Molecular bases of phenotypic and clinical variability in Duchenne and Becker muscular dystrophy

UNIVERSITÀ DEGLI STUDI DI PADOVA

Università degli Studi di Padova

Dipartimento di Scienze Cardiologiche, Toraciche e Vascolari

SCUOLA DI DOTTORATO DI RICERCA IN SCIENZE MEDICHE, CLINICHE E SPERIMENTALI
INDIRIZZO IN NEUROSCIENZE
XXVIII CICLO

"MOLECULAR BASES OF PHENOTYPIC AND CLINICAL VARIABILITY IN DUCHENNE AND BECKER MUSCULAR DYSTROPHY"

Direttore della Scuola: Ch.mo Prof. Gaetano Thiene
Coordinatore d'indirizzo: Ch.ma Prof.ssa Elena Pegoraro
Supervisore: Ch.ma Prof.ssa Elena Pegoraro

Dottorando: Dott. Luca Bello

Index

Abstract 1

Esposizione riassuntiva in lingua italiana 4

Introduction and aims 8

Aim 1: Validation of known genetic modifiers of DMD in the CINRG-DNHS 15
- **Background** 15
- **Methods** 17
- **Results** 21
- **Discussion** 31

Aim 2: DMD genotype-phenotype correlations in the CINRG-DNHS 36
- **Background** 36
- **Methods** 39
- **Results** 42
- **Discussion** 48

Aim 3: Long-term outcomes of glucocorticoid regimens in the CINRG-DNHS 52
- **Background** 52
- **Methods** 54
- **Results** 56
- **Discussion** 63

Aim 4: Genome-wide association study of loss of ambulation in DMD 67
- **Background** 67
- **Methods** 71
- **Results** 75
- **Discussion** 85

Aim 5: Genotype-phenotype correlations in BMD 88
- **Background** 88
- **Methods** 91
- **Results** 94
- **Discussion** 110

Conclusions 117

Acknowledgements 119

References 122

Abstract

Dystrophinopathies are a group of X-linked recessive neuromuscular disorders due to mutations in the *DMD* gene. Truncating mutations, causing dystrophin absence in skeletal and cardiac muscle, cause the more severe form of dystrophinopathy, Duchenne muscular dystrophy (DMD). Conversely, mutations which respect the open reading frame, and give rise to quantitatively or qualitatively altered dystrophin, cause the milder allelic variant known as Becker muscular dystrophy (BMD).

DMD is a devastating disorder. Progressive muscle wasting and weakness causes disability since childhood, and the natural history is characterized by loss of independent ambulation (LoA) around 10 - 15 years of age, and reduced life expectancy because of respiratory and cardiac complications in young adults. Glucocorticoid corticosteroids (GCs) might delay disease progression, and there are promising novel molecular treatments, but a definitive cure remains elusive. Promising molecular treatments include antisense oligonucleotides (AONs) inducing exon skipping in out-of-frame deletions and premature stop codon readthrough compounds.

There is relevant variability in the severity and rate of progression of muscle wasting and weakness in DMD, which is not explained, if not in a minor proportion, by the disease-causing mutation, as all DMD patients have a complete, or near to complete dystrophin defect. Recently, our group and other authors have described genetic modifiers of DMD, i.e. common single nucleotide polymorphisms (SNPs) associated to more or less severe DMD expressivity. These include rs28357094, a SNP in the promoter of the *SPP1* gene, enconding the cytokine osteopontin, and a coding haplotype in the *LTBP4* gene, enconding Latent Transforming growth factor β-Binding Protein 4. These variants modulate the expression (*SPP1*), or alter the aminoacid sequence (for *LTBP4*) of corresponding proteins, both of which are involved in inflammatory and pro-fibrotic pathways. They were both identified by candidate gene approaches (respectively, expression profiling studies and a murine genome scan).

The first aim of this thesis was to provide an independent validation of the genetic association of the *SPP1* rs28357094 SNP and the *LTBP4* haplotype with age at LoA in DMD. This was achieved using data from the Cooperative International Neuromuscular Research Group Duchenne Natural History Study (CINRG-DNHS) of 340 DMD patients from 20 worldwide Centers. In this population, the minor allele G at rs28357094 was associated to a 2-year delay of LoA in CINRG-DNHS participants who had been treated with glucocorticoids ($p < 0.05$), and no significant effect in untreated patients, suggesting that the SNP might be a pharmacodynamic biomarker of GC response. Furthermore, the homozygous *LTBP4* haplotype "IAAM" was associated to a 2-year delay of LoA in participants of European descent ($p < 0.05$), but not in the whole multi-ethinc CINRG-DNHS cohort, highlighting the relevance of population stratification in genetic modifier studies.

The second aim was to test for associations between specific *DMD* mutations and age at LoA in the CINRG-DNHS. We confirmed previous reports that deletions bordering exon 44 (and thus amenable to AON treatment for skipping of this exon), as well as the deletion of exons 3-7, were associated to later LoA ($p < 0.01$ and < 0.05 respectively). These findings have repercussions on clinical trial design and prognosis.

A third aim was to study age at LoA as a long-term outcome of several different GC regimens currently adopted in DMD. In this observational study, we found that the use of daily deflazacort was associated to 2.7-year later LoA than daily prednisone ($p < 0.001$), an unexpected finding that may be confirmed by ongoing randomized trials.

As a fourth aim, we genotyped 175/340 CINRG-DNHS participants with an Exome Chip, including thousands of functional (regulatory or coding) variants, and performed a genome-wide association study (GWAS) of age at LoA in a subgroup of 109 unrelated participants of European ancestry. While no SNP surpassed the Bonferroni-corrected significance threshold, we performed a hypothesis-driven prioritization of findings, focused on inflammatory and pro-fibrotic pathways, and identified a hit in a gene involved in inflammation and cell-mediated immunity. The GWAS association of earlier LoA with the minor allele at the identified locus ($p < 9.9*10^{-5}$) was validated in a

collaborative cohort of 660 DMD patients from the University of Padova, the European Bio-NMD network, and the United Dystrophinopathy Project in the USA ($p < 0.05$).

Finally, the fifth aim focused on BMD, which features a milder, but even more variable clinical picture than DMD. In BMD, most patients have in-frame deletions leading to internally deleted dystrophin protein. We quantified dystrophin by Western Blot, performed a retrospective study of LoA and loss of the ability to run, and a 1-year longitudinal study of motor function (6 Minute Walk Test [6MWT], North Star Ambulatory Assessment [NSAA], timed function tests [TFTs]) in 69 BMD patients at the University of Padova. We found that deletions bordering exon 45 were associated with frequent loss of the ability to run, risk of LoA in adults, and overt muscle weakness; while some deletion groups, like those bordering exon 51, or limited to exon 48, were preserved from these signs of disease progression. This is relevant not only to BMD prognosis and genetic counseling, but also to outcomes of exon skipping AON treatments which aim to reproduce the same deletions at the transcript level in DMD. Furthermore, we observed that NSAA and 6MWT, which we mutuated from DMD studies and had not been applied to BMD, were feasible, clinically meaningful, and able to identify disease progression at 1 year, suggesting their adequacy as outcome measure for future BMD clinical trials.

Altogether, the work presented here provides novel insights into the mechanisms of phenotypic and clinical variability in dystrophinopathy, which will be useful in delivering improved care for these disabling diseases.

Esposizione riassuntiva in lingua italiana

Le distrofinopatie sono un gruppo di malattie legate al cromosoma X in modalità recessiva, dovute a mutazioni nel gene *DMD*. La forma più severa, la distrofia muscolare di Duchenne (DMD), è causata da mutazioni troncanti, che provocano una completa assenza della proteina distrofina nel muscolo scheletrico e cardiaco. Al contrario, la forma allelica più mite, nota come distrofia muscolare di Becker (BMD), è causata da mutazioni che rispettano la cornice di lettura del gene, dando origine a distrofina quantitativamente e/o qualitativamente alterata.

La DMD è una malattia devastante, caratterizzata da una progressiva degenerazione del tessuto muscolare, con deficit di forza e disabilità sin dall'infanzia. La storia naturale è caratterizzata da perdita della deambulazione autonoma attorno all'età di 10 - 15 anni, e da ridotta aspettativa di vita a causa di complicanze respiratorie e cardiache nei giovani adulti. I glucocorticoidi (GC) possono ritardare la progressione della malattia, ma una cura definitiva non è ancora disponibile. Fra i trattamenti molecolari innovativi più promettenti, annoveriamo gli oligonucleotidi antisenso (AON) che inducono l'"exon skipping" di specifici esoni nelle mutazioni "out-of-frame", e i composti che inducono il "readthrough" dei codoni di stop prematuri.

Vi è una rilevante variabilità nella severità e velocità di progressione del deficit di forza e della degenerazione del tessuto muscolare nella DMD, che non si spiega, se non in piccola parte, in base alle diverse mutazioni patogenetiche, dal momento che tutti i pazienti presentano una completa o quasi completa assenza di distrofina.
Recentemente, il nostro gruppo e altri autori hanno descritto modificatori genetici della DMD, cioè polimorfismi di singolo nucleotide (SNP) associati a espressività più o meno severa del fenotipo DMD: lo SNP rs28357094 nel promotere del gene *SPP1*, codificante per la citochina osteopontina, e un aplotipo codificante nel gene *LTBP4* (Transforming growth factor β-Binding Protein 4). Queste varianti modulano l'espressione (*SPP1*) o alterano la sequenza aminoacidica (*LTBP4*) delle rispettive proteine, entrambe le quali sono coinvolte in vie di segnale pro-infiammatorie e pro-fibrotiche. Questi geni sono

stati identificati come modificatori candidati con approcci diversi (rispettivamente, studi di profili di espressione e mappatura genomica di un modello murino).

Il primo obiettivo di questa tesi è stato di ottenere una validazione indipendente dell'associazione genetica dello SNP rs28357094 nel gene *SPP1* e dell'aplotipo *LTBP4* con l'età di perdita della deambulazione nella DMD. Questo risultato è stato conseguito utilizzando dati raccolti nel Duchenne Natural History Study del Cooperative International Neuromuscular Research Group (CINRG-DNHS), condotto su 340 pazienti DMD in 20 Centri in tutto il mondo. In questa coorte, l'allele minore G dello SNP rs28357094 era associato a un prolungamento della deambulazione di 2 anni nei partecipanti al CINRG-DNHS che erano stati trattati con GC ($p < 0.05$), ma a nessun effetto nei partecipanti non trattati. Ciò suggerisce che questo SNP potrebbe essere un biomarcatore farmacodinamico di risposta ai GC. L'aplotipo omozigote "IAAM" di *LTBP4*, invece, era associato a un prolungamento della deambulazione di 2 anni nei partecipanti di origine europea ($p < 0.05$), ma non nell'intera coorte multietnica CINRG-DNHS, evidenziando la rilevanza della stratificazione di popolazione negli studi sui modificatori genetici.

Il secondo obiettivo è stato rivolto alla identificazione di eventuali associazioni fra specifiche mutazioni *DMD* e perdita della deambulazione nel CINRG-DNHS. Le delezioni confinanti con l'esone 44 (e quindi eleggibili per trattamento con AON che provocano lo "skipping" di questo esone) erano associate a perdita della deambulazione più tardiva, così come la delezione degli esoni dal 3 al 7 (rispettivamente $p < 0.01$ e < 0.05). Questi risultati potranno avere ripercussioni rilevanti sulla prognosi e sulla progettazione di trial clinici.

Come terzo obiettivo, abbiamo studiato l'età alla perdita della deambulazione come esito a lungo termine di diversi regimi di GC attualmente utilizzati nella DMD. In questo studio osservazionale, abbiamo osservato che l'uso del deflazacort quotidiano era associato perdita della deambulazione più tardiva di 2.7 anni rispetto al prednisone quotidiano ($p < 0.001$), un risultato inaspettato che potrebbe essere confermato da studi clinici randomizzati attualmente in corso.

Come quarto obiettivo, abbiamo genotipizzato 175/340 partecipanti al CINRG-DNHS con Exome Chip, un chip di genotipizzazione che include migliaia di varianti funzionali (regolatorie e codificanti), e abbiamo condotto un "genome-wide association study" (GWAS) del fenotipo :età alla perdita della deambulazione in un sottogruppo di 109 pazienti di origine europea (o europea-americana) non imparentati fra loro. Dal momento che nessuno SNP sorpassava la soglia di significatività corretta per test multipli secondo Bonferroni, abbiamo priorizzato i risultati focalizzandoci sui geni coinvolti in vie di segnale pro-infiammatorie e pro-fibrotiche. Abbiamo identificato così uno SNP in un gene coinvolto nell'infiammazione e nell'immunità cellulo-mediata. L'associazione dell'allele minore al locus identificato con una più precoce perdita della deambulazione ($p < 9.9*10^{-5}$) è stata successivamente validata in una coorte collaborativa di 660 pazienti DMD raccolta presso l'Università di Padova, il network Bio-NMD, e il network statunitense United Dystrophinopaty Project ($p < 0.05$).

Infine, il quinto obiettivo si è rivolto alla BMD, che presenta un quadro clinico più mite, ma anche più variabile rispetto alla DMD. La maggioranza dei pazienti BMD ha delezioni che rispettano la cornice di lettura con espressione di distrofina internamente deleta. Abbiamo quantificato la distrofina tramite Western Blot, e condotto uno studio retrospettivo della perdita della deambulazione e della capacità di correre, e uno studio longitudinale di 1 anno delle funzioni motorie ("6 Minute Walk Test" [6MWT], "North Star Ambulatory Assessment" [NSAA] e "timed function tests" [TFTs]) in 69 pazienti BMD seguiti presso l'Università di Padova. Abbiamo osservato che le delezioni che terminano sull'esone 45 presentavano frequente perdita della capacità di correre, occasionale perdita della deambulazione e un franco deficit di forza, mentre altri gruppi di delezioni, come quelle che terminano sull'esone 51, o limitate al solo esone 48, presentavano una funzione muscolare migliore. Questi risultati sono rilevanti non solo per la prognosi e la consulenza genetica nella BMD, ma consentono anche di poter predire un possibile esito della terapia con AON per la DMD. Gli AON mirano infatti a riprodurre a livello di trascritto le medesime delezioni osservate nei pazienti BMD. Inoltre, abbiamo osservato che le misure funzionali 6MWT e NSAA, mutuate da studi

sulla DMD e non ancora utilizzate nella BMD, sono di semplice applicazione, clinicamente rilevanti, e in grado di identificare la progressione di malattia a un anno. Questi risultati suggeriscono che esse siano appropriate misure di esito funzionale in futuri studi clinici.

Complessivamente, i lavori qui presentati offrono nuove prospettive sui meccanismi di variabilità fenotipica e clinica nelle distrofinopatie, che risulteranno utili per offrire cure migliori ai pazienti affetti da queste malattie disabilitanti.

Introduction and aims

Dystrophinopathies are a group of genetic disorders mainly affecting skeletal and cardiac muscle, caused by deficiency of the protein dystrophin in the sarcolemma of muscle fibers [Hoffman et al., 1987]. Dystrophin is encoded by the Duchenne muscular dystrophy gene (DMD), linked to the short arm of the X chromosome at the Xp21 locus. Dystrophinopathies are inherited as X-linked recessive disorders, hemizygote males being affected by the disease and heterozygote females being mostly asymptomatic carriers. The mutation rate is high, with about one third of cases with negative familiar history arising from de novo mutations [van Essen et al., 1992].

Duchenne muscular dystrophy (DMD, MIM #310200) is the severe form of dystrophinopathy caused by a complete deficiency of dystrophin protein, affecting 1 in 3800 ~ 5000 newborn males [Mostacciuolo et al., 1987; van Essen et al., 1992; Stark, 2015] and representing one of the most common lethal childhood disorders. Boys affected with DMD suffer loss of independent ambulation around the beginning of the second decade, and death by dilated cardiomyopathy (DCM) and/or respiratory insufficiency in the third or fourth decade. Becker muscular dystrophy (BMD, MIM #300376) is the relatively milder condition caused by the presence of a reduced and/or partially functional dystrophin protein [Monaco et al., 1988; Hoffman et al., 1988; Hoffman et al., 1989]. BMD phenotypes are highly variable, ranging from severe forms, which are similar to DMD, to milder clinical pictures with retainment of independent ambulation through adulthood; the association of DCM is likewise variable, and strongly influences life expectancy. Disease manifestations are relatively rare in DMD/BMD carriers (~10%) [Moser and Emery, 1974], but may include severe phenotypes [Pegoraro et al., 1995].

A partial dystrophin defect may also cause familiar X-linked DCM (MIM #302045) with no substantial muscular abnormalities [Melacini et al., 1996; Ferlini et al., 1999] or mild phenotypes such as quadriceps myopathy, cramp and myalgia syndrome with or

without myoglobinuria [Gospe et al., 1989; Doriguzzi et al., 1993], and asymptomatic elevation of serum creatin kinase (CK) [Morrone et al., 1997].

To date, dystrophinopathies have no definitive cure, but the implementation of a multidisciplinary care protocol comprising a combination of palliative pharmacological interventions (i.e. steroid therapy) [Bushby et al., 2010a], cardiological therapy, antibiotic therapy, ventilatory assistance, physiotherapy, osteoarticular surgery, and screening for prevention of complications, has led to a substantial improvement of life quality and expectancy in DMD and severe forms of BMD [Sejerson et al., 2009; Bushby et al., 2010b]; moreover, increasing knowledge about the molecular basis of dystrophinopathies has offered patients and families accurate genetic counseling, and, more recently, experimental molecular and genetic therapies [Guglieri and Bushby, 2010; Hoffman et al., 2011].

BMD encompasses a wide spectrum of clinical manifestations. The severe end of this spectrum includes patients with a clinical presentation almost indistinguishable from DMD, despite the presence of dystrophin protein detected by biochemical essays [Hoffman et al., 1991]; the mild end comprises adult-onset disorders without any clinical manifestation in the first 3-4 decades of life, independent ambulation being retained even in the old age. Many different intermediate phenotypes are possible. In “typical” forms, the onset of symptoms is around 12 years of age, loss of ambulation may ensue from adolescence onward, and life expectancy is around 50 years [Emery et al., 2002]. The incidence of BMD is about 5 times lower than DMD (1 in ~ 18,000 live born males) [Mostacciuolo et al., 1987], but may be partially underestimated because of undiagnosed mild cases. As in DMD, muscle weakness and wasting are evident proximally more than distally, and usually greater in lower than in upper limbs. The anterior compartment of the thigh is more prominently affected, and exclusively affected in the “quadriceps myopathy” phenotype. Calf hypertrophy is very frequent and joint retractions may develop in the presence of relevant weakness, especially at the ankles. Hyperlordosis and scoliosis are less common than in DMD, and are usually

observed in the severe end of the phenotypical spectrum. Macroglossia is rarely observed. Atypical clinical presentation is much more common in BMD than in DMD: possible examples are cramps and recurrent myoglobinuria [Bushby et al., 1991] or life-threatening adverse reactions to general anesthesia [Bush and Dubowitz, 1991]. Although some authors include pauci- or asymptomatic forms of dystrophinopathy in the BMD spectrum, the nosographic definition of BMD does not correctly apply to patients without actual dystrophic alterations of skeletal muscle, whom may be better described as having mild dystrophinopathy.

Despite DMD patients all carrying (by definition) truncating mutations which cause a complete dystrophin defect, extensive natural history studies have shown considerable inter-patient variation in DMD onset and progression [McDonald et al., 2013; Henricson et al., 2013; Pane et al., 2014a]. Daily treatment with glucocorticoids is considered standard of care [Bushby et al., 2010a], yet there is variable patient-patient response to treatment, both in terms of efficacy (improved gross motor skills) and side effect profiles [Henricson et al., 2013; Bello et al., 2015a]. Also, there is considerable variation in the practice of prescription and use of glucocorticoids [Griggs et al., 2013; Bello et al., 2015b].

Multiple factors contribute to the observed clinical variation in DMD, including standards of care, glucocorticoid use, and genetic background (modifying polymorphisms, ethnicity). These variables are not independent. For example, a patient's geographic location has effects on ethnicity, standards of care, and glucocorticoid use [Bello et al., 2015b]. The inter-dependent nature of these variables can make it challenging to isolate a single variable and define its contribution to disease expressivity. Furthermore, specific types of DMD mutation, despite being predicted to be out-of-frame at the genomic level, may show relatively milder phenotypes because of alternative splicing, as demonstrated for mutations in the N-terminal domain [Muntoni et al., 1994; Winnard et al., 1995; Gualandi et al., 2006], some nonsense

mutations [Flanigan et al., 2011], or deletions bordering exon 44 [van den Bergen et al., 2014a; Pane et al., 2014b].

Despite these various sources of clinical variability and the challenges in identifying their specific roles, the identification of modifier variants of DMD in other genes, acting in epistasis with the disease-causing mutation, is broadly felt to be important, and initial progress has been made using a candidate gene approach [Vo and McNally, 2015]. The *SPP1* gene locus, encoding the osteopontin protein (OPN), was identified as a candidate by mRNA profiling studies of muscle biopsies from clinically mild and severe patients [Pegoraro et al., 2011]. A polymorphism in the promoter of the *SPP1* gene (rs28357094) known to change SPP1 mRNA expression by 5-fold, was found to be associated with age at loss of ambulation (LoA) in an Italian cohort, and grip strength in the Cooperative International Neuromuscular Research Group Duchenne Natural History Study (CINRG-DNHS) [Pegoraro et al., 2011]. The association of rs28357094 with ambulation phenotypes was validated in a second Italian multi-center cohort [Bello et al., 2012]. However, the same association of rs28357094 with LoA was not seen in a European multicenter cohort (Bio-NMD) [van den Bergen et al., 2015] and in a United States multicenter cohort (United Dystrophinopathy Project, UDP) [Flanigan et al., 2013]. As the proportion of patients treated with glucocorticoids vary from cohort to cohort, glucocorticoid treatment is an important covariate to include in statistical models studying genetic modifiers, and the variable rate of glucocorticoid use in different cohorts might confounded validation results. The effects of the rs28357094 *SPP1* locus was also studied in adult volunteer populations, where it was found to be associated with upper arm muscle volume [Hoffman et al., 2013] and response to eccentric activity-induced muscle damage [Barfield et al., 2014] in young adult females.

A second modifier of muscular dystrophy was identified by genetic linkage with disease severity in an outbred murine γ-sarcoglycan deficient model [Heydemann et al., 2009], which pointed to the *Ltbp4* murine gene. This finding translated to human DMD, as a coding haplotype in *LTBP4* was associated with age at LoA [Flanigan et al., 2013]. This

association was replicated in the Bio-NMD cohort [van den Bergen et al., 2015] and in participants of European ancestry in the CINRG cohort [Bello et al., 2015a], although not in an Italian multi-center cohort [Barp et al., 2015].

The candidate gene approach has led to the identification of two robust genetic modifier loci, *SPP1* and *LTBP4*. The products of both these genes are involved in secondary inflammation and tissue remodeling, which appear to be key modifier pathways of muscular dystrophy [Chen at al., 2005; Ceco and McNally, 2014]. *SPP1* is heavily upregulated by activation of NF-κB signaling during bouts of degeneration and regeneration in skeletal muscle [Hoffman et al., 2013]. NF-κB signaling is triggered in DMD by necrotic cells liberating damage-associated molecular patterns, which stimulate Toll-like receptors (TLRs) [Rosenberg et al., 2015]. In turn, upregulated osteopontin modulates TGFβ-mediated signals [Vetrone et al., 2009], which dictate either successful (reparative) or unsuccessful (fibrotic) regeneration. The LTBP4 protein also regulates TGFβ signaling by binding TGFβ in a latent complex in the extracellular matrix [Flanigan et al., 2013], and modifier haplotypes appear to influence susceptibility to proteolitic cleavage and subsequent TGFβ signaling activation [Ceco et al., 2014].

