

**Określenie wpływu nieprawidłowości splicingu mRNA
czynnika transkrypcyjnego NFIX na patomechanizm
dystrofii miotonicznej oraz opracowanie narzędzia terapii
genowej dla tej choroby**

Zuzanna Rogalska

Rozprawa doktorska

Promotor: prof. dr hab. Krzysztof Sobczak



Zakład Ekspresji Genów,
Instytut Biologii Molekularnej i Biotechnologii,
Uniwersytet im. Adama Mickiewicza w Poznaniu

Poznań 2023

**Determination of the impact of splicing abnormalities
of the NFIX transcription factor mRNA
on the pathomechanism of myotonic dystrophy and
development of gene therapy tool for this disease**

Zuzanna Rogalska

PhD thesis

Supervisor: prof. dr hab. Krzysztof Sobczak



Department of Gene Expression,
Institute of Molecular Biology and Biotechnology,
Adam Mickiewicz University

Poznań 2023

I would like to thank:

Prof. Krzysztof Sobczak – for sharing your professional knowledge, assistance, time, and invaluable contribution to my thesis

All my colleagues, lab, and institute members for their help, support, inspiration, advice, and knowledge, especially Piotr Cywoniuk, Kasia Taylor, Ania Baud, Kasia Tutak, and Sandra Wejnerowska.

My family and friends for your invaluable support, love, and encouragement, especially my Parents, sister, and Maciek.

CONTENTS

STRESZCZENIE.....	6
SUMMARY.....	10
1. INTRODUCTION	13
1.1 GENETICS, PATHOMECHANISM AND THERAPY OF MYOTONIC DYSTROPHIES.....	13
1.1.1 <i>Mutations in myotonic dystrophy genes.....</i>	<i>13</i>
1.1.2 <i>Molecular pathomechanism of DM.....</i>	<i>15</i>
1.1.3 <i>Experimental therapies of DM1.....</i>	<i>17</i>
1.2. NUCLEAR FACTOR IX (NFIX) TRANSCRIPTION FACTOR.....	22
1.2.1 <i>Family of NFI transcription factors.....</i>	<i>23</i>
1.2.2 <i>Structure of NFIX gene products.....</i>	<i>24</i>
1.2.3 <i>Expression and function of NFIX gene.....</i>	<i>25</i>
1.2.4 <i>Alternative splicing of NFIX.....</i>	<i>27</i>
1.2.5 <i>Mutations of NFIX associated with human diseases.....</i>	<i>27</i>
1.2.6 <i>Nfix deficient mice.....</i>	<i>28</i>
2. AIMS OF THE STUDY	29
3. CONTRIBUTION OF SPLICING ABNORMALITIES OF NFIX EXON 7 ON PATHOMECHANISM OF MYOTONIC DYSTROPHY.....	30
3.1 RESULTS AND DISCUSSION	30
3.1.1 <i>Changes in alternative splicing of NFIX exon 7 in skeletal muscles of DM patients.....</i>	<i>30</i>
3.1.2 <i>Generation of stable HEK-293 cell line with a functional knockout of NFIX (NFIX-KO).....</i>	<i>34</i>
3.1.3 <i>Generation of genetic constructs for overexpression of two splicing isoforms of NFIX.....</i>	<i>39</i>
3.1.4 <i>Generation of two stable cell lines expressing NFIX+7 or NFIX-7 using Flp-In™ T-REx™ system.....</i>	<i>41</i>
3.1.5 <i>Expression of other members of NFI family is not affected in NFIX-KO and either NFIX isoforms expressing HEK293 cells.....</i>	<i>44</i>
3.1.6 <i>NFIX regulates the expression of genes involved in RNA metabolism in generated HEK293 cells.....</i>	<i>46</i>
3.1.7 <i>Generation of human skeletal muscle cell models with enrichment of either NFIX+7 or NFIX-7 isoform.....</i>	<i>52</i>
3.1.8 <i>Generation of human skeletal muscle cell model with insufficiency of NFIX.....</i>	<i>57</i>
3.1.9 <i>Identification of genes sensitive to the level of NFIX and to its splicing isoforms.....</i>	<i>59</i>
3.1.10 <i>Abnormal splicing of NFIX exon 7 significantly contributes to the gene expression changes in skeletal muscles of DM1 patients.....</i>	<i>63</i>
3.2 DISCUSSION.....	68
3.2.1 <i>NFIX is both a positive and negative regulator of gene expression.....</i>	<i>68</i>
3.2.2 <i>NFIX+7 is a more effective activator of gene expression regulation.....</i>	<i>70</i>
3.2.3 <i>The enrichment of the NFIX+7 isoform leads to the abnormal gene expression observed in DM1.....</i>	<i>72</i>
3.3 CONCLUSIONS.....	76
3.4. MATERIALS AND METHODS	77
3.5 BIBLIOGRAPHY	88
4. SUSTAINABLE RECOVERY OF MBNL ACTIVITY IN AUTOREGULATORY FEEDBACK LOOP IN MYOTONIC DYSTROPHY.....	100

ABBREVIATIONS

5' or 3'-UTRs 5' or 3' untranslated regions
AAV adeno-associated virus
ASOs antisense oligonucleotides
Au arbitrary units au
circRNA circular RNA
CNS central nervous system
CNBP cellular nucleic acid binding protein
CPM Counts per Million
Ct cycle threshold
CTG^{exp} or CCTG^{exp} CTG or CCTG expansion
CTRL control
CQ chloroquine
DM1 and DM2 Myotonic dystrophy type 1 and type 2
DMPK dystrophia myotonica protein kinase
DOX doxycycline
DSB double-strand breaks
ex22 exon 22
FISH fluorescence *in situ* hybridization
FRT FLP recombinase target sites
GO Gene Ontology
HSA human skeletal actin gene
HSkM human skeletal myoblast
KO knockout
MBNLs Muscleblind-like proteins
MEFs mouse embryonic fibroblasts
MYOD myogenic differentiation
MYOG myogenin
NFIX nuclear factor I X gene
Nfix-KO *Nfix* knockout
ns non-significant
PSI percent-spliced-in
PBZ phenylbutazone
RNA-seq RNA sequencing
SD standard deviation
siNFIX siRNA against *NFIX*
WT wild type
ZnF zinc fingers

STRESZCZENIE

Dystrofia miotoniczna typu 1 (DM1) jest dziedziczną, dominującą chorobą genetyczną, spowodowaną ekspansją trójnukleotydowych powtórzeń CTG w genie *DMPK*. Nadmiernie wydłużone powtórzenia CUG (CUG^{exp}) w zmutowanym mRNA są toksycznym produktem zmutowanego genu. Zmutowany mRNA sekwestruje czynniki splicingowe takie jak MBNL, a następnie w wyniku funkcjonalnego niedoboru tych białek wywołuje globalne zmiany w alternatywnym splicingu w mięśniach szkieletowych, sercu i mózgu, czego efektem są objawy charakterystyczne dla DM1. Nieprawidłowości w alternatywnym splicingu są kluczową molekularną przyczyną rozwoju DM1. Profil alternatywny splicingu wielu genów zaangażowanych w homeostazę mięśni i ich prawidłowe funkcjonowanie jest istotnie zaburzony. Niemniej jednak, molekularne konsekwencje zaburzeń alternatywnego splicingu większości genów są nadal nieznane.

Jednym z eksonów zależnych od MBNL, którego efektywność włączenia do mRNA w procesie splicingu jest istotnie podwyższona w mięśniach szkieletowych DM, jest ekson 7 *NFIX*, genu kodującego czynnik transkrypcyjny niezbędny do prawidłowego rozwoju mięśni. Pierwszym celem niniejszej pracy było lepsze zrozumienie patomechanizmu DM w kontekście zaburzeń splicingowych *NFIX* poprzez zrozumienie wpływu włączenia eksonu 7 na aktywność transkrypcyjną *NFIX*. Aby odpowiedzieć na to pytanie, zastosowano dwa alternatywne podejścia. Pierwszym z nich było stworzenie dwóch stabilnych linii komórkowych z indukowalną nadekspresją izoform *NFIX* z eksonem 7 i bez eksonu 7 (*NFIX*+7 lub *NFIX*-7) w oparciu o komórki HEK z funkcjonalnym nokautem endogenego *NFIX* (*NFIX*-KO). Nieoczekiwane wyniki RNA-seq nie wykazały znaczących różnic w aktywności transkrypcyjnej tych dwóch izoform w stworzonych modelach komórkowych, być może z powodu niewłaściwego dopasowania poziomu ekspresji egzogenów (zbyt wysoki). Drugie podejście opierało się na manipulacji włączania eksonu 7 do endogenie ekspresowanego

mRNA *NFIX* przy użyciu antysensownego oligonukleotydu (AON) nakierowanego na połączenie eksonu 7 z intronem 7. W ludzkich komórkach mięśni szkieletowych, w których dominuje izoforma *NFIX*+7, zastosowany AON indukował efektywne pomijanie eksonu 7 i produkcję głównie izoformy *NFIX*-7. Takie podejście umożliwiło wygenerowanie modelu komórkowego w kontekście środowiska, w którym *NFIX* jest naturalnie zaangażowany w proces prawidłowego rozwoju mięśni, ale także w patogenezę DM1. W tych komórkach ekspresja *NFIX* jest stosunkowo wysoka i jest regulowana przez natywny promotor. Wyniki eksperymentów RNA-seq ujawniły setki genów, których ekspresja uległa istotnym zmianom po traktowaniu AON, co sugeruje, że obie izoformy splicingowe - *NFIX*+7 i *NFIX*-7 - różnią się znacząco aktywnością transkrypcyjną. Analiza ontologii genów (GO) wykazała znaczne wzbogacenie klas genów, które są wrażliwe na te dwie izoformy *NFIX*, ale także zaobserwowano istotną zmianę ich ekspresji w mięśniach pacjentów z DM1. Wśród 5641 genów, których ekspresja jest znacząco zmieniona w tkankach DM1 ($P < 0,05$) zidentyfikowano 2070 genów, których ekspresja była znacząco zmieniona w komórkach mięśniowych poddanych działaniu AON zmieniającemu splicing *NFIX*. Głównymi terminami GO wzbogaconymi w obu porównaniach (w pierwszym – geny wrażliwe na poziom *NFIX* oraz na obecność jego izoform splicingowych, w drugim – geny wrażliwe na izoformy splicingowe *NFIX* oraz zmienione w tkankach pacjentów DM1) były geny składników strukturalnych macierzy zewnątrzkomórkowej i geny białek wiążących się z kolagenami. Są to geny niezmiernie istotne dla prawidłowej budowy i aktywności mięśni szkieletowych, a zatem zaburzenie splicingu *NFIX* może stanowić istotny element patogenezy DM1. Podsumowując tę część pracy można stwierdzić, że nieprawidłowa dystrybucja eksonu 7 *NFIX* spowodowana sekwestracją białek MBNL na RNA ze zmutowanym CUG^{exp} wywołuje liczne zmiany ekspresji genów zachodzących w mięśniach szkieletowych, które to mogą być odpowiedzialne za kształtowanie fenotypu chorobowego obserwowanego u pacjentów z DM1.

DM jest chorobą nieuleczalną. Kiedy liczba powtórzeń osiąga wysoki poziom (od setek do tysięcy), występujące objawy kliniczne są cięższe oraz wzrasta wydajność sekwestracji zależnej od długości powtórzeń. Proponowane wcześniej strategie terapeutyczne prowadzące do podwyższenia poziomu MBNL1 za pomocą narzędzi terapii genowej wiążą się z występowaniem wielu niepożądanych konsekwencji. Długotrwała, niekontrolowana nadekspresja MBNL1 u myszy prowadzi bowiem do zmniejszenia masy ciała, zwiększonej śmiertelności i uszkodzenia mięśni, w tym mięśnia sercowego. Terapia genowa prowadząca do podwyższenia poziomu MBNL1 u pacjentów z DM1 wiąże się również z wieloma innymi ograniczeniami. Jednym z nich jest niejednorodność sekwestracji MBNL spowodowana mozaicyzmem somatycznym długości powtórzeń CTG w komórkach tego samego pacjenta oraz pomiędzy różnymi osobami ze zdiagnozowanym DM1. Aby sprostać tym ograniczeniom, zaprojektowałam i wygenerowałam autoregulowany konstrukt do nadekspresji *MBNL1*, który umożliwia produkcję białka MBNL1 dostosowaną do poziomu aktywnej puli endogennych białek MBNL w komórce. Ta strategia terapeutyczna daje więc możliwość przezwyciężenia ograniczeń wynikających z niejednorodności ekspansji powtórzeń CTG i w konsekwencji różnego stopnia niedoboru MBNL w różnych komórkach czy włóknach mięśniowych. Konstrukt ten zawiera sekwencję kodującą MBNL1 przedzieloną fragmentem pre-mRNA *ATP2A1* z alternatywnym eksonem wrażliwym na poziom MBNL, który zawiera kodon stop w ramce odczytu dla MBNL1. Włączenie tego eksonu prowadzi więc do tworzenia nieaktywnej, skróconej formy białka, natomiast jego wyłączenie, następujące przy niedoborze MBNL, powoduje wzrost produkcji w pełni aktywnej formy MBNL1. Takie podejście umożliwia ograniczoną i ściśle kontrolowaną nadekspresję *MBNL1*, w zależności od stopnia niedoboru białka, a zarazem może odpowiadać na niejednorodność długości powtórzeń CUG. Podsumowując tę część pracy można powiedzieć, że przeprowadzone badania wykazały, że nadekspresja MBNL1 ze stworzonego konstrukt genetycznego podlega autoregulacji i ma

potencjał terapeutyczny prowadzący do korygowania nieprawidłowości alternatywnego splicingu, co wykazano dla komórek pochodzących od pacjentów z DM1.

SUMMARY

Myotonic dystrophy type 1 is a hereditary, autosomal disease caused by expansion of trinucleotide CTG repeats in *DMPK* gene. Expanded CUG repeats in mutant mRNA is a major toxic product of mutant gene. It sequester MBNL splicing factors and due to functional insufficiency of these proteins trigger global changes in alternative splicing in skeletal muscles, heart and brain, which result in symptoms characteristic for DM1. The abnormalities in alternative splicing are crucial molecular feature of DM1. The alternative splicing of many genes engaged in muscle homeostasis and proper functioning is impaired. Although, the molecular consequence of majority of genes with altered splicing pattern is still unknown.

One of the MBNL-dependent exons, which inclusion is significantly higher in skeletal muscles of DM, is exon 7 of *NFIX*, a gene encoding for transcription factor essential for muscle development. First aim of this study was to better understand the pathomechanism of DM in case of abnormalities in the splicing of *NFIX*. The examination of whether the contribution of exon 7 has an impact on NFIX transcriptional activity gave a deeper understanding of DM1 pathomechanism which is studied for a long time. To answer this question two alternative approaches were used. First of them was generation of two stable cell lines with inducible overexpression of NFIX isoforms with and without exon 7 (NFIX+7 or NFIX-7) based on a HEK cell with functional knockout of endogenous *NFIX* (*NFIX*-KO). Unexpectedly, RNA-seq results did not show significant differences in transcriptional activity of these two isoforms in developed cellular models, perhaps due to not well adjusted level of overexpression of exogenes (too high). The second approach based on manipulation of exon 7 inclusion of endogenous NFIX mRNA using the antisense oligonucleotide (AON) targeting (AON) targeting exon 7/intron 7 boundary. In human skeletal muscle cells, in which NFIX+7 isoform predominates, this AON induces efficient exon 7 skipping and production of NFIX-7 isoform.

This approach gave the possibility to generate a model with an cellular environment where NFIX is engaged in the developmental process, but also in DM1 pathogenesis. The expression level of NFIX is relatively high in these cells and is driven by the native promoter. The results of RNA-seq revealed hundreds of genes which expression was significantly changed after AON treatment, suggesting that both splicing isoforms, NFIX+7 and NFIX-7, differ significantly in transcriptional activity. Gene ontology (GO) analysis showed significant enrichment of groups of genes that are sensitive to NFIX isoforms and abnormally expressed in DM1 patients. Among 5641 genes which expression is affected in DM1 ($P < 0.05$), 2070 genes were identified which are also sensitive to treatment with AON modulating splicing pattern of NFIX. The major GO molecular function terms enriched in both comparisons (first: genes sensitive to the level of NFIX and its splicing isoforms; second: genes sensitive to the NFIX splicing isoforms and changed in DM1 tissues) were extracellular matrix structural constituent, and collagen binding. These functions are incredibly important for the proper structure and activity of skeletal muscle, therefore *NFIX* splicing abnormality may be essential trigger of DM1 pathogenesis. Taken together, these results showed that abnormal distribution of NFIX exon 7 caused by sequestration of MBNL proteins on CUG^{exp} affects a subset of gene expression changes occurring in DM1 skeletal muscles, which may be responsible for the development of disease phenotype observed in DM1 patients.

DM is an incurable disease. When the number of repeats reach a higher level clinical symptoms are more severe and the sequestration efficiency, which depends on the size expansion, increase. Previously proposed therapeutic strategies leading to increase of the MBNL1 level via gene therapy tools are associated with many undesirable consequences. Uncontrolled MBNL1 overexpression in mice leads to reduced body weight, increased mortality, or muscle damage, including heart muscle. The gene therapy leading to the increase of MBNL1 level is also associated with many limitations. One of them is heterogeneity in sequestration of MBNLs

caused by somatic mosaicism of CTG repeat length in cells of the same patient and between individuals with DM1. To deal with these limitations I designed and generated the autoregulated MBNL1 overexpression construct which enables the production of MBNL1 adjusted to the level of active pool of endogenous MBNLs. It was assumed that this therapeutic strategy gave the possibility to overcome the limitations caused by heterogeneity of CTG repeat expansions and as a consequence different levels of MBNL insufficiency in different cells/myofibers. This construct contains MBNL1-coding sequence separated by the fragment of *ATP2A1* pre-mRNA with MBNL-sensitive alternative exon containing in frame stop codon. The inclusion of this exon leads to the arrangement of the inactive form of the protein but its exclusion, occurring during MBNL insufficiency, gives rise in production of fully active MBNL1. This approach enables the restricted expression of MBNL1 protein only if its level in cell is too low and potentially can be controlled by heterogeneity of CUG^{exp} load. It was shown that expression of MBNL1 assembled from this construct is tunable and has therapeutic potential to correct the alternative splicing abnormalities in DM1 patients-derived cells.

1. INTRODUCTION

1.1 Genetics, pathomechanism and therapy of myotonic dystrophies

Microsatellite sequences are abundant repeated tracts in prokaryotes and eukaryotes. In the human genome occupy about 3% of the entire sequence (Lander et al., 2001; Tóth et al., 2000). Trinucleotide repeats are microsatellites whose abnormal expansion can lead to neuro-muscular, neurodevelopmental, or neurodegenerative disorders like myotonic dystrophies, Huntington's disease, spinobulbar muscular atrophy, several spinocerebellar ataxias, Friedreich ataxia, fragile X syndromes (Lander et al., 2001). The expansion can occur both in coding and non-coding regions like 5' or 3' untranslated regions (5' or 3'-UTRs) and introns. The repeats are usually highly unstable in germline tissue and repeat instability is a function of the length of repeat tract (Fortune, 2000). The situation when repeats enlarge in next-generation and therefore progeny exhibit earlier and more severe disease symptoms are called genetic anticipation (Mirkin, 2007).

When the number of repeats exceeds the specific threshold for each gene, one of two types of mutation of an affected gene can occur. First, loss of gene function, which leads to gene silencing, and second gain of function in which toxic RNA or toxic protein can be a major driver of disease development (Wenstrom, 2002).

1.1.1 Mutations in myotonic dystrophy genes

Myotonic dystrophy type 1 and type 2 (DM1 and DM2) are examples with a RNA gain of function mechanism. The mutant non-coding transcript with exceeded number of repeats, abnormally expanded are retained in the nucleus and sequester particular proteins (Ranum & Cooper, 2006; Wheeler & Thornton, 2007). These diseases are autosomal dominant disorders caused by the expansion of either CTG (CTG^{exp}) repeats in the 3'UTR of the dystrophin myotonia protein kinase (*DMPK*) gene (Fig. 1a) or CCTG (CCTG^{exp}) repeats within the intron

1 of cellular nucleic acid binding protein (*CNBP*) gene, respectively (Brook et al., 1992), (Liquori et al., 2001). The main symptoms of DM1 include muscle weakness and wasting, myotonia, cataracts, cardiac abnormalities, and intellectual disability (Brook et al., 1992), (Liquori et al., 2001).

Healthy individuals have 5-37 CTG repeats in 3'UTR of *DMPK*, and a number between 38-50 is considered as a premutation allele. Premutation carriers have no disease-specific symptoms but their repeat tracts have a tendency to expand during maternal or paternal transmission (Martorell et al., 2001). When the number of repeats is at least 50, the patients may have mild symptoms, however, if they reach hundreds to thousands clinical symptoms are more severe and are diagnosed earlier (Ozinski et al., 2021). Moreover, during the lifespan of individuals, the length of CTG repeats can be significantly different between tissues which is a result of somatic repeat instability (Gourdon et al., 1997; Reddy & Housman, 1997). The variation in repeat length can range from tens to hundreds or even thousands (Monckton et al., 1995). Based on the period of diseased development there are recognized four types of DM1: congenital, onsets at birth, childhood/juvenile, or adult form (Aslanidis et al., 1992). The most severe form is congenital DM, with a high mortality rate during the neonatal period, where children receive CTG repeats expansion from their mothers, who might be asymptomatic premutation carrier (Lanni & Pearson, 2019). This form is usually characterized by the most severe DM1 symptoms like hypotonia, respiratory failure, intellectual disability, and cognitive deficits (De Serres-Bérard et al., 2021).

In DM2 patients the expansion size may reach more than 10,000 repeats (Liquori et al., 2001). Because of somatic repeat expansion, the length of CTG^{exp} and CCTG^{exp} is varied between tissues; the longer repeat expansion is observed in skeletal muscles and heart than in blood (Monckton et al., 1995; Thornton et al., 1994; Zatz et al., 1995). Moreover, the length of CCTG^{exp} may also be deeply heterogeneous in the same patient (Massimiliano Alfano 2022) or

CTG^{exp} maybe even different between parts of the muscle of the same patient (Ballester-Lopez et al., 2020; Monckton et al., 1995).

1.1.2 Molecular pathomechanism of DM

The CUG or CCUG repeat tract in mutant RNA assembles a thermodynamically stable hairpin structure containing repetitive structural motifs with two Watson-Crick C-G base pairs which are interrupted by U•U mismatches (Liquori et al., 2001; TIAN et al., 2000). RNAs with CUG expansion (CUG^{exp}) or CCUG expansion (CCUG^{exp}) interact with many proteins and form mostly multiple nuclear foci. One group of interacting proteins are Muscleblind-like proteins (MBNLs) which are efficiently sequestered within these nuclear structures (Fig. 1b) (Fardaei, 2002) and the sequestration efficiency depends on the size of expansion.

Three MBNL paralogs, MBNL1, MBNL2, and MBNL3, are RNA-binding proteins essential for proper RNA processing. They primarily regulate alternative splicing but also RNA localization, stability, and alternative polyadenylation (Batra et al., 2015; Ho et al., 2004; E. T. Wang et al., 2012). MBNL1 is expressed mostly in adult skeletal muscles, MBNL2 in adult brain, and MBNL3 in the placenta and during muscle cell differentiation. MBNLs regulating hundreds of alternative splicing and polyadenylation facilitate the transition from fetal- to adult-specific mRNA isoforms (Fardaei, 2002; Fernandez-Costa et al., 2011; Kanadia et al., 2003). All MBNL family members recognize their RNA targets through clustered YGCY sequence motifs, where Y represents a pyrimidine, and bind to them by zinc fingers (ZnF) organized in two tandem domains localized at N-terminus (Du et al., 2010; Teplova & Patel, 2008). The location of MBNLs binding site within pre-mRNA targets determines if the regulation of alternative exon is positive or negative. The presence of the preferred sequence motif upstream or within the exon leads to exon exclusion while downstream facilitate exon inclusion (Fig. 2) (Du et al., 2010; Goers et al., 2010).

The sequestration of MBNLs to the pathogenic CUG and CCUG repeats leads to abnormalities of many alternative splicing events, especially in skeletal muscles and heart, some of which are involved in specific disease symptoms (Ho et al., 2004). More recent studies showed the correlation between changes in the splicing pattern of known MBNL targets and DM1 symptoms like muscle weakness and wasting (*BINI*, *DNTA*, *DMD*, *CAPN3*) (Fugier et al., 2011, Nakamori et al., 2008, Nakamori et al., 2013, Lin et al., 2006), myotonia (*CLCNI*) (Lueck et al., 2007), and insulin resistance (*INSR*) (Savkur et al., 2001). However, the effect of abnormal splicing of the majority of other MBNL targets is unknown. One of them is the nuclear factor I X gene (*NFIX*), encoding for a transcription factor involved in neurogenesis and myogenesis (Yamashita et al., 2012). One of two major aims of my project was to understand the consequences of abnormalities in the splicing of *NFIX* exon 7 in DM. Therefore, in the next 1.2 subchapter the structure and function of *NFIX* are described.

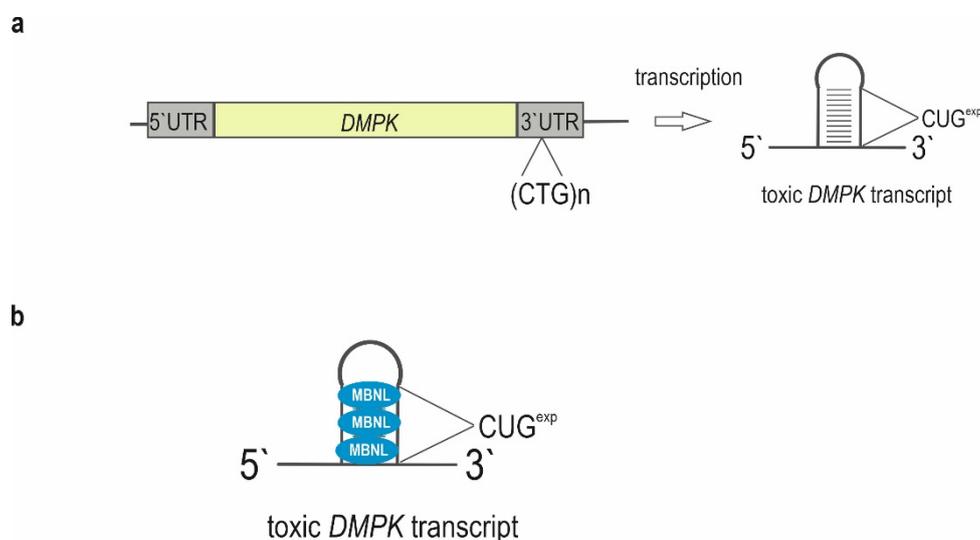


Figure 1 The scheme of the molecular pathomechanism of DM1.

a) DM1 is caused by (CTG)_n expansion within 3'UTR of *DMPK*. The toxic transcript of this gene assembles a thermodynamically stable hairpin structure containing expanded CUG repeats which

b) efficiently bind and sequester MBNL proteins.

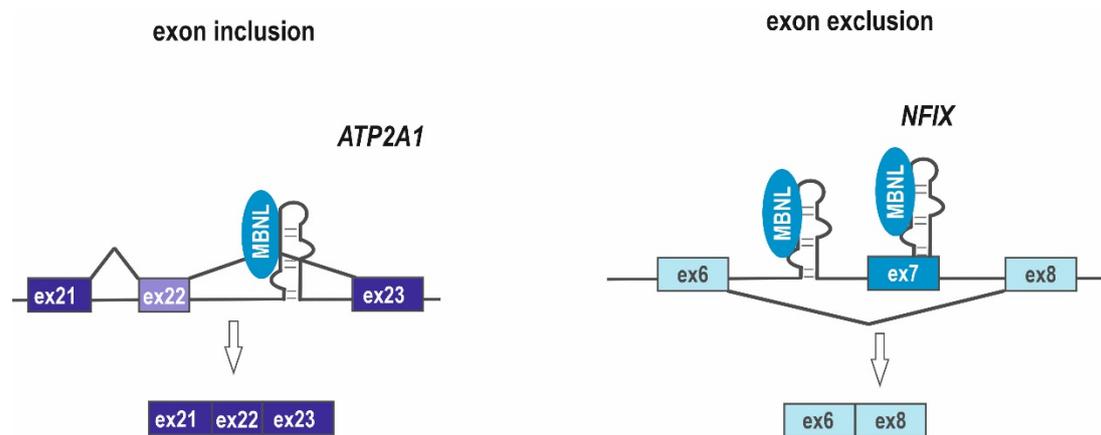


Figure 2. The pattern of alternative splicing regulation by MBNL proteins.

Binding of MBNLs to the sequence downstream of the alternative exon lead to the exon inclusion, as exemplified in left panel for *ATP2A1* exon 22; binding to the upstream sequence, or within the alternative exon, promotes alternative exon exclusion, as exemplified in right panel for *NFIX* exon 7.

1.1.3. Experimental therapies of DM1

Thus far there is no treatment for DM. Nonetheless, there are a few approaches that have been already tested in DM1 models and in a certain way displayed rescue of the disease phenotype. The majority of therapeutic strategies concern enhancing the activity of MBNL protein. It can be achieved by 1) increasing the expression of endogenous *MBNL1* (Fig. 3a, b) (Chen et al., 2016), 2) release of MBNLs from sequestration on the mutant *DMPK* transcript using small molecules (Fig. 3c) or antisense oligonucleotides (ASOs) (Fig. 3d) (Nakamori et al., 2016; Wheeler et al., 2009) or 3) overexpression of exogenous *MBNL1* (Fig. 3e) (Kanadia et al., 2006).

The first mentioned category involves manipulation of the endogenous pool of MBNLs is modifying the activity of MBNL1 promoter using the nonsteroidal anti-inflammatory drugs (NSAIDs). It was shown that phenylbutazone (PBZ), one of the NSAIDs, leads to an increase in *MBNL1* expression, through suppression of methylation region MeR2 with enhancer activity within intron 1 of *MBNL1*.

Subsequently, the improvement of muscle histopathology and rescue of alternative splicing defects were observed (Fig. 3a) (Chen et al., 2016). The next proposed therapy to increase endogenous MBNL1 levels is interfering with the autophagy process using autophagy blocker, the chloroquine (CQ) (Fig. 3a) (Bargiela et al., 2015, 2019). It was proven that in the DM1 model autophagy is increased and may contribute to muscle atrophy (Loro et al., 2010). An alternative approach regards silencing of two microRNAs, *miR-23b* and *miR-218*, which in normal conditions negatively regulate *MBNL1* and *MBNL2* mRNA. The use of particular antagomiRs upregulates the MBNL level and improves muscle defects specific for the DM1 phenotype (Fig. 3b) (Cerro-Herreros et al., 2018, 2020).

The second therapeutic strategy involves the implementation of ligands that binds to the CUG^{exp} and inhibit the formation of MBNL1-toxic CUG repeats complexes (Fig. 3c) (Arambula et al., 2009; Wong et al., 2014). One of the defined compounds is furamidine which demonstrates the reverse of miss-splicing events in the DM1 model and importantly, did not trigger the toxicity (Siboni et al., 2015). The second identified small molecule with proven potential in DM1 therapy is erythromycin, which benefits is the fact that is already used as an antibiotic and well-tolerated in humans. This compound exhibits high affinity to the CUG repeats, leading to the release of MBNL1 and afterward, improves splicing abnormalities and myotonia (Nakamori et al., 2016). What is more, the additive treatment with the application of erythromycin and furamidine leads to the rescue of more MBNL1-sensitive alternative splicing defects compared to the use of one of them (Jenquin et al., 2019). The last discussed, but not least small molecule

which abrogates the formation of MBNL-CUG^{exp} complexes is compound known as 2b. It has the ability to decrease the number of nuclear foci and improve particular splicing events in DM1 cells and a transgenic mouse model of DM1 (Angelbello et al., 2021).

Apart from the application of small molecules to inhibit the formation of pathogenic complexes of MBNLs with CUG^{exp} is the implementation of ASO (Fig. 3d). According to the mode of action they can be divided into two categories: 1) Blocking ASO which binds to their target and inhibit the interaction with other proteins (Aartsma-Rus, 2017) 2) Gapmers, recruiting RNase H to degrade their target (Crooke et al., 1995). One of the examined DM1 therapy candidates is ASO, composed of locked nucleic acids (LNAs), characterized by high binding affinity and also deprived of the ability to activate RNase H. Importantly, it effectively prevents sequestration of MBNL1 to the CUG^{exp} through binding to the repeats. Indeed, both in DM1 cells and mouse models the reduction of CUG^{exp} foci and correction of abnormal splicing events were observed (Wheeler et al., 2009).

The most straightforward approach seems to be the overexpression of exogenous MBNL1 (Fig. 3e). For the first time, it was described in the DM1 mouse model, *HSA-LR*, that expresses transgene with 220 CTG repeats in 3'UTR of human skeletal actin gene (*HSA*). The intramuscular injection of adeno-associated virus (AAV) encoding for MBNL1 induced the alternative splicing correction and rescued muscle hyperexcitability. On the contrary, the normal structure of myofibers was not restored, perhaps because of unsatisfactory overexpression (Kanadia et al., 2006). On a similar theme, recently described, overexpression of MBNL1 Δ -decoy protein that is deprived of C-terminal domain of MBNL1, which absence causes this engineered protein has still a high binding affinity to the CUG^{exp} but has diminished splicing activity (Fig. 3e). The transduction via AAV vectors in the *HSA-LR* mouse model demonstrated rescue of abnormal splicing events, correction of myopathy, and myotonia (Arandel et al., 2022).

The sequestration of MBNLs is directly associated with alternative splicing alterations and subsequently is the reason for many DM1 symptoms. Thus, the restoration of protein function seems to be one of the best targets in therapy. The increasing number of candidates in potential therapy has been already proposed but they correlate with many limitations. First of all, the binding of MBNLs to the CUG^{exp} is highly heterogeneous because of somatic mosaicism of repeat length. This effect is also observed in splicing changes, which are not reproduced in all tissues of the same patient (López-Martínez et al., 2020). Additionally, the approaches regarding enhancing the endogenous MBNL1 level in an indirect way may provoke an off-target effect (Jenquin et al., 2018, 2019). Based on the many types of research carried out so far, it can be assumed that some of above mentioned approaches may be safe and effective in potential DM1 therapy but require more excessive investigation.

DM1 therapeutic strategies - enhance the MBNL level

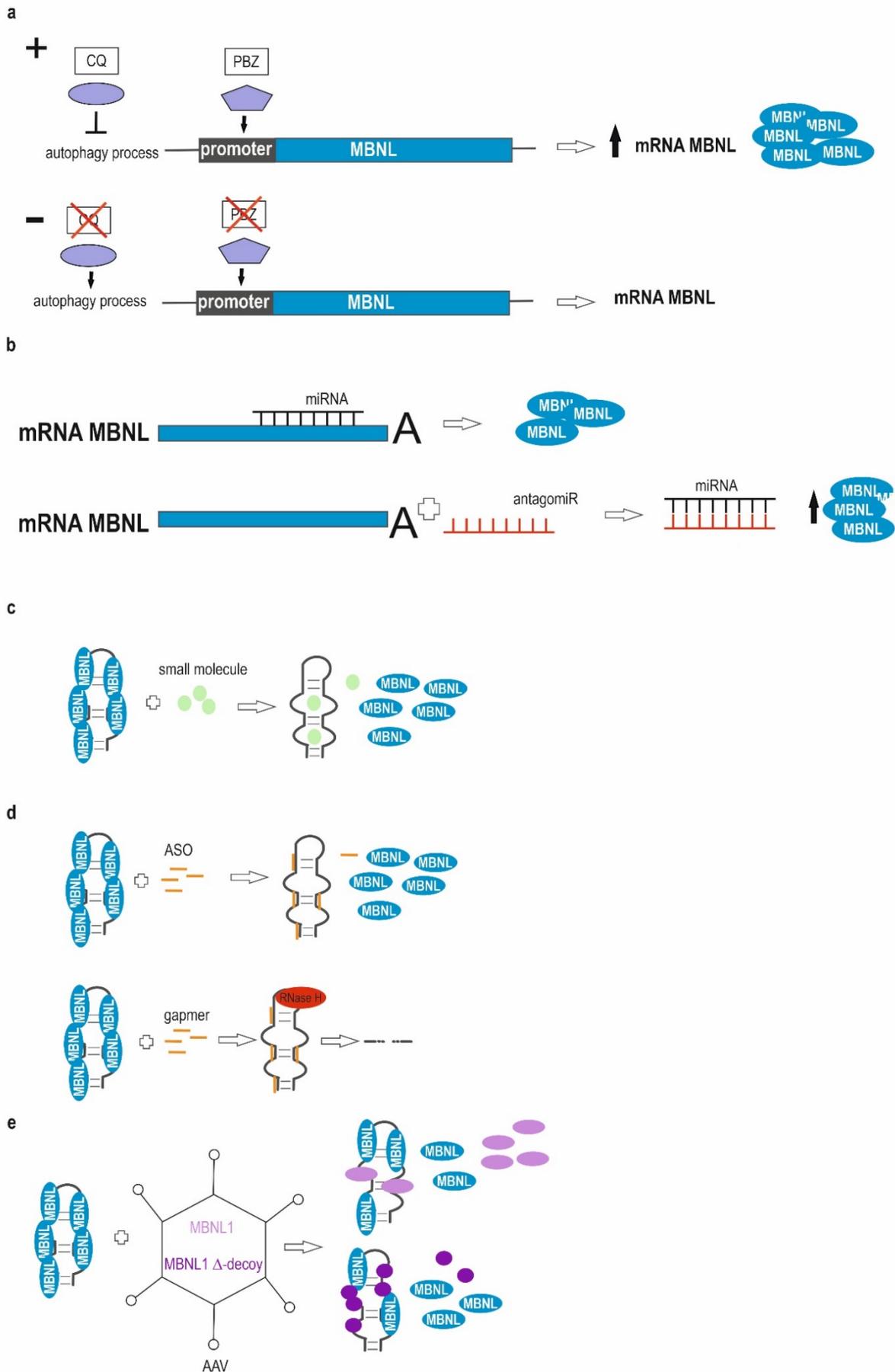


Figure 3. Therapeutic strategies of DM1 leading to the increase of activity of MBNL proteins.

- a) PBZ, one of the NSAIDs drug, activate transcription from the *MBNL1* locus by affecting promoter and CQ, antiautophagic drug, upregulating MBNL level through inhibition the autophagy process in which MBNL is degraded;
- b) silencing the specific miRNA (black) which negatively regulate MBNL1 mRNA (blue) with antagomir (red);
- c) small molecules (green) which bind with high affinity to the CUG^{exp}, inhibit MBNL sequestration and formation of ribonuclear foci;
- d) blocking ASOs (orange), which bind to the CUG^{exp} and block the interaction with MBNL (top graph); Gapmers (orange), which recruiting RNase H (red) to degrade their target
- e) overexpression of exogenous MBNL1 or MBNL1 Δ -decoy from AAV vector. In the case of MBNL1 overexpression, the therapeutic activity has the exogenous fully active MBNL1 which can bind CUG^{exp} and release endogenous MBNLs but also can regulate splicing of target pre-mRNAs (lighter purple). In the second example, MBNL1 Δ -decoy binds with high affinity to the CUG^{exp} and thus releases endogenous MBNLs from sequestration, restoring its function (darker purple).

1.2. Nuclear Factor I X (NFIX) transcription factor

As mentioned above, *NFIX* exon 7 is one of many MBNL-sensitive exons. Abnormal inclusion of this exon is observed in DM1. It was shown that all MBNLs negatively regulate the alternative splicing of exon 7 through binding to the YGCY motifs which are present both in the upstream intron and within the exon 7 (Fig. 4) (Du et al., 2010). Therefore, in tissues with high level of MBNLs exon 7 exclusion mRNA isoforms predominates and in tissue with low level of MBNLs or in DM conditions isoform with exon 7 is produced more frequently. Because one of the aims of this study is to better understand the pathomechanism of DM in case of abnormalities in the splicing of *NFIX*, the structure, and function of this gene is described in more detail below.

activate or repress the expression of the same gene or act in the opposite way (Pérez-Casellas et al., 2009).

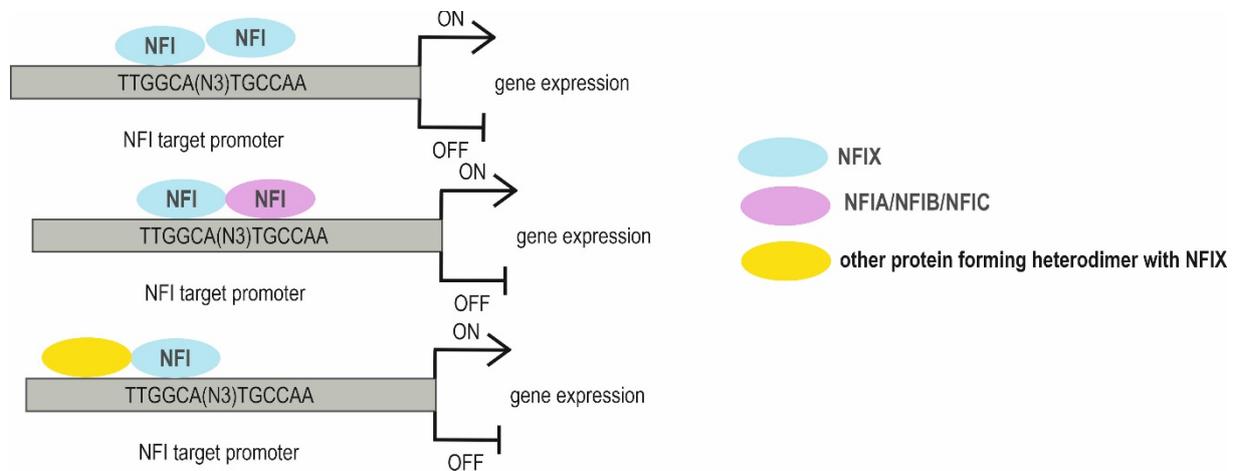


Figure 5. The scheme of a NFI-dependent transcriptional activity.

NFI proteins bind to the same DNA sequence motifs within the promoter of the target gene as homo- or hetero-dimers with other members of NFI family (pink) or different proteins (yellow). After binding they modulate gene expression through either activation or repression of its transcription.

1.2.2 Structure of *NFIX* gene products

NFI proteins are composed of two domains, N-terminal and C-terminal. The first of them is conserved between the four members, mainly encoded by the second exon, which is responsible for DNA binding and dimerization. The stable homo or heterodimers between all of the NFI family members can be formed without leading to the impairment of DNA-binding affinity or specificity (Gronostajski, 2000). The C-terminal domain, encoded by exons 3-11, is obligated to modulate the transcription (Fig. 6). Within this domain there are elements responsible for the activation or repression of transcription. The known activation domain is rich in proline, which stimulates transcription. The interaction with a particular coactivator or corepressor within these elements may lead to a different effect on gene expression. (Gronostajski, 2000; Kruse & Sippel, 1994). The variances between the transactivation domain, resulting from the presence

of alternative splicing exons (Fig. 6), are the reason why four family members demonstrate the different mechanisms of regulation of transcription (Apt et al., 1994).

Moreover, exon 2 of the *NFIX* gene can be a source of circular RNA (circRNA), known as circNFIX. CircRNAs are single-stranded RNAs, which belong to the family of non-coding RNA (ncRNAs) The circulated structure is generated by backsplicing during the maturation of pre-mRNA through joining the 5' and 3' ends of single or multiple exons. Recent studies have shown that circRNAs are engaged in several processes like regulation of transcription, splicing, and translation of small proteins. It was also proved that changes in the expression of this group of RNAs are involved in particular diseases (Greco et al., 2018). The expression of circNFIX is increased in cardiac tissues. It acts as a factor that promotes apoptosis in cardiomyocytes and adverse regulator of cardiac regeneration. These data implicate that the downregulation of this RNA can be a potential therapy for cardiac injury (Cui et al., 2020; Huang et al., 2019).

