancer, and colorectal cancer [6]. In addition, miRNAs have been found to play an essential role as biomarkers in numerous other diseases, such as Alzheimer's disease [7], epilepsy [8], sepsis [9], and in neurodegenerative and neuromuscular disorders [10,11].

Myotonic Dystrophy type I (DM1) is an autosomal dominant muscle dystrophy, with a multi-systemic nature and an estimated prevalence of 1:8000 [12]. Patients display a core set of symptoms, which include muscle weakness, myotonia, respiratory failure, cardiac conduction defects, cataracts, and endocrine disturbances. The DM1 phenotype can be divided into five clinical categories, each with distinct clinical features: congenital, childhood-onset, juvenile-onset, adult-onset, and ate-onset DM1. In congenital DM1, symptoms arise at birth or during the first month of life and include hypotonia, respiratory failure, feeding difficulties, failure to thrive, and clubfoot deformities. Mortality is especially high in the perinatal period due to respiratory failure [13]. The diagnosis of childhood-onset DM1 is often missed due to uncharacteristic symptoms. Affected children show cognitive and learning abnormalities after the first year. None of the 'classic' DM1 symptoms, such as muscle myopathy and myotonia, is present at first, but they often develop in adulthood [14]. Juvenile DM1 has an onset between 10 and 20 years. It spans the continuum between childhood and adult-onset and has some overlapping features with both types. Nevertheless, they differ from childhood-onset DM1 by the more pronounced muscle involvement and from the adult-onset in the often underdiagnosed cognitive impairment and earlier cardiac and respiratory problems [15]. The classic/adult-onset is the most prevalent DM1 phenotype and arises typically around the second or third decade of life. Core features are progressive muscle weakness and myotonia with preferential involvement of the cranial, trunk, and distal limb muscles and early-onset cataracts [16,17]. In addition, cardiac conduction defects are common in adult-onset DM1, which is a major contributor to increased mortality and sudden death in patients [18]. Late-onset or mild DM1 patients usually start to show symptoms after the age of forty. Symptoms include mild muscle weakness, premature cataract, and balding. Not only is the disease heterogeneous in disease onset, within the different subtypes much heterogeneity exists as well in symptom development. Symptom manifestation can range from mild muscle weakness with myotonia to loss of ambulation with cardiac conduction defects and severe cataracts. At the point of diagnosis, it is unclear which subset of symptoms will develop over time.

DM1 is caused by a Cytosine-Thymine-Guanine (CTG) expansion in the in the 3' untranslated region of the myotonic dystrophy protein kinase gene (*DMPK*). Unaffected individuals carry 5–35 CTG repeats, whereas the length of the CTG expansion in patients can range from fifty to thousands of CTGs and has been associated with age of symptom onset and disease severity [19,20]. DM1 is considered an RNA gain-of-function disorder, in which expanded transcripts accumulate as intranuclear RNA foci. These foci can sequester muscleblind-like (MBNL) proteins, which subsequently lead to diminished activity and downstream deregulation of the alternative splicing of several genes [21,22]. In addition, the RNA foci cause hyper-phosphorylation and thus upregulation of the CUG-BP and ETR-3-like factors family member 1 (CELF1) via several signaling kinases [22,23]. CELF1 is also a splice factor, and its inappropriate phosphorylation results in deregulation of downstream target genes. Disruption of the alternative splicing of several genes can be directly correlated to symptoms arising in DM1.

For example, mis-splicing of the *ClC-1* chloride channel leads to reduced chloride conductance in muscle fibers, which is known to produce myotonia [24], and alternative splicing of cardiac troponin T and the insulin receptor is correlated with cardiac defects and diabetes, respectively [21]. However, these spliceopathies cannot explain the full multi-systemic picture found in DM1, and the discoveries of bidirectional transcription of the *DMPK* gene, aberrant DNA methylation, repeat-associated non-ATG translation, and miRNA deregulation have added to the vast complexity of DM1 pathology [25–28].

In DM1, the need for a biomarker that can predict disease outcomes is essential, especially for cardiac and respiratory complications, to improve patient survival. Symptoms and disease progression are highly variable among patients, and life expectancy is significantly decreased, primarily due to respiratory failure and cardiac events [29]. This high variability and the multi-systemic nature of the disease present a unique challenge in terms of disease management. Currently, disease management is mostly focused on symptomatic treatment and preserving function and independence. Current assessment strategies include genetic testing, electromyography, skeletal muscle histopathology, and magnetic resonance. The only circulating marker used to date is creatine kinase (CK), but this is not elevated in all patients and lacks disease specificity, since it is altered in most disorders involving skeletal muscle damage [30], as well as intense physical activity [31]. Other suggested biomarkers are the alternative splicing changes observed in the DM1 skeletal muscle [32]. However, monitoring them would require several invasive muscle biopsies. miRNAs have the potential to help estimate the degree and type of involvement of symptoms before they appear as well as to apply the appropriate preventive measures. For example, several studies have indicated that miRNAs correlate to muscle strength and disease stage in DM1 [33,34]. In addition to disease management, miRNAs could be used as biomarkers for future clinical trials. Currently, outcome measures consist mostly of clinical and functional ones, such as cognitive function assessment by questionnaires, patient-reported outcomes such as DM1-activ, muscle testing by the 6 min walk test, grip strength, Muscular Impairment Rating Scale (MIRS), and myotonia degree [35]. These can be, however, influenced by several confounding factors. For example, the 6 min walk test is not only influenced by muscle strength but also by factors such as the level of fatigue and motivation [36]. These tests also might lack sensitivity to subtle changes happening in the rather short time frame of most clinical trials.

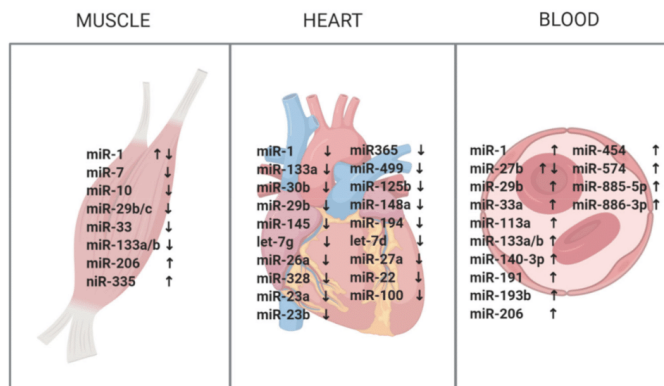
Here we review the current knowledge on the potential of miRNAs as a biomarker for disease development in DM1. Due to the multi-systemic nature of DM1 and the different tissues involved, the advances will be discussed per tissue, and we will touch on the therapeutic potential of miRNAs in DM1.

## 2. miRNAs in Skeletal Muscle

Several endogenous miRNAs are ubiquitously expressed, such as let-7, but a subset of miRNAs is only expressed in certain tissues. An example is the muscle-specific miRNAs (myo-miRs), which are expressed in skeletal muscle and cardiac muscle. They consist of miR-1, miR-133a/b, and miR-206, which are highly enriched in skeletal muscle, whereas miR-1 and miR-133a/b are also highly enriched in cardiac muscle. miR-1 plays an important role in skeletal muscle growth and differentiation by regulating the serum response factor (SRF) and myocyte enhancer factor-2 (MEF2) [37]. Both activate skeletal muscle gene expression in collaboration with myogenic basic helix loop helix proteins. miR-133 was initially thought to drive proliferation by repressing the expression of SRF, indicating an opposing role from miR-1 [38]. However, recent reports suggest a similar regulating function to miR-1, promoting differentiation in a similar way [39]. Both miR-1 and miR-133 are also involved in the modulation of electrical conductance in the heart [40]. miR-206, on the other hand, is solely expressed in skeletal muscle and promotes myoblast differentiation [41]. Later on, miR-208a/b, miR-499, and miR-486 were added to the list of myo-miRs. These new members are striated muscle-specific, and miR-208a is found only in cardiac tissue. Both miR-208 and miR-499 are encoded in slow-twitch, type I myosin genes and are required to establish the slow-twitch fiber phenotype [42]. miR-486 is potentially involved in muscle growth and homeostasis by regulating the Phosphoinositide 3-kinase/Protein kinase B signaling pathway [43]. Myo-miRs have been suggested to play a role in several diseases, including cardiac hypertrophy, cardiac arrhythmias, and muscular disorders [44–47].



Since DM1 is a muscle dystrophy first, and therefore the most affected tissue is the muscle, naturally the first studies on miRNA dysregulation as a potential biomarker for disease were conducted in the muscle (Figure 1, Table 1). It was first studied by Gambardella et al. in 2010 [48]. They focused on myo-miRs, in addition to miR-103 and miR-107, which were proposed to be attractive candidates for binding *DMPK* mRNA [49]. More specifically, a computational analysis revealed that miR-103 and miR-107 would repress the *DMPK* wild-type allele by 15%, and that repression of mRNA by CTG repeat-binding miRNAs would increase with the number of CTGs in the 3'UTR. They proposed a model in which CTG repeat-binding miRNAs, such as miR107 and miR-103, preferentially bind to mutated *DMPK* mRNA, which could have a miRNA-leaching effect and, therefore, implications for DM1 pathology [49]. Only miR-206 was found to be overexpressed in biopsies of DM1 patients. However, the predicted target genes of miR-206, such as Utrophin, were not altered. A potential explanation can be the presence of multiple myofibers in skeletal muscle. For example, miR-206 is known to be highly enriched in regenerating fibers [50], which constitutes only a small portion of the bulk tissue measured. Although miR-206 shows a pronounced increase due to the addition of regenerating fibers, this does not mean that its target genes are equally affected, and minor changes might not be detected in bulk tissue. Interestingly, miR-206 was localized in centralized nuclei and nuclear clumps in DM1 muscle sections, both hallmarks of DM1 histopathology. In the following year, Perbellini et al. [51] conducted a similar study in muscle biopsies from 15 DM1 patients, but with very different results. They broadened their study to 24 miRNA candidates, which have important regulatory roles in skeletal muscle or were found to be dysregulated in Duchenne muscular dystrophy. They found miR-1 and miR-335 to be upregulated and miR-29b/c and miR-33 to be downregulated. Predicted targets involved in muscle development (MEF2a, MET, GATA6, and HAND2), arrhythmia (KCNE1, KCNJ2, CALM2), splicing (SFRS9), and atrophy (DAG1, DIABLO, RET, TRIM63, TGFβ3) of miR-1 and miR-29 were significantly upregulated in DM1 patients, indicating functional relevance and a possible contribution to DM1 pathology. In addition, miR-1 was vastly present in the internal nuclei and in nuclear clumps of DM1 myofibers, similar to the miR-206 distribution found by Gambardella et al. [48]. Although expression levels were found not to be dysregulated, the cellular distribution of miR-206 and miR-133b was severely altered in patients compared to controls, in a similar way as miR-1.



**Figure 1.** A summary of the studied miRNAs in skeletal muscle, heart, and blood. ↑ = upregulation, ↓ = downregulation. Created with [BioRender.com](https://www.biorender.com).

**Table 1.** Overview of the differentially expressed miRNAs analyzed in muscle biopsies.

Study	miRNAs	Muscle Biopsy	Sample Size	Normalization	Targets	Associated Pathological Signs and Mechanisms	
Gambardella 2010 [48]	miR-206	↑	7 DM1 4 ctrls	Hsa-let-7a	MEF2a * <sub>1</sub> MET * <sub>1</sub> GATA6 * <sub>1</sub> HAND2 * <sub>1</sub>	muscle development <sup>1</sup>	
Perbellini 2011 [51]	miR-1 *	↑	15 DM1 14 ctrls	miR-16	KCNE1 * <sub>2</sub>	Arrhythmia <sup>2</sup>	
	miR-29b/c **	↓			CALM2 * <sub>2</sub>		
	miR-335	↑			SFRS9 * <sub>3</sub>		Splicing <sup>3</sup>
	miR-33	↑			DAG1 ** <sub>4</sub> DIABLO ** <sub>4</sub> RET **TRIM63 ** <sub>4</sub> TGFB3 * <sub>4</sub>		Atrophy <sup>4</sup>
Fernandez-Costa 2012 [52]	miR-1 ***	↓	5 DM1 3 ctrls	Sno-RNA RNU48	SOD1 *** <sub>5</sub>	Free radical removal <sup>5</sup>	
	miR-7 ****	↓			SMARCA4 *** <sub>6</sub>		Transcription <sup>6</sup>
	miR-10 *****	↓			NET 1 ** <sub>7</sub>		Apoptosis and signal transduction <sup>7</sup>
		↓			CTSB *** <sub>8</sub> AIC4 *** <sub>9</sub> VCL *** <sub>10</sub> UBE11 **** <sub>11</sub>		Proteolysis <sup>8</sup> Autophagy <sup>9</sup> Cytoskeleton <sup>10</sup> Protein localization/activity regulation <sup>11</sup>
Fritegatto 2017 [53]	miR-1	↓	12 DM1 6 ctrls	Hsa-let-7a			
	miR-133a/b miR206	↑					
Ambrose 2017 [52]	miR-1	↓	9 DM1 9 ctrls	Hsa-let-7a			
	miR-133a miR-29c	↓			Biceps branchii Gastrocnemius Deltoid		

↑ = upregulation, ↓ = downregulation, DM1 = DM1 patients, ctrls = healthy subjects, colors indicate to which miRNA the targets belong. Normalization stands for the miRNA, either endogenous or spike-in, that has been used to normalize the differentially expressed miRNAs. \*, \*\*, \*\*\*, \*\*\*\*, \*\*\*\*\* indicate to which miRNA the targets belong. <sup>1-11</sup> numbers indicates which Associated Pathological Signs and Mechanisms belong to which target.

The paper by Fernandez-Costa et al. [52] adopted a very different approach to studying miRNA dysregulation of DM1 by using a *Drosophila* model, in which CTG expansions were introduced. They found 20 miRNAs to be differentially expressed, of which 19 were downregulated and 1 was upregulated. Only three of those were preserved in humans, miR-1, miR-7, and miR-10, which underwent a validation study in muscle biopsies from five DM1 patients. The downregulation was preserved in humans, and at least seven of their target genes were upregulated. For miR-1, these were the antioxidant enzyme superoxide dismutase 1 (SOD1), the transcriptional regulator SWI/SNF-related, matrix-associated, actin-dependent regulator of chromatin, subfamily A, member 4 (SMARCA4), and the nucleotide exchange factor Neuroepithelial Cell Transforming 1 (NET1). miR-7-upregulated targets included amyloid precursor protein secretase Cathepsin B, autophagy regulator cysteine protease (ATG4), and cytoskeletal protein Vinculin (VCL), and miR-10 caused an upregulation of Ubiquitin-activating enzyme E1 (UBE11), a member of the small ubiquitin-like modifier (SUMO)ylation machinery. This suggests that miRNA dysregulation in DM1 is involved in a wide range of pathological mechanisms. The finding of miR-1 downregulation is in contrast to the results found by Perbellini et al. [51], since they found miR-1 to be upregulated. To further elucidate the mechanism involved in the dysregulation of these miRNAs, they performed several experiments on the pri-miRNA precursors of these miRNAs and found that expanded CTG repeats decreased the levels of the primary precursor of miR-7 in *Drosophila* flies, and this was also observed in skeletal muscle biopsies of DM1 patients. In addition, they found that Muscleblind, the homolog of human MBNL1, is necessary for the regulation of miR-1 and miR-7 in *Drosophila* flies. MBNL1 protein in humans is trapped by RNA foci in DM1 and has been previously described to participate in the biogenesis of miR-1 [1]. Although it will have to be validated in humans, their results in flies suggest that the depletion of MBNL1 in DM1 can be an explanation for the miR-1 downregulation in skeletal muscle biopsies of DM1 patients. They also made the first link to the physiological relevance of miRNA dysregulation, albeit only in flies. Overexpression of miR-10 in the DM1 *drosophila* model increased their lifespan.

The vast difference in results between the studies, which at first glance look very similar, might be explained by the use of different muscle biopsies. The first study was conducted in biopsy material from the vastus lateralis, whereas the second study used material coming from the biceps brachii, and the last study used a mixture of material from the vastus lateralis, biceps brachii, and deltoid. Since in DM1, in general, the observed muscle weakness is prominently distal, and only in later stages of the diseases are the proximal muscles, such as the bicep brachii and the vastus lateralis, involved, it could be that the stage of each individual patient at the moment of biopsy can influence the obtained results. In addition, miRNA expression levels are studied by quantitative PCR (qPCR), where expression levels are normalized against an endogenous miRNA of which the levels are known to be stable or by using a spike-in miRNA. Total miRNA content has been shown to be an experimental variable and adds a systematic bias in miRNA quantification [54]. No gold standard for this type of normalization exists, and groups often use different approaches, decreasing the comparability of their studies. In the case of above studies, all three studies opted for a different normalization approach, a potential explanation for their contradictory results. However, Fritegotto et al. [53] recently published a study in which they adopted a very similar experimental design to Gambardella et al. [48]. They used the same muscle biopsy material and the same housekeeping gene for normalization, but their results match only in part. miR-206 is upregulated, which is in consensus with what Gambardella et al. [48] found in 2010. However, they found additionally that miR-1 and miR-133a/b were significantly decreased. This result is in consensus with the findings of Fernandez-Costa [52], but in vast contrast to the study by Perbellini et al. [51], since they found miR-1 to be upregulated. A possible explanation can be that Perbellini et al. [51] used control subjects that were admitted for suspected neuromuscular disorders. This can skew their results since these 'control' subjects may have an underlying muscular disorder. Interestingly, this study by Fritegotto et al. [53] was the first study to try to correlate their miRNA dysregulation findings to DM1 pathology. A histopathological score was assigned to each muscle biopsy. However, the histopathological score and disease duration did not correlate significantly to miRNA expression levels. They did observe that the lowest levels of miR-1 and

miR133a/b and the highest levels of miR-206 were found in patients with the most severe histopathological score or longest disease duration. Thus, even though Fritegotto et al. [53] and Gambardella et al. [48] adopted the same experimental procedure, their results differ. It might be due to the use of different clinical subtypes of DM1. In Fritegotto et al.'s study [53], the clinical subtypes ranged from childhood to adult onset. Unfortunately, Gambardella et al. [48] only disclosed the current age range and not the age of onset, so the clinical subtype is unknown, but it might be relevant. Childhood-onset DM1 patients display a different array of symptoms compared to adult-onset DM1, suggesting that different pathological mechanisms are at play.

Although studies in skeletal muscle have given us valuable information on miRNA dysregulation in DM1, it can be argued that it does not have the best biomarker potential since obtaining a muscle biopsy is an invasive procedure, especially if you consider that the diagnosis of DM1 can be done with a simple genetic test using a blood sample, which makes muscle biopsies in DM1 scarce and, most of the time, unnecessary. The studies did give us important information on potential targets for therapy development, which will be discussed further on in the review.

### 3. miRNAs in the Heart

Information on miRNA patterns in the heart of DM1 patients is scarce, most likely due to the difficulty of sample collection. To date, two papers have looked into the involvement of miRNAs in cardiac pathology in DM1 (Figure 1, Table 2). First of all, Rau et al. [55] showed in 2010 that miR-1 is downregulated in DM1 patients compared to controls, and that downstream targets of miR-1, the Voltage-dependent L-type calcium channel subunit alpha-1C (CACNA1C) and connexin 43 (GJA1), which are responsible for intracardiomyocyte conductance, are upregulated in heart samples of DM1 patients. The loss of miR-1 was due to MBNL1, which binds to the pre-miR-1, blocking the binding site for LIN-28 and subsequently disrupting the processing of pre-miR-1. The second study, performed by Kalsotra et al. [56], screened a DM1 mouse model for over 500 miRNAs and identified 54 differentially expressed miRNAs. To translate these findings to humans, 22 miRNAs were tested in heart tissue of eight DM1 samples and four controls. Twenty were significantly downregulated in human DM1 heart tissue, including two myo-miRs (miR-1 and miR-133a). Pathway analysis of the misregulated miRNAs revealed the loss of function of the myocyte enhancer factor-2 (Mef2) transcriptional network. In addition, several of the affected miRNAs found in the study have been previously described to produce arrhythmias or fibrotic changes. For example, miR-1, which was identified in both studies, is known to regulate gap junction proteins and cardiac channels, and a reduction in its expression greater than 50% may contribute to the conduction defects found in DM1 [55]. Two other downregulated miRNAs worth mentioning are miR-23a/b, of which postnatal upregulation has been shown to result in CELF1 downregulation [57]. Additionally, upon introduction of expanded CUG repeat RNA into a heart-specific and inducible DM1 mouse model, postnatal upregulation of miR-23a/b is reversed, and CELF1 is subsequently upregulated, proposing an additional mechanism for CELF1 upregulation in DM1.

**Table 2.** Overview of the differentially expressed miRNAs analyzed in heart biopsies.

Study	miRNAs	Heart	Sample Size	Normalization	Targets	Associated Pathological Signs
Rau 2011 [55]	miR-1	↓	5 DM1 8 ctrls	U6 snRNA	CACNA1C <sup>1</sup> GJA1 <sup>2</sup>	Arrhythmias <sup>1</sup> Conduction <sup>2</sup>
	miR-1*	↓				
	miR-133**	↓				
	miR-30b**	↓				
	miR-29b**	↓				
	miR-145	↓				
	let-7g	↓				
	miR-26a	↓				
	miR-328	↓				
	miR-23a	↓				
	miR-23b	↓				
	miR365	↓				
	miR-499	↓				
	miR-125b	↓				
miR-148a	↓					
miR-194	↓					
let-7d	↓					
miR-27a	↓					
miR-22	↓					
miR-100	↓					
Kalsotra 2014 [56]		unknown	8 DM1 4 ctrls	MammU6		Conduction* Fibrosis**

↓ = upregulation, ↓ = downregulation. DM1 = DM1 patients, ctrls = healthy subjects, colors indicate to which miRNA the proposed disease involvement belongs. Normalization stands for the miRNA, either endogenous or spike-in, that has been used to normalize the differentially expressed miRNAs. \*, \*\* indicate to which miRNA the associated pathological signs belong. <sup>1,2</sup> numbers indicates which Associated Pathological Signs belong to which target.

#### 4. miRNAs in Blood and Serum

Although the above studies discussing tissues, muscle, and heart have given us valuable information regarding the involvement of miRNAs in DM1 pathology, due to the invasive nature of sample collection, their biomarker potential is limited. The field, therefore, shifted its focus on circulating miRNAs for biomarkers in DM1 (Figure 1, Table 3). Circulating miRNAs diffuse into the bloodstream from several of the affected tissues, making them excellent candidates as biomarkers. In addition, drawing blood is a simple and much less invasive procedure than obtaining a muscle biopsy.

The first study on circulating miRNAs in DM1 was performed by Perfetti et al. in 2014 [33]. They performed a miRNA panel of 381 miRNAs in twenty-four plasma samples of DM1 patients and twenty-six controls. The fourteen miRNAs that showed a significantly different expression level compared to controls were subjected to validation in a bigger cohort ( $n = 36$ ). In this bigger cohort, nine miRNAs were differentially expressed: miR-133a, miR-193b, miR-191, miR-454, miR-574, miR-885-5p, and miR-886-3p were increased, whereas miR-27b levels were decreased. Pooling the nine miRNAs into a DM1-miRNA score resulted in an accurate discrimination between DM1 patients and controls. miR-133a alone was also able to discriminate between the two populations. Interestingly, both the DM1-miRNA score and miR-133a showed a negative correlation with global muscle strength, assessed by the Medical Research Council (MRC) muscle scale, and significant increases in median miR-133a and DM1-miRNA scores were seen in higher MIRS classes. The following year, Koutsoulidou et al. performed a very similar study in the sera of DM1 patients [34]. The biggest difference is they decided to focus solely on myo-miRNAs and studied the expressions of miR-1, miR-133a, miR133b, and miR-206. miR-133a is the only myo-miRNA that was previously found to be differentially expressed [33]. However, in this study all four myo-miRs were found to be significantly increased in blood sera of twenty-three DM1 patients, and myo-miRs were distinguishable between patients and controls with Receiver Operating Characteristic (ROC) curves ranging from 0.94 to 0.97, or when averaged into a single value with a score of 0.98. In contrast to Perfetti et al. [33], they were unable to find any correlation between the miRNA expression levels and DM1 severity. However, when the groups were stratified, based not on their clinical picture and MRC muscle scale but on their progression in muscle wasting, they found that the four miRNAs were increased in progressive DM1 patients compared to stable patients, and that the miRNAs could discriminate between progressive and stable DM1 patients. Progressive vs. stable was based on whether the patients showed a change in the MRC scale in the past two years. This is the first longitudinal correlation made and a first hint at the real biomarker potential of these miRNAs.

**Table 3.** Overview of the differentially expressed miRNAs analyzed in blood.

Study	miRNAs	Blood	Sample Size	Normalization	Associated Clinical Characteristics
Perfetti 2014 [33]	miR-113a	Plasma	36 DM1 36 ctrls	Cel-miR-39 miR-17-5p miR-106a	Negative correlation to MRC score Positive correlation to MIRS scale
	miR-193b				
	miR-191				
	miR-454				
	miR-574				
	miR-885-5p				
Koutsoulidou 2015 [34]	miR-886-3p	Serum	23 DM1 23 ctrls	miR-16	Increased levels in progressive muscle wasting compared to stable muscle wasting
	miR-27b				
	miR-1				
	miR-133a/b				
	miR-206				
	miR-1				
Perfetti 2016 [38]	miR-133a/b	Plasma	103 DM1 111 ctrls	Cel-miR-39 miR-17-5p miR-106a	Negative correlation to muscle strength Non-significant correlation to CK levels
	miR-206				
	miR-140-3p				
	miR-454				
	miR-574				
	miR-27b				
Koutsoulidou 2017 [59]	miR-1	Serum	63 DM1 63 ctrls	miR-16	Increased levels in progressive muscle wasting compared to stable muscle wasting
	miR-133a/b				
	miR-206				
	miR-113a				
	miR-193b				
	miR-191				
Pegoraro 2020 [51]	miR-574	Serum	9 DM1 7 ctrls	miR-39-3p C. elegans	After a six-week exercise training the 4 miRNAs significantly decreased in parallel with improved muscle function
	miR-885-5p				
	miR-886-3p				
	miR-27b				
	miR-1				
	miR-133a/b				
Ambrose 2017 [52]	miR-206	Whole blood	10 DM1 10 ctrls	Hsa-miR-183	
	miR-133a				
	miR-29b				

↑ = upregulation, ↓ = downregulation, DM1 = DM1 patients, ctrls = healthy subjects, \* only in a subset of patients. Normalization stands for the miRNA, either endogenous or spike-in, that has been used to normalize the differentially expressed miRNAs. General associated clinical characteristics are indicated per study.