However, the candidate gene approach for identification of genetic associations is known to be subject to various biases [Wills et al., 2009]. Genome-wide association studies (GWAS) are broadly felt to show less bias in genetic association discovery. GWAS studies typically employ highly parallel statistical tests of genetic association, where thousands to millions of polymorphic loci are tested simultaneously. In order to withstand multiplicity adjustment of association p-values, large populations have to be studied for typical effect sizes. This requirement works against utilization of GWAS to identify rare monogenic disease modifiers, as collecting and phenotyping large patient populations is challenging, so that specific methods would be needed to overcome this issue.

BMD, compared to DMD, presents milder, but even more variable skeletal muscle wasting and weakness. Genotype-phenotype correlation studies have shown that loss of

functionally crucial actin- or dystroglycan-binding domains, respectively at the N- or C-terminus, can result in DMD-like phenotypes despite detectable protein [Hoffman et al., 1991; Aartsma-Rus et al., 2006; Kesari et al., 2008]. Deletions in the large dystrophin rod domain, which harbours the majority of BMD causing mutations, might differently affect the physical properties of resulting internally deleted dystrophin, depending on the preservation or loss of structural "phase" between spectrin repeats and hinge regions [Kaspar et al., 2009]. Some specific deletions, such as deletions of in-frame exons in the proximal rod domain [Angelini et al., 1994a], and deletions including the hinge 3 domain encoded by exons 50 and 51 [Carsana et al., 2005; Anthony et al., 2011], have been more frequently associated to mild or asymptomatic cases; while frequently observed deletions situated in the DMD mutational hotspot around exon 45-53 [White and Den Dunnen, 2006], but not including exons 50-51, have been linked with a typical clinical picture of BMD [Bushby et al., 1993; Anthony et al., 2014a; van den Bergen et al., 2014b]. Moreover, a quantitative correlation between dystrophin content in muscle and BMD severity has been described, both as a linear or threshold effect [Angelini et al., 1994a; Angelini et al., 1996; Comi et al., 1994; Anthony et al., 2011; van den Bergen et al., 2014b]. Recently, renewed interest has been kindled in this field, as some BMD-causing deletions can be regarded as naturally occurring models of the in-frame deletions produced at the transcript level by splice-modulating antisense oligonucleotides (AONs), employed in the treatment of DMD with the exon skipping approach [Aartsma-Rus, 2012; Arechevala-Gomez et al., 2012].

In summary, the mechanisms underlying phenotype variability in Duchenne and Becker muscular dystrophy are manifold and complex. Their relevance for potential treatments is intuitive, because a molecular mechanism which can modify disease severity could be pharmacologically modulated with the purpose of slowing disease progression. Furthermore, genetic variants associated to faster or slower disease progression could be used as prognostic biomarkers for better design of clinical trials and interpretation of their results. The object of this thesis is to improve our knowledge of these mechanisms. To this end, we formulated the following specific aims:

Aim 1: Validate known genetic modifiers of DMD, i.e. the *SPP1* rs28357094 single nucleotide polymorphism (SNP) and the *LTPB4* "IAAM" haplotype, in the Cooperative International Neuromuscular Research Group Duchenne Natural History Study (CINRG-DNHS).

Aim 2: Assess phenotype differences between participants in the CINRG-DNHS with different types of out-of-frame DMD mutations.

Aim 3: Assess phenotype differences between participants in the CINRG-DNHS who had been treated with different glucocorticoid corticosteroid (GC) regimens.

Aim 4: Identify novel genetic modifier genes of DMD by a genome-wide association study (GWAS) of age at loss of ambulation (LoA) in the CINRG-DNHS.

Aim 5: Assess phenotype differences between BMD patients with different types of in-frame DMD mutations in the BMD cohort referring to the Neuromuscular Clinic at the University of Padova.

Note that while this Introduction provides the background which led to the formulation of the specific Aims, and serves as a preamble to the whole thesis, in each of the following five Chapters, each dedicated to a specific Aim, more focused and detailed background information is provided, so that each Chapter may be read as a stand-alone research paper. Therefore, there might be redundancy between this Introduction and the Background section of each Chapter.

Aim 1: Validation of known genetic modifiers of DMD in the CINRG-DNHS

Background

Duchenne muscular dystrophy (DMD) is caused by the absence of the protein dystrophin in myofibers, due to truncating dystrophin gene mutations [Hoffman et al., 1987]. Despite this homogeneous molecular defect, variability in phenotype severity is commonly observed, e.g. variable age at loss of ambulation (LoA). This is due to environmental factors, such as implementation of standards of care (glucocorticoid corticosteroid [GC] treatment, physical therapy, management of contractures, fracture prevention) [Bushby et al., 2010a-b], and to the genetic background. Two genetic modifiers of DMD, i.e. common polymorphisms that modulate disease severity combined with a pathogenic mutation, have been described: a single nucleotide polymorphism (SNP) in the promoter of the *SPP1* (secreted phosphoprotein 1, or osteopontin) gene, and a coding *LTBP4* (Latent Transforming growth factor β Binding Protein 4) haplotype. The association of the *SPP1* rs28357094 rare G allele with earlier LoA, in a dominant inheritance model, was originally reported in 106 Italian DMD patients [Pegoraro et al., 2011]. *SPP1* encodes an inflammatory cytokine involved in tissue damage response, and is part of the transforming growth factor β (TGFβ) pathway [Ceco and McNally, 2013]. The rs28357094 polymorphism alters transcription, at baseline [Giacopelli et al., 2004] and in response to steroid hormones [Barfield et al., 2014].

The *LTBP4* locus was identified by genome-wide mapping in a murine model of muscular dystrophy [Heydemann et al., 2009]. Subsequently, a *LTBP4* haplotype was associated with variable LoA in 254 patients with severe dystrophinopathy (United Dystrophinopathy Project, UDP) [Flanigan et al., 2013]. The haplotype consists of four coding SNPs in strong linkage disequilibrium (LD), one of which, rs10880, was independently associated with age at LoA. Homozygotes for the minor allele T at

rs10880 (T1140M), in LD with the haplotype IAAM, showed later LoA. The proposed mechanism is that the "IAAM" protein isoform results in a more stable latent TGFβ complex, reducing TGFβ signaling. In the same paper, the authors found no association of *SPP1* genotype with age at LoA.

Validation of genetic associations in independent cohorts is essential to establish genetic modifiers of Mendelian diseases [Nelson and Griggs, 2011], but may be exaggerated or obscured by confounding variables, such as ancestry-dependent differences in allele frequency and haplotype configuration, which associate with variations of standards of care and other environmental factors, and lead to population stratification [Enoch et al., 2006; Tian et al., 2008; Genin et al., 2008]. Disparities in diagnostics [Holtzer et al., 2011], standards of care [Fox et al., 2014], and phenotype severity [Bortolini and Zatz, 1987] between DMD patients of different ethnic backgrounds have in fact been reported. The Cooperative International Neuromuscular Research Group Duchenne Natural History Study (CINRG-DNHS) comprises participants from 20 Centers in 4 continents, constituting an ethnically diverse cohort.

We have expanded analysis of the CINRG-DNHS cohort, from the baseline cross-sectional analysis of grip strength in 156 participants, to a longitudinal study (average follow-up 4 years) of all 340 participants [McDonald et al., 2013; Henricson et al., 2013]. Here we sought to test the effect of *SPP1* and *LTBP4* genotypes on LoA in the CINRG-DNHS population, controlling for GC treatment and population stratification.

Methods

The institutional review board or ethics review board at each participating Institution approved the study protocol, consent and assent documents. Informed consent/assent was obtained for each participant prior to conducting study procedures.

Inclusion and exclusion criteria. The inclusion and exclusion criteria for the CINRG-DNHS have been previously described [McDonald et al., 2013; Henricson et al., 2013]. Recruitment was aimed at obtaining a population representing an age span from very young to adult (age 2-28 years at baseline). On the other hand, recruitment was not specifically aimed at obtaining sub-populations with homogeneous ancestry for genetic association analysis. For all analyses focused on *SPP1* and *LTBP4* genotypes, we excluded patients with no available genomic DNA for SNP genotyping.

LoA and GC treatment definitions. LoA was defined as patient-reported continuous wheelchair use, verified by inability to walk 10 meters unassisted. GC treatment history was recorded both retrospectively at baseline, and longitudinally during the study, and the population was dichotomized into treated at least 1 year with GCs before LoA, and untreated or treated less than 1 year before LoA. This included patients who had gone on and off GC treatment one or more times, but were cumulatively treated for at least 1 year before LoA. The 1-year treatment threshold was chosen based on the clinical rationale that a long term effect of GC treatment, like delaying disease milestones, cannot be reasonably expected from a short-term treatment. Data were also re-analyzed with a 6-month treatment threshold, as in Flanigan et al., in order to compare *LTBP4* data with the same methodology as the original report of this modifier.

Race and ethnicity definitions. Self-identification of participants into one of the following racial categories was recorded: African American, Asian, Caucasian, Mixed, or Other; self-identification into non-Hispanic or Hispanic ethnicity was specified as separate option, according to official U.S. census categories (http://www.whitehouse.gov/omb/fedreg_1997standards/). Here, we clustered this

information into the following groups: African American, Asian, non-Hispanic Caucasian, Hispanic Caucasian, Hispanic (i.e. participants self-identifying their ethnicity as Hispanic, and race as Mixed/Other), and Other (i.e. participants self-identifying their ethnicity as non-Hispanic, and race as Mixed or Other); and distinguished as South Asian those participants recruited at the Study Center in Chennai, India (for these participants no DNA samples were available because of local regulations that did not allow the shipment of DNA for the purposes of our study).

Targeted genotyping. Targeted genotyping was carried out by TaqMan allele discrimination assays. For *SPP1* rs28357094, a dominant model for the minor allele G was adopted [Pegoraro et al., 2011]. For *LTBP4* association studies, we focused mainly on the rs10880 SNP (T1140M), which is in strong LD with the IAAM haplotype, and showed an independent association with LoA in the original report, in a recessive model [Flanigan et al., 2013]. Focusing initially on a single significant SNP allowed inclusion of larger numbers of participants in the analyses. In order to confirm association with the full haplotype in carriers of the rs10880 T allele, we genotyped rs2303729 (V194I), rs1131620 (T787A), and rs1051303 (T820A). *LTBP4* haplotypes were phased with PLINK [Purcell et al., 2007] and median ages at LoA were calculated for all observed haplotype configurations, in all those patients for whom haplotypes could be phased with at least 90% probability.

MDS analysis. Genotype data from the Illumina HumanExome Chip was available for 175 participants. These were not selected by ethnicity, nor any phenotype-related characteristic, but solely on the base of available DNA quantity and quality. While mainly focusing on coding regions, the chip contains ancestry markers and other common variant markers (~30,000) that ensure sufficient genome-wide coverage for multidimensional scaling (MDS) analysis. Chip design information is publicly available at http://genome.sph.umich.edu/wiki/Exome_Chip_Design#Illumina_Exome_Arrays. Genotypes were called from raw intensity data with the Genome Studio software, and data was exported into PLINK format with the dedicated plug-in software by Illumina.

PLINK was used for data cleaning and MDS analysis [Neale and Purcell, 2008]. Subjects with > 10% missed calls and SNPs with > 5% missed calls were removed to ensure data quality. MDS analysis was based on the calculation of genome-wide identity-by-state pairwise distances, on a set of "pruned" genome-wide markers in no significant LD with each other, using the PLINK whole genome association analysis toolset. LD-based pruning parameters were the following: 50 SNP window size, 5 SNP window slide at each step, variance inflation factor threshold = 2. The two highest-ranking principal components were plotted.

Grouping criteria for analyses of race/ethnicity differences in GC treatment and LoA. Median LoA and GC treatment rates were calculated in participants with different self-identified race/ethnicities in the whole DNHS cohort of 340 patients, grouped as defined above.

Grouping criteria for analyses of SNP effect on LoA. SNP effects on LoA were analyzed in the whole cohort with available genomic DNA for genotyping, and in a subcohort of participants of Caucasian ancestry, identified by MDS analysis. In both cohorts, analyses were carried out in three groupings based on GC treatment: all participants regardless of treatment; GC-treated participants (at least 1 year while ambulatory as defined above); GC-untreated participants (including treated less than 1 year while ambulatory, as defined above). Data were also re-analyzed with a 6-month GC treatment threshold as explained above. A flow diagram of participant grouping, with corresponding planned analyses, participant numbers, and GC treatment rates, is shown in Figure 1.

Statistical methods. Rates of GC treatment between self-identified racial and ethnic subgroups were compared by χ^2 test. LoA was studied in a time-to-event model with age as the time variable, and LoA as the failure event. Ambulatory participants were censored at the age of last follow-up. Median ages at LoA in race/ethnicity and genotype subgroups were based on the empiric survival curve from a Kaplan-Meier (KM) curve calculation, and compared by log-rank test. For *SPP1* and *LTBP4* genotypes, patients were grouped based on inheritance models specified above. Additionally,

concurrent effects of GC treatment and SNP genotype were analyzed in a Cox regression model, with a time-varying GC treatment covariate (on/off treatment defined for all participants at the time of each event). Statistical significance was set at $p < 0.05$. Partek Genomics Suite 6.6 (St. Louis, MO) and STATAV13 (College Station, TX) were used for statistical analyses.

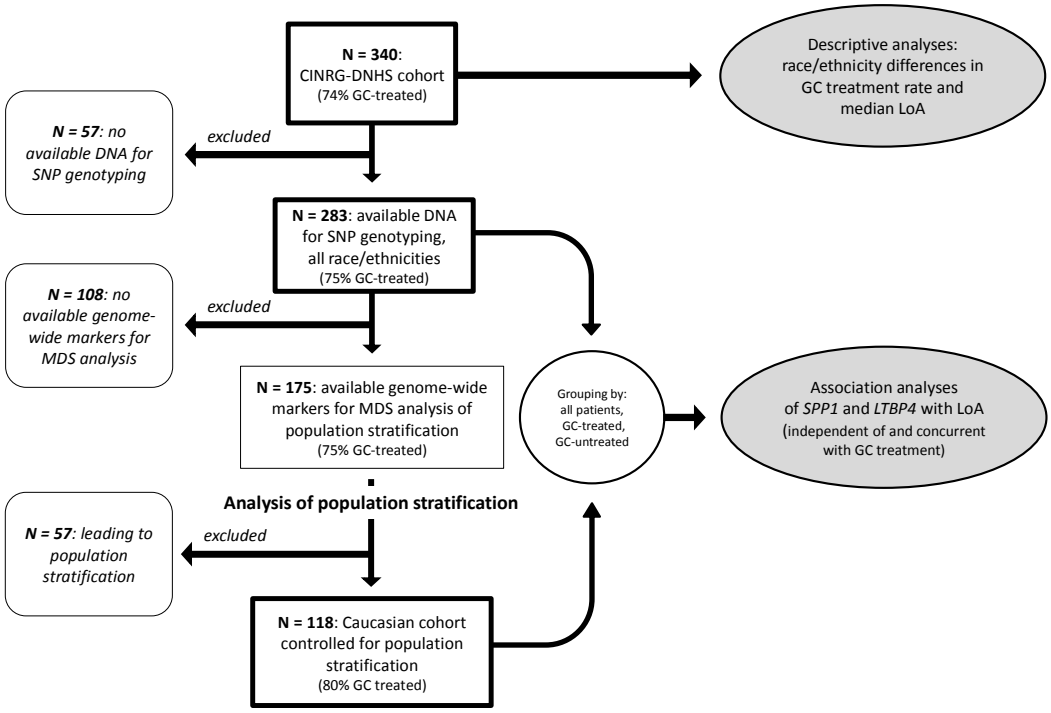

Figure 1. Flow diagram of analysis plan and population grouping. Subgroups included in different analyses are shown, starting from the top with the whole DNHS cohort, and in subsequent steps excluding patients with no available DNA for genotyping; subjects with no available genome-wide markers for multidimensional scaling analysis for population stratification; subjects leading to population stratification. Thick-border boxes indicate groups selected for specific analyses.

Results

Distribution by race/ethnicity and GC treatment. Self-identified race/ethnicity in the 340 participants enrolled in the CINRG-DNHS was distributed as follows: 225 (66%) Caucasian, 23 (7%) Hispanic-Caucasian, 41 (12%) South Asian, 18 (5%) Hispanic, 14 (4%) Asian, 6 (2%) African American, and 13 (4%) Other. Of note, the South Asian group was the only one entirely referring to a single Study Center (Chennai, India). GC treatment was administered for at least 1 year before LoA to 252/340 participants (74%). There were differences in the proportion of participants treated for at least 1 year with GCs, between self-identified racial-ethnic subgroups. Overall, participants self-identifying as Caucasian (both Hispanic and non-Hispanic ethnicity) were more often treated for at least 1 year than other participants: 191/248 (77%) vs. 61/92 (65%), χ^2 p = 0.045 (Table 1). This comparison might be biased by different proportions of younger, ambulatory GC-naïve participants, as opposed to non-ambulatory participants who did not receive treatment before LoA. When analyzing non-ambulatory participants only, we found a trend in the same direction: non-ambulatory Caucasians treated at least 1 year while ambulatory were 113/162 (70%) vs. other ethnicities 39/67 (58%, χ2 p = 0.09, Table 1).

Table 1. Number and percentage of participants treated with GCs at least 1 year while ambulatory, grouped by self-identified race and ethnicity.

Self-identified race and ethnicity		All participants		Non-ambulatory		Ambulatory	
African American		2/6	(33%)	0/4	(0%)	2/2	(100%)
Asian		9/14	(64%)	8/13	(62%)	1/1	(100%)
Caucasian	non-Hispanic	173/225	(77%)	102/146	(70%)	71/79	(90%)
	Hispanic	18/23	(78%)	11/16	(69%)	7/7	(100%)
Hispanic		8/18	(44%)	6/16	(37%)	2/2	(100%)
Other		12/13	(92%)	5/6	(83%)	7/7	(100%)
South Asian		30/41	(73%)	20/28	(71%)	10/13	(77%)

GC: glucocorticoid corticosteroids.

Age at LoA and GC treatment show differences between self-reported ethnic subgroups. Median age at LoA was 12.4 years in the non-Hispanic Caucasian population (n = 225). Compared to this numerically predominant group, median age at LoA was significantly earlier in the Hispanic (n = 18, 9.7 years, log-rank p = 0.003) and South Asian (n = 41, 10.4 years, p < 0.001) subpopulations. Median LoA was earlier in the Asian subpopulation (n = 14, 11.3 years), and later in Hispanic-Caucasian (n = 23, 13.0 years) and African American subpopulations (n = 6, 14.2), but differences with non-Hispanic Caucasians were not statistically significant.

Table 2. Minor allele frequencies (MAFs) for *SPP1* rs28357094 and *LTBP4* rs10880 compared to 1000 Genomes Project MAFs.

Population	*SPP1* rs28357094		*LTBP4* rs10880	
	DNHS	**1000G**	**DNHS**	**1000G**
African American (n = 6)	0.10	0.04 (AFR)	0.30	0.51 (AFR)
Asian (n = 14)	0.08	0.00 (ASN)	0.31	0.29 (ASN)
non-Hispanic Caucasian (n = 225)	0.18	0.24 (EUR)	0.36	0.41 (EUR)
Hispanic Caucasian (n = 23)	0.11	0.24 (EUR)	0.34	0.41 (EUR)
Hispanic (n = 18)	0.18	0.14 (AMR)	0.31	0.27 (AMR)
Other (n = 13)	0.13	0.12	0.31	0.38
Overall	0.17	0.12	0.35	0.38

DNHS: Duchenne Natural History Study. **1000G:** 1000 Genomes Project. Minor allele frequencies for 1000G refer to a continental reference population (in brackets), or to the whole project if not specified otherwise.

Genotyping results. Genomic DNA samples for targeted genotyping were available for 283/340 participants. The 57 patients excluded because of unavailability of DNA samples comprised all 41 patients followed at the study Center in Chennai, India, as regulatory authorities did not allow participation to this part of the study. Minor allele frequencies (MAFs) for *SPP1* rs28357094 and *LTBP4* rs10880 in the CINRG-DNHS population, broken down by ethnic subgroups, are shown in Table 2, compared to MAFs in continental reference populations from the 1,000 Genome Project (http://www.1000genomes.org/). For both SNPs, the MAF in the numerically preponderant Caucasian population was slightly lower than in the 1,000 Genomes

reference EUR population (0.18 vs. 0.24 for *SPP1* rs2835704, and 0.36 vs. 0.41 for *LTBP4* rs10880). For *SPP1* rs28357094, MAFs in Asian and Hispanic populations were higher than reference. These findings might be suggestive of population admixture. Both SNPs were in Hardy-Weinberg equilibrium (HWE). *LTBP4* haplotypes could be phased with 90% probability in 242 participants, including 28/32 "TT" homozygotes for LTBP4 rs10880. Of these, 24/28 were homozygotes for the full IAAM haplotype based on rs2303729, rs1131620, and 1051303 genotypes, while 4/28 participants were heterozygotes for other rare haplotypes (VAAM or VTTM). All SNPs in the *LTBP4* haplotype were in HWE.

Association analyses in the whole genotyped cohort: the effect of *SPP1* on LoA is GC treatment-dependent in the DNHS population. Analyses relative to SPP1 and LTBP4 genotypes were limited to 283 patients with available genomic DNA samples (Figure 1). Of these, 279 (because of limited availability of genomic DNA for a few participants) were successfully genotyped for *SPP1* rs28357094. Median ages at LoA for genotype groups and results of log-rank and Cox regression analyses are summarized in Table 3. Median ages at LoA were 11.8 years in 84 participants carrying the minor allele (TG/GG), and 13.0 years in 195 participants carrying the TT genotype (log-rank $p = 0.048$, Figure 2a). This closely reproduces the methodology of the previously reported association of rs28357094 genotype with LoA in 106 Italian patients [Pegoraro et al., 2011], representing an independent validation of association with this phenotype. In the Cox regression model with GC treatment as time-varying covariate, the hazard ratio (HR) ± standard error (SE) for TG/GG genotype was 1.22 ± 0.20 (p = n.s.). The HR for GC treatment was 0.41 ± 0.07 ($p < 0.001$). In 274 participants genotyped for *LTBP4* rs10880, median ages at LoA were 12.0 years in 242 participants with the CC/CT genotype, and 13.9 years in 32 homozygotes for the minor allele T (log-rank $p = 0.20$, Figure 3a). In the Cox regression model with GC treatment as time-varying covariate, HR for the TT genotype was 0.78 ± 0.18 (p = n.s.). The HR for GC treatment was 0.39 ± 0.06 ($p < 0.001$). In this and the following analyses, participant numbers for the two genotyped SNPs differ slightly, because of limited availability of genomic DNA for a few

participants. These data show directions of association as previously reported (*SPP1* TG/GG genotype: earlier age at LoA; *LTBP4* TT genotype: later age at LoA). *SPP1* was statistically significant in the log-rank comparison of median LoA, but not in the GC-treatment adjusted Cox model, while *LTBP4* did not reach statistical significance. Findings for *LTBP4* were similar for 24/32 rs10880 TT homozygotes carrying the whole IAAM/IAAM haplotype (data not shown). When grouping only GC-treated participants (at least 1 year of treatment while ambulatory), we observed a 1.9-year difference in median LoA between *SPP1* rs28357094 genotypes: median ages at LoA were 12.0 and 13.9 years for n = 63 GG/GT and n = 150 TT respectively (log-rank p = 0.032, Figure 2b). In the Cox regression model with GC treatment as time-varying covariate, the HR for TG/GG genotype was 1.61 ± 0.32 (p = 0.016). The HR for GC treatment was 1.30 ± 0.49 (p = n.s.). Median ages at LoA were identical (10.0 years) for untreated participants with different *SPP1* genotypes (n = 21 and 45 respectively, Figure 2b). This suggests that the *SPP1* locus may be a pharmacodynamic marker for GC response, rather than directly modifying DMD severity. As for the *LTBP4* rs10880 genotype, median ages at LoA in GC-treated participants were 13.3 and 13.9 years for the CC/CT and TT genotype, n = 178 and 27 respectively (log-rank p = n.s., Figure 3b). In the Cox regression model with GC treatment as time-varying covariate, the HR for the TT genotype was 0.74 ± 0.20 (p = n.s.). The HR for GC treatment was 1.08 ± 0.40 (p = n.s.). The number of untreated participants with the rare recessive genotype was too small for a meaningful comparison (n = 64 and 5, Figure 3b).