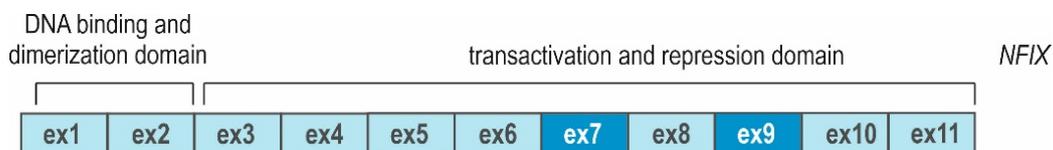


Figure 6. The scheme of a structure of *NFIX* coding sequence.

The exon 1 and 2 encode a part of N-terminal domain responsible for binding to the DNA targets and dimerization; exons from 3 to 11 are part of C-terminal domain, essential for regulation of transcription. The alternative exons are marked with darker blue.

1.2.3 Expression and function of *NFIX* gene

NFI transcription factors are expressed in multiple organs and tissues like the brain, skeletal muscle, lung, liver, intestine, and connective tissue (Gronostajski, 2000). A higher level of this protein was identified in the central nervous system (CNS). The analysis of *NFIX* expression proved that it is expressed in the majority of regions of the developing and adult brain, especially in the cerebral cortex and cerebellum (Piper et al., 2019). Moreover, a higher level of *NFIX* is also identified in skeletal muscles (Messina et al., 2010).

NFI family members are the crucial regulators of cerebral granule neurons development, formation and migration of axons, and dendritogenesis (Wang et al., 2004, 2007). One of the most important events during the development of the nervous system is the differentiation of progenitor cells to the neurons and glia. The abnormalities in this process due to lack of *Nfix* in mice lead to severe consequences (Heng et al., 2014; Piper et al., 2011). NFIX controls the proliferation and migration of the neural progenitor cells using a variety of mechanisms like repression of particular genes, activation of differentiation processes, and even *via* epigenetic controls (Heng et al., 2014; Piper et al., 2011).

NFIX is engaged not only in the development of the brain and spinal cord but also has a crucial role in the development of skeletal muscles (Messina et al., 2010). The occurrence of embryonic fibers is the first wave of myogenesis, the second wave fuses the fetal myoblast and generates secondary fibers (Biressi et al., 2007). The embryonic and fetal myoblast show expression of different specific markers which suggests they are a distinct population with the individual genetic program (Pistocchi et al., 2013). *NFIX* is robustly expressed in the secondary but barely present in the primary myoblasts (Rossi et al., 2016). It was investigated that NFIX activates genes in the development of skeletal muscle in the fetal period but represses genes responsible for myogenesis in the embryonic period (Messina et al., 2010). This proves that NFIX is one of the most important transcription regulators of the switch during muscle formation (Taglietti et al., 2018). The knock-out of *Nfix* in mice induces delayed muscle differentiation and regeneration after injury (Rossi et al., 2016) and triggers the occurrence of muscles characteristic of the embryonic period (Rossi et al., 2017).

Aside from CNS and skeletal muscle development, it was revealed that NFIX is a regulator of mouse hematopoietic stem and progenitor cells. Hematopoietic stem cells with the depleted of *Nfix* (*Nfix*-KO) are not able to persist in bone marrow after transplantation into wild type mice

(Holmfeldt et al., 2013). Recently it was also proven that NFIX is expressed in spermatocytes and is involved in spermatogenesis. The *Nfix*-KO displayed structural defects of spermatocytes and in consequence interrupted spermatogenesis (Davila et al., 2022).

1.2.4 Alternative splicing of NFIX

Depends on tissue and stage of development NFIX protein can occur in different isoforms as a result of alternative splicing of its pre-mRNA. The divergence in the length of the mRNA achieved by the alternative splicing in region downstream exon 2 may lead to differences among the N-terminal domain. So far, at least 3 alternative splicing isoforms of *NFIX* have been discovered in mice and humans (Apt et al., 1994;), (Zhang et al., 2015). The existing isoforms occur by skipping exon 7 or 9, deprived of just exon 7 or exon 9 or both of them (Imbriano & Molinari, 2018). In adult tissue, the isoform with exons 7 and 9 predominates in the brain, in contrast to the muscle where the major mRNA isoform missing these exons (Imbriano & Molinari, 2018; Konieczny et al., 2014; Singh et al., 2011).

1.2.5 Mutations of *NFIX* associated with human diseases

The mutation or aberrant expression of *NFIX* can lead to the disorders (Malan et al., 2010). There are two known syndromes in which *NFIX* is involved. The haploinsufficiency induced by heterozygous point mutation leading to the nonsense-mediated mRNA decay is associated with Malan syndrome. This disorder is characterized by overgrowth, intellectual disability, and macrocephaly (Malan et al., 2010). Importantly, the development of macrocephaly in patients is similar to the development of megalencephaly in mice lacking *Nfix* (Harris et al., 2016). The heterozygous frameshift and splice-site variants mutation occurring downstream of exon 5, does not lead to the degradation of mRNA because the mutant mRNA escape nonsense-mediated decay, but the production of mutant NFIX protein. It is referred to as the second known syndrome called Marshall-Smith (Kooblall et al., 2019). The main symptoms include

advanced bone age, intellectual disability, and difficulties in gain weight (Martinez et al., 2015). (Martinez et al., 2015).

1.2.6 *Nfix* deficient mice

It has been already conducted a variety of experiments on gene deletion in mice to verify the pivotal role of *Nfix* in CNS development (Campbell et al., 2008; Dixon et al., 2013; Driller et al., 2007; Heng et al., 2014, 2015; Piper et al., 2011). The homozygous mice with *Nfix* knockout (*Nfix*-KO) usually die at weaning which makes it impossible to analyze the effects of *Nfix* deletion in adult mice (Harris et al., 2013). Primarily, these mice have problems with eyes opening, (Campbell et al., 2008) and develop brain malformation, spine deformation, hydrocephalus (Driller et al., 2007), and enlarged ventricles (Campbell et al., 2008; Vidovic et al., 2015). Moreover, these mice manifest abnormalities in the formation of the cortex, hippocampus, and dentate gyrus (Campbell et al., 2008; Heng et al., 2014; Piper et al., 2011). Recent studies showed that hydrocephalus is associated with delayed maturation of ependymal cells of the brain in *Nfix*-KO mice. Therefore, these results demonstrate that NFIX is involved in the modulation of ependymal cell adhesion within the lateral ventricles (Harkins et al., 2022). Generally, the knockout of this gene leads to the growth of delayed mice, low weight, and premature death (Campbell et al., 2008; Driller et al., 2007).

Regarding the development of skeletal muscle, the mice with the knockout of *Nfix* are smaller than controls, show disorganization of muscle fibers, and are not able to extend their limbs (Driller et al., 2007; Rossi et al., 2016). These mice also have smaller regenerating myofibers which suggests a shift in balance in the regeneration of skeletal muscle (Rossi et al., 2016, 2017; Taglietti et al., 2018). Moreover, expression of myostatin which is a strong inhibitor of myogenesis is upregulated in these mice, what supports the fact that NFIX directly represses the expression of myostatin (Rossi et al., 2016, 2017).

2. AIMS OF THE STUDY

The main aim of the project was to better understand the mechanisms of gene expression changes in myotonic dystrophy (DM) and to develop a potential therapeutic strategy for DM based on gene therapy tools.

The DM1 is a consequence of the sequestration of MBNL proteins on toxic CUG^{exp} RNA leading to alternative splicing abnormalities of many genes. One of them is gene encoding for NFIX transcription factor, which exon 7 is abnormally included in mRNA in DM muscles. The first hypothesis of the project was that the inclusion of exon 7 affects the transcriptional activity of NFIX. Therefore, the specific aim of the first part of the project was to determine differences in the activity of two NFIX isoforms (+ex7 and –ex7) and define the contribution of missplicing of exon 7 to the gene expression changes observed in the skeletal muscles of DM patients.

One of the potential therapeutic strategies in DM is the rescue of MBNLs from sequestration or overexpression of MBNL, however uncontrolled increase of MBNL activity may have deleterious effect on proper cell function. The second aim of the project was to design and test the autoregulated MBNL1 overexpression system as a potential gene therapy tool for fine-tuning of activity of MBNL pool in DM conditions.

Results relating to the first aim of the project are described in the first part of the dissertation which is in the form of a monograph (Chapter 3). The second part of the thesis is an article by Rogalska & Sobczak published in *Molecular Therapy–Nucleic Acids* in 2022 (Chapter 4).

3. CONTRIBUTION OF SPLICING ABNORMALITIES of *NFIX* EXON 7 ON PATHOMECHANISM OF MYOTONIC DYSTROPHY

3.1 Results and Discussion

The sequestration of MBNLs to the pathogenic CUG repeats is the main reason for alteration in many alternative splicing events, especially in skeletal muscles, the heart and the brain. However, splicing changes caused by MBNLs insufficiency explain only some of the disease symptoms (Ho et al., 2004). It was already shown that the presence of CUG^{exp} induces also changes in gene expression in different animal models (Du et al., 2010; Qawasmi et al., 2019; Todd et al., 2013). DM1 patients demonstrate overall transcriptomic alterations in skeletal muscle (Todorow et al., 2021, E. T. Wang et al., 2019), heart (E. T. Wang et al., 2019), and brain (Otero et al., 2021). The analysis of DM1 patient biopsies revealed that less than 4% of the significantly differentiated genes demonstrated missplicing events (Braun et al., 2022).

Among genes in which significant changes in alternative splicing are observed in DM animal models and DM1 patients are transcription factors. One of them is *NFIX*, the gene involved in myogenesis, in which the abnormal inclusion of exon 7 is detected in DM1, but so far molecular and phenotypic consequences of this abnormality are unknown (Yamashita et al., 2012). This is the main reason why the question regarding contribution of changes in distribution of *NFIX* splicing isoforms on gene expression in DM was raised in this project.

3.1.1. Changes in alternative splicing of *NFIX* exon 7 in skeletal muscles of DM patients

NFIX is expressed in the majority of organs and tissues, but a higher expression is observed in skeletal muscle (Fig. 7a) (Gronostajski, 2000). It consists of 11 exons. Many different mRNA isoforms can be formed due to alternative splicing and alternative transcription initiation. Inclusion of two alternative exons (exon 7 and exon 9) leading to the assembly of three different *NFIX* protein isoforms (Y. Zhou et al., 2014). Other mRNA variants of this gene can produce

protein isoforms differing in a N-terminal sequence due to translation initiation at alternate start codons or a frameshift caused by the exclusion of alternative exons (Fig. 7b).

As mentioned in the Introduction, MBNL proteins negatively regulate alternative splicing of *NFIX* exon 7. In DM1 and DM2, toxic RNAs containing hundreds or thousands of CUG or CCUG repeats, sequester MBNL proteins (Du et al., 2010). As a consequence, the abnormal splicing of hundreds of genes, including the *NFIX* exon 7 is observed. Importantly, there is no changes in expression level of *NFIX* (Fig. 7c). In skeletal muscles of DM1 and DM2 patients the isoform with exon 7 inclusion (**NFIX+7**) predominates, in contrast to the non-DM individuals in which major mRNA isoform lacks exon 7 (**NFIX-7**) (Fig. 7d). Depends on clinical manifestation the missplicing of exon 7 is different in different patients and correlates with disease severity (Nakamori et al., 2013). Similar abnormalities, but with smaller changes in PSI value, were observed in the skeletal muscles of DM1 mouse model, *HSA-LR*, showing consistency with the above results (Fig. 7e). *HSA-LR* is a mouse model expressing the transgene with 220 CTG repeats in 3'UTR of human skeletal actin gene (*HSA*). All these results showed that switching to the **NFIX+7** isoform is characteristic of skeletal muscles expressing expanded CUG^{exp} (Fig. 7d, e).

The splicing pattern of *Nfix* exon 7 also changes during mouse skeletal muscle development, with a similar distribution of both isoforms in the transition from prenatal to the postnatal stage (embryonic day 18.5 and postnatal day 1) and significant enrichment of *Nfix-7* isoform in the adult mouse (12 weeks) (Fig. 7f). It was already demonstrated that isoforms deprived of both alternative exons or just exon 9 predominate in mouse fetal myoblast. However, the isoform lacking exons 7 and 9 is a key regulator of the transcriptional switch from the embryonic to the fetal period during muscle development (Imbriano & Molinari, 2018).

To achieve broader insight into the impact of exon 7 on the transcriptional activity of NFIX in this project two different experimental approaches were used: (1) generation of the stable cell line with functional knockout (KO) of *NFIX* and recovery of this transcription factor activity in a stable cell line expressing either NFIX+7 or NFIX-7 isoform (describe in Subchapter 3.1.2, 3.1.3 and 3.1.4) and (2) generation of human skeletal muscle cell models with enrichment of either NFIX+7 or NFIX-7 isoform using splicing switching antisense oligonucleotides (AONs) blockers which can induce a skipping of the targeted exon 7 (Subchapter 3.1.7).

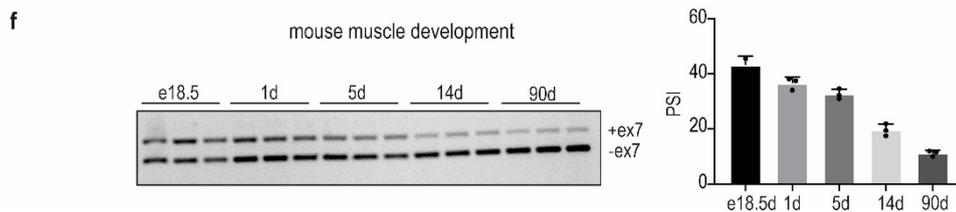
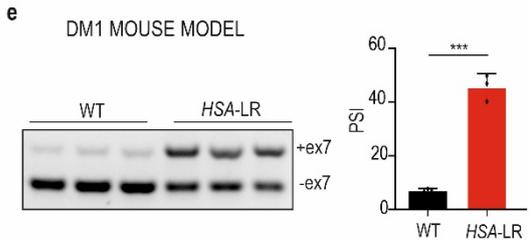
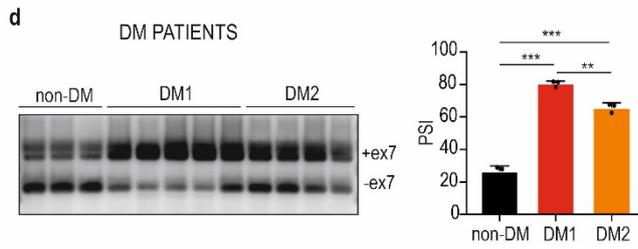
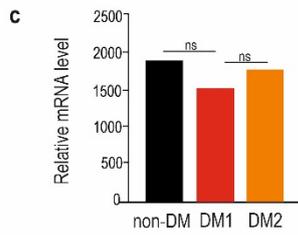
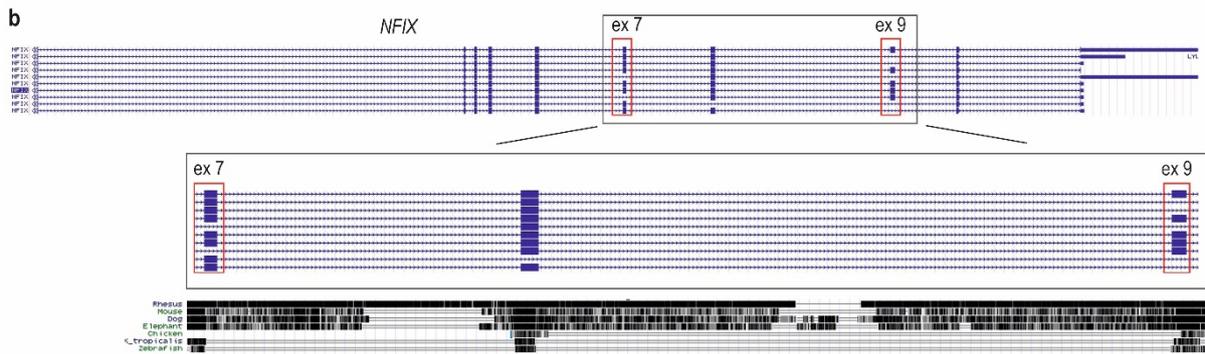
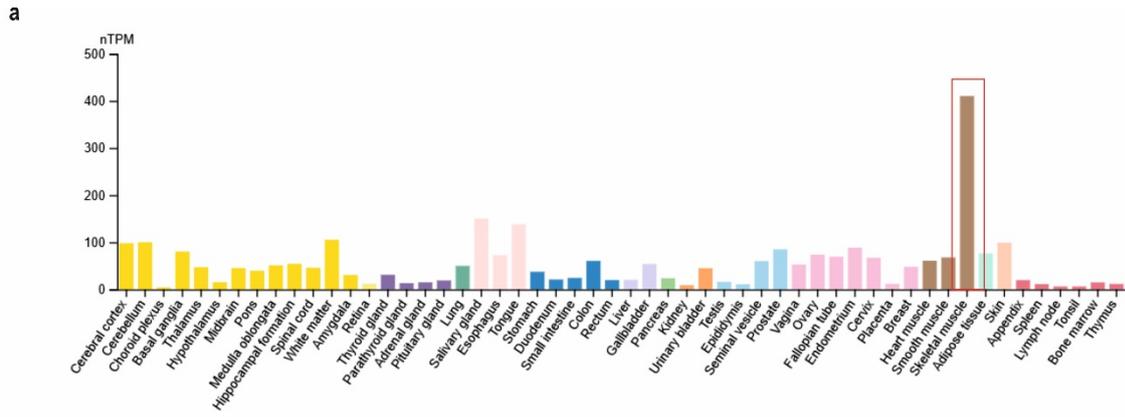


Figure 7. *NFIX* structure and splicing patterns of exon 7.

a) *NFIX* RNA expression pattern presented as tissue specificity (source: <https://www.proteinatlas.org/>).

b) Structure of *NFIX* gene with different possible transcript variants resulting from the changes in transcription start site, splicing of alternative exon 7 and exon 9, or frameshift caused by the exclusion of alternative exon 9. The alternative exons are indicated by red frames (source: <https://www.ncbi.nlm.nih.gov/>).

c) Average *NFIX* expression level in skeletal muscle biopsies from DM patients and healthy individuals from data obtained from whole transcriptomic analysis. Bars represent average expression from 7 independent probes with SD; unpaired Student's *t*-test; ns, non-significant.

d) Alternative splicing profile of *NFIX* exon 7 in skeletal muscles of three normal adults (non-DM), five DM1, and four DM2 patients analyzed by RT-PCR. The percentage of mRNA isoform with the inclusion of alternative exon was calculated using the of percent-spliced-in (PSI) parameter which demonstrates the portion of mRNA with included alternative exon. Bars represent average PSI from 3-5 independent biopsy samples with standard deviation (SD); unpaired Student's *t*-test; ** $P < 0.01$; *** $P < 0.001$.

e) Results of RT-PCR analysis for alternatively spliced *Nfix* exon 7 in skeletal muscle of wild type (WT) and *HSA-LR* mice. Bars represent average PSI from 3 independent mice (dots) with SD; unpaired Student's *t*-test; *** $P < 0.001$.

f) Results of RT-PCR analysis of changes in *Nfix* exon 7 splicing pattern during mouse skeletal muscle development; embryonic day 18.5 (e18.5), postnatal day 1 (1d), day 5 (5d), 2 weeks (2w) and 14 weeks (14w) old mice; bars represent average PSI from three independent experiments with standard deviation (SD).

3.1.2 Generation of stable HEK-293 cell line with a functional knockout of *NFIX* (*NFIX-KO*)

In the first step of the first approach, a knockout of *NFIX* (*NFIX-KO*) in the T-REx™HEK-293 cell line, characterized by the majority of isoforms without exon 7 was generated (Fig. 8a).

This cell line was chosen because of the possibility of further genetic manipulation due to the presence of sites for homologous recombination in the cell genome. The generation of the *NFIX-KO* cell line was an essential step in further understanding the contribution of *NFIX* activity to specific cellular pathways.

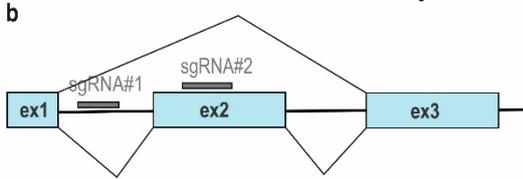
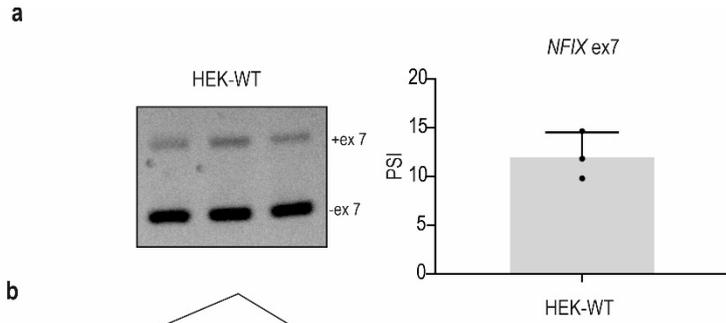
The *NFIX*-KO cell line was obtained using CRISPR-Cas9 technology. At first, the essential gene fragment for the NFIX proper functioning was chosen to induce *NFIX* deactivation. This is the exon 2 sequence encoding a domain responsible for binding to DNA. Designed two sgRNAs directed Cas9 protein to the flanking sequence upstream of and within the exon 2, which in turn caused double-strand breaks (DSB) and delete part of the exon 2 from the *NFIX* gene (Fig. 8b). The removal of this exon leads to an open reading frame shift and termination of translation due to the appearance of premature STOP codon in exon 4. Even if the protein is formed, is in the truncated version and lacks the ability for DNA binding and therefore, is deprived of the regulation of transcription activity.

Genotyping of clonal cell line obtained after treatment with designed gRNAs and Cas9 showed deletion of the part of the exon 2 (Fig. 8c). Results of RT-PCR assay demonstrated exclusion of exon 2 resulting from the lack of 3'-splice site in intron 1 (Fig. 8d). The immunofluorescence (Fig. 8e) and western blot (Fig. 8f) analysis showed no signal from NFIX in *NFIX*-KO cells what suggest a functional knockout of this gene.

To further support obtained results, expression of a few known targets of NFIX transcription factor: *NCAMI* (Heng et al., 2014), *SOX3* (Vidovic et al., 2015), and *IGFBP5* (Pérez-Casellas et al., 2009) was tested. Real-time RT-qPCR analysis confirmed that deletion of part of *NFIX* exon 2 leads to the downregulation of *IGFBP5* and upregulation of *NCAMI* and *SOX3* (Fig. 8g left panel). The same effect was observed for the expression of *Igfbp5* in skeletal muscles and brain of the *Nfix*-KO mouse model (Fig. 8g; right panel), which was obtained also by exon 2 excision and in consequence disruption of *Nfix* function (Campbell et al., 2008).

Moreover, expression of circRNA being a product of backsplicing of *NFIX* exon 2 (*circNFIX*) (Fig. 9a) was observed in neither *NFIX*-KO cells (Fig. 9b) nor *Nfix*-KO mice (Fig. 9c). The sequence of circNFIX is firmly conserved among humans and mice (Huang et al., 2019).

Results of this part of the project confirmed the functional knockout of *NFIX* in T-Rex™-293 cell: NFIX protein was not detected (Fig. 8e, f), expression of NFIX targets and the level of circ*NFIX* changed significantly (Fig. 8g and 9). These changes are similar to those observed for previously generated *Nfix*-KO mouse model (Huang et al., 2019).



ttgtgggggaggaggaggaggagaaggcggtttccttgccccccctctaacgctgttttcccttctctttccccctcatcccgtctccccctcccgtccccctgccccgcagatgctccggct
 tgccgcctgcagGATGAGTTCCACCCGTTTCATCGAGGCACTGCTGCCTCACGTCCGCGCTTTCTCCTACACCTGGTTCAACCTGCAGGCGCGGA
 AGCGCAAGTACTTCAAGAAGCATGAAAAGCGGATGTCGAAGGACGAGGAGCGGGCGGTGAAGGACGAGCTGCTGGGCGAGAAGCCCGA
 GATCAAGCAGAAGTGGGCATCCCGGCTGCTGGCCAAGCTGCGCAAGGACATCCGGCCCCGAGTTCCGCGAGGACTTCGTGCTGACCATCA
 CGGGCAAGAAGCCCCCTGCTGCGTGCTCTCCAACCCCGACCAGAAGGGCAAGATCCGCGGATTGACTGCCTGCGCCAGGCTGACAA
 GGTGTGGCGGCTGGACCTGGTCATGGTGATTTGTTTAAAGGGATCCCCCTGAAAGTACTGATGGGGAGCGGCTCTACAAGTCGCCTCA
 GTGCTCGAACCCCGGCTGTGCGTCCAGCCACATCACATTGGAGTCACAATCAAAGAAGTGGATCTTATCTGGCTTACTTTGTCCACACTC
 CGGtaggtcgttcaaccattttccctctcattttatttctgtctggcattgttctgtttatgttctctaattccaagcgataactcgccatggcctaactggtgatgcc

intron sequences

PAM sequence

sgRNA sequences

exon 2 sequence

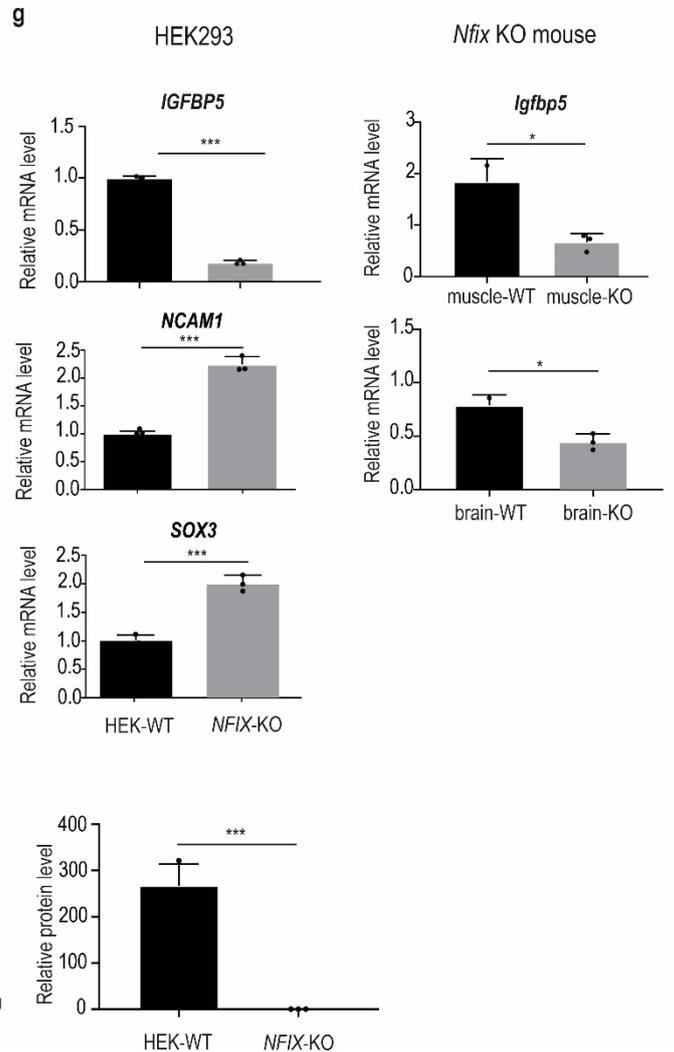
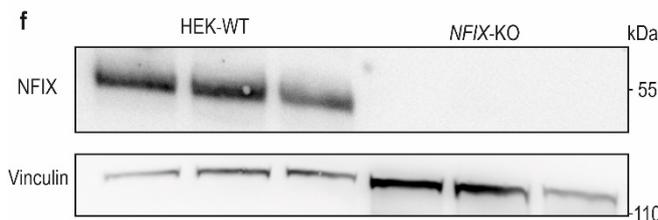
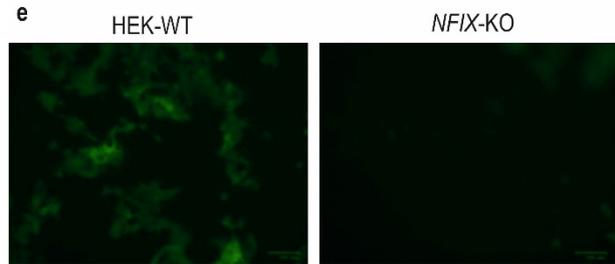
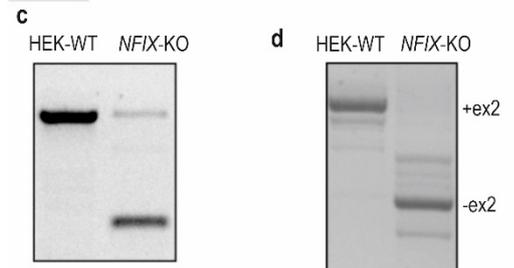


Figure 8. Functional knockout of *NFIX* in T-RexTM-293 cell line.

a) Results of RT-PCR analysis of the splicing pattern of *NFIX* exon 7 in T-RexTM-293 cells. Bar represents the average PSI from three independent experiments with standard deviation (SD).

b) The scheme and the sequence of a fragment of the *NFIX* gene showing the binding sites of designed sgRNAs. The first one was complementary to the sequence upstream of exon 2, and the second one was within exon 2. The expected endonuclease cleavage site is located 3 nucleotides upstream PAM (Protospacer Adaptive Motif) sequence. Exon 2 is a long sequence of 532 nucleotides (blue) encoding for more than 177 amino acids. Deletion of this exon is out-of-frame mutation.

c) Results of genotyping showed efficient deletion of the DNA sequence containing fragment of intron 1/exon 2 of *NFIX* and **d)** results of RT-PCR analysis showed exclusion of exon 2 from *NFIX*. WT, wild-type cells.

e) Representative fluorescence microscopy images of immunofluorescence analysis of WT and *NFIX*-KO T-RexTM-293 cells using primary anti-NFIX and secondary FITC-labelled anti-rabbit antibodies (green); scale bar=100µm.

f) Results of western blot for NFIX protein (anti-NFIX antibody). The results are averages from n=3 independent experiments normalized to Vinculin; unpaired Student's t-test; *** P < 0.001.

g) Results of RT-qPCR analysis of expression level of genes positively or negatively regulated by NFIX in T-RexTM-293 cells (left panel) and in *Nfix*-KO mouse (right panel). Bars represent the average from n=3 independent experiments (dots) for each experimental condition with SD normalized to *ACTB*; unpaired Student's t-test: *, P < 0.05; ***, P < 0.001.

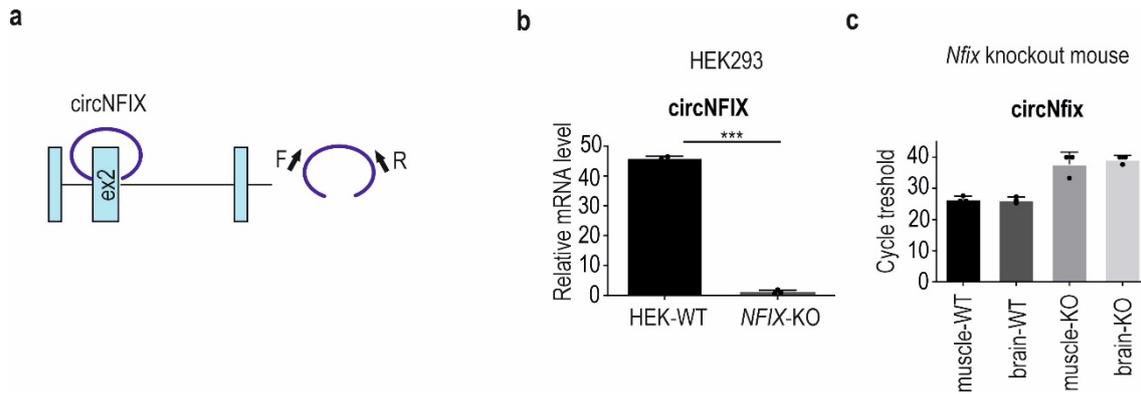


Figure 9. Effect of *NFIX*-KO on the circ*NFIX*.

a) The scheme of circ*NFIX*, which is a product of backsplicing of *NFIX* ex2, and the arrangement of forward and reverse primers used to amplify this sequence.

b) Results of RT-qPCR analysis in WT and *NFIX*-KO T-RexTM-293 cells revealed significant decrease of circ*NFIX* in KO cells. Bars represent the average from n=3 independent experiments (dots) for each experimental condition with SD normalized to *ACTB*; unpaired Student's *t*-test: ***, $P < 0.001$.

c) Results of cycle threshold (Ct) revealed much higher value of Ct for tissues from *Nfix*-KO compare to wild type mice. Bars represent average from n=3 independent experiments (dots) for each experimental condition with SD normalized to *ACTB*.

3.1.3 Generation of genetic constructs for overexpression of two splicing isoforms of *NFIX*

For the overexpression of *NFIX*+7 or *NFIX*-7 isoforms, two different genetic constructs were generated. Each *NFIX* isoform sequence was cloned to the pcDNATM5/FRT/TO expression vector which allows for inducible exogen expression. The hybrid human cytomegalovirus/Tet operator 2 (CMV)/TetO2 promoter present in this vector is activated by tetracycline/doxycycline.

Next, the transcriptional activity of two isoforms was evaluated. T-RExTM-293 cells were co-transfected with different amounts of a construct encoding one of two *NFIX* isoforms and with a plasmid expressing luciferase reporter gene under the control of *NFIX*-dependent promoter

fragment, coming from *IGFBP5*. Hence, the activity of luciferase corresponds to the NFIX transcriptional activity. Indeed, increasing the dose of constructs overexpressing NFIX+7 or NFIX-7 lead to an increase in the activity of luciferase, which suggests that both isoforms activated the *IGFBP5* promoter; however, NFIX-7 showed significantly higher activity (Fig. 10). The obtained result was an encouragement to further examine differences in transactivation potential of two NFIX splicing isoforms.

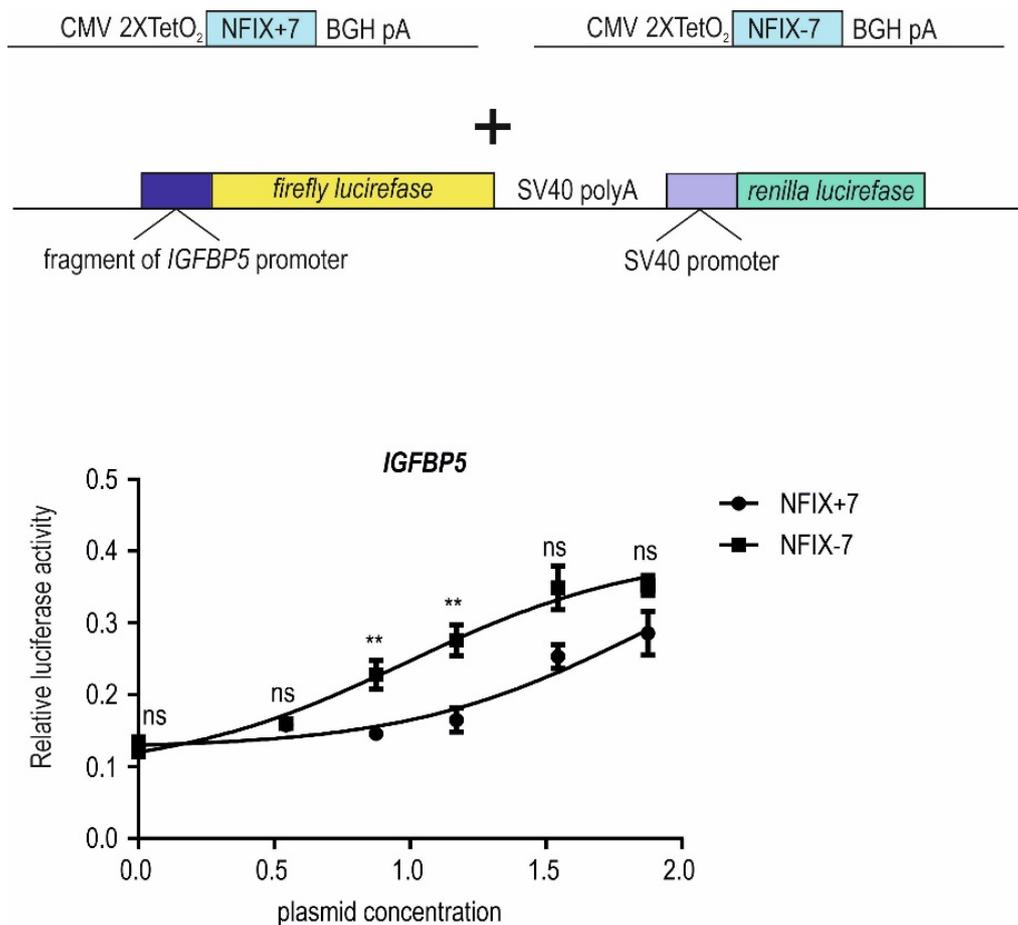


Figure 10. Differences in transcriptional activity of two NFIX isoforms.

T-REx™-293 cells were co-transfected with a distinct concentration of pcDNA™5/FRT/TO with *NFIX+7* or *NFIX-7* (ng/ml; log scale), a plasmid expressing firefly luciferase gene under the control of fragment of *IGFBP5* promoter, and renilla luciferase sequence used for normalization. Lines represent average of luciferase activity (arbitrary units; au) from n=3 independent experiments for each group normalized to renilla luciferase; unpaired Student's t-test: **, P < 0.01; ns, non-significant.

3.1.4 Generation of two stable cell lines expressing NFIX+7 or NFIX-7 using Flp-In™ T-REx™ system

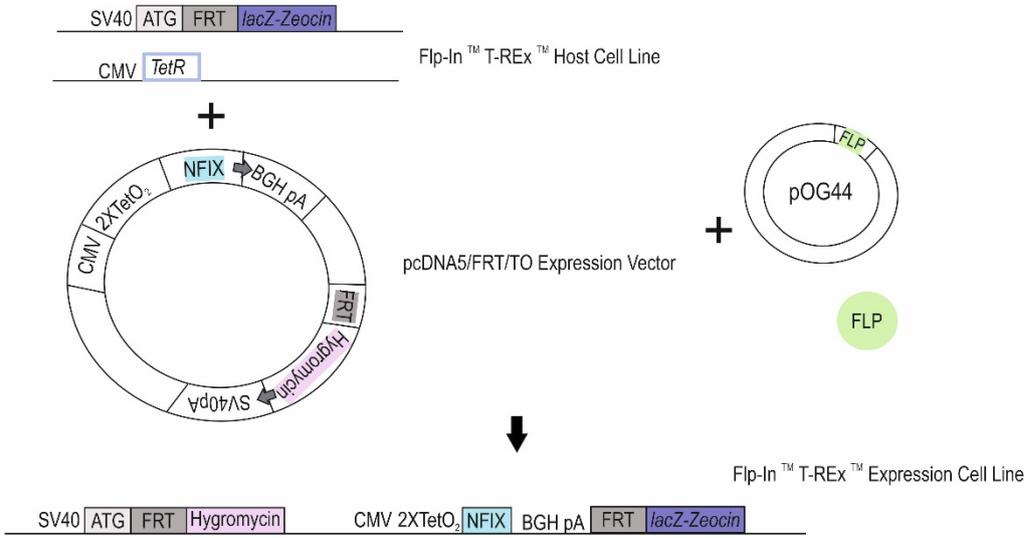
To test the hypothesis that sequence encoded by exon 7 has the potential impact on the transcriptional activity of NFIX, two stable cell lines expressing a particular splicing isoform were generated in the background of the homozygotic *NFIX*-KO. Cell lines were generated by using site-specific recombination between two FLP recombinase target (FRT) sites, which are present in a single copy in the genome of T-REx™-293 cells and in the pcDNA™5/FRT/TO expression plasmid. This recombination is mediated by Flp recombinase. The cells were co-transfected with pcDNA™5/FRT/TO encoding *NFIX+7* or *NFIX-7* and with the pOG44, the Flp recombinase expressing plasmid. The selection of positive clones was achieved with the use of hygromycin, as the plasmid for NFIX overexpression contains the hygromycin resistance gene sequence that is deprived of START codon and promoter. Thus, the cells acquire long-term resistance to a certain antibiotic only if the plasmid is integrated into the selected site within a genome of cells and in this way obtain START codon (Fig. 11a). The expression of the gene of interest is possible in this system through the presence of TetO₂ sequence inserted into the CMV promoter, which is the binding site for Tet repressor. Tet repressor is stably expressed by T-REx™-293 host cells. The binding of Tet repressor to the promoter represses the transcription of transgene. The doxycycline (alternative of tetracycline) added to the cells (DOX- treated cell) binds with high affinity to the Tet repressor leading to the induction of transgene expression (Fig. 11b).

The generation of these cell lines was supported by the fact that a stable cell line is genetically homogenous because of the obtainment population from a single clone and the expression level of the transgene will be regulated due to the application of tetracycline sensitive promoter. The western blot analysis was performed to confirm the expression of *NFIX+7* and *NFIX-7* upon treatment with doxycycline (analog of tetracycline), which is an inductor of transgene

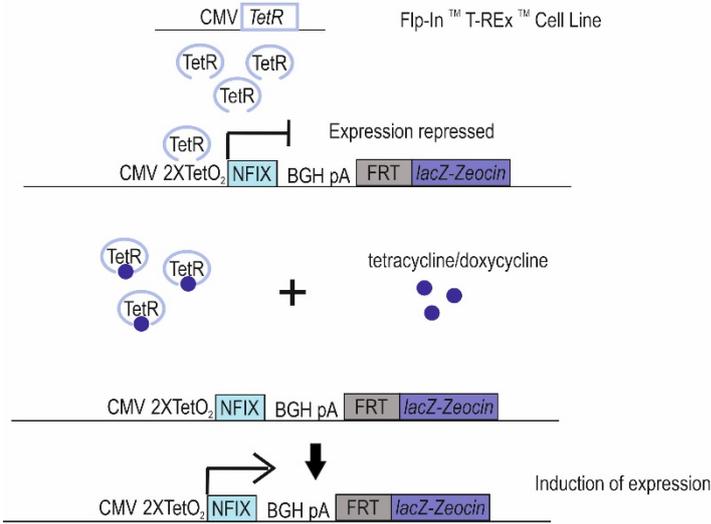
expression. The level of NFIX is sensitive to doxycycline concentration (Fig. 11c). In the absence of doxycycline in the culture medium it was no expression of the transgene, the dose of 250 ng/ml slightly increased amount of NFIX, and the last used dose (500 ng/ml) led to the highest level of both NFIX+7 and NFIX-7 (Fig. 11c).

To support the results of the western blot, the expression of *IGFBP5*, well known NFIX target, was tested. Real-time RT-PCR analysis confirmed that both isoforms of NFIX caused the increase of *IGFBP5* expression to the level observed in WT background cells (Fig. 11d).