Both studies mentioned above repeated their analysis in bigger cohorts: Perfetti et al. [58] chose to include the three additional miRNAs found by Koutsoulidou et al. [34] into this validation study ( $n = 103$ ). Out of these twelve combined miRNAs, eight were found to be differentially expressed: miR-1, miR-133a/b, miR-206, miR-140-3p, miR-454, and miR-574 were increased, and miR-27b was decreased. This means that four out of nine of the originally found were not validated in this bigger cohort. Notably, the myo-miRs were able to differentiate between patients and controls, but better results were obtained when all the miRNAs were combined into a DM1-miRNA score, when myo-miRs scores were pooled, or when levels of miR-133a and miR-133b were averaged into a miR-133a/b score. This time they were again able to find a significant, negative correlation with muscle strength and a weak, direct correlation with CK values. An interesting point raised was that miR-133b was significantly higher in female DM1 patients compared to DM1 males, but this difference between gender was not found in healthy subjects or in the original study. Of note, they included DM2 patients in this study as well to see whether the findings in DM1 were translatable to DM2, and all except miR-27b were found to be deregulated in thirty DM2 patients. Koutsoulidou et al. focused first on the four myo-miRs previously reported [34] and validated their differential expression in a cohort of 63 DM1 patients. The previously reported increased levels in progressive DM1 patients compared to stable DM1 patients were also validated [59]. In addition to validating their previous results, they also considered the eight additional miRNAs found by Perfetti et al. in plasma [33]. Six out of eight were found to be significantly increased, whereas no differences were found in miR-454 expression, and miR-27b, previously reported to be significantly decreased in DM1 patients, was shown to be significantly increased. A novel finding was the encapsulation of these myo-miRs in EVs: it was reported that the majority of the differentially expressed myo-miRs in DM1 patients could be found in EVs. The levels of the four myo-miRs were significantly higher in EVs isolated from sera of the DM1 patients compared to controls and seem to discriminate between progressive and stable DM1 patients. However, caution should be taken with regards to these results. It has been shown that polymer-based methods to precipitate EVs (such as the Exoquick™ Exosome Precipitation solution used in this study) do not exclusively precipitate EVs, and co-isolation of other molecules, including RNA-Protein complexes, is a possibility [60]. Especially in the case of miRNAs, it has been shown that these kits may preferentially precipitate non-exosomal miRNAs [61]. This means that the miRNAs found in this precipitate may not exclusively originate from EVs or in fact might primarily be non-exosomal.

Although some discrepancies can be found between the studies of the two groups above, they both show a clear miRNA deregulation pattern in the sera or plasma of DM1 patients. This is in vast contrast to another study by Fernandez-Costa et al. [62] in 2016, which was unable to validate six serum miRNAs as DM1 biomarkers. A panel of 175 known serum miRNAs was tested in ten DM1 patients and controls, of which six potential candidates were chosen, but they were unable to validate these candidates by qPCR. The strongest two candidates, miR-21 and miR-130a, were further tested in a bigger cohort of twenty-one DM1 patients, but no differences were found.

Several correlations between miRNA expression levels and DM1 clinical phenotype have been made, and a recent paper has shown how miRNA expression levels can be a biomarker for rehabilitation [63]. After a six-week exercise rehabilitation training, a significant decrease was found in four myo-miRs, namely miR-1, miR-133a/b, and miR-206, in parallel with an improvement in muscle function. Of note, when the four miRNAs were first analyzed for differences between patients and controls, only miR-1 and miR-206 were significantly increased, while miR133a/b were upregulated only in a subset of the patients. The study does, however, show us the potential of myo-miRs in serum as a rehabilitation biomarker, which would be a nice addition to the now used clinical outcomes, which often lack sensitivity and reproducibility and can be influenced by the multi-systemic nature of the disease.

To date, only one study has looked at the serum and muscle miRNA levels simultaneously [64]. Since the myo-miRs in blood are thought to be a representation of the miRNA levels in muscle, one study used a slightly different approach in that they used whole blood instead of plasma or serum. They found the levels of circulating miR-133a, miR-29b, and miR-33a to be increased in DM1



patients. However, in muscle tissue two myo-miRs (miR-1 and miR-133a) and miR29c were found to be downregulated, which indicates that circulating myo-miRs are not reflective of the situation in the muscle. The increase in miR-133a was previously described by several studies [58,59,63], whereas the other two miRNAs are novel findings. Their results in muscle are similar to the findings of two other studies, although the downregulation of miR-29c has not been reported before [52,53].

Similar to the findings in muscle, several studies focusing on miRNA expression levels in blood show no homogeneous findings. Again, differences in sample sizes and normalization methods chosen could be to blame. For example, Perfetti et al. [33,58] decided to use a combination of three different normalizers, Cel-miR-39, miR-17-5p, and miR-106a, whereas Koutsoulidou et al. [34,59] used miR-16 to normalize their expression profiles. The study by Fernandez-Costa et al. [62], which failed to validate any of the previous findings, used the mean Ct value of miRNAs detected for normalization. Both Perfetti et al. [58] and Koutsoulidou et al. [59] decided to validate their results in a bigger cohort, and both found different results, showing the importance and influence of sample size on experiments. Another important difference between the studies is the use of different blood components (i.e., serum, plasma, or whole blood). It has been shown that the miRNA expression levels differ between serum and plasma, for example due to the activation and release of miRNAs from platelets during collection of plasma [65].

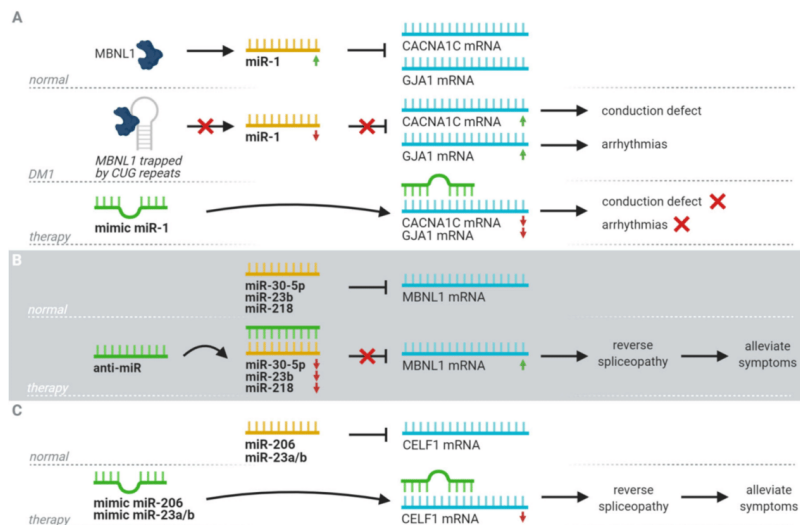
Several things have to be taken into account with regards to extracellular miRNAs. Although some miRNAs are tissue-specific, there is still a major overlap between the expression of miRNAs. For example, although miRNA-206 is highly enriched in skeletal muscle, it can also be detected in cardiac muscle, and its origin might be from both. The same holds true for miR-1 and miR-133, which are highly enriched in both tissues. The multi-systemic nature of DM1 may add additional complexity to determining the biomarker potential of extracellular miRNAs, as it will be extremely difficult to distinguish between the origins of the extracellular miRNAs. Specifically, for myo-miRs, their (possibly low) disease specificity might be a cause for concern. Several studies have shown that extracellular myo-miRs are elevated in different muscular pathologies, including DMD and limb-girdle muscular dystrophies [11,66], which might indicate that elevated extracellular myo-miRs is a general marker of muscle pathology. Furthermore, the expression levels in the muscle are not always reflected in that found in serum or plasma. It has been shown that the increase in extracellular myo-miRs in serum and plasma is, in part, due to selective release of certain miRNAs during muscle growth and regeneration, instead of solely a passive leakage from damaged muscle and, therefore, more likely a complex function of the regenerative/degenerative status of the muscle, overall muscle mass, and tissue expression levels, which can have implications for their biomarker potential [67,68].

## 5. Therapeutic Potential

The dysregulation patterns of miRNAs found in DM1 and their links to the clinical phenotype also show their potential as therapeutics. The search for a therapeutic, targeting miRNAs in DM1, has focused mainly on miRNAs associated with the two splice factors involved in DM1 pathology, namely MBNL1 and CELF1 (Figure 2).

Four different miRNAs (miR-1, miR-30-5p, miR-23b, and miR-218) associated with the splice factor MBNL1 have been studied as therapeutic targets. As previously mentioned, pre-miR-1 processing is regulated by MBNL1. Sequestration of MBNL1 by toxic RNA results in downregulation of miR-1 and subsequent upregulation of downstream targets of miR-1. This, in turn, is linked to several DM1 symptoms, for example, one of the miR-1 targets is CACNA1C, which encodes for calcium channels in the heart. Misregulation may contribute to arrhythmias observed in DM1 patients [55]. However, no steps toward upregulating miR-1 to assess its therapeutic potential have been taken. miR-30-5p is another miRNA associated with the MBNL1 family, since it is a direct repressor of MBNL1-3 translation. A study showed that introducing a miR-30-5p mimic into C2C12 muscle cells downregulated the expression of MBNL and deregulated downstream targets of MBNL, including Trim55 (involved in sarcomere assembly) and Insulin Receptor (binding of insulin to the receptor initiates a downstream pathway that regulates muscle development) [69]. This shows an interesting possibility of using

anti-miRs to repress miR-30-5p expression, increase MBNL1 protein levels, and alleviate associated DM1 symptoms. The last two miRNAs linked to MBNL1 were found by Cerro-Herreros et al. in 2018 [70]. miR-23b and miR-218 were identified as repressors of MBNL1 translation. Anti-miR treatment of each miRNA independently in DM1 myoblasts and DM1 mice increased MBNL1 levels, and in the latter this was linked to improvements in spliceopathy profile, histopathological signs, and functional myotonia without toxicity [70]. More recently, the same group has conducted a more in-depth study on the therapeutic potential of an anti-miR against miR-23b. They showed that the therapeutic effect is dose-dependent and has long-lasting effects. Subcutaneous administration of anti-miR 23b in human skeletal actin long repeat mice upregulated the expression of MBNL1 and rescued splicing alterations, grip strength, and myotonia [71].



**Figure 2.** An overview of the suggested therapeutic potential of miRNAs in Myotonic Dystrophy type I (DM1) by targeting downstream targets of MBNL1, upregulating mRNA levels of MBNL1, or downregulating mRNA levels of CELF1 by the use of miRNA mimics or anti-miRs. (A) The use of an miR-1 mimic to circumvent the detrimental effect of MBNL1 sequestration, resulting in a downregulation of downstream targets and symptom alleviation. (B) The use of an anti-miR for miRNA inhibitors of MBNL1 to increase expression levels in DM1 and alleviate symptoms. (C) The use of a miRNA mimic of inhibitors of CELF1 to reduce its expression and alleviate symptoms in DM1. MBNL1 = muscleblind-like 1, CELF1 = CUG-BP and ETR-3-like factors family member 1, CACNA1C = Voltage-dependent L-type calcium channel subunit alpha-1C, GJA1 = connexin 43. Created with [BioRender.com](https://www.biorender.com).

For the splice factor CELF1, the focus has been primarily on miR-206, which may modulate CELF1 overexpression. The introduction of miR-206 mimics into cells overexpressing CELF1 inhibited CELF1 expression and improved the myoblast fusion index and myotube area [72]. Another pair of miRNAs of note are miR-23a/b; they have shown to be downregulated by introducing expanded repeats, which leads to the upregulation of CELF1. This is another potential therapeutic ability worth exploring [56].

A recent study has discovered a new potential therapeutic target, namely miR-7. miR-7 was previously described to be downregulated in a DM1 *Drosophila* model and in muscle biopsies from patients [52]. In the current study, overexpression of miR-7 resulted in the rescue of DM1 myoblast

fusion defects and myotube growth by repressing autophagy and the ubiquitin-proteasome system [73]. This improvement was independent of MBNL1. These results provide evidence for a new therapeutic candidate for muscle dysfunction in DM1.

## 6. Conclusions and Future Perspectives

miRNAs have great potential to become diagnostic tools in several diseases, including cancer, neurodegenerative diseases, and neuromuscular diseases. However, to date, insufficient knowledge and a lack of conclusive results to clarify the role of miRNA in disease diagnosis have held back the implementation of miRNA biomarkers in clinical settings [4]. In the past two decades, extensive research has been conducted in the miRNA expression profiles of DM1 patients and their biomarker potential. Several expression profiles have been found to be able to distinguish between DM1 patients and healthy subjects, and even a link has been made between progressive and non-progressive muscle wasting in DM1 patients. However, discrepancy between studies still exists. One of the main causes of the discrepancies seems to be the normalization process necessary for qPCR, for which no gold standard exists. A way around this might be the use of a fairly novel technique, the digital droplet PCR, which produces an absolute quantity, eliminating the need for normalization [74]. This will help in the comparability of results between groups. Further research, especially longitudinal studies, are needed to unravel the true biomarker potential of miRNAs in DM1 to see whether they can help in the prediction of disease progression and/or in the prediction of treatment efficacy.

**Author Contributions:** Conceptualization, E.K. and G.N.-G.; writing-original draft preparation, E.K.; writing-review and editing, E.K., G.N.-G., A.B.-L., V.A.-G., A.M.-P.; supervision, G.N.-G.; funding acquisition, G.N.-G. All authors have read and agreed to the published version of the manuscript.

**Funding:** G.N.G. and V.A.G. declare grants from Instituto de Salud Carlos III (Grant Numbers: PII5/01756; P118/00114 and P18/00713), Madrid, Spain, and AFM Telethon (Trampoline grant number #21108), France. E.K. is funded by the “La Caixa” Foundation (ID 100010434), fellowship code LCF/BQ/DI18/11660019, co-funded by the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement n°713673. A.B.L. is funded by an FI Agaur fellowship FL B 01090 and by SGR 1520. G.N.G. and V.A.G. are supported by a Miguel Servet research contracts (ISCIII CP12/03057, CP14/00032, CPII17/00004, CPII19/00021, and FEDER). V.A.G. also acknowledges support from Ikerbasque.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Guo, H.; Ingolia, N.T.; Weissman, J.S.; Bartel, D.P. Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* **2010**, *466*, 835–840. [[CrossRef](#)]
- Berezikov, E.; Guryev, V.; van de Belt, J.; Wienholds, E.; Plasterk, R.H.A.; Cuppen, E. Phylogenetic Shadowing and Computational Identification of Human microRNA Genes. *Cell* **2005**, *120*, 21–24. [[CrossRef](#)]
- Shivdasani, R.A. MicroRNAs: Regulators of gene expression and cell differentiation. *Blood* **2006**, *108*, 3646–3653. [[CrossRef](#)] [[PubMed](#)]
- Condrat, C.E.; Thompson, D.C.; Barbu, M.G.; Bugnar, O.L.; Boboc, A.; Cretoiu, D.; Suci, N.; Cretoiu, S.M.; Voinea, S.C. miRNAs as Biomarkers in Disease: Latest Findings Regarding Their Role in Diagnosis and Prognosis. *Cells* **2020**, *9*, 276. [[CrossRef](#)] [[PubMed](#)]
- Makarova, J.A.; Shkurnikov, M.U.; Wicklein, D.; Lange, T.; Samatov, T.R.; Turchinovich, A.A.; Tonevitsky, A.G. Intracellular and extracellular microRNA: An update on localization and biological role. *Prog. Histochem. Cytochem.* **2016**, *51*, 33–49. [[CrossRef](#)] [[PubMed](#)]
- Sun, Z.; Shi, K.; Yang, S.; Liu, J.; Zhou, Q.; Wang, G.; Song, J.; Li, Z.; Zhang, Z.; Yuan, W. Effect of exosomal miRNA on cancer biology and clinical applications. *Mol. Cancer* **2018**, *17*. [[CrossRef](#)]
- Wiedrick, J.T.; Phillips, J.I.; Lusardi, T.A.; McFarland, T.J.; Lind, B.; Sandau, U.S.; Harrington, C.A.; Lapidus, J.A.; Galasko, D.R.; Quinn, J.F.; et al. Validation of MicroRNA Biomarkers for Alzheimer’s Disease in Human Cerebrospinal Fluid. *J. Alzheim. Dis.* **2019**, *67*, 875–891. [[CrossRef](#)] [[PubMed](#)]
- Pitkänen, A.; Ekolle Ndode-Ekane, X.; Lapinlampi, N.; Puhakka, N. Epilepsy biomarkers—Toward etiology and pathology specificity. *Neurobiol. Dis.* **2019**, *123*, 42–58. [[CrossRef](#)]
- Kingsley, S.M.K.; Bhat, B.V. Role of microRNAs in sepsis. *Inflamm. Res.* **2017**, *66*, 553–569. [[CrossRef](#)]

10. Viswambaran, V.; Thanseem, I.; Vasu, M.M.; Poovathinal, S.A.; Anitha, A. miRNAs as biomarkers of neurodegenerative disorders. *Biomark. Med.* **2017**, *11*, 151–167. [[CrossRef](#)]
11. Coenen-Stass, A.M.L.; Wood, M.J.A.; Roberts, T.C. Biomarker Potential of Extracellular miRNAs in Duchenne Muscular Dystrophy. *Trends Mol. Med.* **2017**, *23*, 989–1001. [[CrossRef](#)] [[PubMed](#)]
12. Harper, P.S. *Major Problems in Neurology: Myotonic Dystrophy*, 3rd ed.; WB Saunders: London, UK, 2001.
13. Campbell, C.; Sherlock, R.; Jacob, P.; Blayney, M. Congenital myotonic dystrophy: Assisted ventilation duration and outcome. *Pediatrics* **2004**, *113*, 811–816. [[CrossRef](#)] [[PubMed](#)]
14. Douniol, M.; Jacquette, A.; Cohen, D.; Bodeau, N.; Rachidi, L.; Angeard, N.; Cuisset, J.-M.; Vallée, L.; Eymard, B.; Plaza, M.; et al. Psychiatric and cognitive phenotype of childhood myotonic dystrophy type 1. *Dev. Med. Child Neurol.* **2012**, *54*, 905–911. [[CrossRef](#)] [[PubMed](#)]
15. De Antonio, M.; Dogan, C.; Hamroun, D.; Mati, M.; Zerrouki, S.; Eymard, B.; Katsahian, S.; Bassez, G. French Myotonic Dystrophy Clinical Network Unravelling the myotonic dystrophy type 1 clinical spectrum: A systematic registry-based study with implications for disease classification. *Rev. Neurol. (Paris)* **2016**, *172*, 572–580. [[CrossRef](#)] [[PubMed](#)]
16. Mathieu, J.; Allard, P.; Potvin, L.; Prévost, C.; Bégin, P. A 10-year study of mortality in a cohort of patients with myotonic dystrophy. *Neurology* **1999**, *52*, 1658–1662. [[CrossRef](#)]
17. De Die-Smulders, C.E.; Höweler, C.J.; Thijs, C.; Mirandolle, J.F.; Anten, H.B.; Smeets, H.J.; Chandler, K.E.; Geraedts, J.P. Age and causes of death in adult-onset myotonic dystrophy. *Brain* **1998**, *121*, 1557–1563. [[CrossRef](#)]
18. Groh, W.J.; Groh, M.R.; Saha, C.; Kincaid, J.C.; Simmons, Z.; Ciafaloni, E.; Pourmand, R.; Otten, R.F.; Bhakta, D.; Nair, G.V.; et al. Electrocardiographic abnormalities and sudden death in myotonic dystrophy type 1. *N. Engl. J. Med.* **2008**, *358*, 2688–2697. [[CrossRef](#)]
19. Groh, W.J.; Groh, M.R.; Shen, C.; Monckton, D.G.; Bodkin, C.L.; Pascuzzi, R.M. Survival and CTG repeat expansion in adults with myotonic dystrophy type 1. *Muscle Nerve* **2011**, *43*, 648–651. [[CrossRef](#)]
20. Logigian, E.L.; Moxley, R.T.; Blood, C.L.; Barbieri, C.A.; Martens, W.B.; Wiegner, A.W.; Thornton, C.A.; Moxley, R.T. Leukocyte CTG repeat length correlates with severity of myotonia in myotonic dystrophy type 1. *Neurology* **2004**, *62*, 1081–1089. [[CrossRef](#)]
21. Meola, G.; Cardani, R. Myotonic dystrophies: An update on clinical aspects, genetic, pathology, and molecular pathomechanisms. *Biochim. Biophys. Acta Mol. Basis Dis.* **2015**, *1852*, 594–606. [[CrossRef](#)]
22. López-Martínez, A.; Soblechero-Martín, P.; De-La-puente-ovejero, L.; Nogales-Gadea, G.; Arechavala-Gomez, V. An overview of alternative splicing defects implicated in myotonic dystrophy type i. *Genes* **2020**, *11*, 1109. [[CrossRef](#)]
23. Kuyumcu-Martinez, N.M.; Wang, G.S.; Cooper, T.A. Increased Steady-State Levels of CUGBP1 in Myotonic Dystrophy 1 Are Due to PKC-Mediated Hyperphosphorylation. *Mol. Cell* **2007**, *28*, 68–78. [[CrossRef](#)] [[PubMed](#)]
24. Mankodi, A.; Takahashi, M.P.; Jiang, H.; Beck, C.L.; Bowers, W.J.; Moxley, R.T.; Cannon, S.C.; Thornton, C.A. Expanded CUG repeats trigger aberrant splicing of ClC-1 chloride channel pre-mRNA and hyperexcitability of skeletal muscle in myotonic dystrophy. *Mol. Cell* **2002**, *10*, 35–44. [[CrossRef](#)]
25. Sicot, G.; Gourdon, G.; Gomes-Pereira, M. Myotonic dystrophy, when simple repeats reveal complex pathogenic entities: New findings and future challenges. *Hum. Mol. Genet.* **2011**, *20*, R116–R123. [[CrossRef](#)] [[PubMed](#)]
26. Mahadevan, M.S. Myotonic dystrophy: Is a narrow focus obscuring the rest of the field? *Curr. Opin. Neurol.* **2012**, *25*, 609–613. [[CrossRef](#)] [[PubMed](#)]
27. Gudde, A.E.E.G.; van Heeringen, S.J.; de Oude, A.I.; van Kessel, I.D.G.; Estabrook, J.; Wang, E.T.; Wieringa, B.; Wansink, D.G. Antisense transcription of the myotonic dystrophy locus yields low-abundant RNAs with and without (CAG)<sub>n</sub> repeat. *RNA Biol.* **2017**, *14*, 1374–1388. [[CrossRef](#)] [[PubMed](#)]
28. Zu, T.; Gibbens, B.; Doty, N.S.; Gomes-Pereira, M.; Huguet, A.; Stone, M.D.; Margolis, J.; Peterson, M.; Markowski, T.W.; Ingram, M.A.C.; et al. Non-ATG-initiated translation directed by microsatellite expansions. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 260–265. [[CrossRef](#)]
29. Wahbi, K.; Babuty, D.; Probst, V.; Wissocque, L.; Labombarda, F.; Porcher, R.; Bécane, H.M.; Lazarus, A.; Béhin, A.; Laforêt, P.; et al. Incidence and predictors of sudden death, major conduction defects and sustained ventricular tachyarrhythmias in 1388 patients with myotonic dystrophy type 1. *Eur. Heart J.* **2017**, *38*, 751–758. [[CrossRef](#)]
30. Falsaperla, R.; Parano, E.; Romano, C.; Praticò, A.D.; Pavone, P. HyperCKemia as a biomarker for muscular diseases. *Clin. Ter.* **2010**, *161*, 185–187.
31. Koch, A.J.; Pereira, R.; Machado, M. The creatine kinase response to resistance exercise. *J. Musculoskelet. Neu. Interact.* **2014**, *14*, 68–77.