Table 3. Kaplan-Meier and Cox regression models for LoA: median age at LoA, log-rank p-values, Cox Hazard Ratios and p-values by *SPP1-LTBP4* genotypes and GC treatment.

Whole genotyped DNHS cohort (n = 283*)

		All participants				GC-treated				GC-untreated		
		n	**median age at LoA (yrs)**	**p-value (KM log-rank)**	**HR (95% CI) and p-value (Cox)**	**n**	**median age at LoA (yrs)**	**p-value (KM log-rank)**	**HR (95% CI) and p-value (Cox)**	**n**	**median age at LoA (yrs)**	**p-value (KM log-rank)**
SPP1 rs28357094	TT	195	**13.0**	**p = 0.048**	1.22 (0.89-1.68)	150	**13.9**	**p = 0.032**	**1.61 (1.09-2.37)**	45	10.0	0.6
	TG/GG	84	**11.8**		p = 0.22	63	**12.0**		**p = 0.016**	21	10.0	
LTBP4 rs10880	CC/CT	242	12.0	p = 0.20	0.78 (0.49-1.24)	178	13.3	p = 0.27	0.74 (0.44-1.26)	64	10	-
	TT	32	13.9		p - 0.29	27	13.9		p = 0.27	5	9.1	

Caucasian cohort, controlled for population stratification by MDS (n = 118*)

		All participants				GC-treated				GC-untreated		
		n	**median age at LoA (yrs)**	**p-value (KM log-rank)**	**HR (95% CI) and p-value (Cox)**	**n**	**median age at LoA (yrs)**	**p-value (KM log-rank)**	**HR (95% CI) and p-value (Cox)**	**n**	**median age at LoA (yrs)**	**p-value (KM log-rank)**
SPP1 rs28357094	TT	81	**13.9**	**p = 0.047**	1.54 (0.93-2.54)	67	**13.9**	p = 0.07	**1.85 (1.01-3.38)**	14	10.0	0.7
	TG/GG	35	**12.0**		p = 0.09	26	**12.0**		**0.047**	9	9.0	
LTBP4 rs10880	CC/CT	103	**12.6**	**p = 0.024**	0.49 (0.23-1.07)	80	**13.8**	**p = 0.046**	0.47 (0.20-1.09)	23	10	-
	TT	12	**15.0**		p = 0.07	11	**16.0**		p = 0.08	*1*	*N/A*	

DNHS: Duchenne Natural History Study. ***** Total n may not correspond exactly to genotype group n because of few ungenotyped patients (limited DNA availability). **MDS:** multidimensional scaling analysis. **GC-treated:** glucocorticoid corticosteroid treatment at least 1 year while ambulatory. **GC-untreated:** no or < 1-year treatment while ambulatory. **LoA:** loss of ambulation. **KM:** Kaplan-Meier survival analysis with log-rank comparison of median age at LoA. **HR:** Hazard Ratio for genotype in Cox regression model with GC-treatment as a time-varying covariate. **Bold figures** indicate statistically significant effect of genotype on LoA.

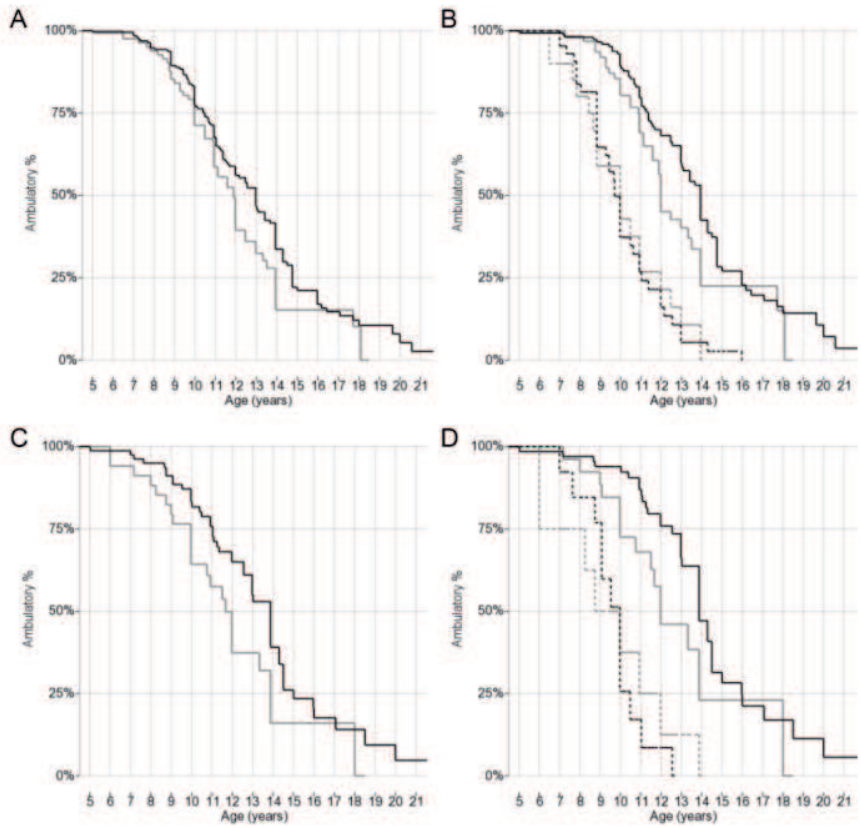

Figure 2. Kaplan-Meier plots of age at loss of ambulation grouped by *SPP1* rs28357094 genotype. A) All patients genotyped for *SPP1* rs28357094, including all races and ethnicities (n = 279), grouped 2 ways by rs28357094 genotype (TT black, TG/GG grey). B) All patients genotyped for *SPP1* rs28357094, including all races and ethnicities (n = 279), grouped 4 ways by rs28357094 genotype (TT: black, TG/GG: grey) and GC treatment (at least 1 year while ambulatory: continuous; less than 1 year or untreated: dashed). C) Caucasian cohort controlled for population stratification and genotyped for *SPP1* rs28357094 (n = 116), grouped 2 ways by rs28357094 genotype (TT: black, TG/GG: grey). D) Caucasian cohort controlled for population stratification and genotyped for *SPP1* rs28357094 (n = 116), grouped 4 ways by rs28357094 genotype (TT: black, TG/GG: grey) and GC treatment (at least 1 year while ambulatory: continuous; less than 1 year or untreated: dashed).

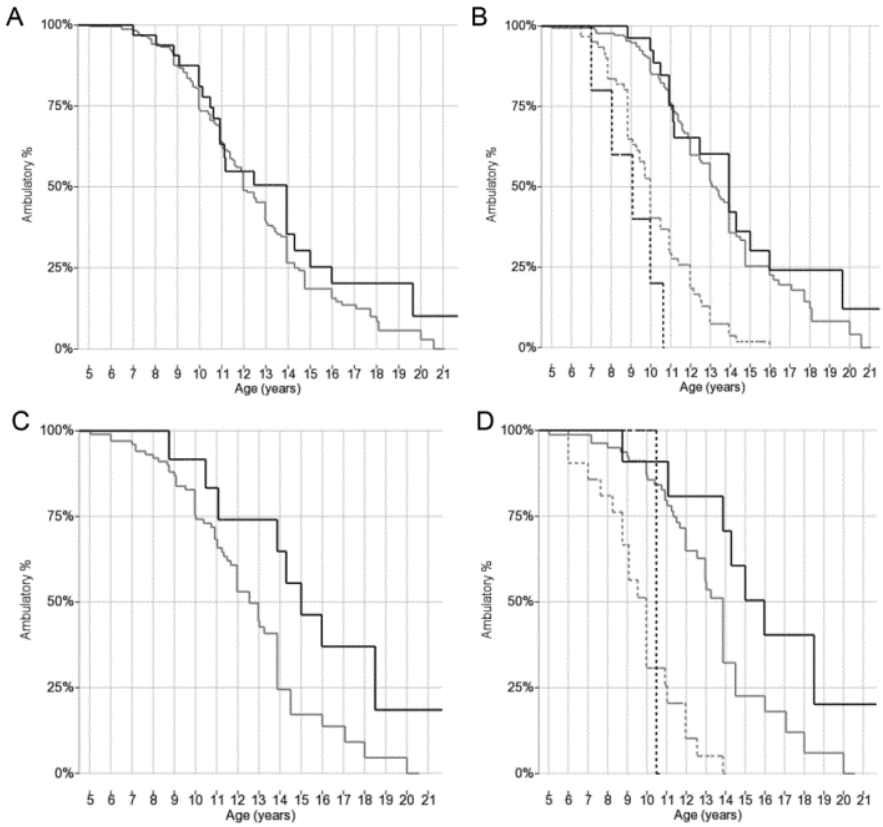

Figure 3. Kaplan-Meier plots of age at loss of ambulation grouped by *LTBP4* rs10880 genotype. A) All patients genotyped for *LTBP4* rs10880, including all races and ethnicities (n = 274), grouped 2 ways by rs10880 genotype (TT black, CC/CT grey). B) All patients genotyped for *LTBP4* rs10880, including all races and ethnicities (n = 274), grouped 4 ways by rs10880 genotype (TT: black, CC/CT: grey) and GC treatment (at least 1 year while ambulatory: continuous; less than 1 year or untreated: dashed). C) Caucasian cohort controlled for population stratification and genotyped for *LTBP4* rs10880 (n = 115), grouped 2 ways by rs10880 genotype (TT: black, CC/CT: grey). D) Caucasian cohort controlled for population stratification and genotyped for *LTBP4* rs10880 (n = 115), grouped 4 ways by rs10880 genotype (TT: black, CC/CT: grey) and GC treatment (at least 1 year while ambulatory: continuous; less than 1 year or untreated: dashed).

MDS analysis shows admixture and population stratification. MDS analysis was based on the calculation of identity-by-state pairwise distances, performed on 175 participants with available genome-wide markers. Compared to 108 patients excluded from this analysis because of unavailability of genome-wide markers (DNA quantity and quality not sufficient for SNP chip analysis), there were no significant differences in GC treatment rate (75.0% vs 76.5%) or median age at LoA (13.0 vs. 12.0 years, log-rank p = 0.12). MDS analysis identified a first principal component with lower values for participants of European ancestry. This component is plotted on the x axis in Figure 4. The y axis represents the second principal component. On the left side of the plot (low values of first principal component) 118 participants self-identifying mostly as non-Hispanic Caucasian (n = 115), and rarely as Hispanic-Caucasians (n = 2), or Other (n = 1), are clustered closely together, indicating a subcohort of relatively homogeneous European ancestry. Of the remaining 57 patients with higher values of the first component, 12 self-identified as non-Hispanic Caucasian, appearing as "outliers" on the right side of the plot, and indicating admixture and population stratification within self-identified Caucasian participants.

Association analyses in the Caucasian cohort controlled for population stratification lead to validation of both *SPP1* and *LTBP4* association with age of LoA. As MDS analysis showed population stratification within self-identified racial-ethnic groups, in order to adjust for population-related confounding factors in genetic association, we restricted subsequent analyses to 118 patients showing no evidence of population stratification (Figure 1). Of these, 116 were successfully genotyped for *SPP1* rs28357094. Median ages at LoA were 12.0 and 13.9 years, n = 35 and 81 for TG/GG and TT respectively (log rank p = 0.047, Figure 2c). In the Cox regression model with GC treatment as time-varying covariate, the HR for TG/GG genotype was 1.54 ± 0.17 (p = 0.09), and the HR for GC treatment was 0.26 ± 0.07 (p < 0.001).

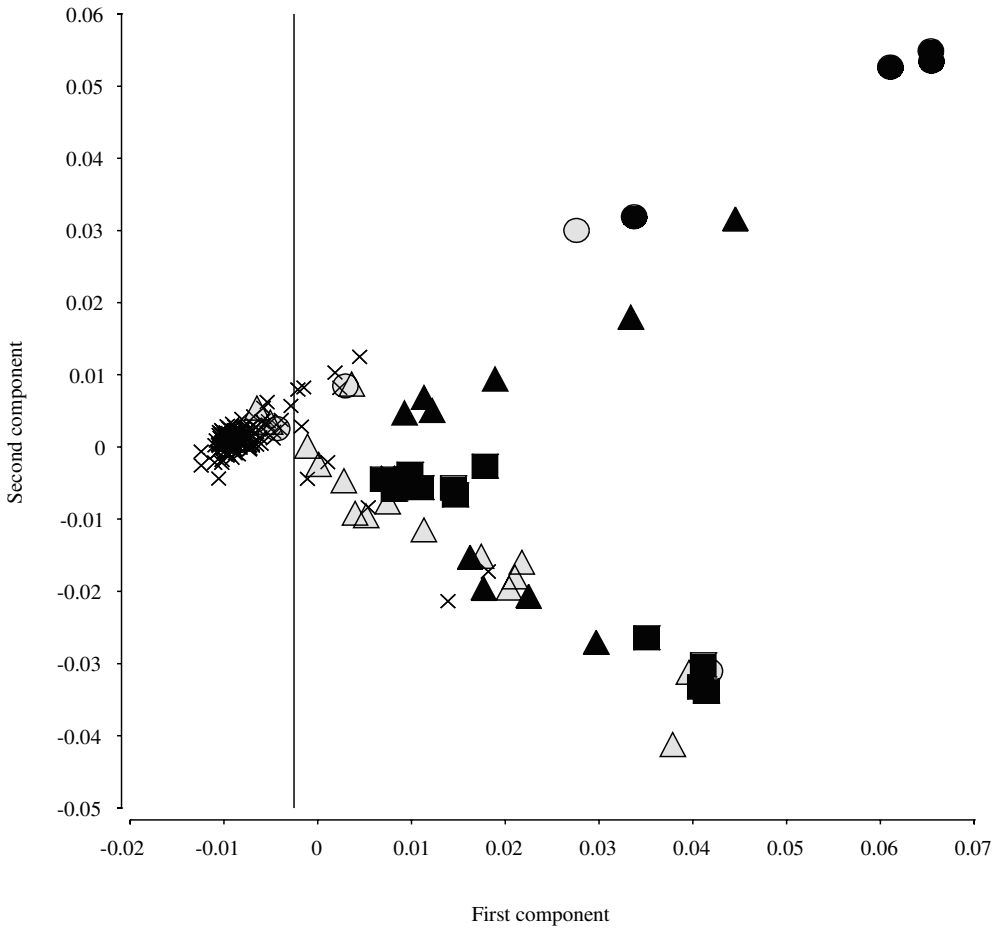

Figure 4. Cartesian plot of Multidimensional Scaling analysis of genome-wide markers population stratification. Values of the 2 highest ranking components are shown (1st on the x axis and 2nd on the y axis). Shape and color of the markers indicates self-identified ethnicity. Participants self-identifying as non-Hispanic Caucasian, indicated by "x" shaped markers, form a cluster with low values of the first component (< -0.0025, vertical cut-off line). Forty-five participants with other self-identified races and ethnicities are mostly positioned right of the cut-off line: African American (full circles), Asian (full squares), Hispanic Caucasian (empty triangles), Hispanic (full triangles), and Other (empty circles). Twelve participants self-identifying as non-Hispanic Caucasians appear as outliers, while 3 participants self-identifying as Hispanic Caucasian or Other cluster together with non-Hispanic Caucasians and are included in subsequent analyses.

Of the participants described above, 93 were GC-treated for at least 1 year while ambulatory. For these patients, KM median ages at LoA were 12.0 and 13.9 years for TG/GG and TT genotypes, n = 26 and 67 (log-rank p = 0.07, Figure 2d). When applying the Cox regression model with GC treatment as time-varying covariate, HR for TG/GG genotype was 1.85 ± 0.57 (p = 0.047). The HR for GC treatment was 0.72 ± 0.39 (p = n.s.). In 23 GC-untreated patients, KM median ages at LoA were 9.0 and 10.0 years for TG/GG and TT (n = 9 and 14 respectively, p = n.s., Figure 2d).

Of 118 participants in the Caucasian cohort controlled for population stratification, 115 were genotyped for *LTBP4* rs10880. KM curves plotted for this group showed delayed median LoA (15.0 years) in 12 participants carrying the TT genotype, in contrast to 103 carrying the CC/CT genotype (12.6 years, log-rank p = 0.024, Figure 3c). Of these 12 participants, 9 were homozygotes for the full IAAM haplotype, while 3 were heterozygotes for IAAM and other rare *LTBP4* haplotypes (VAAM and VTTM). In the Cox regression model with GC treatment as time-varying covariate, the HR for TT genotype was 0.49 ± 0.19 (p = 0.07). The HR for GC treatment was 0.26 ± 0.07 (p < 0.001).

Of the participants described in the previous paragraph, 91 were GC-treated for at least 1 year before LoA. Within this group, KM median age at LoA was 16.0 years for TT genotype and 13.8 for CC/CT, n = 11 and 80 (log-rank p-value = 0.046, Figure 3d). The Cox regression model with GC treatment as time-varying covariate showed a HR for TT genotype of 0.47 ± 0.20 (p = 0.08). The HR for GC treatment was 0.75 ± 0.40 (p = n.s.). The presence of just one GC-untreated participant with rs10880 TT genotype precludes statistical analysis of GC-untreated participants for *LTBP4* in this subgroup.

Taken together, these findings support the protective effect of the rs10880 TT genotype described by Flanigan et al.

Discussion

We aimed to study the effect of two genetic modifiers on DMD phenotype, using age at LoA as a disease severity indicator, in the CINRG-DNHS cohort. This cohort includes 340 participants followed longitudinally for an average of four years, and the study design and baseline data have been recently reported [McDonald et al., 2013; Henricson et al., 2013]. We grouped participants by genotype at two loci associated with age at LoA in DMD (*SPP1* rs283570944 and *LTBP4* haplotype). We had previously reported a cross-sectional analysis of grip strength as a function of *SPP1* genotype in a subset of this cohort (n = 156), not stratified for ethnicity. Here, we report time-to-event analyses for age at LoA in the complete CINRG-DNHS cohort (except 57 participants with unavailable DNA samples) for both the *SPP1* and *LTBP4* loci, controlling for population stratification and GC treatment as possible confounders.

It is well established that different ethnic groups show different MAFs for any specific genetic polymorphism, as well as different LD between genetic markers and functional variants. This can lead to hidden population stratification even within self-identified racial/ethnic groups, and thus to false positive or false negative findings in genetic association. The CINRG-DNHS cohort recruited participants from 20 clinical centers in four different continents, and is ethnically heterogeneous, although with a majority of Caucasian participants. Both previous reports [Fox et al., 1987; Kenneson et al., 2006; Holtzer et al., 2011], and observed tendencies of the phenotype to differ between ethnic groups in our data, further stress the importance of accounting for population stratification issues. On the other hand, because of inherent study design characteristics of the CINRG-DNHS, which did not purposely recruit representative racial/ethnic subgroups, conclusive statements cannot be made about racial/ethnic disparities in DMD, based on our data.

The second potential “confounding factor”, GC treatment, is probably the single environmental factor most heavily affecting age at LoA in DMD [DeSilva et al., 1987; Angelini et al., 1994b; McAdam et al., 2012; Ricotti et al., 2013a; Bello et al., 2015b]. We

accounted for it by implementation in time-to-event models, as a grouping criterion for KM analyses and a time-varying covariate for Cox regression. Because of CINRG-DNHS population characteristics, i.e. a numerical predominance of GC-treated participants, our findings can be generalized more confidently to GC-treated DMD populations. Nevertheless, this is more relevant for patient care and clinical trials, as GC treatment is considered a standard of care [Bushby et al., 2010a], and often an inclusion criterion in innovative clinical trials.

We first studied SNP effects in the entire cohort with available DNA samples, without grouping for GC treatment or ethnicity. The *SPP1* rs28357094 G allele, in a dominant model, was associated to 1.2-year earlier median LoA. This reproduces the methodology of the original report [Pegoraro et al., 2011] and represents an independent validation of association with LoA. The recessive *LTBP4* rs10880 T allele, in close LD with the IAAM haplotype, showed a direction of association as previously reported [Flanigan et al., 2013], but not a statistically significant difference.

We then grouped participants genotyped for *SPP1* and *LTBP4* by GC treatment. In our baseline analysis of 156 DNHS participants, the association of *SPP1* genotype with grip strength showed the largest effect in GC-treated participants [Pegoraro et al., 2011]; and an 80-patient Italian cohort, in which the association was established with longitudinal changes of ambulation-related functional measures [Bello et al., 2012], was almost entirely GC-treated. In line with these previous findings, which suggest a stronger modifier role of *SPP1* in GC-treated patients, an effect of *SPP1* genotype on LoA in the CINRG-DNHS cohort was observed in the GC-treated subgroup (Figure 2b). This finding supports a role of *SPP1* rs28357094 as a modulator of GC response in DMD, rather than of disease progression itself. This is in concordance with several preclinical studies of *SPP1* promoter function: the minor G allele at rs28357094 decreases transcriptional activity of the gene at baseline [Giacopelli et al., 2004], but shows a three-fold increase in gene expression in response to steroids, whereas the common allele leaves expression unchanged by steroids [Barfield et al., 2014]. Consistent with a

steroid-induced alteration of *SPP1* expression, differences in *SPP1* mRNA levels between genotypes were not found in vivo in DMD diagnostic muscle biopsies, obtained prior to GC treatment [Piva et al., 2012]. Evidence of a sexually dimorphic effect of rs28357094 genotype on muscle size and remodeling in Caucasian women [Hoffman et al., 2013], and of an increased transcriptional response of the *SPP1* promoter to estrogen stimuli [Barfield et al., 2014], is consistent with a pharmacodynamic role of this genetic biomarker. GCs are well-known transcriptional regulators of inflammation-related genes [Rhen and Cidlovski, 2015], both directly, through positive or negative GC-responsive elements (GREs), and indirectly, through suppression of other transcription factors (e.g. NF-κB); and these mechanisms are relevant to GC efficacy, and possibly side effects in DMD [Fisher et al., 2015; Heier et al., 2013]. The *SPP1* promoter is predicted to contain both GREs and NF-κB responsive sites, and further studies are needed to dissect these mechanisms both *in vitro* and *in vivo*. To our knowledge, this is the first demonstration of a pharmacodynamic biomarker for response to GCs, and this may be relevant to other conditions where GCs are standard of care. Our data, combined with the recent publication by Barfield et al., suggest that the 20-30% of DMD patients with the rare *SPP1* allele are "poor responders" to GCs, possibly with an altered balance between beneficial and side effects. Future studies are needed to confirm this association, before routine genotyping of *SPP1* is considered as part of standard of care in DMD.