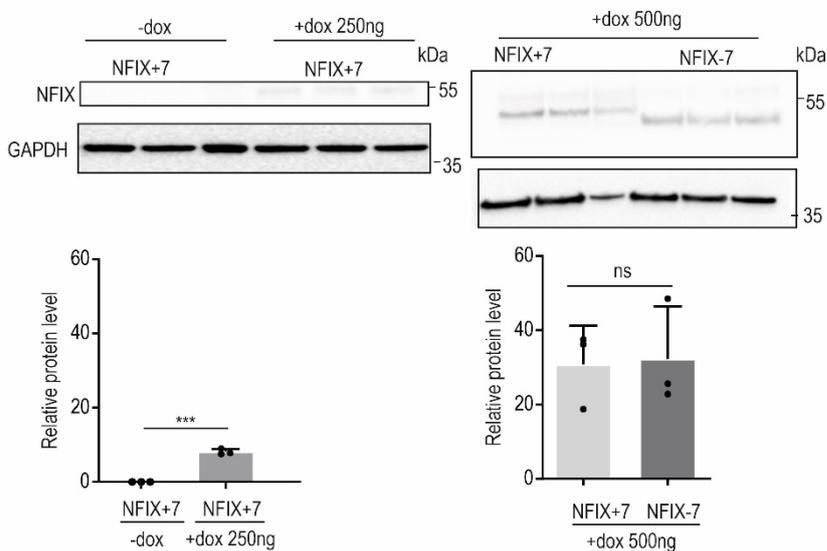
a



b



c



d

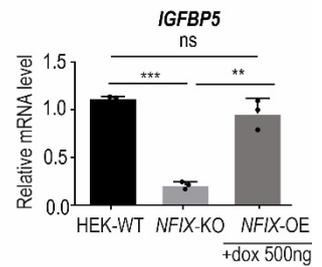


Figure 11. Characteristics of stable cell line expressing NFIX+7 or NFIX-7 isoform.

a) Diagram of Flp-In™ T-REx™ system. The host cells containing FRT site are co-transfected with expression constructs pcDNA™5/FRT/TO *NFIX+7* or *NFIX-7* and with pOG44 plasmids. Incorporation of the expression construct into the FRT locus leads to the acquisition of hygromycin resistance by cells due to the proper organization of open reading frame for hygromycin resistant gene (located in construct) in a context of the START codon and promoter sequence (located in host genome).

b) Diagram of Flp-In™ T-REx™ system showing chemical induction of transgene expression. TetR is a protein constitutively expressed in Flp-In™ T-REx™ cells. It binds to the TetO sequence and represses the transgene transcription. Tetracycline (or alternative doxycycline) if added to the culture medium binds TetR, which leads to its dissociation from TetO and activation of transgene transcription.

c) Results of western blot analysis of two NFIX splicing isoforms in stable T-REx™-293 cell line before and after treatment with doxycycline in two different concentrations: 250 ng/ml and 500 ng/ml. Anti-FLAG antibody staining was carried out as FLAG sequence is fused to the C-terminal end of NFIX in both constructs. Bars represent the average signal for n=3 independent experiments normalized to GAPDH; unpaired Student's t-test; *P<0.05; ** P<0.01; *** P<0.001; ns, non-significant.

d) Results of RT-qPCR analysis of the level of *IGFBP5* mRNA in three T-REx™-293 cell lines (WT, *NFIX-KO*, and with stable NFIX overexpression, after treatment with doxycycline in concentrations 500 ng/ml). Bars represent average from n=3 independent experiments (dots) for each experimental condition with SD normalized to *TOP1*; unpaired Student's t-test; ** P<0.01; *** P<0.001; ns, non-significant.

3.1.5 Expression of other members of NFI family is not affected in *NFIX-KO* and either *NFIX* isoforms expressing HEK293 cells

Recent research has shown that transcription factors from the NFI family have overlapping functions and can regulate mutual targets in the developing nervous system (Fraser et al., 2020). Therefore, it was crucial to see whether the expression of other members of the NFI family can be affected by *NFIX* knockout or overexpression (Fig. 12). The results of RT-PCR assays showed no changes in either *NFIX-KO* or *NFIX+7* T-REx™-293 cells (Fig. 12a). Moreover, in muscles of *Nfix-KO* the level of expression of *Nfia*, *Nfib*, and *Nfic* is also unaffected (Fig. 12b).

All these data suggest that changes in the expression of NFIX target genes are not a result of the differences in the activity of other NFI factors in generated cell models.

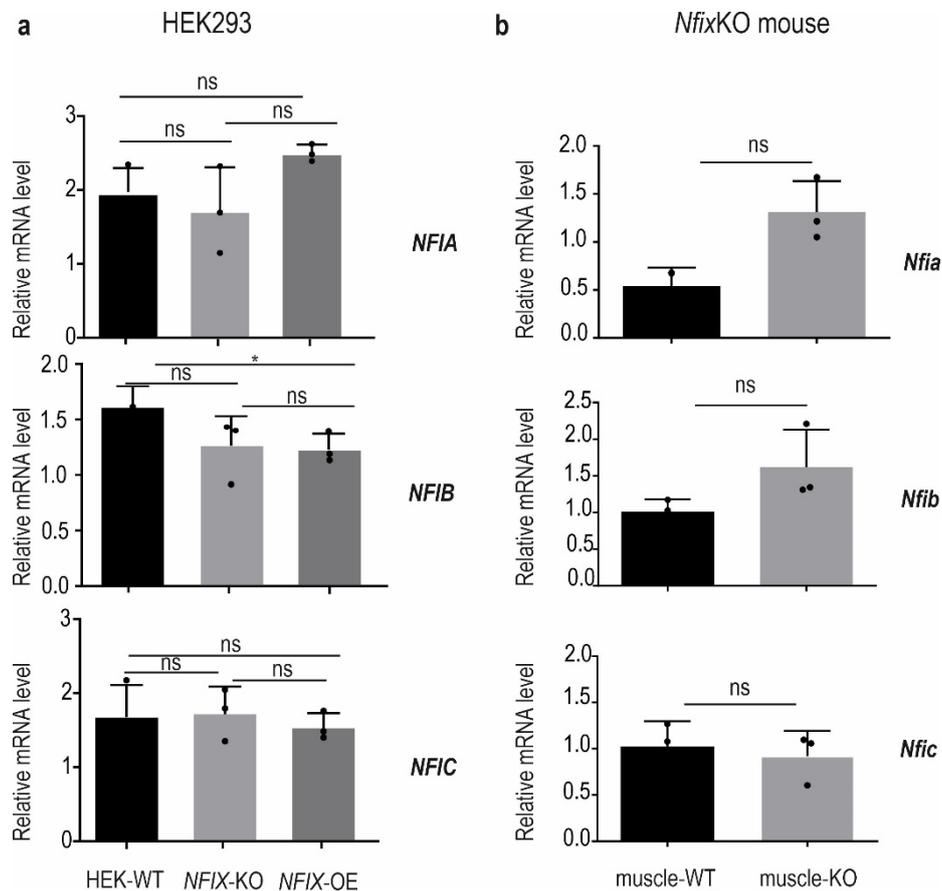


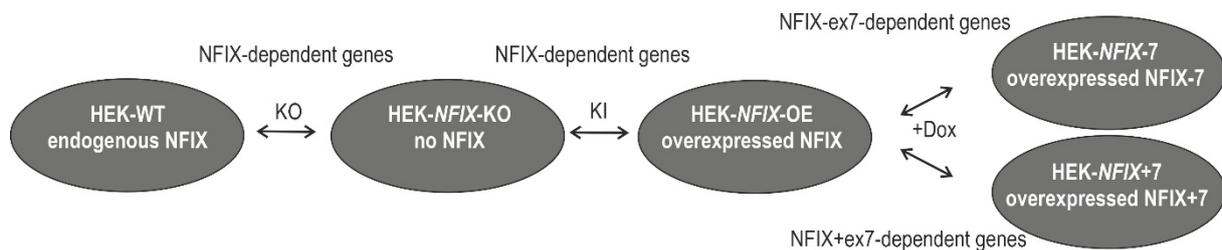
Figure 12. The expression level of NFI family members in generated HEK293 cell models.

a) Results of RT-qPCR analysis of expression level of *NFIA*, *NFIB*, and *NFIC* in stable T-REx™-293 cell lines. Bars represent the average signal from n=4 for *NFIA* and *NFIB* and n=3 for *NFIC* independent experiments for each group normalized to *ACTB*.

b) Results of RT-qPCR analysis of expression level of *Nfia*, *Nfib*, and *Nfic* in *Nfix*-KO mouse tissue. Bars represent the average signal from n=3 independent experiments for each group normalized to *Gapdh*; **a** and **b** unpaired Student's *t*-test: *, $P < 0.05$ ns, non-significant.

3.1.6 NFIX regulates the expression of genes involved in RNA metabolism in generated HEK293 cells

To identify genes sensitive to the level of NFIX the whole transcriptome-based RNA sequencing (RNA-seq) was performed for WT, *NFIX*-KO cells, cells without induction (*NFIX*-OE), and cells with induced expression of NFIX isoforms (*NFIX*+7 and *NFIX*-7). The results of RNA-seq for both *NFIX* transgene expression demonstrate a small, but statistically sufficient, increase of mRNA in samples before doxycycline treatment and a much higher increase after doxycycline treatment support transgene expression (Fig. 13a). The increase of mRNA before adding the doxycycline may be a reason for leaky gene expression which can occur even without induction of transcription. This is the reason why the samples in further comparisons are not divided into before and after doxycycline treatment (*NFIX*-OE).



For the first comparison, *NFIX*-KO vs WT, differential expression was found for 5,020 genes and for the second, *NFIX*-KO vs *NFIX*+/-7 (*NFIX*-OE), 203 genes (if $P_{adj} < 0.01$) or 1049 (if $P_{adj} < 0.05$) from total 12,676 genes with average expression cutoffs above 1 Counts per Million (CPM) (Fig. 13b). Differences that were common for both comparisons revealed a list of 157 (or 777 if less stringent parameters were used) genes sensitive to NFIX level (Fig. 13c). The number of genes which expression is changed in *NFIX*-OE compare to *NFIX*-KO samples suggesting that promoter leakage is observed which means that transcription from the inactive promoter, without inductor (doxycycline) addition occur. Importantly, the expression of almost all these genes changed in opposite directions in *NFIX*-KO and *NFIX*-OE cells. In this

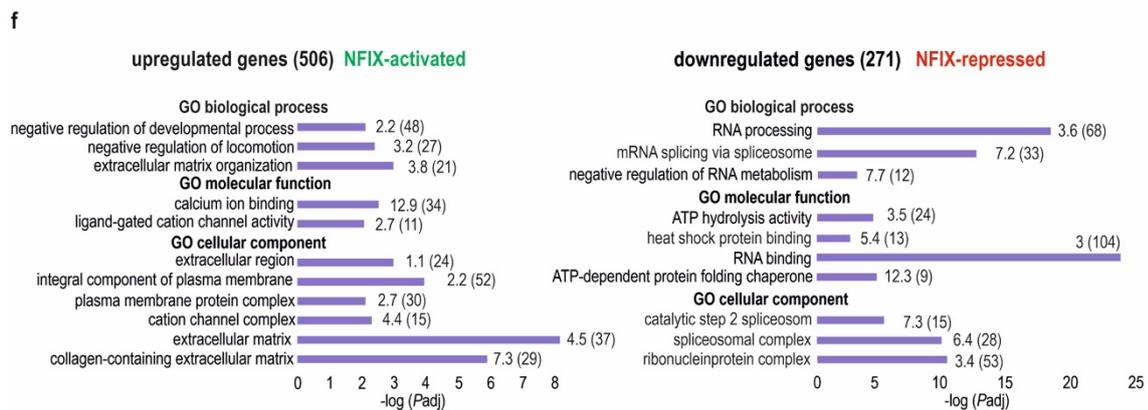
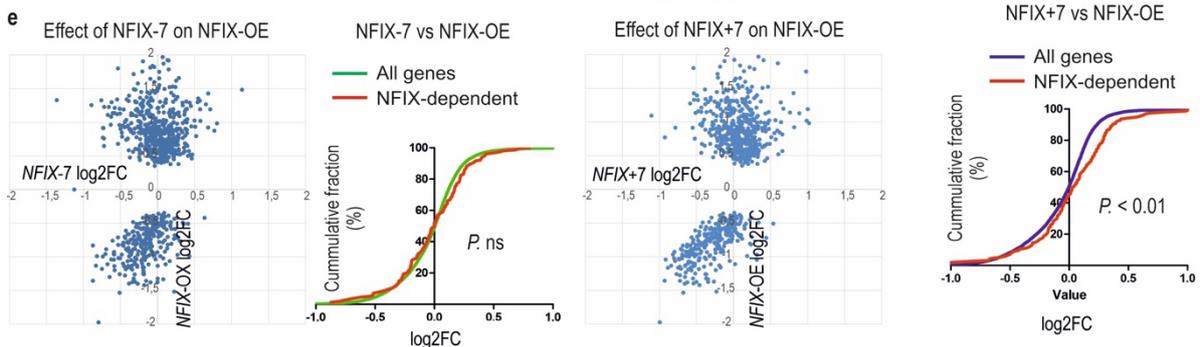
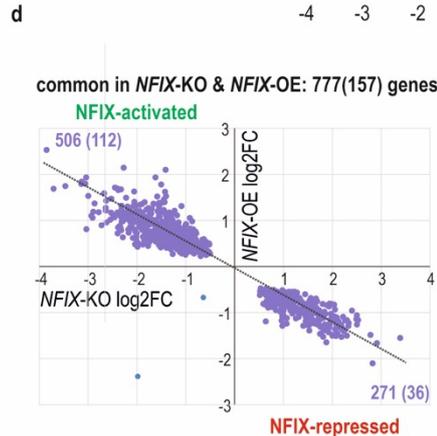
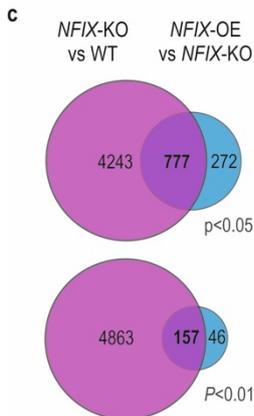
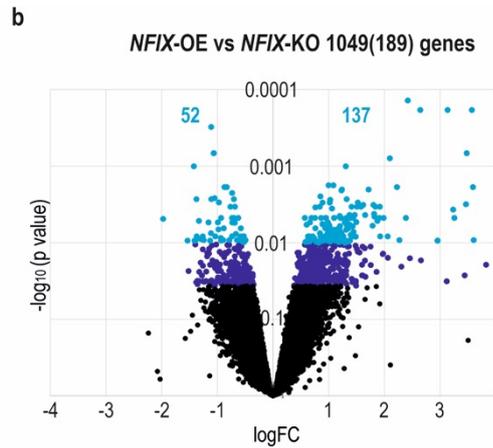
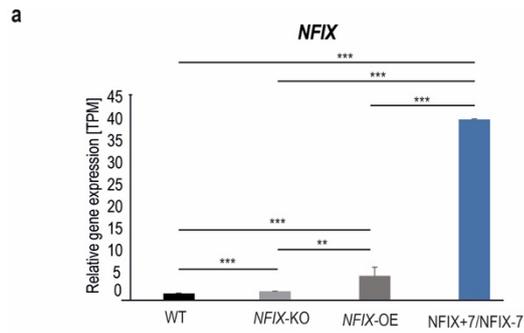
population, the genes negatively regulated in *NFIX*-KO and positively regulated in *NFIX*-OE cells predominated, which suggests that *NFIX* is mostly a positive regulator of gene expression (Fig. 13d).

Unexpectedly, however, just a few significant differences were identified between cells overexpressing different splicing isoforms, either *NFIX*+7 or *NFIX*-7, in doxycycline-treated cells. The vast majority of *NFIX*-sensitive genes were changed for DOX-treated cells overexpressing either *NFIX*+7 or *NFIX*-7 isoform compared to control cells that were not treated with DOX. The comparative analysis of *NFIX*+7 or *NFIX*-7 and *NFIX*-OE (cells not treated with doxycycline) shows that the addition of an inductor increases the activation or repression of particular genes for one or both isoforms. In these analyzed groups, most genes do not respond to the increasing level of doxycycline, thus their expression levels do not change upon the increase of *NFIX* level (Fig. 13e). In the observed increase of expression after doxycycline treatment for a group of 157 *NFIX*-sensitive genes, *NFIX*+7 seems to act as a more active isoform than *NFIX*-7 (Fig. 13e; plots with a cumulative fraction). These may suggest that these two isoforms differ in transcriptional activity; however, the observation needs to be confirmed in a study using another cellular model, which is described below.

To classify functional classes of identified 777 *NFIX*-sensitive genes to the Gene Ontology (GO) terms, the analysis powered by PANTHER (v17.0) was performed. To this analysis the greater group of genes, with less stringent parameters were used. The *NFIX*-sensitive genes were divided into two groups: upregulated (506) and downregulated (271) by *NFIX* ($P_{adj} < 0.05$). For genes upregulated by *NFIX*, the analysis revealed significant enrichment for the class of genes involved in, among others, the extracellular matrix ($P=2.14\cdot 10^{-9}$; fold enrichment 4.5) (Fig. 13f). On the other hand for genes downregulated by *NFIX*, GO-analysis revealed significant enrichment in the category, among others, RNA binding ($P=5\cdot 10^{-22}$; fold enrichment

2.99), of which mostly includes genes involved in RNA metabolism and RNA processing ($P=5E-17$; fold enrichment 3.64 (Fig. 13f).

To further confirm if the NFIX is engaged in the regulation of RNA metabolism genes a real-time RT-PCR analysis was performed to validate expression changes of selected genes. The majority (5/6) of the analyzed genes encoding for RNA binding proteins showed an increase in expression in *NFIX*-KO cells compared to WT cells, which confirmed the results of RNA-seq experiments. Accordingly, the rescue to the level observed in WT cells was observed for all tested genes in samples with NFIX overexpression, indicating negative regulation of these genes by NFIX (Fig. 13g; left panel). Additionally, the expression level of the same genes in *Nfix*-KO mice (Fig. 13g right panel) was checked. The analysis proved that lack of NFIX leads to the downregulation of these targets also in mouse tissue. All these analyses revealed a new function of NFIX. This transcription factor is a negative regulator of a set of genes involved in the regulation of RNA metabolism.



g

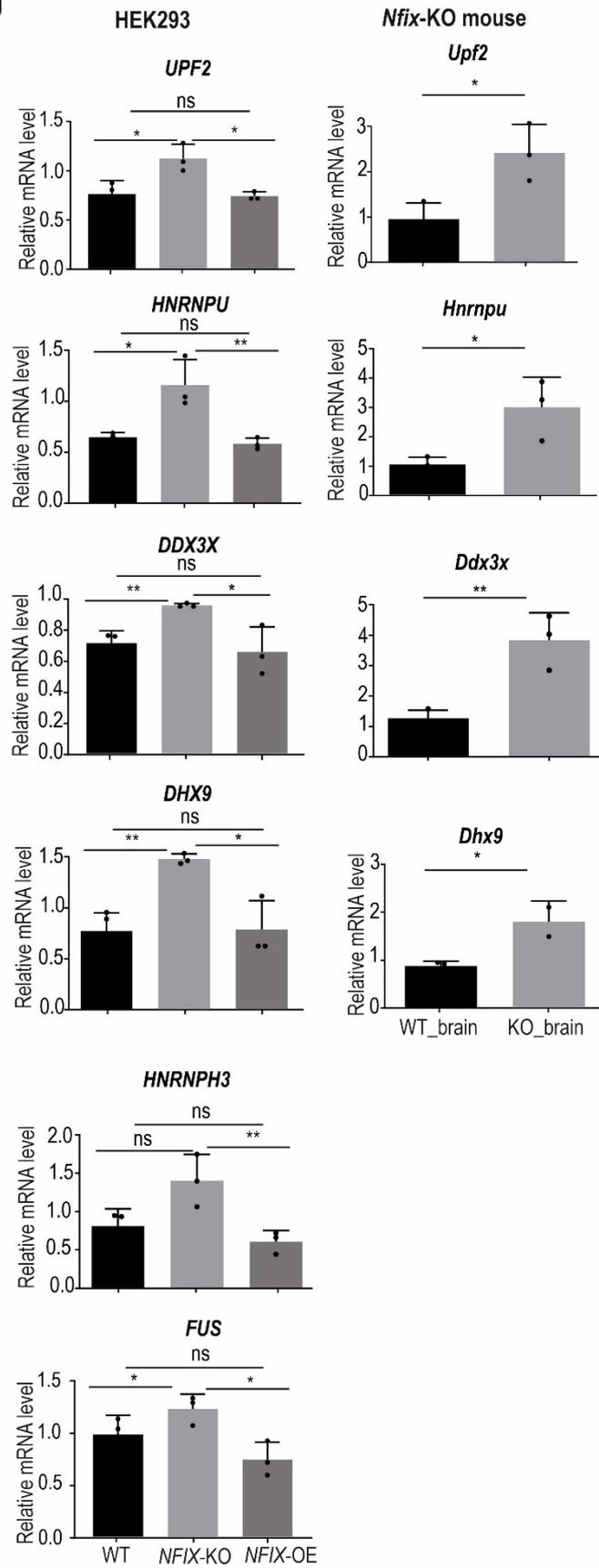


Figure 13. Identification of genes regulated by NFIX in HEK-293 cells.

a) Gene expression levels (TPM) from results of RNA-seq for WT, *NFIX*-KO cells, and cells before induction with doxycycline (*NFIX*-OE) and after doxycycline treatment with induced expression of NFIX isoforms (*NFIX*+7 and *NFIX*-7). Bars represent average expression from 3 replicates with SD; unpaired Student's *t*-test; ** $P < 0.01$; *** $P < 0.001$.

b) The volcano plot showing differentially expressed genes with statistical significance (blue) in comparison of *NFIX*-KO vs *NFIX*-OE.

c) Venn diagram showing overlapping genes (157 or 777 depends on threshold of P_{adj} value, either 0.01 or 0.05 respectively) which are significantly misregulated in both the absence of *NFIX* (*NFIX*-KO) and *NFIX* overexpression (*NFIX*-OE).

d) XY-plot of gene expression changes in both *NFIX*-KO and *NFIX*-OE (cells with basal overexpression of NFIX) with the indicated number of upregulated and downregulated genes. The genes with the higher fold change (above 2-fold) are indicated in parenthesis.

e) XY-plot of gene expression changes observed in Dox treated *NFIX*-7 or *NFIX*+7 cells (high expression of either isoform) and *NFIX*-OE (cells with basal expression of NFIX; without addition of doxycycline, an activator of transgene expression). It shows increasing activity (both activation and repression) of overexpressed isoforms on expression pattern of genes sensitive to NFIX. Most of these genes remain on unchanged level before and after doxycycline treatment. On a side of each plot the juxtaposition of expression changes for either all genes or selected 157 *NFIX*-sensitive genes in Dox treated *NFIX*-7 or *NFIX*+7 cells is presented. Significant increase of *NFIX*-sensitive genes after doxycycline treatment in *NFIX*+7 cells suggests that *NFIX*+7 isoform is a stronger transactivator. The graph shows a statistically significant difference *** $P < 0.001$, based on Mann-Whitney test.

f) Results of Gene Ontology enrichment analysis for 777 *NFIX*-sensitive genes common for both comparisons (*NFIX*-KO vs WT and *NFIX*-KO vs *NFIX*-OE). The selected terms of GO Biological process, GO Molecular function, GO cellular component for upregulated genes (410) and downregulated genes (273) are shown due to the high P value. The numbers behind the bars indicate enrichment (fold change), and the numbers in the parentheses indicate the number of genes for each term.

g) Results of real-time RT-PCR analysis of expression of six chosen genes coding for RNA binding proteins in stable T-REx™-293 cell lines (left panel) and four genes in *Nfix*-KO brain (right panel). Bars represent the average from $n=3$ (with the exception of $n=2$ for *Dhx9* KO_brain) independent experiments for each group normalized to *ACTB* in T-REx™-293 cells and to *Gapdh* in *Nfix*-KO brain; unpaired Student's *t*-test; * $P < 0.05$; ** $P < 0.01$; ns, non-significant.

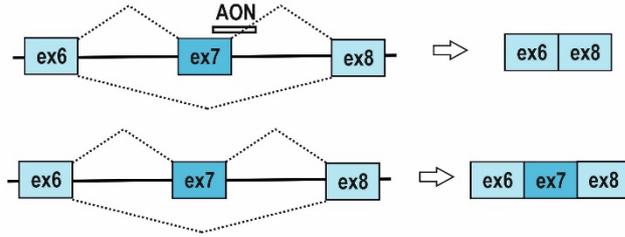
3.1.7 Generation of human skeletal muscle cell models with enrichment of either NFIX+7 or NFIX-7 isoform

The second approach to study potential differences in gene expression regulation by two splicing isoforms of NFIX (NFIX+7 and NFIX-7) was based on the establishment of muscle cell models through manipulation of splicing exon 7 of endogenous *NFIX*. Muscle cells (differentiated and undifferentiated myoblasts) were chosen. The choice of muscle cells is supported by the fact that NFIX is engaged in muscle development and the other partners which interact with NFIX are present at a sufficient level (Messina et al., 2010).

Cells were treated with antisense oligonucleotides (AONs) blockers which are complementary to the 5'-splice sites of alternative exon 7 and which can induce a skipping of targeted exon (Fig. 14a). The designed AON has 13 residues with a mixture of locked nucleic acid (LNA) and DNA residues arranged alternately. The scramble AON with the same chemical modifications was used as a negative control. This approach allowed us to achieve cell models expressing mostly NFIX+7 if treated with control AON or mostly NFIX-7 if treated with NFIX-specific AON. The splice switch activity of the used AON was specified for 5 different cell lines. Three of them were muscle cells: undifferentiated human skeletal myoblast (HSkM), HSkM after 4 days of differentiation into myotubes, and mouse C₂C₁₂ myoblasts. Moreover, two mouse embryonic fibroblasts (MEFs) with different statuses of expression of MBNL genes were also tested: MEFs isolated from wild type animals and mice with full knockout of both *Mbnl1* and *Mbnl2* (MEF-1&2KO). RT-PCR analyses showed that this AON changes the NFIX isoform distribution, *via* efficient exclusion of exon 7 in each tested cell line (Fig. 14b, c). The desired effect was achieved even in the cells with a similar distribution of both splicing isoforms (MEF WT) and in the cells where isoform with exon 7 predominates (HSkM, MEF-1&2KO). Even in the cell line where isoform without exon 7 predominates (C₂C₁₂), the small effect of exon 7 skipping was observed (Fig. 14c). In further experiments, HSkM cells were used because

it allows us to achieve conditions similar to skeletal muscles where NFIX plays a crucial role in the development processes. Moreover, this cell line is characterized by expression mostly NFIX+7. The deltaPSI, the parameter which measures the difference between the PSIs of two compared groups, was higher for HSkM after 4 days of differentiation (deltaPSI=48) than for nondifferentiated cells (deltaPSI=42). This was another strong argument in favor to use this cell line in further studies.

a



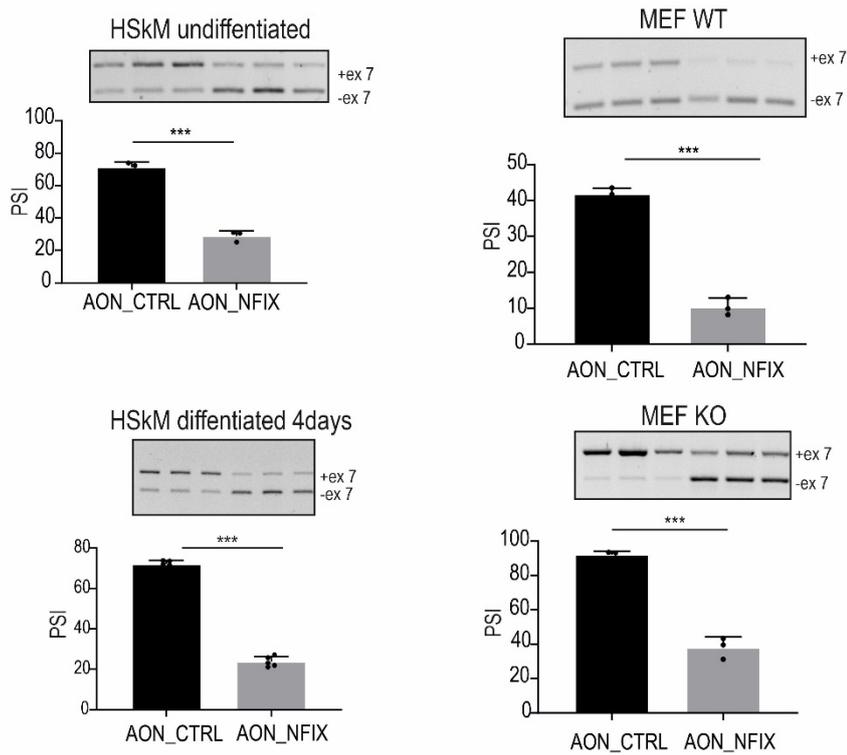
ACCCGGCTTCTCTAAAGAAGTCAGGAAAGCTGGACTTCTGCAGTGCCTCTCCTCTCAGGGCAGCTCCCCGGCATGGCTTTCA
 CCCACCACCCGCTGCCTGTGCTTGTGGAGT**CAGACCAGgtgag**aaatggggggccccggagggggcagttggggaggtgctga
 gtatctaggggcagctggccaggtgaagggccagagctgaccccgagtgacagccccacgctctccaagtgggccaaggtgcagcagcgggt
 gggcatgggagccggcccctgtcacgtagcacctcatcgaggacc

 fragment of intron 7 sequence

 AON binding site

 exon 7 sequence

b



c

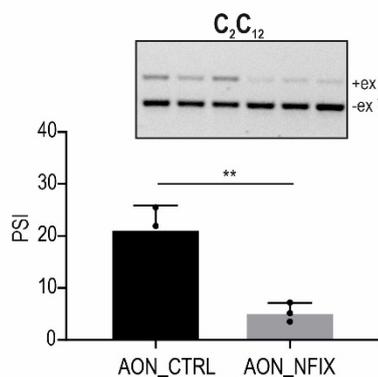


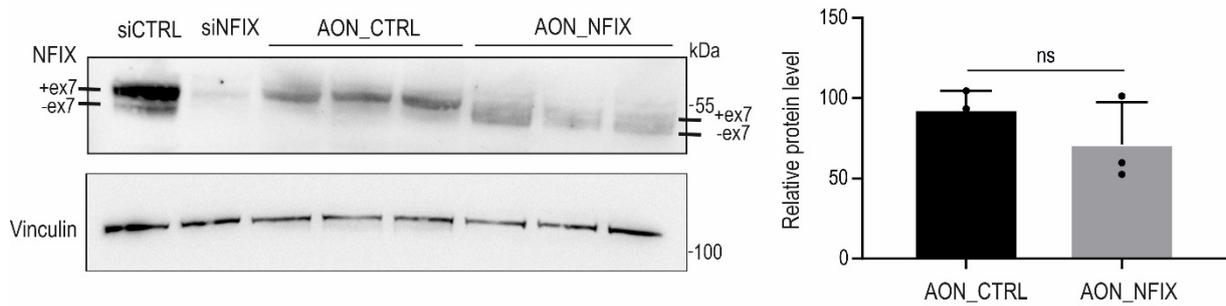
Figure 14. Effect of antisense splicing switchers on inclusion of *NFIX* exon 7

a) The scheme and the sequence of a fragment of the *NFIX* gene showing the binding site for AON within the 5'-splice site which in consequence should lead to the exon exclusion. The AON binding site is indicated by a black frame.

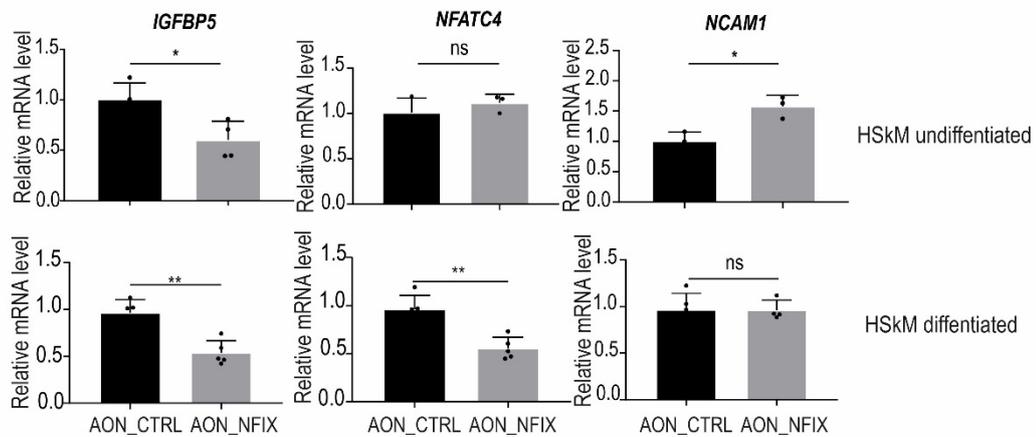
b), c) RT-PCR analysis showing splicing pattern of alternative exon 7 in 4 different cell line after treatment with control (AON_CTRL) or *NFIX*-targeting AONs (AON_NFIX) at a concentration of 100 nM, for 48 h. Bars represent average PSI from 3 independent experiments (dots) with SD; unpaired Student's *t*-test; ** $P < 0.01$; *** $P < 0.001$.

To further confirm that AON induced a skipping of targeted exon the western blot analysis was performed. The change in the splicing pattern of exon 7 induced by AON treatment on RNA level is positively correlated with distribution of protein isoforms of *NFIX* (Fig. 15a). After treatment with *NFIX*-specific AON the *NFIX*-7 isoform predominates. To achieve insight into potential differences in transcriptional activity of two *NFIX* isoforms a real-time RT-PCR analysis of the expression changes of three known *NFIX* targets, namely *IGFBP5* (Pérez-Casellas et al., 2009), *NFATC4* (Messina et al., 2010) and *NCAMI* (Heng et al., 2014), were performed for both HSkM undifferentiated and differentiated myoblasts treated with either control or *NFIX*-specific AONs. The results showed that the expression of most of these genes significantly changes after AON_NFIX treatment what suggests that inclusion of exon 7 may affect the transcriptional activity of *NFIX* (Fig. 15b). Furthermore, gene expression of two markers of muscle cell differentiation, namely myogenic differentiation (*MYOD*) and myogenin (*MYOG*) was evaluated in HSkM cells after 4 days differentiation to check if treatment with AON could have had an impact on myogenesis (Fig. 15c). It appeared that AON treatment did not significantly change the expression of these genes, suggesting that the applied approach does not disturb the differentiation process of skeletal muscle cells (Fig. 15c).

a



b



c

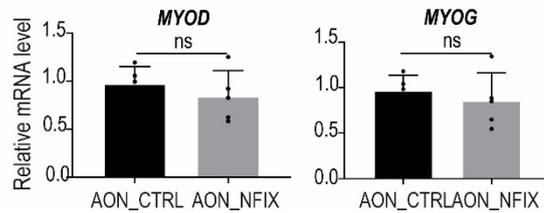


Figure 15. Differences in expression of genes regulated by NFIX after AON_NFIX treatment.

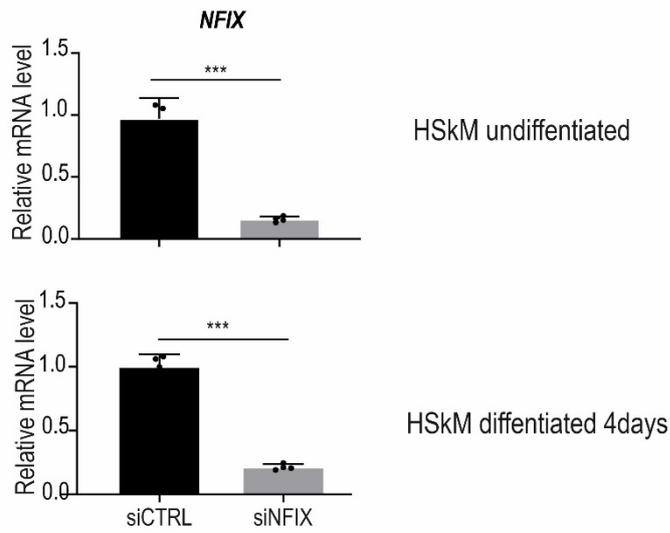
a) Results of western blot analysis showing the level of NFIX (anti-NFIX antibody) in differentiated HSkM cells treated with AON_CTRL or AON_NFIX. Cells treated with siNFIX or siCTRL were used as a control of specificity of antibody used in western blotting. Vinculin was used as loading control. The results are averages from n=3 independent experiments normalized to Vinculin; unpaired Student's t-test; ns, non-significant.

b) Results of real-time PCR analysis of expression level of *IGFBP5*, *NFATC4*, and *NCAM1* in undifferentiated and differentiated HSkM cells, and c) *MYOG* and *MYOD* just in differentiated HSkM, after AON treatment. No changes in *MYOG* and *MYOD* expression suggest that AON_NFIX treatment does not affect myogenesis. Bars represent the average signal from 3 independent experiments (dots) with SD, normalized to *GAPDH*; unpaired Student's *t*-test; * $P < 0.05$; ** $P < 0.01$; ns, non-significant.

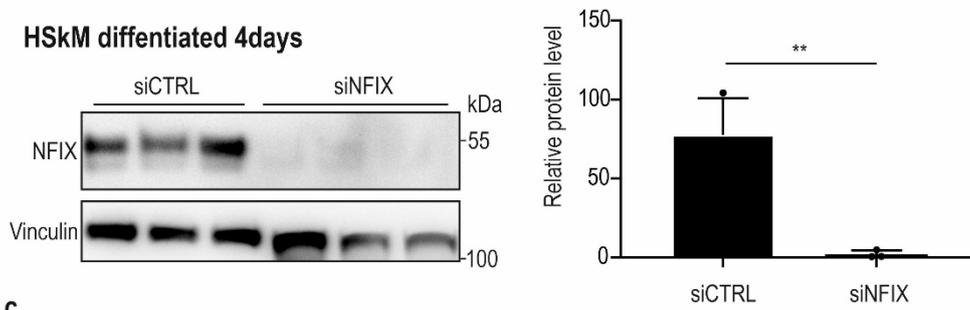
3.1.8 Generation of human skeletal muscle cell model with insufficiency of NFIX

Next, the cell model with the knockdown of *NFIX* was generated to achieve broader insight into the targets of *NFIX* and affected pathways to which *NFIX* may be contributed. The HSkM cells were transfected with siRNA against *NFIX* (siNFIX) or control siRNA. Next, the efficient knockdown on the mRNA level (Fig. 16a) in undifferentiated and differentiated HSkM cells on *NFIX* protein level (Fig. 16b), both in undifferentiated and differentiated HSkM cells, was confirmed. Moreover, analysis of two *NFIX* targets showed significant expression changes of these genes in cells treated with siNFIX (Fig. 16c).

a



b



c

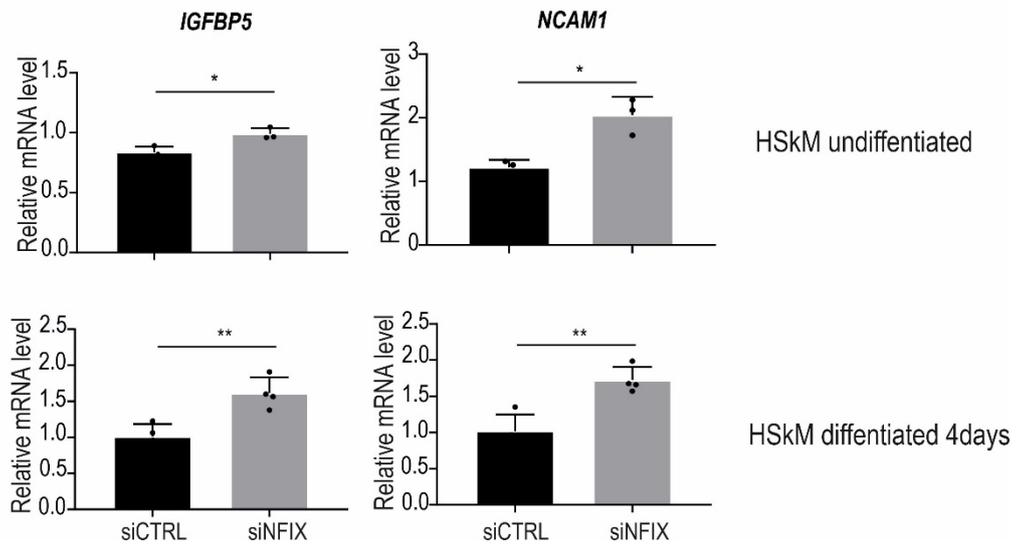


Figure 16. Confirmation of efficient RNAi-induced knockdown of *NFIX* in muscle cells.

a) Results of real-time PCR analysis of expression level of *NFIX* in undifferentiated and differentiated HSkM cells, after siNFIX treatment. Bars represent the average signal from 3 independent experiments (dots) with SD, normalized to *GAPDH*; unpaired Student's *t*-test; ** $P < 0.01$; *** $P < 0.001$.

b) Results of western blot analysis showing the level of NFIX (anti-NFIX antibody) in differentiated HSkM cells treated with siNFIX or siCTRL. Vinculin was used as loading control. The results are averages from $n=3$ independent experiments normalized to Vinculin; unpaired Student's *t*-test; ** $P < 0.01$ for comparison of siCTRL vs siNFIX groups.

c) Results of real-time PCR analysis of expression level of *IGFBP5*, and *NCAM1* in undifferentiated and differentiated HSkM cells, after siNFIX treatment. Bars represent the average signal from 3 independent experiments (dots) with SD, normalized to *GAPDH*; unpaired Student's *t*-test; * $P < 0.05$; ** $P < 0.01$.

3.1.9 Identification of genes sensitive to the level of NFIX and to its splicing isoforms

To investigate potential NFIX targets in the context of differences in activity of two NFIX splicing isoforms (+/-ex7) the whole transcriptome-based RNA sequencing was implemented. Four groups of samples were prepared: cells treated with either control siRNA (siCTRL), siNFIX, control AON (AON_CTRL), or AON_NFIX. The total number of genes identified in RNA-seq data was 10,602 with average expression cutoffs above 1 CPM. Among genes that changed in the absence of NFIX (siNFIX) were 2,250 upregulated and 3,911 downregulated genes ($P_{adj} < 0.01$). Among genes that are sensitive to the exon 7 splicing (AON_NFIX) was 2,761 upregulated and 2,947 downregulated genes ($P_{adj} < 0.01$). Interestingly, as many as 2,040 genes were significantly misregulated ($P_{adj} < 0.01$) both in cells with insufficiency of NFIX (siNFIX) and with enrichment of NFIX-7 isoform (AON_NFIX), which suggests that the expression of this group of genes is sensitive to both the level of NFIX and contribution of its splicing isoforms (Fig. 17a). Among this group, 309 genes were upregulated and 456 genes were downregulated both after treatment with siNFIX and AON_NFIX, 430 were upregulated

after siNFIX and downregulated after AON_NFIX treatment and 875 were upregulated after AON_NFIX and downregulated after siNFIX treatment (Fig. 17b). Last group contains the highest number of genes, however, the greatest fold change is observed in the group of genes positively regulated by NFIX and NFIX+7 isoform (456 genes). In the vast majority of cases, the exclusion of exon 7 leads to a decrease in NFIX activity. Considering the fold change cutoffs above and below 2 ($\log_2FC > 1$ or < -1) the number of identified genes was reduced to 154 ($P_{adj} < 0.01$) (Fig. 17a). The division of this group to the upregulated and downregulated genes reveals that most of these genes are downregulated after siNFIX and AON-NFIX treatment. The transition from NFIX+7 to NFIX-7 (after AON treatment) leads to a comparable effect after insufficiency of NFIX (after siNFIX treatment). Taking this into account, it can be concluded that NFIX+7 is a more efficient activator, but there are also different cases observed where NFIX-7 acts as a stronger activator or repressor (Fig. 17b).

The group of 2,040 genes without fold change cutoff criterion was taken to the gene ontology analysis and also grouped into the upregulated (1,184) and downregulated (886) genes after AON treatment. The analysis for genes downregulated after treatment with AON and significantly changed in NFIX insufficiency, representing genes whose expression is enhanced by NFIX+7 isoform, demonstrated the enrichment of genes involved in an extracellular matrix structural constituent ($P=1.40E-03$; fold enrichment 3.32), external encapsulating region ($P=1.66E-13$, fold enrichment 3.20), extracellular matrix ($P=4.64E-10$; fold enrichment 4.27), collagen-containing extracellular matrix ($P=5.71E-05$; fold enrichment 3.72), extracellular region ($P=5.31E-04$; fold enrichment 1.65) and external side of plasma membrane ($P=2.79E-03$; fold enrichment 3.70). The analysis for upregulated genes after treatment with AON and also significantly changed in NFIX insufficiency demonstrated the enrichment of genes involved in a cell adhesion ($P=7.90E-03$; fold enrichment 2.54), cell surface ($P=7.90E-03$; fold enrichment 2.54), focal adhesion ($P=1.50E-03$; fold enrichment 1.90), actin cytoskeleton

($P=7.10E-03$; fold enrichment 1.86), extracellular exosome ($P=5.80E-03$; fold enrichment 1.44) (Fig, 17c).