32. Nakamori, M.; Sobczak, K.; Puwanant, A.; Welle, S.; Eichinger, K.; Pandya, S.; Dekdebrun, J.; Heatwole, C.R.; McDermott, M.P.; Chen, T.; et al. Splicing biomarkers of disease severity in myotonic dystrophy. *Ann. Neurol.* **2013**, *74*, 862–872. [[CrossRef](#)] [[PubMed](#)]
33. Perfetti, A.; Greco, S.; Bugiardi, E.; Cardani, R.; Gaia, P.; Gaetano, C.; Meola, G.; Martelli, F. Plasma microRNAs as biomarkers for myotonic dystrophy type 1. *Neuromuscul. Disord.* **2014**, *24*, 509–515. [[CrossRef](#)] [[PubMed](#)]
34. Koutsoulidou, A.; Kyriakides, T.C.; Papadimas, G.K.; Christou, Y.; Kararizou, E.; Papanicolaou, E.Z.; Phylactou, L.A. Elevated muscle-specific miRNAs in serum of myotonic dystrophy patients relate to muscle disease progress. *PLoS ONE* **2015**, *10*, e0125341. [[CrossRef](#)] [[PubMed](#)]
35. Gagnon, C.; Heatwole, C.; Hébert, L.J.; Hogrel, J.-Y.; Laberge, L.; Leone, M.; Meola, G.; Richer, L.; Sansone, V.; Kierkegaard, M. Report of the third outcome measures in myotonic dystrophy type 1 (OMMYD-3) international workshop. *J. Neuromuscul. Dis.* **2018**, *5*, 523–537. [[CrossRef](#)] [[PubMed](#)]
36. Alfano, L.; Lowes, L.; Berry, K.; Flanigan, K.; Cripe, L.; Mendell, J. Role of motivation on performance of the 6-minute walk test in boys with Duchenne muscular dystrophy. *Dev. Med. Child Neurol.* **2015**, *57*, 57–58. [[CrossRef](#)]
37. Wu, N.; Gu, T.; Lu, L.; Cao, Z.; Song, Q.; Wang, Z.; Zhang, Y.; Chang, G.; Xu, Q.; Chen, G. Roles of miRNA-1 and miRNA-133 in the proliferation and differentiation of myoblasts in duck skeletal muscle. *J. Cell. Physiol.* **2019**, *234*, 3490–3499. [[CrossRef](#)] [[PubMed](#)]
38. Chen, J.F.; Mandel, E.M.; Thomson, J.M.; Wu, Q.; Callis, T.E.; Hammond, S.M.; Conlon, F.L.; Wang, D.Z. The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nat. Genet.* **2006**. [[CrossRef](#)] [[PubMed](#)]
39. Feng, Y.; Niu, L.L.; Wei, W.; Zhang, W.Y.; Li, X.Y.; Cao, J.H.; Zhao, S.H. A feedback circuit between miR-133 and the ERK1/2 pathway involving an exquisite mechanism for regulating myoblast proliferation and differentiation. *Cell Death Dis.* **2013**, *4*. [[CrossRef](#)]
40. Chistiakov, D.A.; Orekhov, A.N.; Bobryshev, Y.V. Cardiac-specific miRNA in cardiogenesis, heart function, and cardiac pathology (with focus on myocardial infarction). *J. Mol. Cell. Cardiol.* **2016**, *94*, 107–121. [[CrossRef](#)]
41. Hak, K.K.; Yong, S.L.; Sivaprasad, U.; Malhotra, A.; Dutta, A. Muscle-specific microRNA miR-206 promotes muscle differentiation. *J. Cell Biol.* **2006**, *174*, 677–687. [[CrossRef](#)]
42. Van Rooij, E.; Quiat, D.; Johnson, B.A.; Sutherland, L.B.; Qi, X.; Richardson, J.A.; Kelm, R.J.; Olson, E.N. A Family of microRNAs Encoded by Myosin Genes Governs Myosin Expression and Muscle Performance. *Dev. Cell* **2009**, *17*, 662–673. [[CrossRef](#)] [[PubMed](#)]
43. Small, E.M.; O'Rourke, J.R.; Moresi, V.; Sutherland, L.B.; McAnally, J.; Gerard, R.D.; Richardson, J.A.; Olson, E.N. Regulation of PI3-kinase/Akt signaling by muscle-enriched microRNA-486. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 4218–4223. [[CrossRef](#)]
44. Carè, A.; Catalucci, D.; Felicetti, F.; Bonci, D.; Addario, A.; Gallo, P.; Bang, M.L.; Segnalini, P.; Gu, Y.; Dalton, N.D.; et al. MicroRNA-133 controls cardiac hypertrophy. *Nat. Med.* **2007**, *13*, 613–618. [[CrossRef](#)] [[PubMed](#)]
45. Yang, B.; Lin, H.; Xiao, J.; Lu, Y.; Luo, X.; Li, B.; Zhang, Y.; Xu, C.; Bai, Y.; Wang, H.; et al. The muscle-specific microRNA miR-1 regulates cardiac arrhythmogenic potential by targeting GJA1 and KCNJ2. *Nat. Med.* **2007**, *13*, 486–491. [[CrossRef](#)]
46. Greco, S.; De Simone, M.; Colussi, C.; Zaccagnini, G.; Fasanaro, P.; Pescatori, M.; Cardani, R.; Perbellini, R.; Isaia, E.; Sale, P.; et al. Common micro-RNA signature in skeletal muscle damage and regeneration induced by Duchenne muscular dystrophy and acute ischemia. *FASEB J.* **2009**, *23*, 3335–3346. [[CrossRef](#)]
47. Eisenberg, I.; Eran, A.; Nishino, I.; Moggio, M.; Lamperti, C.; Amato, A.A.; Lidov, H.G.; Kang, P.B.; North, K.N.; Mitrani-Rosenbaum, S.; et al. Distinctive patterns of microRNA expression in primary muscular disorders. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 17016–17021. [[CrossRef](#)]
48. Gambardella, S.; Rinaldi, F.; Lepore, S.M.; Viola, A.; Loro, E.; Angelini, C.; Vergani, L.; Novelli, G.; Botta, A. Overexpression of microRNA-206 in the skeletal muscle from myotonic dystrophy type 1 patients. *J. Transl. Med.* **2010**, *8*, 48. [[CrossRef](#)]
49. Hon, L.S.; Zhang, Z. The roles of binding site arrangement and combinatorial targeting in microRNA repression of gene expression. *Genome Biol.* **2007**, *8*. [[CrossRef](#)]
50. Yuasa, K.; Hagiwara, Y.; Ando, M.; Nakamura, A.; Takeda, S.; Hijikata, T. MicroRNA-206 is highly expressed in newly formed muscle fibers: Implications regarding potential for muscle regeneration and maturation in muscular dystrophy. *Cell Struct. Funct.* **2008**, *33*, 163–169. [[CrossRef](#)]

51. Perbellini, R.; Greco, S.; Sarra-Ferraris, G.; Cardani, R.; Capogrossi, M.C.; Meola, G.; Martelli, F. Dysregulation and cellular mislocalization of specific miRNAs in myotonic dystrophy type 1. *Neuromuscul. Disord.* **2011**, *21*, 81–88. [[CrossRef](#)]
52. Fernandez-Costa, J.M.; Garcia-Lopez, A.; Zuñiga, S.; Fernandez-Pedrosa, V.; Felipo-Benavent, A.; Mata, M.; Jaka, O.; Aiastui, A.; Hernandez-Torres, F.; Aguado, B.; et al. Expanded CTG repeats trigger miRNA alterations in *Drosophila* that are conserved in myotonic dystrophy type 1 patients. *Hum. Mol. Genet.* **2013**, *22*, 704–716. [[CrossRef](#)] [[PubMed](#)]
53. Fritegatto, C.; Ferrati, C.; Pegoraro, V.; Angelini, C. Micro-RNA expression in muscle and fiber morphometry in myotonic dystrophy type 1. *Neurol. Sci.* **2017**. [[CrossRef](#)] [[PubMed](#)]
54. Roberts, T.C.; Coenen-Stass, A.M.L.; Wood, M.J.A. Assessment of RT-qPCR normalization strategies for accurate quantification of extracellular microRNAs in murine Serum. *PLoS ONE* **2014**, *9*. [[CrossRef](#)] [[PubMed](#)]
55. Rau, F.; Freyermuth, F.; Fugier, C.; Villemin, J.P.; Fischer, M.C.; Jost, B.; Dembele, D.; Gourdon, G.; Nicole, A.; Duboc, D.; et al. Misregulation of miR-1 processing is associated with heart defects in myotonic dystrophy. *Nat. Struct. Mol. Biol.* **2011**, *18*, 840–845. [[CrossRef](#)] [[PubMed](#)]
56. Kalsotra, A.; Singh, R.K.; Gurha, P.; Ward, A.J.; Creighton, C.J.; Cooper, T.A. The Mef2 transcription network is disrupted in myotonic dystrophy heart tissue, dramatically altering miRNA and mRNA expression. *Cell Rep.* **2014**, *6*, 336–345. [[CrossRef](#)] [[PubMed](#)]
57. Kalsotra, A.; Wang, K.; Li, P.F.; Cooper, T.A. MicroRNAs coordinate an alternative splicing network during mouse postnatal heart development. *Genes Dev.* **2010**, *24*, 653–658. [[CrossRef](#)]
58. Perfetti, A.; Greco, S.; Cardani, R.; Fossati, B.; Cuomo, G.; Valaperta, R.; Ambroggi, F.; Cortese, A.; Botta, A.; Mignarri, A.; et al. Validation of plasma microRNAs as biomarkers for myotonic dystrophy type 1. *Sci. Rep.* **2016**, *6*, 38174. [[CrossRef](#)] [[PubMed](#)]
59. Koutsoulidou, A.; Photiades, M.; Kyriakides, T.C.; Georgiou, K.; Prokopi, M.; Kapnisis, K.; Lusakowska, A.; Nearchou, M.; Christou, Y.; Papadimas, G.K.; et al. Identification of exosomal muscle-specific miRNAs in serum of myotonic dystrophy patients relating to muscle disease progress. *Hum. Mol. Genet.* **2017**, *26*, 3285–3302. [[CrossRef](#)]
60. Lötvall, J.; Hill, A.F.; Hochberg, F.; Buzás, E.I.; Di Vizio, D.; Gardiner, C.; Gho, Y.S.; Kurochkin, I.V.; Mathivanan, S.; Quesenberry, P.; et al. Minimal experimental requirements for definition of extracellular vesicles and their functions: A position statement from the International Society for Extracellular Vesicles. *J. Extracell. Vesicles* **2014**, *1*, 1–6. [[CrossRef](#)]
61. Karttunen, J.; Heiskanen, M.; Navarro-Ferrandis, V.; Das Gupta, S.; Lipponen, A.; Puhakka, N.; Rilla, K.; Koistinen, A.; Pitkänen, A. Precipitation-based extracellular vesicle isolation from rat plasma co-precipitate vesicle-free microRNAs. *J. Extracell. Vesicles* **2019**, *8*. [[CrossRef](#)]
62. Fernandez-Costa, J.M.; Llamusi, B.; Bargiela, A.; Zulaica, M.; Alvarez-Abril, M.C.; Perez-Alonso, M.; De Munain, A.L.; Lopez-Castel, A.; Artero, R. Six serum miRNAs fail to validate as myotonic dystrophy type 1 biomarkers. *PLoS ONE* **2016**, *11*, e0150501. [[CrossRef](#)] [[PubMed](#)]
63. Pegoraro, V.; Cudia, P.; Baba, A.; Angelini, C. MyomiRNAs and myostatin as physical rehabilitation biomarkers for myotonic dystrophy. *Neurol. Sci.* **2020**, 1–8. [[CrossRef](#)] [[PubMed](#)]
64. Ambrose, K.K.; Ishak, T.; Lian, L.H.; Goh, K.J.; Wong, K.T.; Ahmad-Annuar, A.; Thong, M.K. Deregulation of microRNAs in blood and skeletal muscles of myotonic dystrophy type 1 patients. *Neurol. India* **2017**, *65*, 512–517. [[CrossRef](#)] [[PubMed](#)]
65. Keller, A.; Meese, E. Can circulating miRNAs live up to the promise of being minimal invasive biomarkers in clinical settings? *Wiley Interdiscip. Rev. RNA* **2016**, *7*, 148–156. [[CrossRef](#)] [[PubMed](#)]
66. Vignier, N.; Amor, F.; Fogel, P.; Duvallet, A.; Poupiot, J.; Charrier, S.; Arock, M.; Montus, M.; Nelson, I.; Richard, I.; et al. Distinctive Serum miRNA Profile in Mouse Models of Striated Muscular Pathologies. *PLoS ONE* **2013**, *8*. [[CrossRef](#)]
67. Coenen-Stass, A.M.L.; Betts, C.A.; Lee, Y.F.; Mäger, I.; Turunen, M.P.; El Andaloussi, S.; Morgan, J.E.; Wood, M.J.A.; Roberts, T.C. Selective release of muscle-specific, extracellular microRNAs during myogenic differentiation. *Hum. Mol. Genet.* **2016**, *25*, 3960–3974. [[CrossRef](#)]
68. Coenen-Stass, A.M.L.; Sork, H.; Gatto, S.; Godfrey, C.; Bhomra, A.; Krjutškov, K.; Hart, J.R.; Westholm, J.O.; O'Donovan, L.; Roos, A.; et al. Comprehensive RNA-Sequencing Analysis in Serum and Muscle Reveals Novel Small RNA Signatures with Biomarker Potential for DMD. *Mol. Ther. Nucl. Acids* **2018**, *13*, 1–15. [[CrossRef](#)]

69. Zhang, B.W.; Cai, H.F.; Wei, X.F.; Sun, J.J.; Lan, X.Y.; Lei, C.Z.; Lin, F.P.; Qi, X.L.; Plath, M.; Chen, H. mir-30-5p regulates muscle differentiation and alternative splicing of muscle-related genes by targeting MBNL. *Int. J. Mol. Sci.* **2016**, *17*, 182. [[CrossRef](#)]
70. Cerro-Herreros, E.; Sabater-Arcis, M.; Fernandez-Costa, J.M.; Moreno, N.; Perez-Alonso, M.; Llamusi, B.; Artero, R. MiR-23b and miR-218 silencing increase Muscleblind-like expression and alleviate myotonic dystrophy phenotypes in mammalian models. *Nat. Commun.* **2018**, *9*. [[CrossRef](#)]
71. Cerro-Herreros, E.; González-Martínez, I.; Moreno-Cervera, N.; Overby, S.; Pérez-Alonso, M.; Llamusi, B.; Artero, R. Therapeutic Potential of AntagomiR-23b for Treating Myotonic Dystrophy. *Mol. Ther. Nucleic Acids* **2020**, *21*, 837–849. [[CrossRef](#)]
72. Koutalianos, D.; Koutsoulidou, A.; Mastrogiannopoulos, N.P.; Furling, D.; Phylactou, L.A. MyoD transcription factor induces myogenesis by inhibiting Twist-1 through miR-206. *J. Cell Sci.* **2015**, *128*, 3631–3645. [[CrossRef](#)] [[PubMed](#)]
73. Sabater-Arcis, M.; Bargiela, A.; Furling, D.; Artero, R. miR-7 Restores Phenotypes in Myotonic Dystrophy Muscle Cells by Repressing Hyperactivated Autophagy. *Mol. Ther. Nucleic Acids* **2020**, *19*, 278–292. [[CrossRef](#)] [[PubMed](#)]
74. Cheng, Y.; Dong, L.; Zhang, J.; Zhao, Y.; Li, Z. Recent advances in microRNA detection. *Analyst* **2018**, *143*, 1758–1774. [[CrossRef](#)] [[PubMed](#)]

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



**GENERAL  
DISCUSSION**





The main goal of this doctoral thesis was to have an in-depth look in potential disease modifiers of DM1 and to identify their link to DM1 pathology and clinical phenotype. The findings laid out in this dissertation have added novel findings, confirmed existing reports, and called into question previously reported findings. Five potential disease modifiers were the focus of this dissertation and several links to DM1 phenotype have been made and are discussed below.

### DM1 VARIANT REPEATS

---

The effect of variant repeat patterns on the DM1 clinical phenotype is still poorly understood. Although variant repeats have been previously linked to disease phenotype, the results are quite controversial. Nevertheless, the most observed finding to date is a delay in age of onset in variant repeat carrying DM1 patients compared to pure repeat carrying DM1 patients [26,34,189,190,193]. Our results support this observation, with an age of onset >50 years for the three sisters and an asymptomatic son at the age of 35. However, the also often reported milder phenotype that accompanies the delay in onset was absent in our interrupted family. Several studies have reported less severe muscle weakness, lower degree of myotonia and better respiratory function [26,44,189,190,193,194]. In our reported DM1 family, we observed a quite severe clinical manifestation after initial age of onset in the three sisters with cardiac and respiratory dysfunction, requiring a pacemaker and mechanical ventilation. Although the majority of studies reported a milder phenotype, there are reports indicating either no differences or a worsening of symptoms in variant repeat carrying DM1 patients [189,191]. Of note, several atypical traits were observed in the three sisters of the interrupted family, including proximal limb weakness, severe axial involvement and the absence of a myopathic face, despite the presence of moderate facial weakness. Proximal weakness is more suggestive of DM2 and the dropped-head observed in one of the sisters is more resembling limb-girdle muscle dystrophy. An atypical phenotype has been previously observed, including a DM2-like phenotype [26,191].

CTG expansion intergenerational analysis revealed a contraction and an expansion of the CTG repeat in the interrupted family. The expansion was accompanied by anticipation as the next generation had an earlier age of onset. The patient with the observed contraction is still asymptomatic and therefore no links to disease severity can be made. Variant repeat carrying individuals have shown a more stable intergenerational transmission, with either similar or contracted transmitted CTG repeats [26,43,192,195]. Although the overall consensus is a stabilizing effect, anticipation has been observed in several interrupted DM1 families [26,34,43,189,190].

Variant repeats in our cohort displayed a prevalence of ~10% on individual basis and 3% on family level. This is in consensus with previously reported findings, where the prevalence is estimated to be between 3 to 5% [26,43,189–191]. The variants found were CCGs at the 3' end region, which is one of the most common found variant repeats to date

and have been reported as single repeats, in CCGCTG hexamers or as small or large (CCG)<sub>n</sub> arrays [26,43,189–191]. The studied interrupted family members in our cohort showed all previously described CCG variant repeats (single repeats, hexamers and small (CCG)<sub>n</sub> arrays) in the expanded alleles, which were both changed and preserved upon generational transmission.

Altogether, our study addressing the effect of variant repeats on DM1 corroborates the general findings that their prevalence is between 3 to 5% and that at clinical level, variant repeats delay the DM1 age of onset compared to pure repeat carrying patients. However, our results did not show a milder phenotype, except for the asymptomatic patient, which interestingly showed a contracted repeat, in line with the overall consensus that variant repeats have a stabilizing effect on CTG repeat. Due to the small number of interrupted patients present in the DM1 population, it is hard to perform accurate genotype-phenotype correlations and there is still much uncertainty. Studies with larger DM1 cohorts, preferably with DM1 families, are needed to better understand the phenotypic consequences of variant repeat patterns and to study their effect on intergenerational transmissions of the *DMPK* expanded allele, and to address whether variant repeats have an impact in the epigenetic regulation of the *DMPK* locus.

#### DM1-AS TRANSCRIPTION AND RAN TRANSLATION

---

In 2005, antisense transcription was first reported in DM1 and the first report on RAN translation was in 2011, both were considered potential new disease modifiers in DM1. However, limited new data on antisense transcription, and no new data on RAN translation has been published since their first reports and sample collection in the initial reports was limited. Thanks to the precious collection of DM1 patient-derived cells available to use, we addressed the presence of antisense transcription and polyGln RAN protein in three primary cell cultures of patients with DM1, namely myoblast, skin fibroblast and lymphoblastoid cell lines, in order to further elucidate its contribution to DM1 pathology.

Lower levels of DM1-AS expression were observed in DM1 patients compared to controls, and the primer combination encompassing the repeat region seemed to favor the wild-type alleles. Only one patient showed a band that could correspond to the expanded repeat and this patient carried the smallest CTG expansion. Hence, it could be possible that the much larger expanded repeats could not be detected by this method, which in turn could be a potential explanation for the lower levels of expression seen in patients compared to controls. However, the lower levels were also observed with the other two primer combinations that did not encompass the expanded repeat, making it highly unlikely that the lower levels seen were solely due to the binding of the wild-type allele. DM1-AS expression has only been studied by a handful of other groups. These studies revealed either no changes or higher levels of DM1-AS expression, which is in vast contrast to the results obtained from our study. Moreover, Gudde and collaborators showed, when stratified based on inferred

MBNL concentration, that the most severely affected patients showed a three-fold increase in DM1-AS expression compared to controls [218]. We were unable to do such stratification, but we have extensive knowledge on the clinical phenotype of our DM1 cohort, although upon revision, a correlation between expression levels and clinical phenotype could not be found. Our results disagree with the previous study, but the sample size of our cohort was rather small for clinical comparisons, hindering the analysis. To better determine whether DM1-AS transcript expression is linked to disease severity, a bigger cohort is needed.

The presence of DM1-AS transcripts in DM1 cells does not necessarily mean that these transcripts can reach the cytoplasm and be RAN-translated. Subcellular fractionation and FISH antisense RNA foci analysis revealed the presence of DM1-AS transcripts in the cytoplasm of both patients and controls, with a higher percentage in the nuclear fraction, and a part of these DM1-AS transcripts contained the expanded repeat. The presence of cytoplasmic DM1-AS transcripts and the presence of antisense RNA foci were previously described [216–218]. Gudde and collaborators showed the presence of DM1-AS transcripts as a heterogeneous pool of transcripts with and without the expanded repeat in the cytoplasmic fraction of myoblasts [218], corroborating our results.

However, the polyGln RAN protein was undetectable in all three of our primary cell cultures in both protein blots, which revealed a 42 kD polyGln-containing protein with the two commercial anti-polyGln antibodies, and in immunofluorescence, which showed infrequent staining of the nucleus and an aggregate around the nucleus. The 42 kD protein might be TBP and in fact, the original immunogen for the 1C2 antibody was the general transcription factor TBP, which contains a 38-Gln stretch and therefore matches our results. It was shown, however, that TBP will always appear in blots, but the binding of longer stretches of PolyGln will be favored [327]. Accordingly, a certain subset of lymphoblastoids did show a band that might correspond to the polyGln RAN protein, but the custom  $\alpha$ -DM1 antibody showed a range of non-specific bands in both patients and controls and we were therefore unable to determine the origin of this protein with certainty. In addition, it is difficult to know the exact size of the polyGln RAN protein produced by the DM1-AS, as the disease is prone to somatic mosaicism [31,32]. Nevertheless, Zu and collaborators showed a band just below 60 kD in a patient carrying 85 CTG·CAGs [221]. Our patients carried expansions much larger than that, and when estimating the molecular weight based on the CTG expansion size, it was possible to have polyGln RAN proteins in the range we found within the lymphoblastoid cell lines. This will remain, however, hypothetical, as it seems we do not have a proper functional custom DM1 polyGln RAN antibody and no positive control available to test its functionality. The aggregate found with immunofluorescence co-localized with the Golgi apparatus, which might indicate the detection of another endogenous polyGln-containing protein. For example, ataxin-2, the product of the SCA2 gene, contains 22 glutamines and resides in the Golgi apparatus [328]. In addition, our immunofluorescence did show staining of the nucleus at high antibody concentrations, which might be due to

binding of the transcription factor TBP, also detected by the immunoblots (42 kD band). Taken together, this would mean that both commercial anti-polyGln antibodies bind to several endogenous polyGln-containing proteins, especially at higher antibody concentrations. However, the previous reported polyGln RAN proteins were not found, even though two antibodies were the same used as in the previous report and similar cells were used. Our DM1-AS results suggested that the presence of DM1-AS transcripts containing the expanded repeat in the cytoplasm of DM1 cells is quite a rare occurrence. This highly affects the chance of producing polyGln RAN proteins. In addition, polyGln-containing proteins are very common in healthy subjects. Taking these two notions together, it might be plausible that with current techniques, sensitivity is too low to detect such low quantities of the polyGln RAN protein, which in addition is hindered by the presence of other polyGln containing proteins.

Although we were unable to detect polyGln RAN proteins in our DM1 cells, much progress has been made in other repeat expansion disorders displaying RAN translation, which could help in the field of DM1. Seven expansion disorders have been added since the first discovery of RAN translation in SCA8 and DM1 [222,238,241,244–246,329]. Of these, SCA8, SCA3 and HD are the three repeat expansion disorders in which the RAN proteins originate from a CAG expansion, and can therefore result in polyGln RAN proteins. Interestingly, *in vivo*, none of the diseases show polyGln RAN proteins, but instead produce poly-alanine, and for HD additionally poly-serine RAN proteins. It might be interesting to include custom antibodies for the two additional homo-polymeric protein possibilities with regard to DM1.

Taken together, DM1-AS transcript levels were lower in patients compared to controls and were present in both the nucleus and the cytoplasm of DM1 cells. Further research into the effect of expanded DM1-AS transcripts on MBNL1 sequestration and overall DM1 pathology is needed to understand their link to clinical phenotype. Only a small portion of the DM1-AS transcripts contained the expanded repeat, substantially lowering the possibility of RAN translation in DM1. The polyGln RAN protein was not present in patient-derived DM1 cells, or was present in such low quantities that it is below the detection limit of the currently available techniques, suggesting RAN translation might not play a major role in DM1 pathology. The development of more sensitive techniques and antibodies with higher specificity are needed to understand the real contribution of RAN translation to DM1.

### DNA METHYLATION PROFILES ACROSS THE *DMPK* LOCUS

---

The fourth disease modifier studied was DNA methylation and the overall goal of this study was to investigate the contribution of epigenetics to DM1 pathology, by analyzing the DNA methylation profiles across the *DMPK* locus in several tissues and tissue-derived cells in all clinical subtypes of DM1. So far, all the published studies in DM1 have only addressed the DNA methylation pattern in the region containing the expanded repeat, which

overlaps with a big CpG island (named CpGi 374), and where CTCF can be bound at two distinct places (named CTCF1 and CTCF2). This was considered particularly relevant because the *DMPK* gene is found in a chromosomal location containing a high density of genes, being most of them almost exclusively expressed in testis, and there is overlapping between the 3' UTR of one gene with the regulatory regions of the next. The *SIX5* transcription start site is less than 400bp from the 3' end region of the *DMPK*, and more importantly, the *SIX5* enhancer/promoter resides in CpGi 374. The gene adjacent to the 5' end region of the *DMPK* gene, *DMWD*, is also within 500 bp from the gene (Figure 10 introduction). Importantly, CTCF is a transcription factor that can function as an insulator and hypermethylation of the CTCF regions could inhibit CTCF-binding and disrupt the insulator element altering gene expression [292].