Flanigan et al. described a larger effect of the protective *LTBP4* variant in GC-treated patients, than in untreated. In the CINRG-DNHS cohort, when stratifying by GC treatment and *LTBP4* genotype, we did not observe differences in genotype effect between the treated and untreated populations. Re-analyzing data with the same GC-treatment threshold (at least 6 months before LoA) as in the original report did not modify these findings (data not shown).

In order to adjust for potential population stratification bias, we performed MDS analysis on a subgroup of 175 participants. This showed that the correspondence between self-identified ethnicity and unbiased grouping determined by MDS was

partial, with several outlier self-identifying Caucasian participants (admixture). Subsequently, we selected MDS rather than self-identification as a method to adjust for population stratification. A limitation of this study is the unavailability of genome-wide markers in the whole cohort, which would have allowed the selection of a larger homogeneous subcohort. However, the reason for exclusion of 108 patients was technical (DNA sample quality and quantity) and not linked to any clinically relevant variables. Furthermore, we verified that excluded patients did not significantly differ in terms of GC treatment and age at LoA (data not shown). Thus, we expect the cohort analyzed by MDS to be representative of the whole CINRG-DNHS.

In the smaller, but more homogeneous MDS-selected Caucasian subcohort, *LTBP4* rs10880 was confirmed as a strong modifier of ambulatory function in DMD, with a median age at LoA in carriers of the protective genotype of 15.0 years overall, and 16.0 with GC treatment (Figure 3c-d), which positions > 50% of these patients within the "intermediate dystrophinopathy" clinical spectrum. The stronger association of rs10880 in Caucasians did not seem to be in direct relation with a stronger LD with the full IAAM haplotype, although numbers are too small to reach a conclusion (i.e. very few patients with rare haplotypes). It could also be hypothesized that in Caucasians the T allele for rs10880 might be in stronger LD with another unrecognized, functional variant, than in patients with different ancestries. On the other hand, Flanigan et al. did not find additional nonsynonymous coding variation by re-sequencing 40 chromosomes and querying the 1000 Genome database; and conducted in vitro experiments showing an effect of the coding haplotype on TGFβ signaling, in conditions of equal LTBP4 protein expression. As for *SPP1* rs28357094, the effect on age at LoA in the Caucasian subcohort appears relatively smaller in magnitude than that of the *LTBP4* haplotype. This is similar to what was described in the single Center cohort from Padova [Peograro et al., 2011], which was recruited from a homogeneous Caucasian population (great majority from north-eastern Italy) with a predominance of GC-treated participants.

Recently, another genetic modifier study in DMD was published by a collaborative European group [van den Bergen et al., 2015], further confirming the effect of the *LTBP4* IAAM haplotype in delaying LoA. On the other hand, the *SPP1* association was not replicated in 336 patients, of whom 102 had been treated with GCs for at least 1 year while ambulatory. If *SPP1* were indeed a modifier of GC response, as our association data and in vitro findings suggest, this low GC treatment rate might have limited the power of this part of the study.

Lastly, while the genetic modifiers described here seem to explain some of the variance of the LoA phenotype in DMD, both from our data and from case reports of outlier DMD phenotypes [Zatz et al., 2014] it appears that several other, yet uncharacterized genetic factors must be at play.

In conclusion, our findings show that *SPP1* rs28357094 acts as a modifier of the long-term effect of GC treatment in the CINRG-DNHS. Furthermore, we confirm that *LTBP4* rs10880 modifies age at LoA in DMD. Our data also stress the importance of adjusting for GC treatment and population substructure in genetic association studies in DMD. These findings are relevant for future analyses of observational and interventional studies involving international, multicentric, ethnically diverse cohorts.

Aim 2: DMD genotype-phenotype correlations in the CINRG-DNHS

Background

Duchenne muscular dystrophy (DMD) is caused by truncating mutations in the DMD gene, leading to the absence of dystrophin protein [Hoffman et al., 1987]. DMD presents with muscle weakness and wasting in early childhood, and progresses to loss of independent ambulation (LoA), which, according to a classic clinical description [Darras et al., 2000] ensues by the age of 13. The milder allelic form Becker muscular dystrophy (BMD), caused by non-truncating mutations [Hoffman et al., 1989], is characterized by LoA after the age of 16, while intermediate (IMD) forms show LoA between 13 and 16 years.

Improvements in standards of care [Bushby et al., 2010a-b] have delayed disease milestones in DMD [Henricson et al., 2013; Bello et al., 2015b], so that distinctions between "responder" DMD, IMD, and severe BMD have become blurred: currently, age at LoA in DMD may span from before 10 years up to 18 years in responders to glucocorticoids. As, by definition, DMD-causing mutations alter the open reading frame and abolish dystrophin expression, this wide variability has been partly explained with differences in care and treatment [Henricson et al., 2013; Ricotti et al., 2013a; Bello et al., 2015b], and trans-acting genetic modifiers [Pegoraro et al., 2011; Bello et al., 2012; Flanigan et al., 2013; van den Bergen et al., 2015; Bello et al., 2015a].

Nevertheless, mutations predicted to be out-of-frame at the genomic level may sometimes give rise to small amounts of dystrophin. For instance, mutations in the 5' region often escape the reading frame rule [Aartsma-Rus et al., 2006; Kesari et al., 2008], probably by means of translation reinitiation from downstream alternative promoters, as suggested for the relatively frequent out-of-frame deletion of exons 3-7 [Muntoni et al., 1994; Winnard et al., 1995; Gualandi et al., 2006]. Moreover, some

DMD exons are probably skipped physiologically at a low rate, explaining patients with relatively mild phenotypes who carry nonsense mutations situated within in-frame exons [Flanigan et al., 2011], and single- or multi-exon deletions bordering exon 44 [van den Bergen et al., 2014a; Pane et al., 2014b]. Despite "rescued" dystrophin expression being very low, often below the sensitivity threshold of standard diagnostic assays such as immunohistochemistry (IHC) or Western Blot (WB), it seems to be sufficient to substantially modulate disease progression, stabilizing function and delaying LoA.

Natural history studies can aid in deciphering these genotype-phenotype correlations, which are relevant for the design and interpretation of interventional studies. In particular, several novel DMD treatments are mutation-specific, which makes it necessary to distinguish specific mutation subgroups in order to correctly interpret natural history data.

A promising mutation-specific therapeutic approach is antisense oligonucleotide (AON) mediated "exon skipping". AONs aim at restoring the reading frame in patients with out-of-frame single- or multi-exon deletions. Proof-of-concept evidence of dystrophin restoration, and promising safety and efficacy data have been provided by phase I-II trials of AONs based on the 2'-O-Methyl-phosphorothioate [van Deutekom et al., 2007; Goemans et al., 2011; Flanigan et al., 2014; Voit et al., 2014] and morpholino [Kinali et al., 2009; Cirak et al., 2011; Mendell et al., 2013] chemistries. Currently, AONs targeting exons 43, 44, 51, and 53 are being evaluated in clinical trials (NCT02310906, NCT00159250, NCT02255552, NCT02500381 NCT01957059, NCT02329769). Despite initial and continuing successes, a phase III trial of a 2'-O-Methyl-phosphorothioate exon 51 skipping compound (Drisapersen, NCT01254019) failed to achieve a significant functional benefit. This raised questions regarding dystrophin restoration levels [Lu et al., 2014; Hoffman and McNally, 2014], as well as trial design and selection of clinical endpoints. The clinical trials for the exon 51 skipping compound based on the morpholino chemistry (Eteplirsen) are still under clinical evaluation and continue to be promising. Very recently, data supporting its efficacy was provided based on comparison

with an Italian multicenter natural history cohort, stratified by DMD mutation [Mendell et al., 2015].

A different mutation-specific therapeutic approach is represented by small molecules promoting ribosomal readthrough of premature termination codons caused by nonsense mutations. In this field, Translarna® has shown promising results on clinical endpoints in a multicenter phase IIb trial [Buschby et al., 2014], and was provisionally approved in 2014 by the European Medicine Agency (EMA) for the treatment of nonsense-mediated DMD [Haas et al., 2015]. Results of a confirmatory phase III trial (NCT01826487) have very recently been made public, further suggesting some clinical efficacy (http://ir.ptcbio.com/releasedetail.cfm?ReleaseID=936905).

In this context, the aim of the present study was to analyze correlations between different truncating DMD gene mutations and loss of ambulation (LoA) in the Cooperative International Neuromuscular Research Group Duchenne Natural History Study (CINRG-DNHS), a large prospective DMD natural history study [McDonald et al., 2013], with particular attention to mutation groups amenable to novel molecular treatments.

Methods

CINRG-DNHS inclusion criteria. Participants in the "parent" CINRG-DNHS (distinguished from a recruiting extension, NCT00468832) were recruited between 2006 and 2009, and inclusion criteria have been described [McDonald et al., 2013]. Importantly, patients were excluded if they had a proximal (5' of exon 25) out-of-frame mutation and a BMD phenotype (because of frequent violations to the reading frame rule in this region); and could be included on the grounds of a typical DMD phenotype, even if they had an in-frame *DMD* mutation, or no demonstrated *DMD* mutation but abnormal dystrophin IHC or WB. Average follow-up was 4 years.

Additional inclusion criteria. In order to group participants by *DMD* mutation type and amenability to molecular therapies, we further selected participants with available evidence of a *DMD* genetic mutation and the following characteristics: single- or multi-exon out-of-frame deletion with univocally defined exon boundaries (MLPA or multiplex PCR with amplification of immediately flanking exons); single- or multi-exon out-of-frame duplication confirmed by MLPA; small out-of-frame *DMD* mutation (insertion, deletion); nonsense mutation; splicing mutation.

DMD mutation studies. Diagnostic genetic testing for causative *DMD* mutations was performed at local institutions as part of the standard diagnostic work-up, and reviewed by a central CINRG genetic counselor, who reviewed rearrangement exon boundaries and unified mutation nomenclature (following HGVS recommendations) and reference sequences (genomic NG_012232.1, transcript NM_004006.2).

Grouping of DMD mutations. The grouping rationale was to describe the natural history of LoA in groups of participants with typical vs. atypical phenotypes, and/or amenability to mutation-specific molecular therapies. We defined the following categories: out-of-frame deletions amenable to skipping of (1) exon 44, (2) exon 45, (3) exon 51 (4) exon 53, (5) deletion of exons 3-7; (6) out-of-frame deletions not amenable to skipping of

exons 44, 45, 51, or 53; (7) out-of-frame duplications; (8) nonsense mutations; and (9) out-of-frame small insertions or deletions.

Genetic modifier genotyping. The parent CINRG-DNHS cohort showed significant effects of two known genetic modifiers of ambulation in DMD [Bello et al., 2015a]: the SNPs rs28357094 in the *SPP1* promoter [Pegoraro et al., 2011] and rs10880 in the coding portion of the *LTBP4* gene [Flanigan et al., 2013]. Genotyping and grouping by inheritance models was as described previously [Bello et al., 2015a].

Loss of ambulation. We defined LoA as patient- and/or caregiver-reported age at continuous wheelchair use, approximated to the nearest month, and verified by inability to walk 10 meters via the site CINRG Clinical Evaluator testing when possible.

Statistical analysis. We performed a time-to-event analysis of LoA with age (years) as time variable and LoA as event. Median age at LoA and corresponding 95% confidence intervals (CIs) were estimated by plotting empirical Kaplan-Meier curves for participant groups defined by mutation (as described above), and by GC treatment administered while ambulatory, with the following grouping: untreated (or short treatment of < 1 year); prednisone or prednisolone; deflazacort. Participants switching between prednisone/prednisolone and deflazacort were grouped accordingly to the drug administered for the longest time. We used Cox proportional hazard models to estimate and compare age-related risks of LoA. Covariates included *DMD* mutations as described above, and time-varying GC drug (prednisone/prednisolone or deflazacort) and weight-adjusted dose as described elsewhere [Bello et al., 2015b]. Note that GC treatment as a time-varying covariate in the Cox proportional hazard analysis is independent of the grouping of individual patients, used for the empirical Kaplan-Meier estimation of median age at LoA. The same analyses were repeated adding covariates for *SPP1* and *LTBP4* genotype; this was performed separately, because DNA for genotyping of both SNPs was not available in 41 participants largely due to regulatory issues in some countries which did not allow DNA shipping. Statistical significance was set at $p < 0.05$. All analyses were performed with the “survival” package in R, version 3.2.1.

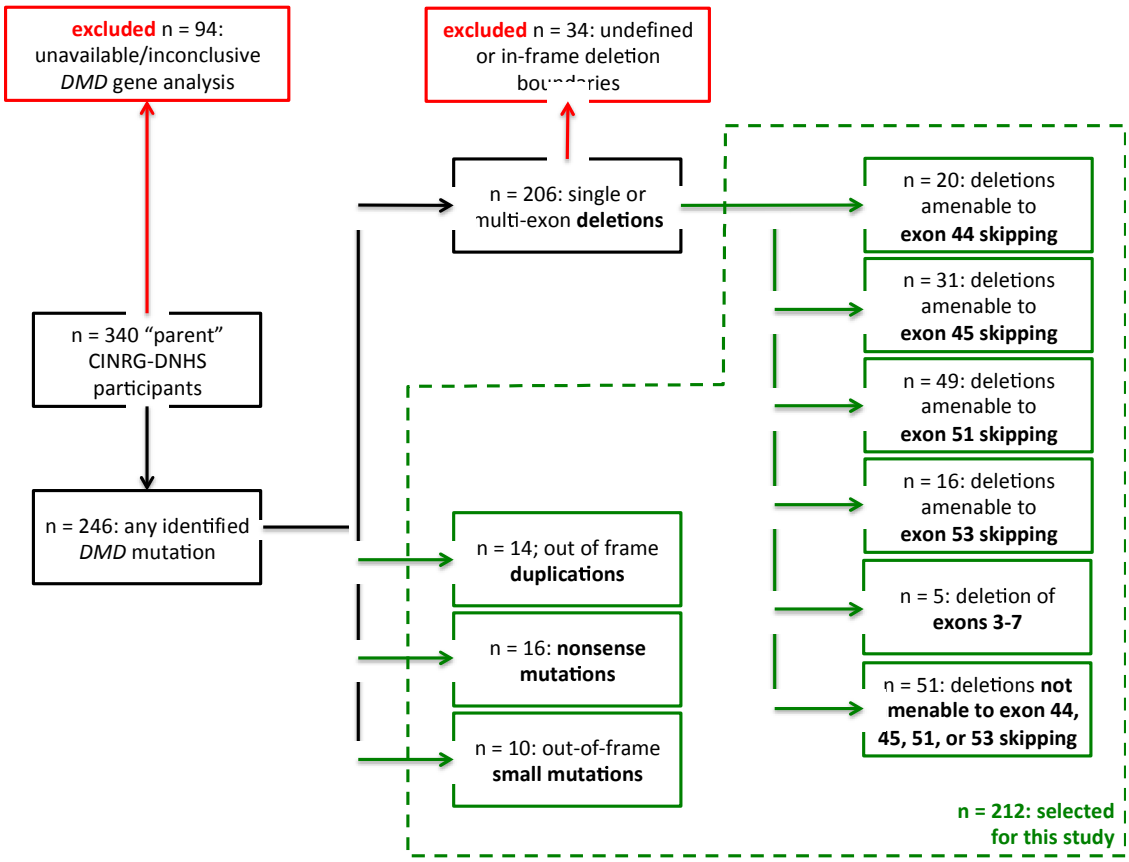

Figure 1: flow chart for participant selection from the "parent" (i.e. not including a currently recruiting extension) CINRG Duchenne Natural History Study (DNHS) for the analyses presented in this paper.

Results

Selected cohort. Inclusion criteria described in Methods led to the selection of 212/340 (62.3%) participants to the "parent" CINRG-DNHS (Figure 1). Their distribution by mutation and exon skipping amenability is described in Table I. Of note, there were 49/212 participants (23.1%) with deletions amenable to skipping of exon 51, 31/212 (14.6%) of exon 45, 20/212 (9.4%) of exon 44, 16/212 (7.5%) of exon 53, and 16/212 (7.5%) with nonsense mutations.

Age at LoA by mutation group. Median age at LoA with 95 % CIs by mutation group is shown in Table I, and Kaplan-Meier plots of LoA by age and mutation group are shown in Figure 2. The largest group (n = 51), comprising participants with out-of-frame deletions not amenable to skipping of exons 44, 45, 51, and 53, showed a median age at LoA of 12.7 years (95% Confidence Interval [CI] 11 - 14 years). The groups with later LoA included deletion of exons 3-7 (n = 5, median age at LoA 15.2 years, 95% CI 9 - undetermined), and deletions amenable to exon 44 skipping (n = 20, median 14.8 years, 95% CI 12 - undetermined).

Cox regression analysis of LoA. Cox regression analysis including GC treatment covariates showed that the delay of LoA was statistically significant in both the group with deletion of exons 3-7 (Hazard Ratio [HR] 0.24, 95% CI 0.07 - 0.82, p = 0.02) and the group with deletions amenable to skipping of exon 44 (HR 0.34, 95% CI 0.15 - 0.74, p = 0.007). Cox regression parameters for all other *DMD* mutations did not differ significantly from the reference group (Table I). As previously described in the whole CINRG-DNHS population[7], treatment with both prednisone/prednisolone or deflazacort were strongly associated with later LoA, with a lower HR (i.e. later LoA) for deflazacort (HR 0.34 and 0.22 respectively, p < 0.0001 for both), while there was no significant independent effect of GC dose in this population. In 171/212 participants with available modifier genotypes, amenability to skipping of exon 44 remained significantly associated with later LoA (HR 0.24, 95% CI 0.10 - 0.61, p = 0.006), and so was deletion of exons 3-7 (HR 0.21, 95% CI 0.05 - 0.92, p = 0.01), after adjusting for *SPP1* and *LTBP4*

genotypes as additional covariates. Modifier genotype effects were in the same direction as previously described, but not statistically significant in this multivariate model in this population (data not shown).

Table I: Participant distributions, median age at LoA, and Cox regression parameters for the time-to-event analysis of LoA

Cox regression factor	Level of factor	n	%	Median age (years) at LoA (95% CI)	HR (95% CI)	p-value
***DMD* mutation**	Exon 44 skipping amenable del	20	9%	14.8 (12 - ∞)	0.34 (0.15 - 0.74)	***0.007***
	Exon 45 skipping amenable del	31	15%	12.0 (11.0 - 13.5)	1.16 (0.66 - 2.06)	n.s.
	Exon 51 skipping amenable del	49	23%	11.6 (10.4 - 12)	0.84 (0.51 - 1.40)	n.s.
	Exon 53 skipping amenable del	16	8%	14.0 (9.0 - 15.0)	0.79 (0.39 - 1.57)	n.s.
	Deletion of exons 3-7	5	2%	15.2 (9.0 - ∞)	0.24 (0.07 - 0.82)	***0.02***
	Other out-of-frame deletion	51	24%	12.7 (11.0 - 14.0)	1 *	-
	Out-of-frame duplication	14	7%	12.7 (8.0 - ∞)	1.10 (0.50 - 2.41)	0.65
	Nonsense mutation	16	8%	11.1 (10.0 - 18.1)	0.64 (0.32 - 1.27)	n.s.
	Other frameshifting small mut	10	5%	14.0 (9.7 - ∞)	0.76 (0.29 - 1.97)	n.s.
GC drug‡	Untreated (or treated < 1 year)	55	26%	9.7 (9.0 - 11.0)	1 *	-
	Prednisone or prednisolone[§]	63	30%	12.0 (11.3 - 14.0)	0.34 (0.20 - 0.57)	***<0.0001***
	Deflazacort[¶]	94	44%	14.0 (13.7 - 15.0)	0.22 (0.12 - 0.40)	***<0.0001***
GC dose		-		-	1.12 (0.77 - 1.63)	n.s.
Total		212	100%	12.0 (11.5 - 13.0)	-	-

LoA: loss of independent ambulation. **CI**: confidence interval. **HR**: hazard ratio. **del**: deletion. **mut**: mutation. ***DMD***: dystrophin gene. **∞**: upper CI undetermined because of low numerosity. *****A HR of 1 is given for factor levels which are taken as reference in the Cox regression model. **GC**: glucocorticoid corticosteroid. **‡** For the calculation of median LoA, participants are grouped in treated less than 1 year, treated at least 1 year and > 50% of the time with prednisone or prednisolone, and treated at least 1 year and > 50% of the time with deflazacort; while HRs are calculated for time-varying covariates, independent of the grouping of individual participants. **§** Including participant switching between drugs who were > 50% of the time on prednisone or prednisolone while

ambulatory. ¶ Including participant switching between drugs who were > 50% of the time on deflazacort while ambulatory.