Next, several genes from these groups were validated using real-time RT-PCR analysis (Fig. 17d). The obtained results confirmed RNA-seq data showing changes in expression level in the absence of NFIX (siNFIX) and with changes of the distribution of NFIX exon 7 (AON_NFIX). The insufficient level of NFIX (siNFIX) leads both to downregulation and upregulation of gene expression suggesting that NFIX is a positive and negative regulator of certain groups of genes. The analysed genes were divided into NFIX-activated and NFIX-repressed (decrease or increase of mRNA level in case of NFIX insufficiency, respectively). These results highlight that NFIX+7 is a more effective activator or repressor because of observed more genes with significantly changed expression with higher fold change (Fig. 17b). It is consistent with results from HEK-293 cells (Fig. 13e). These results indicate that NFIX+7 is more effective in regulating transcription than NFIX-7, which is highlighted in gene expression activation. It also should be said that both isoforms have different targets for which they perform as stronger regulator. In the majority, NFIX+7 is a stronger activator than NFIX-7 but for particular genes, the NFIX-7 acts as a more effective activator or repressor because of gene preference.

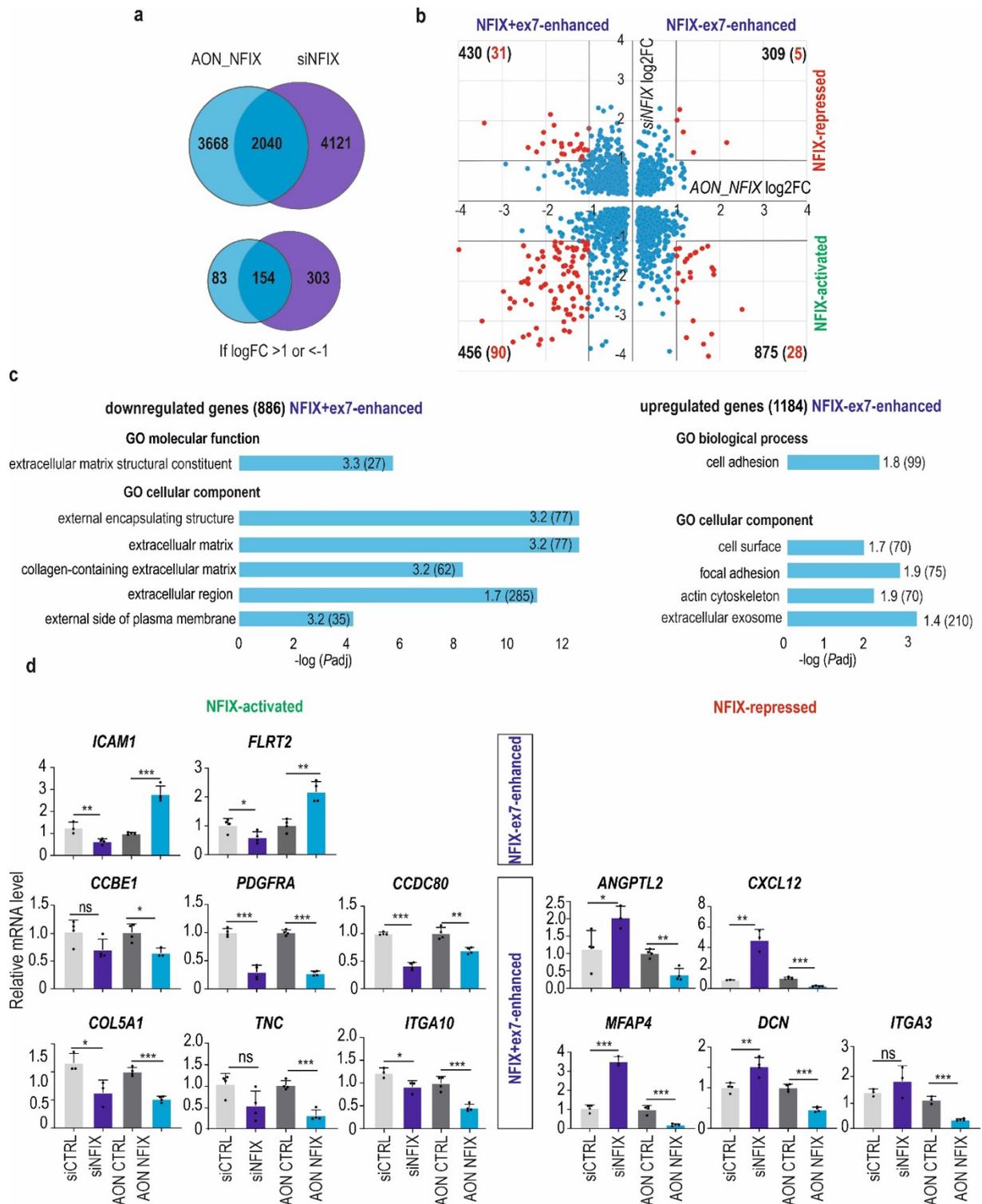


Figure 17. Results of RNA-seq analysis for cells treated with siNFIX or AON_NFIX.

a) Venn diagrams showing genes that are significantly misregulated ($P < 0.01$) in the absence of *NFIX* (siNFIX) and after AON_NFIX treatment (upper diagram) and the same gene groups but with log fold change cutoffs above 1 and below -1 (lower diagram).

b) XY-plot of gene expression values upon siNFIX and AON_NFIX treatment with the indicated number of upregulated and downregulated genes. The genes with the higher fold change above 2-fold are indicated in red.

c) Results of Gene Ontology enrichment analysis performed on 2,040 genes changed both in siNFIX and AON_NFIX treatment, divided into upregulated genes (1,184) and downregulated genes (886) after AON treatment. The selected terms of GO molecular function, GO Biological process, and GO cellular component for upregulated genes and downregulated genes are presented due to the low P value. The numbers behind the bars indicate enrichment and the numbers in the parentheses indicate a number of genes.

d) Results of real-time PCR analysis of the expression level of 13 selected genes from the class which represents enrichment in GO analysis. The genes were chosen because they demonstrated significant changes in expression level. They were divided into NFIX-activated (8 genes) and NFIX-repressed (5 genes), decrease and increase of expression level after siNFIX treatment, respectively. Bars represent the average signal from 3 independent experiments (dots) with SD, normalized to *GAPDH*; unpaired Student's *t*-test; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; ns, non-significant.

3.1.10 Abnormal splicing of *NFIX* exon 7 significantly contributes to the gene expression changes in skeletal muscles of DM1 patients

As described above, aberrant splicing of exon 7 of *NFIX* is observed in DM1 (Fig. 7b). The insufficiency of MBNL caused by its sequestration on CUG^{exp} RNA leads to abnormal inclusion of exon 7. Therefore, NFIX+ex7 isoform predominates in skeletal muscles of DM patients compare to normal muscles, where NFIX-ex7 is major isoform. In the last part of the project expression changes of previously identified group of genes sensitive to isoforms of NFIX transcription factor was analyzed in a list of genes abnormally expressed in DM1 to characterize contribution of abnormally spliced isoforms of NFIX in disease development. To answer this question the results of previously published whole transcriptomic data obtained for skeletal muscle biopsies from DM1 patients (Nakamori et al., 2013) were juxtaposed with the results of RNA-seq experiments for muscle cells with different NFIX splicing profile obtained in this study.

In total a list of 10,602 genes was represented in both data sets (only genes with average expression cutoffs above 1 CPM were selected for further analysis). The splicing pattern of muscle cells treated with AON_NFIX reflects those observed in healthy individuals (NFIX-ex7 predominates), while cell treated with AON_CTRL reflects those observed in DM1 patients (NFIX+ex7 predominates). The analysis revealed 2,070 genes whose expression is changed in DM1 tissues and in cells treated with AON_NFIX, and 861 genes with changed expression in DM1 tissue and both in cells with insufficiency of NFIX (siNFIX) and with enrichment of NFIX-7 isoform ($P_{adj} < 0.05$) (Fig. 18a). From genes whose expression is changed significantly in DM1 tissues ($P_{adj} < 0.01$) the great amount of genes are also abnormally expressed in AON_NFIX and siNFIX treated cells ($P < 0.01$) (Fig. 18b). These 861 genes were divided into two subgroups: (i) upregulated in muscles of DM1 patients ($P_{adj} < 0.05$) and downregulated in cells treated with AON_NFIX ($P < 0.05$) and (ii) genes that are downregulated in DM1 patients and upregulated in cells with enrichment of NFIX-7 (Fig. 18c).

After that, we focus on these groups of genes because they may consist of genes whose expression may be changed in DM1 patients as a result of abnormal distribution of NFIX isoforms due to abnormal splicing of exon 7. Thus, the gene ontology analysis to get a broader insight into the classification of these genes was performed. The analysis for the first group (DM1 upregulated and AON_NFIX downregulated – NFIX+7-enhanced genes) revealed significant enrichment of genes involved in an collagen containing extracellular matrix ($P = 3.32E-08$; fold enrichment 4.48), endoplasmic reticulum lumen ($P = 9.04E-05$; fold enrichment 4.31), extracellular exosome ($P = 1.70E-04$; fold enrichment 1.98), angiogenesis ($P = 6.86E-02$; fold enrichment 3.38), cell migration ($P = 7.78E-03$; fold enrichment 2.53), collagen binding ($P = 1.02E-03$; fold enrichment 8.46) and extracellular matrix structural constituent ($P = 7.12E-07$; fold enrichment 6.82) (Fig. 18d). This analysis also confirmed results described in previous subchapter regarding processes in which genes responding to exon 7

alternative splicing are engaged (Fig. 17c, d). The analysis for the second group (DM1 downregulated and AON_NFIX upregulated) did not show any significant enrichments.

Next, three out of twelve genes which expression changes were confirmed in the previous subchapter (Fig. 17d) were tested using real-time RT-PCR analysis also in *HSA*-LR mice (Fig. 18e). The obtained results confirmed RNA-seq data showing changes in expression level in the absence of NFIX (siNFIX treated cells) and in alterations in the presence of exon 7 of NFIX (AON_NFIX treated cells). Results obtained for this model, in which NFIX+7 predominates (Fig. 8a), are consistent with gene expression patterns in samples after AON_NFIX treatment where NFIX-7 isoform predominates (reverse direction of gene expression changes). The significant changes observed in *HSA*-LR were demonstrated in less amount (3 out of 12) splicing events which can be a reason for the smaller disproportion of NFIX isoforms in *HSA*-LR mice than in DM1 patients (Fig. 7c). Cumulatively, all these data strongly suggest that significant fraction of impaired gene expression in DM1 skeletal muscles is related to disruption of altered splicing of exon 7 of *NFIX*.

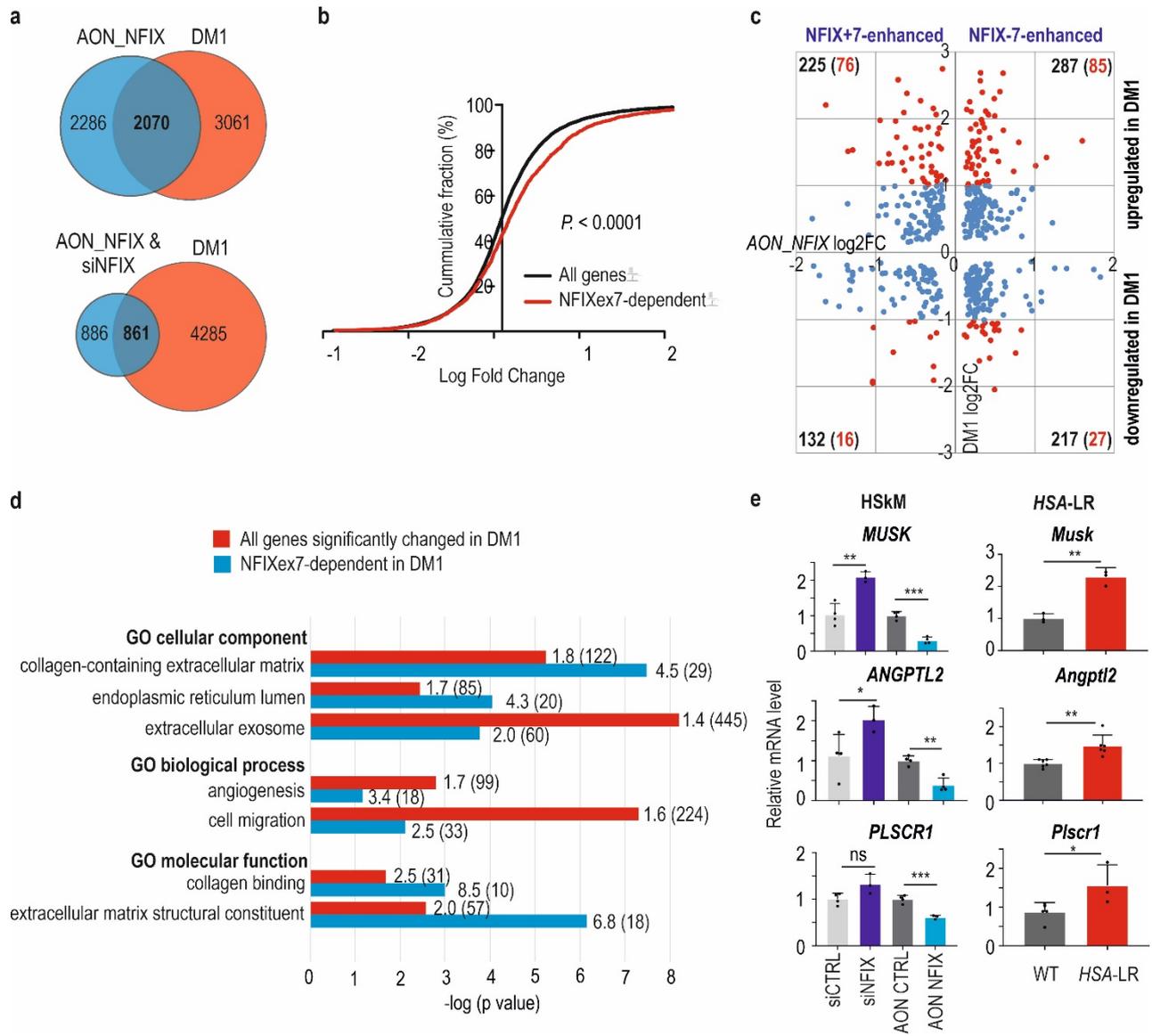


Figure 18. Changes in gene expression observed in DM1 skeletal muscles partially are a consequence of abnormal NFIX isoform distribution.

a) Venn diagram showing the overlapping genes after AON_NFIX treatment (difference in distribution of NFIX splicing isoform) and in patients with DM1 (enrichment of NFIX+7 isoform) (top diagram), and both in cells with insufficiency of NFIX (siNFIX) and with enrichment of NFIX-7 isoform and in patients with DM1 (bottom diagram).

b) The juxtaposition of changes in gene expression in DM1 tissues (black line) and changes of NFIX-dependent genes in DM1 (red line). Genes were selected based on experiments performed in cells treated with both siNFIX and AON_NFIX (NFIXex7-dependent). The graph shows a statistically significant difference $*** P < 0.001$, based on Mann-Whitney test.

c) XY-plot of gene expression changes in DM1 tissue (y axis) and upon AON_NFIX treatment (x axis) with the indicated number of upregulated and downregulated genes. The number of

genes with changed expression are indicated for each quarter of the chart. The genes with the higher fold change (Log₂FC higher than 1 and lower than -1) are indicated in red in parentheses.

d) Gene Ontology enrichment analysis for genes downregulated after AON_NFIX treatment and changed in DM1 (861 genes). The red bars represent all genes significantly changed in DM1, blue bars represent NFIXex7-dependent genes in DM1. The representative GO Cellular Component terms and GO Molecular function terms are shown. The groups of genes which are presented on these graphs were chosen due to the low P value and proven gene expression changes in DM1 patients. The red bars represent genes with significantly changed expression in DM1, and the blue bars represent genes with significantly changed expression of genes dependent on NFIXex7 in DM1 (cells treated with siNFIX and AON_NFIX). The numbers behind the bars indicate enrichment, and the numbers in the parentheses indicate a number of genes.

e) Results of real-time PCR analysis of expression three out of twelve chosen NFIX targets which displayed enrichment in GO analysis in HSkM cells and in *HSA*-LR mice. In *HSA*-LR the gene expression is upregulated (majority of NFIX+7), whereas after AON_NFIX (majority of NFIX-7) gene expression is downregulated. Bars represent the average signal from n=3-4 independent experiments (dots) for each group with SD, normalized to *GAPDH*; unpaired Student's *t*-test; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; ns, non-significant.

3.2 DISCUSSION

NFIX is a transcription factor that regulates the expression of many genes which are involved in different cellular processes (Gronostajski, 2000). It modulates (activates or represses) the transcription as homodimers or heterodimers with another member from the NFI family or with different proteins (Gronostajski, 2000). Interestingly, the recent study on human transcription factor (TF) interactions showed that most (118) of the detected (202) TF-TF interactions were interactions with NFI family members. This result proves that NFI family members are engaged in multiple cellular processes with other TFs (Göös et al., 2022). It was already proven that NFIX is engaged in the development of the brain (Campbell et al., 2008), skeletal muscle (Messina et al., 2010), hematopoiesis (Holmfeldt et al., 2013) and spermatogenesis (Davila et al., 2022). The common point from all of these studies is the intermediary role of NFIX in stem cell biology like proliferation and differentiation (Piper et al., 2019). Experiments performed in this study allowed to uncover new NFIX functions. Moreover, they showed that changes in the alternative exon 7 splicing pattern impact activity of this transcription factor and contribute to pathogenesis of myotonic dystrophy.

3.2.1 NFIX is both a positive and negative regulator of gene expression

NFIX is one of the crucial regulators among transcription factors, of temporal progression of muscle generation in mice (Rossi et al., 2016). The majority of previously described studies of NFIX function were mainly conducted on transgenic animals with *Nfix*-KO. The studies carried out so far provide insight into NFIX function in case of its insufficiency. The genetic knockout of *NFIX* in mice or cell lines gives general information about NFIX function. However, analysis of gene expression changes performed for these models does not always show the direct targets of NFIX but can also result from secondary consequences of introduced knockout. Moreover, the mice with knockout of any of the members of the *Nfi* family exhibit a lot of similarities. Both knockouts of *Nfia*, *Nfib*, or *Nfix* lead to the inappropriate development of the hippocampus

and delayed differentiation into mature neurons and glia (Heng et al., 2014; Piper et al., 2010, 2014). It was revealed that a significant group of target genes is common among the members (Fraser et al., 2020). The observed changes not always may be the result of a lack of *Nfix* because of the possible mutual replenishment. To overcome these limitations, the model of two stable cell lines with inducible overexpression of NFIX isoforms with and without exon 7 (NFIX+7 or NFIX-7) based on a HEK cell line with a functional knockout of endogenous *NFIX* (*NFIX*-KO) was generated. Inducible gene expression activated by doxycycline allowed to reduce secondary consequences of introduced knockout. Results of this part of the project confirmed the functional knockout of *NFIX* in T-REx™-293 cell: NFIX protein was not detected (Fig. 8e, f), expression of NFIX targets and the level of circ*NFIX* changed significantly (Fig. 8g and 9b). These changes are similar to those observed for the *Nfix*-KO mouse model (Fig. 9c). The obtained results indicate interesting evidence of this transcription factor. Firstly, the expression level of other NFI factors is not affected after changes in NFIX expression level (*NFIX*-KO and either in *NFIX*-OE) (Fig. 12). The performed RNA-seq analysis revealed that about 40% from total identified genes with average expression cutoffs above 1(CPM) are changed in the comparison *NFIX*-KO vs WT, and just about 1% of genes in the comparison *NFIX*-KO vs *NFIX*-OE, but most of them are shared for both comparisons (Padj < 0.01) (Fig 13c). Surprisingly, RNA-seq results did not show significant differences in the transcriptional activity of the two isoforms (NFIX+7 and NFIX-7). The potential explanation for it may be too high level of NFIX expression of exogenous, even without chemical induction and the expression of two isoforms might be different in both cell lines, despite the same dosage of doxycycline, which was used as inductor of transgene expression. Moreover, this model did not ensure the natural promoter and environment for NFIX. HEK cells are not the environment where NFIX is highly expressed and engaged in the developmental process, thus many partner proteins which can be associated with different functions of NFIX isoforms may not be

expressed. Nevertheless, it was revealed that a low level of NFIX overexpression, before the addition the doxycycline (*NFIX*-OE) lead to a significant change in gene expression of several hundred genes (Fig. 13d). However, the increase of NFIX overexpression (after the addition of doxycycline) lead to further activation or repression of particular genes, however, most of the changes are not statistically significant (Fig. 13e). This analysis demonstrated that NFIX acts as an activator or repressor of many genes, nonetheless the positive regulation predominates (about 70-75% genes sensitive to NFIX) and these changes show higher fold change (Fig. 13d). Moreover, the increasing level of NFIX+7 enhanced transcription of NFIX-dependent genes more efficiently than isoform lacking exon 7 (Fig. 13e).

The Gene Ontology (GO) analysis, for genes upregulated by NFIX, revealed significant enrichment for the class of genes involved in, among others, an extracellular region which was already shown (Fig. 13f) (B. Zhou et al., 2015). The GO analysis, for downregulated genes shows that NFIX is a negative regulator of genes associated with RNA metabolism (Fig. 13f and g). These experiments revealed a new function of the transcription factor.

3.2.2 NFIX+7 is a more effective activator of gene expression regulation

As mentioned above, NFIX has a clearly apparent effect on myogenesis. It targets many different myogenic regulators, mostly as a heterodimer with other transcription factors. One of the key examples is the formation of the complex with Protein kinase C theta (PRKCC) for activation of Myocyte Enhancer Factor 2A (*Mef2a*), the protein encoded by this gene is a transcription factor that activates many muscle-specific genes and is engaged in muscle development (Messina et al., 2010). Another example in which NFIX acts as a heterodimer in myogenesis is an interaction with SRY-box transcription factor 6 (SOX6) to repress *Myh7*, encoding for myosin heavy chain, present at a low level in secondary fibers during fetal myogenesis (Taglietti et al., 2016). On the other hands, NFIX may also act as homodimer

without other partners. One of best known example is the direct repression of myostatin gene (MSTN), which is a critical regulator of prenatal and postnatal myogenesis (Rossi et al., 2016). Moreover, the duration and the degree of NFIX expression are crucial for controlling the proper timing of myogenic differentiation (Xiao et al., 2022). These results show that changes in the expression of the above genes is an effect of direct or indirect NFIX regulation. Moreover, the studies conducted so far were mainly focused on NFIX deficiency models, without distinguishing between distinct role of protein isoforms, including the two splicing isoforms (NFIX+7 or NFIX-7), analyzed in this study.

The previous research and those performed in this study revealed that both in embryonic and postnatal development and also during myogenic differentiation the changes in splicing pattern of *NFIX* exon 7 are observed (Fig. 7f). Taking this into account the human skeletal muscle cell models with enrichment of either NFIX+7 or NFIX-7 isoform were generated for broader understating of participation of changes in splicing pattern of exon 7 in transcriptional regulation. This approach was based on manipulation of exon 7 inclusion of endogenous *NFIX* using the antisense oligonucleotides (AON)s inducing efficient exon 7 skipping and production of NFIX-7 (Fig. 14) as major isoform in human skeletal muscle cells in which NFIX+7 predominates. These AONs belong to the class of modified nucleic acid analogues which do not induce target RNA degradation. They bind to the RNA by Watson-Crick base pairing and have the ability to switch splicing by interfering with proteins involved in the splicing machinery (Kole et al., 2012). This type of AONs was already used in many studies to broaden insight into particular splicing variants. One of them is the implantation of the splice-switching AONs to study *MBNL1* splicing isoform, lacking exon 7. This study revealed the function of MBNL1 lacking exon 7 protein, which is an antisurvival factor with a defined tumour suppressive role (Tabaglio et al., 2018). Another application of splice-switching AONs regards *CLCN1* which aberrant splicing lead to one of the DM symptoms – myotonia. This study

revealed the differences in the splicing pattern of *CLCN1* between mice and humans (Nakamura et al., 2016).

The results of RNA-seq analysis performed in this study highlight that there are at least a few hundred genes, which are NFIX targets and are sensitive to the inclusion of exon 7, suggesting that both isoforms differ in activity. Also, the cell model with the knockdown of *NFIX* was generated (using siNFIX) to broaden insight into the genes regulated by NFIX. From the total number of genes identified in RNA-seq results, about one third of genes, with average expression cutoffs above 1(CPM), are changed in the absence of NFIX ($P < 0.01$). Interestingly, about half of genes from this group is significantly misregulated ($P < 0.01$) both in cells treated with siNFIX (absence of NFIX) and with AON_NFIX (majority of NFIX-7 isoform). Because of gene preference, for particular genes, NFIX+7 is a more effective transcription regulator (activator or repressor), for others the NFIX-7. The obtained results indicate that NFIX+7 is a more efficient activator, because of greater fold change values (Fig 17b). Moreover, it was revealed that genes that are dependent on NFIX splicing isoforms are involved in functions that were previously mentioned: the extracellular matrix, extracellular region, cell adhesion, (B. Zhou et al., 2015), and actin cytoskeleton (Liu et al., 2020). Additionally, this approach uncovered so far unknown NFIX functions. These results give a broad insight into the NFIX function, which can be part of the large network of proteins that are engaged in for example communication between cells, growth, migration, or adhesion.

3.2.3 The enrichment of the NFIX+7 isoform leads to the abnormal gene expression observed in DM1

The main symptoms of DM include muscle weakness and wasting. They are associated with the abnormal splicing of many genes. *NFIX* is one of the MBNL-sensitive genes, in which alternative exon 7 is predominantly included in skeletal muscles of DM. There are many genes

whose aberrant splicing has an impact on DM symptoms. The Bridging Integrator-1 (*BINI*) encodes a protein required for the proper biogenesis of transverse tubules. Exclusion of exon 11 from its pre-mRNA due to MBNL sequestration leads to T tubule abnormalities and muscle weakness (Fugier et al., 2011). Another one, Dystrobrevin- α (*DTNA*) which exclusion of exon 11A and 12 alters the signaling pathway that has an impact on muscle growth (Nakamori et al., 2008). Myotonia is associated with membrane hyperexcitability triggered by the chloride channel which is deprived of its normal function. This protein is encoded by *CLCN1* which inclusion of exon7a (Lueck et al., 2007) and intron 2 in pre-mRNA in DM conditions leads to the occurrence of premature stop codon and formation of truncated protein (Mankodi et al., 2002). The Ryanodine Receptor 1 (*RYR1*) is involved in intracellular calcium homeostasis which is disturbed in DM1 due to exclusion of exon 70 from the pre-mRNA and contributes to the reduced contraction strength of the muscle (Kimura et al., 2005; Nakamori et al., 2013). Another example of an abnormal splicing event that affects calcium homeostasis and then muscle degeneration is the exclusion of exon 22 in *ATP2A1* pre-mRNA (Kimura et al., 2005). Dystrophin is a protein that maintains the muscle structure, the exclusion of exon 78 from *DMD* pre-mRNA coding dystrophin causes instability of the sarcolemma and subsequently has an impact on muscle weakness (Nakamori et al., 2013; Rau et al., 2015). The decrease of protease Calpain3 activity encoded by *CAPN3* due to the exclusion of exon 13 in from its pre-mRNA also correlates with muscle weakness (Lin et al., 2006; López-Martínez et al., 2020). Another DM1 symptom is insulin resistance, which occurs as a result of the exclusion of exon 11 from the *INSR* pre-mRNA which encodes for Insulin Receptor. The abnormal splicing leads to the formation of proteins characterized by a lower response to insulin (Savkur et al., 2001). For the majority of genes with altered splicing pattern, the molecular consequence is still unknown. So far phenotypic consequences of abnormal splicing of *NFIX* pre-mRNA were not described. The splicing defect of *NFIX*, but also abnormalities of more than 20 other genes was correlated

with muscle weakness of DM1 patients (Nakamori et al., 2013). Therefore, the examination of changes in NFIX exon 7 seems to be essential to better understand the pathomechanism of DM.

Taking into consideration the above reports, the generation of alternative models for a better understanding of the NFIX function seemed to be essential. Especially, the fact that the understating of the function of a particular splicing isoforms is fundamental in broadening the knowledge of myotonic dystrophy pathomechanism.

The juxtaposition of whole transcriptomic data obtained for skeletal muscle biopsies from DM1 patients (Nakamori et al., 2013) with the results of RNA-seq experiments with different distributions of NFIX splicing isoforms revealed that the large number of genes whose expression changed significantly in DM1 tissues ($P < 0.05$) are also abnormally expressed in cells with NFIX-7 enrichment ($P < 0.01$) (Fig. 18a). Gene Ontology analysis showed significant enrichment of groups of genes which expression is also changed in DM1 patients. The analysis of the group of genes that are downregulated in DM1 patients (enrichment of NFIX+7) and upregulated in cells with enrichment of NFIX-7 (AON_NFIX treatment) revealed enrichment of genes involved in, among others, collagen containing extracellular matrix, extracellular matrix structural constituent (Fig. 18d). The obtained results are consistent with the previous data of gene expression analysis in DM1 myoblast and myotubes (Todorow et al., 2021), and heart and muscle biopsies from DM1 individuals (E. T. Wang et al., 2019).

Taken together, these results showed that abnormal distribution of NFIX exon 7 caused a change in gene expression in DM1 skeletal muscles, which may be responsible for the development of disease phenotype observed in DM1 patients.

The comparison of non-disease controls between proliferating myoblast and post-mitotic myotubes showed changes in gene expression involved in cell adhesion or extracellular matrix organization, the same groups identified in analysis with different distributions of NFIX

splicing isoforms. Abnormal expression of genes from these groups may have an impact on appropriate differentiation, development of muscle, and regeneration (Todorow et al., 2021). Similarly, the comparison of DM1 tissue with unaffected individuals obtained from a series of heart and muscle autopsies and biopsies demonstrated changes in gene expression involved also in the extracellular matrix, and regulation of cell-substrate adhesion (E. T. Wang et al., 2019).

In summary it may be said that two isoforms show significant differences in activity of specific groups of genes. Some genes are more efficiently regulated by NFIX+7 and some by NFIX-7. They are negative regulators of genes involved, for instance, in cell adhesion, structural constituent of chromatin, focal adhesion, and positive regulators of genes involved, for instance, in the extracellular matrix, collagen-containing extracellular matrix, external side of plasma membrane. The presence of exon 7 has an impact on gene expression changes which is probably correlated with DM1 symptoms. Higher incidence of NFIX+7 may be important driver of increased expression of extracellular matrix proteins in DM, what may impact connection of myofibers in skeletal muscles. The confirmation of the correlation between DM1 symptoms and changes in NFIX isoform distribution requires further investigation.

3.3 Conclusions

1. Two stable cell lines with inducible overexpression of NFIX isoforms with and without exon 7 (NFIX+7 or NFIX-7) based on the HEK cell line with a functional knockout of endogenous *NFIX* (*NFIX*-KO) were generated. As revealed from RNA-seq experiments majority of genes is positively regulated by NFIX (ca. 70-75%). NFIX acts as a repressor for about 25-30% of genes whose expression increases in cells with NFIX insufficiency (*NFIX*-KO) (Fig. 13). These results confirm that NFIX is both positive and negative regulator of gene expression.

2. Second cellular model generated in this project showed significant differences in transcription regulation by the two NFIX splicing isoforms. Antisense oligonucleotide (AON) induces efficient exon 7 skipping and production of NFIX-7 isoform in human skeletal muscle cells. The RNA-seq analysis revealed that generally NFIX+7 isoform is stronger activator as positively regulates higher number of genes and with greater fold change, however, for individual genes, NFIX-7 isoform is more active (Fig. 17). This result is consistent with the described above observation from HEK-derived cell models indicating the majority of NFIX-sensitive genes are positively regulated by this transcription factor.

3. In skeletal muscles of DM1 patients the NFIX+7 splicing isoform predominates while in normal muscles the NFIX-7 is the major isoform. Comparison of previously published results of differential expression in DM1 and results obtained in this project revealed that significant fraction of genes abnormally expressed in DM1 is sensitive to NFIX+7 (Fig. 18). The major GO terms enriched for genes sensitive to the NFIX splicing isoforms and changed in DM1 tissues are functions that are essential for the proper structure and activity of skeletal muscles, e.g. extracellular matrix structural constituent and collagen binding. This result suggests that changes in the splicing pattern of *NFIX* exon 7 caused by sequestration of MBNL proteins on toxic CUG^{exp} RNA in DM1 is important trigger of molecular pathomechanism of the disease.

3. 4. Materials and Methods

Table 1. sgRNA sequences

sgRNA	sequence
sgRNA#1	5'-TCCCTCGCCCCGCATGCTCC- 3'
sgRNA#2	5'-GCGGATGACTGCCTGCGCC- 3'

Table 2. siRNA and AONs sequences

siRNA	sequence
siCtrl sense	5' -UCGAAGUAAUCCGCGUACGdT-3'
siCtrl antisense	5' -CGUACGCGGAAUACUUCGAdT-3'
siNFIx sense	5' -CCUCUGAUAAGAUCGACAAdT-3'
siNFIx antisense	5' -UUGUCGAUCUUAUCAGAGGdT-3'
AON CTRL	5'- GCTTATCGTCGCC -3'
AON NFIx	5'- CTCACCTGGTCTG -3'

dT; deoxythymidine overhang, AON composed of LNA units (bolded) and DNA, all positions were phosphorothioated

Table 3. PCR primers used for alternative splicing assays and gene expression analyses;

PCR target	Primer set	sequence
<i>IGFBP5 promoter</i>	IGF5prF	5'-ACGCGTCCCTTGCCCCTTTCTTACAT-3'
	IGF5prR	5'-CATATGCCTGCAGAAGTTTCCAAAGAG-3'
<i>NFIx ex7</i>	NFIx_F	5'-GAGCCCTGTTGATGACGTGTTCTA -'3
	NFIx_R	5'-CTGCACAAACTCCTTCAGTGAGTC - 3'

Nfix ex7	Nfix_F	5'-GAGTCCAGTAGATGATGTGTTCTA - 3'
	Nfix_R	5'-CTGCACAAACTCCTTCAGCGAGTC - 3'
<i>NFIX</i> ex2 (genotyping)	NFxint1F	5'-GCTGCTTTTCTCCTTCTCTC - 3'
	NfxInt2R	5'-GGAGACAGAAGACAGGAAAA - 3'
<i>NFIX</i> ex2 (mRNA)	NXhoRSF	5'-ATATGATCTCGAAGCTATG - 3'
	NF3UTRr1	5'-GGGGGATTTTTCCATTCT - 3'
circNFIX	circNFIX_F	5'-TTCCTACACCTGGTTCAA - 3'
	circNFIX_R	5'-CAATGTGATGTGGCTGGA - 3'
<i>IGFBP5</i>	IGFBP5_F	5'-ACTGTGACCGCAAAGGATTC - 3'
	IGFBP5_F	5'-CTGTCTGAAGGCGTGGCACT - 3'
<i>Igfbp5</i>	Igfbp5_F	5'-ACTGTGACCGCAAAGGATTC - 3'
	Igfbp5_R	5'-CTGTCTGAAGGCGTGGCACT - 3'
GAPDH	GAPDH_F	5'-GAGTCAACGGATTTGGTCGT - 3'
	GAPDH_R	5'-TTGATTTTGGAGGGATCTCG - 3'
ACTB	ACTB_F	5'-TCCCTGGAGAAGAGCTACGA - 3'
	ACTB_R	5'-AGCACTGTGTTGGCGTACAG - 3'
<i>SOX3</i>	SOX3_F	5'-AACGCGTTCATGGTGTGGTC - 3'
	SOX3_R	5'-CCGGGTACTCCTTCATGTGC - 3'
<i>NFATC4</i>	NFATC4_F	5'-CCAGACTCCAAGGTGGTGTT - 3'
	NFATC4_R	5'-CTTTGCAGATCACAGGCAGA - 3'
<i>NFIA</i>	NFIA_F	5'-CGCCCGGCAGTTATGTATTC - 3'
	NFIA_R	5'-ATTTTCGTTTTTCGGGCCTGC - 3'
<i>Nfia</i>	Nfia_F	5'-AGGGACAAGCATGGGTGTTT - 3'
	Nfia_R	5'-AAGCATGGGTACACTGTCCA - 3'

<i>NFIB</i>	NFIB_F	5'-CAAAGTTTCGCCTGCGATCA - 3'
	NFIB_R	5'-GCGCTGGGAAAGTTCAAGGT - 3'
<i>Nfib</i>	Nfib_F	5'-TCCTGCCAAGAATCCTCCAG - 3'
	Nfib_R	5'-TCGGTGGAGAAGACAGCGACC-3'
<i>NFIC</i>	NFIC_F	5'-GGATGTATTCGTCCCCGCTC-3'
	NFIC_R	5'-GTTGAACCAGGTGTAGGCCGA-3'
<i>Nfic</i>	Nfic_F	5'-GACCTGTACCTGGCCTACTTTG-3'
	Nfic_R	5'-CACACCTGACGTGACAAAGCTC-3'
UPF2	UPF2_F	5'-CAGAGTCAAGCTGAGGGCA -3'
	UPF_R	5'-CGCTCAGCTGGCATTATGTG-3'
Upf2	Upf2_F	5'-AGCCCGAGCGCTGGAGT-3'
	Upf2_R	5'-CGCTCAGCTGGCATTATGTG-3'
HNRNPU	HNRNPU_F	5'-GCTATCCATACCCTCGTGCC-3'
	HNRNPU_R	5'-CCACTGGTTGTAGCCCTGAG-3'
Hnrnpu	Hnrnpu_F	5'-GCTATCCATACCCACGTGGC-3'
	Hnrnpu_R	5'-CCACTGGTTGTAGCCCTGAG-3'
DDX3X	DDX3X_F	5'-ATGGCTTGTGCCCAAACAG-3'
	DDX3X_R	5'-CGCCTGGACCATCTGAATAAA-3'
Ddx3x	Ddx3x_F	5'-ATGGCTTGTGCCCAAACAG-3'
	Ddx3x_R	5'-TAGCCCTCAGAGCTTCTCCT-3'
DHX9	DHX9_F	5'-TTGGCAGTACACGGTATGGA-3'
	DHX9_R	5'-ATAGCCTCCACCAACACCTG-3'
Dhx9	Dhx9_F	5'-GGTTCCCAAGGAGGCTACAG-3'
	Dhx9_R	5'-CCATAACCCCTTCGCTGGAA-3'

HNRNPH3	HNRNPH3_F	5'-CTCCTGCGAGAGGACGTT-3'
	HNRNPH3_R	5'-GGCGCCAATTCCTTTCACC-3'
FUS	FUS_F	5'-AGCAGTGGTGGCTATGAACC-3'
	FUS_R	5'-ATGACGTGATCCTTGGTCCC-3'
MYOD	MYOD_F	5'-TACAGTGGCGACTCAGATGC-3'
	MYOD_R	5'-AGATGCGCTCCACTATGCT-3'
MYOG	MYOG_F	5'-CTTGCTCAGCTCCCTCAACC-3'
	MYOG_R	5'-GGTGTTAGCCTTATGTGAATGG-3'
ICAM1	ICAM1_F	5'-ACGGAGCTCCCAGTCCTAAT-3'
	ICAM1_R	5'-CTCCTTCTGGGGAAAGGCAG-3'
FLRT2	FLRT2_F	5'-GGATCAGGGTGGCAGTTCTC-3'
	FLRT2_R	5'-ATCTGCAATCTGACGGCTGC3'
ANGPTL2	ANGPTL2_F	5'-GAGAACACCAACCGCCTCAT-3'
	ANGPTL2_R	5'-GCCAGTATTCGCCGTCAATG-3'
Angptl2	Angptl2_F	5'-GACATGACCCTGGAGGTTGG-3'
	Angptl2_R	5'-CCAGCCAGTACTACCATCG-3'
CXCL12	CXCL12_F	5'-GGACTTTCCGCTAGACCCAC-3'
	CXCL12_R	5'-GTCCTCATGGTTAAGGCCCC-3'
MUSK	MUSK_F	5'-GGAAGTTGAGGTTTTTGCCAGG-3'
	MUSK_R	5'-AGTGCAGGGTCACAAAGGAG-3'
Musk	Musk_F	5'-AATGGAAGGAAAGGCCACC-3'
	Musk_R	5'-GTTTGGAATGTCCGCGCTC-3'
MFAP4	MFAP4_F	5'-CTGACAGCATGAAGGCACTC-3'
	MFAP4_R	5'-GGGGTAGATGAGGTACACGC-3'

PLSCR1	PLSCR1_F	5'-CCCTTTTCCGGCTGACTTCT-3'
	PLSCR1_R	5'-CGAGACAAGGTCCAGAGAGC-3'
Plscr1	Plscr1_F	5'-AGCTGCTGTTCCGACATTGA-3'
	Plscr1_R	5'-AGGTCTAGCGGGAAGTGGAT-3'
DCN	DCN_F	5'-TCCTTTCCACACCTGCAAAC-3'
	DCN_R	5'-GCCTCTCTGTTGAAACGGTC-3'
ITGA3	ITGA3_F	5'-CCCAGAGGACCAAGGAAACC-3'
	ITGA3_R	5'-CTCCTGGCTCAGCAAGAACA-3'

NFIX ex7/ *Nfix* ex7 - alternative splicing of exon 7; *NFIX* ex2 (genotyping) – confirmation of *NFIX* knockout on DNA; *NFIX* ex2 (mRNA) - confirmation of *NFIX* knockout on cDNA (deletion of exon 2)

Genetic constructs

The vectors used to incorporate sgRNA sequences (Table 1) were obtained through cloning appropriate DNA sgRNA#1 fragment to the pSpCas9n(BB)-2A-Puro (PX462) V2.0 and sgRNA#2 fragment to the pSpCas9n(BB)-2A-GFP (PX458) (Addgene). Both of them already consist of Cas9-encoding sequence. One of them contains one sgRNA-encoding gene and puromycin resistance gene which allows for selection on positive clones, the second consists of second sgRNA and GFP sequence. Luciferase reporter assays were performed using modified pmirGLO vector (Promega) by replacement of original promoter with fragment of *IGFBP5* promoter sequence (DNA fragment obtained in PCR using primer pair for *IGFBP5* promoter fragment Table 3). Two stable cell lines expressing a particular *NFIX* isoform were generated through co-transfection with pOG44 Flp-Recombinase Expression Vector which was a gift (Dr Adama Ciesiołka, Adam Mickiewicz University), and pcDNA5/FRT/TO encoding the FRT site

and hygromycin resistance gene (Thermo Fisher Scientific), with cloned isoform NFIX+7 or NFIX-7 under the tetracycline regulated promoter.

Construction of stable cell lines that Express one of two NFIX splicing isoforms

The stable cell lines were generated using Flp-In expression system and commercially available Flp-In™ T-REx™ 293 host cell line according to the manufacturer's instructions (Thermo Fisher Scientific) in the background of the *NFIX*-KO. The positive clones which contain NFIX sequences incorporated into the genome were selected with hygromycin. Hygromycin (Thermo Fisher Scientific) in a final concentration of 50 µg/ml was added to the cells 48h after transfection, the selection lasted for 3 weeks. The plasmid for NFIX overexpression contains the hygromycin resistance gene sequence that is deprived of codon start and promoter. Thus, the cells acquire long-term resistance to a certain antibiotic only if the plasmid is integrated into the genome of cells. The desired cell line was expanded from one cell to obtain a genetically homogenous population.