Our results comparing blood samples highlight a gain of DNA methylation in the CTCF1 region of DM1 developmental cases, which is more prevalent with earlier onset and correlates with higher disease severity and CTG expansion length. The upward gradient observed, has been described previously [55,297–299], but the trend was not as clear as ours, since hypermethylation was more concentrated towards the congenital subtype. Additionally, fifty percent of the CTCF1 hypermethylated cases showed also hypermethylation in the CTCF2 region. However, CTCF2 methylation without CTCF1 hypermethylation was not observed, which may indicate that the beginning of the aberrant DNA methylation is not random and it spreads beyond the CTG repeat only in certain cases/conditions. Several studies have shown that in non-affected samples, CTCF binds strongly to CTCF1, whereas the CTCF2 binding site is more controversial showing either weaker or no CTCF binding [135,293,295,330]. The preferential hypermethylation observed at the CTCF1 region, mainly in the congenital cases, can therefore potentially be more disruptive and a bigger contributor to DM1 pathology. The CpGs in the binding site region of CTCF1 and CTCF2 were both hypermethylated in our methylated samples, as well as in other studies, which could potentially inhibit CTCF-binding and disrupt the hypothesized insulator element between the *SIX5* enhancer and *DMPK* promoter, resulting in increased *DMPK* levels, meanwhile reducing *SIX5* expression, and worsening DM1 pathology [7,292]. However, some reports showed no differences in *DMPK* and *SIX5* expression levels between methylated and non-methylated samples [298] and Yanovsky-Dagan and collaborators identified a region upstream of CTCF1 binding site that when hypermethylated decreased *SIX5* expression [135]. Additionally, in DM1 adult samples a reduction of *DMPK* levels has extensively been reported, making this theory only applicable for DM1 developmental cases, which are in fact the ones that showed CTCF1 hypermethylation.

Our study is the first to report a more severe disease manifestation at skeletal muscle, cardiac and cognitive level in hypermethylated DM1 childhood cases, as phenotype associations have so far only been performed in adults [294,301]. Furthermore, we found an association with CTG expansion length, where longer CTG expansions are more likely to be

hypermethylated. This is in concordance with previous results [32,296,298,301], although other studies have suggested no such correlations exist or they may not be absolute [55,299]. Additionally, we have found an increased, but not exclusive, maternal transmission in the hypermethylated cases. Barbé and collaborators have suggested the presence of a parent-of-origin effect, where DNA methylation may account for the maternal bias for CDM1 transmission, the larger maternal CTG expansions, age of onset and clinical phenotype [55]. Although we do see a similar trend, as all our CDM1 cases are both maternally transmitted and hypermethylated, for the other clinical phenotypes, parental inheritance is not a good predictor for methylation status or vice versa. This is strengthened by the observation of Morales and collaborators, where a large family showed several paternally transmitted methylated cases [297]. It therefore may be a good diagnostic indicator during prenatal screening, but less efficient as a general disease marker. Of note, a recent study in DM1 spermatozoa found that DNA methylation was not affecting sperm viability and these spermatozoa were compatible with "in vitro" fertilization [331]. These findings go against the hypothesis that reduced survival is associated with methylated spermatozoa, preventing the transmission of CDM1 [7], and therefore, other explanations for this maternal bias should be explored.

Our study cohort included several families, giving us the opportunity to study inheritance of the DNA methylation profiles. We found the CTCF methylation status of our patients was not inheritable, as several unmethylated mothers gave birth to methylated cases. This is in accordance with previous studies [55,297]. Interestingly, we observed that the offspring of methylated mothers carried contractions of the CTG expansion, while the offspring of unmethylated mothers carried expansions of the repeat. This suggests that although the bigger CTG expansion sizes are associated with hypermethylation, when a methylated parent passes on the methylation status, it coincides with the transference of a smaller CTG expansion. Some authors have evaluated the effect of DNA methylation on the stability of the CTG expansion repeat [332]. When using bacterial and primate cellular models of 83 to 100 CTG repeat expansions, DNA methylation was found to be associated with a stabilization of the repeat size. However, caution must be taken with these observations, as the sample size in their study and our study was low and additional studies are needed to further elucidate this observation.

As mentioned before, all previous DNA methylation studies were focused in the CTG repeat region, located in the CpGi 374. However, the *DMPK* locus harbors three more CpG islands, which have never been investigated in DM1 samples. Interestingly, our results showed a novel tissue-specific epigenetic landscape in muscle samples for CpGi 43 and CTCF1, which brings new insights in DM1 epigenetics. Recently, Buckley and collaborators reported the existence of a *DMPK* alternative promoter, which overlaps with CpGi 43, as well as cell type-dependent differences in promoter usage according to epigenetic features [293]. The data presented by Buckley and collaborators showed a predominant use of the canonical

upstream promoter in skeletal muscle and myogenic cells from control individuals [293]. Conversely, in blood, a predominant usage of the alternative *DMPK* promoter was suggested [293]. Interestingly, our study in DM1 muscle samples showed a specific demethylation of this alternative promoter located at CpGi 43 in skeletal muscle tissue and muscle-derived cells. This could potentially alter chromatin conformation and result in a shift of the promoter usage from the strongest/canonical one to the weak/alternative promoter, decreasing *DMPK* expression levels in DM1 myogenic samples.

Additionally, this DM1 tissue-dependent demethylation was accompanied by a gain of methylation in the CTCF1, but not the CTCF2 region. However, the CpG sites residing in the CTCF1 binding site remained unmethylated in muscle samples. This may imply that although there is a disease-specific gain of methylation for the CTCF1 region in muscle, the CTCF binding site is not disrupted, allowing CTCF binding. Nevertheless, CTCF1 hypermethylation might affect other chromatin interactions, in turn affecting gene expression. Buckley and collaborators showed that CTCF regions, as well as intragenic regions of *DMWD* and radial spoke head 6 homolog A genes (*RSPH6A* genes), located next to the *DMPK* gene, display enhancer chromatin features in control muscle cells [293]. Our findings showing the hypermethylation of the CTCF1 region in DM1 muscle samples may indicate loss of chromatin interactions between the *DMPK* promoter and these potential myogenic enhancers. Interestingly, Brouwer and collaborators showed an increase of the H3K9me3 chromatin repressive mark, together with gain of DNA methylation, in the CTCF1 region (and to a lesser extent in CTCF2) in heart tissue of DM1 mice, which correlated with decreased *DMPK* expression [295]. However, to demonstrate that DNA methylation changes in the CTCF1 region in DM1 skeletal muscles could alter chromatin interactions in the *DMPK* locus, further experiments are needed.

Our study addressed for first time the DNA methylation status of patients-derived DM1 cells. Our results showed that most of the tissue-derived primary cells showed very similar DNA methylation patterns as the tissue of origin. However, in some cases we observed a slightly higher increase in methylation levels in cells versus the corresponding tissues. This can be explained by the observation that cellular models, especially immortalized cell lines or primary cell cultures that have been in culture for a substantial amount of time, can increase DNA methylation levels [253,333,334], and/or because of the purity of cell cultures compared to tissues containing distinct cell types. Overall, our results showed that the DM1 patient-derived cells preserve the genetic and epigenetic features, which make them excellent models to study DM1 pathology.

Altogether, our results showed a distinct DNA methylation profile across DM1 tissues and uncovered a novel and dual epigenetic signature involving gain of DNA methylation in the flanking region of CTG expansion, accompanied by specific DNA demethylation in the *DMPK* gene body of DM1 samples, which highlighted the contribution of epigenetic changes



to DM1 pathology. Further research is needed to understand the impact of this dual myogenic epigenetic signature on gene expression and DM1 pathology.

#### THE BIOMARKER POTENTIAL OF miRNAs

---

The last disease modifier studied are miRNA expression profiles. miRNAs have great potential to become diagnostic tools in several diseases, including cancer, neurodegenerative diseases, and neuromuscular diseases. However, to date, insufficient knowledge and a lack of conclusive results to clarify the role of miRNA in disease diagnosis have held back the implementation of miRNA biomarkers in clinical settings [313]. In the past two decades, extensive research has been conducted in the miRNA expression profiles of DM1 patients and their biomarker potential. Several expression profiles have been found to be able to distinguish between DM1 patients and healthy subjects, and even a link has been made between progressive and non-progressive muscle wasting in DM1 patients. However, discrepancy between studies still exists.

The vast difference in results between the studies, which at first glance look very similar in set-up, might be explained by the use of different tissues of origin. Studies in muscle use biopsy material from the vastus lateralis [314,316], the biceps brachii [317], or a mixture of material from the vastus lateralis, biceps brachii, and deltoid [318]. Since in DM1, in general, the observed muscle weakness is prominently distal, and only in later stages of the diseases are the proximal muscles, such as the bicep brachii and the vastus lateralis, involved, it could be that the stage of each individual patient at the time of biopsy can influence the obtained results. The studies that focus on blood use different blood components as well, i.e., serum, plasma or whole blood. It has been shown that the miRNA expression levels differ between serum and plasma, for example due to the activation and release of miRNAs from platelets during collection of plasma [335]. Moreover, different clinical subtypes are used across the reported studies or clinical subtypes are not mentioned. The different clinical subtypes can potentially differ in their miRNA profiles. Childhood-onset DM1 patients display a different array of symptoms compared to adult-onset DM1, suggesting that different pathological mechanisms are at play.

Another reason for the discrepancies found could be the use of quantitative PCR (qPCR) to study miRNA expression levels, where expression levels are normalized against an endogenous miRNA of which the levels are known to be stable or by using a spike-in miRNA. Total miRNA content has been shown to be an experimental variable and adds a systematic bias in miRNA quantification [336]. No gold standard for this type of normalization exists, and groups often use different approaches, decreasing the comparability of their studies. A way around the normalization process might be the use of a fairly novel technique, the digital droplet PCR, which produces an absolute quantity, eliminating the need for normalization [337]. This will help in the comparability of results between groups. Sample size and type of patient cohort seems to be of influence as well. Both Koutsoulidou and collaborators [323]

and Perfetti and collaborators [322] tried to validate their initial findings in a bigger cohort, but obtained different results, highlighting the importance of sample size and the specific group of patients included [324,325].

Further research, especially longitudinal studies, are needed to unravel the true biomarker potential of miRNAs in DM1 and to see whether they can help in the prediction of disease progression and/or in the prediction of treatment efficacy.

## CONCLUDING REMARKS

---

The results lined out in this doctoral thesis have highlighted the complexity of DM1 pathology. Here, we have addressed the contribution of genetic and epigenetic modifiers to DM1 pathology. Our results have shed light on the contribution of variant repeats, antisense transcription, RAN translation and changes in DNA methylation and miRNA expression levels to DM1 pathology. There is no doubt that a full understanding of the entire DM1 pathology is needed to be able to find new therapeutic targets for this incurable disease, but also to find prognostic markers for disease development to facilitate symptom management. Although further research is needed, the data presented in this dissertation will contribute to a better understanding of this complex and multi-systemic disease.





# CONCLUSIONS



**Chapter I**

- Variant repeats can be linked to a delay of disease onset, but not necessarily a milder phenotype, as an aging-related severe disease manifestation might be present.
- Atypical features can be present in variant carrying individuals, hampering disease diagnosis.
- Intergenerational transmission can result in both contractions and expansions, the latter linked to anticipation.

**Chapter II**

- DM1-AS transcript levels are lower in DM1 patients compared to controls, and can be found in both nucleus and cytoplasm as a heterogeneous pool with and without the expanded repeat.
- Only a small portion of the DM1-AS transcripts contained the expanded repeat, substantially lowering the possibility of RAN translation in DM1.
- The polyGln RAN protein is not present in patient-derived DM1 cells, or is present in such low quantities that it is below the detection limit of the currently available techniques. This raises the question whether RAN translation contributes to DM1 pathology.

**Chapter III**

- A gain of DNA methylation is present in the CTCF1 region of DM1 developmental cases, which is more prevalent with earlier onset.
- Childhood hypermethylated patients present a more severe muscular, cardiac and cognitive manifestation of the disease.

## Conclusions

- Longer CTG expansions are associated with a higher chance of hypermethylation.
- Hypermethylation of the CTCF1 regions is linked to increased, but not exclusive, maternal transmission of the expanded allele.
- DNA methylation status in DM1 patients is not inheritable. Parental hypermethylation is associated with CTG size contractions, instead of expansions, in offspring.
- There is a muscle-specific dual epigenetic signature involving gain of DNA methylation in the flanking region of the CTG expansion, accompanied by specific DNA demethylation present in the *DMPK* gene body of DM1 samples.

## Chapter IV

- Extensive research into the miRNA expression profiles of DM1 patients and their biomarker potential reveal a global deregulation of miRNA expression in DM1.
- miRNA expression profiles can distinguish between DM1 patients and healthy subjects, and can be linked to clinical phenotype.
- Further research, especially longitudinal studies, are needed to unravel the true disease modulating and biomarker potential of miRNAs in DM1 to see whether they can help in the prediction of disease progression and/or in the prediction of treatment efficacy.



# **BIBLIOGRAPHY**





1. Harper PS *Major Problems in Neurology: Myotonic Dystrophy*; 3rd ed.; WB Saunders: London, UK, 2001;
2. Johnson, N.E.; Butterfield, R.J.; Mayne, K.; Newcomb, T.; Imburgia, C.; Dunn, D.; Duval, B.; Feldkamp, M.L.; Weiss, R.B. Population-Based Prevalence of Myotonic Dystrophy Type 1 Using Genetic Analysis of Statewide Blood Screening Program. *Neurology* **2021**, *96*, e1045–e1053, doi:10.1212/WNL.0000000000011425.
3. Mishra, S.K.; Singh, S.; Lee, B.; Khosa, S.; Moheb, N.; Tandon, V.A. "Dystrophia Myotonica" and the Legacy of Hans Gustav Wilhelm Steinert. *Ann. Indian Acad. Neurol.* **2018**, *21*, 116, doi:10.4103/AIAN.AIAN\_182\_17.
4. Peric, S.; Vujnic, M.; Dobricic, V.; Marjanovic, A.; Basta, I.; Novakovic, I.; Lavrnica, D.; Rakocevic-Stojanovic, V. Five-year study of quality of life in myotonic dystrophy. *Acta Neurol. Scand.* **2016**, *134*, 346–351, doi:10.1111/ANE.12549.
5. Mathieu, J.; Allard, P.; Potvin, L.; Prévost, C.; Bégin, P. A 10-year study of mortality in a cohort of patients with myotonic dystrophy. *Neurology* **1999**, *52*, 1658–62.
6. De Antonio, M.; Dogan, C.; Hamroun, D.; Mati, M.; Zerrouki, S.; Eymard, B.; Katsahian, S.; Bassez, G.; French Myotonic Dystrophy Clinical Network Unravelling the myotonic dystrophy type 1 clinical spectrum: A systematic registry-based study with implications for disease classification. *Rev. Neurol. (Paris)*. **2016**, *172*, 572–580, doi:10.1016/j.neurol.2016.08.003.
7. Lanni, S.; Pearson, C.E. Molecular genetics of congenital myotonic dystrophy. *Neurobiol. Dis.* **2019**, *132*, doi:10.1016/j.nbd.2019.104533.
8. Campbell, C.; Levin, S.; Siu, V.M.; Venance, S.; Jacob, P. Congenital myotonic dystrophy: Canadian population-based surveillance study. *J. Pediatr.* **2013**, *163*, doi:10.1016/J.JPEDS.2012.12.070.
9. Echenne, B.; Bassez, G. Congenital and infantile myotonic dystrophy. *Handb. Clin. Neurol.* **2013**, *113*, 1387–1393, doi:10.1016/B978-0-444-59565-2.00009-5.
10. Ho, G.; Carey, K.A.; Cardamone, M.; Farrar, M.A. Myotonic dystrophy type 1: Clinical manifestations in children and adolescents. *Arch. Dis. Child.* **2019**, *104*, 48–52, doi:10.1136/ARCHDISCHILD-2018-314837.
11. Johnson, N.E.; Butterfield, R.; Berggren, K.; Hung, M.; Chen, W.; Dibella, D.; Dixon, M.; Hayes, H.; Pucillo, E.; Bounsanga, J.; et al. Disease burden and functional outcomes in congenital myotonic dystrophy: A cross-sectional study. *Neurology* **2016**, *87*, 160–167, doi:10.1212/WNL.0000000000002845.
12. Ho, G.; Cardamone, M.; Farrar, M. Congenital and childhood myotonic dystrophy: Current aspects of disease and future directions. *World J. Clin. Pediatr.* **2015**, *4*, 66–80, doi:10.5409/wjcp.v4.i4.66.
13. Campbell, C.; Sherlock, R.; Jacob, P.; Blayney, M. Congenital myotonic dystrophy: assisted ventilation duration and outcome. *Pediatrics* **2004**, *113*, 811–6.
14. Meola, G.; Cardani, R. Myotonic dystrophies: An update on clinical aspects, genetic, pathology, and molecular pathomechanisms. *Biochim. Biophys. Acta - Mol. Basis Dis.* **2015**, *1852*, 594–606, doi:10.1016/j.bbdis.2014.05.019.
15. Echenne, B.; Rideau, A.; Roubertie, A.; Sébire, G.; Rivier, F.; Lemieux, B. Myotonic dystrophy type I in childhood Long-term evolution in patients surviving the neonatal period. *Eur. J. Paediatr. Neurol.* **2008**, *12*, 210–223,

- doi:10.1016/J.EJPN.2007.07.014.
16. Ekström, A.B.; Hakenäs-Plate, Louise; Tulinius, M.; Wentz, E. Cognition and adaptive skills in myotonic dystrophy type 1: A study of 55 individuals with congenital and childhood forms. *Dev. Med. Child Neurol.* **2009**, *51*, 982–990, doi:10.1111/j.1469-8749.2009.03300.x.
  17. Douniol, M.; Jacqueline, A.; Cohen, D.; Bodeau, N.; Rachidi, L.; Angeard, N.; Cuisset, J.-M.; Vallée, L.; Eymard, B.; Plaza, M.; et al. Psychiatric and cognitive phenotype of childhood myotonic dystrophy type 1. *Dev. Med. Child Neurol.* **2012**, *54*, 905–11, doi:10.1111/j.1469-8749.2012.04379.x.
  18. Udd, B.; Krahe, R. The myotonic dystrophies: molecular, clinical, and therapeutic challenges. *Lancet Neurol.* **2012**, *11*, 891–905, doi:10.1016/S1474-4422(12)70204-1.
  19. Groh, W.J.; Groh, M.R.; Saha, C.; Kincaid, J.C.; Simmons, Z.; Ciafaloni, E.; Pourmand, R.; Otten, R.F.; Bhakta, D.; Nair, G. V; et al. Electrocardiographic abnormalities and sudden death in myotonic dystrophy type 1. *N. Engl. J. Med.* **2008**, *358*, 2688–97, doi:10.1056/NEJMoa062800.
  20. Johnson, N.E.; Abbott, D.; Cannon-Albright, L.A. Relative risks for comorbidities associated with myotonic dystrophy: A population-based analysis. *Muscle Nerve* **2015**, *52*, 659–661, doi:10.1002/MUS.24766.
  21. Heatwole, C.; Bode, R.; Johnson, N.; Quinn, C.; Martens, W.; McDermott, M.P.; Rothrock, N.; Thornton, C.; Vickrey, B.; Victorson, D.; et al. Patient-reported impact of symptoms in myotonic dystrophy type 1 (PRISM-1). *Neurology* **2012**, *79*, 348–357, doi:10.1212/WNL.0B013E318260CBE6.
  22. Bellini, M.; Biagi, S.; Stasi, C.; Costa, F.; Mumolo, M.G.; Ricchiuti, A.; Marchi, S. Gastrointestinal manifestations in myotonic muscular dystrophy. *World J. Gastroenterol.* **2006**, *12*, 1821–8.
  23. Rönblom, A.; Forsberg, H.; Danielsson, Å. Gastrointestinal symptoms in myotonic dystrophy. *Scand. J. Gastroenterol.* **1996**, *31*, 654–657, doi:10.3109/00365529609009145.
  24. Johnson, N.E. Myotonic Muscular Dystrophies. *Continuum (Minneapolis, Minn.)* **2019**, *25*, 1682–1695, doi:10.1212/CON.0000000000000793.
  25. Martorell, L.; Monckton, D.G.; Sanchez, A.; Lopez de Munain, A.; Baiget, M. Frequency and stability of the myotonic dystrophy type 1 premutation. *Neurology* **2001**, *56*, 328–335, doi:10.1212/WNL.56.3.328.
  26. Pešović, J.; Perić, S.; Brkušanić, M.; Brajušković, G.; Rakočević-Stojanović, V.; Savić-Pavičević, D. Molecular genetic and clinical characterization of myotonic dystrophy type 1 patients carrying variant repeats within DMPK expansions. *Neurogenetics* **2017**, *18*, 207–218, doi:10.1007/s10048-017-0523-7.
  27. Abbruzzese, C.; Porrini, S.C.; Mariani, B.; Gould, F.K.; Mcabney, J.P.; Monckton, D.G.; Ashizawa, T.; Giacanelli, M. Instability of a premutation allele in homozygous patients with myotonic dystrophy type 1. *Ann. Neurol.* **2002**, *52*, 435–441, doi:10.1002/ana.10304.
  28. Barceló, J.M.; Mahadevan, M.S.; Tsilfidis, C.; Mackenzie, A.E.; Korneluk, R.G. Intergenerational stability of the myotonic dystrophy protomutation. *Hum. Mol. Genet.* **1993**, *2*, 705–709, doi:10.1093/HMG/2.6.705.

29. Peterlin, B.; Logar, N.; Zidar, J. CTG repeat analysis in lymphocytes, muscles and fibroblasts in patients with myotonic dystrophy. *Pflugers Arch.* **1996**, *431*, R199-200, doi:10.1007/bf02346337.
30. Ballester-Lopez, A.; Koehorst, E.; Linares-Pardo, I.; Núñez-Manchón, J.; Almendrote, M.; Lucente, G.; Arbex, A.; Alonso, C.P.; Lucia, A.; Monckton, D.G.; et al. Preliminary findings on ctg expansion determination in different tissues from patients with myotonic dystrophy type 1. *Genes (Basel)*. **2020**, *11*, 1–8, doi:10.3390/genes11111321.
31. Morales, F.; Couto, J.M.; Higham, C.F.; Hogg, G.; Cuenca, P.; Braida, C.; Wilson, R.H.; Adam, B.; Del Valle, G.; Brian, R.; et al. Somatic instability of the expanded CTG triplet repeat in myotonic dystrophy type 1 is a heritable quantitative trait and modifier of disease severity. *Hum. Mol. Genet.* **2012**, *21*, 3558–3567, doi:10.1093/HMG/DDS185.
32. Morales, F.; Vásquez, M.; Corrales, E.; Vindas-Smith, R.; Santamaría-Ulloa, C.; Zhang, B.; Siritto, M.; Estecio, M.R.; Krahe, R.; Monckton, D.G. Longitudinal increases in somatic mosaicism of the expanded CTG repeat in myotonic dystrophy type 1 are associated with variation in age-at-onset. *Hum. Mol. Genet.* **2020**, *29*, 2496–2507, doi:10.1093/hmg/ddaa123.
33. Overend, G.; Légaré, C.; Mathieu, J.; Bouchard, L.; Gagnon, C.; Monckton, D.G. Allele length of the DMPK CTG repeat is a predictor of progressive myotonic dystrophy type 1 phenotypes. *Hum. Mol. Genet.* **2019**, *28*, 2245–2254, doi:10.1093/hmg/ddz055.
34. Cumming, S.A.; Jimenez-Moreno, C.; Okkersen, K.; Wenninger, S.; Daidj, F.; Hogarth, F.; Littleford, R.; Gorman, G.; Bassez, G.; Schoser, B.; et al. Genetic determinants of disease severity in the myotonic dystrophy type 1 OPTIMISTIC cohort. *Neurology* **2019**, *93*, E995–E1009, doi:10.1212/WNL.0000000000008056.
35. Anvret, M.; Ahlberg, G.; Grandell, U.; Hedberg, B.; Johnson, K.; Edström, L. Larger expansions of the CTG repeat in muscle compared to lymphocytes from patients with myotonic dystrophy. *Hum. Mol. Genet.* **1993**, *2*, 1397–400, doi:10.1093/hmg/2.9.1397.
36. Thornton, C.A.; Johnson, K.; Moxley, R.T. Myotonic dystrophy patients have larger CTG expansions in skeletal muscle than in leukocytes. *Ann. Neurol.* **1994**, *35*, 104–7, doi:10.1002/ana.410350116.
37. Ohya, K.; Tachi, N.; Kon, S. ichiro; Kikuchi, K.; Chiba, S. Somatic cell heterogeneity between DNA extracted from lymphocytes and skeletal muscle in congenital myotonic dystrophy. *Jpn. J. Hum. Genet.* **1995**, *40*, 319–326, doi:10.1007/BF01900598.
38. Martorell, L.; Monckton, D.G.; Gamez, J.; Johnson, K.J.; Gich, I.; Lopez De Munain, A.; Baiget, M. Progression of somatic CTG repeat length heterogeneity in the blood cells of myotonic dystrophy patients. *Hum. Mol. Genet.* **1998**, *7*, 307–312, doi:10.1093/HMG/7.2.307.
39. Martorell, L.; Martinez, J.M.; Carey, N.; Johnson, K.; Baiget, M. Comparison of CTG repeat length expansion and clinical progression of myotonic dystrophy over a five year period. *J. Med. Genet.* **1995**, *32*, 593–596, doi:10.1136/JMG.32.8.593.
40. Monckton, D.G.; Wong, L.J.; Ashizawa, T.; Caskey, C.T. Somatic mosaicism, germline expansions, germline reversions and intergenerational reductions in myotonic dystrophy males: small pool PCR analyses. *Hum. Mol. Genet.* **1995**, *4*, 1–8, doi:10.1093/hmg/4.1.1.