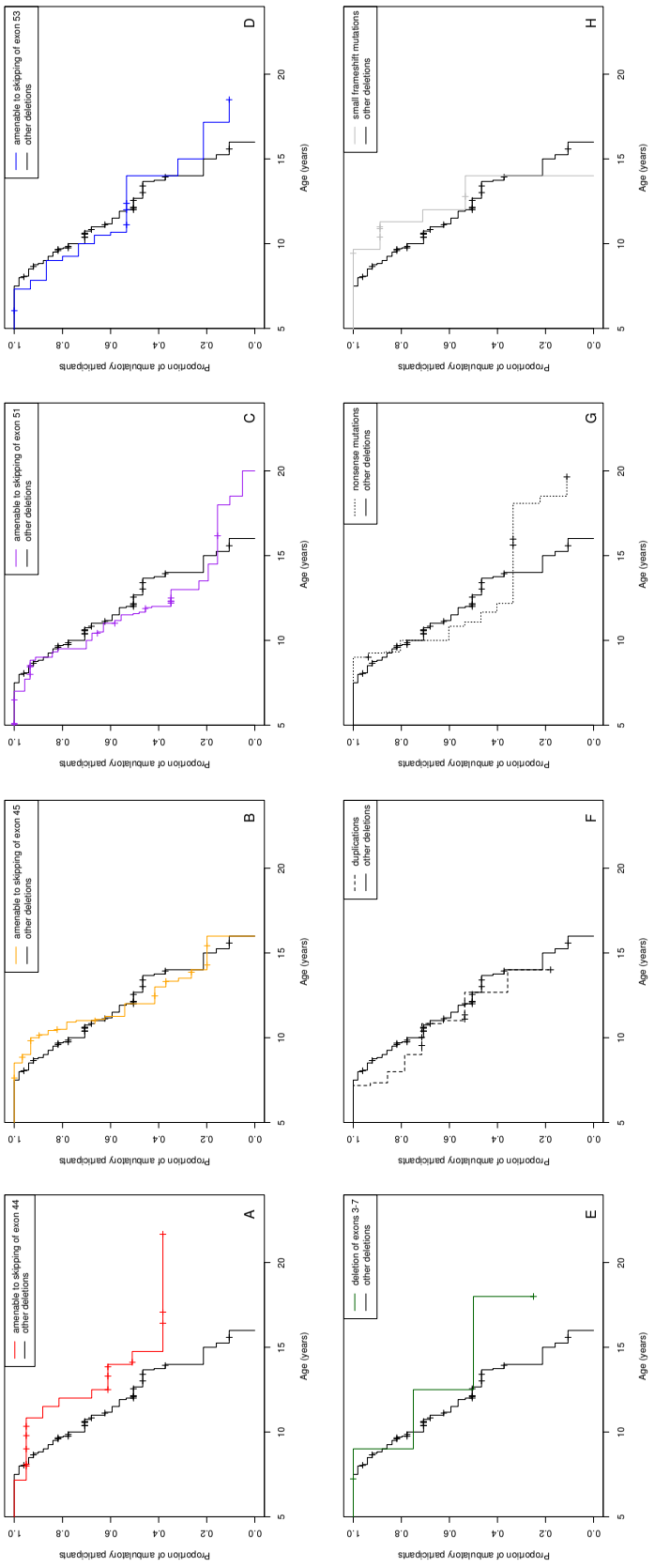
A
amenable to skipping of exon 44
other deletions
B
amenable to skipping of exon 45
other deletions
C
amenable to skipping of exon 51
other deletions
D
amenable to skipping of exon 53
other deletions
E
deletion of exons 3-7
other deletions
F
duplications
other deletions
G
nonsense mutations
other deletions
H
small frameshift mutations
other deletions
Proportion of ambulatory participants
Age (years)

Figure 2: Kaplan-Meier plots of loss of ambulation by mutation group. Plots for patients with (A) deletions amenable to skipping of exon 44, (B) deletions amenable to skipping of exon 45, (C) deletions amenable to skipping of exon 51, (D) deletions amenable to skipping of exon 53, (E) deletion of exons 3-7, (F) single- or multi-exon duplications, (G) nonsense mutation, and (H) small frameshift mutations are compared in each graph with the reference group of patients with single-or multi-exon deletions not amenable to skipping of exons 44, 45, 51, or 53 ("other deletions").

Discussion

In the present study, we describe significant differences in age at LoA between DMD patients with different *DMD* mutations. LoA is a major natural history milestone in DMD, and a reliable overall indicator of the severity of disease progression, with a strong correlation with the longitudinal changes of functional measures commonly adopted in clinical trials [Mazzone et al., 2013; Pane et al., 2014a].

We observed an approximately 2-year delay of median LoA in 20 participants who had mutations amenable to exon 44 skipping. Similar results were observed in a retrospective genotype-phenotype association study in Dutch subjects [van den Bergen et al., 2015]. Most DMD patients with exon 44 skipping eligibility carried the relatively frequent single-exon deletion of exon 45: 60% in our cohort and 66% in the cohort reported by van den Bergen et al. This mutation has been long known to induce endogenous skipping of the adjacent exon 44, resulting in traces of dystrophin expression [Prior et al., 1997]. Traces of dystrophin were reported in 3/6 DNHS participants with exon 44 skipping eligibility by IHC, and 0/4 by WB. Although these are diagnostic protein studies collected from retrospective laboratory reports, and therefore not directly comparable to dystrophin quantitation techniques recommended in clinical trials [Anthony et al., 2014b], it can be inferred from this observation that limited amounts of dystrophin, only detectable by a more sensitive technique such as IHC, are still able to induce a relevant delay of disease milestones. This is in keeping with observations in the dystrophin/utrophin double knockout mouse, an animal model of severe dystrophinopathy, in which re-expression of small amounts of dystrophin improves muscle pathology and function [van Putten et al., 2013].

Recently, a Japanese research group identified a splicing silencer effect of the intronic junction sequence from a DMD patient with exon 45 deletion, which could promote skipping of exon 44 [Dwianingsih et al., 2014]. Interestingly, we observed an intermediate DMD/BMD phenotype also in two patients with out-of-frame multi-exon deletions proximal to exon 44 (del 10-43 and 38-43), and a different deletion breakpoint in intron 43 instead of 44. These two patients were still ambulatory at the ages of 21

and 16 years with prednisone treatment, suggesting a similar splicing silencer mechanism for the intron 43 breakpoint, or different molecular mechanisms. One of them had traces of dystrophin identified by IHC, but not WB.

A distinct, albeit small group of participants with a milder phenotype in the CINRG-DNHS is represented by 5 DMD patients carrying the deletion of exons 3-7, previously described as an exception to the reading frame rule [Muntoni et al., 1994; Winnard et al., 1995; Gualandi et al., 2006]. IHC showed traces if dystrophin in 1/3 participants with this mutation, and WB in 0/3 (the three participants who underwent IHC and the three who underwent WB are not the same). However, the one patient with reported dystrophin traces lost ambulation at the age of 9 years, despite high-dose weekend prednisone since the age of 4. A 5' internally deleted dystrophin protein might be less efficient in rescuing the phenotype, because of the disruption of the functionally relevant N-terminal actin-binding domain, as well as other genetic and environmental confounders, may influence the clinical outcome. Larger case series are needed to fully understand the phenotypes associated with the deletion of exons 3-7. Notably, some IMD cases with this deletion might have been excluded due to the CINRG-DNHS criteria (see Methods).

These findings have several potential repercussions for clinical trials. First, some patients eligible for exon 44 skipping might show stabilized function even in the placebo group, especially in short studies, making small drug effects challenging to identify. Second, among DMD patients with the same exon skipping eligibility there might be some who activate endogenoues exon skipping mechanisms, and some who do not. Therefore, accurate quantitation of baseline dystrophin for each patient participating in dystrophin-restoring trials is of paramount importance.

Our data did not confirm reports that DMD patients with deletions amenable to exon 53 skipping might present earlier LoA [Servais et al., 2015], as these patients appeared to be positioned in the "typical" range for age at LoA in DMD. In order to reproduce the methodology of the cited study by Servais and colleagues, we calculated mean age at LoA in non-ambulatory patients only, which was 11.4 years in 10 non-ambulatory

CINRG-DNHS participants amenable to exon 53 skipping, vs. 8.9 years in 13 patients in the cited study. A higher rate of GC treatment in our subjects likely plays a relevant role in this difference. Further studies on functional measures may confirm the tendency to reduced upper limb function described by the same authors in this mutation group.

Ten participants carried a single-exon deletion of exon 52, which could be theoretically amenable to both exon 51 and 53 skipping. Here, we grouped these participants together with deletions amenable to exon 51 skipping, a therapeutic approach which has reached more advanced stages of clinical trials. Interestingly, these exon 52 deleted participants had early LoA, median age being 10.0 versus 11.9 years in all other exon 51 skipping eligible participants (log-rank p = 0.016). This genotype-phenotype association needs independent validation, but if confirmed, it could be relevant for the interpretation of the results of AON clinical trials targeting exon 51 and 53.

A different group of participants amenable to molecular treatment were 16 participants carrying nonsense mutations. This group is important as ambulatory DMD patients above the age of 5 are currently being prescribed Translarna® in several European countries, under the provisional approval of the EMA. While the median age at LoA in this group was similar to the "reference" population carrying non-exon-skipping-eligible mutations, there were 5 "outlier" cases showing prolonged ambulation (close to or beyond the age of 16), consistent with IMD, as shown by a rightward shift of the third and fourth quartiles of the Kaplan Meier plot (Figure 2, panel G). *DMD* nonsense mutations have already been described in association with IMD/BMD phenotypes. Furthermore, it has been shown that exons where IMD/BMD nonsense mutations occur are usually in-frame, situated in the functionally dispensable rod domain, and defined by weaker splice signals, explaining easier induction of endogenous exon skipping [Flanigan et al., 2011]. In fact, of five "IMD" participants with nonsense mutations in the CINRG-DNHS, three carried stop codons within in-frame exons (14, 29, and 30), but two carried stop codons in out-of-frame exons: exon 45, again suggesting alternative splicing in this region, and exon 69, suggesting escape from mRNA nonsense-mediated decay in a distally located mutation, giving rise to a C-terminally truncated protein. Interestingly,

another participant who lost ambulation early (10 years of age) carried a proximal nonsense mutation (c.9 G>T, p.Trp3*) with a described founder effect in North America. This is in contrast to the previously described association of this mutation to a mild BMD phenotype [Flanigan et al., 2009], probably caused by dystrophin rescue by downstream translation reinitiation [Wein et al., 2014]. Patients with a definitely mild BMD phenotype would not have been included in the DNHS because of clinical exclusion criteria; nevertheless, it could be relevant to stratify nonsense mutation DMD patients by stop codon position (in-frame or out-of-frame exon) in clinical trials. Currently, Translarna® is only approved by the EMA for nonsense-mediated DMD in Europe, but cases such as those described here show that the distinction with nonsense-mediated IMD/BMD might be blurry and hard to define.

A limitation of this natural history study is the use of DNA mutation data derived from clinical records. Although all CINRG clinical sites actively pursue adherence to the modern standards of care in DMD diagnosis, the identification of the causative mutation (especially sub-exonic small mutations), was not possible in some cases. As single- or multi-exon deletions are easier to identify, the proportions of DMD patients amenable to skipping of individual exons might be slightly inflated in this report. Future studies including full *DMD* sequencing at the genomic and/or RNA level, as well as next generation sequencing (NGS) studies of intronic deletion/duplication breakpoints in selected cases, might further refine genotype/phenotype correlations.

In conclusion, we provide mutation-specific natural history data regarding LoA in DMD, carefully adjusting for the effect of other disease-modifying variables, which is relevant for the design and interpretation of clinical trials for innovative therapeutics in DMD. Importantly, DMD patients with deletions amenable to exon 44 skipping, exon 3-7 deletion, and point mutations within in-frame exons should be excluded from natural history-derived placebo groups for the evaluation of AONs targeting rare deletions, and their distribution in treated/placebo groups should be carefully balanced in clinical trials of non-dystrophin-restoring treatments.

Aim 3: Long-term outcomes of glucocorticoid regimens in the CINRG-DNHS

Background

Duchenne muscular dystrophy (DMD) is caused by *DMD* gene mutations leading to the absence of dystrophin in skeletal muscle [Hoffman et al., 1987]. While dystrophin-restoring treatments are being developed, the only recommended [Moxley et al., 2005; Bushby et al., 2010a] pharmacological intervention is glucocorticoid corticosteroid (GC) treatment [Drachman et al., 1974; Brooke et al., 1987; Mendell et al., 1989; Fenichel et al., 1991; Griggs et al., 1991; Angelini et al., 1994b; Biggar et al., 2001; Manzur et al., 2008] with prednisone or its active metabolite prednisolone (PRED), or deflazacort (DFZ). Mechanisms of action include anti-inflammation/immunosuppression [Kissel et al., 1991], membrane stabilization [Jacobs et al., 1996], enhanced regeneration [Sklar and Brown, 1991; Anderson et al., 2000; Dadgar et al., 2014], and gene expression modulation [Fisher et al., 2005]. Side effects are common but usually manageable [Bushby et al., 2010a-b].

Long-term efficacy of GCs in delaying loss of independent ambulation (LoA) and other "disease milestones", although well described [Angelini et al., 1994b; De Silva et al., 1987; McAdam et al., 2012; Kim et al., 2015], is supported by lower-class evidence than short-term effects on muscle strength [Manzur et al., 2008]. Baseline data from the Cooperative International Neuromuscular Research Group Duchenne Natural History Study (CINRG-DNHS) [Henricson et al., 2013] showed that participants were more often ambulatory at age ≥ 10 years, if currently treated with GCs. Here, we expand to a longitudinal time-to-event analysis of GC regimen effects on LoA.

Prescribed GC regimens are manifold in DMD [Griggs et al., 2013], few studies directly comparing PRED vs. DFZ [Reitter, 1995; Bonifati et al., 2000]. There is evidence of better efficacy of daily GCs [Manzur et al., 2008], but several alternative regimens are applied

(e.g. weekend [Escolar et al., 2011], 10-days-on/10-days-off [Ricotti et al., 2013a]). A global, randomized, blinded trial of daily prednisone, daily DFZ, and 10-days-on/10-days-off prednisone is underway (www.for-dmd.org). In parallel, novel "dissociative steroids" aim to a broader therapeutic window, by separating pharmacodynamic mechanisms responsible for efficacy and side effects [Hoffman et al., 2012; Heier et al., 2013]. Before randomized trial results and innovative treatments become available, natural history studies can provide useful information regarding different GC regimens in DMD.

Methods

Informed consent. The institutional review board or ethics review board at each participating institution approved the study protocol, consent and assent documents. Informed consent/assent was obtained for each participant prior to conducting study procedures.

Study population. We present data from 340 DMD patients, aged 2-28 years, enrolled in the parent CINRG-DNHS (distinguished from a currently recruiting extension, http://clinicaltrials.gov/show/NCT00468832). Inclusion criteria have been described [McDonald et al., 2013].

GC treatment. At baseline and follow-up visits, we recorded time of beginning/discontinuation, drug, dose, and pattern of administration of previous and current GC regimens.

GC dose. Dose data was converted to % ratios of recommended doses for PRED (0.75 mg/kg/day) and DFZ (0.9 mg/kg/day).

Definition of LoA. Age at LoA was defined by patient-reported continuous wheelchair use, confirmed by inability to walk 10 meters unaided.

Grouping by GC treatment relative to LoA. GC regimens < 1 month were ignored. For comparisons of median LoA between GC-treated and untreated participants, we considered "GC-treated" only those patients who had been administered GCs for ≥ 1 year, starting ≥ 1 year before LoA; the rationale being that a long-term effect cannot be attained with a short-term treatment.

Grouping by GC regimen for Kaplan-Meier analyses. Because of a low number of participants subject to intermittent regimens (10-days-on/10-days-off, 10 days/month, every other day [QOD], 5 days/week) we grouped these regimens together. We

analyzed the high-dose (10 mg/kg/week) 2 days/week ("weekend") regimen separately, because of pharmacologically different properties of this treatment.

Cox regression analyses of PRED and DFZ use, regimen and dose. As many participants changed drugs, regimen, and doses during treatment, all these variables were evaluated for concurrent effects as time-varying covariates in a Cox regression model, independent of grouping of individual participants (see also statistical analysis paragraph below).

Overlap with CINRG clinical trials. Twenty-nine participants were rolled over into the DNHS from a CINRG clinical trial of daily vs. weekend prednisone [Escolar et al., 2011].

Genetic modifiers. *SPP1* rs28357094 and *LTBP4* rs10880 were genotyped as described [Bello et al, 2015a]

Side effects. We report frequency of physician-reported side effects in participants treated with GCs while ambulatory.

Statistical analysis. Average GC dose was compared between drug-regimen subgroups by Mann-Whitney U test, while cumulative dose and age at start of treatment were compared by Student's t-test. LoA was studied as event in a time-to-event model, with age as time variable, and censoring of ambulatory participants at the age of last follow-up. Median ages at LoA, calculated from empiric Kaplan-Meier curves, were compared by log-rank test. A Cox regression model was devised with the following time-varying covariates: GC drug (untreated, PRED, or DFZ), GC regimen (untreated, daily, low-dose intermittent, or weekend), and mg/kg/day dose, adjusting for random effects depending on CINRG study Site. Hazard Ratios (HR) were calculated for each covariate, with untreated as reference (HR = 1) for categorical covariates. A linear test compared covariate levels within the Cox regression model. Statistical significance was set at $p < 0.05$. Frequency of adverse effects between regimens was compared by χ^2 test. STATA V13 and Partek GS 6.6 were used for analyses.

Results

Age, follow-up and ambulatory status. At last follow-up (data updated through December 2013), average age was 15.7 ± 5.6 years (range 4.5 – 33.1), and average follow-up 3.8 ± 1.8 years; 111 participants were ambulatory (32.6%), while 229 (67.4%) had lost ambulation. Average age at last follow-up was 11.2 ± 3.1 years in ambulatory participants, and 18.0 ± 5.3 years in non-ambulatory.

Distribution by GC treatment while ambulatory. Sixty-three participants (18.5%) were untreated while ambulatory (including one patient treated with a non-GC anabolic steroid). At last follow-up, 54 of these were non-ambulatory, and 9 ambulatory and GC-naïve. Conversely, 277 participants (81.4%) were treated with GCs while ambulatory. Of these, 175 were non-ambulatory, and 98 still ambulatory at last follow-up. A ≥1 year GC treatment was administered while ambulatory to 252 participants (74.1% total, 91.0% treated). Average ± standard deviation (SD) duration of treatment while ambulatory was 4.0 ± 3.3 years, ranging 0.1 – 18.3 years.

GC treatment and age at baseline. Average age at baseline was higher in patients treated <1 year or untreated while ambulatory, vs. treated ≥1 year while ambulatory (15.1 ± 6.4 vs. 10.9 ± 5.2 years, $p < 0.001$), reflecting increased implementation of GC treatment as a standard of care in younger participants.

GC treatment and LoA. Kaplan-Meier analysis showed that median LoA was 3 years later in participants treated ≥1 year while ambulatory vs. untreated or treated < 1 year (13.0 vs. 10.0 years, n = 252 vs. 88, log-rank $p < 0.0001$, Figure 1A).

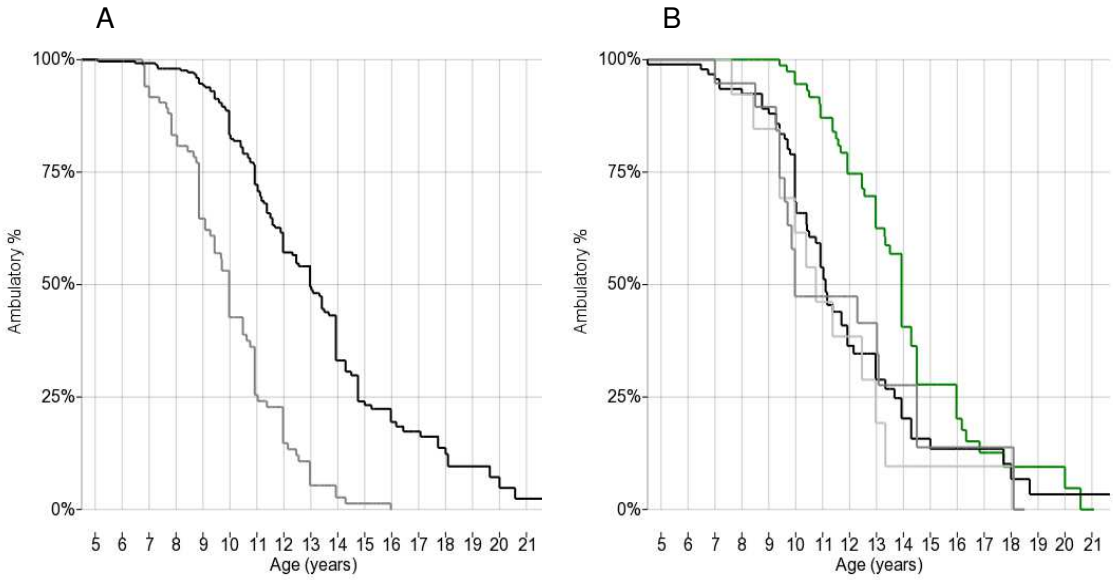

Figure 1. Kaplan-Meier plots of the proportion of ambulatory participants relative to age (years), grouped by GC treatment. A) Participants treated at least 1 year while ambulatory (n = 252, black line) vs. treated less or untreated (n = 88, grey line). B) Participants treated with the most common drug-regimen combinations: daily PRED (n = 94, black line), high-dose 2 days/week PRED (n = 19, dark grey line), low-dose intermittent PRED (n = 14, light grey line), and daily DFZ (n = 80, green line).

Table I. Distribution by GC regimen administered while ambulatory, with average daily dose, average age at start of treatment, and median age at LoA for each regimen.

Drug	Regimen	n	%	Dose† ± SD	Cumulative GC dose 1 = 1 year @ PRED 0.75 or DFZ 0.9 mg/kg/day	Start age (years) ± SD	*SPP1* rs28357094 TT genotype§ (%)	*LTBP4* rs10880 TT genotype§ (%)	Median age at LoA (years)
PRED	Daily	94	33.9%	75% ± 17%	2.96	6.6 ± 1.9	70.5%	14.7%	11.2
DFZ	Daily	80	28.9%	83% ± 15%**	4.73***	7.2 ± 2.0*	74.4%	9.3%	13.9***
PRED	Switched	23	8.3%	94% ± 37%*	4.30**	7.0 ± 2.0	58.8%	17.6%	11.6
Switched	Daily	21	7.6%	71% ± 16%	3.87	6.2 ± 2.3	68.7%	18.7%	13.4*
PRED	High-dose 2 days/week	19	6.9%	131% ± 36%***	5.64***	7.0 ± 2.1	76.5%	0.0%	10.0
Switched	Switched	15	5.4%	85% ± 26%	5.75***	5.2 ± 1.5	75%	20.0%	14.0**
DFZ	Switched	8	2.9%	82% ± 14%	3.64	6.2 ± 1.7	71.4%	0.0%	16.0
PRED	5 days/week	5	1.8%	71% ± 14%	1.88	8.0 ± 1.1	80.0%	20.0%	10.7‡
PRED	Every other day	4	1.4%	38% ± 9%	1.86	9.1 ± 1.9			
PRED	10 days on/off	2	0.7%	47% ± 4%	1.03	9.4 ± 0.4			
PRED	10 days/month	2	0.7%	50% ± 24%	0.27	6.1 ± 0.4			
DFZ	High-dose 2 days/week	2	0.7%	136% ± 10%	4.11	11.5 ± 2.9	-	-	-
DFZ	Every other day	1	0.4%	65% ± 0%	6.22	3.6 ± 0.0	-	-	-
PRED	Twice daily	1	0.4%	48% ± 0%	1.59	6.9 ± 0.0	-	-	-

n: participant number. **†** Dose is indicated as % of standard mg/kg/daily (0.75 mg/kg for PRED or 0.9 mg/kg for DFZ as applicable). **SD:** standard deviation. **LoA:** loss of ambulation. **PRED:** prednisone or prednisolone. **DFZ:** deflazacort. ***** Log-rank test vs. daily PRED $p < 0.05$. ****** p-value vs. daily PRED $p < 0.01$. ******* Log-rank test vs. daily PRED $p < 0.001$. ‡ Data for grouped low-dose intermittent PRED regimens, log-rank p vs. daily PRED n.s. **§** Genotypes were not available for all participants.

Distribution by GC regimen while ambulatory (Table I). As previously reported [Griggs et al., 2013], there was major variation in GC regimen prescription. Fourteen distinct regimens of PRED or DFZ were observed. PRED was administered while ambulatory to 150 participants (54.1% of treated), and DFZ to 91 (32.9%). Of 36 (13.0%) participants switching between drugs while ambulatory, 35 switched from PRED to DFZ (one later switching back to PRED), and one from DFZ to PRED. GCs were administered daily to 195 participants (70.4%), 2 days/week to 21 (7.6%), intermittently (including 10-days-on/10-days-off, 10 days/month, 5 days/week, QOD) to 14 (5.1%), and twice daily to one. Forty-six participants switched between regimens while ambulatory: 22 from non-daily to daily, 19 from daily to non-daily, and 5 between non-daily regimens.