Cell culture and transfection

The Flp-In™ T-REx™ 293 cells were grown in a high glucose DMEM medium with L-glutamine (Biowest) supplemented with 10% foetal bovine serum (Thermo Fisher Scientific), 1% antibiotic/antimycotic (Life Technologies) at 37 °C in 5% CO₂. Human Skeletal Myoblast (HSkM) cells, purchased from Gibco (Thermo Fisher), were grown in HAM F-10 medium (Lonza) supplemented with 20% FBS (BioWest), 2 mM l-glutamine, 100 U/ml penicillin, and 100 µg/ml streptomycin (Invitrogen) at 37 °C, 5% CO₂. Then transfer to a 2% horse serum-containing medium to induce differentiation and fusion. The doxycycline was added to induce gene expression in 250-1000 ng/ml concentration and cells were harvested after 48 h. For transfection cells were plated in 6-well plates or 96-well plates (for Luciferase activity assay) and transfected at ~80% confluency using Lipofectamine 3000 (Thermo Fisher Scientific)

according to the manufacturer's instructions. The total amount of DNA added to the cells was 2 µg/ml of cell culture medium in a 1:1 ratio in the case of two vectors with sgRNA sequences and a 9:1 ratio of pOGG4 and particular pcDNA5/FRT/TO plasmids, respectively. pcDNA5/FRT/TO and pmirGLO were added to cells in 4:1 ratio. The transfection with siRNA against 3'UTR fragment of *NFIX* sequence or control siRNA (synthesized by Sigma-Aldrich) and AON targeted exon complementary to the 5' splice sites of exon 7 and AON CTRL (synthesized by Genomed) was performed using Lipofectamine™ RNAiMAX (Thermo Fisher Scientific) at 50 nM for siRNA or 100 nM concentration for AONs (sequence of siRNA duplexes and AONs are specified in Table 2). After 24 h post-transfection, the medium was replaced to induce differentiation of HSkM myoblasts. After 4 days the cells were harvested.

RT-PCR splicing assays and qPCR gene expression analyses

The cells were harvested using TRIzol™ Reagent (Thermo Fisher Scientific) and total RNA was isolated using Total RNA Zol-Out™ D (A&A Biotechnology) according to the manufacturer's protocol. The cDNA was synthesized using High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific) with Random Primers (Thermo Fisher Scientific) according to the manufacturer's protocol. PCR was performed using GoTaq DNA Polymerase (Promega). PCR products were separated in 1-2% agarose gel with ethidium bromide. The images were captured using G:Box EF2 (Syngene) and analyzed using GeneTools (Syngene). PSI was calculated based on signals of two bands, corresponding to PCR product containing or missing alternative exons, according to the following formula (isoform with included exon*100)/(isoforms with included exon+excluded exon). Real-time RT-qPCRs were performed in a QuantStudio™ 7 Flex System (Thermo Fisher Scientific) using Maxima SYBR Green/ROX qPCR Master Mix (Thermo Fisher Scientific) according to the manufactures' instructions. Targets were amplified with primers listed in Table 3 at 58–60°C annealing temperature. Ct values were normalized against *GAPDH* or *ACTN*. Fold differences in

expression level were calculated according to the $2^{-\Delta\Delta Ct}$ method (Livak & Schmittgen, 2001).

All used primers are listed in Table 3

Western blot

Cells were lysed with RIPA buffer [150 mM NaCl, 50 mM Tris-HCl pH 8.0, 1 mM ethylenediaminetetraacetic acid (EDTA), 0.5% NP-40, 0.5% Triton X-100, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate (SDS)] supplemented with Halt™ Protease Inhibitor Cocktail (Thermo Fisher Scientific). Lysates were incubated on ice and vortexed followed by centrifugation at $15,000 \times g$ at 4°C for 15 min. Concentration of protein in cell extracts were measured by Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific). Samples were heated standard sample buffer. 95°C for 5 min. Electrophoresis and wet transfer were performed with the use of the Mini-PROTEAN Tetra System (Bio-Rad). 25-40 µg of protein extracts were separated by Bolt 4-12% Bis-Tris Plus (Thermo Fisher Scientific) and transferred to nitrocellulose membrane (Sigma-Aldrich) (1 h, 100 V) in Laemmli buffer with 20% methanol. Membranes were blocked for 1 h in 5% Skim Milk Powder (Sigma-Aldrich) in TBST buffer [(Tris-buffered saline (TBS), 0.1% Tween-20]. Membranes were incubated with a primary antibody against FLAG (A8592, Sigma-Aldrich) 1:1,000, GAPDH (sc-47724, Santa Cruz), NFIX (HPA007533, Atlas Antibodies) 1:1000 or Vinculin (sc-73614 HRP, SantaCruz Biotechnology) 1:2500, in 5% Skim Milk Powder in TBST for 1 h at RT. Membranes were washed in TBST and incubated with secondary antibodies conjugated with horseradish peroxidase, anti-mouse (A9044, Sigma-Aldrich) 1:20,000 or anti-rabbit (A9169, Sigma-Aldrich) 1:10,000 in TBST for 1 h at RT. Membranes were again washed in TBST and detected using Immobilon Forte Western HRP substrate (Sigma-Aldrich). Images were captured using G:Box Chemi-XR5 (Syngene) and quantified using Multi Gauge 3.0 software (Fujifilm).

Fluorescence microscopy

For immunostaining HEK-WT and HEK-KO cells were fixed for 15 min. with 4% paraformaldehyde PFA 48 h post transfection and blocked for 1 h in 1% BSA diluted in PBS-Tween (0.1%; PBS-T). Incubation with Anti-NFIX/CTF antibody (ab231324) (1:100) was conducted O/N at 4°C in the blocking solution. Secondary goat anti-mouse FITC-labeled antibody (1:400; Jackson ImmunoResearch Laboratories) was applied for 1 h at RT in PBS-T. Images were taken with Axio Observer.Z1 microscope equipped with AxioCam MRm camera, filter set 09 or 10 (GFP) A-Plan 10×/0.25 Ph1 objective (Zeiss), and AxioVs40 module.

Luciferase activity assay

All measurements of luciferase activity were carried out after 24 h of transfection with a Dual-Luciferase Reporter Assay System (Promega) according to the manufacturer's manual with Infinity 200 PRO (Tecan) equipment. pmirGLO plasmids, except for the sequence of firefly luciferase, have a *Renilla* luciferase open reading frame under a control of different promoter, which is used as transfection and loading control. *Luc2* activity is shown as a ratio of signal from Firefly and Renilla.

RNA-seq and bioinformatics analysis

Total RNA was isolated with the use of Total RNA Zol-Out™ D (A&A Biotechnology) according to the manufacturer's instructions with DNase I treatment. RNA quantity was checked with the use of Qubit® RNA BR Assay Kit and Qubit Fluorometer (Thermo Fisher Scientific) according to manufacturer's instructions, RNA quality was checked with the use of RNA 6000 Nano Kit and 2100 Bioanalyzer System (Agilent) according to the manufacturer's instructions. Library preparation and total RNA sequencing (2×100 bp) were performed by CeGaT (Germany) with the use of 100 ng of RNA, TruSeq Stranded Total RNA kit, and NovaSeq 6000. Demultiplexing of the sequencing reads was performed with Illumina bcl2fastq

(2.19). Adapters were trimmed with Skewer (version 0.2.2) (Jiang et al., 2014). Reference *Mus musculus* (GRCm38, primary assembly) genome and annotations were downloaded from Ensembl (ver. 91 or 93). Quality and adapter trimming of short reads was performed using Trimmomatic 0.39 (Bolger et al., 2014). Short reads matching known rRNA sequences were removed using HISAT2 2.1.0 aligner (Kim et al., 2015). Read quality reports before and after quality filtering were prepared using FastQC 0.11.5 software (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Filtered reads were aligned to the reference genomes using STAR 2.7.1a algorithm (Dobin et al., 2013). Read mapping reports were created using Qualimap 2.2.2 software (Okonechnikov et al., 2016). RSEM 1.3.0 (RNA-Seq by Expectation Maximization (Li & Dewey, 2011) was used to quantify the expression values of genes. Hierarchical clustering of RNA-seq samples (Pearson correlation metric, centroid linkage) based on the expression values of all genes was performed using standard R 3.3.1 functions (R Core Team 2016) and variance stabilizing transformation provided by DESeq2 package (Love et al., 2014). Differential expression analysis between designated groups of samples was performed using voom+limma (Ritchie et al., 2015) pipeline with settings described in (Law et al., 2018). The edgeR 3.14.0 package was used for initial data filtering and normalization using the TMM method (Law et al., 2018). Voom function from the limma 3.28.21 package (Law et al., 2018) was employed to estimate the mean-variance relationship in the data. Afterward, linear modeling and empirical Bayes moderation were applied to identify differentially expressed genes between groups of interest as described in ref. (Law et al., 2018) *P* value was generated using moderated *t*-statistic and adjusted for multiple testing using Benjamini–Hochberg’s method (adjusted *p*-value, adj.P.Val). adj.P.Val threshold of 0.01 and fold change threshold of 1.1 were used during the analysis.

Gene ontology analysis

For gene ontology (GO) analysis, data from differential expression analysis from RNA-seq experiments of *NFIX*-KO, *NFIX*-OE HEK293 cells and siCTRL, siNFIX, AON CTRL or AON NFIX treated HSkM cells and microarray experiments of muscle biopsies from DM patients and healthy controls (Nakamori et al., 2013) were used. List of differentially expressed genes expressed on arbitrarily determined level (AveExpr > 1; AveExpr – average expression across all samples, in log₂ CPM) were prepared. The significantly upregulated or downregulated genes were identified based on adjusted *P* value (Adj.P.Val either ≤ 0.05 or ≤ 0.01 depends on experiments what was specified in the figure legends; Adj.P.Val – Benjamini–Hochberg false discovery rate adjusted *P* value). A reference list of genes was prepared from all genes expressed in HEK293 cells, HSkM cells or muscle biopsies from DM patients on an arbitrarily determined level (AveExpr > 1 CPM). The analysis was performed with the use of PANTHER14.1 program. The analysis type was PANTHER Overrepresentation Test, annotation version and the release date was GO database released 2022-03-15. Annotation datasets were GO biological process complete, GO molecular function complete, and biological process. The statistical test was Fisher's exact test with the use of Bonferroni correction for multiple testing.

Statistical analysis

Group data are expressed as the means ± standard deviation (SD). The statistical significance was determined by unpaired, two-tailed Student's *t*-test using Prism software version 8 (GraphPad): **P* < 0.05; ** *P* < 0.01; *** *P* < 0.001; ns, non-significant. All analyses based on at least three independent biological replicates (exceptions are indicated in the figure legends) and whole experiments were repeated at least twice to confirm obtained results

3.5 BIBLIOGRAPHY

- Aartsma-Rus, A. (2017). FDA Approval of Nusinersen for Spinal Muscular Atrophy Makes 2016 the Year of Splice Modulating Oligonucleotides. *Nucleic Acid Therapeutics*, 27(2), 67–69. <https://doi.org/10.1089/nat.2017.0665>
- Angelbello, A. J., Benhamou, R. I., Rzuczek, S. G., Choudhary, S., Tang, Z., Chen, J. L., Roy, M., Wang, K. W., Yildirim, I., Jun, A. S., Thornton, C. A., & Disney, M. D. (2021). A Small Molecule that Binds an RNA Repeat Expansion Stimulates Its Decay via the Exosome Complex. *Cell Chemical Biology*, 28(1), 34–45.e6. <https://doi.org/10.1016/j.chembiol.2020.10.007>
- Apt, D., Liu, Y., & Bernard, H.-U. (1994a). Cloning and functional analysis of spliced isoforms of human nuclear factor I-X: interference with transcriptional activation by NFI/CTF in a cell-type specific manner. *Nucleic Acids Research*, 22(19), 3825–3833. <https://doi.org/10.1093/nar/22.19.3825>
- Arambula, J. F., Ramisetty, S. R., Baranger, A. M., & Zimmerman, S. C. (2009). A simple ligand that selectively targets CUG trinucleotide repeats and inhibits MBNL protein binding. *Proceedings of the National Academy of Sciences*, 106(38), 16068–16073. <https://doi.org/10.1073/pnas.0901824106>
- Arandel, L., Matloka, M., Klein, A. F., Rau, F., Sureau, A., Ney, M., Cordier, A., Kondili, M., Polay-Espinoza, M., Naouar, N., Ferry, A., Lemaitre, M., Begard, S., Colin, M., Lamarre, C., Tran, H., Buée, L., Marie, J., Sergeant, N., & Furling, D. (2022). Reversal of RNA toxicity in myotonic dystrophy via a decoy RNA-binding protein with high affinity for expanded CUG repeats. *Nature Biomedical Engineering*. <https://doi.org/10.1038/s41551-021-00838-2>
- Aslanidis, C., Jansen, G., Amemiya, C., Shutler, G., Mahadevan, M., Tsilfidis, C., Chen, C., Alleman, J., Wormskamp, N. G. M., Vooijs, M., Buxton, J., Johnson, K., Smeets, H. J. M., Lennon, G. G., Carrano, A. v., Korneluk, R. G., Wieringa, B., & de Jong, P. J. (1992). Cloning of the essential myotonic dystrophy region and mapping of the putative defect. *Nature*, 355(6360), 548–551. <https://doi.org/10.1038/355548a0>
- 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., Chojnacki, J., Lucia, A., Pintos-Morell, G., Coll-Cantí, J., Martínez-Piñeiro, A., Ramos-Fransi, A., & Nogales-Gadea, G. (2020). Three-dimensional imaging in myotonic dystrophy type 1. *Neurology Genetics*, 6(4), e484. <https://doi.org/10.1212/NXG.0000000000000484>
- Bargiela, A., Cerro-Herreros, E., Fernandez-Costa, J. M., Vilchez, J. J., Llamusi, B., & Artero, R. (2015). Increased autophagy and apoptosis contribute to muscle atrophy in a myotonic dystrophy type 1 *Drosophila* model. *Disease Models & Mechanisms*, 8(7), 679–690. <https://doi.org/10.1242/dmm.018127>
- Bargiela, A., Sabater-Arcis, M., Espinosa-Espinosa, J., Zulaica, M., Lopez de Munain, A., & Artero, R. (2019). Increased Muscleblind levels by chloroquine treatment improve myotonic dystrophy type 1 phenotypes in in vitro and in vivo models. *Proceedings of the National Academy of Sciences*, 116(50), 25203–25213. <https://doi.org/10.1073/pnas.1820297116>
- Batra, R., Manchanda, M., & Swanson, M. S. (2015). Global insights into alternative polyadenylation regulation. *RNA Biology*, 12(6), 597–602. <https://doi.org/10.1080/15476286.2015.1040974>

- Biressi, S., Molinaro, M., & Cossu, G. (2007). Cellular heterogeneity during vertebrate skeletal muscle development. *Developmental Biology*, *308*(2), 281–293. <https://doi.org/10.1016/j.ydbio.2007.06.006>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Braun, M., Shoshani, S., & Tabach, Y. (2022). Transcriptome changes in DM1 patients' tissues are governed by the RNA interference pathway. *Frontiers in Molecular Biosciences*, *9*. <https://doi.org/10.3389/fmolb.2022.955753>
- 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., Sohn, R., Zeman, B., Snell, R. G., Rundle, S. A., Crow, S., Davies, J., Shelbourne, P., Buxton, J., Jones, C., ... Housman, D. E. (1992). Molecular basis of myotonic dystrophy: Expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell*, *68*(4), 799–808. [https://doi.org/10.1016/0092-8674\(92\)90154-5](https://doi.org/10.1016/0092-8674(92)90154-5)
- Campbell, C. E., Piper, M., Plachez, C., Yeh, Y. T., Baizer, J. S., Osinski, J. M., Litwack, E. D., Richards, L. J., & Gronostajski, R. M. (2008a). The transcription factor Nfix is essential for normal brain development. *BMC Developmental Biology*, *8*, 1–18. <https://doi.org/10.1186/1471-213X-8-52>
- Cerro-Herreros, E., González-Martínez, I., Moreno-Cervera, N., Overby, S., Pérez-Alonso, M., Llamusi, B., & Artero, R. (2020). Therapeutic Potential of AntagomiR-23b for Treating Myotonic Dystrophy. *Molecular Therapy - Nucleic Acids*, *21*, 837–849. <https://doi.org/10.1016/j.omtn.2020.07.021>
- Cerro-Herreros, E., Sabater-Arcis, M., Fernandez-Costa, J. M., Moreno, N., Perez-Alonso, M., Llamusi, B., & Artero, R. (2018). miR-23b and miR-218 silencing increase Muscleblind-like expression and alleviate myotonic dystrophy phenotypes in mammalian models. *Nature Communications*, *9*(1), 2482. <https://doi.org/10.1038/s41467-018-04892-4>
- Chen, G., Masuda, A., Konishi, H., Ohkawara, B., Ito, M., Kinoshita, M., Kiyama, H., Matsuura, T., & Ohno, K. (2016). Phenylbutazone induces expression of MBNL1 and suppresses formation of MBNL1-CUG RNA foci in a mouse model of myotonic dystrophy. *Scientific Reports*, *6*(1), 25317. <https://doi.org/10.1038/srep25317>
- Crooke, S. T., Lemonidis, K. M., Neilson, L., Griffey, R., Lesnik, E. A., & Monia, B. P. (1995). Kinetic characteristics of *Escherichia coli* RNase H1: cleavage of various antisense oligonucleotide-RNA duplexes. *Biochemical Journal*, *312*(2), 599–608. <https://doi.org/10.1042/bj3120599>
- Cui, X., Dong, Y., Li, M., Wang, X., Jiang, M., Yang, W., Liu, G., Sun, S., & Xu, W. (2020). A circular RNA from NFIX facilitates oxidative stress-induced H9c2 cells apoptosis. *In Vitro Cellular and Developmental Biology - Animal*, *56*(9), 715–722. <https://doi.org/10.1007/s11626-020-00476-z>
- Davila, R. A., Spiller, C., Harkins, D., Harvey, T., Jordan, P. W., Gronostajski, R. M., Piper, M., & Bowles, J. (2022). Deletion of NFIX results in defective progression through meiosis within the

- mouse testis. *Biology of Reproduction*, 106(6), 1191–1205.
<https://doi.org/10.1093/biolre/ioac049>
- De Serres-Bérard, T., Pierre, M., Chahine, M., & Puymirat, J. (2021). Deciphering the mechanisms underlying brain alterations and cognitive impairment in congenital myotonic dystrophy. *Neurobiology of Disease*, 160, 105532. <https://doi.org/10.1016/j.nbd.2021.105532>
- Dixon, C., Harvey, T. J., Smith, A. G., Gronostajski, R. M., Bailey, T. L., & Piper, M. (2013). Nuclear Factor One X Regulates Bobby Sox During Development of the Mouse Forebrain. *Cellular and Molecular Neurobiology*, 33(7), 867–873. <https://doi.org/10.1007/s10571-013-9961-4>
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1), 15–21. <https://doi.org/10.1093/bioinformatics/bts635>
- Driller, K., Pagenstecher, A., Uhl, M., Omran, H., Berlis, A., Gründer, A., & Sippel, A. E. (2007). Nuclear Factor I X Deficiency Causes Brain Malformation and Severe Skeletal Defects. *Molecular and Cellular Biology*, 27(10), 3855–3867. <https://doi.org/10.1128/MCB.02293-06>
- Du, H., Cline, M. S., Osborne, R. J., Tuttle, D. L., Clark, T. A., Donohue, J. P., Hall, M. P., Shiue, L., Swanson, M. S., Thornton, C. A., & Ares, M. (2010a). Aberrant alternative splicing and extracellular matrix gene expression in mouse models of myotonic dystrophy. *Nature Structural & Molecular Biology*, 17(2), 187–193. <https://doi.org/10.1038/nsmb.1720>
- Fardaei, M. (2002). Three proteins, MBNL, MBLL and MBXL, co-localize in vivo with nuclear foci of expanded-repeat transcripts in DM1 and DM2 cells. *Human Molecular Genetics*, 11(7), 805–814. <https://doi.org/10.1093/hmg/11.7.805>
- Fernandez-Costa, J. M., Llamusi, M. B., Garcia-Lopez, A., & Artero, R. (2011). Alternative splicing regulation by Muscleblind proteins: from development to disease. *Biological Reviews*, 86(4), 947–958. <https://doi.org/10.1111/j.1469-185X.2011.00180.x>
- Fortune, M. T. (2000). Dramatic, expansion-biased, age-dependent, tissue-specific somatic mosaicism in a transgenic mouse model of triplet repeat instability. *Human Molecular Genetics*, 9(3), 439–445. <https://doi.org/10.1093/hmg/9.3.439>
- Fraser, J., Essebier, A., Brown, A. S., Davila, R. A., Harkins, D., Zalucki, O., Shapiro, L. P., Penzes, P., Wainwright, B. J., Scott, M. P., Gronostajski, R. M., Bodén, M., Piper, M., & Harvey, T. J. (2020a). Common Regulatory Targets of NFIA, NFIX and NFIB during Postnatal Cerebellar Development. *Cerebellum*, 19(1), 89–101. <https://doi.org/10.1007/s12311-019-01089-3>
- Fugier, C., Klein, A. F., Hammer, C., Vassilopoulos, S., Ivarsson, Y., Toussaint, A., Tosch, V., Vignaud, A., Ferry, A., Messaddeq, N., Kokunai, Y., Tsuburaya, R., de la Grange, P., Dembele, D., Francois, V., Precigout, G., Boulade-Ladame, C., Hummel, M.-C., de Munain, A. L., ... Charlet-Berguerand, N. (2011). Misregulated alternative splicing of BIN1 is associated with T tubule alterations and muscle weakness in myotonic dystrophy. *Nature Medicine*, 17(6), 720–725. <https://doi.org/10.1038/nm.2374>
- Goers, E. S., Purcell, J., Voelker, R. B., Gates, D. P., & Berglund, J. A. (2010). MBNL1 binds GC motifs embedded in pyrimidines to regulate alternative splicing. *Nucleic Acids Research*, 38(7), 2467–2484. <https://doi.org/10.1093/nar/gkp1209>

- Göös, H., Kinnunen, M., Salokas, K., Tan, Z., Liu, X., Yadav, L., Zhang, Q., Wei, G.-H., & Varjosalo, M. (2022). Human transcription factor protein interaction networks. *Nature Communications*, *13*(1), 766. <https://doi.org/10.1038/s41467-022-28341-5>
- Greco, S., Cardinali, B., Falcone, G., & Martelli, F. (2018). Circular RNAs in Muscle Function and Disease. *International Journal of Molecular Sciences*, *19*(11), 3454. <https://doi.org/10.3390/ijms19113454>
- Gronostajski, R. M. (2000). Roles of the NFI/CTF gene family in transcription and development. *Gene*, *249*(1–2), 31–45. [https://doi.org/10.1016/S0378-1119\(00\)00140-2](https://doi.org/10.1016/S0378-1119(00)00140-2)
- Harkins, D., Harvey, T. J., Atterton, C., Miller, I., Currey, L., Oishi, S., Kasherman, M., Davila, R. A., Harris, L., Green, K., Piper, H., Parton, R. G., Thor, S., Cooper, H. M., & Piper, M. (2022). Hydrocephalus in Nfix^{-/-} Mice Is Underpinned by Changes in Ependymal Cell Physiology. *Cells*, *11*(15), 2377. <https://doi.org/10.3390/cells11152377>
- Harris, L., Dixon, C., Cato, K., Heng, Y. H. E., Kurniawan, N. D., Ullmann, J. F. P., Janke, A. L., Gronostajski, R. M., Richards, L. J., Burne, T. H. J., & Piper, M. (2013). Heterozygosity for Nuclear Factor One X Affects Hippocampal-Dependent Behaviour in Mice. *PLoS ONE*, *8*(6), e65478. <https://doi.org/10.1371/journal.pone.0065478>
- Harris, L., Zalucki, O., Gobius, I., McDonald, H., Osinki, J., Harvey, T. J., Essebier, A., Vidovic, D., Gladwyn-Ng, I., Burne, T. H., Heng, J. I., Richards, L. J., Gronostajski, R. M., & Piper, M. (2016). Transcriptional regulation of intermediate progenitor cell generation during hippocampal development. *Development*, *143*(24), 4620–4630. <https://doi.org/10.1242/dev.140681>
- Heng, Y. H. E., McLeay, R. C., Harvey, T. J., Smith, A. G., Barry, G., Cato, K., Plachez, C., Little, E., Mason, S., Dixon, C., Gronostajski, R. M., Bailey, T. L., Richards, L. J., & Piper, M. (2014a). NFIX Regulates Neural Progenitor Cell Differentiation During Hippocampal Morphogenesis. *Cerebral Cortex*, *24*(1), 261–279. <https://doi.org/10.1093/cercor/bhs307>
- Heng, Y. H. E., Zhou, B., Harris, L., Harvey, T., Smith, A., Horne, E., Martynoga, B., Andersen, J., Achimastou, A., Cato, K., Richards, L. J., Gronostajski, R. M., Yeo, G. S., Guillemot, F., Bailey, T. L., & Piper, M. (2015). NFIX Regulates Proliferation and Migration Within the Murine SVZ Neurogenic Niche. *Cerebral Cortex*, *25*(10), 3758–3778. <https://doi.org/10.1093/cercor/bhu253>
- Ho, T. H., Charlet-B, N., Poulos, M. G., Singh, G., Swanson, M. S., & Cooper, T. A. (2004). Muscleblind proteins regulate alternative splicing. *The EMBO Journal*, *23*(15), 3103–3112. <https://doi.org/10.1038/sj.emboj.7600300>
- Holmfeldt, P., Pardieck, J., Saulsberry, A. C., Nandakumar, S. K., Finkelstein, D., Gray, J. T., Persons, D. A., & McKinney-Freeman, S. (2013). Nfix is a novel regulator of murine hematopoietic stem and progenitor cell survival. *Blood*, *122*(17), 2987–2996. <https://doi.org/10.1182/blood-2013-04-493973>
- Huang, S., Li, X., Zheng, H., Si, X., Li, B., Wei, G., Li, C., Chen, Y., Chen, Y., Liao, W., Liao, Y., & Bin, J. (2019). Loss of Super-Enhancer-Regulated circRNA Nfix Induces Cardiac Regeneration after Myocardial Infarction in Adult Mice. *Circulation*, *139*(25), 2857–2876. <https://doi.org/10.1161/CIRCULATIONAHA.118.038361>

- Imbriano, C., & Molinari, S. (2018). Alternative Splicing of Transcription Factors Genes in Muscle Physiology and Pathology. *Genes*, *9*(2), 107. <https://doi.org/10.3390/genes9020107>
- Jenquin, J. R., Coonrod, L. A., Silverglate, Q. A., Pellitier, N. A., Hale, M. A., Xia, G., Nakamori, M., & Berglund, J. A. (2018). Furamidine Rescues Myotonic Dystrophy Type I Associated Mis-Splicing through Multiple Mechanisms. *ACS Chemical Biology*, *13*(9), 2708–2718. <https://doi.org/10.1021/acscchembio.8b00646>
- Jenquin, J. R., Yang, H., Huigens III, R. W., Nakamori, M., & Berglund, J. A. (2019). Combination Treatment of Erythromycin and Furamidine Provides Additive and Synergistic Rescue of Mis-splicing in Myotonic Dystrophy Type 1 Models. *ACS Pharmacology & Translational Science*, *2*(4), 247–263. <https://doi.org/10.1021/acspsci.9b00020>
- Jiang, H., Lei, R., Ding, S.-W., & Zhu, S. (2014). Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics*, *15*(1), 182. <https://doi.org/10.1186/1471-2105-15-182>
- Kanadia, R. N., Shin, J., Yuan, Y., Beattie, S. G., Wheeler, T. M., Thornton, C. A., & Swanson, M. S. (2006). Reversal of RNA missplicing and myotonia after muscleblind overexpression in a mouse poly(CUG) model for myotonic dystrophy. *Proceedings of the National Academy of Sciences*, *103*(31), 11748–11753. <https://doi.org/10.1073/pnas.0604970103>
- Kanadia, R. N., Urbinati, C. R., Crusselle, V. J., Luo, D., Lee, Y.-J., Harrison, J. K., Oh, S. P., & Swanson, M. S. (2003). Developmental expression of mouse muscleblind genes Mbnl1, Mbnl2 and Mbnl3. *Gene Expression Patterns*, *3*(4), 459–462. [https://doi.org/10.1016/S1567-133X\(03\)00064-4](https://doi.org/10.1016/S1567-133X(03)00064-4)
- Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nature Methods*, *12*(4), 357–360. <https://doi.org/10.1038/nmeth.3317>
- Kimura, T., Nakamori, M., Lueck, J. D., Pouliquin, P., Aoike, F., Fujimura, H., Dirksen, R. T., Takahashi, M. P., Dulhunty, A. F., & Sakoda, S. (2005). Altered mRNA splicing of the skeletal muscle ryanodine receptor and sarcoplasmic/endoplasmic reticulum Ca²⁺-ATPase in myotonic dystrophy type 1. *Human Molecular Genetics*, *14*(15), 2189–2200. <https://doi.org/10.1093/hmg/ddi223>
- Kole, R., Krainer, A. R., & Altman, S. (2012). RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nature Reviews Drug Discovery*, *11*(2), 125–140. <https://doi.org/10.1038/nrd3625>
- Konieczny, P., Stepniak-Konieczna, E., & Sobczak, K. (2014). MBNL proteins and their target RNAs, interaction and splicing regulation. *Nucleic Acids Research*, *42*(17), 10873–10887. <https://doi.org/10.1093/nar/gku767>
- Kooblall, K., Stevenson, M., Stewart, M., Szoke-Kovacs, Z., Hough, T., Leng, H., Horwood, N., Vincent, T., Hennekam, R., Potter, P., Cox, R., Brown, S., Wells, S., Teboul, L., & Thakker, R. (2019). A mouse model generated by CRISPR-Cas9 with a frameshift mutation in the nuclear factor 1/X (*NFIX*) gene has phenotypic features reported in Marshall-Smith Syndrome (MSS) patients. *Endocrine Abstracts*. <https://doi.org/10.1530/endoabs.65.OC3.1>

- Kruse, U., & Sippel, A. E. (1994). Transcription factor nuclear factor I proteins form stable homo- and heterodimers. *FEBS Letters*, *348*(1), 46–50. [https://doi.org/10.1016/0014-5793\(94\)00585-0](https://doi.org/10.1016/0014-5793(94)00585-0)
- Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., ... Morgan, M. J. (2001). Initial sequencing and analysis of the human genome. *Nature*, *409*(6822), 860–921. <https://doi.org/10.1038/35057062>
- Lanni, S., & Pearson, C. E. (2019). Molecular genetics of congenital myotonic dystrophy. *Neurobiology of Disease*, *132*, 104533. <https://doi.org/10.1016/j.nbd.2019.104533>
- Law, C. W., Alhamdoosh, M., Su, S., Dong, X., Tian, L., Smyth, G. K., & Ritchie, M. E. (2018). RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. *F1000Research*, *5*, 1408. <https://doi.org/10.12688/f1000research.9005.3>
- Li, B., & Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, *12*(1), 323. <https://doi.org/10.1186/1471-2105-12-323>
- Lin, X., Miller, J. W., Mankodi, A., Kanadia, R. N., Yuan, Y., Moxley, R. T., Swanson, M. S., & Thornton, C. A. (2006). Failure of MBNL1-dependent post-natal splicing transitions in myotonic dystrophy. *Human Molecular Genetics*, *15*(13), 2087–2097. <https://doi.org/10.1093/hmg/ddl132>
- Liquori, C. L., Ricker, K., Moseley, M. L., Jacobsen, J. F., Kress, W., Naylor, S. L., Day, J. W., & Ranum, L. P. W. (2001). Myotonic Dystrophy Type 2 Caused by a CCTG Expansion in Intron 1 of ZNF9. *Science*, *293*(5531), 864–867. <https://doi.org/10.1126/science.1062125>
- Liu, Z., Ge, R., Zhou, J., Yang, X., Cheng, K. K., Tao, J., Wu, D., & Mao, J. (2020). Nuclear factor IX promotes glioblastoma development through transcriptional activation of Ezrin. *Oncogenesis*, *9*(4), 39. <https://doi.org/10.1038/s41389-020-0223-2>
- Livak, K. J., & Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods*, *25*(4), 402–408. <https://doi.org/10.1006/meth.2001.1262>
- López-Martínez, A., Soblechero-Martín, P., de-la-Puente-Ovejero, L., Nogales-Gadea, G., & Arechavala-Gomez, V. (2020). An Overview of Alternative Splicing Defects Implicated in Myotonic Dystrophy Type I. *Genes*, *11*(9), 1109. <https://doi.org/10.3390/genes11091109>
- Loro, E., Rinaldi, F., Malena, A., Masiero, E., Novelli, G., Angelini, C., Romeo, V., Sandri, M., Botta, A., & Vergani, L. (2010). Normal myogenesis and increased apoptosis in myotonic dystrophy type-1 muscle cells. *Cell Death & Differentiation*, *17*(8), 1315–1324. <https://doi.org/10.1038/cdd.2010.33>
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. <https://doi.org/10.1186/s13059-014-0550-8>
- Lueck, J. D., Mankodi, A., Swanson, M. S., Thornton, C. A., & Dirksen, R. T. (2007). Muscle Chloride Channel Dysfunction in Two Mouse Models of Myotonic Dystrophy. *Journal of General Physiology*, *129*(1), 79–94. <https://doi.org/10.1085/jgp.200609635>

- Malan, V., Rajan, D., Thomas, S., Shaw, A. C., Louis dit Picard, H., Layet, V., Till, M., van Haeringen, A., Mortier, G., Nampoothiri, S., Pušeljčić, S., Legeai-Mallet, L., Carter, N. P., Vekemans, M., Munnich, A., Hennekam, R. C., Colleaux, L., & Cormier-Daire, V. (2010). Distinct Effects of Allelic NFIX Mutations on Nonsense-Mediated mRNA Decay Engender Either a Sotos-like or a Marshall-Smith Syndrome. *The American Journal of Human Genetics*, *87*(2), 189–198. <https://doi.org/10.1016/j.ajhg.2010.07.001>
- Mankodi, A., Takahashi, M. P., Jiang, H., Beck, C. L., Bowers, W. J., Moxley, R. T., Cannon, S. C., & Thornton, C. A. (2002). Expanded CUG Repeats Trigger Aberrant Splicing of ClC-1 Chloride Channel Pre-mRNA and Hyperexcitability of Skeletal Muscle in Myotonic Dystrophy. *Molecular Cell*, *10*(1), 35–44. [https://doi.org/10.1016/S1097-2765\(02\)00563-4](https://doi.org/10.1016/S1097-2765(02)00563-4)
- Martinez, F., Marín-Reina, P., Sanchis-Calvo, A., Perez-Aytés, A., Oltra, S., Roselló, M., Mayo, S., Monfort, S., Pantoja, J., & Orellana, C. (2015). Novel mutations of NFIX gene causing Marshall-Smith syndrome or Sotos-like syndrome: one gene, two phenotypes. *Pediatric Research*, *78*(5), 533–539. <https://doi.org/10.1038/pr.2015.135>
- Martorell, L., Monckton, D. G., Sanchez, A., Lopez de Munain, A., & Baiget, M. (2001). Frequency and stability of the myotonic dystrophy type 1 premutation. *Neurology*, *56*(3), 328–335. <https://doi.org/10.1212/WNL.56.3.328>
- Messina, G., Biressi, S., Monteverde, S., Magli, A., Cassano, M., Perani, L., Roncaglia, E., Tagliafico, E., Starnes, L., Campbell, C. E., Grossi, M., Goldhamer, D. J., Gronostajski, R. M., & Cossu, G. (2010). Nfix Regulates Fetal-Specific Transcription in Developing Skeletal Muscle. *Cell*, *140*(4), 554–566. <https://doi.org/10.1016/j.cell.2010.01.027>
- Mirkin, S. M. (2007). Expandable DNA repeats and human disease. *Nature*, *447*(7147), 932–940. <https://doi.org/10.1038/nature05977>
- Monckton, D. G., Wong, L.-J. C., 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>
- Nakamori, M., Kimura, T., Kubota, T., Matsumura, T., Sumi, H., Fujimura, H., Takahashi, M. P., & Sakoda, S. (2008). Aberrantly spliced -dystrobrevin alters -syntrophin binding in myotonic dystrophy type 1. *Neurology*, *70*(9), 677–685. <https://doi.org/10.1212/01.wnl.0000302174.08951.cf>
- Nakamori, M., Sobczak, K., Moxley, R. T., & Thornton, C. A. (2009). Scaled-down genetic analysis of myotonic dystrophy type 1 and type 2. *Neuromuscular Disorders*, *19*(11), 759–762. <https://doi.org/10.1016/j.nmd.2009.07.012>
- Nakamori, M., Sobczak, K., Puwanant, A., Welle, S., Eichinger, K., Pandya, S., Dekdebrun, J., Heatwole, C. R., McDermott, M. P., Chen, T., Cline, M., Tawil, R., Osborne, R. J., Wheeler, T. M., Swanson, M. S., Moxley, R. T., & Thornton, C. A. (2013a). Splicing biomarkers of disease severity in myotonic dystrophy. *Annals of Neurology*, *74*(6), 862–872. <https://doi.org/10.1002/ana.23992>

- Nakamori, M., Taylor, K., Mochizuki, H., Sobczak, K., & Takahashi, M. P. (2016a). Oral administration of erythromycin decreases <sc>RNA</sc> toxicity in myotonic dystrophy. *Annals of Clinical and Translational Neurology*, 3(1), 42–54. <https://doi.org/10.1002/acn3.271>
- Nakamura, T., Ohsawa-Yoshida, N., Zhao, Y., Koebis, M., Oana, K., Mitsuhashi, H., & Ishiura, S. (2016). Splicing of human chloride channel 1. *Biochemistry and Biophysics Reports*, 5, 63–69. <https://doi.org/10.1016/j.bbrep.2015.11.006>
- Okonechnikov, K., Conesa, A., & García-Alcalde, F. (2016). Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics*, 32(2), 292–294. <https://doi.org/10.1093/bioinformatics/btv566>
- Otero, B. A., Poukalov, K., Hildebrandt, R. P., Thornton, C. A., Jinnai, K., Fujimura, H., Kimura, T., Hagerman, K. A., Sampson, J. B., Day, J. W., & Wang, E. T. (2021). Transcriptome alterations in myotonic dystrophy frontal cortex. *Cell Reports*, 34(3), 108634. <https://doi.org/10.1016/j.celrep.2020.108634>
- Pérez-Casellas, L. A., Wang, X., Howard, K. D., Rehage, M. W., Strong, D. D., & Linkhart, T. A. (2009). Nuclear Factor I transcription factors regulate IGF binding protein 5 gene transcription in human osteoblasts. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1789(2), 78–87. <https://doi.org/10.1016/j.bbagr.2008.08.013>
- Piper, M., Barry, G., Harvey, T. J., McLeay, R., Smith, A. G., Harris, L., Mason, S., Stringer, B. W., Day, B. W., Wray, N. R., Gronostajski, R. M., Bailey, T. L., Boyd, A. W., & Richards, L. J. (2014). NFIB-Mediated Repression of the Epigenetic Factor Ezh2 Regulates Cortical Development. *Journal of Neuroscience*, 34(8), 2921–2930. <https://doi.org/10.1523/JNEUROSCI.2319-13.2014>
- Piper, M., Barry, G., Hawkins, J., Mason, S., Lindwall, C., Little, E., Sarkar, A., Smith, A. G., Moldrich, R. X., Boyle, G. M., Tole, S., Gronostajski, R. M., Bailey, T. L., & Richards, L. J. (2010). NFIA Controls Telencephalic Progenitor Cell Differentiation through Repression of the Notch Effector Hes1. *Journal of Neuroscience*, 30(27), 9127–9139. <https://doi.org/10.1523/JNEUROSCI.6167-09.2010>
- Piper, M., Gronostajski, R., & Messina, G. (2019). Nuclear Factor One X in Development and Disease. In *Trends in Cell Biology* (Vol. 29, Issue 1). <https://doi.org/10.1016/j.tcb.2018.09.003>
- Piper, M., Harris, L., Barry, G., Heng, Y. H. E., Plachez, C., Gronostajski, R. M., & Richards, L. J. (2011a). Nuclear factor one X regulates the development of multiple cellular populations in the postnatal cerebellum. *The Journal of Comparative Neurology*, 519(17), 3532–3548. <https://doi.org/10.1002/cne.22721>
- Pistocchi, A., Gaudenzi, G., Foglia, E., Monteverde, S., Moreno-Fortuny, A., Pianca, A., Cossu, G., Cotelli, F., & Messina, G. (2013). Conserved and divergent functions of Nfix in skeletal muscle development during vertebrate evolution. *Development*, 140(7), 1528–1536. <https://doi.org/10.1242/dev.076315>
- Pjanic, M., Pjanic, P., Schmid, C., Ambrosini, G., Gaussin, A., Plasari, G., Mazza, C., Bucher, P., & Mermod, N. (2011). Nuclear factor I revealed as family of promoter binding transcription activators. *BMC Genomics*, 12. <https://doi.org/10.1186/1471-2164-12-181>

- Pjanic, M., Schmid, C. D., Gaussin, A., Ambrosini, G., Adamcik, J., Pjanic, P., Plasari, G., Kerschgens, J., Dietler, G., Bucher, P., & Mermod, N. (2013). Nuclear Factor I genomic binding associates with chromatin boundaries. *BMC Genomics*, *14*(1), 99. <https://doi.org/10.1186/1471-2164-14-99>
- Qawasmi, L., Braun, M., Guberman, I., Cohen, E., Naddaf, L., Mellul, A., Matilainen, O., Roitenberg, N., Share, D., Stupp, D., Chahine, H., Cohen, E., Garcia, S. M. D. A., & Tabach, Y. (2019). Expanded CUG Repeats Trigger Disease Phenotype and Expression Changes through the RNAi Machinery in *C. elegans*. *Journal of Molecular Biology*, *431*(9), 1711–1728. <https://doi.org/10.1016/j.jmb.2019.03.003>
- Ranum, L. P. W., & Cooper, T. A. (2006). RNA-MEDIATED NEUROMUSCULAR DISORDERS. *Annual Review of Neuroscience*, *29*(1), 259–277. <https://doi.org/10.1146/annurev.neuro.29.051605.113014>
- Rau, F., Lainé, J., Ramanoudjame, L., Ferry, A., Arandel, L., Delalande, O., Jollet, A., Dingli, F., Lee, K.-Y., Peccate, C., Lorain, S., Kabashi, E., Athanasopoulos, T., Koo, T., Loew, D., Swanson, M. S., le Rumeur, E., Dickson, G., Allamand, V., ... Furling, D. (2015). Abnormal splicing switch of DMD's penultimate exon compromises muscle fibre maintenance in myotonic dystrophy. *Nature Communications*, *6*(1), 7205. <https://doi.org/10.1038/ncomms8205>
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, *43*(7), e47–e47. <https://doi.org/10.1093/nar/gkv007>
- Rossi, G., Antonini, S., Bonfanti, C., Monteverde, S., Vezzali, C., Tajbakhsh, S., Cossu, G., & Messina, G. (2016a). Nfix Regulates Temporal Progression of Muscle Regeneration through Modulation of Myostatin Expression. *Cell Reports*, *14*(9). <https://doi.org/10.1016/j.celrep.2016.02.014>
- Rossi, G., Bonfanti, C., Antonini, S., Bastoni, M., Monteverde, S., Innocenzi, A., Saclier, M., Taglietti, V., & Messina, G. (2017). Silencing Nfix rescues muscular dystrophy by delaying muscle regeneration. *Nature Communications*, *8*(1), 1055. <https://doi.org/10.1038/s41467-017-01098-y>
- Savkur, R. S., Philips, A. v., & Cooper, T. A. (2001). Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy. *Nature Genetics*, *29*(1), 40–47. <https://doi.org/10.1038/ng704>
- Siboni, R. B., Bodner, M. J., Khalifa, M. M., Docter, A. G., Choi, J. Y., Nakamori, M., Haley, M. M., & Berglund, J. A. (2015). Biological Efficacy and Toxicity of Diamidines in Myotonic Dystrophy Type 1 Models. *Journal of Medicinal Chemistry*, *58*(15), 5770–5780. <https://doi.org/10.1021/acs.jmedchem.5b00356>
- Singh, S. K., Wilczynska, K. M., Grzybowski, A., Yester, J., Osrah, B., Bryan, L., Wright, S., Griswold-Prenner, I., & Kordula, T. (2011). The Unique Transcriptional Activation Domain of Nuclear Factor-I-X3 Is Critical to Specifically Induce Marker Gene Expression in Astrocytes. *Journal of Biological Chemistry*, *286*(9), 7315–7326. <https://doi.org/10.1074/jbc.M110.152421>
- Tabaglio, T., Low, D. H., Teo, W. K. L., Goy, P. A., Cywoniuk, P., Wollmann, H., Ho, J., Tan, D., Aw, J., Pavesi, A., Sobczak, K., Wee, D. K. B., & Guccione, E. (2018). MBNL1 alternative