41. Morales, F.; Vásquez, M.; Santamaría, C.; Cuenca, P.; Corrales, E.; Monckton, D.G. A polymorphism in the MSH3 mismatch repair gene is associated with the levels of somatic instability of the expanded CTG repeat in the blood DNA of myotonic dystrophy type 1 patients. *DNA Repair (Amst)*. **2016**, *40*, 57–66, doi:10.1016/j.dnarep.2016.01.001.
42. Peric, S.; Pesovic, J.; Savic-Pavicevic, D.; Stojanovic, V.R.; Meola, G. Molecular and Clinical Implications of Variant Repeats in Myotonic Dystrophy Type 1. *Int. J. Mol. Sci.* **2021**, *23*, doi:10.3390/IJMS23010354.
43. Braida, C.; Stefanatos, R.K.A.A.; Adam, B.; Mahajan, N.; Smeets, H.J.M.M.; Niel, F.; Goizet, C.; Arveiler, B.; Koenig, M.; Lagier-Tourenne, C.; et al. Variant CCG and GGC repeats within the CTG expansion dramatically modify mutational dynamics and likely contribute toward unusual symptoms in some myotonic dystrophy type 1 patients. *Hum. Mol. Genet.* **2010**, *19*, 1399–1412, doi:10.1093/hmg/ddq015.
44. Cumming, S.A.; Hamilton, M.J.; Robb, Y.; Gregory, H.; McWilliam, C.; Cooper, A.; Adam, B.; McGhie, J.; Hamilton, G.; Herzyk, P.; et al. De novo repeat interruptions are associated with reduced somatic instability and mild or absent clinical features in myotonic dystrophy type 1. *Eur. J. Hum. Genet.* **2018**, doi:10.1038/s41431-018-0156-9.
45. Ashizawa, T.; Anvret, M.; Baiget, M.; Barceló, J.M.; Brunner, H.; Cobo, A.M.; Dallapiccola, B.; Fenwick, R.G.; Grandell, U.; Harley, H.; et al. Characteristics of Intergenerational Contractions of the CTG Repeat in Myotonic Dystrophy. *Am. J. Hum. Genet.* **1994**, *54*, 414.
46. Harley, H.G.; Rundle, S.A.; MacMillan, J.C.; Myring, J.; Brook, J.D.; Crow, S.; Reardon, W.; Fenton, I.; Shaw, D.J.; Harper, P.S. Size of the unstable CTG repeat sequence in relation to phenotype and parental transmission in myotonic dystrophy. *Am. J. Hum. Genet.* **1993**, *52*, 1164–74.
47. Ashizawa, T.; Dunne, P.W.; Ward, P.A.; Seltzer, W.K.; Richards, C.S. Effects of the sex of myotonic dystrophy patients on the unstable triplet repeat in their affected offspring. *Neurology* **1994**, *44*, 120–2, doi:10.1212/wnl.44.1.120.
48. Lavedan, C.; Hofmann-Radvanyi, H.; Shelbourne, P.; Rabes, J.P.; Duros, C.; Savoy, D.; Dehaupas, I.; Luce, S.; Johnson, K.; Junien, C. Myotonic dystrophy: size- and sex-dependent dynamics of CTG meiotic instability, and somatic mosaicism. *Am. J. Hum. Genet.* **1993**, *52*, 875–883.
49. López De Munain, A.; Cobo, A.M.; Poza, J.J.; Navarrete, D.; Martorell, L.; Palau, F.; Emparanza, J.I.; Baiget, M. Influence of the sex of the transmitting grandparent in congenital myotonic dystrophy. *J. Med. Genet.* **1995**, *32*, 689–691, doi:10.1136/JMG.32.9.689.
50. Morales, F.; Vásquez, M.; Cuenca, P.; Campos, D.; Santamaría, C.; Del Valle, G.; Brian, R.; Sittenfeld, M.; Monckton, D.G. Parental age effects, but no evidence for an intrauterine effect in the transmission of myotonic dystrophy type 1. *Eur. J. Hum. Genet.* **2015**, *23*, 646, doi:10.1038/EJHG.2014.138.
51. Martorell, L.; Gámez, J.; Cayuela, M.L.; Gould, F.K.; McAbney, J.P.; Ashizawa, T.; Monckton, D.G.; Baiget, M. Germline mutational dynamics in myotonic dystrophy type 1 males: allele length and age effects. *Neurology* **2004**, *62*, 269–274, doi:10.1212/WNL.62.2.269.
52. Novelli, G.; Gennarelli, M.; Menegazzo, E.; Angelini, C.; Dallapiccola, B. Discordant clinical outcome in myotonic dystrophy relatives showing (CTG) $n > 700$  repeats. *Neuromuscul. Disord.* **1995**, *5*, 157–159, doi:10.1016/0960-8966(94)00044-A.

53. Cobo, A.M.; Poza, J.J.; Martorell, L.; de Munain, I.; Emparanza, J.I.; Baiget, M.; Cobo J J Poza, A.M.; Lopez de Munain I Emparanza, A.J.; Creu Sant Pau, S.; Martorell M Baiget, S.L. Contribution of molecular analyses to the estimation of the risk of congenital myotonic dystrophy., doi:10.1136/jmg.32.2.105.
54. Koch, M.C.; Grimm, T.; Harley, H.G.; Harper, P.S. Genetic Risks for Children of Women with Myotonic Dystrophy. *Am. J. Hum. Genet* **1991**, *48*, 1084–1091.
55. Barbé, L.; Lanni, S.; López-Castel, A.; Franck, S.; Spits, C.; Keymolen, K.; Seneca, S.; Tomé, S.; Miron, I.; Letourneau, J.; et al. CpG Methylation, a Parent-of-Origin Effect for Maternal-Biased Transmission of Congenital Myotonic Dystrophy. *Am. J. Hum. Genet.* **2017**, *100*, 488–505, doi:10.1016/j.ajhg.2017.01.033.
56. Jansen, G.; Willems, P.; Coerwinkel, M.; Nillesen, W.; Smeets, H.; Vits, L.; Howeler, C.; Brunner, H.; Wieringa, B. Gonosomal mosaicism in myotonic dystrophy patients: involvement of mitotic events in (CTG)<sub>n</sub> repeat variation and selection against extreme expansion in sperm. *Am. J. Hum. Genet.* **1994**, *54*, 575.
57. De Temmerman, N.; Sermon, K.; Seneca, S.; De Rycke, M.; Hilven, P.; Lissens, W.; Van Steirteghem, A.; Liebaers, I. Intergenerational instability of the expanded CTG repeat in the DMPK gene: studies in human gametes and preimplantation embryos. *Am. J. Hum. Genet.* **2004**, *75*, 325–329, doi:10.1086/422762.
58. Martorell, L.; Monckton, D.G.; Gamez, J.; Baiget, M. Complex patterns of male germline instability and somatic mosaicism in myotonic dystrophy type 1. *Eur. J. Hum. Genet.* **2000**, *8*, 423–430, doi:10.1038/SJ.EJHG.5200478.
59. Seriola, A.; Spits, C.; Simard, J.P.; Hilven, P.; Haentjens, P.; Pearson, C.E.; Sermon, K. Huntington's and myotonic dystrophy hESCs: down-regulated trinucleotide repeat instability and mismatch repair machinery expression upon differentiation. *Hum. Mol. Genet.* **2011**, *20*, 176–185, doi:10.1093/HMG/DDQ456.
60. Brook, J.D.; McCurrach, M.E.; Harley, H.G.; Buckler, A.J.; Church, D.; Aburatani, H.; Hunter, K.; Stanton, V.P.; Thirion, J.P.; Hudson, T.; et al. Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* **1992**, *68*, 799–808, doi:10.1016/0092-8674(92)90154-5.
61. Fu, Y.H.; Friedman, D.L.; Richards, S.; Pearlman, J.A.; Gibbs, R.A.; Pizzuti, A.; Ashizawa, T.; Perryman, M.B.; Scarlato, G.; Fenwick, R.G.; et al. Decreased expression of myotonin-protein kinase messenger RNA and protein in adult form of myotonic dystrophy. *Science* **1993**, *260*, 235–238, doi:10.1126/SCIENCE.8469976.
62. Berul, C.I.; Maguire, C.T.; Aronovitz, M.J.; Greenwood, J.; Miller, C.; Gehrmann, J.; Housman, D.; Mendelsohn, M.E.; Reddy, S. DMPK dosage alterations result in atrioventricular conduction abnormalities in a mouse myotonic dystrophy model. *J. Clin. Invest.* **1999**, *103*, doi:10.1172/JCI5346.
63. Jansen, G.; Groenen, P.J.T.A.; Bächner, D.; Jap, P.H.K.; Coerwinkel, M.; Oerlemans, F.; Van Den Broek, W.; Gohlsch, B.; Pette, D.; Plomp, J.J.; et al. Abnormal myotonic dystrophy protein kinase levels produce only mild myopathy in mice. *Nat. Genet.* **1996**, *13*, 316–324, doi:10.1038/NG0796-316.
64. Reddy, S.; Smith, D.B.J.; Rich, M.M.; Lefterovich, J.M.; Reilly, P.; Davis, B.M.; Tran, K.; Rayburn, H.; Bronson, R.; Cros, D.; et al. Mice lacking the myotonic dystrophy protein kinase develop a late onset progressive myopathy. *Nat. Genet.* **1996**, *13*, 325–335, doi:10.1038/ng0796-325.
65. Davis, B.M.; McCurrach, M.E.; Taneja, K.L.; Singer, R.H.; Housman, D.E. Expansion

- of a CUG trinucleotide repeat in the 3' untranslated region of myotonic dystrophy protein kinase transcripts results in nuclear retention of transcripts. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 7388–93.
66. Mankodi, A.; Logigian, E.; Callahan, L.; McClain, C.; White, R.; Henderson, D.; Krym, M.; Thornton, C.A. Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat. *Science* **2000**, *289*, 1769–1772, doi:10.1126/SCIENCE.289.5485.1769.
  67. Seznec, H.; Agbulut, O.; Sergeant, N.; Savouret, C.; Ghestem, A.; Tabti, N.; Willer, J.C.; Ourth, L.; Duros, C.; Brisson, E.; et al. Mice transgenic for the human myotonic dystrophy region with expanded CTG repeats display muscular and brain abnormalities. *Hum. Mol. Genet.* **2001**, *10*, 2717–2726, doi:10.1093/HMG/10.23.2717.
  68. Mahadevan, M.S.; Yadava, R.S.; Yu, Q.; Balijepalli, S.; Frenzel-Mccardell, C.D.; Bourne, T.D.; Phillips, L.H. Reversible model of RNA toxicity and cardiac conduction defects in myotonic dystrophy. *Nat. Genet.* **2006**, *38*, 1066–1070, doi:10.1038/ng1857.
  69. Storbeck, C.; Drmanic, S.; Daniel, K.; Waring, J.D.; Jirik, F.R.; Parry, D.J.; Ahmed, N.; Sabourin, L.A.; Ikeda, J.E.; Korneluk, R.G. Inhibition of myogenesis in transgenic mice expressing the human DMPK 3'-UTR. *Hum. Mol. Genet.* **2004**, *13*, 589–600, doi:10.1093/hmg/ddh064.
  70. Napierała, M.; Krzyzosiak, W.J. CUG repeats present in myotonin kinase RNA form metastable "slippery" hairpins. *J. Biol. Chem.* **1997**, *272*, 31079–31085, doi:10.1074/JBC.272.49.31079.
  71. Holt, I.; Mittal, S.; Furling, D.; Butler-Browne, G.S.; Brook, J.D.; Morris, G.E. Defective mRNA in myotonic dystrophy accumulates at the periphery of nuclear splicing speckles. *Genes Cells* **2007**, *12*, 1035–1048, doi:10.1111/J.1365-2443.2007.01112.X.
  72. Fardaei, M.; Rogers, M.T.; Thorpe, H.M.; Larkin, K.; Hamshere, M.G.; Harper, P.S.; Brook, J.D. Three proteins, MBNL, MBLL and MBXL, co-localize in vivo with nuclear foci of expanded-repeat transcripts in DM1 and DM2 cells. *Hum. Mol. Genet.* **2002**, *11*, 805–14.
  73. Kuyumcu-Martinez, N.M.; Wang, G.S.; Cooper, T.A. Increased Steady-State Levels of CUGBP1 in Myotonic Dystrophy 1 Are Due to PKC-Mediated Hyperphosphorylation. *Mol. Cell* **2007**, *28*, 68–78, doi:10.1016/j.molcel.2007.07.027.
  74. Tapial, J.; Ha, K.C.H.; Sterne-Weiler, T.; Gohr, A.; Braunschweig, U.; Hermoso-Pulido, A.; Quesnel-Vallières, M.; Permanyer, J.; Sodaei, R.; Marquez, Y.; et al. An atlas of alternative splicing profiles and functional associations reveals new regulatory programs and genes that simultaneously express multiple major isoforms. *Genome Res.* **2017**, *27*, 1759–1768, doi:10.1101/GR.220962.117.
  75. Irimia, M.; Blencowe, B.J. Alternative splicing: Decoding an expansive regulatory layer. *Curr. Opin. Cell Biol.* **2012**, *24*, 323–332, doi:10.1016/j.ceb.2012.03.005.
  76. López-Martínez, A.; Soblechero-Martín, P.; De-La-puente-ovejero, L.; Nogales-Gadea, G.; Arechavala-Gomez, V. An overview of alternative splicing defects implicated in myotonic dystrophy type I. *Genes (Basel)*. **2020**, *11*, 1–27.
  77. Pascual, M.; Vicente, M.; Monferrer, L.; Artero, R. The Muscleblind family of proteins: An emerging class of regulators of developmentally programmed

- alternative splicing. *Differentiation* **2006**, *74*, 65–80, doi:10.1111/j.1432-0436.2006.00060.x.
78. Kanadia, R.N.; Johnstone, K.A.; Mankodi, A.; Lungu, C.; Thornton, C.A.; Esson, D.; Timmers, A.M.; Hauswirth, W.W.; Swanson, M.S. A muscleblind knockout model for myotonic dystrophy. *Science* **2003**, *302*, 1978–1980, doi:10.1126/SCIENCE.1088583.
79. Lee, K.S.; Cao, Y.; Witwicka, H.E.; Tom, S.; Tapscott, S.J.; Wang, E.H. RNA-binding protein Muscleblind-like 3 (MBNL3) disrupts myocyte enhancer factor 2 (Mef2) {beta}-exon splicing. *J. Biol. Chem.* **2010**, *285*, 33779–33787, doi:10.1074/JBC.M110.124255.
80. Cass, D.; Hotchko, R.; Barber, P.; Jones, K.; Gates, D.P.; Berglund, J.A. The four Zn fingers of MBNL1 provide a flexible platform for recognition of its RNA binding elements. *BMC Mol. Biol.* **2011**, *12*, doi:10.1186/1471-2199-12-20.
81. Tran, H.; Gourrier, N.; Lemercier-Neuillet, C.; Dhaenens, C.M.; Vautrin, A.; Fernandez-Gomez, F.J.; Arandel, L.; Carpentier, C.; Obriot, H.; Eddarkaoui, S.; et al. Analysis of exonic regions involved in nuclear localization, splicing activity, and dimerization of Muscleblind-like-1 isoforms. *J. Biol. Chem.* **2011**, *286*, 16435–16446, doi:10.1074/JBC.M110.194928.
82. Terenzi, F.; Ladd, A.N. Conserved developmental alternative splicing of muscleblind-like (MBNL) transcripts regulates MBNL localization and activity. *RNA Biol.* **2010**, *7*, 43–55, doi:10.4161/RNA.7.1.10401.
83. Warf, M.B.; Berglund, J.A. MBNL binds similar RNA structures in the CUG repeats of myotonic dystrophy and its pre-mRNA substrate cardiac troponin T. *RNA* **2007**, *13*, 2238–2251, doi:10.1261/RNA.610607.
84. Yuan, Y.; Compton, S.A.; Sobczak, K.; Stenberg, M.G.; Thornton, C.A.; Griffith, J.D.; Swanson, M.S. Muscleblind-like 1 interacts with RNA hairpins in splicing target and pathogenic RNAs. *Nucleic Acids Res.* **2007**, *35*, 5474–5486, doi:10.1093/nar/gkm601.
85. Charlet-B., N.; Savkur, R.S.; Singh, G.; Philips, A. V.; Grice, E.A.; Cooper, T.A. Loss of the muscle-specific chloride channel in type 1 myotonic dystrophy due to misregulated alternative splicing. *Mol. Cell* **2002**, *10*, 45–53, doi:10.1016/S1097-2765(02)00572-5.
86. Mankodi, A.; Takahashi, M.P.; Jiang, H.; Beck, C.L.; Bowers, W.J.; Moxley, R.T.; Cannon, S.C.; Thornton, C.A. Expanded CUG repeats trigger aberrant splicing of CIC-1 chloride channel pre-mRNA and hyperexcitability of skeletal muscle in myotonic dystrophy. *Mol. Cell* **2002**, *10*, 35–44.
87. Fugier, C.; Klein, A.F.; Hammer, C.; Vassilopoulos, S.; Ivarsson, Y.; Toussaint, A.; Tosch, V.; Vignaud, A.; Ferry, A.; Messaddeq, N.; et al. Misregulated alternative splicing of BIN1 is associated with T tubule alterations and muscle weakness in myotonic dystrophy. *Nat. Med.* **2011**, *17*, 720–725, doi:10.1038/NM.2374.
88. Tang, Z.Z.; Yarotsky, V.; Wei, L.; Sobczak, K.; Nakamori, M.; Eichinger, K.; Moxley, R.T.; Dirksen, R.T.; Thornton, C.A. Muscle weakness in myotonic dystrophy associated with misregulated splicing and altered gating of Ca(V)1.1 calcium channel. *Hum. Mol. Genet.* **2012**, *21*, 1312–1324, doi:10.1093/HMG/DDR568.
89. Ho, T.H.; Charlet-B., N.; Poulos, M.G.; Singh, G.; Swanson, M.S.; Cooper, T.A. Muscleblind proteins regulate alternative splicing. *EMBO J.* **2004**, *23*, 3103–3112, doi:10.1038/SJ.EMBOJ.7600300.



## Bibliography

90. Savkur, R.S.; Philips, A. V.; Cooper, T.A. Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy. *Nat. Genet.* **2001**, *29*, 40–47, doi:10.1038/NG704.
91. Freyermuth, F.; Rau, F.; Kokunai, Y.; Linke, T.; Sellier, C.; Nakamori, M.; Kino, Y.; Arandel, L.; Jollet, A.; Thibault, C.; et al. Splicing misregulation of SCN5A contributes to cardiac-conduction delay and heart arrhythmia in myotonic dystrophy. *Nat. Commun.* **2016**, *7*, doi:10.1038/NCOMMS11067.
92. Philips, A. V.; Timchenko, L.T.; Cooper, T.A. Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. *Science (80-. )*. **1998**, *280*, 737–741, doi:10.1126/science.280.5364.737.
93. Charizanis, K.; Lee, K.Y.; Batra, R.; Goodwin, M.; Zhang, C.; Yuan, Y.; Shiue, L.; Cline, M.; Scotti, M.M.; Xia, G.; et al. Muscleblind-like 2-mediated alternative splicing in the developing brain and dysregulation in myotonic dystrophy. *Neuron* **2012**, *75*, 437–450, doi:10.1016/J.NEURON.2012.05.029.
94. Goodwin, M.; Mohan, A.; Batra, R.; Lee, K.Y.; Charizanis, K.; Fernández Gómez, F.J.; Eddarkaoui, S.; Sergeant, N.; Buée, L.; Kimura, T.; et al. MBNL Sequestration by Toxic RNAs and RNA Misprocessing in the Myotonic Dystrophy Brain. *Cell Rep.* **2015**, *12*, 1159–1168, doi:10.1016/J.CELREP.2015.07.029.
95. Fernandez-Gomez, F.; Tran, H.; Dhaenens, C.M.; Caillet-Boudin, M.L.; Schraen-Maschke, S.; Blum, D.; Sablonnière, B.; Buée-Scherrer, V.; Buee, L.; Sergeant, N. Myotonic Dystrophy: an RNA Toxic Gain of Function Tauopathy? In *Advances in Experimental Medicine and Biology*; Springer, 2019; Vol. 1184, pp. 207–216.
96. Timchenko, N.A.; Cai, Z.J.; Welm, A.L.; Reddy, S.; Ashizawa, T.; Timchenko, L.T. RNA CUG repeats sequester CUGBP1 and alter protein levels and activity of CUGBP1. *J. Biol. Chem.* **2001**, *276*, 7820–6, doi:10.1074/jbc.M005960200.
97. Jones, K.; Jin, B.; Iakova, P.; Huichalaf, C.; Sarkar, P.; Schneider-Gold, C.; Schoser, B.; Meola, G.; Shyu, A. Bin; Timchenko, N.; et al. RNA Foci, CUGBP1, and ZNF9 are the primary targets of the mutant CUG and CCUG repeats expanded in myotonic dystrophies type 1 and type 2. *Am. J. Pathol.* **2011**, *179*, 2475–2489, doi:10.1016/J.AJPATH.2011.07.013.
98. Ward, A.J.; Rimer, M.; Killian, J.M.; Dowling, J.J.; Cooper, T.A. CUGBP1 overexpression in mouse skeletal muscle reproduces features of myotonic dystrophy type 1. *Hum. Mol. Genet.* **2010**, *19*, 3614–3622, doi:10.1093/HMG/DDQ277.
99. Ladd, A.N. CUG-BP, Elav-like family (CELF)-mediated alternative splicing regulation in the brain during health and disease. *Mol. Cell. Neurosci.* **2013**, *56*, 456–464, doi:10.1016/J.MCN.2012.12.003.
100. Orengo, J.P.; Chambon, P.; Metzger, D.; Mosier, D.R.; Snipes, G.J.; Cooper, T.A. Expanded CTG repeats within the DMPK 3' UTR causes severe skeletal muscle wasting in an inducible mouse model for myotonic dystrophy. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 2646–2651, doi:10.1073/PNAS.0708519105.
101. Kalsotra, A.; Xiao, X.; Ward, A.J.; Castle, J.C.; Johnson, J.M.; Burge, C.B.; Cooper, T.A. A postnatal switch of CELF and MBNL proteins reprograms alternative splicing in the developing heart. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 20333–20338, doi:10.1073/PNAS.0809045105.
102. Lin, X.; Miller, J.W.; Mankodi, A.; Kanadia, R.N.; Yuan, Y.; Moxley, R.T.; Swanson, M.S.; Thornton, C.A. Failure of MBNL1-dependent post-natal splicing transitions in

- myotonic dystrophy. *Hum. Mol. Genet.* **2006**, *15*, 2087–2097, doi:10.1093/HMG/DDL132.
103. Moraes, K.C.M.; Wilusz, C.J.; Wilusz, J. CUG-BP binds to RNA substrates and recruits PARN deadenylase. *RNA* **2006**, *12*, 1084–1091, doi:10.1261/RNA.59606.
  104. Iakova, P.; Wang, G.L.; Timchenko, L.; Michalak, M.; Pereira-Smith, O.M.; Smith, J.R.; Timchenko, N.A. Competition of CUGBP1 and calreticulin for the regulation of p21 translation determines cell fate. *EMBO J.* **2004**, *23*, 406–417, doi:10.1038/SJ.EMBOJ.7600052.
  105. Timchenko, N.A.; Patel, R.; Iakova, P.; Cai, Z.J.; Quan, L.; Timchenko, L.T. Overexpression of CUG triplet repeat-binding protein, CUGBP1, in mice inhibits myogenesis. *J. Biol. Chem.* **2004**, *279*, 13129–13139, doi:10.1074/JBC.M312923200.
  106. Wang, E.T.; Cody, N.A.L.; Jog, S.; Biancolella, M.; Wang, T.T.; Treacy, D.J.; Luo, S.; Schroth, G.P.; Housman, D.E.; Reddy, S.; et al. Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. *Cell* **2012**, *150*, 710–724, doi:10.1016/J.CELL.2012.06.041.
  107. Paul, S.; Dansithong, W.; Kim, D.; Rossi, J.; Webster, N.J.G.; Comai, L.; Reddy, S. Interaction of muscleblind, CUG-BP1 and hnRNP H proteins in DM1-associated aberrant IR splicing. *EMBO J.* **2006**, *25*, 4271–83, doi:10.1038/sj.emboj.7601296.
  108. Ravel-Chapuis, A.; Bélanger, G.; Yadava, R.S.; Mahadevan, M.S.; DesGroseillers, L.; Côté, J.; Jasmin, B.J. The RNA-binding protein Staufen1 is increased in DM1 skeletal muscle and promotes alternative pre-mRNA splicing. *J. Cell Biol.* **2012**, *196*, 699–712, doi:10.1083/JCB.201108113.
  109. Bachinski, L.L.; Baggerly, K.A.; Neubauer, V.L.; Nixon, T.J.; Raheem, O.; Siritto, M.; Unruh, A.K.; Zhang, J.; Nagarajan, L.; Timchenko, L.T.; et al. Most expression and splicing changes in myotonic dystrophy type 1 and type 2 skeletal muscle are shared with other muscular dystrophies. *Neuromuscul. Disord.* **2014**, *24*, 227–240, doi:10.1016/J.NMD.2013.11.001.
  110. Orenco, J.P.; Ward, A.J.; Cooper, T.A. Alternative splicing dysregulation secondary to skeletal muscle regeneration. *Ann. Neurol.* **2011**, *69*, 681–690, doi:10.1002/ANA.22278.
  111. de Haro, M.; Al-Ramahi, I.; De Gouyon, B.; Ukani, L.; Rosa, A.; Faustino, N.A.; Ashizawa, T.; Cooper, T.A.; Botas, J. MBNL1 and CUGBP1 modify expanded CUG-induced toxicity in a Drosophila model of myotonic dystrophy type 1. *Hum. Mol. Genet.* **2006**, *15*, 2138–2145, doi:10.1093/HMG/DDL137.
  112. Garcia-Lopez, A.; Monferrer, L.; Garcia-Alcover, I.; Vicente-Crespo, M.; Alvarez-Abril, M.C.; Artero, R.D. Genetic and chemical modifiers of a CUG toxicity model in Drosophila. *PLoS One* **2008**, *3*, doi:10.1371/JOURNAL.PONE.0001595.
  113. Klesert, T.R.; Cho, D.H.; Clark, J.I.; Maylie, J.; Adelman, J.; Snider, L.; Yuen, E.C.; Soriano, P.; Tapscott, S.J. Mice deficient in Six5 develop cataracts: implications for myotonic dystrophy. *Nat. Genet.* **2000**, *25*, 105–109, doi:10.1038/75490.
  114. Guiraud-Dogan, C.; Huguet, A.; Gomes-Pereira, M.; Brisson, E.; Bassez, G.; Junien, C.; Gourdon, G. DM1 CTG expansions affect insulin receptor isoforms expression in various tissues of transgenic mice. *Biochim. Biophys. Acta* **2007**, *1772*, 1183–1191, doi:10.1016/J.BBADIS.2007.08.004.
  115. Vignaud, A.; Ferry, A.; Huguet, A.; Baraibar, M.; Trollet, C.; Hyzewicz, J.; Butler-