Median LoA by regimen (Table I). The most frequently used treatment protocol (daily PRED, n = 94) was associated with a median age at LoA of 11.2 years. Median LoA was later in participants taking daily DFZ (13.9 years, n = 80, log-rank p = 0.0001), in "switchers" from daily PRED to daily DFZ (14.0 years, n = 21, log-rank p = 0.03), and "switchers" between different drugs and regimens (14.0 years, n = 15, log-rank p = 0.009). LoA in participants taking other regimens did not differ significantly from daily PRED. Kaplan-Meier plots of LoA for the most common regimens (daily PRED, daily DFZ, weekend PRED, and intermittent PRED) are shown in Figure 1B.

Dose (Table I). Average dose of daily PRED administered while ambulatory (n = 94) was 75% ± 17% of recommended, lower than daily DFZ (83 ± 15%, n = 80, p = 0.002). Doses for weekend regimens (and "switchers" to-from weekend) were higher (see Table I) because of the different protocol (10 mg/kg/week = 1.42 mg/kg/day).

Age at start of treatment (Table I). Average age at start of GC treatment (excluding treatments started after LoA) was 6.8 ± 2.1 years (range 2.0 - 14.2). Daily PRED was started earlier than daily DFZ (6.6 ± 1.9 vs. 7.2 ± 2.0 years, p = 0.03).

***SPP1* and *LTBP4* genotypes.** There were no significant differences in the frequencies of modifier genotypes between drug-regimen groups (Table I).

Time-varying Cox regression analysis of PRED vs. DFZ, regimen and dose (Table II). A Cox regression model was used to test concurrent, independent effects on LoA of several time-varying factors: use of PRED or DFZ; use of daily, low-dose intermittent, or high-dose weekend regimens; and average daily dose. The HR ± standard error (SE) associated to PRED was 0.498 ± 0.080, $p < 0.001$. DFZ treatment was associated to a lower HR (later LoA): HR 0.294 ± 0.053, $p < 0.001$. The linear test between covariate levels indicated that this difference was statistically significant ($p = 0.003$). HRs for different administration regimens were 0.382 ± 0.058 for daily, 0.362 ± 0.119 for intermittent, and 0.508 ± 0.135 for high-dose 2 days/week. None of the differences between regimens were statistically significant in this model (few participants treated non-daily). HR for dose was 0.392 ± 0.070, $p < 0.001$. Note that all Cox regression coefficients (Table II) are referred to covariate effects (drug, regimen, or dose) in the time-varying model, independent of grouping of individual participants by treatment (as in the Kaplan-Meier analyses); subsequently, data from “switcher” participants is included in Cox analyses. Also, the 1-year treatment threshold described above applies to log-rank tests of treated vs. untreated, and not to Cox regression results described in this paragraph.

Table II. Parameters for the time-varying Cox regression analysis of effects of GC drugs, regimens, and dose on LoA.

Covariates	Levels	HR	SE	p-value	95% CI	Linear tests between covariate levels
Drug	Untreated	1*	-	-	-	PRED vs. DFZ: **p = 0.003**
	PRED	0.498	0.080	**< 0.001**	0.363 – 0.683	
	DFZ	0.294	0.053	**< 0.001**	0.207 – 0.419	
Regimen	Untreated	1*	-	-	-	Daily vs. 2 days/week: p = 0.27 Daily vs. intermittent: p = 0.86 2 days/week vs. intermittent: p=0.38
	Daily	0.382	0.058	**< 0.001**	0.285 – 0.515	
	2 days/week	0.508	0.135	**0.011**	0.301 – 0.856	
	Intermittent	0.362	0.119	**0.002**	0.190 – 0.689	
Dose	% of standard	0.392	0.070	**<0.001**	0.277 – 0.553	-

HR: Hazard Ratio. **SE:** Standard Error. **CI:** Confidence Interval. **PRED:** prednisone or prednisolone. **DFZ:** deflazacort. ***** Untreated was used as reference in the model (HR set at 1).

Frequency of side effects (Table III). Side effect frequency was calculated in 277 participants (86.2%) with any treatment duration while ambulatory. Weight gain (65%), cushingoid appearance (55%), growth delay (37%), behavior changes (37%), low bone mass density (BMD) and/or fracture (22%), cataracts (15%), and skin abnormalities (13%) were most commonly reported. Some frequencies might be underestimated, because side effects were recorded only for the 3 most recent GC regimens before study baseline. We chose daily PRED, the most frequently prescribed regimen, as reference for comparisons. Weight gain frequency was similar for daily DFZ and daily PRED, but daily DFZ showed higher incidence of cushingoid appearance (72% vs. 50%, p = 0.002), growth delay (60% vs. 27%, p < 0.0001), and cataracts (29% vs. 5%, p < 0.0001). Behavior changes were more common in “switchers” between different drugs (p = 0.048), between different administration regimens (p = 0.04), or both (p = 0.001), suggesting that behavior disturbances might often induce clinicians and families to modify the

treatment. Reported growth delay was strikingly more frequent, other than in participants consistently on DFZ, also in "switchers" between drugs (p = 0.006 for daily treatment and p=0.03 for others), confirming a strong association between DFZ and stunted growth. On the contrary, growth delay was rare (5% vs. 27%, p = 0.04) with weekend GCs. Cataracts were more frequent, other than with daily DFZ, also in "switchers" (p < 0.0001). Skin abnormalities were more frequent with weekend GCs (p = 0.004). Finally, low-dose intermittent regimens showed a lower incidence of most side effects. This was statistically significant only for weight gain (23% vs. 67%, p = 0.002) and cushingoid appearance (0.004), arguably due to low numerosity in this group (n = 13).

Table III. Frequency of physician-reported side effects in participants treated with different GCs regimens while ambulatory.

Drug	Regimen	n	Weight gain	Cushingoid	Behavior change	Growth delay	Cataracts	Low BMD or fracture	Skin abnormalities	Hirsutism	Stomach pain	Anxiety or depression	Gastric ulcer	Diabetes	Headache	Sleep disturbance	Hypertension
PRED	Daily	94	67%	50%	30%	27%	5%	22%	11%	10%	2%	3%	3%	1%	1%	1%	0%
DFZ	Daily	80	63%	**72%****	33%	**60%*****	**29%*****	25%	8%	5%	3%	0%	0%	0%	3%	0%	1%
PRED	Switched	23	70%	48%	**52%***	17%	4%	9%	13%	0%	4%	0%	0%	4%	0%	0%	0%
Switched	Daily	21	76%	62%	**52%***	**57%****	14%	24%	24%	0%	0%	0%	10%	0%	0%	0%	0%
PRED	High-dose 2 days/week	19	79%	37%	42%	**5%***	11%	26%	**37%****	0%	0%	0%	0%	0%	0%	5%	0%
Switched	Switched	15	80%	67%	**73%***	**53%***	**40%*****	40%	27%	13%	0%	7%	0%	0%	0%	0%	0%
PRED	Low-dose intermittent‡	13	**23%****	**8%****	15%	8%	0%	23%	8%	8%	8%	8%	0%	8%	0%	0%	0%
DFZ	Switched	8	38%	25%	50%	25%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
All	All	277†	65%	55%	37%	37%	15%	22%	13%	6%	2%	2%	2%	1%	1%	1%	0%

n: participant number. **BMD:** Bone Mass Density. **PRED:** prednisone or prednisolone. **DFZ:** deflazacort. **†**All patients treated while ambulatory, not exactly equal to sum of other values because of a few patients on different, rarely prescribed regimens. **‡** Low-dose intermittent includes 10 days on/off, 10 days/month, 5 days/week, and every other day. ***** Chi-square p-value < 0.05, ****** < 0.01, and ******* < 0.001 compared to daily PRED.

Discussion

The long-term effect of GC treatment in prolonging independent ambulation in DMD, demonstrated by several previous studies [Angelini et al., 1994b; De Silva et al., 1987; McAdam et al., 2012; Manzur et al., 2008], is confirmed by data from the CINRG-DNHS presented here, with an estimated 3-year median delay of LoA. While virtually no untreated participants were able to walk beyond the age of 14, this was possible for approximately a third of GC-treated participants in the DNHS. However, because of inherent limitations of an observational, non-randomized study, these estimates of GC effect magnitude might be inflated. Recent years have seen a parallel increase in the frequency of GC prescription for DMD, and in the implementation of other standards of care such as physical therapy, management of joint contractures, and bone fracture prevention. In fact, CINRG-DNHS participants who did not receive GCs while ambulatory were significantly older, on average, than participants who did, denoting this "historical" improvement in care. It is not possible, in an observational, non-randomized study, to clearly discern how much of the observed LoA delay is actually caused by GCs, and how much by other treatments. Nevertheless, GC treatment was probably the single most important factor in this modification of the natural history of DMD.

PRED and DFZ regimens administered to CINRG-DNHS participants during the ambulatory phase of the disease were manifold, recapitulating a well-described variation in practice [Griggs et al., 2013]. The recent observational study from the North Star network [Ricotti et al., 2013a] reported on a cohort mostly treated with PRED, and compared daily and intermittent (mainly 10-days-on/10-days-off) regimens. The distribution of GC regimens was different in the CINRG-DNHS: a substantial part of the population was on DFZ, and daily regimens were preponderant.

Few studies have directly compared PRED and DFZ [Manzur et al., 2008]. Based on these, the two drugs appeared comparable in efficacy, the main differences residing in tolerability [Bonifati et al., 2000]. Therefore, we were surprised to observe a more than 2-year later median age at LoA between participants treated with daily DFZ compared to daily PRED (Kaplan-Meier analysis), and a significant reduction of estimated yearly LoA

risk with DFZ (Cox regression). This may be partly explained by higher average dosing in the DFZ group, in turn determined by more aggressive treatment, or, hypothetically, by a more favorable tolerability profile requiring less dose tapering. However, not only did we not observe a reduced incidence of weight gain with DFZ, but most common side effects were more frequent, suggesting that clinicians prescribing DFZ used higher doses in spite of side effects, and/or there was higher adherence to treatment. Earlier commencement of treatment, another hypothetical cause of increased efficacy, cannot be invoked to explain the better outcome in DFZ treated patients: on the contrary, daily PRED was started significantly earlier than daily DFZ in the CINRG-DNHS population. As it is common in clinical practice to start treatment when motor function reaches a plateau, DFZ treatment may have been started later because of a later plateau of motor function, which denotes in itself a milder disease progression. Furthermore, as many clinicians refrained from incrementing the dose with growth, as a means of managing side effects, participants started younger on PRED may have received lower cumulative doses, as the starting dose, calculated on a lower weight, was left unchanged in subsequent years. In summary, there was a strong association of DFZ with later LoA in the CINRG-DNHS, but this cannot be taken as conclusive evidence for a greater long-term efficacy.

DFZ is not commercially available in the United States, where many CINRG sites are located, and more expensive than prednisone, implying that its use may be associated to higher standards of care and possibly adherence. Nevertheless, it remains possible that DFZ does possess, because of uncharacterized pharmacodynamic mechanisms, a greater long-term efficacy than PRED, which could not be ascertained by previous short-term studies. The results of the time-varying Cox regression analysis (adjusted for dose as an independent factor, and for random effects of study Site to account for standards of care), appear to support an independent beneficial effect of DFZ. We also excluded differences in genetic modifier polymorphism frequency in the *SPP1* and *LTBP4* genes, which, as we recently reported, have a significant effect in this population [Bello et al., 2015a]. Randomized clinical trials, such as FOR-DMD, will shed more light on these

issues. Until then, some consideration should be given to stratifying clinical trial cohorts by DFZ/PRED treatment.

Data regarding non-daily GC regimens in the CINRG-DNHS is complex to analyze, due to their fragmentation, and the common practice of switching regimens as a means of tapering or adapting doses. HRs for daily vs. weekend regimens were not significantly different, consistent with findings of equivalence in quantitative muscle strength in a previous CINRG clinical trial [Escolar et al., 2011]. On the other hand, low-dose intermittent regimens (e.g. 10-days-on/10-days-off) were seldom used within CINRG, so that a conclusive comparison between these regimens and daily, such as recently published by the North Star clinical network [Ricotti et al., 2013a], cannot be obtained from CINRG-DNHS data.

Two common side effects of chronic GC treatment in the pediatric population, cushingoid appearance and growth stunting, were significantly more frequent with daily DFZ than daily PRED. Again, this might be explained at least in part by higher dosing, or possibly adherence. The previously reported higher incidence of cataracts with DFZ [McAdam et al., 2012] is confirmed by our data. On the other hand, we did not observe a lower frequency of weight increase with DFZ, as previously suggested [Bonifati et al., 2000], although it may be argued that if DFZ was dosed higher, a similar incidence of weight gain might still be the expression of better weight control with DFZ. Low-dose intermittent regimens (despite small participant numbers) showed lower frequencies of most side effects, as previously reported [Ricotti et al., 2013a]. The tolerability profile of weekend PRED appeared comparable to the daily regimen, as previously shown by a CINRG clinical trial [Escolar et al., 2011], except for less frequent growth stunting.

Consistent with comments following publication of GC treatment data from the North Star network [Ricotti et al., 2013a-b; Dubowitz, 2013a-b], growth stunting appeared to be associated with later LoA. Indeed, patients treated with daily DFZ showed both the latest LoA, and the most frequent growth stunting. It is difficult to discern from observational data whether a biomechanical advantage from short stature might play a causative role in delaying LoA, or if prolonged ambulation and short stature are simply

concurrent effects of treatment. An answer to this question might be provided by systematic and longitudinal correlations of stature and functional measures (e.g. strength, speed). From a clinical standpoint, the greatest consideration should be given to the impact that stunted growth, together with the frequently associated pubertal delay, has on the quality of life and self-image of DMD patients, in an effort to tailor GC treatment on the individual expectations and needs of each patient.

In conclusion, we provide evidence that GCs are effective in delaying LoA in patients with DMD, and that DFZ might be associated with greater long-term efficacy (i.e. later LoA), despite more frequent side effects. The observed better long-term outcome of DFZ might be at least partly due to higher dosing, higher adherence, and better standards of care. Nonetheless, stratification by PRED or DFZ treatment might be considered in clinical trials, in order to account for variability of weakness progression. This study emphasizes the need for further randomized, blinded, longitudinal trials of different GC regimens in DMD.

Aim 4: Genome-wide association study of loss of ambulation in DMD

Background

Duchenne muscular dystrophy is a monogenic X-linked disease affecting 1/3800-5000 males in all world populations [Mostacciuolo et al., 1987; Stark, 2015]. The *DMD* gene is the largest known in any genome, with 79 exons distributed over 2.3 megabases of Xp21 [Koenig et al., 1987]. *DMD* codes for the dystrophin protein, a membrane cytoskeletal component of myogenic cells [Hoffman et al., 1987]. Loss-of-function mutations of the *DMD* gene causing DMD lead to dystrophin deficiency in skeletal and cardiac cells. Dystrophin-deficiency, in turn, leads to myofiber membrane instability, bouts of myofiber necrosis, and eventual failed regeneration of skeletal muscle leading to muscle wasting and an early death.

Extensive natural history studies have shown considerable inter-patient variation in DMD onset and progression [McDonald et al., 2013; Henricson et al., 2013; Pane et al., 2014a]. Daily treatment with glucocorticoids (GCs) is considered standard of care [Bushby et al., 2010a], yet there is variable patient-patient response to treatment, both in terms of efficacy (improved gross motor skills) and side effect profiles [Henricson et al., 2013; Bello et al., 2015b]. Also, there is considerable variation in the practice of prescription and use of glucocorticoids [Griggs et al., 2013; Bello et al., 2015b].

Multiple factors contribute to the observed clinical variation in DMD, including standards of care, GC use, and genetic background (modifying polymorphisms, ethnicity). These variables are not independent. For example, a patient's geographic location has effects on ethnicity, standard of care, glucocorticoid use, and genetic modifiers [Bello et al., 2015a]. The inter-dependent nature of these variables can make it challenging to isolate a single variable and define its contribution to disease expressivity.

Despite these challenges, the identification of genetic modifiers of Duchenne muscular dystrophy is broadly felt to be important, and initial progress has been made using a candidate gene approach [Vo and McNally, 2015]. The *SPP1* gene locus, encoding osteopontin protein (OPN) was identified as a candidate by mRNA profiling studies of muscle biopsies from clinically mild and severe patients [Pegoraro et al., 2011]. A polymorphism in the promoter of the *SPP1* gene (rs28357094) known to change *SPP1* mRNA expression by 5-fold [Giacopelli et al., 2004], was found to be associated with age at loss of ambulation (LoA) in an Italian cohort, and grip strength in the Cooperative International Neuromuscular Research Group Duchenne Natural History Study (CINRG DNHS) [Pegoraro et al., 2011]. The association of rs28357094 with ambulation phenotypes was validated in a second Italian multi-center cohort [Bello et al., 2012], as well as in the CINRG cohort [Bello et al. 2015a]. However, the same association of rs28357094 with LoA was not seen in a European multicenter cohort (Bio-NMD) [van den Bergen et al., 2015] and in a United States multicenter cohort (United Dystrophinopathy Project, UDP) [Flanigan et al., 2013]. More detailed studies of rs28357094 within an expanded CINRG cohort found that both ethnicity and steroid use altered findings of genotype/phenotype associations, and it was concluded that the rs28357094 *SPP1* locus was likely a pharmacogenetic marker, influencing patient response to GC treatment [Bello et al. 2015b]. As the proportion of patients treated with GC vary from cohort to cohort, glucocorticoid treatment is an important covariate to include in statistical models studying genetic modifiers. The effects of the rs28357094 *SPP1* locus was also studied in adult volunteer populations, where it was found associated with upper arm muscle volume [Hoffman et al., 2013] and with response to eccentric activity-induced muscle damage [Barfield et al. 2014] in young adult females.

A second modifier of muscular dystrophy was identified by genetic linkage with disease severity in an outbred murine γ-sarcoglycan deficient model [Heydemann et al., 2009], which pointed to the *Ltbp4* murine gene. This finding translated to human DMD, as a coding haplotype in *LTBP4* was associated with age at LoA [Flanigan et al., 2013]. This association was replicated in the Bio-NMD cohort [van den Bergen et al., 2015] and in

participants of European ancestry in the CINRG cohort [Bello et al., 2015a], although not in an Italian multi-center cohort [Barp et al., 2015].

The candidate gene approach has led to the identification of two robust genetic modifier loci, *SPP1* and *LTBP4*. The products of both these genes are involved in secondary inflammation and tissue remodeling, which appear to be key modifier pathways of muscular dystrophy [Chen at al., 2005; Ceco and McNally, 2014]. *SPP1* is heavily upregulated by activation of NF-κB signaling during bouts of degeneration and regeneration in skeletal muscle [Hoffman et al., 2013]. NF-κB signaling is triggered in DMD by necrotic cells liberating damage-associated molecular patterns, which stimulate Toll-like receptors (TLRs) [Rosenberg et al., 2015]. In turn, upregulated osteopontin modulates TGFβ-mediated signals [Vetrone et al., 2009], which dictate either successful (reparative) or unsuccessful (fibrotic) regeneration. The LTBP4 protein also regulates TGFβ signaling by binding TGFβ in a latent complex in the extracellular matrix [Flanigan et al., 2013], and the modifier haplotype appears to influence susceptibility to proteolitic cleavage and subsequent TGFβ signaling activation [Ceco et al., 2014].

However, the candidate gene approach for identification of genetic associations is known to be subject to various biases [Wills et al., 2009]. Genome-wide association studies (GWAS) are broadly felt to show less bias in genetic association discovery. GWAS studies typically employ highly parallel statistical tests of genetic association, where thousands to millions of polymorphic loci are tested simultaneously. In order to withstand multiplicity adjustment of association p-values, large populations have to be studied for typical effect sizes. This requirement works against utilization of GWAS to identify rare monogenic disease modifiers.

In order to work around this problem, we took a “hybrid” approach, in between a “hypothesis free” GWAS and a candidate gene study. We ran a GWAS of age at LoA in a small (for GWAS standards) “discovery” cohort, adjusting for glucocorticoid treatment. High-ranking loci were then prioritized based on the TGFβ and TLR-NF-κB pathways, adjusting multiplicity correction to 3 genes functionally annotated to be involved in these pathways. We then moved to additional testing in four independent validation

cohorts. This led to the identification of a novel robust genetic modifier of DMD, the *CD40* locus.

Methods

Informed consent and ethics approvals. All patients or their legal guardians consented to the use of genetic and clinical data for research purposes, and study procedures were reviewed by local IRB/Ethics Committees, as previously reported for CINRG [McDonald et al., 2013], Bio-NMD [van den Bergen et al., 2015], Padova [Pegoraro et al., 2011], and UDP [Flanigan et al., 2013].

Inclusion criteria. Inclusion criteria have been previously reported for CINRG [McDonald et al., 2013], Bio-NMD [van den Bergen et al., 2015], Padova [Pegoraro et al., 2011], and UDP [Flanigan et al., 2013]. Collectively, criteria were aimed at the inclusion of "typical" DMD patients with truncating *DMD* mutations, except for the UDP study, which explicitly included intermediate muscular dystrophy and Becker muscular dystrophy patients. These were excluded from validation analyses presented here.

Exome chip genotyping and data cleaning. Exome chip genotyping and data cleaning methods in the CINRG-DNHS cohorts have been described [Bello et al., 2015a]. Briefly, genotyping with the Illumina (San Diego, CA) HumanExome chip was performed in 175/340 CINRG-DNHS participants of different ethnicities, who were selected on the basis of available DNA quantity and quality, and did not differ for clinical or demographic features. Genotype calling was performed with the Genome Studio software, and genotype data were exported into PLINK format with the dedicated plug-in software by Illumina. Data cleaning was performed by PLINK and included missing call thresholds of 0.01 for both individuals and SNPs; a heterozygosity threshold of ± 4 standard deviations from the mean; check for cryptic duplicates and relatedness in an IBS matrix (PIHAT threshold of 0.1). A subcohort of 109 unrelated individuals of European/European American descent was selected by multidimensional scaling (MDS) analysis of Exome Chip genotypes as described [Bello et al., 2015a], and MDS was repeated in this population to check for residual population stratification or outliers.

Association of LoA with Exome Chip genotypes. Genome-wide association with age at LoA was tested within a Cox proportional hazards model. The dependent variable

(phenotype) was age at LoA, with censoring of participants who were ambulatory at the last evaluation. The independent variables were genotypes (additive inheritance model) of 27,027 Exome Chip SNPs with MAF > 0.05, and GC treatment coded as a binary categorical covariate: treated at least 1 year while ambulatory vs. treated < 1 year or untreated while ambulatory. The Cox proportional hazards test was performed by plugging the R function "coxph" into PLINK via the Rserve package. A Bonferroni corrected p-value of $1.8*10^{-6}$ (0.05/27,027 SNPs, "exome-wide" significance) was set for this experiment. QQ and Manhattan plots were created with the "qqman" package [Turner, 2014] in R.