- splicing isoforms play opposing roles in cancer. *Life Science Alliance*, 1(5), e201800157. <https://doi.org/10.26508/lsa.201800157>
- Taglietti, V., Angelini, G., Mura, G., Bonfanti, C., Caruso, E., Monteverde, S., le Carrou, G., Tajbakhsh, S., Relaix, F., & Messina, G. (2018). RhoA and ERK signalling regulate the expression of the myogenic transcription factor Nfix. *Development*. <https://doi.org/10.1242/dev.163956>
- Taglietti, V., Maroli, G., Cermenati, S., Monteverde, S., Ferrante, A., Rossi, G., Cossu, G., Beltrame, M., & Messina, G. (2016). Nfix Induces a Switch in Sox6 Transcriptional Activity to Regulate MyHC-I Expression in Fetal Muscle. *Cell Reports*, 17(9), 2354–2366. <https://doi.org/10.1016/j.celrep.2016.10.082>
- Teplova, M., & Patel, D. J. (2008). Structural insights into RNA recognition by the alternative-splicing regulator muscleblind-like MBNL1. *Nature Structural & Molecular Biology*, 15(12), 1343–1351. <https://doi.org/10.1038/nsmb.1519>
- Thornton, C. A., Johnson, K., & Moxley, R. T. (1994). Myotonic dystrophy patients have larger CTG expansions in skeletal muscle than in leukocytes. *Annals of Neurology*, 35(1), 104–107. <https://doi.org/10.1002/ana.410350116>
- TIAN, B., WHITE, R. J., XIA, T., WELLE, S., TURNER, D. H., MATHEWS, M. B., & THORNTON, C. A. (2000). Expanded CUG repeat RNAs form hairpins that activate the double-stranded RNA-dependent protein kinase PKR. *RNA*, 6(1), S1355838200991544. <https://doi.org/10.1017/S1355838200991544>
- Todd, P. K., Ackall, F. Y., Hur, J., Sharma, K., Paulson, H. L., & Dowling, J. J. (2013). Transcriptional changes and developmental abnormalities in a zebrafish model of myotonic dystrophy type 1. *Disease Models & Mechanisms*. <https://doi.org/10.1242/dmm.012427>
- Todorow, V., Hintze, S., Kerr, A. R. W., Hehr, A., Schoser, B., & Meinke, P. (2021). Transcriptome Analysis in a Primary Human Muscle Cell Differentiation Model for Myotonic Dystrophy Type 1. *International Journal of Molecular Sciences*, 22(16), 8607. <https://doi.org/10.3390/ijms22168607>
- Tóth, G., Gáspári, Z., & Jurka, J. (2000). Microsatellites in Different Eukaryotic Genomes: Survey and Analysis. *Genome Research*, 10(7), 967–981. <https://doi.org/10.1101/gr.10.7.967>
- Vidovic, D., Harris, L., Harvey, T. J., Evelyn Heng, Y. H., Smith, A. G., Osinski, J., Hughes, J., Thomas, P., Gronostajski, R. M., Bailey, T. L., & Piper, M. (2015a). Expansion of the lateral ventricles and ependymal deficits underlie the hydrocephalus evident in mice lacking the transcription factor NFIX. *Brain Research*, 1616, 71–87. <https://doi.org/10.1016/j.brainres.2015.04.057>
- 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., Lécuyer, E., & Burge, C. B. (2012). Transcriptome-wide Regulation of Pre-mRNA Splicing and mRNA Localization by Muscleblind Proteins. *Cell*, 150(4), 710–724. <https://doi.org/10.1016/j.cell.2012.06.041>
- Wang, E. T., Treacy, D., Eichinger, K., Struck, A., Estabrook, J., Olafson, H., Wang, T. T., Bhatt, K., Westbrook, T., Sedehizadeh, S., Ward, A., Day, J., Brook, D., Berglund, J. A., Cooper, T.,

- Housman, D., Thornton, C., & Burge, C. (2019). Transcriptome alterations in myotonic dystrophy skeletal muscle and heart. *Human Molecular Genetics*, 28(8), 1312–1321. <https://doi.org/10.1093/hmg/ddy432>
- Wang, W., Mullikin-Kilpatrick, D., Crandall, J. E., Gronostajski, R. M., Litwack, E. D., & Kilpatrick, D. L. (2007). Nuclear factor I coordinates multiple phases of cerebellar granule cell development via regulation of cell adhesion molecules. *Journal of Neuroscience*, 27(23), 6115–6127. <https://doi.org/10.1523/JNEUROSCI.0180-07.2007>
- Wang, W., Stock, R. E., Gronostajski, R. M., Wong, Y. W., Schachner, M., & Kilpatrick, D. L. (2004). A Role for Nuclear Factor I in the Intrinsic Control of Cerebellar Granule Neuron Gene Expression. *Journal of Biological Chemistry*, 279(51), 53491–53497. <https://doi.org/10.1074/jbc.M410370200>
- Wenstrom, K. D. (2002). Fragile X and other trinucleotide repeat diseases. *Obstetrics and Gynecology Clinics of North America*, 29(2), 367–388. [https://doi.org/10.1016/S0889-8545\(01\)00005-5](https://doi.org/10.1016/S0889-8545(01)00005-5)
- Wheeler, T. M., Sobczak, K., Lueck, J. D., Osborne, R. J., Lin, X., Dirksen, R. T., & Thornton, C. A. (2009a). Reversal of RNA Dominance by Displacement of Protein Sequestered on Triplet Repeat RNA. *Science*, 325(5938), 336–339. <https://doi.org/10.1126/science.1173110>
- Wheeler, T. M., & Thornton, C. A. (2007). Myotonic dystrophy: RNA-mediated muscle disease. *Current Opinion in Neurology*, 20(5), 572–576. <https://doi.org/10.1097/WCO.0b013e3282ef6064>
- Wong, C.-H., Nguyen, L., Peh, J., Luu, L. M., Sanchez, J. S., Richardson, S. L., Tuccinardi, T., Tsoi, H., Chan, W. Y., Chan, H. Y. E., Baranger, A. M., Hergenrother, P. J., & Zimmerman, S. C. (2014). Targeting Toxic RNAs that Cause Myotonic Dystrophy Type 1 (DM1) with a Bisamidinium Inhibitor. *Journal of the American Chemical Society*, 136(17), 6355–6361. <https://doi.org/10.1021/ja5012146>
- Xiao, D., Caldow, M., Kim, H. J., Blazej, R., Koopman, R., Manandi, D., Parker, B. L., & Yang, P. (2022). Time-resolved phosphoproteome and proteome analysis reveals kinase signaling on master transcription factors during myogenesis. *IScience*, 25(6), 104489. <https://doi.org/10.1016/j.isci.2022.104489>
- Yamashita, Y., Matsuura, T., Shinmi, J., Amakusa, Y., Masuda, A., Ito, M., Kinoshita, M., Furuya, H., Abe, K., Ibi, T., Sahashi, K., & Ohno, K. (2012). Four parameters increase the sensitivity and specificity of the exon array analysis and disclose 25 novel aberrantly spliced exons in myotonic dystrophy. *Journal of Human Genetics*, 57(6), 368–374. <https://doi.org/10.1038/jhg.2012.37>
- Zatz, M., Passos-Bueno, M. R., Cerqueira, A., Marie, S. K., Vainzof, M., & Pavanello, R. C. M. (1995). Analysis of the CTG repeat in skeletal muscle of young and adult myotonic dystrophy patients: when does the expansion occur? *Human Molecular Genetics*, 4(3), 401–406. <https://doi.org/10.1093/hmg/4.3.401>
- Zhang, X., Zhou, Y., Pan, C., Lei, C., Dang, R., Chen, H., & Lan, X. (2015). Novel alternative splice variants of NFIX and their diverse mRNA expression patterns in dairy goat. *Gene*, 569(2), 250–258. <https://doi.org/10.1016/j.gene.2015.05.062>

- Zhou, B., Osinski, J. M., Mateo, J. L., Martynoga, B., Sim, F. J., Campbell, C. E., Guillemot, F., Piper, M., & Gronostajski, R. M. (2015). Loss of NFIX Transcription Factor Biases Postnatal Neural Stem/Progenitor Cells Toward Oligodendrogenesis. *Stem Cells and Development*, 24(18), 2114–2126. <https://doi.org/10.1089/scd.2015.0136>
- Zhou, Y., Cai, H., Xu, Y., Sun, J., Lan, X., Lei, C., & Chen, H. (2014). Novel isoforms of the bovine *Nuclear factor I/X (CCAAT-binding transcription factor)* transcript products and their diverse expression profiles. *Animal Genetics*, 45(4), 581–584. <https://doi.org/10.1111/age.12177>

4. SUSTAINABLE RECOVERY OF MBNL ACTIVITY IN AUTOREGULATORY FEEDBACK LOOP IN MYOTONIC DYSTROPHY

Sustainable recovery of MBNL activity in autoregulatory feedback loop in myotonic dystrophy

Zuzanna Rogalska¹ and Krzysztof Sobczak¹

¹Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Uniwersytetu Poznańskiego 6, 61-614 Poznań, Poland

Muscleblind-like proteins (MBNLs) are RNA-binding proteins essential for the developmental regulation of various processes including alternative splicing. Their activity is misregulated in myotonic dystrophy type 1 (DM1), an incurable genetic, neuromuscular disorder caused by uncontrolled expansion of CTG repeats. Mutant RNAs containing hundreds or thousands of repeats efficiently sequester MBNL proteins. As a consequence, global alternative splicing abnormalities are induced. Importantly, the size of expansion differs significantly not only between patients but also between different parts of the same muscle as a consequence of somatic expansion. One of the potential therapeutic strategies in DM is overexpression of MBNLs. However, gene therapy tools might induce excessive activity of MBNLs, what in turn might change the metabolism of many RNAs. To overcome these limitations, we designed an autoregulated MBNL1 overexpression system. The genetic construct contains an MBNL1-coding sequence separated by the fragment of *ATP2A1* pre-mRNA with an MBNL-sensitive alternative exon containing stop codon in the reading frame of MBNL1. Inclusion of this exon leads to the arrangement of an inactive form of the protein, but exclusion gives rise to fully active MBNL1. This approach enables the autoregulation of the amount of overexpressed MBNL1 with high dynamic range which ensures a homogeneous level of this protein in cells treated with the genetic construct. We demonstrated beneficial effects of an autoregulated construct on alternative splicing patterns in DM1 models and cells derived from patients with DM1.

INTRODUCTION

Three *Muscleblind*-like genes (*MBNLs*) encode for RNA-binding proteins essential for regulation of various processes of RNA metabolism including alternative splicing, polyadenylation, RNA localization, and stability.^{1–3} MBNL1 is mainly expressed in skeletal muscles, MBNL2 in brain, and MBNL3 in placenta and during muscle cell differentiation, where they regulate hundreds of alternative splicing events leading to the expression of adult-specific mRNA isoforms.^{4–6} All family members bind to RNA via four zinc fingers (ZFs) organized in two tandems connected by a long linker. MBNLs recognize their targets through closely organized multiple 5'-YGCY-3' sequence motifs, where Y represents a pyrimidine.^{7,8} The location of the MBNL-bind-

ing sites on target RNAs determines if the alternative exon is included or excluded. The presence of the preferred sequence motif upstream of or within the exon leads to the exon exclusion while its presence downstream facilitates exon inclusion.^{3,8,9}

Activity of MBNLs is significantly misregulated in many diseases, including myotonic dystrophy (DM). DM1 and DM2 are autosomal-dominant genetic multisystemic disorders with symptoms mainly affecting skeletal and cardiac muscles and the central nervous system.¹⁰ They are caused by uncontrolled expansion of either CTG in the 3' untranslated region (3' UTR) of the dystrophin protein kinase (*DMPK*) gene or CCTG repeats within the intron of the cellular nucleic-acid-binding protein (*CNBP*) gene, respectively.^{11,12} RNA containing expanded CUG (CUG^{exp}) or CCUG (CCUG^{exp}) form thermodynamically stable long hairpin structures that efficiently bind and sequester MBNL proteins from nucleoplasm and, together, form nuclear foci.^{13,14} Efficiency of sequestration depends on the length of CUG^{exp} and CCUG^{exp}, which is significantly variable in different patients, ranging from hundreds to thousands of repeats.¹¹ As a consequence, different severities of global abnormalities of alternative splicing and alternative polyadenylation are observed, which is further associated with disease severity and progression.^{8,11,15,16} Generally, the activity of MBNLs in skeletal muscles and heart is significantly lowered to the level observed in newborns. Moreover, the length of CUG^{exp} in DM1 significantly differs in different tissues, which is a consequence of somatic repeat expansion occurring during the lifespan of patients.^{17,18} More extensive somatic expansion is observed in heart and skeletal muscles than in blood and many other tissues.^{18–20} The length of repeats can be different even in nuclei of the same muscle fiber, which reflects somatic mosaicism of CUG^{exp}.^{21,22} Moreover, the length of CCUG^{exp} may be also deeply heterogeneous in the same patient, as indicated by recent studies showing variability in repeat length and repeat composition.^{23,24}

Received 18 March 2022; accepted 31 October 2022;
<https://doi.org/10.1016/j.omtn.2022.10.023>

Correspondence: Krzysztof Sobczak, Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Uniwersytetu Poznańskiego 6, 61-614 Poznań, Poland.

E-mail: ksobczak@amu.edu.pl

Hence, the sequestration of MBNL proteins is highly heterogeneous in both DM forms, even in the same tissue of the same patient.

So far, there are no cures for these diseases; however, a few therapeutic approaches were tested in cell and animal models of DM, and some of them demonstrated rescue of disease phenotype. The therapeutic strategies for DM1 may be grouped into three main categories: (1) induction of degradation of mutant RNA containing CUG^{exp}, (2) the release of MBNLs from pathogenic sequestration, and (3) the increase of MBNL expression. The second category involves the application of CUG^{exp}-specific small molecules or antisense oligonucleotides (ASOs), which can inhibit the formation of the pathogenic RNA-MBNL complex and lead to the improvement of alternative splicing defects.^{25–28}

The third mentioned category involves either manipulation of an endogenous pool of MBNLs, e.g., the use of non-steroidal anti-inflammatory drugs that modify activity of *MBNL1* promoter,²⁹ or antagonomiRs targeting *miR-23b* or *miR-218* to increase translation from *MBNL1* and *MBNL2* mRNAs,²⁸ or the application of gene therapy tools for MBNL overexpression.

Proof of concept for gene therapy was described for the DM1 mouse model, *HSA-LR*, expressing a transgene with 220 CTG repeats in 3' UTR of human skeletal actin gene (*HSA*). The MBNL1-encoding transgene was delivered to the *HSA-LR* mouse via transduction with adeno-associated virus (AAV).³⁰ Twenty-three weeks after intramuscular AAV injection, the MBNL1 overexpression rescued the splicing defects of many MBNL-sensitive genes. At the same time, muscle hyperexcitability was missing. Otherwise, the normal structure of myofibers was not recovered, suggesting that the overexpression was insufficient to rescue this phenotype.³⁰ More recently, it was shown that truncated MBNL1 overexpressed from the AAV vector, which preserved both ZFs tandems but is deprived of the C-terminal domain, binds to the CUG^{exp} with high affinity, leading to the release endogenous MBNL proteins from sequestration and then rescue of disease phenotypes in the DM1 mouse model.³¹ In other research, it was shown that the cross of *HSA-LR* with the mouse model with multisystemic overexpression of MBNL1 or overexpression restricted just to the skeletal muscles (*MBNL1-OE*) showed the rescue of DM-like defects, including the decrease of percentage of fibers with central nuclei reflecting myopathy changes. On the other hand, the long-term, multisystemic overexpression of MBNL1 in wild-type mice and another DM1 mouse model, DM200⁺, which utilizes the doxycycline-dependent promoter that controls the expression of a transgene with 200 CTG repeats in the 3' UTR of the *DMPK* gene, led to reduced body weight and increased mortality.^{32,33} Moreover, intramuscular injection of AAV encoding the full length of MBNL1 to the wild-type (WT) mice induced formation of muscle fibers with internal nuclei, which indicates muscle damage.³¹ These results suggest that there are some limitations in therapeutic strategies against DM based on the uncontrolled overexpression of MBNLs, especially considering the significant variability in the size and expression level of CUG^{exp} in different tissues of patients with DM.

To overcome the limitations caused by heterogeneity of CTG/CCTG repeat expansions and, consequently, different levels of MBNL insufficiency in different cells/myofibers, we designed and tested the autor-regulated MBNL1 overexpression constructs, which enable the significant production of MBNL1 only if its level in the cell is too low and potentially can be controlled by heterogeneity of the CUG^{exp} load. We demonstrated that the level of the protein assembling from the construct is homogeneous from cell to cell, is controlled by a pool of available MBNLs, and has therapeutic potential to correct the alternative splicing abnormalities in cellular models of DM1. Considering DM1 and DM2 acquire the same mechanism of MBNL sequestration and, in consequence, share similar missplicing events, the potential therapeutics may be suitable for both.³⁴

RESULTS

Constructs for autoregulated overexpression of MBNL1auto

For the autoregulated overexpression of MBNL1, we designed the hybrid genetic construct MB22#1, which contains an MBNL1-encoding sequence separated by a fragment of the *ATP2A1* gene containing MBNL-sensitive alternative exon 22 (ex22) and neighboring introns (Figures 1A and S1). We decided to choose ex22 because of its high sensitivity to MBNL regulation. In skeletal muscles, the isoform with the inclusion of ex22 predominates, but in affected muscles of patients with DM1, different levels of ex22 exclusion are observed (Figure 1B), which is correlated with disease severity.³⁵ The sequence of the *ATP2A1* gene fragment was incorporated between ex2 and ex3 of *MBNL1* so as to not disturb the structure of any of the ZF tandems. Alternative ex22 is positively regulated by all MBNL paralogs.³⁶ We hypothesized that the inclusion of ex22 may lead to the production of a truncated, inactive form of the protein because of the presence of in-frame stop codon (Figure 1A). Therefore, in cells with low levels of MBNLs, ex22 can be excluded during pre-mRNA maturation, and the fully active form of MBNL1 protein (MBNL1auto) can be assembled. We also designed two other constructs by replacing native MBNL-binding sequences in intron 22 with sequences showing different sensitivities to MBNLs.³⁷ First, MB22#2 contains the sequence with four consecutive 5'-UGCU motifs as an MBNL-binding site. Second, control MB22-del has a deletion of the native MBNL-binding site of *ATP2A1* (Figure 1A). Moreover, at the C terminus of the MBNL1auto sequence, either FLAG tag or GFP tag was added. We hypothesized that these constructs allow the adjustment of different amounts of MBNL1auto due to different efficiency of ex22 inclusion.

To test this hypothesis, we co-transfected COS7 cells with one of three generated MB22 constructs containing different MBNL-sensitive elements and either MBNL1-GFP (without [w/o] autoregulatory cassette) or as a control GFP expressing vector. The RT-PCR base splicing assay showed that the percentage of mRNA isoforms with the exclusion of ex22 depends on the presence of an MBNL-sensitive RNA regulatory element (MB22#1, -#2, or -del) and the level of the MBNL pool in cells. We decided to measure the ex22 exclusion, as this isoform can produce functional MBNL1auto. As expected, the exclusion rate of ex22 was higher for MB22#1 and MB22#2 constructs in cells with a basal pool of endogenous MBNLs and significantly

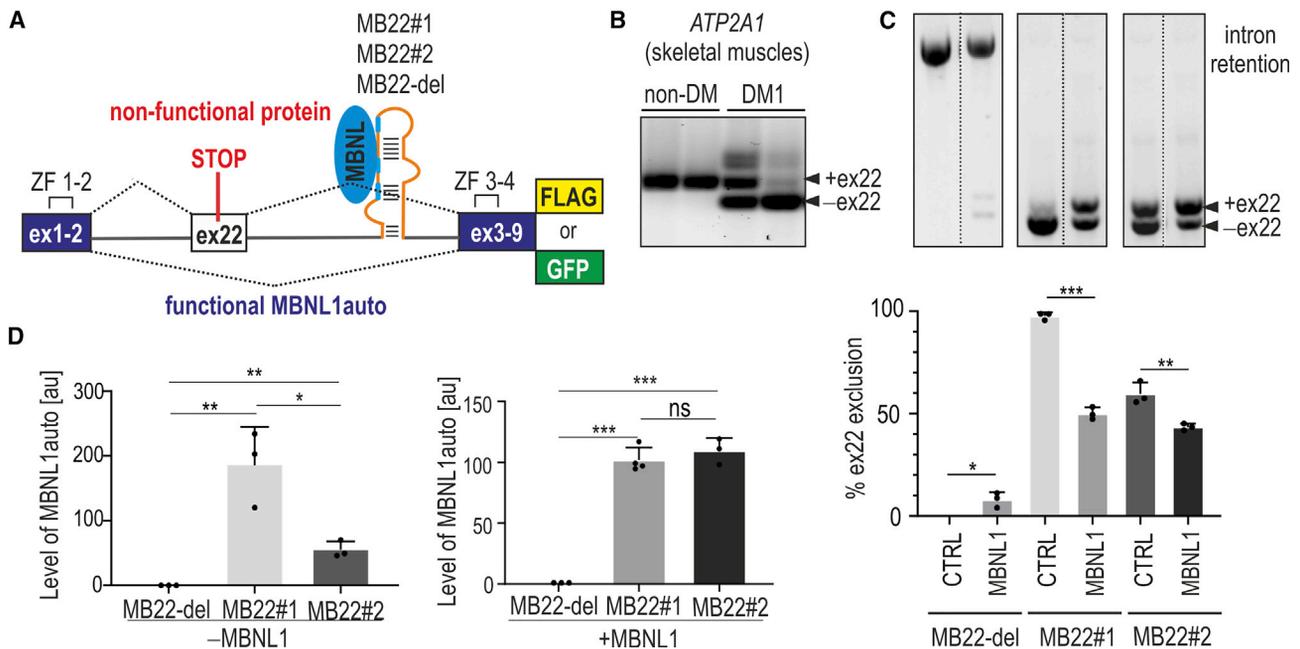


Figure 1. Inclusion of the alternative exon to mRNA encoding MBNL1auto depends on the MBNL-binding RNA regulatory motif

(A) The scheme of the genetic construct for autoregulated overexpression of MBNL1auto. The *MBNL1* coding sequence was divided into two parts separated by an intron21/exon (ex)22/intron22 sequence from human *ATP2A1*. Ex1-2 is a sequence of the first two exons of *MBNL1*, containing the first zinc finger tandem (ZF1–ZF2); ex3-9 is a sequence of cDNA of ex3–ex9 with ZF3 and ZF4. *ATP2A1* fragment contains the alternative ex22 flanked from both sides by introns and 7 bp of ex21 and 2 bp of ex23 to maintain correct splicing regulation and to keep the open reading frame of MBNL1auto. The inclusion of ex22, positively regulated by MBNLs, leads to premature translation termination and the arrangement of an inactive form of protein. The MB22 construct contains the wild-type sequence recognized by MBNLs (blue) within intron22 of *ATP2A1* (MB22#1). This sequence was replaced by 4xUGCU MBNL-binding motif (MB22#2) or completely removed (MB22-del; orange). GFP and FLAG are tags located in the frame of MBNL1auto. (B) Alternative splicing profile of *ATP2A1* ex22 in two normal adult skeletal muscles (non-DM) and two different skeletal muscles from patients with DM1 (DM1) analyzed by RT-PCR. Isoforms with and without ex22 are marked. (C) Results of RT-PCR analysis of ex22 exclusion in cells transfected with MB22 constructs containing different MBNL-sensitive elements (MB22#1, -#2, and -del) and with MBNL1-GFP (MBNL1) or GFP (CTRL) overexpression. The percentage of alternative ex22 exclusion reflects mRNA isoforms coding for the active form of the protein. Bars represent average from $n = 3$ independent experiments (dots) for each experimental condition with standard deviation (SD). (D) Results of western blot analysis showing the level of MBNL1auto in cells transfected with either MB22-del, -#1, or -#2 (left; –MBNL1) or co-transfected with these three constructs and MBNL1-GFP (right; +MBNL1). Anti-FLAG antibody staining was carried out as FLAG sequence is fused to the C-terminal end of MBNL1auto in each MB22 construct. Bars represent average signal (arbitrary units [a.u.]) from $n = 3–4$ independent experiments for each group normalized to mCherry. Co-transfection with mCherry expression vector was utilized as a normalization control of transfection. (B and C) Unpaired Student's *t* test; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; ns, non-significant.

decreased with MBNL1-GFP overexpression (Figure 1C). Unexpectedly, the level of mRNA with excluded ex22 from the control MB22-del construct is very low because the isoform with retention of both introns predominates.

Next, we analyzed the level of MBNL1auto in cells transfected with one of three autoregulated MB22 constructs (Figure 1D, left graph) or co-transfected with MB22s and non-autoregulated MBNL1-GFP (Figure 1D, right graph) using western blot. A higher amount of protein was produced in cells treated with MB22#1. On the other hand, a similar amount of MBNL1auto protein was detected in cells co-transfected with either MB22#1 or MB22#2 and MBNL1-GFP. These results are consistent with a RT-PCR-base splicing assay showing the percentage of mRNA isoforms with exclusion of ex22 (Figure 1C). The highest ex22 exclusion was in cells treated with MB22#1, suggesting the highest potential to produce the full length of MBNL1auto. Furthermore, we evaluated the total pool of MBNL1 after MB22 over-

expression. The western blot analysis showed about a 2-times increase of the MBNL1 pool compared with the endogenous level of this protein (Figure S2). MBNL1auto was not detected in cells transfected with MB22-del (Figure 1D), which corroborated the results of mRNA splicing analysis (Figure 1C).

All these experiments demonstrate that autoregulated MB22 constructs are capable of overexpressing different levels of MBNL1auto in cells with different pools of MBNL proteins.

The effect of a pool of MBNLs on production of MBNL1auto from the autoregulated MB22 construct

By design, the amount of produced MBNL1auto should also differ in cells with different endogenous MBNL pools. Therefore, we overexpressed MB22#1 in three cell lines with different levels of MBNL pools (total of MBNL1, -2, and -3): in the COS7 line and the HEK293 line, with relatively high and low endogenous pools of MBNLs,

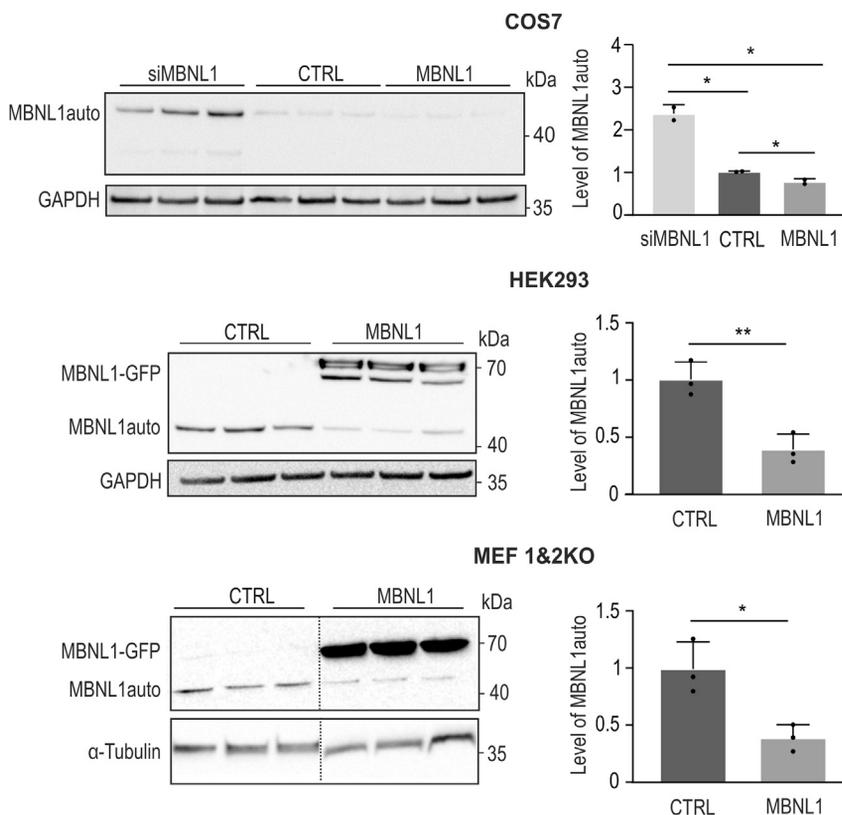


Figure 2. Biosynthesis of the MBNL1auto depends on the available pool of MBNLs

Western blot analysis of MBNL1auto in COS7, HEK-293, and MEF cells with *Mbnl1* and *Mbnl2* knockout. Bars represent average signal from $n = 3$ independent experiments for each group normalized to GAPDH in COS7 and HEK-293 cells and to α -tubulin in MEF cells. Anti-FLAG antibody staining was carried out as FLAG sequence is fused to the C-terminal end of both constructs. Unpaired Student's *t* test was used to calculate statistical significance: * $p < 0.05$; ** $p < 0.01$.

respectively, and mouse embryonic fibroblasts with full knockout of *Mbnl1* and *Mbnl2* (MEF-1&2KO).³⁸ These cells were co-transfected with either MBNL1-GFP to increase the pool of MBNLs or GFP used as a negative control (Figure 2). Overexpressed MBNL1-GFP, similar to endogenous MBNLs, can bind to the regulatory element within the intron of MB22#1 pre-mRNA, leading to ex22 inclusion and premature termination of translation (ex22 contains a stop codon), which decreases the level of MBNL1auto. Moreover, COS7 cells were also transfected with small interfering RNA (siRNA) against *MBNL1* (siMBNL1), targeting the sequence within the 3' UTR to not disturb expression of the MB22 construct. As expected, siMBNL1 treatment significantly increased the level of MBNL1auto by about 2.5 times. On the other hand, the overexpression of MBNL1-GFP has a negative effect on the level of protein from the MB22#1 construct in each tested cell line (Figure 2), leading to an about 3-fold decrease of MBNL1auto in HEK293 and MEF-1&2KO cells. Taken together, these results indicate that production of MBNL1auto strongly depends on the changing level of the MBNL pool in a broad range of MBNL concentrations.

The level of MBNL1auto is homogeneous in cells

To achieve insight into the ability of the MBNL1auto protein to autoregulate, we co-transfected COS7 with different amounts of construct encoding either MBNL1auto or MBNL1-GFP (100–1,000 ng) and minigenes containing MBNL-sensitive alternative exons (*MBNL1* ex5, *Nfix* ex7, *Atp2a1* ex22) (Figures 3A and 3B). This approach

gave us the comparison of MBNL1 proteins that came from either autoregulated or non-autoregulated constructs on depth and dynamics of alternative splicing regulation. Results of RT-PCR assay showed that the percentage of inclusion of ex22 from MB22#1 is indeed not sensitive to the amount of transfected construct (Figure 3A). Moreover, increasing amounts of transfected MB22#1 did not significantly differ in the regulation of three other alternative exons (Figure 3B). In contrast, MBNL1-GFP showed concentration-dependent splicing regulation, reaching saturation for the highest amount of overexpression construct (Figure 3B). A different pattern of ex22 inclusion from *Atp2a1* minigene does not demonstrate better activity of MBNL1-GFP than the MBNL1auto, as this minigene responds to low levels of MBNL1 overexpression. Collectively, these results suggest that expression of MBNL1auto can reach a certain maximal level independently from amount of MB22 genetic construct delivered into cells.

To directly measure the diversity of MBNL1auto-GFP production in MB22#2-GFP-transfected COS7 cells, we utilized the flow cytometry approach (Figure S3A). Quantification on resolution of a single cell performed in two independent experiments revealed that the level of MBNL1auto-GFP is more homogeneous (4-times difference in a GFP signal between the 25th and 75th percentiles of analyzed cells) than the level of MBNL1-GFP (13-times difference) which was used as a negative control (Figures 3C and S3B). We confirmed an equal level of MBNL1auto-GFP using confocal microscopy (Figure 3D). Taken together, these findings revealed that MBNL1auto-GFP shows homogeneous expression in transfected cells.

MBNL1auto corrects alternative splicing abnormalities in DM1 cell models

To investigate therapeutic potential of MB22, we first studied if the MBNL1auto-GFP protein can bind to the CUG^{exp} repeats. COS7 cells were co-transfected with a MB22#2 construct expressing MBNL1auto fused with GFP and either a construct expressing a mutant *DMKP* mRNA fragment containing ex11–ex15 with 960 interrupted CUG repeats in the 3' UTR (CUG₉₆₀) or a normal *DMPK* fragment without

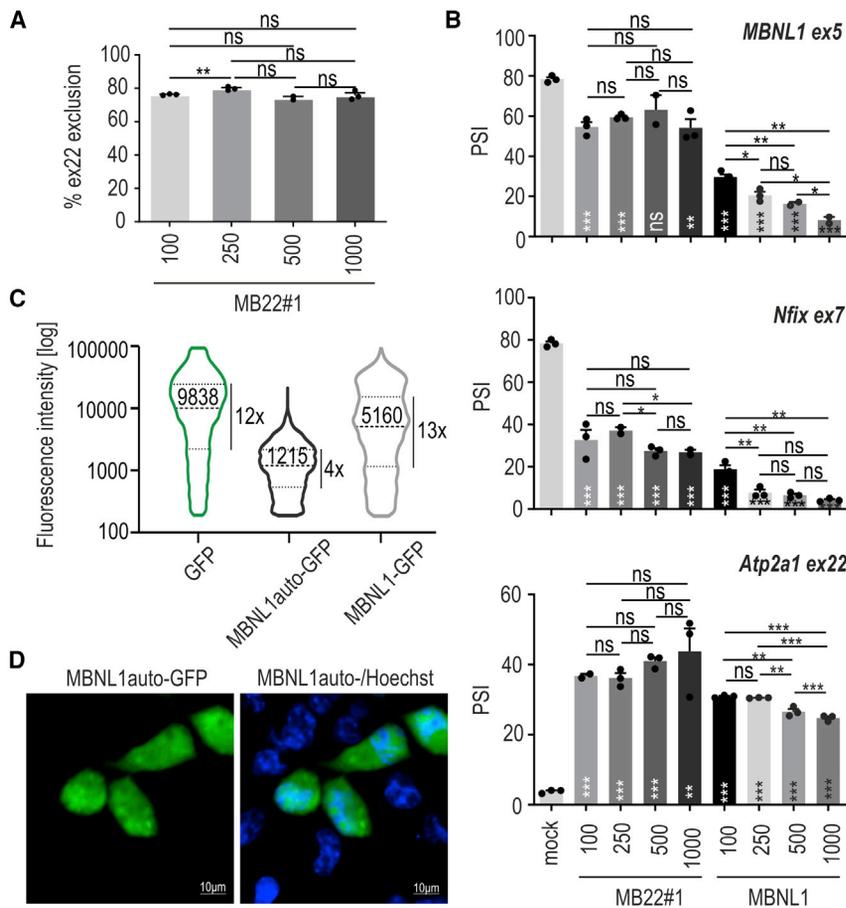


Figure 3. The autoregulatory potential and expression homogeneity of MB22 construct

(A) The percentage of alternative ex22 exclusion from mRNA encoding for MBNL1auto after treatment of COS7 cells with different amounts of MB22#1, ranging from 100 to 1,000 ng/mL culture medium; $n = 3$. (B) The dose-dependent inclusion of alternative exons of three different MBNL-sensitive minigenes after treatment of COS7 cells with MB22#1- or MBNL1-GFP-overexpression construct (ng/mL); $n = 3$. (A and B) Unpaired Student's t test; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; ns, non-significant. Stars placed on bars denote statistical significance compared with the mock. (C) Violin plots showing the distribution of cells with different fluorescent signals of either GFP, MBNL1auto-GFP, or MBNL1-GFP proteins. COS7 cells were transfected with adequate vectors 48 h prior to flow cytometry analysis. Median fluorescent intensity (black solid line) and 25th and 75th percentiles of signal (dashed lines) are shown. Fold change between 25th and 75th percentiles of signal for all analyses is also indicated. Cells with signal below 200 were rejected from analyses based on results for control experiment for mock-transfected cells. GFP and MBNL1-GFP without autoregulatory cassette state as negative controls. Graphs represent values from $n = 4$ independent biological replicates for each experimental condition; $N = 13,942$ (GFP), $N = 6,760$ (MBNL1auto-GFP), and $N = 10,793$ (MBNL1-GFP) cells. (D) Representative confocal microscopy image showing localization of MBNL1auto-GFP (green) in COS7 cells transfected with MB22#2-GFP. Nuclei were stained with Hoechst (blue); scale bar, 10 μm.

repeats (CUG₀). In cells with CUG₉₆₀, the fluorescent signal from MBNL1auto-GFP protein is located mostly in the ribonuclear foci. In contrast, diffuse distribution and, again, homogeneous GFP signals between cells were observed in a control experiment in cells with CUG₀ (Figure 4A). Moreover, we also performed RNA fluorescence *in situ* hybridization (FISH), which also confirmed co-localization of MBNL1auto-GFP with CUG^{exp}, using a probe labeled with Cy-3, which detects CUG^{exp} (Figure S4). These experiments showed that MBNL1auto, by binding to CUG^{exp} in cell nuclei, can replace endogenous MBNLs from sequestration.

Therefore, next we looked at the potential of MBNL1auto on the correction of MBNL-sensitive alternative splicing in two DM1 cell models. In the first, COS7 cells were co-transfected with MB22#1 and mutant or normal *DMPK*-expressing constructs (CUG₉₆₀ or CUG₀) together with two MBNL-sensitive minigenes (*Nfix* ex7, *Atp2a1* ex22). RT-PCR analyses showed that the presence of MBNL1auto leads to the partial, but significant, rescue of all tested DM1-specific alternative splicing events (Figure 4B). To further support these results, we also utilized other cellular models with siRNA-induced insufficiency of MBNLs (*siMBNL1*; as described above). Cells with silenced endogenous *MBNL1* were co-transfected with MB22#1

and the same splicing minigenes (*Nfix* ex7, *Atp2a1* ex22). We confirmed that *siMBNL1* efficiently knocked down *MBNL1* (Figure S5) and significantly increased the level of MBNL1auto (Figure 2). Then, using splicing-specific RT-PCR assays, we demonstrated that MBNL1auto rescued the pathogenic missplicing triggered by *MBNL1* deficiency (Figure 4C).

Lentiviral-based production of MBNL1auto corrects splicing abnormalities in cells derived from patients with DM1

The promising results obtained in the two DM models described above encouraged us to assess the therapeutic potential of MB22 in two different cells derived from patients with DM1 that expressed *DMPK* transcripts with very long CUG^{exp} (DM1-1 and DM1-2). Previously, we showed that these cells manifest DM1-specific molecular phenotypes like the formation of CUG^{exp} nuclear foci, sequestration of MBNLs, and missplicing of several transcripts.^{25,39} In this study, we utilized lentiviral vectors encoding MBNL1auto-GFP from autoregulatory MB22#2-GFP or GFP used as a negative control. The fusion with GFP was used to monitor the efficacy of transduction. The fibroblast from a healthy individual was used as a control cell line (non-DM). Twelve days after transduction, we checked the expression level of the *DMPK* gene to assess the potential effect of the production of

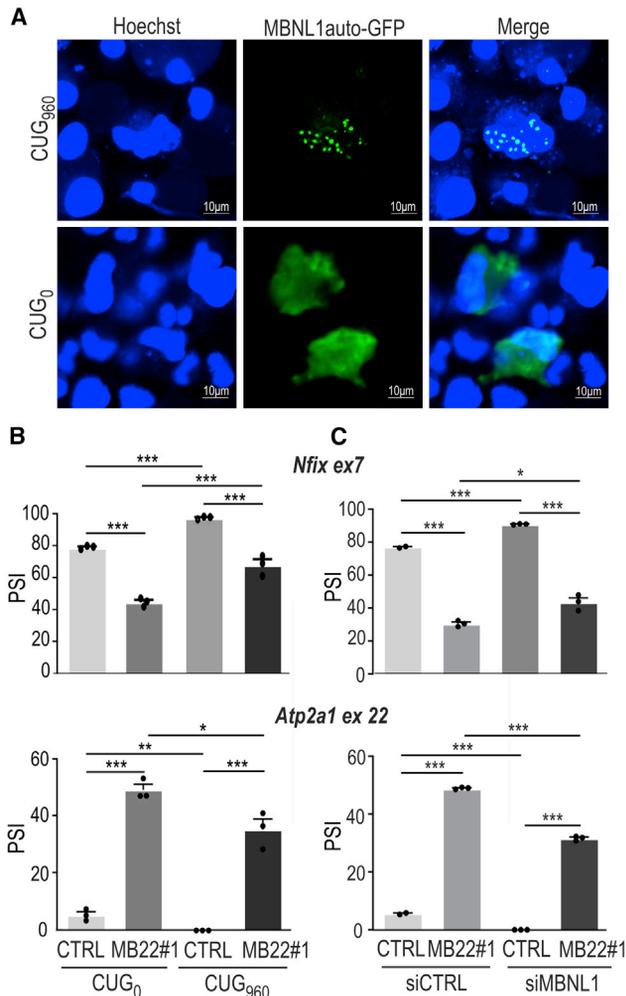


Figure 4. The therapeutic potential of MB22 autoregulated construct in DM models

(A) Representative confocal images showing the nuclear foci containing MBNL1auto-GFP in cells expressing a mutant *DMPK* fragment with CUG₉₆₀ (top panel). In cells expressing a normal *DMPK* fragment without CUG repeats, MBNL1auto-GFP is distributed equally (bottom panel); scale bar, 10 μ m. (B) Results of RT-PCR analyses showing changes in the regulation of two MBNL-dependent exons from *Nfix* ex7 and *Atp2a1* ex22 minigenes. Splicing changes are expressed as the percent spliced in (PSI). Cells were co-transfected with MB22#1 or control GFP construct (CTRL) and either mutant (CUG₉₆₀) or normal (CUG₀) *DMPK*-fragment-expressing constructs. (C) As in (B) but for cells treated with either control siRNA (siCTRL) or siRNA targeting the 3' UTR of *MBNL1* (siMBNL1). These cells were then transfected with either GFP or MB22#1 construct. The percentage of mRNA isoform with inclusion of alternative exon was calculated using the inverse of PSI parameter, which demonstrates the portion of mRNA with an included alternative exon. Bars represent average PSI from three independent experiments (with exception of n = 2 for siCTRL + GFP condition) with SD. (B and C) Unpaired Student's t test was used to calculate statistical significance. *p < 0.05; **p < 0.01; ***p < 0.001; ns, non-significant.

MBNL1auto. The quantitative real-time RT-PCR analysis showed no differences in the steady-state level of *DMPK* mRNA in samples treated with different vectors (Figure 5A). Next, we evaluated alterna-

tive splicing changes of eight mRNAs that are known to be MBNL sensitive: *INSR* ex11, *FLNB* ex31, *MYO5A* ex33, *MBNL2* ex5, *MBNL2* ex7, *MBNL1* ex1, *NCOR2* ex19, and *PHKA1* ex19. All of them showed significant correction in both DM1 cell lines (Figures 5B and S6). Importantly, no significant splicing changes were observed in the non-DM1 cell line treated with MB22#2-GFP lentivirus (the only exception is a small change of ex19 inclusion in *PHKA1*). Cumulatively, all these data strongly suggest that the MBNL1auto protein, whose production is autoregulated on the level of alternative splicing of its mRNA, has the ability to rescue alternative splicing alterations in different DM1 models.