- Browne, G.; Puymirat, J.; Gourdon, G.; Furling, D. Progressive skeletal muscle weakness in transgenic mice expressing CTG expansions is associated with the activation of the ubiquitin-proteasome pathway. *Neuromuscul. Disord.* **2010**, *20*, 319–325, doi:10.1016/J.NMD.2010.03.006.
116. Gomes-Pereira, M.; Foiry, L.; Nicole, A.; Huguet, A.; Junien, C.; Munnich, A.; Gourdon, G. CTG trinucleotide repeat “big jumps”: large expansions, small mice. *PLoS Genet.* **2007**, *3*, 0488–0491, doi:10.1371/JOURNAL.PGEN.0030052.
117. Lueck, J.D.; Mankodi, A.; Swanson, M.S.; Thornton, C.A.; Dirksen, R.T. Muscle chloride channel dysfunction in two mouse models of myotonic dystrophy. *J. Gen. Physiol.* **2007**, *129*, 79–94, doi:10.1085/JGP.200609635.
118. Lee, J.E.; Bennett, C.F.; Cooper, T.A. RNase H-mediated degradation of toxic RNA in myotonic dystrophy type 1. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 4221–4226, doi:10.1073/PNAS.1117019109/SUPPL\_FILE/PNAS.201117019SI.PDF.
119. Rolfsmeier, M.L.; Lahue, R.S. Stabilizing Effects of Interruptions on Trinucleotide Repeat Expansions in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **2000**, *20*, 173–180, doi:10.1128/MCB.20.1.173-180.2000/ASSET/CBB9B070-61BF-4C6D-8E5B-7E1CC90C805C/ASSETS/GRAPHIC/MB0101262005.JPEG.
120. Miller, J.W.; Urbinati, C.R.; Teng-Umnuay, P.; Stenberg, M.G.; Byrne, B.J.; Thornton, C.A.; Swanson, M.S. Recruitment of human muscleblind proteins to (CUG)<sub>n</sub> expansions associated with myotonic dystrophy. *EMBO J.* **2000**, *19*, 4439–48, doi:10.1093/emboj/19.17.4439.
121. Grande, V.; Hathazi, D.; O’Connor, E.; Marteau, T.; Schara-Schmidt, U.; Hentschel, A.; Gourdon, G.; Nikolenko, N.; Lochmüller, H.; Roos, A. Dysregulation of GSK3β-Target Proteins in Skin Fibroblasts of Myotonic Dystrophy Type 1 (DM1) Patients. *J. Neuromuscul. Dis.* **2021**, *8*, 603–619, doi:10.3233/JND-200558.
122. Xia, G.; Ashizawa, T. Dynamic changes of nuclear RNA foci in proliferating DM1 cells. *Histochem. Cell Biol.* **2015**, *143*, 557–64, doi:10.1007/s00418-015-1315-5.
123. Loro, E.; Rinaldi, F.; Malena, A.; Masiero, E.; Novelli, G.; Angelini, C.; Romeo, V.; Sandri, M.; Botta, A.; Vergani, L. Normal myogenesis and increased apoptosis in myotonic dystrophy type-1 muscle cells. *Cell Death Differ.* **2010**, *17*, 1315–1324, doi:10.1038/cdd.2010.33.
124. Buj-Bello, A.; Furling, D.; Tronchère, H.; Laporte, J.; Lerouge, T.; Butler-Browne, G.S.; Mandel, J.L. Muscle-specific alternative splicing of myotubularin-related 1 gene is impaired in DM1 muscle cells. *Hum. Mol. Genet.* **2002**, *11*, 2297–2307, doi:10.1093/HMG/11.19.2297.
125. Langlois, M.A.; Lee, N.S.; Rossi, J.J.; Puymirat, J. Hammerhead ribozyme-mediated destruction of nuclear foci in myotonic dystrophy myoblasts. *Mol. Ther.* **2003**, *7*, 670–680, doi:10.1016/S1525-0016(03)00068-6.
126. Furling, D.; Doucet, G.; Langlois, M.A.; Timchenko, L.; Belanger, E.; Cossette, L.; Puymirat, J. Viral vector producing antisense RNA restores myotonic dystrophy myoblast functions. *Gene Ther.* **2003**, *10*, 795–802, doi:10.1038/sj.gt.3301955.
127. Larsen, J.; Pettersson, O.J.; Jakobsen, M.; Thomsen, R.; Pedersen, C.B.; Hertz, J.M.; Gregersen, N.; Corydon, T.J.; Jensen, T.G. Myoblasts generated by lentiviral mediated MyoD transduction of myotonic dystrophy type 1 (DM1) fibroblasts can be used for assays of therapeutic molecules. *BMC Res. Notes* **2011**, *4*, doi:10.1186/1756-0500-4-490.

128. Provenzano, C.; Cappella, M.; Valaperta, R.; Cardani, R.; Meola, G.; Martelli, F.; Cardinali, B.; Falcone, G. CRISPR/Cas9-Mediated Deletion of CTG Expansions Recovers Normal Phenotype in Myogenic Cells Derived from Myotonic Dystrophy 1 Patients. *Mol. Ther. - Nucleic Acids* **2017**, *9*, 337–348, doi:10.1016/J.OMTN.2017.10.006/ATTACHMENT/5B437ED5-57A2-4389-81C8-047CF0B4DB52/MMC1.PDF.
129. Arandel, L.; Espinoza, M.P.; Matloka, M.; Bazinet, A.; De Dea Diniz, D.; Naouar, N.; Rau, F.; Jollet, A.; Edom-Vovard, F.; Mamchaoui, K.; et al. Immortalized human myotonic dystrophy muscle cell lines to assess therapeutic compounds. *Dis. Model. Mech.* **2017**, *10*, 487–497, doi:10.1242/DMM.027367.
130. Pantic, B.; Borgia, D.; Giunco, S.; Malena, A.; Kiyono, T.; Salvatori, S.; De Rossi, A.; Giardina, E.; Sangiuolo, F.; Pegoraro, E.; et al. Reliable and versatile immortal muscle cell models from healthy and myotonic dystrophy type 1 primary human myoblasts. *Exp. Cell Res.* **2016**, *342*, 39–51, doi:10.1016/J.YEXCR.2016.02.013.
131. O’Leary, D.A.; Vargas, L.; Sharif, O.; Garcia, M.E.; Sigal, Y.J.; Chow, S.K.; Schmedt, C.; Caldwell, J.S.; Brinker, A.; Engels, I.H. Hts-compatible patient-derived cell-based assay to identify small molecule modulators of aberrant splicing in myotonic dystrophy type 1. *Curr. Chem. Genomics* **2010**, *4*, 9–18, doi:10.2174/1875397301004010009.
132. Marteyn, A.; Maury, Y.; Gauthier, M.M.; Lecuyer, C.; Vernet, R.; Denis, J.A.; Pietu, G.; Peschanski, M.; Martinat, C. Mutant human embryonic stem cells reveal neurite and synapse formation defects in type 1 myotonic dystrophy. *Cell Stem Cell* **2011**, *8*, 434–444, doi:10.1016/J.STEM.2011.02.004.
133. Denis, J.A.; Gauthier, M.; Rachdi, L.; Aubert, S.; Giraud-Triboulet, K.; Poydenot, P.; Benchoua, A.; Champon, B.; Maury, Y.; Baldeschi, C.; et al. mTOR-dependent proliferation defect in human ES-derived neural stem cells affected by myotonic dystrophy type 1. *J. Cell Sci.* **2013**, *126*, 1763–1772, doi:10.1242/JCS.116285.
134. Thomson, J.A.; Itskovitz-Eldor, J.; Shapiro, S.S.; Waknitz, M.A.; Swiergiel, J.J.; Marshall, V.S.; Jones, J.M. Embryonic Stem Cell Lines Derived from Human Blastocysts. *Science (80- )*. **1998**, *282*, 1145–1147, doi:10.1126/SCIENCE.282.5391.1145.
135. Yanovsky-Dagan, S.; Avitzour, M.; Altarescu, G.; Renbaum, P.; Eldar-Geva, T.; Schonberger, O.; Mitrani-Rosenbaum, S.; Levy-Lahad, E.; Birnbaum, R.Y.; Gepstein, L.; et al. Uncovering the Role of Hypermethylation by CTG Expansion in Myotonic Dystrophy Type 1 Using Mutant Human Embryonic Stem Cells. *Stem Cell Reports* **2015**, *5*, 221–231, doi:10.1016/j.stemcr.2015.06.003.
136. Takahashi, K.; Yamanaka, S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* **2006**, *126*, 663–676, doi:10.1016/J.CELL.2006.07.024/ATTACHMENT/A7BA2E0F-99EF-4A86-88AA-418202149347/MMC1.PDF.
137. Yu, J.; Vodyanik, M.A.; Smuga-Otto, K.; Antosiewicz-Bourget, J.; Frane, J.L.; Tian, S.; Nie, J.; Jonsdottir, G.A.; Ruotti, V.; Stewart, R.; et al. Induced pluripotent stem cell lines derived from human somatic cells. *Science (80- )*. **2007**, *318*, 1917–1920, doi:10.1126/SCIENCE.1151526/SUPPL\_FILE/YU.SOM.REVISION.1.PDF.
138. Gao, Y.; Guo, X.; Santostefano, K.; Wang, Y.; Reid, T.; Zeng, D.; Terada, N.; Ashizawa, T.; Xia, G. Genome Therapy of Myotonic Dystrophy Type 1 iPS Cells for Development of Autologous Stem Cell Therapy. *Mol. Ther.* **2016**, *24*, 1378–1387, doi:10.1038/MT.2016.97/ATTACHMENT/0EB0CF1B-4094-44E1-819D-0D64FE1CA154/MMC1.ZIP.

139. Wang, Y.; Wang, Z.; Sun, H.; Shi, C.; Yang, J.; Liu, Y.; Liu, H.; Zhang, S.; Zhang, L.; Xu, Y.; et al. Generation of induced pluripotent stem cell line(ZZUI006-A)from a patient with myotonic dystrophy type 1. *Stem Cell Res.* **2018**, *32*, 61–64, doi:10.1016/J.SCR.2018.08.013.
140. Ueki, J.; Nakamori, M.; Nakamura, M.; Nishikawa, M.; Yoshida, Y.; Tanaka, A.; Morizane, A.; Kamon, M.; Araki, T.; Takahashi, M.P.; et al. Myotonic dystrophy type 1 patient-derived iPSCs for the investigation of CTG repeat instability. *Sci. Rep.* **2017**, *7*, doi:10.1038/SREP42522.
141. Martineau, L.; Racine, V.; Benichou, S.A.; Puymirat, J. Lymphoblastoids cell lines – Derived iPSC line from a 26-year-old myotonic dystrophy type 1 patient carrying (CTG)200 expansion in the DMPK gene: CHUQi001-A. *Stem Cell Res.* **2018**, *26*, 103–106, doi:10.1016/J.SCR.2017.12.010.
142. Serenó, L.; Coma, M.; Rodríguez, M.; Sánchez-Ferrer, P.; Sánchez, M.B.; Gich, I.; Agulló, J.M.; Pérez, M.; Avila, J.; Guardia-Laguarta, C.; et al. A novel GSK-3beta inhibitor reduces Alzheimer's pathology and rescues neuronal loss in vivo. *Neurobiol. Dis.* **2009**, *35*, 359–367, doi:10.1016/J.NBD.2009.05.025.
143. Wei, C.; Jones, K.; Timchenko, N.A.; Timchenko, L. GSK3 $\beta$  is a new therapeutic target for myotonic dystrophy type 1. *Rare Dis. (Austin, Tex.)* **2013**, *1*, e26555, doi:10.4161/RDIS.26555.
144. Wang, M.; Weng, W.-C.; Stock, L.; Lindquist, D.; Martinez, A.; Gourdon, G.; Timchenko, N.; Snape, M.; Timchenko, L. Correction of Glycogen Synthase Kinase 3 $\beta$  in Myotonic Dystrophy 1 Reduces the Mutant RNA and Improves Postnatal Survival of DMSXL Mice. *Mol. Cell. Biol.* **2019**, *39*, doi:10.1128/MCB.00155-19.
145. Horrigan, J.; Gomes, T.B.; Snape, M.; Nikolenko, N.; McMorn, A.; Evans, S.; Yaroshinsky, A.; Della Pasqua, O.; Oosterholt, S.; Lochmüller, H. A Phase 2 Study of AMO-02 (Tideglusib) in Congenital and Childhood-Onset Myotonic Dystrophy Type 1 (DM1). *Pediatr. Neurol.* **2020**, *112*, 84–93, doi:10.1016/J.PEDIATRNEUROL.2020.08.001.
146. Jenquin, J.R.; Yang, H.; Huigens, R.W.; Nakamori, M.; Berglund, J.A. Combination Treatment of Erythromycin and Furamidine Provides Additive and Synergistic Rescue of Mis-Splicing in Myotonic Dystrophy Type 1 Models. *ACS Pharmacol. Transl. Sci.* **2019**, *2*, 247–263, doi:10.1021/ACSPTSCI.9B00020.
147. Nakamori, M.; Taylor, K.; Mochizuki, H.; Sobczak, K.; Takahashi, M.P. Oral administration of erythromycin decreases RNA toxicity in myotonic dystrophy. *Ann. Clin. Transl. Neurol.* **2016**, *3*, 42–54, doi:10.1002/ACN3.271.
148. Arandel, L.; Matloka, M.; Klein, A.F.; Rau, F.; Sureau, A.; Ney, M.; Cordier, A.; Kondili, M.; Polay-Espinoza, M.; Naouar, N.; et al. Reversal of RNA toxicity in myotonic dystrophy via a decoy RNA-binding protein with high affinity for expanded CUG repeats. *Nat. Biomed. Eng.* **2022**, *6*, 207–220, doi:10.1038/S41551-021-00838-2.
149. Laustriat, D.; Gide, J.; Barrault, L.; Chautard, E.; Benoit, C.; Auboeuf, D.; Boland, A.; Battail, C.; Artiguenave, F.; Deleuze, J.F.; et al. In Vitro and In Vivo Modulation of Alternative Splicing by the Biguanide Metformin. *Mol. Ther. - Nucleic Acids* **2015**, *4*, e262, doi:10.1038/MTNA.2015.35.
150. Takarada, T.; Nishida, A.; Takeuchi, A.; Lee, T.; Takeshima, Y.; Matsuo, M. Resveratrol enhances splicing of insulin receptor exon 11 in myotonic dystrophy type 1 fibroblasts. *Brain Dev.* **2015**, *37*, 661–668, doi:10.1016/J.BRAINDEV.2014.11.001.

151. Witherspoon, L.; O'Reilly, S.; Hadwen, J.; Tasnim, N.; Mackenzie, A.; Farooq, F. Sodium Channel Inhibitors Reduce DMPK mRNA and Protein. *Clin. Transl. Sci.* **2015**, *8*, 298–304, doi:10.1111/CTS.12275.
152. Costantini, A.; Trevi, E.; Pala, M.I.; Fancellu, R. Can long-term thiamine treatment improve the clinical outcomes of myotonic dystrophy type 1? *Neural Regen. Res.* **2016**, *11*, 1487–1491, doi:10.4103/1673-5374.191225.
153. Kubowicz, P.; Zelazczyk, D.; Pekala, E. RNAi in clinical studies. *Curr. Med. Chem.* **2013**, *20*, 1801–1816, doi:10.2174/09298673113209990118.
154. Chery, J. RNA therapeutics: RNAi and antisense mechanisms and clinical applications. *Postdoc J. a J. Postdr. Res. Postdr. Aff.* **2016**, *4*, doi:10.14304/SURYA.JPR.V4N7.5.
155. Langlois, M.A.; Boniface, C.; Wang, G.; Alluin, J.; Salvaterra, P.M.; Puymirat, J.; Rossi, J.J.; Lee, N.S. Cytoplasmic and nuclear retained DMPK mRNAs are targets for RNA interference in myotonic dystrophy cells. *J. Biol. Chem.* **2005**, *280*, 16949–16954, doi:10.1074/JBC.M501591200.
156. Sobczak, K.; Wheeler, T.M.; Wang, W.; Thornton, C.A. RNA interference targeting CUG repeats in a mouse model of myotonic dystrophy. *Mol. Ther.* **2013**, *21*, 380–387, doi:10.1038/MT.2012.222.
157. Chan, J.H.P.; Lim, S.; Wong, W.S.F. Antisense oligonucleotides: from design to therapeutic application. *Clin. Exp. Pharmacol. Physiol.* **2006**, *33*, 533–540, doi:10.1111/J.1440-1681.2006.04403.X.
158. González-Barriga, A.; Kranzen, J.; Croes, H.J.E.; Bijl, S.; Van Den Broek, W.J.A.A.; Van Kessel, I.D.G.; Van Engelen, B.G.M.; Van Deutekom, J.C.T.; Wieringa, B.; Mulders, S.A.M.; et al. Cell membrane integrity in myotonic dystrophy type 1: implications for therapy. *PLoS One* **2015**, *10*, doi:10.1371/JOURNAL.PONE.0121556.
159. Bisset, D.R.; Stepniak-Konieczna, E.A.; Zavaljevski, M.; Wei, J.; Carter, G.T.; Weiss, M.D.; Chamberlain, J.R. Therapeutic impact of systemic AAV-mediated RNA interference in a mouse model of myotonic dystrophy. *Hum. Mol. Genet.* **2015**, *24*, 4971–4983, doi:10.1093/HMG/DDV219.
160. Prakash, T.; Bhat, B. 2-Modified Oligonucleotides for Antisense Therapeutics. *Curr. Top. Med. Chem.* **2007**, *7*, 641–649, doi:10.2174/156802607780487713.
161. Khvorova, A.; Watts, J.K. The chemical evolution of oligonucleotide therapies of clinical utility. *Nat. Biotechnol.* **2017**, *35*, 238–248, doi:10.1038/NBT.3765.
162. Wheeler, T.M.; Sobczak, K.; Lueck, J.D.; Osborne, R.J.; Lin, X.; Dirksen, R.T.; Thornton, C.A. Reversal of RNA dominance by displacement of protein sequestered on triplet repeat RNA. *Science (80-. )*. **2009**, *325*, 336–339, doi:10.1126/SCIENCE.1173110/SUPPL\_FILE/WHEELER-SOM.PDF.
163. Mulders, S.A.M.; Van Den Broek, W.J.A.A.; Wheeler, T.M.; Croes, H.J.E.; Van Kuik-Romeijn, P.; De Kimpe, S.J.; Furling, D.; Platenburg, G.J.; Gourdon, G.; Thornton, C.A.; et al. Triplet-repeat oligonucleotide-mediated reversal of RNA toxicity in myotonic dystrophy. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 13915–13920, doi:10.1073/PNAS.0905780106/SUPPL\_FILE/0905780106SI.PDF.
164. González-Barriga, A.; Mulders, S.A.M.; Van De Giessen, J.; Hooijer, J.D.; Bijl, S.; Van Kessel, I.D.G.; Van Beers, J.; Van Deutekom, J.C.T.; Franssen, J.A.M.; Wieringa, B.; et al. Design and analysis of effects of triplet repeat oligonucleotides

- in cell models for myotonic dystrophy. *Mol. Ther. - Nucleic Acids* **2013**, *2*, e81, doi:10.1038/MTNA.2013.9/ATTACHMENT/7F2352B7-7FD7-456D-90DA-AB6B85568CC9/MMC11.PDF.
165. Nakamori, M.; Gourdon, G.; Thornton, C.A. Stabilization of Expanded (CTG)<sup>n</sup>(CAG)<sup>m</sup> Repeats by Antisense Oligonucleotides. *Mol. Ther.* **2011**, *19*, 2222–2227, doi:10.1038/MT.2011.191.
166. Wojtkowiak-Szlachcic, A.; Taylor, K.; Stepniak-Konieczna, E.; Sznajder, L.J.; Mykowska, A.; Sroka, J.; Thornton, C.A.; Sobczak, K. Short antisense-locked nucleic acids (all-LNAs) correct alternative splicing abnormalities in myotonic dystrophy. *Nucleic Acids Res.* **2015**, *43*, 3318–3331, doi:10.1093/nar/gkv163.
167. Pandey, S.K.; Wheeler, T.M.; Justice, S.L.; Kim, A.; Younis, H.S.; Gattis, D.; Jauvin, D.; Puymirat, J.; Swayze, E.E.; Freier, S.M.; et al. Identification and Characterization of Modified Antisense Oligonucleotides Targeting DMPK in Mice and Nonhuman Primates for the Treatment of Myotonic Dystrophy Type 1. *J. Pharmacol. Exp. Ther.* **2015**, *355*, 329–340, doi:10.1124/JPET.115.226969.
168. Jauvin, D.; Chrétien, J.; Pandey, S.K.; Martineau, L.; Revillod, L.; Bassez, G.; Lachon, A.; McLeod, A.R.; Gourdon, G.; Wheeler, T.M.; et al. Targeting DMPK with Antisense Oligonucleotide Improves Muscle Strength in Myotonic Dystrophy Type 1 Mice. *Mol. Ther. - Nucleic Acids* **2017**, *7*, 465–474, doi:10.1016/J.OMTN.2017.05.007.
169. Manning, K.S.; Rao, A.N.; Castro, M.; Cooper, T.A. BNANC Gapmers Revert Splicing and Reduce RNA Foci with Low Toxicity in Myotonic Dystrophy Cells. *ACS Chem. Biol.* **2017**, *12*, 2503–2509, doi:10.1021/ACSCHEMBIO.7B00416/SUPPL\_FILE/CB7B00416\_SI\_001.PDF.
170. Lo Scudato, M.; Poulard, K.; Sourd, C.; Tomé, S.; Klein, A.F.; Corre, G.; Huguet, A.; Furling, D.; Gourdon, G.; Buj-Bello, A. Genome Editing of Expanded CTG Repeats within the Human DMPK Gene Reduces Nuclear RNA Foci in the Muscle of DM1 Mice. *Mol. Ther.* **2019**, *27*, 1372–1388, doi:10.1016/J.YMTHE.2019.05.021.
171. Batra, R.; Nelles, D.A.; Roth, D.M.; Krach, F.; Nutter, C.A.; Tadokoro, T.; Thomas, J.D.; Sznajder, Ł.J.; Blue, S.M.; Gutierrez, H.L.; et al. The sustained expression of Cas9 targeting toxic RNAs reverses disease phenotypes in mouse models of myotonic dystrophy type 1. *Nat. Biomed. Eng.* **2020**, *5*, 157–168, doi:10.1038/s41551-020-00607-7.
172. Zhang, B.W.; Cai, H.F.; Wei, X.F.; Sun, J.J.; Lan, X.Y.; Lei, C.Z.; Lin, F.P.; Qi, X.L.; Plath, M.; Chen, H. mir-30-5p regulates muscle differentiation and alternative splicing of muscle-related genes by targeting MBNL. *Int. J. Mol. Sci.* **2016**, *17*, doi:10.3390/ijms17020182.
173. Cerro-Herreros, E.; Sabater-Arcis, M.; Fernandez-Costa, J.M.; Moreno, N.; Perez-Alonso, M.; Llamusi, B.; Artero, R. MiR-23b and miR-218 silencing increase Muscblind-like expression and alleviate myotonic dystrophy phenotypes in mammalian models. *Nat. Commun.* **2018**, *9*, doi:10.1038/s41467-018-04892-4.
174. Cerro-Herreros, E.; González-Martínez, I.; Moreno-Cervera, N.; Overby, S.; Pérez-Alonso, M.; Llamusi, B.; Artero, R. Therapeutic Potential of AntagomiR-23b for Treating Myotonic Dystrophy. *Mol. Ther. - Nucleic Acids* **2020**, *21*, 837–849, doi:10.1016/j.omtn.2020.07.021.
175. Koutalios, D.; Koutsoulidou, A.; Mastrogiannopoulos, N.P.; Furling, D.; Phylactou, L.A. MyoD transcription factor induces myogenesis by inhibiting Twist-1 through miR-206. *J. Cell Sci.* **2015**, *128*, 3631–3645, doi:10.1242/jcs.172288.