Prioritization of GWAS results. In order to select nominally significant SNPs with a suggestive association p-value in the GWAS, we performed a hypothesis-driven prioritization. The hypothesis was that the NF-κB and TGFβ signaling pathways modulate pathology downstream of dystrophin deficiency in skeletal muscle. Genes included in these pathways were downloaded via the Prowler® online software from the Panther classification system (Gene Ontology). NF-κB pathway genes include: *ADAM8, ADIPOQ, AGER, AGT, AIM2, AJUBA, ALK, AMH, ANKRD42, APOL3, AR, ARHGEF2, ATP2C1, BCL10, BCL3, BIRC2, BIRC3, BMP7, BRD4, BST2, BTK, BTRC, C18orf32, C1QTNF3, C9orf89, CAMK2A, CANT1, CAPN3, CARD10, CARD11, CARD14, CARD8, CARD9, CASP1, CASP10, CASP8, CAT, CC2D1A, CCL19, CCL21, CCR7, CD27, CD36, CD40, CFLAR, CHI3L1, CHUK, CIB1, CLEC6A, CLOCK, CLU, COPS8, CTH, CTNNB1, CXXC5, CYLD, DAB2IP, DNAJA3, ECM1, ECT2, EDA, EDA2R, EDAR, EEF1D, EIF2AK2, ERC1, ESR1, F2R, F2RL1, FADD, FAF1, FASLG, FBXW11, FER, FKBP1A, FLNA, FYN, G3BP2, GJA1, GOLT1B, GPR89A, GPRC5B, GSTP1, HMOX1, HSPB1, HTR2B, ICAM1, IKBIP, IKBKAP, IKBKB, IKBKE, IKBKG, IL10, IL12B, IL18, IL18R1, IL1B, IL1RL1, IL23A, IL6, IL6R, INS, IRAK1, IRAK2, IRAK3, IRAK4, IRF3, KRAS, LGALS1, LGALS9, LITAF, LPAR1, LTBR, LTF, LURAP1, MALT1, MAP3K13, MAP3K14, MAP3K3, MAP3K7, MAS1, MAVS, MIB2, MID2, MIER1, MTDH, MTPN, MUL1, MYD88, NDFIP1, NDFIP2, NEK6, NFKB1, NFKB2, NFKBIA, NFKBIB, NFKBID, NLRC3, NLRC4, NLRP12, NLRP3, NLRP6, NLRX1, NOD1, NOD2, NPM1, NTRK1, NUP62, OLFM4, OTUD7B, PAMP, PARK2, PDPK1, PELI1, PELI2, PER1, PIDD1, PIM2, PINK1, PLA2G1B, PLEKHG5, PLK2,*

PPM1A, PPM1B, PPP4C, PPP5C, PRDX1, PRDX3, PRDX4, PRKCB, PRKCE, PRKCH, PRKCI, PRKCQ, PRKCZ, PRKD1, PRKD2, PSMA6, PSMD10, PTGS2, PYCARD, RASSF2, RBCK1, RC3H1, RC3H2, REL, RELA, RHEBL1, RHOA, RHOC, RHOH, RIPK1, RIPK2, RIPK3, RIPK4, RNF25, RNF31, RORA, RPS27A, RPS3, RPS6KA4, RPS6KA5, S100A12, S100A13, S100A4, S100A8, S100A9, S100B, SASH1, SECTM1, SHARPIN, SHISA5, SIRT1, SLC20A1, SLC35B2, SLC44A2, SPHK1, SQSTM1, STAT1, TAB1, TAB2, TAB3, TAK1, TBK1, TERF2IP, TFG, TGFB1, TGFBR3, TGM2, TICAM1, TICAM2, TIRAP, TLE1, TLR2, TLR3, TLR4, TLR6, TLR7, TLR9, TMED4, TMEM101, TMEM9B, TNF, TNFAIP3, TNFRSF10A, TNFRSF10B, TNFRSF11A, TNFRSF19, TNFRSF1A, TNFSF10, TNFSF11, TNFSF14, TNFSF15, TNFSF18, TNIP1, TNIP2, TNIP3, TRADD, TRAF1, TRAF2, TRAF3IP2, TRAF4, TRAF5, TRAF6, TRIM13, TRIM14, TRIM15, TRIM22, TRIM25, TRIM32, TRIM37, TRIM38, TRIM40, TRIM5, TRIM52, TRIM59, TRIM62, TRIM8, TRIP6, TSPAN6, UACA, UBA52, UBB, UBC, UBD, UBE2N, UBE2V1, UNC5CL, VAPA, WLS, WNT5A, ZC3HAV1, ZDHHC13, ZDHHC17, ZFAND6, ZFP91, ZMYND11, ZNF268. TGFβ pathway genes include: *ACVR1, ACVR1B, ACVR1C, ACVR2A, ACVR2B, ACVRL1, AMHR2, ATF2, BAMBI, BMP1, BMP10, BMP15, BMP2, BMP3, BMP4, BMP5, BMP6, BMP7, BMP8A, BMP8B, BMPR1A, BMPR1B, BMPR2, CITED1, CITED2, CITED4, CREBBP, DCP1A, DCP1B, EP300, FKBP1A, FOSL1, FOXH1, GDF1, GDF10, GDF11, GDF15, GDF2, GDF3, GDF5, GDF6, GDF7, GDF9, GDNF, HRAS, INHBA, INHBB, INHBC, INHBE, JUN, JUNB, JUND, KRAS, LEFTY1, LEFTY2, MAP3K7, MAP3K7CL, MAPK1, MAPK10, MAPK11, MAPK12, MAPK13, MAPK14, MAPK3, MAPK8, MAPK9, MSTN, NODAL, NRAS, RRAS, RRAS2, SKI, SKIL, SMAD1, SMAD2, SMAD3, SMAD4, SMAD5, SMAD6, SMAD7, SMAD9, SMURF1, SMURF2, SNIP1, SOD1, TAB1, TGFB1, TGFB2, TGFB3, TGFBR1, TGFBR2, TLL1, TLL2, ZFYVE9*. We selected 506 SNPs in the GWAS whose genomic coordinates lie within, or 10,000 kb up- or downstream from these genes. Note that while the Exome Chip contains mostly exonic SNPs, it also includes intronic or intergenic SNPs, mainly selected because of previous GWAS hits or other evidence of a functional regulatory role (chip design details at http://genome.sph.umich.edu/wiki/Exome_Chip_Design). We defined a "suggestive" Bonferroni-corrected significance threshold as p = 0.05/506 prioritized SNPs = $9.9*10^{-5}$.

Validation cohort and methods. Initial validation of prioritized SNPs below the suggestive p-value threshold was performed in 108 CINRG participants who were not included in the GWAS experiment because of insufficient DNA quantity, but still had available DNA for targeted genotyping by TaqMan allele discrimination assays. As this was a multi-ethnic cohort, at risk for population stratification bias, and no genome-wide markers were available for MDS analysis, we restricted validation analyses to 76 participants self-identifying with the Caucasian race and non-Hispanic ethnicity. Subsequently, we pooled clinical data (ambulatory vs. non-ambulatory; age at LoA or last follow-up; GC treated or untreated before LoA) from the following independent cohorts: Bio-NMD cohort [van den Bergen et al., 2015] consisting of 246 DMD patients from the Centers of Ferrara, Leiden, London, Montpellier, and Newcastle, including mostly patients of European descent; 95 Italian DMD patients followed at the University of Padova [Pegoraro et al., 2011]; 243 DMD patients (only those carrying truncating *DMD* mutations) from the US-based UDP cohort [Flanigan et al., 2013]. The statistical test used for validation was the same as in the GWAS, except for the following differences in the definition of the binary GC treatment covariate: in the Bio-NMD and Padova cohort, patients with any treatment duration before LoA were classified as "treated" (detailed treatment duration or dates were not available); in the UDP cohort, patients with at least 6 months of GC treatment before LoA were classified as "treated". Furthermore, a categorical covariate for Center (CINRG, Bio-NMD Ferrara, Bio-NMD Leiden, Bio-NMD London, Bio-NMD Montpellier, Bio-NMD Newcastle, Padova, UDP) was added to the Cox proportional hazards model in the pooled validation analysis. Both additive and dominant inheritance models were tested in validation. Statistical significance was set at $p < 0.05$ for validation.

Results

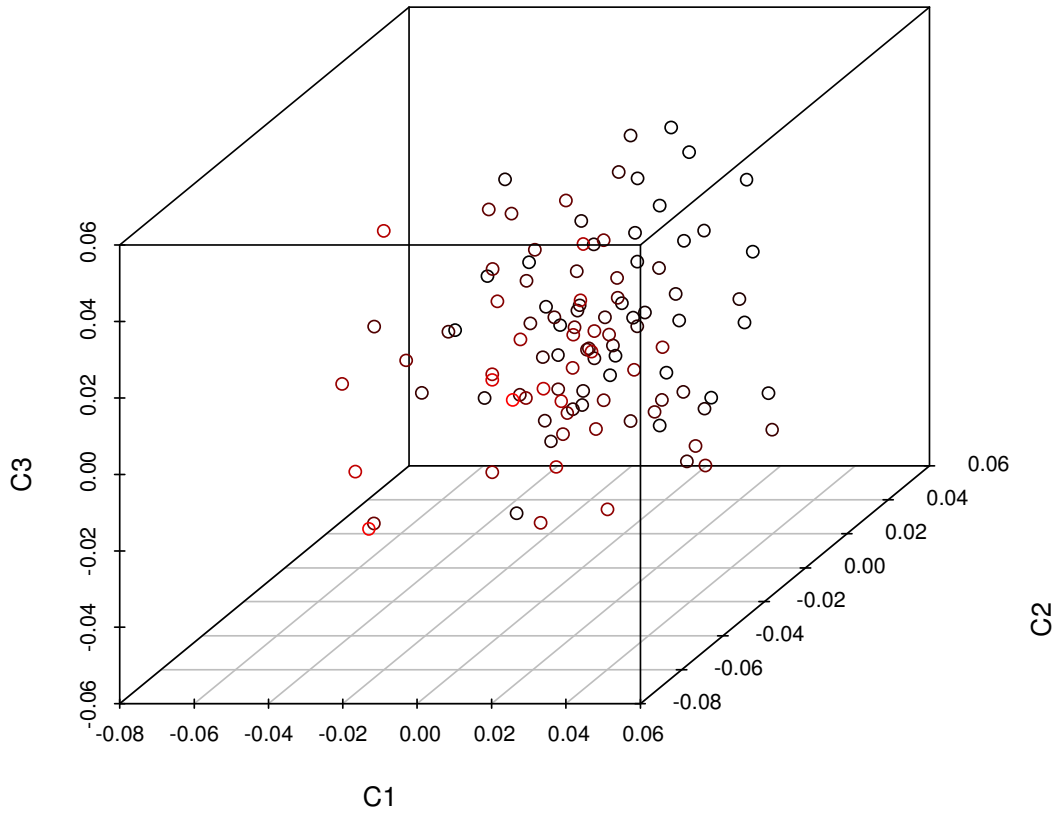

Figure 1. Tridimensional scatter plot of the first 3 principal components obtained by multidimensional scaling of genome-wide genotypes in 109 unrelated CINRG-DNHS participants of European/European American descent (CINRG Exome Chip cohort). A color gradient of red to black highlights negative to positive values along the C2 axis. No clustering of subgroups or extreme outlier individuals are observed.

Of 340 CINRG-DNHS participants from 20 worldwide Centers, 175 (selected only on the basis of DNA availability) were genotyped with the Illumina HumanExome chip, focused on common and rare functional variants in gene-coding regions. For *SPP1* and *LTBP4* validation, a sub-cohort of European/European-American ancestry had been selected by

multidimensional scaling (MDS) to reduce population stratification bias [Bello et al., 2015a]. Our initial GWAS included 109 unrelated participants in this sub-cohort. Repeated MDS confirmed no gross stratification (Figure 1).

The statistical test in the GWAS was a Cox proportional hazards model using age as time variable, LoA as event, GC treatment (defined in Methods) as a binary covariate, and an additive genotype model. The survival approach, unusual for GWAS, allowed inclusion of ambulatory ("censored") participants. We applied a MAF threshold of > 0.05, as the very low numerosity precluded single-SNP or groupwise analyses of rare variants. These would require larger populations, or genotyping/sequencing of phenotypic extremes.

The quantile-quantile (QQ) plot of observed p-values (Figure 2) excluded gross systematic bias ($\lambda_{GC} = 1.09$). In the Manhattan plot (Figure 3) no SNP reached the "exome-wide" significance threshold of $p = 1.8*10^{-6}$ (Bonferroni correction for 27,027 SNPs). Top p-value annotations are shown in Table 1.

Due to recognized low statistical power in the initial GWAS, lack of "exome-wide" significance was expected. Nevertheless, some nominally significant p-values might hold true association. We decided to prioritize SNPs for independent validation. To this end, we introduced the hypothesis that SNPs within, or close to (10,000 kb) genes in the NF-κB and TGFβ signaling pathways would be enriched for true associations. These are extensively studied inflammatory and pro-fibrotic pathways in dystrophin deficiency [Chen et al., 2005; Ceco and McNally, 2013; Rosenberg et al., 2015]. Furthermore, both known modifiers, which derived from unbiased hypothesis-generating experiments, i.e. expression profiles for *SPP1* [Pegoraro et al., 2011], murine genome mapping for *LTBP4* [Flanigan et al., 2013), converged into these pathways.

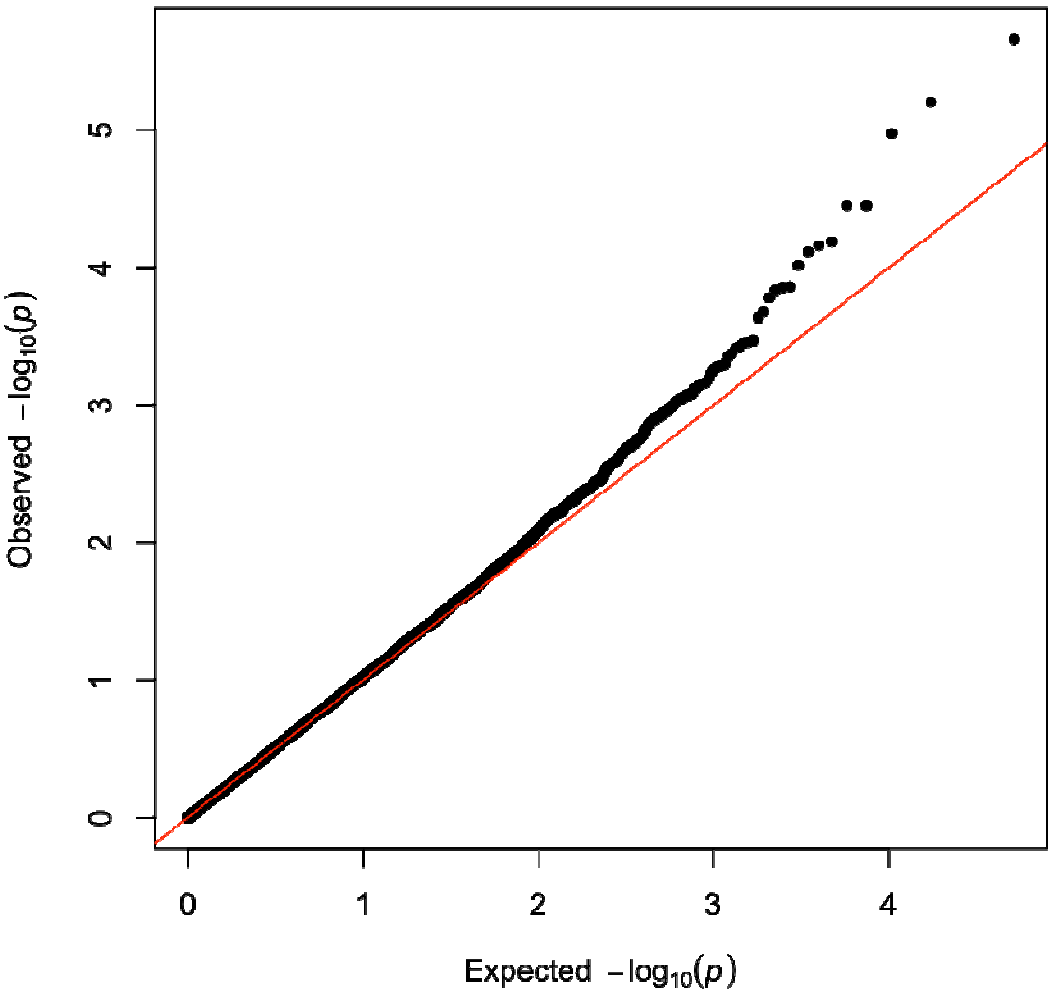

Figure 2. Quantile-quantile plot of p-values observed in the GWAS experiment of age at LoA with Exome Chip genotypes (27,027 SNPs with MAF > 0.05).

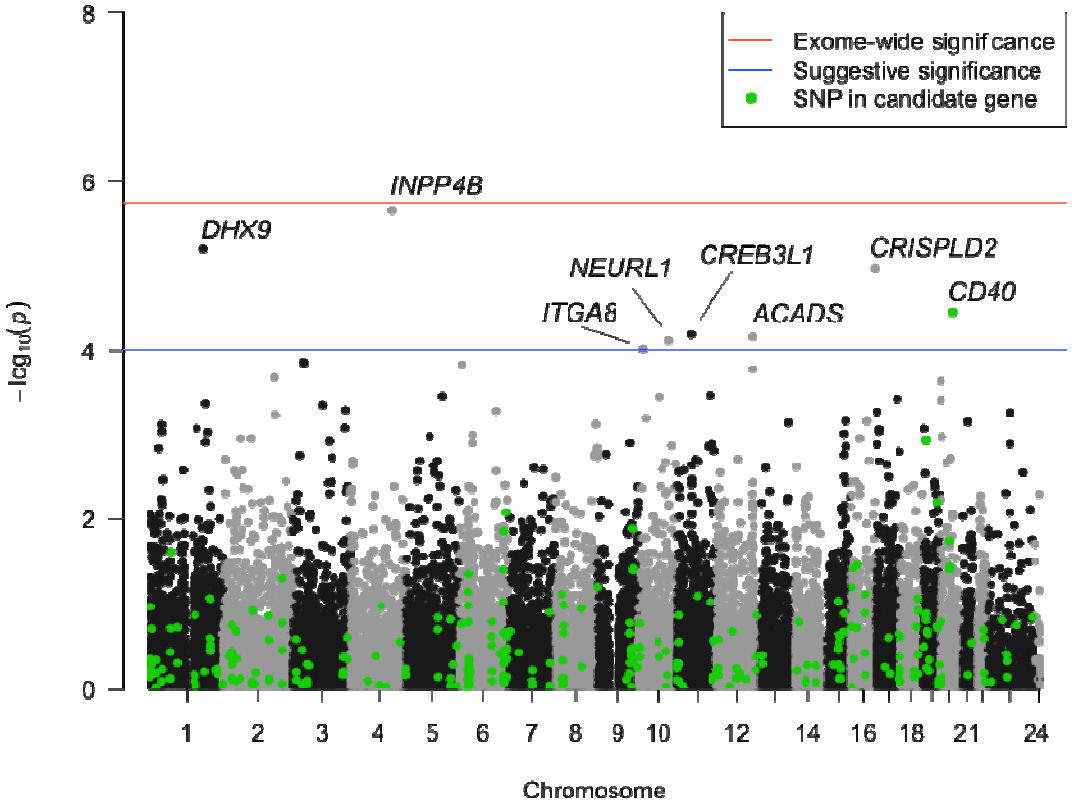

Figure 3. Manhattan plot of GWAS of age at loss of ambulation (Cox proportional hazards with glucocorticoid treatment as a covariate) for 27,027 Exome Chip SNPs with MAF > 0.05. SNPs within (<10,000 kb) candidate genes in the NF-κB and TGFβ pathways are highlighted in green. The red horizontal line (p = 1.8*10-6) refers to Bonferroni correction for 27,027 SNPs, while the blue line (p = 9.9*10-5) refers to Bonferroni correction for 506 SNPs within (<10,000 kb) candidate genes. Top p-values are annotated with gene names (also see Table 1).

Table 1. SNPs showing top p-values in the GWAS of age at loss of ambulation in 109 unrelated participants of European ancestry in the CINRG-DNHS

SNP	Chr	BP	Alleles	Minor allele	MAF	BP from gene	Gene	Mutation	GWAS p-value	Expressed in 17 DMD muscle biopsies	Expressed in 6 normal muscle biopsies	Expression probeset
rs34561493	4	143043397	A/G	A	0.09	0	*INPP4B*	Synonymous S673S	2.18E-06	Yes (17/17)	Yes (6/6)	205376_at
rs4275414	1	182854200	A/G	A	0.27	0	*DHX9*	Intronic	6.28E-06	Yes (17/17)	Yes (6/6)	212107_s_at
rs72799568	16	84902483	T/A	A	0.06	0	*CRISPLD2*	Missense M294L	1.06E-05	Yes (17/17)	Yes (6/6)	221541_at
rs4810485	20	44747947	A/C	A	0.28	0	*CD40*	Intronic	3.53E-05	Yes (17/17)	Yes (6/6)	35150_at
rs6074022	20	44740196	A/G	G	0.28	6710 (5')	*CD40*	Promoter	3.53E-05	Yes (17/17)	Yes (6/6)	35150_at
rs35652107	11	46339011	A/G	A	0.07	0	*CREB3L1*	Missense A411T	6.48E-05	Yes (13/17)	No (0/6)	213059_at
rs2014355	12	121175524	A/G	G	0.22	0	*ACADS*	Intronic	6.90E-05	Yes (17/17)	Yes (6/6)	202366_at
rs2281859	10	105271758	A/G	G	0.36	0	*NEURL1*	Intron - nc transcript	7.69E-05	Yes (17/17)	Yes (6/6)	204889_s_at
rs9333269	10	15649698	A/C	C	0.08	0	*ITGA8*	Missense Q581P	9.61E-05	Yes (17/17)	Yes (5/6)	235666_at

SNP: single nucleotide polymorphism. **GWAS:** genome wide association study. **CINRG-DNHS:** Cooperative International Neuromuscular Research Group Duchenne Natural History Study. **Chr:** chromosome. **BP:** base pair position. **MAF:** observed minor allele frequency. **DMD:** Duchenne muscular dystrophy. **nc:** non-coding.

We prioritized 506 SNPs in 392 NF-κB/TGFβ-related genes (see text) hence a "suggestive" Bonferroni-corrected threshold of $p = 9.9*10^{-5}$, surpassed only by two neighboring SNPs in perfect LD (rs6074022 and rs4810485), situated respectively 6710 bp upstream and within the *CD40* gene on the long arm of chromosome 20. Median age at LoA in carriers of at least one copy of the minor allele was 2.8 years earlier (Figure 4,

panel A), per-copy Hazard Ratio (HR) 2.10 (95% Confidence Interval [CI] 1.45 - 3.04), p = $3.5*10^{-5}$. Thus, this locus was selected for independent validation.

CD40, also known as *TNFRSF5* (Tumor Necrosis Factor Receptor Super-Family member 5), was included in the prioritization list as a NF-κB pathway component, and encodes a costimulatory protein for the Th polarization of T cells, found on the surface of antigen presenting cells (APCs). The SNPs rs6074022 and rs4810485 are located in the promoter and first intron respectively, tagging a haplotype spanning the 5' region of the gene. *CD40* is expressed in healthy and dystrophic muscle (positive present call analysis on 17/17 DMD muscle biopsy samples and 6/6 healthy muscle samples, probeset 35150_at). Genotypes were in Hardy-Weinberg equilibrium (HWE), and the MAF of 28% was close to expected for European ancestry (24% in 1000 Genomes [1000G] CEU).