DISCUSSION

Gene therapy is promising strategy for many incurable diseases. It involves the delivery of new genetic material to patient cells to prevent or slow down development of a particular disease. This kind of treatment enables the efficient cure for some monogenic disorders, including spinal muscular atrophy (Zolgensma).⁴⁰ Overexpression can rescue proper protein function but, when uncontrolled, may also trigger undesirable effects.^{41,42} Therefore, an important limitation to overcome in the gene therapy approach is to adjust the level of overexpressed protein to prevent toxicity but enable the therapeutic effect.

During different stages of tissue differentiation and development, the level of the MBNL pool is crucial for proper alternative splicing regulation. Expression of *MBNL1* and *MBNL2* increases during both embryonic and postnatal stages of development, with the highest level of MBNL1 in adult muscles and the highest MBNL2 in the adult brain.^{43,44} All three MBNL paralogs regulate the same splicing events; however, among them, fast and slow responders exist. Some alternative exons respond to low and some to high levels of MBNLs, and the depth of exon inclusion/exclusion could be regulated in a high range of concentration of these proteins.^{45,46} Therefore, fine-tuning of the MBNL pool is important for proper function of different tissues.

In DM, the activity of all MBNL paralogs is significantly lower due to sequestration of these proteins on toxic CUG^{exp} or CCUG^{exp},^{14,15} and the increase of MBNLs is considered a potential therapeutic strategy, including the application of gene therapy tools.⁴⁷ It was shown that AAV-based production of the MBNL1 protein or its truncated variant rescued muscle hyperexcitability and splicing defects of many MBNL-sensitive genes.^{30,31} Moreover, expression of *MBNL1* from a transgene in knockin mouse models also improved the DM-specific phenotypes.³² On the other hand, overexpression of full-length MBNL1 induced by intramuscular injection of AAV vector in WT mice showed significant muscle damage.³¹ Also, transgenic mice overexpressing MBNL1 from a transgene demonstrated significantly decreased body weight, grip strength, run distance, and heart failure compared with WT littermates³³ or even reduced survival.³² Moreover, uncontrolled and unbalanced overexpression might lead to the excessive activity of MBNLs in some treated muscle fibers or other tissues and, consequently, change the metabolism of many RNA.

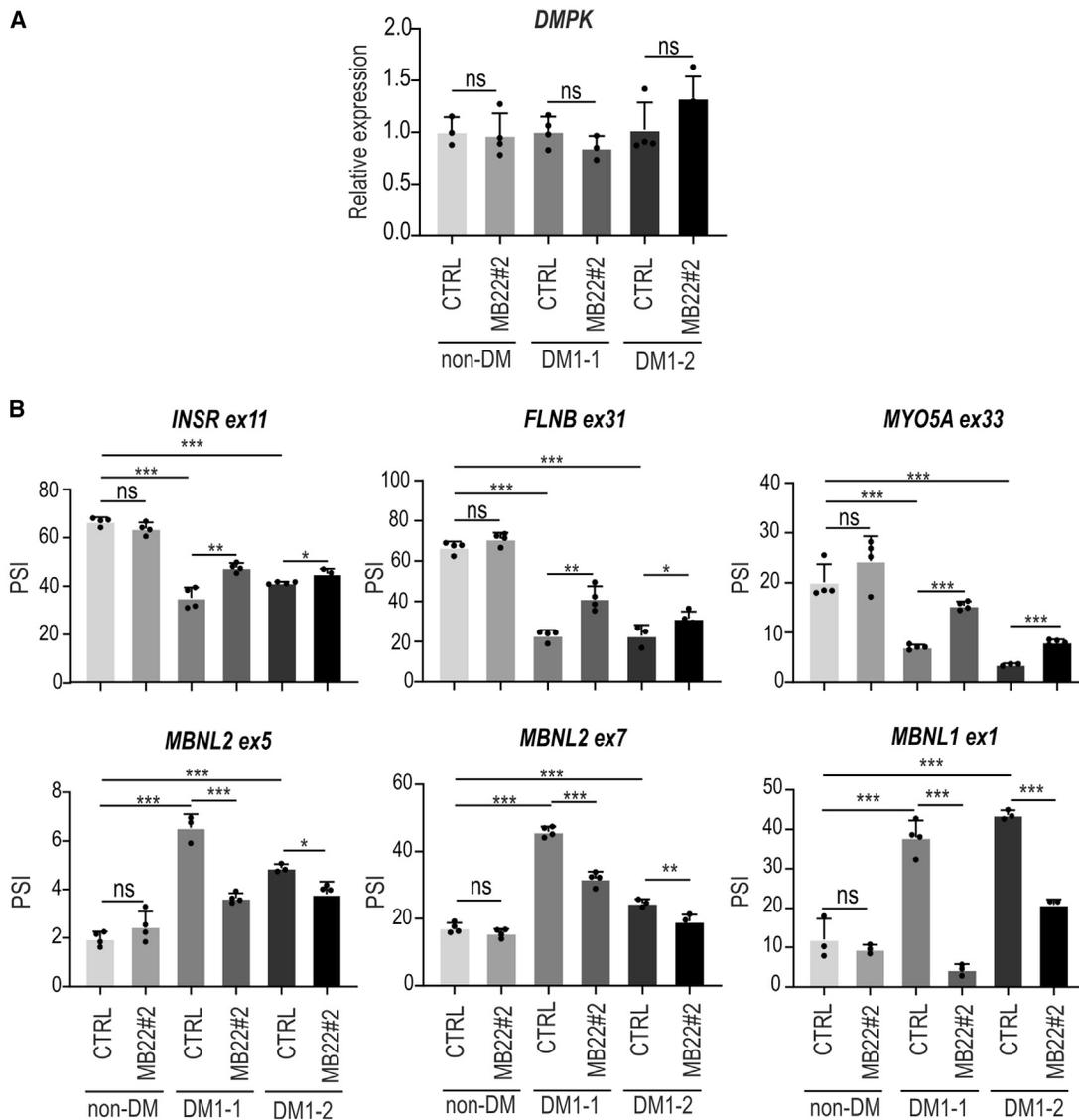


Figure 5. Correction of pathogenic missplicing in DM1 cells treated with MB22 lentiviruses

(A) Results of quantitative real-time RT-PCR analysis showing relative expression of *DMPK* (normalized to *GAPDH*) in three different cell lines: fibroblasts derived from healthy individual (non-DM) and two patients with DM1 (DM1-1 and DM1-2) treated with either control (CTRL) lentiviral vector or lentivirus containing the MB22#2-GFP sequence 12 days from cell transduction. (B) Results of RT-PCR-based analyses of alternative splicing changes in cells described in (A). Changes in the inclusion of positively (top panel) and negatively (bottom panel) regulated MBNL-dependent alternative exons are shown for six transcripts affected in DM1. Splicing changes are expressed as PSI. Bars represent average from 3 to 4 independent experiments (dots); unpaired Student's *t* test; **p* < 0.05; ***p* < 0.01; ****p* < 0.001; ns, non-significant.

Therefore, in this study we decided to develop the self-regulating MB22 overexpression construct for fine-tuning MBNL1 auto production, which depends on the endogenous MBNL pool available in the cell. In designing MBNL1 auto, we selected the sequence of the 41 kDa isoform of MBNL1. This isoform is deprived of alternative ex5, which is responsible for nuclear localization of the protein, but contains alternative ex7, which increases splicing activity of the protein.^{4,48,49} Importantly, this isoform is localized in both the nucleus and cytoplasm and can regulate not only alternative splicing but also other RNA metabolism stages³ and is one of major isoforms present in skel-

etal and cardiac muscles in adults.⁴⁶ It was already demonstrated that the MBNL1 isoform with ex7 can efficiently bind to CUG^{exp} and CCUG^{exp}.⁴⁶

Previous studies of ours and others showed that activity of MBNLs can be autoregulated on the level of alternative splicing.^{4,50} Therefore, we decided to design the autoregulatory MB22 construct, whose expression is sensitive to the MBNL pool, by adding the intron/ex22/intron sequence from *ATP2A1* between the sequence encoding MBNL1 (Figure 1A). MBNL-dependent alternative splicing of pre-mRNA from this

construct is regulated by the efficiency of the inclusion of alternative ex22 containing an in-frame stop codon, and the inclusion of this exon leads to the production of a non-functional short protein.

As proposed in this study, an autoregulated MB22 construct enables the restricted expression of MBNL1auto (Figure 2), which is sufficient for the rescue of DM-specific alternative splicing defects in cellular models with insufficiency of MBNLs induced by either silencing of *MBNL1* or overexpression of toxic CUG^{exp} (Figure 4).

In muscles of patients with DM1 and DM2, the size of CTG or CCTG repeat expansion differs significantly between patients but also between different parts of the same muscle of the same individual, due to somatic instability of repeat tract.^{12,18,21,22} Hence, the sequestration of MBNL proteins is highly heterogeneous. Therefore, we hypothesized that unequal sequestration of the MBNL pool can be buffered by self-regulated overexpression of MBNL1auto. This was partially demonstrated in experiments with transfection-based and uncontrolled delivery of the MB22-GFP construct to model cells (Figure 3). Using conventional overexpression systems, we can expect high heterogeneity in the number of transcripts and proteins generated from the transgene, e.g., as a result of different numbers of copies of plasmid per single cell. Quantitative flow cytometry experiments and microscopic analyses (Figures 3C, S3B, and S3D), as well as monitoring of MBNL-sensitive splicing patterns (Figure 3B), showed that the production of MBNL1auto is homogeneous from cell to cell even after imbalanced delivery of the MB22 genetic construct. Importantly, the buffering level of MBNL1auto leads to significant correction of splicing abnormalities in cellular models of DM1 (Figures 4 and 5).

The therapeutic effect would be caused partially by binding of MBNL1auto to CUG^{exp} and replacement of the endogenous pool of MBNLs from sequestration (Figures 4A and S4) and partially by the splicing activity of MBNL1auto itself. We can predict that the required level of MBNL1auto in skeletal muscles of DM would be different than in the model cells used in this study. Therefore, in further studies, the MBNL-sensitive RNA regulatory element in MB22 can be easily replaced, giving the opportunity to better adjust the buffering level of the MBNL1auto. The regulatory element with lower affinity to MBNL proteins can be used to reduce sensitivity of ex22 inclusion and, consequently, increase the levels of arranged protein, keeping still the control on excess activity of MBNLs. Moreover, the use of a well-selected tissue-specific promoter and a proper delivery system could be adjusted.

Taken together, our results highlight the utility of autoregulated overexpression of MBNL1auto as a potential therapeutic tool in DM1, DM2, and other diseases in which MBNL proteins are sequestered or insufficient.

MATERIALS AND METHODS

Genetic constructs

The MB22 construct was prepared by PCR amplification of an *ATP2A1* gene fragment from human genomic DNA (for primers,

see Table S1) and cloning into the previously described pEGFP-C1-MBNL1-41 vector for MBNL1 overexpression⁴⁶ with removed GFP sequence. The *MBNL1* sequence contains alternative ex7 and is deprived of alternative ex5, which contains a nuclear localization signal (Figure S1). The MB22-del and MB22#2 constructs were obtained by deletion of the fragment with MBNL-binding motifs and the replacement of the WT fragment with the 4xUGCU sequence, respectively. The MB22#2-GFP was prepared by an amplified EGFP sequence from pEGFP-C1 (CloneTech) and fusion to the C-terminal end of *MBNL1*. All the above constructs were prepared using the In-Fusion Cloning method. The desired fragments were amplified using CloneAmp HiFi PCR Premix (TakaraBio) and cloned using NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs) following the manufacturer's instructions. The pEGFP-C1-MBNL1-41 vector used in this study (referred to as MBNL1-GFP overexpression vector) was previously described.⁵⁰ The DT960 (CUG₉₆₀) and DT0 (CUG₆) vectors were a gift (Prof. Thomas Cooper, Baylor College of Medicine) and were already described.^{25,51} The pEGFP-C1 vector (CloneTech) was used as a control and is referred to as GFP. The *Atp2a1* ex22⁵² and *Nfix* ex7 minigenes, gifts from Prof. Manuel Ares, University of California, Santa Cruz, were described earlier.^{8,52} The *MBNL1* ex5 minigene is described in the manuscript (Taylor K. et al., unpublished data). The sequences of primers used in cloning are listed in Table S1. The lentiviral vectors were prepared by an external company (Viral Core Facility) by cloning the sequence of the MB22#2-GFP construct under the control of a cytomegalovirus (CMV) promoter. The lentiviral vector (BL-0533) containing a GFP sequence under the same promoter was used as a control.

Cell culture, transfection, and transduction

The monkey COS7, human HEK293, and mouse embryonic fibroblast (MEF) cells were grown in a high-glucose DMEM medium with L-glutamine (Biowest) supplemented with 10% fetal bovine serum (Thermo Fisher Scientific) and 1% antibiotic/antimycotic (Life Technologies) at 37°C in 5% CO₂. Fibroblasts derived from patients with DM1 (cell lines GM04033 and GM03989 expressing *DMPK* transcripts with ~1,000 and ~2,000 CUG repeats, respectively) and control fibroblasts derived from non-DM1 patient (cell line GM07492) were purchased from the Coriell Cell Repositories. Fibroblasts were grown in Eagle's minimal essential medium (EMEM) (Biowest) supplemented with 10% fetal bovine serum (FBS) (Thermo Fisher Scientific), 1% antibiotic antimycotic (Life Technologies), and 1% non-essential amino acids solution (Sigma) in 5% CO₂ at 37°C. MEF 1&2KO cells were gifts from Maurice Swanson at the University of Florida. For transfection, cells were plated in 12-well plates and transfected at ~80% confluency using Lipofectamine 3000 (Thermo Fisher Scientific) according to the manufacturer's instructions. Single transfection with siRNA against the 3' UTR fragment of the *MBNL1* sequence or control siRNA (synthesized by Sigma-Aldrich) was performed using Lipofectamine RNAiMAX (Thermo Fisher Scientific) at 50 nM concentration (sequence of siRNA duplexes are specified in Table S2). After 24 h, cells were transfected with plasmids. Co-transfection was conducted in a 1:1 ratio with MB22 construct and pEGFP-C1 or pEGFP-C1-MBNL1-41 vector; the same ratio was used in the

experiments with DT960 vector. The co-transfection with minigenes was made in a 1:10 ratio (minigene: construct). For all experiments, the total amount of DNA added to the cells was 1 $\mu\text{g}/\text{mL}$ of cell culture medium, except for experiments with different plasmid concentrations, which are specified in figure legend (Figure 3B). The cells were harvested 48 h after transfection or 72 h after the experiment with siRNA treatment. The samples referred to as a mock are lipofectamine-treated cells. For transduction, cells were plated in 48-well plates. Lentiviral vectors with a concentration of 10^7 particles/mL at a multiplicity of infection (MOI) 1 for control and MOI 10 for MB22#2-GFP in an appropriate amount of growth medium were delivered into the cells. After 24 h, the medium was replaced. The cells were harvested after 12 days.

Alternative splicing and real-time qPCR analysis

The cells were harvested using TRIzol Reagent (Thermo Fisher Scientific), and total RNA was isolated using Total RNA Zol-Out D (A&A Biotechnology) according to the manufacturer's protocol. cDNA was synthesized using TranScriba Kit (A&A Biotechnology) with Random Primers (A&A Biotechnology) according to the manufacturer's protocol. PCR was performed using GoTaq DNA Polymerase (Promega), and primers are listed in Table S3. PCR products were separated in 1%–2% agarose gel with ethidium bromide. The images were captured using G:Box EF2 (Syngene) and analyzed using GeneTools (Syngene) (Figures S7A, S7B, S9A, S9B, S10, and S12). Percent sliced in (PSI) was calculated based on signals of two bands, corresponding to the PCR product containing or missing alternative exons, according to the following formula (isoform with included exon*100)/(isoforms with included exon + excluded exon). Quantitative real-time RT-PCRs were performed in a QuantStudio 7 Flex System (Thermo Fisher Scientific) using Maxima SYBR Green/ROX qPCR Master Mix (Thermo Fisher Scientific) according to the manufacturer's instructions. Targets were amplified with primers listed in Table S3 at 58°C–60°C annealing temperature. Ct values were normalized against *GAPDH*. Fold differences in expression level were calculated according to the $2^{-\Delta\Delta\text{Ct}}$ method.

Western blot

Cells were lysed with radioimmunoprecipitation assay (RIPA) buffer (150 mM NaCl, 50 mM Tris-HCl [pH 8.0], 1 mM ethylenediaminetetraacetic acid [EDTA], 0.5% NP-40, 0.5% Triton X-100, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate [SDS]) supplemented with Halt Protease Inhibitor Cocktail (Thermo Fisher Scientific). Lysates were incubated on ice and vortexed followed by centrifugation at $15,000 \times g$ at 4°C for 15 min. Concentration of protein in cell extracts were measured by Pierce BCA Protein Assay Kit (Thermo Fisher Scientific). Samples were heated with standard sample buffer at 95°C for 5 min. Electrophoresis and wet transfer were performed with the use of the Mini-PROTEAN Tetra System (Bio-Rad). Protein extracts (25–40 μg) were separated on 10% SDS polyacrylamide gels in Laemmli buffer and transferred to nitrocellulose membrane (Sigma-Aldrich) (1 h, 100 V) in Laemmli buffer with 20% methanol. Membranes were blocked for 1 h in 5% Skim Milk Powder (Sigma-Aldrich) in TBST buffer (Tris-buffered saline

[TBS], 0.1% Tween 20). Membranes were incubated with a primary antibody against FLAG (A8592, Sigma-Aldrich) 1:1,000, *GAPDH* (sc-47724, Santa Cruz) 1:10,000, or mCherry (5F8, ChromoTek) 1:1,000 in 5% Skim Milk Powder in TBST for 1 h at room temperature. Membranes were washed in TBST and incubated with secondary antibodies conjugated with horseradish peroxidase, anti-mouse (A9044, Sigma-Aldrich) 1:20,000, or anti-rat (ab6836, abcam) 1:8,000 in TBST for 1 h at room temperature. Membranes were again washed in TBST and detected using Immobilon Forte Western HRP substrate (Sigma-Aldrich). Images were captured using G:Box Chemi-XR5 (Syngene) (Figures S7B, S8, and S11) and quantified using Multi Gauge 3.0 software (Fujifilm).

Microscopic analysis

The confocal microscopy was used to analyze sequestration of MBNL1 auto-GFP to the CUG^{exp} ribonuclear foci and to monitor homogeneity of the level of this protein in cells. COS7 cells were plated in μ -Slide 8 well (ibidi) after 48 h posttransfection with MB22#2-GFP, and (CUG₉₆₀) or (CUG₀) medium was replaced with PBS with diluted Hoechst (Thermo Fisher Scientific), 1:2,000, and incubated for 5 min, protected from light. Images were captured with the Nikon A1Rsi confocal microscope with Nikon Apo $\times 40$ WI λ S DIC N2 objective. GFP, Hoechst, and Cy3 were excited with 488 nm Argon-Ion and 405 and 561 nm diode lasers, respectively. For detection, dichroic mirrors of 405/488/561 nm with spectral filters of 525/50, 595/50, and 450/50 nm were used.

Flow cytometry

For flow cytometry experiments, cells were analyzed 48 h posttransfection with either GFP, MB22#2-GFP, or MBNL1-GFP plasmid. Culture medium was removed, and cells were washed with PBS, trypsinized, collected, and centrifuged for 5 min at $300 \times g$. Cell pellet was suspended in 400 μL PBS. The 100 μL cell suspension was diluted with 100 μL PBS and analyzed with guava easyCyteTM HT flow cytometer and guavaSoft software (Luminex). GFP fluorescence was excited by a 488 nm laser and detected at 525/30 nm. The threshold for GFP-positive cells was set based on the signal from mock-transfected cells. For each sample, 5,000 events were collected. The single event was referred to as single cell. The gating strategy is presented in Figure S3A.

FISH

RNA FISH and immunofluorescence (IF) cells were fixed in 2% PFA/PBS at room temperature for 10 min and washed three times in PBS. Pre-hybridization was performed in 30% formamide and $2 \times$ SSC for 10 min, followed by hybridization in buffer containing 30% formamide, $2 \times$ SSC, 0.02% BSA, 66 $\mu\text{g}/\text{mL}$ yeast tRNA, 10% dextran sulfate, 2 mM vanadyl ribonucleoside complex, and 2 ng/ μL DNA/LNA probe (CAG)₆-CA. The probe was labeled at the 5' end with Cy3 and modified at positions 2, 5, 8, 13, 16, and 19 with LNA. Posthybridization washing was done in 30% formamide and $2 \times$ SSC at 45°C for 30 min followed by $1 \times$ SSC at 37°C for the next 30 min. Slides were mounted in Vectashield medium (Vector Laboratories, Burlingame, CA, USA) with DAPI.

Statistical analysis

Group data are expressed as the means \pm standard deviation (SD). The statistical significance was determined by unpaired, two-tailed Student's *t* test using Prism software v.8 (GraphPad): **p* < 0.05; ***p* < 0.01; ****p* < 0.001; ns, non-significant. All analyses are based on at least three independent biological replicates (exceptions are indicated in the figure legends), and whole experiments were repeated at least twice to confirm obtained results.

DATA AVAILABILITY

This study did not generate/analyze datasets or code.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.omtn.2022.10.023>.

ACKNOWLEDGMENTS

We thank Dr. Katarzyna Taylor for sharing the *MBNL1* ex5 minigene and Dr. Wojciech Kwiatkowski for assistance with confocal microscopy. This work was supported by the Foundation for Polish Science TEAM program co-financed by the European Union within the European Regional Development Fund (POIR.04.04.00-00-5C0C/17-00 to K.S.), Polish National Science Center (2020/38/A/NZ3/00498 to K.S.), the Ministry of Science and Higher Education of the Republic of Poland, from the quality-promoting subsidy, under the Leading National Research Center (KNOW) program for the years 2012–2017 (KNOW RNA Research Center in Poznan; 01/KNOW2/2014). Funding for open access charge: Initiative of Excellence—Research University (05/IDUB/2019/94) at Adam Mickiewicz University, Poznan, Poland.

AUTHOR CONTRIBUTIONS

Conception and design of the experiments, K.S.; performed the experiments, Z.R.; analysis of data, K.S. and Z.R.; writing and revising the manuscript and figures, K.S. and Z.R.

DECLARATION OF INTERESTS

The authors declare no conflicts of interest.

REFERENCES

- Ho, T.H., Charlet-B, N., Poulos, M.G., Singh, G., Swanson, M.S., and Cooper, T.A. (2004). Muscleblind proteins regulate alternative splicing. *EMBO J.* 23, 3103–3112. <https://doi.org/10.1038/sj.emboj.7600300>.
- Batra, R., Manchanda, M., and Swanson, M.S. (2015). Global insights into alternative polyadenylation regulation. *RNA Biol.* 12, 597–602. <https://doi.org/10.1080/15476286.2015.1040974>.
- 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. (2012). Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. *Cell* 150, 710–724. <https://doi.org/10.1016/j.cell.2012.06.041>.
- Fernandez-Costa, J.M., Llamusi, M.B., Garcia-Lopez, A., and Artero, R. (2011). Alternative splicing regulation by Muscleblind proteins: from development to disease. *Biol. Rev. Camb. Philos. Soc.* 86, 947–958. <https://doi.org/10.1111/j.1469-185X.2011.00180.x>.
- Kanadia, R.N., Urbinati, C.R., Crusselle, V.J., Luo, D., Lee, Y.J., Harrison, J.K., Oh, S.P., and Swanson, M.S. (2003). Developmental expression of mouse muscleblind genes Mbnl1, Mbnl2 and Mbnl3. *Gene Expr. Patterns* 3, 459–462. [https://doi.org/10.1016/S1567-133X\(03\)00064-4](https://doi.org/10.1016/S1567-133X(03)00064-4).
- Fardaei, M., Rogers, M.T., Thorpe, H.M., Larkin, K., Hamshere, M.G., Harper, P.S., and Brook, J.D. (2002). Three proteins, MBNL, MBLL and MBXL, co-localize in vivo with nuclear foci of expanded-repeat transcripts in DM1 and DM2 cells. *Hum. Mol. Genet.* 11, 805–814. <https://doi.org/10.1093/hmg/11.7.805>.
- Teplova, M., and Patel, D.J. (2008). Structural insights into RNA recognition by the alternative-splicing regulator muscleblind-like MBNL1. *Nat. Struct. Mol. Biol.* 15, 1343–1351. <https://doi.org/10.1038/nsmb.1519>.
- Du, H., Cline, M.S., Osborne, R.J., Tuttle, D.L., Clark, T.A., Donohue, J.P., Hall, M.P., Shiue, L., Swanson, M.S., Thornton, C.A., and Ares, M., Jr. (2010). Aberrant alternative splicing and extracellular matrix gene expression in mouse models of myotonic dystrophy. *Nat. Struct. Mol. Biol.* 17, 187–193. <https://doi.org/10.1038/nsmb.1720>.
- Goers, E.S., Purcell, J., Voelker, R.B., Gates, D.P., and Berglund, J.A. (2010). MBNL1 binds GC motifs embedded in pyrimidines to regulate alternative splicing. *Nucleic Acids Res.* 38, 2467–2484. <https://doi.org/10.1093/nar/gkp1209>.
- Harper, P.S. (2001). *Myotonic dystrophy*, 3rd ed. (WB Saunders).
- 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. (1992). Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* 68, 799–808. [https://doi.org/10.1016/0092-8674\(92\)90154-5](https://doi.org/10.1016/0092-8674(92)90154-5).
- Liquori, C.L., Ricker, K., Moseley, M.L., Jacobsen, J.F., Kress, W., Naylor, S.L., Day, J.W., and Ranum, L.P. (2001). Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of *ZNF9*. *Science* 293, 864–867. <https://doi.org/10.1126/science.1062125>.
- Miller, J.W., Urbinati, C.R., Teng-Umuay, P., Stenberg, M.G., Byrne, B.J., Thornton, C.A., and Swanson, M.S. (2000). Recruitment of human muscleblind proteins to (CUG)*n* expansions associated with myotonic dystrophy. *EMBO J.* 19, 4439–4448. <https://doi.org/10.1093/emboj/19.17.4439>.
- Sobczak, K., de Mezer, M., Michlewski, G., Krol, J., and Krzyzosiak, W.J. (2003). RNA structure of trinucleotide repeats associated with human neurological diseases. *Nucleic Acids Res.* 31, 5469–5482. <https://doi.org/10.1093/nar/gkg766>.
- Fardaei, M., Larkin, K., Brook, J.D., and Hamshere, M.G. (2001). In vivo co-localisation of MBNL protein with DMPK expanded-repeat transcripts. *Nucleic Acids Res.* 29, 2766–2771. <https://doi.org/10.1093/nar/29.13.2766>.
- Jiang, H., Mankodi, A., Swanson, M.S., Moxley, R.T., and Thornton, C.A. (2004). Myotonic dystrophy type 1 is associated with nuclear foci of mutant RNA, sequestration of muscleblind proteins and deregulated alternative splicing in neurons. *Hum. Mol. Genet.* 13, 3079–3088. <https://doi.org/10.1093/hmg/ddh327>.
- Martorell, L., Monckton, D.G., Gamez, J., Johnson, K.J., Gich, I., Lopez De Munain, A., and Baiget, M. (1998). Progression of somatic CTG repeat length heterogeneity in the blood cells of myotonic dystrophy patients. *Hum. Mol. Genet.* 7, 307–312. <https://doi.org/10.1093/hmg/7.2.307>.
- Wong, L.J., Ashizawa, T., Monckton, D.G., Caskey, C.T., and Richards, C.S. (1995). Somatic heterogeneity of the CTG repeat in myotonic dystrophy is age and size dependent. *Am. J. Hum. Genet.* 56, 114–122.
- Thornton, C.A., Johnson, K., and Moxley, R.T. (1994). Myotonic dystrophy patients have larger CTG expansions in skeletal muscle than in leukocytes. *Ann. Neurol.* 35, 104–107. <https://doi.org/10.1002/ana.410350116>.
- Zatz, M., Passos-Bueno, M.R., Cerqueira, A., Marie, S.K., Vainzof, M., and Pavanello, R.C. (1995). Analysis of the CTG repeat in skeletal muscle of young and adult myotonic dystrophy patients: when does the expansion occur? *Hum. Mol. Genet.* 4, 401–406. <https://doi.org/10.1093/hmg/4.3.401>.
- Monckton, D.G., Wong, L.J., Ashizawa, T., and Caskey, C.T. (1995). Somatic mosaicism, germline expansions, germline reversions and intergenerational reductions in myotonic dystrophy males: small pool PCR analyses. *Hum. Mol. Genet.* 4, 1–8. <https://doi.org/10.1093/hmg/4.1.1>.
- 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. (2020). Three-dimensional imaging in myotonic dystrophy type 1. *Neurol. Genet.* 6, 4844. <https://doi.org/10.1212/NXG.0000000000000484>.

23. Jenquin, J.R., O'Brien, A.P., Poukalov, K., Lu, Y., Frias, J.A., Shorrock, H.K., Richardson, J.L., Mazdiyasi, H., Yang, H., Huigens, R.W., et al. (2022). Molecular characterization of myotonic dystrophy fibroblast cell lines for use in small molecule screening. *iScience* 25, 104198. <https://doi.org/10.1016/j.isci.2022.104198>.
24. Alfano, M., de Antoni, L., Centofanti, F., Visconti, V.V., Maestri, S., Degli Esposti, C., Massa, R., D'Apice, M.R., Novelli, G., Delledonne, M., et al. (2022). Characterization of full-length CNBP expanded alleles in myotonic dystrophy type 2 patients by Cas9-mediated enrichment and nanopore sequencing. *Elife* 11, e80229. <https://doi.org/10.7554/eLife.80229>.
25. Wojtkowiak-Szlachcic, A., Taylor, K., Stepniak-Konieczna, E., Sznajder, L.J., Mykowska, A., Sroka, J., Thornton, C.A., and Sobczak, K. (2015). Short antisense-locked nucleic acids (all-LNAs) correct alternative splicing abnormalities in myotonic dystrophy. *Nucleic Acids Res.* 43, 3318–3331. <https://doi.org/10.1093/nar/gkv163>.
26. Nakamori, M., Taylor, K., Mochizuki, H., Sobczak, K., and Takahashi, M.P. (2016). Oral administration of erythromycin decreases RNA toxicity in myotonic dystrophy. *Ann. Clin. Transl. Neurol.* 3, 42–54. <https://doi.org/10.1002/acn3.271>.
27. Christou, M., Wengel, J., Sokratous, K., Kyriacou, K., Nikolaou, G., Phylactou, L.A., and Mastroyiannopoulos, N.P. (2020). Systemic evaluation of chimeric LNA/2'-O-methyl steric blockers for myotonic dystrophy type 1 therapy. *Nucleic Acid Ther.* 30, 80–93. <https://doi.org/10.1089/nat.2019.0811>.
28. Wheeler, T.M., Sobczak, K., Lueck, J.D., Osborne, R.J., Lin, X., Dirksen, R.T., and Thornton, C.A. (2009). Reversal of RNA dominance by displacement of protein sequestered on triplet repeat RNA. *Science* 325, 336–339. <https://doi.org/10.1126/science.1173110>.
29. Chen, G., Masuda, A., Konishi, H., Ohkawara, B., Ito, M., Kinoshita, M., Kiyama, H., Matsuura, T., and Ohno, K. (2016). Phenylbutazone induces expression of MBNL1 and suppresses formation of MBNL1-CUG RNA foci in a mouse model of myotonic dystrophy. *Sci. Rep.* 6, 25317–25411. <https://doi.org/10.1038/srep25317>.
30. Kanadia, R.N., Shin, J., Yuan, Y., Beattie, S.G., Wheeler, T.M., Thornton, C.A., and Swanson, M.S. (2006). Reversal of RNA missplicing and myotonia after muscleblind overexpression in a mouse poly(CUG) model for myotonic dystrophy. *Proc. Natl. Acad. Sci. USA* 103, 11748–11753. <https://doi.org/10.1073/pnas.0604970103>.
31. Arandel, L., Matloka, M., Klein, A.F., Rau, F., Sureau, A., Ney, M., Cordier, A., Kondili, M., Polay-Espinoza, M., Naouar, N., et al. (2022). Reversal of RNA toxicity in myotonic dystrophy via a decoy RNA-binding protein with high affinity for expanded CUG repeats. *Nat. Biomed. Eng.* 6, 207–220. <https://doi.org/10.1038/s41551-021-00838-2>.
32. Chamberlain, C.M., and Ranum, L.P.W. (2012). Mouse model of muscleblind-like 1 overexpression: skeletal muscle effects and therapeutic promise. *Hum. Mol. Genet.* 21, 4645–4654. <https://doi.org/10.1093/hmg/dds306>.
33. Yadava, R.S., Kim, Y.K., Mandal, M., Mahadevan, K., Gladman, J.T., Yu, Q., and Mahadevan, M.S. (2019). MBNL1 overexpression is not sufficient to rescue the phenotypes in a mouse model of RNA toxicity. *Hum. Mol. Genet.* 28, 2330–2338. <https://doi.org/10.1093/hmg/ddz065>.
34. Meola, G. (2013). Clinical aspects, molecular pathomechanisms and management of myotonic dystrophies. *Acta Myol.* 32, 154–165.
35. Nakamori, M., Sobczak, K., Puwanant, A., Welle, S., Eichinger, K., Pandya, S., Dekdebrun, J., Heatwole, C.R., McDermott, M.P., Chen, T., et al. (2013). Splicing biomarkers of disease severity in myotonic dystrophy. *Ann. Neurol.* 74, 862–872. <https://doi.org/10.1002/ana.23992>.
36. Hino, S.I., Kondo, S., Sekiya, H., Saito, A., Kanemoto, S., Murakami, T., Chihara, K., Aoki, Y., Nakamori, M., Takahashi, M.P., and Imaizumi, K. (2007). Molecular mechanisms responsible for aberrant splicing of SERCA1 in myotonic dystrophy type 1. *Hum. Mol. Genet.* 16, 2834–2843. <https://doi.org/10.1093/hmg/ddm239>.
37. Taylor, K., Sznajder, L.J., Cywoniuk, P., Thomas, J.D., Swanson, M.S., and Sobczak, K. (2018). MBNL splicing activity depends on RNA binding site structural context. *Nucleic Acids Res.* 46, 9119–9133. <https://doi.org/10.1093/nar/gky565>.
38. Batra, R., Charizanis, K., Manchanda, M., Mohan, A., Li, M., Finn, D.J., Goodwin, M., Zhang, C., Sobczak, K., Thornton, C.A., and Swanson, M.S. (2014). Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-mediated disease. *Mol. Cell* 56, 311–322. <https://doi.org/10.1016/j.molcel.2014.08.027>.
39. Stepniak-Konieczna, E., Konieczny, P., Cywoniuk, P., Druzewska, J., and Sobczak, K. (2020). AON-induced splice-switching and DMPK pre-mRNA degradation as potential therapeutic approaches for myotonic dystrophy type 1. *Nucleic Acids Res.* 48, 2531–2543. <https://doi.org/10.1093/nar/gkaa007>.
40. Day, J.W., Finkel, R.S., Chiriboga, C.A., Connolly, A.M., Crawford, T.O., Darras, B.T., Iannaccone, S.T., Kuntz, N.L., Peña, L.D.M., Shieh, P.B., et al. (2021). Onasemnogene abeparvovec gene therapy for symptomatic infantile-onset spinal muscular atrophy in patients with two copies of SMN2 (STRIVE): an open-label, single-arm, multi-centre, phase 3 trial. *Lancet Neurol.* 20, 284–293. [https://doi.org/10.1016/S1474-4422\(21\)00001-6](https://doi.org/10.1016/S1474-4422(21)00001-6).
41. Belbellaa, B., Reutenauer, L., Messaddeq, N., Monassier, L., and Puccio, H. (2020). High levels of frataxin overexpression lead to mitochondrial and cardiac toxicity in mouse models. *Mol. Ther. Methods Clin. Dev.* 19, 120–138. <https://doi.org/10.1016/j.omtm.2020.08.018>.
42. Yue, Y., Wasala, N.B., Bostick, B., and Duan, D. (2016). 100-fold but not 50-fold dystrophin overexpression aggravates electrocardiographic defects in the mdx model of Duchenne muscular dystrophy. *Mol. Ther. Methods Clin. Dev.* 3, 16045. <https://doi.org/10.1038/mtm.2016.45>.
43. Charizanis, K., Lee, K.Y., Batra, R., Goodwin, M., Zhang, C., Yuan, Y., Shiue, L., Cline, M., Scotti, M.M., Xia, G., et al. (2012). Muscleblind-like 2-mediated alternative splicing in the developing brain and dysregulation in myotonic dystrophy. *Neuron* 75, 437–450. <https://doi.org/10.1016/j.neuron.2012.05.029>.
44. Lin, X., Miller, J.W., Mankodi, A., Kanadia, R.N., Yuan, Y., Moxley, R.T., Swanson, M.S., and Thornton, C.A. (2006). Failure of MBNL1-dependent post-natal splicing transitions in myotonic dystrophy. *Hum. Mol. Genet.* 15, 2087–2097. <https://doi.org/10.1093/hmg/ddl132>.
45. Wagner, S.D., Struck, A.J., Gupta, R., Farnsworth, D.R., Mahady, A.E., Eichinger, K., Thornton, C.A., Wang, E.T., and Berglund, J.A. (2016). Dose-dependent regulation of alternative splicing by MBNL proteins reveals biomarkers for myotonic dystrophy. *PLoS Genet.* 12, e1006316–e1006324. <https://doi.org/10.1371/journal.pgen.1006316>.
46. Sznajder, L.J., Michalak, M., Taylor, K., Cywoniuk, P., Kabza, M., Wojtkowiak-Szlachcic, A., Matloka, M., Konieczny, P., and Sobczak, K. (2016). Mechanistic determinants of MBNL activity. *Nucleic Acids Res.* 44, 10326–10342. <https://doi.org/10.1093/nar/gkw915>.
47. Izzo, M., Battistini, J., Provenzano, C., Martelli, F., Cardinali, B., and Falcone, G. (2022). Molecular therapies for myotonic dystrophy type 1: from small drugs to gene editing. *Int. J. Mol. Sci.* 23, 4622. <https://doi.org/10.3390/ijms23094622>.
48. López-Martínez, A., Soblechero-Martín, P., De-La-puente-ovejero, L., Nogales-Gadea, G., and Arechavala-Gomez, V. (2020). An overview of alternative splicing defects implicated in myotonic dystrophy type i. *Genes* 11, E1109–E1127. <https://doi.org/10.3390/genes11091109>.
49. Tran, H., Gourrier, N., Lemerrier-Neuillet, C., Dhaenens, C.M., Vautrin, A., Fernandez-Gomez, F.J., Arandel, L., Carpentier, C., Obriot, H., Eddarkaoui, S., et al. (2011). Analysis of exonic regions involved in nuclear localization, splicing activity, and dimerization of muscleblind-like-1 isoforms. *J. Biol. Chem.* 286, 16435–16446. <https://doi.org/10.1074/jbc.M110.194928>.
50. Konieczny, P., Stepniak-Konieczna, E., Taylor, K., Sznajder, L.J., and Sobczak, K. (2017). Autoregulation of MBNL1 function by exon 1 exclusion from MBNL1 transcript. *Nucleic Acids Res.* 45, 1760–1775. <https://doi.org/10.1093/nar/gkw1158>.
51. Ho, T.H., Savkur, R.S., Poulos, M.G., Mancini, M.A., Swanson, M.S., and Cooper, T.A. (2005). Colocalization of muscleblind with RNA foci is separable from mis-regulation of alternative splicing in myotonic dystrophy. *J. Cell Sci.* 118, 2923–2933. <https://doi.org/10.1242/jcs.02404>.
52. Cywoniuk, P., Taylor, K., Sznajder, L.J., and Sobczak, K. (2017). Hybrid splicing mini-gene and antisense oligonucleotides as efficient tools to determine functional protein/RNA interactions. *Sci. Rep.* 7, 17587–17614. <https://doi.org/10.1038/s41598-017-17816-x>.