176. Castel, A.L.; Overby, S.J.; Artero, R. MicroRNA-based therapeutic perspectives in myotonic dystrophy. *Int. J. Mol. Sci.* **2019**, *20*.
177. Lagrue, E.; Dogan, C.; De Antonio, M.; Audic, F.; Bach, N.; Barnerias, C.; Bellance, R.; Cances, C.; Chabrol, B.; Cuisset, J.M.; et al. A large multicenter study of pediatric myotonic dystrophy type 1 for evidence-based management. *Neurology* **2019**, *92*, E852–E865, doi:10.1212/WNL.0000000000006948.
178. Hamshere, M.G.; Harley, H.; Harper, P.; Brook, J.D.; Brookfield, J.F. Myotonic dystrophy: the correlation of (CTG) repeat length in leucocytes with age at onset is significant only for patients with small expansions. *J. Med. Genet.* **1999**, *36*, 59–61.
179. Savić, D.; Rakocvic-Stojanovic, V.; Keckarevic, D.; Culjkovic, B.; Stojkovic, O.; Mladenovic, J.; Todorovic, S.; Apostolski, S.; Romac, S. 250 CTG repeats in DMPK is a threshold for correlation of expansion size and age at onset of juvenile-adult DM1. *Hum. Mutat.* **2002**, *19*, 131–9, doi:10.1002/humu.10027.
180. Mahadevan, M.; Tsilfidis, C.; Sabourin, L.; Shutler, G.; Amemiya, C.; Jansen, G.; Neville, C.; Narang, M.; Barceló, J.; O'Hoy, K.; et al. Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. *Science* **1992**, *255*, 1253–5, doi:10.1126/science.1546325.
181. Wong, L.J.C.; Ashizawa, T.; Monckton, D.G.; Caskey, C.T.; Richards, C.S. Somatic heterogeneity of the CTG repeat in myotonic dystrophy is age and size dependent. *Am. J. Hum. Genet.* **1995**, *56*, 114–122.
182. Chong-Nguyen, C.; Wahbi, K.; Algalarrondo, V.; Bécane, H.M.; Radvanyi-Hoffman, H.; Arnaud, P.; Furling, D.; Lazarus, A.; Bassez, G.; Béhin, A.; et al. Association between Mutation Size and Cardiac Involvement in Myotonic Dystrophy Type 1: An Analysis of the DM1-Heart Registry. *Circ. Cardiovasc. Genet.* **2017**, *10*, doi:10.1161/CIRCGENETICS.116.001526.
183. Merlevede, K.; Vermander, D.; Theys, P.; Legius, E.; Ector, H.; Robberecht, W. Cardiac involvement and CTG expansion in myotonic dystrophy. *J. Neurol.* **2002**, *249*, 693–698, doi:10.1007/S00415-002-0692-6.
184. Hogrel, J.Y.; Ollivier, G.; Ledoux, I.; Hébert, L.J.; Eymard, B.; Puymirat, J.; Bassez, G. Relationships between grip strength, myotonia, and CTG expansion in myotonic dystrophy type 1. *Ann. Clin. Transl. Neurol.* **2017**, *4*, 921–925, doi:10.1002/ACN3.496.
185. Kroksmark, A.K.; Ekström, A.B.; Björck, E.; Tulinius, M. Myotonic dystrophy: muscle involvement in relation to disease type and size of expanded CTG-repeat sequence. *Dev. Med. Child Neurol.* **2005**, *47*, 478–485, doi:10.1017/S0012162205000927.
186. Winblad, S.; Lindberg, C.; Hansen, S. Cognitive deficits and CTG repeat expansion size in classical myotonic dystrophy type 1 (DM1). *Behav. Brain Funct.* **2006**, *2*, doi:10.1186/1744-9081-2-16.
187. Groh, W.J.; Groh, M.R.; Shen, C.; Monckton, D.G.; Bodkin, C.L.; Pascuzzi, R.M. Survival and CTG repeat expansion in adults with myotonic dystrophy type 1. *Muscle Nerve* **2011**, *43*, 648–651, doi:10.1002/mus.21934.
188. Ballester-Lopez, A.; Linares-Pardo, I.; Koehorst, E.; Núñez-Manchón, J.; Pintos-Morell, G.; Coll-Cantí, J.; Almendrote, M.; Lucente, G.; Arbex, A.; Magaña, J.J.; et al. The need for establishing a universal CTG sizing method in myotonic dystrophy type 1. *Genes (Basel)*. **2020**, *11*, 1–9, doi:10.3390/genes11070757.



## Bibliography

189. Musova, Z.; Mazanec, R.; Krepelova, A.; Ehler, E.; Vales, J.; Jaklova, R.; Prochazka, T.; Koukal, P.; Marikova, T.; Kraus, J.; et al. Highly unstable sequence interruptions of the CTG repeat in the myotonic dystrophy gene. *Am. J. Med. Genet. Part A* **2009**, *149*, 1365–1369, doi:10.1002/ajmg.a.32987.
190. Botta, A.; Rossi, G.; Marcaurelio, M.; Fontana, L.; D'Apice, M.R.; Brancati, F.; Massa, R.; G Monckton, D.; Sanguuolo, F.; Novelli, G. Identification and characterization of 5' CCG interruptions in complex DMPK expanded alleles. *Eur. J. Hum. Genet.* **2017**, *25*, 1–5, doi:10.1038/ejhg.2016.148.
191. Santoro, M.; Masciullo, M.; Pietrobono, R.; Conte, G.; Modoni, A.; Bianchi, M.L.E.; Rizzo, V.; Pomponi, M.G.; Tasca, G.; Neri, G.; et al. Molecular, clinical, and muscle studies in myotonic dystrophy type 1 (DM1) associated with novel variant CCG expansions. *J. Neurol.* **2013**, *260*, 1245–1257, doi:10.1007/s00415-012-6779-9.
192. Tomé, S.; Dandelot, E.; Dogan, C.; Bertrand, A.; Geneviève, D.; Péréon, Y.; DM contraction study group, M.; Simon, M.; Bonnefont, J.-P.; Bassez, G.; et al. Unusual association of a unique CAG interruption in 5' of DM1 CTG repeats with intergenerational contractions and low somatic mosaicism. *Hum. Mutat.* **2018**, *39*, 970–982, doi:10.1002/humu.23531.
193. Miller, J.N.; Van Der Plas, E.; Hamilton, M.; Kosciak, T.R.; Gutmann, L.; Cumming, S.A.; Monckton, D.G.; Nopoulos, P.C. Variant repeats within the DMPK CTG expansion protect function in myotonic dystrophy type 1. *Neurol. Genet.* **2020**, *6*, doi:10.1212/NXG.0000000000000504.
194. Wenninger, S.; Cumming, S.A.; Gutschmidt, K.; Okkersen, K.; Jimenez-Moreno, A.C.; Daidj, F.; Lochmüller, H.; Hogarth, F.; Knoop, H.; Bassez, G.; et al. Associations Between Variant Repeat Interruptions and Clinical Outcomes in Myotonic Dystrophy Type 1. *Neurol. Genet.* **2021**, *7*, e572, doi:10.1212/NXG.0000000000000572.
195. Pešović, J.; Perić, S.; Brkušanin, M.; Brajušković, G.; Rakoč Ević -Stojanović, V.; Savić-Pavić Ević, D. Repeat Interruptions Modify Age at Onset in Myotonic Dystrophy Type 1 by Stabilizing DMPK Expansions in Somatic Cells. *Front. Genet.* **2018**, *9*, 1–14, doi:10.3389/FGENE.2018.00601.
196. Mangin, A.; de Pontual, L.; Tsai, Y.C.; Monteil, L.; Nizon, M.; Boisseau, P.; Mercier, S.; Ziegler, J.; Harting, J.; Heiner, C.; et al. Robust detection of somatic mosaicism and repeat interruptions by long-read targeted sequencing in myotonic dystrophy type 1. *Int. J. Mol. Sci.* **2021**, *22*, 1–24, doi:10.3390/ijms22052616.
197. Cumming, S.A.; Oliwa, A.; Stevens, G.; Ballantyne, B.; Mann, C.; Razvi, S.; Longman, C.; Monckton, D.G.; Farrugia, M.E. A DM1 patient with CCG variant repeats: Reaching the diagnosis. *Neuromuscul. Disord.* **2021**, *31*, 232–238, doi:10.1016/J.NMD.2020.12.005.
198. Eriksson, M.; Hedberg, B.; Carey, N.; Ansved, T. Decreased DMPK transcript levels in myotonic dystrophy 1 type IIA muscle fibers. *Biochem. Biophys. Res. Commun.* **2001**, *286*, 1177–82, doi:10.1006/bbrc.2001.5516.
199. Carango, P.; Noble, J.E.; Marks, H.G.; Funanage, V.L. Absence of myotonic dystrophy protein kinase (DMPK) mRNA as a result of a triplet repeat expansion in myotonic dystrophy. *Genomics* **1993**, *18*, 340–348, doi:10.1006/geno.1993.1474.
200. Bhagwati, S.; Ghatpande, A.; Leung, B. Normal levels of DM RNA and myotonin protein kinase in skeletal muscle from adult myotonic dystrophy (DM) patients. *Biochim. Biophys. Acta* **1996**, *1317*, 155–157, doi:10.1016/S0925-4439(96)00057-9.

201. Frisch, R.; Singleton, K.R.; Moses, P.A.; Gonzalez, I.L.; Carango, P.; Marks, H.G.; Funanage, V.L. Effect of triplet repeat expansion on chromatin structure and expression of DMPK and neighboring genes, SIX5 and DMWD, in myotonic dystrophy. *Mol. Genet. Metab.* **2001**, *74*, 281–291, doi:10.1006/MGME.2001.3229.
202. Korade-Mirnic, Z.; Tarleton, J.; Servidei, S.; Casey, R.R.; Gennarelli, M.; Pegoraro, E.; Angelini, C.; Hoffman, E.P. Myotonic dystrophy: Tissue-specific effect of somatic CTG expansions on allele-specific DMAHP/SIX5 expression. *Hum. Mol. Genet.* **1999**, *8*, 1017–1023, doi:10.1093/hmg/8.6.1017.
203. Krol, J.; Fiszler, A.; Mykowska, A.; Sobczak, K.; de Mezer, M.; Krzyzosiak, W.J. Ribonuclease dicer cleaves triplet repeat hairpins into shorter repeats that silence specific targets. *Mol. Cell* **2007**, *25*, 575–586, doi:10.1016/J.MOLCEL.2007.01.031.
204. Castro, A.F.; Loureiro, J.R.; Bessa, J.; Silveira, I. Antisense Transcription across Nucleotide Repeat Expansions in Neurodegenerative and Neuromuscular Diseases: Progress and Mysteries. *Genes (Basel)*. **2020**, *11*, 1–30, doi:10.3390/GENES11121418.
205. He, Y.; Vogelstein, B.; Velculescu, V.E.; Papadopoulos, N.; Kinzler, K.W. The antisense transcriptomes of human cells. *Science* **2008**, *322*, 1855–1857, doi:10.1126/SCIENCE.1163853.
206. Faghihi, M.A.; Wahlestedt, C. Regulatory roles of natural antisense transcripts. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 637–643, doi:10.1038/NRM2738.
207. Werner, A.; Sayer, J.A. Naturally occurring antisense RNA: function and mechanisms of action. *Curr. Opin. Nephrol. Hypertens.* **2009**, *18*, 343–349, doi:10.1097/MNH.0B013E32832CB982.
208. Barman, P.; Reddy, D.; Bhaumik, S.R. Mechanisms of Antisense Transcription Initiation with Implications in Gene Expression, Genomic Integrity and Disease Pathogenesis. *Non-coding RNA* **2019**, *5*, doi:10.3390/NCRNA5010011.
209. Pelechano, V.; Steinmetz, L.M. Gene regulation by antisense transcription. *Nat. Rev. Genet.* **2013**, *14*, 880–893, doi:10.1038/NRG3594.
210. Katayama, S.; Tomaru, Y.; Kasukawa, T.; Waki, K.; Nakanishi, M.; Nakamura, M.; Nishida, H.; Yap, C.C.; Suzuki, M.; Kawai, J.; et al. Antisense transcription in the mammalian transcriptome. *Science* **2005**, *309*, 1564–1566, doi:10.1126/SCIENCE.1112009.
211. Derrien, T.; Johnson, R.; Bussotti, G.; Tanzer, A.; Djebali, S.; Tilgner, H.; Guernec, G.; Martin, D.; Merkel, A.; Knowles, D.G.; et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* **2012**, *22*, 1775–1789, doi:10.1101/GR.132159.111.
212. Ozsolak, F.; Kapranov, P.; Foissac, S.; Kim, S.W.; Fishilevich, E.; Monaghan, A.P.; John, B.; Milos, P.M. Comprehensive polyadenylation site maps in yeast and human reveal pervasive alternative polyadenylation. *Cell* **2010**, *143*, 1018–1029, doi:10.1016/J.CELL.2010.11.020.
213. Chung, D.W.; Rudnicki, D.D.; Yu, L.; Margolis, R.L. A natural antisense transcript at the Huntington's disease repeat locus regulates HTT expression. *Hum. Mol. Genet.* **2011**, *20*, 3467–3477, doi:10.1093/HMG/DDR263.
214. Li, P.P.; Sun, X.; Xia, G.; Arbez, N.; Paul, S.; Zhu, S.; Peng, H.B.; Ross, C.A.; Koepfen, A.H.; Margolis, R.L.; et al. ATXN2-AS, a gene antisense to ATXN2, is associated with spinocerebellar ataxia type 2 and amyotrophic lateral sclerosis.

- Ann. Neurol.* **2016**, *80*, 600–615, doi:10.1002/ANA.24761.
215. Cho, D.H.; Thienes, C.P.; Mahoney, S.E.; Analau, E.; Filippova, G.N.; Tapscott, S.J. Antisense transcription and heterochromatin at the DM1 CTG repeats are constrained by CTCF. *Mol. Cell* **2005**, *20*, 483–489, doi:10.1016/j.molcel.2005.09.002.
216. Huguet, A.; Medja, F.; Nicole, A.; Vignaud, A.; Guiraud-Dogan, C.; Ferry, A.; Decostre, V.; Hogrel, J.Y.; Metzger, F.; Hoeflich, A.; et al. Molecular, Physiological, and Motor Performance Defects in DMSXL Mice Carrying >1,000 CTG Repeats from the Human DM1 Locus. *PLoS Genet.* **2012**, *8*, doi:10.1371/journal.pgen.1003043.
217. Michel, L.; Huguet-Lachon, A.; Gourdon, G. Sense and antisense DMPK RNA foci accumulate in DM1 tissues during development. *PLoS One* **2015**, *10*, doi:10.1371/journal.pone.0137620.
218. Gudde, A.E.E.G.; van Heeringen, S.J.; de Oude, A.I.; van Kessel, I.D.G.; Estabrook, J.; Wang, E.T.; Wieringa, B.; Wansink, D.G. Antisense transcription of the myotonic dystrophy locus yields low-abundant RNAs with and without (CAG)<sub>n</sub> repeat. *RNA Biol.* **2017**, *14*, 1374–1388, doi:10.1080/15476286.2017.1279787.
219. Hsu, R.J.; Hsiao, K.M.; Lin, M.J.; Li, C.Y.; Wang, L.C.; Chen, L.K.; Pan, H. Long tract of untranslated CAG repeats is deleterious in transgenic mice. *PLoS One* **2011**, *6*, doi:10.1371/JOURNAL.PONE.0016417.
220. Yu, Z.; Teng, X.; Bonini, N.M. Triplet repeat-derived siRNAs enhance RNA-mediated toxicity in a Drosophila model for myotonic dystrophy. *PLoS Genet.* **2011**, *7*, doi:10.1371/JOURNAL.PGEN.1001340.
221. Zu, T.; Gibbens, B.; Doty, N.S.; Gomes-Pereira, M.; Huguet, A.; Stone, M.D.; Margolis, J.; Peterson, M.; Markowski, T.W.; Ingram, M.A.C.; et al. Non-ATG-initiated translation directed by microsatellite expansions. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 260–5, doi:10.1073/pnas.1013343108.
222. Banez-Coronel, M.; Ayhan, F.; Tarabochia, A.D.; Zu, T.; Perez, B.A.; Tusi, S.K.; Pletnikova, O.; Borchelt, D.R.; Ross, C.A.; Margolis, R.L.; et al. RAN Translation in Huntington Disease. *Neuron* **2015**, *88*, 667–677, doi:10.1016/j.neuron.2015.10.038.
223. Maeda, M.; Taft, C.S.; Bush, E.W.; Holder, E.; Bailey, W.M.; Neville, H.; Perryman, M.B.; Bies, R.D. Identification, tissue-specific expression, and subcellular localization of the 80- and 71-kDa forms of myotonic dystrophy kinase protein. *J. Biol. Chem.* **1995**, *270*, 20246–20249, doi:10.1074/JBC.270.35.20246.
224. Lam, L.T.; Pham, Y.C.N.; thi Man, N.; Morris, G.E. Characterization of a monoclonal antibody panel shows that the myotonic dystrophy protein kinase, DMPK, is expressed almost exclusively in muscle and heart. *Hum. Mol. Genet.* **2000**, *9*, 2167–2173, doi:10.1093/HMG/9.14.2167.
225. Kaliman, P.; Llagostera, E. Myotonic dystrophy protein kinase (DMPK) and its role in the pathogenesis of myotonic dystrophy 1. *Cell. Signal.* **2008**, *20*, 1935–1941, doi:10.1016/J.CELLSIG.2008.05.005.
226. Salvatori, S.; Fanin, M.; Trevisan, C.P.; Furlan, S.; Reddy, S.; Nagy, J.I.; Angelini, C. Decreased expression of DMPK: correlation with CTG repeat expansion and fibre type composition in myotonic dystrophy type 1. *Neurol. Sci.* **2005**, *26*, 235–242, doi:10.1007/S10072-005-0466-X.
227. Paul Mounsey, J.; Mistry, D.J.; Ai, C.W.; Reddy, S.; Randall Moorman, J. Skeletal

- muscle sodium channel gating in mice deficient in myotonic dystrophy protein kinase. *Hum. Mol. Genet.* **2000**, *9*, 2313–2320, doi:10.1093/OXFORDJOURNALS.HMG.A018923.
228. Benders, A.A.G.M.; Groenen, P.J.T.A.; Oerlemans, F.T.J.J.; Veerkamp, J.H.; Wieringa, B. Myotonic dystrophy protein kinase is involved in the modulation of the Ca<sup>2+</sup> homeostasis in skeletal muscle cells. *J. Clin. Invest.* **1997**, *100*, 1440–1447, doi:10.1172/JCI119664.
229. Berul, C.I.; Maguire, C.T.; Gehrmann, J.; Reddy, S. Progressive atrioventricular conduction block in a mouse myotonic dystrophy model. *J. Interv. Card. Electrophysiol.* **2000**, *4*, 351–358, doi:10.1023/A:1009842114968.
230. Llagostera, E.; Catalucci, D.; Marti, L.; Liesa, M.; Camps, M.; Ciaraldi, T.P.; Kondo, R.; Reddy, S.; Dillmann, W.H.; Palacin, M.; et al. Role of myotonic dystrophy protein kinase (DMPK) in glucose homeostasis and muscle insulin action. *PLoS One* **2007**, *2*, doi:10.1371/JOURNAL.PONE.0001134.
231. Jackson, R.J.; Hellen, C.U.T.; Pestova, T. V. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 113–127, doi:10.1038/nrm2838.
232. Kozak, M. Features in the 5' non-coding sequences of rabbit alpha and beta-globin mRNAs that affect translational efficiency. *J. Mol. Biol.* **1994**, *235*, 95–110, doi:10.1016/S0022-2836(05)80019-1.
233. Kozak, M. Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87*, 8301–8305, doi:10.1073/pnas.87.21.8301.
234. Kozak, M. Influences of mRNA secondary structure on initiation by eukaryotic ribosomes. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83*, 2850–2854, doi:10.1073/PNAS.83.9.2850.
235. Kozak, M. Influence of mRNA secondary structure on binding and migration of 40S ribosomal subunits. *Cell* **1980**, *19*, 79–90, doi:10.1016/0092-8674(80)90390-6.
236. Lozano, G.; Martínez-Salas, E. Structural insights into viral IRES-dependent translation mechanisms. *Curr. Opin. Virol.* **2015**, *12*, 113–120, doi:10.1016/J.COVIRO.2015.04.008.
237. Zu, T.; Pattamatta, A.; Ranum, L.P.W. Repeat-Associated Non-ATG Translation in Neurological Diseases. *Cold Spring Harb. Perspect. Biol.* **2018**, *10*, a033019, doi:10.1101/cshperspect.a033019.
238. Ash, P.E.A.; Bieniek, K.F.; Gendron, T.F.; Caulfield, T.; Lin, W.-L.; DeJesus-Hernandez, M.; van Blitterswijk, M.M.; Jansen-West, K.; Paul, J.W.; Rademakers, R.; et al. Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron* **2013**, *77*, 639–46, doi:10.1016/j.neuron.2013.02.004.
239. Zu, T.; Liu, Y.; Bañez-Coronel, M.; Reid, T.; Pletnikova, O.; Lewis, J.; Miller, T.M.; Harms, M.B.; Falchook, A.E.; Subramony, S.H.; et al. RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, doi:10.1073/pnas.1315438110.
240. Mori, K.; Weng, S.M.; Arzberger, T.; May, S.; Rentzsch, K.; Kremmer, E.; Schmid, B.; Kretzschmar, H.A.; Cruts, M.; Van Broeckhoven, C.; et al. The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in

- FTLD/ALS. *Science* **2013**, *339*, 1335–1338, doi:10.1126/SCIENCE.1232927.
241. Todd, P.K.; Oh, S.Y.; Krans, A.; He, F.; Sellier, C.; Frazer, M.; Renoux, A.J.; Chen, K.; Scaglione, K.M.; Basrur, V.; et al. CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome. *Neuron* **2013**, *78*, 440–55, doi:10.1016/j.neuron.2013.03.026.
242. Buijsen, R.A.M.; Visser, J.A.; Kramer, P.; Severijnen, E.A.W.F.M.; Gearing, M.; Charlet-Berguerand, N.; Sherman, S.L.; Berman, R.F.; Willemsen, R.; Hukema, R.K. Presence of inclusions positive for polyglycine containing protein, FMRpolyG, indicates that repeat-associated non-AUG translation plays a role in fragile X-associated primary ovarian insufficiency. *Hum. Reprod.* **2016**, *31*, 158, doi:10.1093/HUMREP/DEV280.
243. Davies, J.E.; Rubinsztein, D.. Polyalanine and polyserine frameshift products in Huntington's disease. *J. Med. Genet.* **2006**, *43*, 893–896, doi:10.1136/JMG.2006.044222.
244. Ishiguro, T.; Sato, N.; Ueyama, M.; Fujikake, N.; Sellier, C.; Kanegami, A.; Tokuda, E.; Zamiri, B.; Gall-Duncan, T.; Mirceta, M.; et al. Regulatory Role of RNA Chaperone TDP-43 for RNA Misfolding and Repeat-Associated Translation in SCA31. *Neuron* **2017**, *94*, 108-124.e7, doi:10.1016/j.neuron.2017.02.046.
245. Soragni, E.; Petrosyan, L.; Rinkoski, T.; Wieben, E.; Baratz, K.; Fautsch, M.; Gottesfeld, J. Repeat-Associated Non-ATG (RAN) Translation in Fuchs' Endothelial Corneal Dystrophy. *Invest. Ophthalmol. Vis. Sci.* **2018**, *59*, 1888–1896, doi:10.1167/IOVS.17-23265.
246. Zu, T.; Cleary, J.D.; Liu, Y.; Bañez-Coronel, M.; Bubenik, J.L.; Ayhan, F.; Ashizawa, T.; Xia, G.; Clark, H.B.; Yachnis, A.T.; et al. RAN Translation Regulated by Muscleblind Proteins in Myotonic Dystrophy Type 2. *Neuron* **2017**, *95*, 1292-1305.e5, doi:10.1016/j.neuron.2017.08.039.
247. Cox, D.C.; Cooper, T.A. Non-canonical RAN Translation of CGG Repeats Has Canonical Requirements. *Mol. Cell* **2016**, *62*, 155–156, doi:10.1016/j.molcel.2016.04.004.
248. Kearse, M.G.; Green, K.M.; Krans, A.; Rodriguez, C.M.; Linsalata, A.E.; Goldstrohm, A.C.; Todd, P.K. CGG Repeat-Associated Non-AUG Translation Utilizes a Cap-Dependent Scanning Mechanism of Initiation to Produce Toxic Proteins. *Mol. Cell* **2016**, *62*, 314–322, doi:10.1016/J.MOLCEL.2016.02.034.
249. Cleary, J.D.; Ranum, L.P.W. New developments in RAN translation: insights from multiple diseases. *Curr Opin Genet Dev* **2018**, 125–134, doi:10.1016/j.gde.2017.03.006.New.
250. Sellier, C.; Buijsen, R.A.M.; He, F.; Natla, S.; Jung, L.; Tropel, P.; Gaucherot, A.; Jacobs, H.; Meziane, H.; Vincent, A.; et al. Translation of Expanded CGG Repeats into FMRpolyG Is Pathogenic and May Contribute to Fragile X Tremor Ataxia Syndrome. *Neuron* **2017**, *93*, 331–347, doi:10.1016/J.NEURON.2016.12.016.
251. May, S.; Hornburg, D.; Schludi, M.H.; Arzberger, T.; Rentzsch, K.; Schwenk, B.M.; Grässer, F.A.; Mori, K.; Kremmer, E.; Banzhaf-Strathmann, J.; et al. C9orf72 FTLD/ALS-associated Gly-Ala dipeptide repeat proteins cause neuronal toxicity and Unc119 sequestration. *Acta Neuropathol.* **2014**, *128*, 485–503, doi:10.1007/S00401-014-1329-4.
252. Zhang, Y.J.; Gendron, T.F.; Grima, J.C.; Sasaguri, H.; Jansen-West, K.; Xu, Y.F.; Katzman, R.B.; Gass, J.; Murray, M.E.; Shinohara, M.; et al. C9ORF72 poly(GA)

- aggregates sequester and impair HR23 and nucleocytoplasmic transport proteins. *Nat. Neurosci.* **2016**, *19*, 668–677, doi:10.1038/NN.4272.
253. Berger, S.L.; Kouzarides, T.; Shiekhattar, R.; Shilatifard, A. An operational definition of epigenetics. *Genes Dev.* **2009**, *23*, 781, doi:10.1101/GAD.1787609.
254. Surace, A.E.A.; Hedrich, C.M. The Role of Epigenetics in Autoimmune/Inflammatory Disease. *Front. Immunol.* **2019**, *10*, doi:10.3389/FIMMU.2019.01525.
255. Dawson, M.A.; Kouzarides, T. Cancer epigenetics: from mechanism to therapy. *Cell* **2012**, *150*, 12–27, doi:10.1016/J.CELL.2012.06.013.
256. Coppedè, F. Epigenetics of neuromuscular disorders. *Epigenomics* **2020**, *12*, 2125–2139, doi:10.2217/EPI-2020-0282.
257. Evans-Galea, M. V.; Hannan, A.J.; Carroddus, N.; Delatycki, M.B.; Saffery, R. Epigenetic modifications in trinucleotide repeat diseases. *Trends Mol. Med.* **2013**, *19*, 655–663, doi:10.1016/J.MOLMED.2013.07.007.
258. He, F.; Todd, P.K. Epigenetics in nucleotide repeat expansion disorders. *Semin. Neurol.* **2011**, *31*, 470–483, doi:10.1055/S-0031-1299786.
259. Visconti, V.V.; Centofanti, F.; Fittipaldi, S.; Macri, E.; Novelli, G.; Botta, A. Epigenetics of Myotonic Dystrophies: A Minireview. *Int. J. Mol. Sci.* **2021**, *22*, 12594, doi:10.3390/IJMS222212594.
260. Nageshwaran, S.; Festenstein, R. Epigenetics and Triplet-Repeat Neurological Diseases. *Front. Neurol.* **2015**, *6*, 262, doi:10.3389/FNEUR.2015.00262.
261. Razin, A.; Cedar, H. DNA methylation and gene expression. *Microbiol. Rev.* **1991**, *55*, 173–175, doi:10.1128/MR.55.3.451-458.1991.
262. Moore, L.D.; Le, T.; Fan, G. DNA methylation and its basic function. *Neuropsychopharmacology* **2013**, *38*, 23–38, doi:10.1038/NPP.2012.112.
263. Ehrlich, M.; Gama-Sosa, M.A.; Huang, L.H.; Midgett, R.M.; Kuo, K.C.; Mccune, R.A.; Gehrke, C. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res.* **1982**, *10*, 2709–2721, doi:10.1093/NAR/10.8.2709.
264. Edwards, J.R.; Yarychivska, O.; Boulard, M.; Bestor, T.H. DNA methylation and DNA methyltransferases. *Epigenetics Chromatin* **2017**, *10*, doi:10.1186/S13072-017-0130-8.
265. Monk, M. Epigenetic programming of differential gene expression in development and evolution. *Dev. Genet.* **1995**, *17*, 188–197, doi:10.1002/DVG.1020170303.
266. Monk, M.; Boubelik, M.; Lehnert, S. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* **1987**, *99*, 371–382, doi:10.1242/DEV.99.3.371.
267. Brandeis, M.; Ariel, M.; Cedar, H. Dynamics of DNA methylation during development. *BioEssays* **1993**, *15*, 709–713, doi:10.1002/bies.950151103.
268. Angeloni, A.; Bogdanovic, O. Sequence determinants, function, and evolution of CpG islands. *Biochem. Soc. Trans.* **2021**, *49*, 1109–1119, doi:10.1042/BST20200695.

269. Deaton, A.M.; Bird, A. CpG islands and the regulation of transcription. *Genes Dev.* **2011**, *25*, 1010–1022, doi:10.1101/GAD.2037511.
270. Mohn, F.; Weber, M.; Rebhan, M.; Roloff, T.C.; Richter, J.; Stadler, M.B.; Bibel, M.; Schübeler, D. Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol. Cell* **2008**, *30*, 755–766, doi:10.1016/J.MOLCEL.2008.05.007.
271. Zeng, Y.; Chen, T. DNA Methylation Reprogramming during Mammalian Development. *Genes (Basel)*. **2019**, *10*, doi:10.3390/GENES10040257.
272. Illingworth, R.S.; Gruenewald-Schneider, U.; Webb, S.; Kerr, A.R.W.; James, K.D.; Turner, D.J.; Smith, C.; Harrison, D.J.; Andrews, R.; Bird, A.P. Orphan CpG islands identify numerous conserved promoters in the mammalian genome. *PLoS Genet.* **2010**, *6*, doi:10.1371/JOURNAL.PGEN.1001134.
273. Rugowska, A.; Starosta, A.; Konieczny, P. Epigenetic modifications in muscle regeneration and progression of Duchenne muscular dystrophy. *Clin. Epigenetics* **2021**, *13*, doi:10.1186/S13148-021-01001-Z.
274. Nan, X.; Cross, S.; Bird, A. Gene silencing by methyl-CpG-binding proteins. *Novartis Found. Symp.* **1998**, *214*, doi:10.1002/9780470515501.CH2.
275. Nan, X.; Ng, H.H.; Johnson, C.A.; Laherty, C.D.; Turner, B.M.; Eisenman, R.N.; Bird, A. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **1998**, *393*, 386–389, doi:10.1038/30764.
276. Bostick, M.; Jong, K.K.; Estève, P.O.; Clark, A.; Pradhan, S.; Jacobsen, S.E. UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **2007**, *317*, 1760–1764, doi:10.1126/SCIENCE.1147939.
277. Prokhortchouk, A.; Hendrich, B.; Jørgensen, H.; Ruzov, A.; Wilm, M.; Georgiev, G.; Bird, A.; Prokhortchouk, E. The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. *Genes Dev.* **2001**, *15*, 1613–1618, doi:10.1101/gad.198501.
278. Ehrlich, M. DNA hypermethylation in disease: mechanisms and clinical relevance. *Epigenetics* **2019**, *14*, 1141–1163, doi:10.1080/15592294.2019.1638701.
279. Laird, P.W. Principles and challenges of genome-wide DNA methylation analysis. *Nat. Rev. Genet.* **2010**, *11*, 191–203, doi:10.1038/NRG2732.
280. Youssoufian, H.; Mulder, C. Detection of methylated sequences in eukaryotic DNA with the restriction endonucleases SmaI and XmaI. *J. Mol. Biol.* **1981**, *150*, 133–136, doi:10.1016/0022-2836(81)90328-4.
281. Butkus, V.; Petrauskienė, L.; Maneliene, Z.; Klimašauskas, S.; Laučys, V.; Janulaitis, A. Cleavage of methylated CCCGGG sequences containing either N4-methylcytosine or 5-methylcytosine with MspI, HpaII, SmaI, XmaI and Cfr9I restriction endonucleases. *Nucleic Acids Res.* **1987**, *15*, 7091, doi:10.1093/NAR/15.17.7091.
282. Aberg, K.A.; Chan, R.F.; Xie, L.; Shabalín, A.A.; van den Oord, E.J.C.G. Methyl-CpG-Binding Domain Sequencing: MBD-seq. *Methods Mol. Biol.* **2018**, *1708*, 171–189, doi:10.1007/978-1-4939-7481-8\_10.
283. Jørgensen, H.F.; Adie, K.; Chaubert, P.; Bird, A.P. Engineering a high-affinity methyl-CpG-binding protein. *Nucleic Acids Res.* **2006**, *34*,

- doi:10.1093/NAR/GKL527.
284. Waalwijk, C.; Flavell, R.A. MspI, an isoschizomer of hpaII which cleaves both unmethylated and methylated hpaII sites. *Nucleic Acids Res.* **1978**, *5*, 3231–3236, doi:10.1093/NAR/5.9.3231.
285. Darst, R.P.; Pardo, C.E.; Ai, L.; Brown, K.D.; Kladdé, M.P. Bisulfite sequencing of DNA. *Curr. Protoc. Mol. Biol.* **2010**, *Chapter 7*, doi:10.1002/0471142727.MB0709S91.
286. Li, Y.; Tollefsbol, T.O. DNA methylation detection: bisulfite genomic sequencing analysis. *Methods Mol. Biol.* **2011**, *791*, 11–21, doi:10.1007/978-1-61779-316-5\_2.
287. Tost, J.; Dunker, J.; Gut, I.G. Analysis and quantification of multiple methylation variable positions in CpG islands by Pyrosequencing™. <https://doi.org/10.2144/03351md02> **2018**, *35*, 152–156, doi:10.2144/03351MD02.
288. Colella, S.; Shen, L.; Baggerly, K.A.; Issa, J.P.J.; Krahe, R. Sensitive and quantitative universal Pyrosequencing™ methylation analysis of CpG sites. <https://doi.org/10.2144/03351md01> **2018**, *35*, 146–150, doi:10.2144/03351MD01.
289. Reed, K.; Poulin, M.L.; Yan, L.; Parissenti, A.M. Comparison of bisulfite sequencing PCR with pyrosequencing for measuring differences in DNA methylation. *Anal. Biochem.* **2010**, *397*, 96–106, doi:10.1016/J.AB.2009.10.021.
290. Kurdyukov, S.; Bullock, M. DNA Methylation Analysis: Choosing the Right Method. *Biology (Basel)*. **2016**, *5*, doi:10.3390/BIOLOGY5010003.
291. Bastiaan Holwerda, S.J.; de Laat, W. CTCF: the protein, the binding partners, the binding sites and their chromatin loops. *Philos. Trans. R. Soc. B Biol. Sci.* **2013**, *368*, doi:10.1098/RSTB.2012.0369.
292. Filippova, G.N.; Thienes, C.P.; Penn, B.H.; Cho, D.H.; Hu, Y.J.; Moore, J.M.; Klesert, T.R.; Lobanenko, V. V.; Tapscott, S.J. CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. *Nat. Genet.* **2001**, *28*, 335–343, doi:10.1038/ng570.
293. Buckley, L.; Lacey, M.; Ehrlich, M. Epigenetics of the myotonic dystrophy-associated DMPK gene neighborhood. *Epigenomics* **2016**, *8*, 13–31, doi:10.2217/EPI.15.104/SUPPL\_FILE/EPI-08-13-S2.XLSX.
294. Breton, É.; Légaré, C.; Overend, G.; Guay, S.P.; Monckton, D.; Mathieu, J.; Gagnon, C.; Richer, L.; Gallais, B.; Bouchard, L. DNA methylation at the DMPK gene locus is associated with cognitive functions in myotonic dystrophy type 1. *Epigenomics* **2020**, *12*, 2051–2064, doi:10.2217/EPI-2020-0328.
295. Brouwer, J.R.; Huguet, A.; Nicole, A.; Munnich, A.; Gourdon, G. Transcriptionally repressive chromatin remodelling and CpG methylation in the presence of expanded CTG-repeats at the DM1 locus. *J. Nucleic Acids* **2013**, *2013*, doi:10.1155/2013/567435.
296. Hildonen, M.; Knak, K.L.; Dunø, M.; Vissing, J.; Tümer, Z. Stable Longitudinal Methylation Levels at the CpG Sites Flanking the CTG Repeat of DMPK in Patients with Myotonic Dystrophy Type 1. *Genes (Basel)*. **2020**, *11*, 1–13, doi:10.3390/GENES11080936.



297. Morales, F.; Corrales, E.; Zhang, B.; Vásquez, M.; Santamaría-Ulloa, C.; Quesada, H.; Sirito, M.; Estecio, M.R.; Monckton, D.G.; Krahe, R. Myotonic dystrophy type 1 (DM1) clinical sub-types and CTCF site methylation status flanking the CTG expansion are mutant allele length-dependent. *Hum. Mol. Genet.* **2021**, doi:10.1093/HMG/DDAB243.
298. Santoro, M.; Fontana, L.; Masciullo, M.; Bianchi, M.L.E.; Rossi, S.; Leoncini, E.; Novelli, G.; Botta, A.; Silvestri, G. Expansion size and presence of CCG/CTC/CGG sequence interruptions in the expanded CTG array are independently associated to hypermethylation at the DMPK locus in myotonic dystrophy type 1 (DM1). *Biochim. Biophys. Acta - Mol. Basis Dis.* **2015**, *1852*, 2645–2652, doi:10.1016/J.BBADIS.2015.09.007.
299. Spits, C.; Seneca, S.; Hilven, P.; Liebaers, I.; Sermon, K. Methylation of the CpG sites in the myotonic dystrophy locus does not correlate with CTG expansion size or with the congenital form of the disease. *J. Med. Genet.* **2010**, *47*, 700–703, doi:10.1136/JMG.2009.074211.
300. Steinbach, P.; Gläser, D.; Vogel, W.; Wolf, M.; Schwemmle, S. The DMPK gene of severely affected myotonic dystrophy patients is hypermethylated proximal to the largely expanded CTG repeat. *Am. J. Hum. Genet.* **1998**, *62*, 278–285, doi:10.1086/301711.
301. Légaré, C.; Overend, G.; Guay, S.-P.; Monckton, D.G.; Mathieu, J.; Gagnon, C.; Bouchard, L. DMPK gene DNA methylation levels are associated with muscular and respiratory profiles in DM1. *Neurol. Genet.* **2019**, *5*, e338, doi:10.1212/NXG.0000000000000338.
302. Castel, A.L.; Nakamori, M.; Tomé, S.; Chitayat, D.; Gourdon, G.; Thornton, C.A.; Pearson, C.E. Expanded CTG repeat demarcates a boundary for abnormal CpG methylation in myotonic dystrophy patient tissues. *Hum. Mol. Genet.* **2011**, *20*, 1–15, doi:10.1093/hmg/ddq427.
303. Turker, M.S. Gene silencing in mammalian cells and the spread of DNA methylation. *Oncogene 2002 2135* **2002**, *21*, 5388–5393, doi:10.1038/sj.onc.1205599.
304. Bannister, A.J.; Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res. 2011 213* **2011**, *21*, 381–395, doi:10.1038/cr.2011.22.
305. Wang, Y.H.; Amirhaeri, S.; Kang, S.; Wells, R.D.; Griffith, J.D. Preferential nucleosome assembly at DNA triplet repeats from the myotonic dystrophy gene. *Science* **1994**, *265*, 669–671, doi:10.1126/SCIENCE.8036515.
306. Wang, Y.H.; Griffith, J. Expanded CTG triplet blocks from the myotonic dystrophy gene create the strongest known natural nucleosome positioning elements. *Genomics* **1995**, *25*, 570–573, doi:10.1016/0888-7543(95)80061-P.
307. Volle, C.B.; Delaney, S. CAG/CTG repeats alter the affinity for the histone core and the positioning of DNA in the nucleosome. *Biochemistry* **2012**, *51*, 9814–9825, doi:10.1021/BI301416V.
308. Otten, A.D.; Tapscott, S.J. Triplet repeat expansion in myotonic dystrophy alters the adjacent chromatin structure. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 5465–5469, doi:10.1073/PNAS.92.12.5465.
309. Sorek, M.; Cohen, L.R.Z.; Meshorer, E. Open chromatin structure in PolyQ disease-related genes: a potential mechanism for CAG repeat expansion in the normal human population. *NAR genomics Bioinforma.* **2019**, *1*, doi:10.1093/NARGAB/LQZ003.

310. Esteller, M. Non-coding RNAs in human disease. *Nat. Rev. Genet.* **2011**, *12*, 861–874, doi:10.1038/nrg3074.
311. Guo, H.; Ingolia, N.T.; Weissman, J.S.; Bartel, D.P. Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* **2010**, *466*, 835–840, doi:10.1038/nature09267.
312. Berezikov, E.; Guryev, V.; van de Belt, J.; Wienholds, E.; Plasterk, R.H.A.; Cuppen, E. Phylogenetic Shadowing and Computational Identification of Human microRNA Genes. *Cell* **2005**, *120*, 21–24, doi:10.1016/j.cell.2004.12.031.
313. Condrat, C.E.; Thompson, D.C.; Barbu, M.G.; Bugnar, O.L.; Boboc, A.; Cretoiu, D.; Suci, N.; Cretoiu, S.M.; Voinea, S.C. miRNAs as Biomarkers in Disease: Latest Findings Regarding Their Role in Diagnosis and Prognosis. *Cells* **2020**, *9*, 276, doi:10.3390/cells9020276.
314. Fritegatto, C.; Ferrati, C.; Pegoraro, V.; Angelini, C. Micro-RNA expression in muscle and fiber morphometry in myotonic dystrophy type 1. *Neurol. Sci.* **2017**, doi:10.1007/s10072-017-2811-2.
315. Ambrose, K.K.; Ishak, T.; Lian, L.H.; Goh, K.J.; Wong, K.T.; Ahmad-Annuar, A.; Thong, M.K. Deregulation of microRNAs in blood and skeletal muscles of myotonic dystrophy type 1 patients. *Neurol. India* **2017**, *65*, 512–517, doi:10.4103/neuroindia.NI\_237\_16.
316. Gambardella, S.; Rinaldi, F.; Lepore, S.M.; Viola, A.; Loro, E.; Angelini, C.; Vergani, L.; Novelli, G.; Botta, A. Overexpression of microRNA-206 in the skeletal muscle from myotonic dystrophy type 1 patients. *J. Transl. Med.* **2010**, *8*, 48, doi:10.1186/1479-5876-8-48.
317. Perbellini, R.; Greco, S.; Sarra-Ferraris, G.; Cardani, R.; Capogrossi, M.C.; Meola, G.; Martelli, F. Dysregulation and cellular mislocalization of specific miRNAs in myotonic dystrophy type 1. *Neuromuscul. Disord.* **2011**, *21*, 81–88, doi:10.1016/j.nmd.2010.11.012.
318. Fernandez-Costa, J.M.; Garcia-Lopez, A.; Zuñiga, S.; Fernandez-Pedrosa, V.; Felipo-Benavent, A.; Mata, M.; Jaka, O.; Aiastui, A.; Hernandez-Torres, F.; Aguado, B.; et al. Expanded CTG repeats trigger miRNA alterations in *Drosophila* that are conserved in myotonic dystrophy type 1 patients. *Hum. Mol. Genet.* **2013**, *22*, 704–716, doi:10.1093/hmg/dds478.
319. Rau, F.; Freyermuth, F.; Fugier, C.; Villemin, J.P.; Fischer, M.C.; Jost, B.; Dembele, D.; Gourdon, G.; Nicole, A.; Duboc, D.; et al. Misregulation of miR-1 processing is associated with heart defects in myotonic dystrophy. *Nat. Struct. Mol. Biol.* **2011**, *18*, 840–845, doi:10.1038/nsmb.2067.
320. Kalsotra, A.; Singh, R.K.; Gurha, P.; Ward, A.J.; Creighton, C.J.; Cooper, T.A. The Mef2 transcription network is disrupted in myotonic dystrophy heart tissue, dramatically altering miRNA and mRNA expression. *Cell Rep.* **2014**, *6*, 336–345, doi:10.1016/j.celrep.2013.12.025.
321. Cacchiarelli, D.; Legnini, I.; Martone, J.; Cazzella, V.; D’Amico, A.; Bertini, E.; Bozzoni, I. miRNAs as serum biomarkers for Duchenne muscular dystrophy. *EMBO Mol. Med.* **2011**, *3*, 258, doi:10.1002/EMMM.201100133.
322. Perfetti, A.; Greco, S.; Bugiardini, E.; Cardani, R.; Gaia, P.; Gaetano, C.; Meola, G.; Martelli, F. Plasma microRNAs as biomarkers for myotonic dystrophy type 1. *Neuromuscul. Disord.* **2014**, *24*, 509–515, doi:10.1016/j.nmd.2014.02.005.

## Bibliography

323. Koutsoulidou, A.; Kyriakides, T.C.; Papadimas, G.K.; Christou, Y.; Kararizou, E.; Papanicolaou, E.Z.; Phylactou, L.A. Elevated muscle-specific miRNAs in serum of myotonic dystrophy patients relate to muscle disease progress. *PLoS One* **2015**, *10*, 1–20, doi:10.1371/journal.pone.0125341.
324. Perfetti, A.; Greco, S.; Cardani, R.; Fossati, B.; Cuomo, G.; Valaperta, R.; Ambrogio, F.; Cortese, A.; Botta, A.; Mignarri, A.; et al. Validation of plasma microRNAs as biomarkers for myotonic dystrophy type 1. *Sci. Rep.* **2016**, *6*, 38174, doi:10.1038/srep38174.
325. Koutsoulidou, A.; Photiades, M.; Kyriakides, T.C.; Georgiou, K.; Prokopi, M.; Kapnisis, K.; Lusakowska, A.; Nearchou, M.; Christou, Y.; Papadimas, G.K.; et al. Identification of exosomal muscle-specific miRNAs in serum of myotonic dystrophy patients relating to muscle disease progress. *Hum. Mol. Genet.* **2017**, *26*, 3285–3302, doi:10.1093/hmg/ddx212.
326. Pegoraro, V.; Cudia, P.; Baba, A.; Angelini, C. MyomiRNAs and myostatin as physical rehabilitation biomarkers for myotonic dystrophy. *Neurol. Sci.* **2020**, doi:10.1007/s10072-020-04409-2.
327. Trottier, Y.; Lutz, Y.; Stevanin, G.; Imbert, G.; Devys, D.; Cancel, G.; Saudou, F.; Weber, C.; David, G.; Tora, L.; et al. Polyglutamine expansion as a pathological epitope in Huntington's disease and four dominant cerebellar ataxias. *Nature* **1995**, *378*, 403–6, doi:10.1038/378403a0.
328. Huynh, D.P.; Yang, H.T.; Vakharia, H.; Nguyen, D.; Pulst, S.M. Expansion of the polyQ repeat in ataxin-2 alters its Golgi localization, disrupts the Golgi complex and causes cell death. *Hum. Mol. Genet.* **2003**, *12*, 1485–1496, doi:10.1093/hmg/ddg175.
329. Toulouse, A.; Au-Yeung, F.; Gaspar, C.; Roussel, J.; Dion, P.; Rouleau, G.A. Ribosomal frameshifting on MJD-1 transcripts with long CAG tracts. *Hum. Mol. Genet.* **2005**, *14*, 2649–2660, doi:10.1093/HMG/DDI299.
330. Buendía, G.A.R.; Leleu, M.; Marzetta, F.; Vanzan, L.; Tan, J.Y.; Ythier, V.; Randal, E.L.; Marque, A.C.; Baubec, T.; Murr, R.; et al. Three-dimensional chromatin interactions remain stable upon CAG/CTG repeat expansion. *Sci. Adv.* **2020**, *6*, doi:10.1126/SCIADV.AAZ4012.
331. Yanovsky-Dagan, S.; Cohen, E.; Megalli, P.; Altarescu, G.; Schonberger, O.; Eldar-Geva, T.; Epsztejn-Litman, S.; Eiges, R. DMPK hypermethylation in sperm cells of myotonic dystrophy type 1 patients. *Eur. J. Hum. Genet.* **2021**, doi:10.1038/S41431-021-00999-3.
332. Nichol, K.; Pearson, C.E. CpG methylation modifies the genetic stability of cloned repeat sequences. *Genome Res.* **2002**, *12*, 1246–1256, doi:10.1101/GR.74502.
333. Antequera, F.; Boyes, J.; Bird, A. High levels of de novo methylation and altered chromatin structure at CpG islands in cell lines. *Cell* **1990**, *62*, 503–514, doi:10.1016/0092-8674(90)90015-7.
334. Jones, P.A.; Wolkowicz, M.J.; Rideout, W.M.; Gonzales, F.A.; Marziasz, C.M.; Coetzee, G.A.; Tapscott, S.J. De novo methylation of the MyoD1 CpG island during the establishment of immortal cell lines. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87*, 6117–6121, doi:10.1073/PNAS.87.16.6117.
335. Keller, A.; Meese, E. Can circulating miRNAs live up to the promise of being minimal invasive biomarkers in clinical settings? *Wiley Interdiscip. Rev. RNA* **2016**, *7*, 148–156, doi:10.1002/wrna.1320.

336. Roberts, T.C.; Coenen-Stass, A.M.L.; Wood, M.J.A. Assessment of RT-qPCR normalization strategies for accurate quantification of extracellular microRNAs in murine Serum. *PLoS One* **2014**, *9*, doi:10.1371/journal.pone.0089237.
337. Cheng, Y.; Dong, L.; Zhang, J.; Zhao, Y.; Li, Z. Recent advances in microRNA detection. *Analyst* 2018, *143*, 1758–1774.

*Some of the images in this dissertation are created with Biorender.com*





# FUNDING



This doctoral thesis would not have been possible without the support of several funding bodies:

The project that gave rise to these results received the support of a fellowship from "la Caixa" Foundation (ID100010434), fellowship code LCF/BQ/IN18/11660019.



This project received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 713673



Emma Agathe Koehorst received a travel grant from the COST Action CA17103 'Delivery of Antisense RNA Therapeutics (DARTER)' to attend the COST ACTION meeting in Bilbao.



This project received funding from the Research Projects Institute of Health Carlos III, grant numbers PI15/01756 and PI18/00713, named 'Application of more sensitive techniques of genetic diagnosis, study of phenotype and prognosis modulators in patients with myotonic dystrophies' and 'DIMINUTOS, Childhood and adult myotonic dystrophy: evaluation of new treatments and pathogenicity through genetic, epigenetic and molecular imaging analysis', respectively.



This project received funding from the Ministerio de Ciencia e Innovación, Grant number PID2020-118730RB-I00



This project received funding from the AFM-Telethon, grant project #23557









# **APPENDIX**



APPENDIX I. LIST OF OTHER PUBLICATIONS

---

---

**1. The Need for Establishing a Universal CTG Sizing Method in Myotonic Dystrophy Type 1**

*Authors:* Alfonsina Ballester-Lopez, Ian Linares-Pardo, **Emma Koehorst**, Judit Núñez-Manchón, Guillem Pintos-Morell, Jaume Coll-Cantí, Miriam Almendrote, Giuseppe Lucente, Andrea Arbex, Jonathan J Magaña, Nadia M Murillo-Melo, Alejandro Lucia, Darren G Monckton, Sarah A Cumming, Alba Ramos-Fransi, Alicia Martínez-Piñeiro, Gisela Nogales-Gadea

*Genes (Basel).* 2020 Jul 7;11(7):757. doi: 10.3390/genes11070757

**2. Preliminary Findings on CTG Expansion Determination in Different Tissues from Patients with Myotonic Dystrophy Type 1**

*Authors:* Alfonsina Ballester-Lopez, **Emma Koehorst**, Ian Linares-Pardo, Judit Núñez-Manchón, Miriam Almendrote, Giuseppe Lucente, Andrea Arbex, Carles Puente, Alejandro Lucia, Darren G Monckton, Sarah A Cumming, Guillem Pintos-Morell, Jaume Coll-Cantí, Alba Ramos-Fransi, Alicia Martínez-Piñeiro, Gisela Nogales-Gadea

*Genes (Basel).* 2020 Nov 7;11(11):1321. doi: 10.3390/genes11111321

**3. Three-dimensional imaging in myotonic dystrophy type 1: Linking molecular alterations with disease phenotype**

*Authors:* Alfonsina Ballester-Lopez, Judit Núñez-Manchón, **Emma Koehorst**, Ian Linares-Pardo, Miriam Almendrote, Giuseppe Lucente, Nicolau Guanyabens, Marta Lopez-Osias, Adrián Suárez-Mesa, Shaliza Ann Hanick, Jakub Chojnacki, Alejandro Lucia, Guillem Pintos-Morell, Jaume Coll-Cantí, Alicia Martínez-Piñeiro, Alba Ramos-Fransi, Gisela Nogales-Gadea

*Neurol Genet.* 2020 Jul 21;6(4):e484. doi: 10.1212/NXG.000000000000484



April 2022

Barcelona