OMTN, Volume 30

Supplemental information

**Sustainable recovery of MBNL activity
in autoregulatory feedback loop
in myotonic dystrophy**

Zuzanna Rogalska and Krzysztof Sobczak

Table S1. Primers used for genetic construct preparation.

name of primer	sequence 5' -> 3'
SRCglo_F	CCATTACAACCCGTGCAAGAGGGTAAGGAGT
SRCglo_R	AACTGAAAACATTGGATCTGGAAGGGGAAA
MB22mut_F	TGCTTGCTTGCTTGCTTCCAGTCAGGGTGGGCCG
MB22mut_R	GCAAGCAAGCAAGCAACGGCTCCAGGTGGAGCTGC
MB22del_F	CAAGGTCAGGGCTGCAGTGGGGGGGG
MB22del_R	GCAGCCCTGACCTTGGATGCGACGCG
GFPop_F	TAGGGATCCACCGGATCT
GFPop_R	CTTATCGTCGTCATCCTTGTAATCT
GFPvec_F	GATGACGACGATAAGGTGAGCAAGGGCGAGGAGC
GFPvec_R	TCCGGTGGATCCCTACTTGTACAGCTCGTCCATGCC

Table S2. siRNA sequences

siRNA	sequence 5' -> 3'
siCtrl sense	p- UCGAAGUAUCCGCGUACGdTdT
siCtrl antisense	p- CGUACGCGGAAUACUUCGAdTdT
siMBNL1 sense	p- GGACGAGGUCAUUAGCCAuTdTdT
siMBNL1 antisense	p- AUGGCUAAUGACCUCGUCCdTTdT

Table S3. PCR primers used for alternative splicing assays and real-time qPCR analysis

PCR target	Primer set	sequence 5' -> 3'
MB22 ex22	pEGFP2_F	ACCGGACTCAGATCTCGAATG
	MBex3_R	ACCAGGCTTGGAGAAACAG
Minigene <i>Atp2a1</i> ex22	SrcMins_F	GATCTTCAAGCTCCGGGCCCTG
	SrcMins_R	AGCAATCAGCTAGTCAGTTGCC
Minigene <i>Nfix</i> ex7	pDUP51_F	GCAACCTCAAACAGACACCA
	pDUP51_R	AGCTTGTCACAGTGCAGCTC
Minigene <i>MBNL1</i> ex5	54ex_F	GCTGCCCAATACCAGGTCAAC
	pcDNA RT-spec_R	AAAGGACAGTGGGAGTGGC
<i>INSR</i> ex11	INSR_F	CCAAAGACAGACTCTCAGAT
	INSR_R	AACATCGCCAAGGGACCTGC
<i>FLNB</i> ex31	FLNB_F	GCTTCGGTGGTGTGATATTC
	FLNB_R	GTCACTCACTGGGACATAGG
<i>MYO5A</i> ex22	MYO5A_F	GAACAACCGACAGCAGCAG
	MYO5A_R	TTACGGACCGTCTTATCCTG
<i>MBNL2</i> ex5	MBNL2ex5_F	ATTTTCACCCTCCTGCACAC
	MBNL2ex5_R	CAAGACGCTGGGGTTAAAGA
<i>MBNL2</i> ex7	MBNL2ex7_F	TCCTTTACCAAAGAGACAAGCAC
	MBNL2ex7_R	CTCAATGCAGATTCTTGGCATTCC
	MBNL1ex1_F	CAGCGACATGCAACAGTCTT

<i>MBNL1</i> ex1	MBNL1ex1_R	TGTCAGCAGGATGAGCAAAC
<i>NCOR2</i> ex19	NCOR2_F	ACACCCACAACCGGAATGAGCCTG
	NCOR2_R	GGACTTGGCTTTTCGGCTGCTG
<i>PHKA1</i> ex19	PHKA1_F	TGCACACACTTGAGCTTCATGGA
	PHKA1_R	AAAGTCCACCTCCCCAGACTGGTC
<i>GAPDH</i>	GAPDH_F	GAGTCAACGGATTTGGTCGT
	GAPDH_R	TTGATTTTGGAGGGATCTCG
<i>DMPK</i> UTR	DMPK_F	GCGATCTCTGCCTGCTTACT
	DMPK_R	GTCCTAGGTGGGGACAGACA

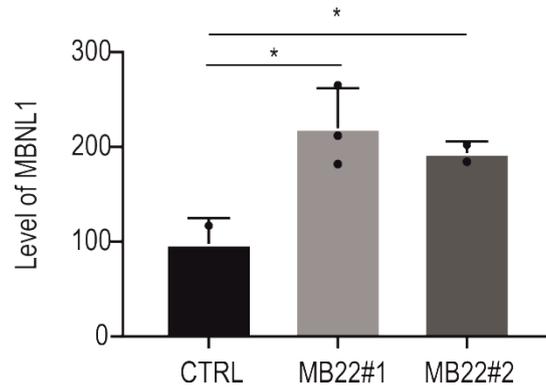
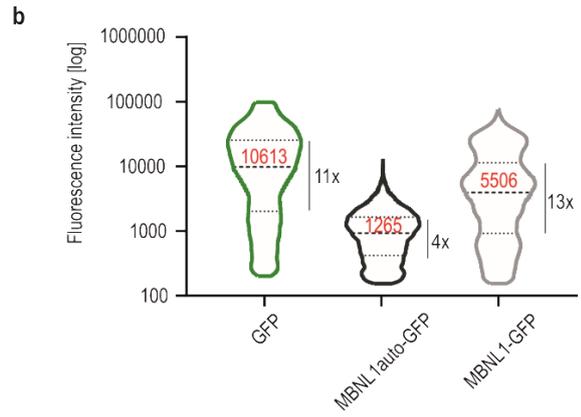
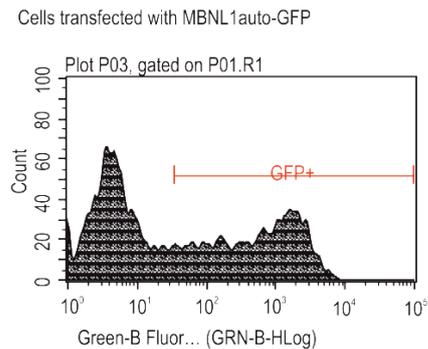
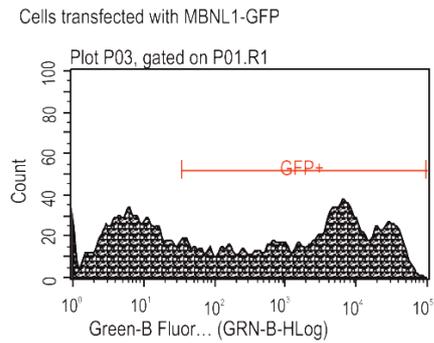
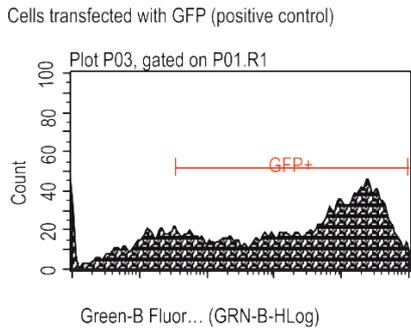
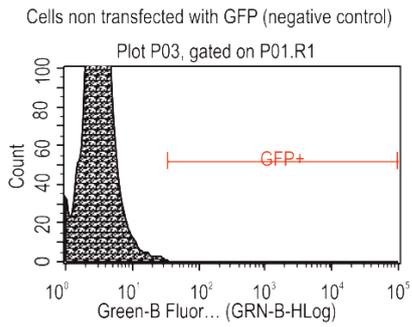
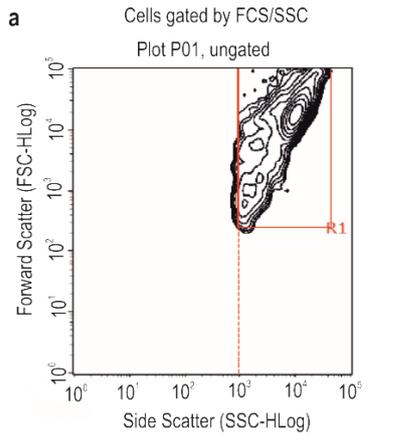


Figure S2. Total level of MBNL1 in MB22 treated cells (related to Fig. 1d).

Results of western blot analysis showing the level of total pool of MBNL1 in cells co-transfected with either MB22#1 or MB22#2 or GFP (CTRL). Anti-MBNL1 antibody staining was carried out. Bars represent average signal from n=3 independent experiments for each group normalized to mCherry. Co-transfection with mCherry expression vector was utilized as a normalization control of transfection. The MB22 treated cells were compared to GFP plasmid treated cells (CTRL); unpaired Student's *t*-test; * $P < 0.05$.



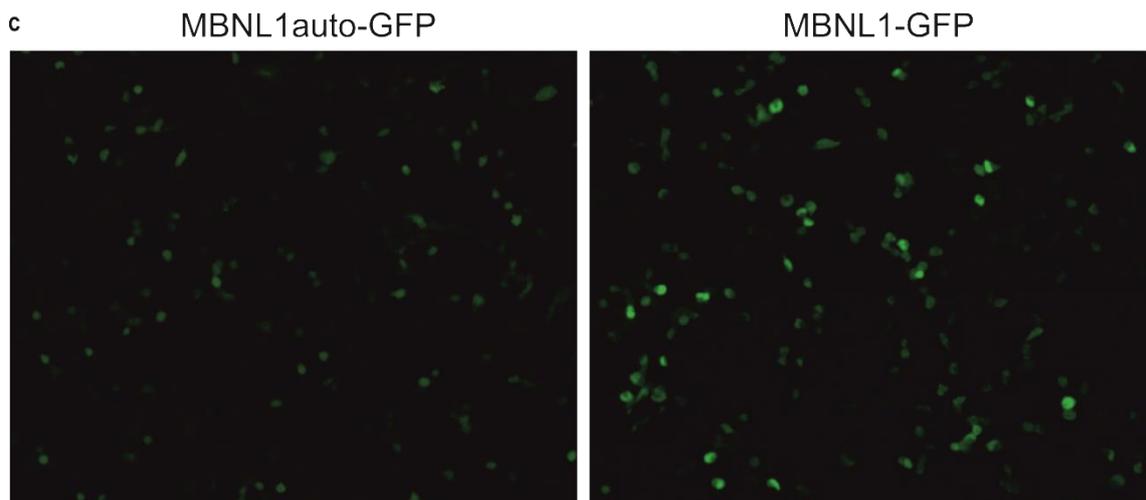


Figure S3. Representative flow cytometry gating and distribution of cells with different fluorescent signal (related to Fig. 3d).

a) For flow cytometry data analysis cells were first gated by Forward Scatter and Side Scatter (FSC/SSC). To distinguish between cells positive and negative for GFP channel, control cell sample (cells non-transfected with GFP referred as mock) were analyzed and utilized as background. Cells transfected with pEGFP-C1 plasmid with strong expression of GFP were also analyzed as positive control. Cells transfected with MBNL1-GFP and MB22#2-GFP represent fluorescence signal from MBNL1-GFP and MBNL1auto-GFP, respectively.

b) Violin plots showing the distribution of cells with different fluorescent signal of either GFP, MBNL1auto-GFP or MBNL1-GFP proteins. COS7 cells were transfected with adequate vectors 48 h prior to flow cytometry analysis. Median fluorescent intensity (black solid line) and 25th and 75th percentile of signal (dashed line) are shown. Fold-change between 25th and 75th percentile of signal for all analyses is also indicated. Cells with signal below 200 were rejected from analyses based on results for control experiment for mock-transfected cells. Signals below 200 were rejected based on analysis of non-transfected cells; Graphs represent values from n=4 independent biological replicates for each experimental condition; N=12916 (GFP), N=6781 (MBNL1auto-GFP) and N=10496 (MBNL1-GFP) cells. The obtained results from GFP, MBNL1-GFP and MBNL1auto-GFP were compared with mock treated as a control; Unpaired Student's *t*-test; **P* < 0.05; ** *P* < 0.01; *** *P* < 0.001; ns, non-significant.

c) Representative confocal microscopy images showing localization of MBNL1auto-GFP and MBNL1-GFP in COS7 cells transfected with either autoregulated MB22#2-GFP or conventional MBNL1-GFP constructs, respectively.

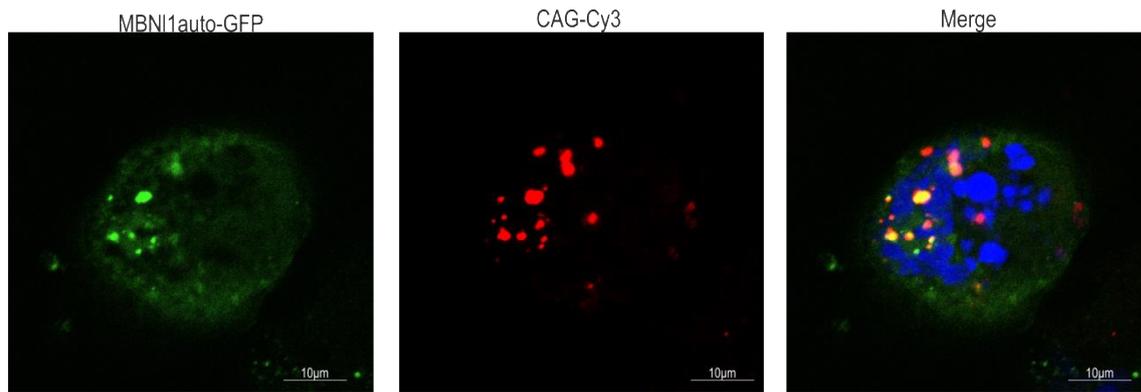


Figure S4. MBNL1auto colocalize with CUG^{exp} in MB22 treated cells (related to Fig. 4a).

Representative confocal images of FISH analysis to visualize the CUG^{exp} nuclear foci containing MBNL1auto in cells co-transfected with plasmid expressing mutant *DMPK* fragment containing CUG₉₆₀ with MB22#2-GFP construct. RNA FISH was performed with DNA/LNA probes (CAG)₆-CA labeled at the 5'-end with Cy3 (CAG-Cy3) and MBNL1auto-GFP was visualized with GFP-specific filter (MBNL1auto-GFP); scale bar, 10µm.

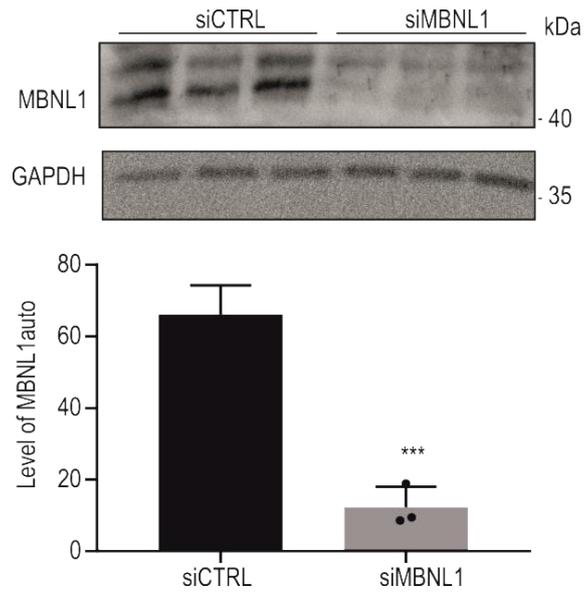


Figure S5. Confirmation of efficient knockdown of *MBNL1* (related to Fig. 4b).

Results of western blot analysis showing the level MBNL1 in COS7 cells treated with *siMBNL1* or *siCTRL*. Quantification of protein level (anti-MBNL1 antibody) is normalized to GAPDH. The results are averages from n=3 independent experiments; unpaired Student's *t-test*; *** $P < 0.001$ for comparison of *siCTRL* vs *siMBNL1* groups.

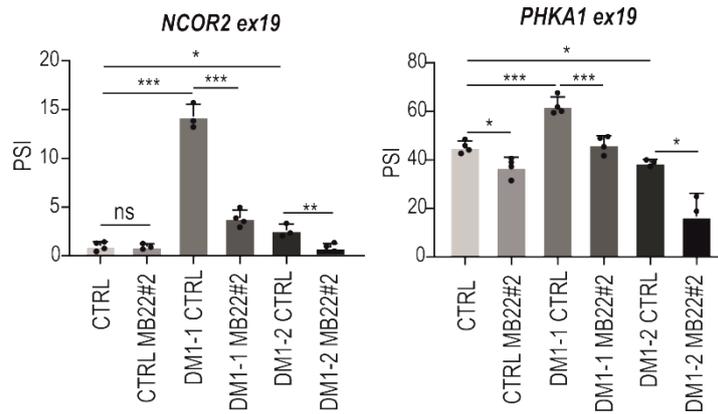
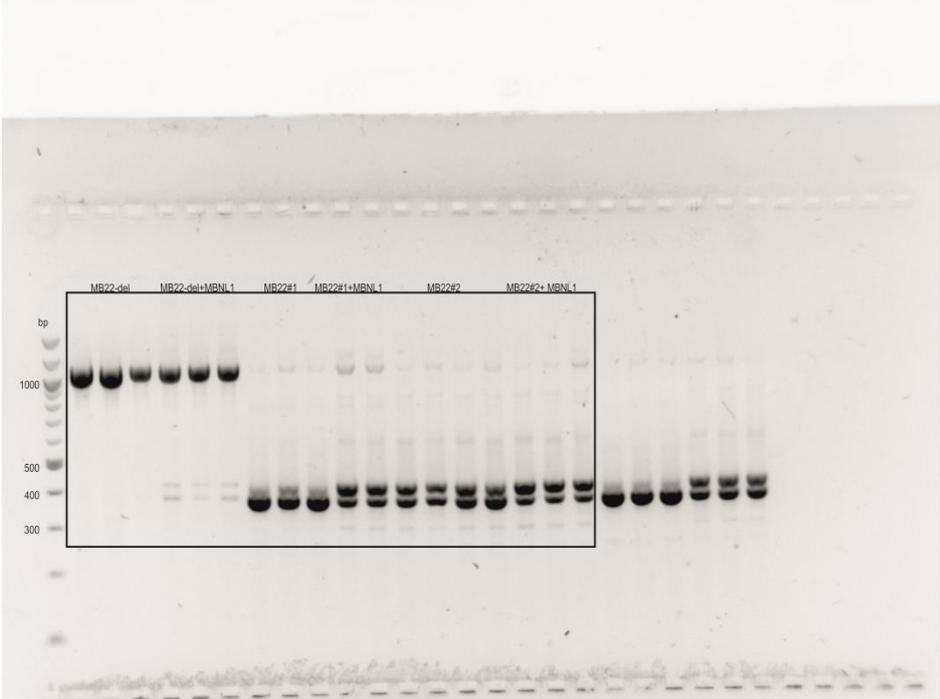


Figure S6. Correction of pathogenic mis-splicing in DM1 cells treated with MB22 lentiviruses (related to Fig. 5b).

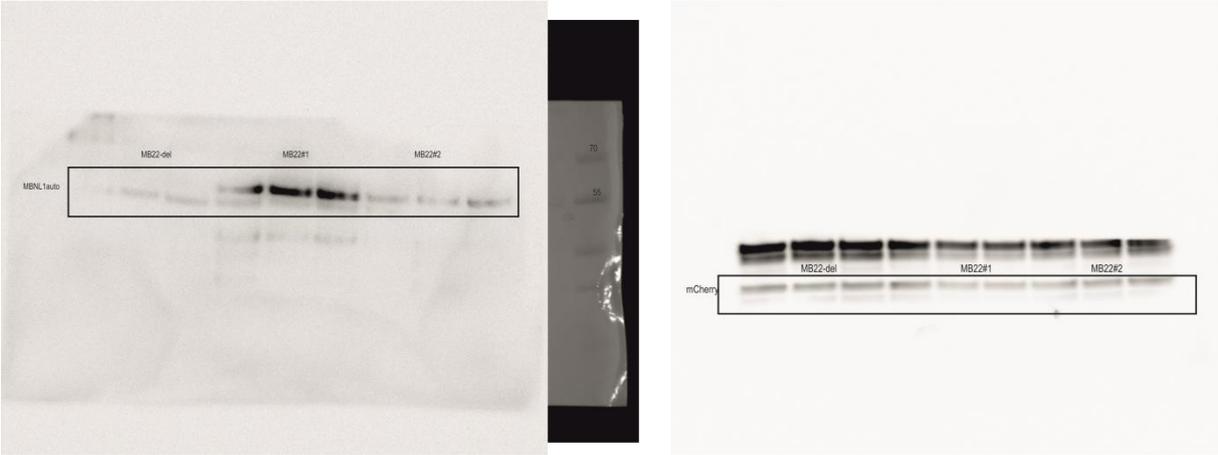
Results of RT-PCR analyses showing changes in the regulation of two negatively regulated MBNL-dependent exons. Splicing changes are expressed as PSI. Three different cell line: fibroblast from healthy individuals (non-DM), and two DM1 patients (DM1-1 and DM1-2) were treated with control lentiviral vector (CTRL) or with lentivirus containing MB22#2-GFP sequence. Bars represent average from four independent experiments (dots); unpaired Student's *t*-test; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; ns, non-significant.

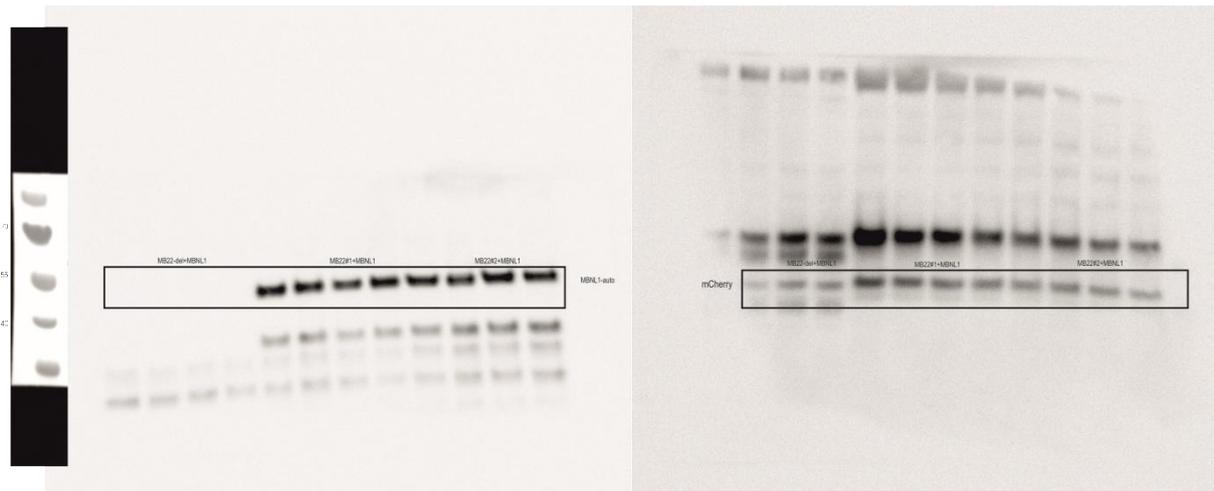
Supplementary files

a



b





c

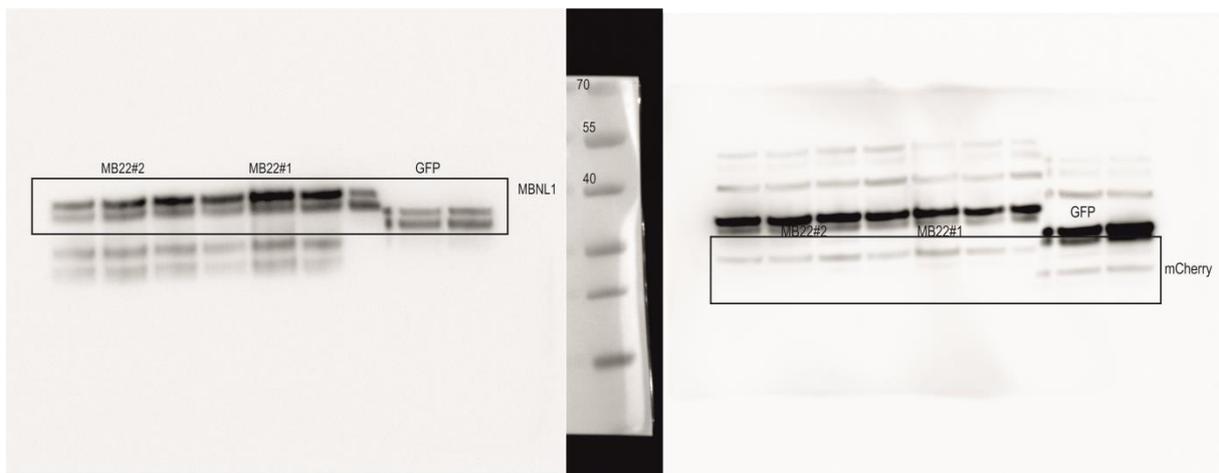
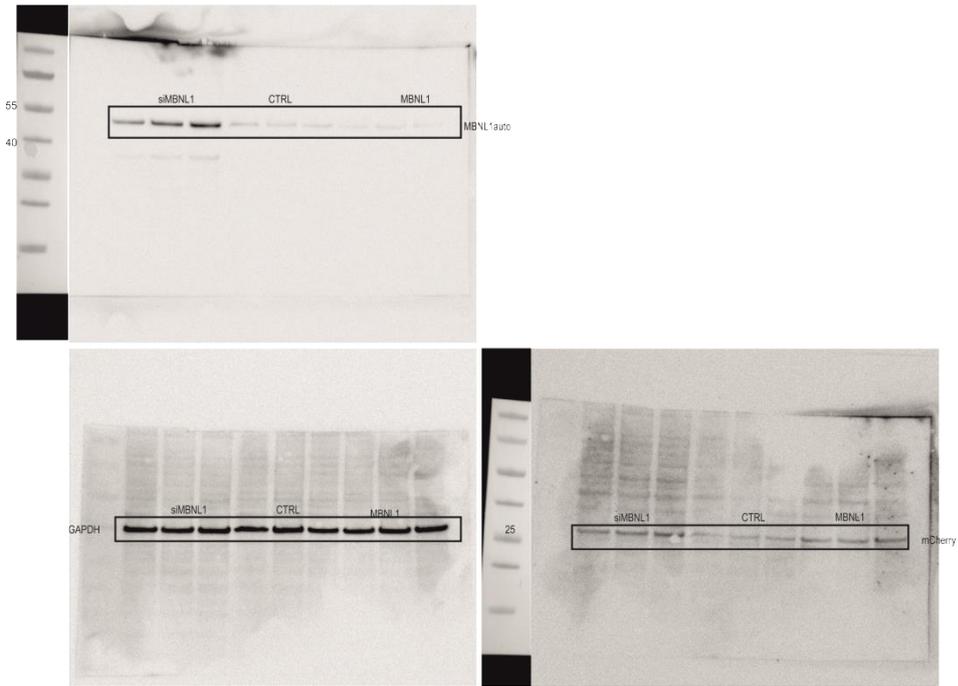


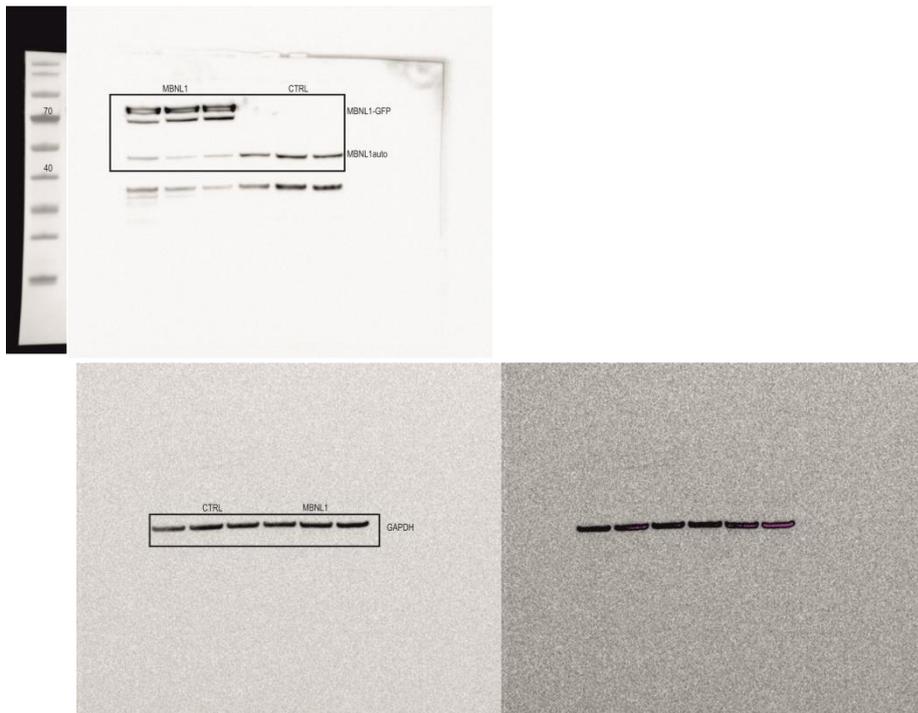
Figure S7 (related to Fig. 1c, d and Fig. S2).

a) Full agarose gel image for RT-PCR products results show in Fig. 1c. Fragments presented in Fig. 1c are indicated by black rectangles. **b)** Full blot image for western blot results show in Fig. 1d. Fragments presented in Fig. 1d are indicated by black rectangles. **c)** Full blot image for western blot results show in Fig. S2. Fragments presented in Fig. S2 are indicated by black rectangles.

COS7



HEK293



MEF 1&2KO

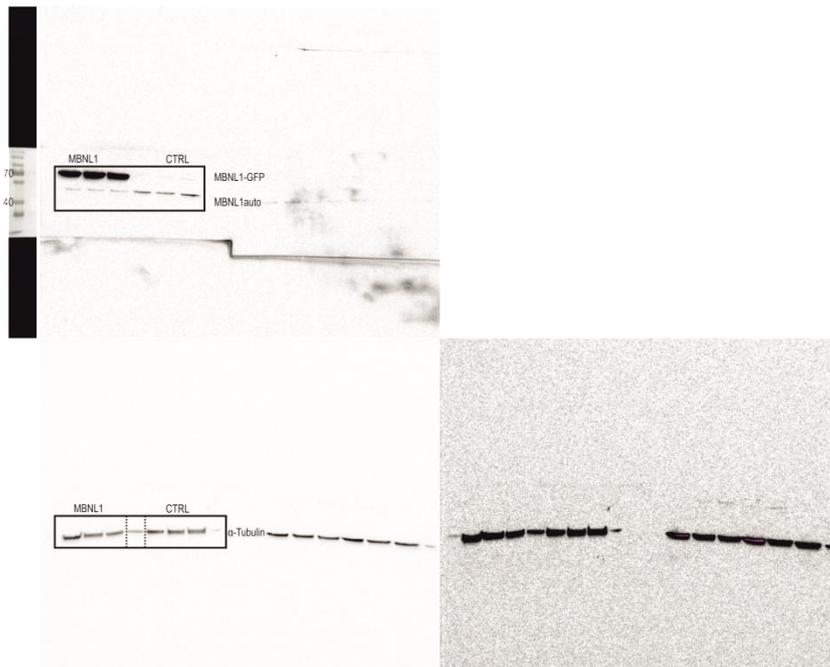
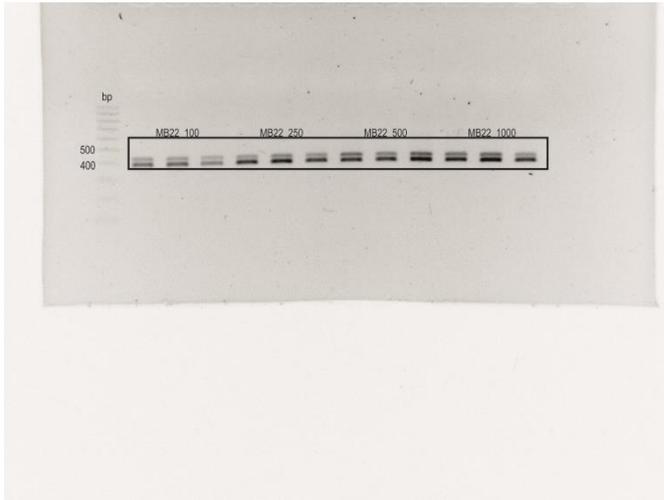


Figure S8 (related to Fig. 2).

Full blot images for western blot results show in Fig. 2. Fragments presented in Fig. 2 are indicated by black rectangles. Additional full blot image with GAPDH for HEK239 and α -Tubulin MEF 1&2KO was added with adjustment of contrast to see blot boundaries.

a



b

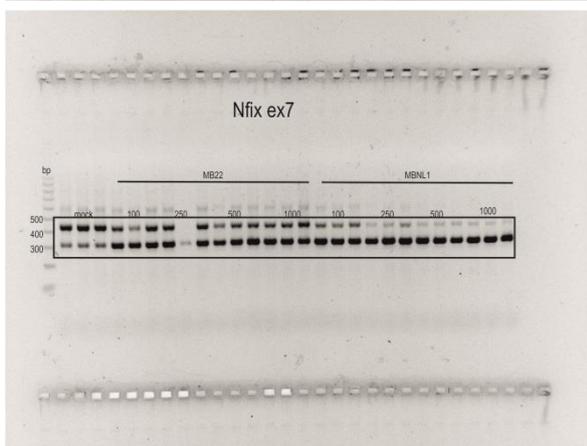
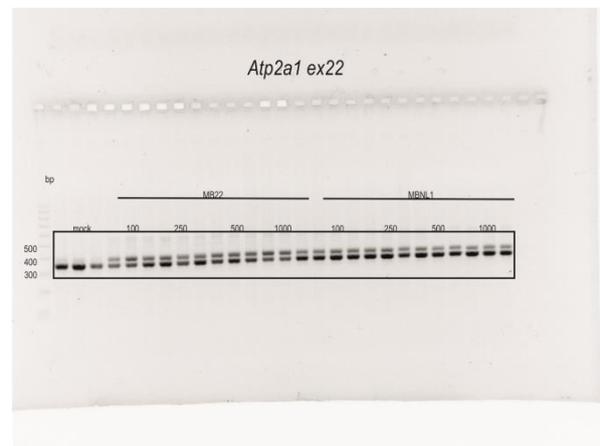
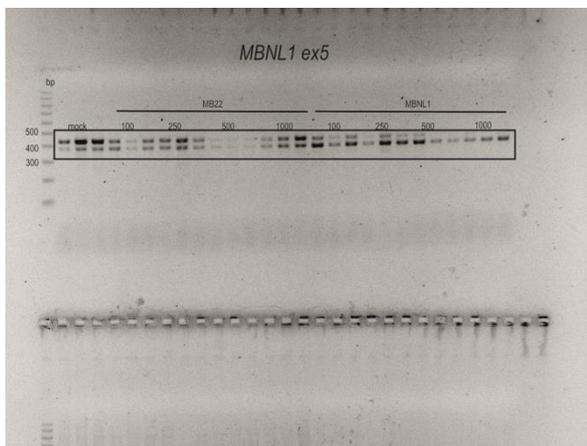
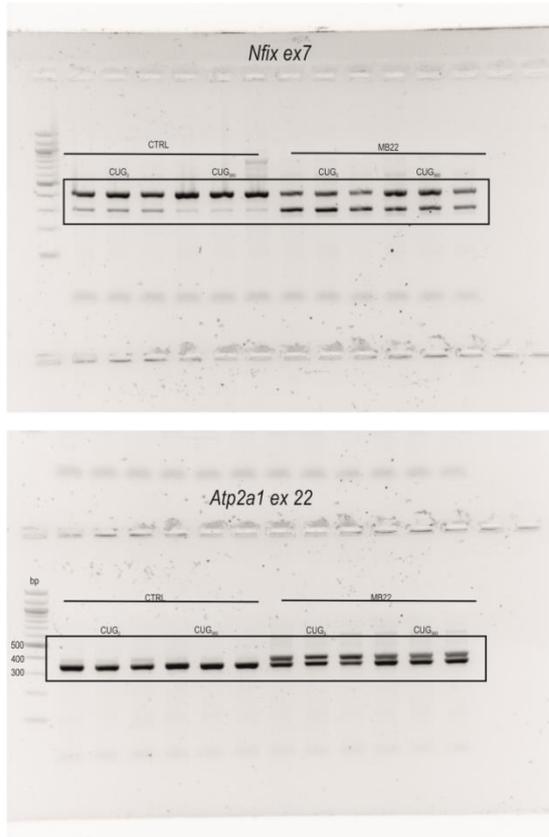


Figure S9 (related to Fig. 3a, b).

a) Full agarose gel image for RT-PCR results show in Fig. 3a. Fragments presented in Fig. 3a are indicated by black rectangles. **b)** Full agarose gel image for RT-PCR results show in Fig. 3b. Fragments presented in Fig. 3b are indicated by black rectangles.

a



b

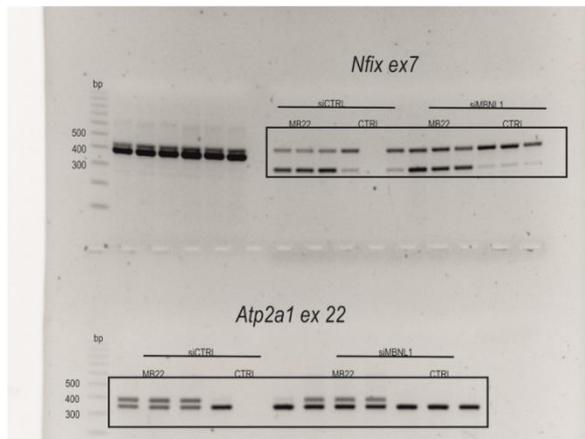


Figure S10 (related to Fig. 4b, c).

a) Full agarose gel image for RT-PCR results show in Fig. 4b. Fragments presented in Fig. 4b are indicated by black rectangles. **b)** Full agarose gel image for RT-PCR results show in Fig. 4c. Fragments presented in Fig. 4c are indicated by black rectangles.

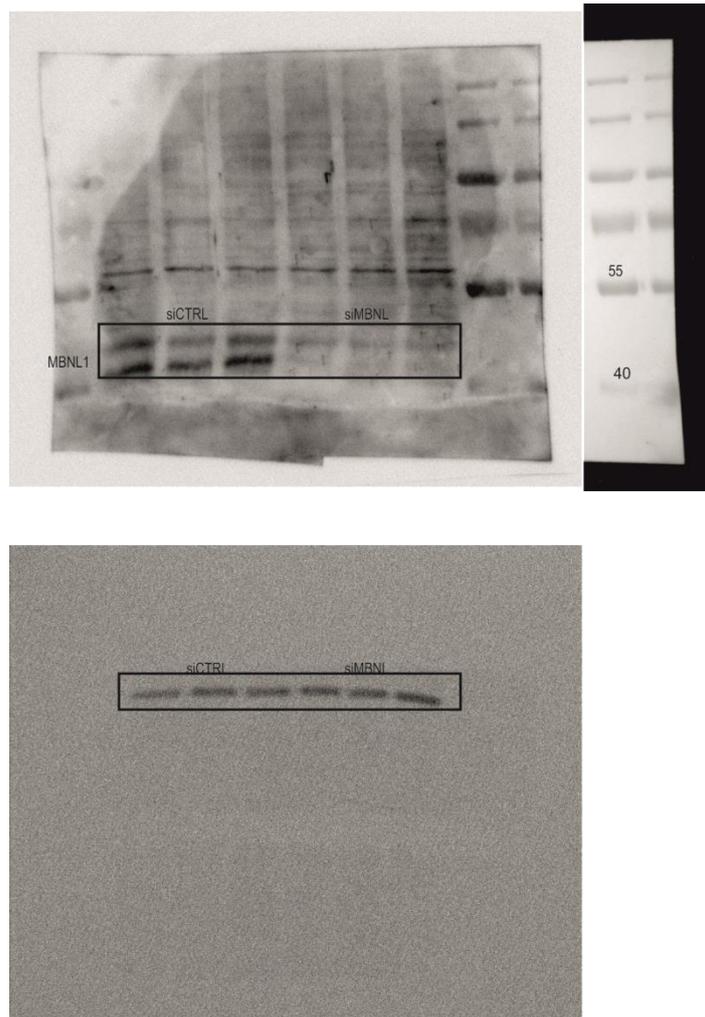


Figure S11 (related to Fig. S5).

Full blot images for western blot results show in Fig. S5. Fragments presented in Fig. S5 are indicated by black rectangles.

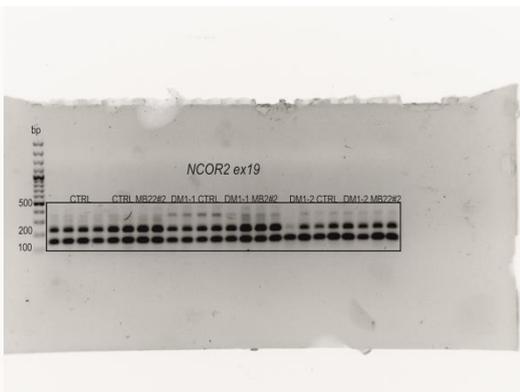
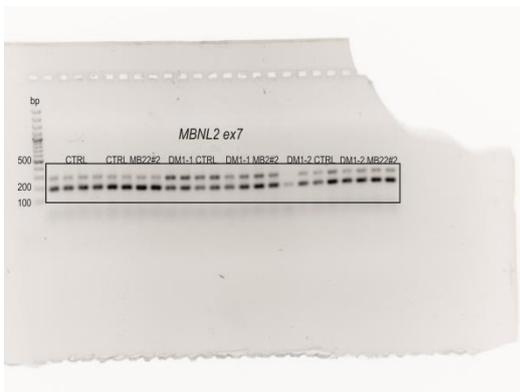
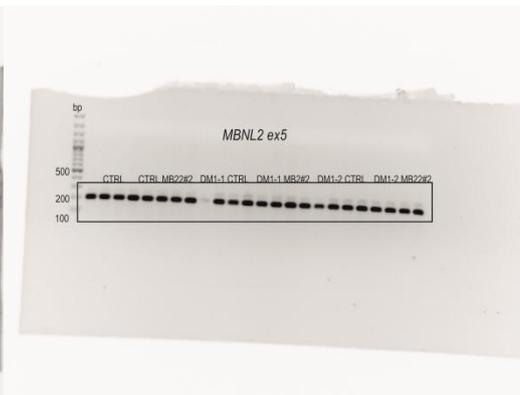
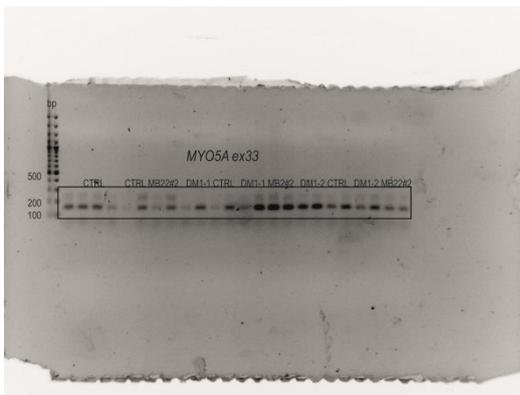
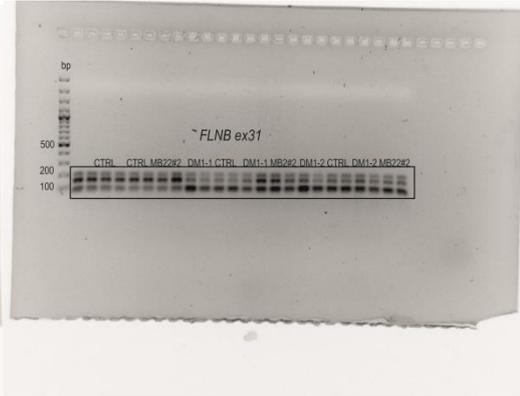
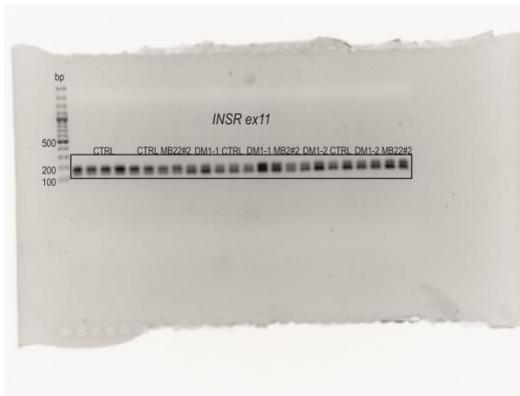


Figure S12 (related to Fig. 5b and Fig. S6).

Full agarose gel image for RT-PCR results show in Fig. 5b and Fig. S6. Fragments presented in Fig. 5a and Fig. S6 are indicated by black rectangles.



Poznań, 12.05.2023

Prof. dr hab. Krzysztof Sobczak, prof. UAM
Zakład Ekspresji Genów

OŚWIADCZENIE

Oświadczam, że w pracy autorstwa Rogalska Z, Sobczak K. zatytułowanej „Sustainable recovery of MBNL activity in autoregulatory feedback loop in myotonic dystrophy”, a opublikowanej w *Molecular Therapy-Nucleic Acids*, 2022; 30: 438-448. doi: 10.1016/j.omtn.2022.10.023. mój udział polegał na stworzeniu koncepcji pracy, pomocy w zaprojektowaniu eksperymentów i interpretacji uzyskanych wyników oraz w przygotowaniu oraz edycji manuskryptu.

Krzysztof Sobczak

Poznań, 2.08.2023

Zuzanna Rogalska

Zakład Ekspresji Genów

OŚWIADCZENIE

Oświadczam, że w pracy autorstwa Rogalska Z., Sobczak K. zatytułowanej „Sustainable recovery of MBNL activity in autoregulatory feedback loop in myotonic dystrophy” a opublikowanej w *Molecular Therapy-Nucleic Acids*, 2022; 30: 438-448. Doi: 10.1016/j.omtn.2022.10.023, mój udział polegał na współudziale w projektowaniu eksperymentów, wykonaniu ich, analizie ich oraz przygotowaniu manuskryptu.

Zuzanna Rogalska
Zuzanna Rogalska

Oświadczenie potwierdzam
K. Sobczak

Funding

The work was supported by the following funding:

1. Faculty of Biology's "Dean's grant" (GDWB-04/2020) to Zuzanna Rogalska
2. MINIGRANT POWER, Interdisciplinary doctoral studies at the Faculty of Biology, Adam Mickiewicz University (POWR.03.02.00-00-I006/17) to Zuzanna Rogalska
3. ID-UB minigrant for doctoral students (017/02/SNP/0018) to Zuzanna Rogalska
4. The Foundation for Polish Science grant (POIR.04.04.00-00-5C0C/17-00) to Krzysztof Sobczak
5. The Polish National Science Centre, grants (UMO-2020/38/A/NZ3/00498) to Krzysztof Sobczak

Academic achievements

During the implementation of the doctoral thesis, I was a co-author of the following publications:

1. Rogalska, Z., & Sobczak, K. (2022). Sustainable recovery of MBNL activity in autoregulatory feedback loop in myotonic dystrophy. *Molecular Therapy - Nucleic Acids*, 30, 438–448. <https://doi.org/10.1016/j.omtn.2022.10.023>
2. Augustyniak, A., Szymański, T., Porzucek, F., Mieloch, A. A., Semba, J. A., Hubert, K. A., Grajek, D., Krela, R., Rogalska, Z., Zalc-Budziszewska, E., Wysocki, S., Sobczak, K., Kuczyński, L., & Rybka, J. D. (2023). A cohort study reveals different dynamics of SARS-CoV-2-specific antibody formation after Comirnaty and Vaxzevria vaccination. *Vaccine*, 41(34), 5037–5044. <https://doi.org/10.1016/j.vaccine.2023.06.008>