Targeted genotyping (TaqMan) of rs1883832 (C/T, minor allele T) was used for validation. This SNP is situated in the *CD40* 5' UTR between rs4810485 and rs6074022, in perfect LD with both (confirmed by TaqMan in the Exome Chip cohort).

The first validation step was performed in 108 CINRG-DNHS participants, not overlapping the Exome Chip cohort. Carriers of the T allele showed 1.5-year earlier LoA, Cox proportional hazards p = 0.07 for additive and 0.02 for dominant genotype model (data not shown). Although these data pointed towards validation of rs1883832 as a modifier, we were aware of potential population stratification bias, as some non-European ancestries, e.g. East Asian, show higher rs1883832 MAF (1000G), and earlier LoA in the CINRG-DNHS [Bello et al., 2015a]. Thus, we limited validation to 76/108 participants of self-identified non-Hispanic European race/ethnicity. In these, the T allele was associated with 0.8-year earlier LoA, p = n.s. (Figure 4, panel B).

Subsequently, we expanded validation studies to other independent DMD cohorts, using the same statistical model (except minor differences in GC treatment covariate definition, see Methods). In the Bio-NMD cohort (n = 246), the T allele was associated with 0.6-year earlier median LoA (Figure 4, panel C), p = 0.08 (additive) and 0.0496 (dominant). In the Padova cohort (n = 95), the T allele was associated with a 0.3-year

earlier median LoA (Figure 4, panel D), p = n.s.. Finally, in the UDP cohort (n = 243 DMD, mostly European/American) the T allele was associated with a 0.5-year earlier median LoA (Figure 4, panel E), p = 0.13 (additive) and 0.04 (dominant).

The pooled validation cohorts comprised 660 DMD patients, 50% treated with GCs, 47% untreated, and 8% with unknown GC treatment status. The *CD40* rs1883832 SNP was in HWE, with a MAF of 28% (close to expected for European ancestry). The T allele was associated overall with a 1-year earlier median LoA (Figure 4, panel F), per-copy HR = 1.16 (95% CI 1.02 - 1.32), p = 0.02 (additive model), and HR 1.31 (95% CI 1.10 - 1.56), p = 0.002 in the dominant model. This pooled validation analysis included a covariate to adjust for differences between cohorts (see Methods). Survival analysis parameters for all cohorts are presented in Table 2.

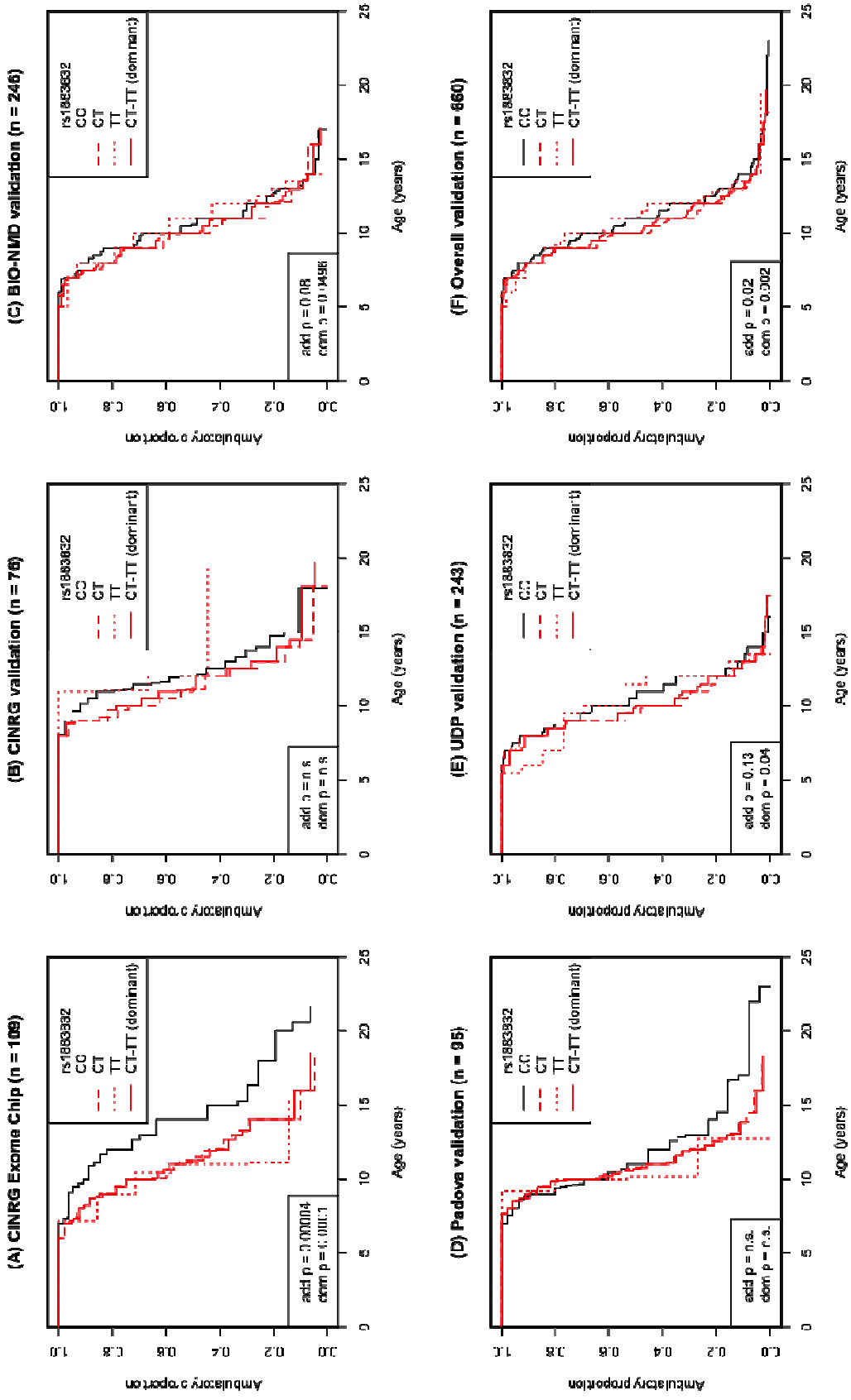
(A) CINRG Exome Chip (n = 109)
(B) CINRG validation (n = 76)
(C) BIO-NMD validation (n = 246)
(D) Padova validation (n = 95)
(E) UDP validation (n = 243)
(F) Overall validation (n = 660)
rs1883832
CC
CT
TT
CT-TT (dominant)
Ambulatory proportion
Age (years)

Figure 4. Kaplan-Meier plots of age at loss of ambulation by *CD40* rs1883832 genotype (additive and dominant models) in the CINRG Exome Chip cohort (panel A) and validation cohorts: CINRG self-identified Caucasian validation cohort (panel B), Bio-NMD cohort (panel C), Padova cohort (panel D), UDP cohort (panel E), and overall validation cohort (sum of CINRG validation Bio-NMD, Padova, and UDP, panel F). "add": additive inheritance model. "dom": dominant inheritance model.

Table 2. Parameters for Kaplan-Meier and Cox proportional Hazards analyses of age at loss of ambulation by *CD40* rs1883832 genotype in the CINRG Exome Chip and validation cohorts.

Cohort	Kaplan-Meier analysis parameters						Cox proportional hazards parameters			
	Parameter	Genotype					Parameter	Covariate		
		CC	CT	TT	CT and TT (dominant)	Total		Additive genotype	Dominant genotype	GC treatment
CINRG Exome Chip cohort	**n**	56	44	9	53	109	**HR**	2.10	2.64	0.16
	Median age at LoA (years)	14.0	11.3	11.0	11.2	13.0	**95% CI**	1.45 - 3.04	1.60 - 4.35	0.09 - 0.29
	95% CI	13.0 - 15.2	10.0 - 13.2	9.0 - NA	10.4 - 13.0	12.0 - 14.0	**p-value**	*0.00004*	*0.0001*	*< 0.0001*
CINRG validation cohort	**n**	42	28	6	34	76	**HR**	1.21	1.02	0.29
	Median age at LoA (years)	12.0	11.0	12.0	11.2	12.0	**95% CI**	0.69 - 2.11	0.67 - 1.55	0.15 - 0.55
	95% CI	11.6 - 13.8	10.0 - 13.0	11.1 - NA	10.5 - 13.0	11.2 - 12.5	**p-value**	n.s.	n.s.	*0.0002*
BIO-NMD validation cohort	**n**	118	98	30	128	246	**HR**	1.22	1.36	0.31
	Median age at LoA (years)	10.6	10.0	11.0	10.0	10.5	**95% CI**	0.98 - 1.51	1.00 - 1.84	0.21 - 0.44
	95% CI	10.0 - 11.0	9.6 - 10.9	10.0 - 12.5	10.0 - 11.0	10.0 - 11.0	**p-value**	0.08	*0.0496*	*< 0.0001*
Padova validation cohort	**n**	47	40	8	48	95	**HR**	1.20	1.19	0.41
	Median age at LoA (years)	11.0	10.8	10.2	10.7	11.0	**95% CI**	0.81 - 1.79	0.75 - 1.90	0.25 - 0.67
	95% CI	10.0 - 13.0	10.2 - 11.9	10.0 - NA	10.2 - 11.9	10.3 - 12.0	**p-value**	n.s.	n.s.	*0.0004*
UDP validation cohort	**n**	139	91	13	104	243	**HR**	1.18	1.32	0.68
	Median age at LoA (years)	10.5	9.5	11.5	10.0	10.0	**95% CI**	0.96 - 1.45	1.01 - 1.73	0.52 - 0.89
	95% CI	10.0 - 11.0	9.0 - 10.0	9.5 - NA	9.0 - 10.0	10.0 - 10.5	**p-value**	0.13	0.038	*< 0.0001*
Overall validation cohort*	**n**	346	257	57	314	660	**HR**	1.16	1.31	0.483
	Median age at LoA (years)	11.0	10.0	11.1	10.0	10.6	**95% CI**	1.02 - 1.32	1.10 - 1.56	0.40 - 0.58
	95% CI	10.5 - 11.0	10.0 - 10.5	10.2 - 12.0	10.0 - 10.5	10.2 - 11.0	**p-value**	*0.02*	*0.002*	*0.005*

CINRG: Cooperative International Neuromuscular Research Group. **GC:** glucocorticoid corticosteroids. **LoA:** loss of ambulation. **CI:** confidence interval. **HR:** Hazard Ratio.

Discussion

The hypothesis-prioritized GWAS suggested a modifier effect of the *CD40* locus, which was subject to a strong independent validation in a large independent, collaborative cohort. Because of a "winner's curse" effect, effect size is substantially smaller in the validation than GWAS cohorts (1 vs. 2.8 years); the smaller size effect in the larger validation cohort is probably the most realistic estimate. The GWAS was run with an additive inheritance model, representing the most balanced hypothesis when testing thousands of loci in parallel, as a complete dominant effect is rare for common SNPs. Nevertheless, median LoA by genotype in studied populations suggests a dominant model for a damaging effect of the T allele on the ambulation phenotype in DMD.

The C>T transition at rs1883832, adjacent to translation start in the 5' UTR, disrupts a translationally relevant Kozak sequence, and the promoter SNP rs6074022 (in perfect LD with rs1883832) influences *CD40* transcriptional activity [Gandhi et al., 2010]. Furthermore, the minor allele at this locus has been associated to enhanced alternative splicing of a Δexon6 secreted isoform, which might act as a decoy receptor [Onouchi et al., 2012]. Overall the minor allele, which we observed in association to earlier LoA in DMD, seems to downregulate CD40 signaling by multiple transcriptional and post-transcriptional mechanisms [Field et al., 2015]. The list of described genetic associations at this locus is long [Jacobson et al., 2005; Pineda et al., 2008; Raychaudhuri et al., 2008; Skibola et al., 2008; van der Linden et al., 2009; del Río-Espínola et al., 2010; Blanco-Kelly et al., 2010; Rodríguez-Rodríguez et al., 2010; Gandhi et al., 2010; Orozco et al., 2010; Wang et al., 2011; Nieters et al., 2011; Shuang et al., 2011; Tanizawa et al., 2011; Chen et al., 2012; Li et al., 2012; Yang et al., 2012; García-Bermúdez et al., 2012; Onouchi et al., 2012; Joo et al., 2013; Sokolova et al., 2013; Yun et al., 2014; İnal et al., 2015; Jiang et al., 2015; Wu et al., 2015; Chen et al., 2015; Panach et al., 2015; Field et al., 2015], and includes GWA and candidate gene studies of inflammatory diseases (but also of bone density, atherosclerosis, lymphoma, etc.), with bidirectional odds ratios. This bidirectionality (risk or protective factor) suggests that *CD40* signaling regulates

immunity in a complex manner, which might enhance or dampen immune and inflammatory processes based on the physiopathological context.

In DMD, there is a well-established role of T cells in the pathogenesis [Gussoni et al., 1994; Morrison et al., 2000; Morrison et al., 2005; Farini et al., 2007; Pinto-Mariz et al., 2010; Villalta et al., 2014; Cascabulho et al., 2012] and response to glucocorticoids [Kissel et al., 1991], both muscle fibers and immune cells being able to act as APCs and present antigens to T cells [Sugiura et al., 2000]. CD40 is upregulated in inflammatory muscle diseases, influencing chemokine production, and expressed on both infiltrating cells and muscle fibers [Sugiura et al., 2000], a mechanism which might also regulate secondary inflammation in muscular dystrophy. Functional studies are warranted to investigate specific molecular mechanisms underlying this association.

The only previous study identifying a modifier gene of a Mendelian disease phenotype with a genome-wide approach was the discovery of *DCTN4* variants as a risk factor for P. Aeruginosa infection in cystic fibrosis [Emond et al., 2012]. This study was based on exome sequencing of extreme phenotypes (early vs. late or no infection) in 91 patients, focused on moderately rare and rare variants (MAF < 12.5%), adopted a gene-wise burden test, and also relied on subsequent validation. We illustrate an alternative approach, which takes into account that even common functional variants may have strong modifier effects in Mendelian diseases, when the affected gene is pathologically upregulated. These different approaches may prove complementary in explaining variable expressivity of rare monogenic diseases.

In conclusion, we report a novel modifier locus of the rare disease DMD, identified by an initial underpowered GWAS focused on common functional variants, hypothesis-based prioritization of GWAS p-values, and validation in independent cohorts, in an international collaborative effort which collectively represents the largest DMD genetic association study so far. Reduced CD40-mediated cell-cell signaling in carriers of the minor rs1883832 allele might precipitate failed regeneration and fibrosis, and this pathway represents a novel therapeutic target in DMD. This study represents a

paradigm for the investigation of common functional variants as modifiers of rare monogenic diseases.

Aim 5: Genotype-phenotype correlations in BMD

Background

Becker muscular dystrophy (BMD) is an X-linked neuromuscular disorder caused by non-truncating mutation in the *DMD* gene, leading to quantitatively and/or qualitatively altered dystrophin protein being expressed in skeletal and cardiac muscle [Hoffman et al., 1989]. Most frequently, causative mutations are single- or multi-exon deletions (approximately 70%), while single or multi-exon duplications and small (sub-exonic) mutations represent approximately 15% each [Darras et al., 2000].

The typical clinical picture of BMD comprises proximal muscle wasting and weakness, presenting in adolescents or young adults, more evident at the thigh extensors and pelvic girdle than the shoulder girdle; pronounced calf hypertrophy; possible loss of motor functions, such as rising from the floor or walking independently, occurring years or decades after onset; and frequent dilated cardiomyopathy, with no obvious correlation with the severity of skeletal muscle involvement [Darras et al., 2000]. BMD can be distinguished from the more severe allelic disorder, Duchenne muscular dystrophy (DMD), which is characterized at the molecular level by truncating mutations and abolished dystrophin expression [Hoffman et al., 1987]. A classic clinical definition classifies as BMD those patients who remain ambulatory after the age of 16 years, while DMD patients lose ambulation by the age of 13 [Darras et al., 2000]. However, this general principle is challenged by the existence of intermediate forms (IMD) and by the improvement of standards of care in DMD, so that the distinction between severe cases of BMD and “outlier” DMD/IMD patients may be blurred. Conversely, at the mild end of the BMD spectrum are patients who present with muscle hypertrophy (prominently of the calves) and/or elevated creatin kinase (CK) levels, but virtually no muscle wasting or weakness [Angelini et al., 1994a].

Genotype-phenotype correlation studies have shown that loss of functionally crucial actin- or dystroglycan-binding domains, respectively at the N- or C-terminus, can result in DMD-like phenotypes despite detectable protein [Hoffman et al., 1991; Aartsma-Rus et al., 2006; Kesari et al., 2008]. Deletions in the large dystrophin rod domain, which harbours the majority of BMD causing mutations, might differently affect the physical properties of resulting internally deleted dystrophin, depending on the preservation or loss of structural "phase" between spectrin repeats and hinge regions [Kaspar et al., 2009]. Some specific deletions, such as deletions of in-frame exons in the proximal rod domain [Angelini et al., 1994a], and deletions including the hinge 3 domain encoded by exons 50 and 51 [Carsana et al., 2005; Anthony et al., 2011], have been more frequently associated to mild or asymptomatic cases; while frequently observed deletions situated in the DMD mutational hotspot around exon 45-53 [White and Den Dunnen, 2006], but not including exons 50-51, have been linked with a typical clinical picture of BMD [Bushby et al., 1993; Anthony et al., 2014a; van den Bergen et al., 2014b]. Moreover, a quantitative correlation between dystrophin content in muscle and BMD severity has been described, both as a linear or threshold effect [Angelini et al., 1994a; Angelini et al., 1996; Comi et al., 1994; Anthony et al., 2011; van den Bergen et al., 2014b]. Recently, renewed interest has been kindled in this field, as some BMD-causing deletions can be regarded as naturally occurring models of the in-frame deletions produced at the transcript level by splice-modulating antisense oligonucleotides (AONs), employed in the treatment of DMD with the exon skipping approach [Aartsma-Rus, 2012; Arechevala-Gomez et al., 2012].

While establishing and refining genotype-phenotype correlations, and describing natural history is relevant for prognosis and counseling, the longitudinal description of validated, clinically meaningful outcome measures is needed for the optimal design of clinical trials for upcoming therapeutic interventions. While this has been the object of increasing attention in DMD [Mazzone et al. 2011; Mazzone et al., 2013; Henricson et al., 2013; McDonald et al., 2013; Pane et al., 2014a], there is scarce data describing the application of standardized functional measures in BMD, with a cross-sectional or

longitudinal approach. We mutuated outcome measures currently established in natural history studies and clinical trials for ambulatory DMD patients: the Six Minute Walk Test (6MWT), measuring the walking distance a patient is able to cover in 6 minutes [McDonald et al., 2010], the North Star Ambulatory Assessment (NSAA), a scale scoring the ability to perform motor tasks of everyday life with or without compensation [Mayhew et al., 2011; Scott et al., 2012], and three timed function tests (TFTs): run/walk 10 m, rise from the floor, climb four standard steps. These were evaluated at baseline and after 1 year in a population of BMD patients referring to our Center.

Methods

Informed consent. The study was approved by the local Ethics committee, and all patients or their legal guardians consented to study procedures.

Inclusion criteria. We selected male patients with a diagnosis of BMD based on the following criteria: muscle tissue Western Blot (WB) or immunofluorescence (IF) showing reduced dystrophin relative to control, in the presence of a *DMD* gene mutation; or altered molecular weight dystrophin identified by WB; or evidence of an in-frame *DMD* gene mutation.

Dystrophin quantification. Dystrophin Western blot was performed following SDS-PAGE, according to methods described elsewhere [Angelini et al., 1996], using a monoclonal antibody against the C-terminus and visualization on X-ray films by ECL-chemioluminescence method (Amersham). Control and patient samples were loaded in adjacent lanes to determine both molecular weight and relative abundance. Dystrophin quantity was determined by densitometric analysis of dystrophin bands, normalized to myosin bands in the post-transfer Coomassie blue stained gels, with subtraction of background.

Functional measures. 6MWT, NSAA, and TFTs (run or walk 10 m, rise from the floor starting in a straight-leg sitting position, climb 4 standard stars) were performed by trained neuromuscular physicians in compliance with validated protocols [Mazzone et al., 2009; Mazzone et al., 2010; McDonald et al., 2010; Mayhew et al., 2011], at baseline and after 12 ± 1 months. For the purposes of cross-sectional analyses (descriptive statistics, correlations) of baseline functional measures, patients who were unable to perform tasks because of loss of function (e.g. 6MWT and run/walk velocity in non-ambulatory) were scored as zero. For the purposes of longitudinal analyses of functional changes after 1 year, only patients who were able to perform tasks at baseline were included.

Disease milestones. Loss of independent ambulation (LoA) was defined as continuous wheelchair use, and loss of the ability to run (LoR) was defined as the inability to accelerate significantly from normal gait and lift both feet off the ground. The times at loss of ability to climb stairs and rise from the ground was also noted in the patient history, but were not analyzed because many patients were unable to date these events with precision.

Grouping of patients for statistical analyses. We grouped patients with *DMD* deletions predicted to result at the transcript level from exon skipping with antisense oligonucleotides currently being experimented in DMD clinical trials. These included the deletions of exons 45-47, 45-48, 45-49, and 45-55, hereafter called collectively as "del 45-x", as models of exon 45 skipping (28 patients), and the deletions of exons 34-51, 45-51, 48-51, and 50-51, hereafter called collectively "del x-51", as models of exon 51 skipping (10 patients). While the deletion of exons 45-51 might result at the transcript level from the skipping of both exons 45 and 51, we included this deletion in the "del x-51" group because the 45-50 deletion is a more common cause of DMD than the 46-51 deletion [White and den Dunnen, 2006]. Furthermore, as the isolated deletion of exon 48 was the most frequent mutation not included in the groups described above (10 patients), we considered it as a separate group for statistical analyses ("del 48"). This resulted in four mutation groups: "del 45-x", "del 48", "del x-51", and "other". For graphical representation and summary statistics of functional measures associated to different levels of dystrophin quantity, we grouped patients with available WB dystrophin quantification into 3 groups, showing 0-33%, 34-66%, and 67-100% dystrophin quantity relative to control.

Statistical analyses. Differences in age between mutation groups were tested by Student's t test, while dystrophin levels were compared by Mann-Whitney U test. Time-to-event analyses of loss of ambulation and ability to run were performed, with loss of function as event, age as time variable, and censoring patients able to walk/run at last follow-up. Age at loss of function was compared between groups by log-rank test, and

hazard ratios (HRs) associated to WB dystrophin levels were estimated by Cox regression. Baseline correlations between dystrophin quantity and functional measures, between age and functional measures, and between different functional measures were assessed by Spearman's rank correlation coefficient. Baseline functional differences between mutation groups were tested by Kruskal-Wallis rank sum test. Significance of functional changes after 1 year in the whole cohort and individual mutation groups was tested by Friedman rank sum test. Correlation between dystrophin quantity and functional change at 1 year was tested by Spearman's rank correlation coefficient. Statistical significance was set at $p < 0.05$. Statistical analyses were performed with R version 3.2.1. Power calculations were performed with PS version 3.0.43 [Dupont and Plummer, 1990], using the paired design t-test option.

Results

Patients. We recruited 69 BMD patients aged 6 to 69 years (mean ± standard deviation [SD] 33.2 ± 16.5, Table I), from 61 unrelated families. Familial cases included one family with three affected brothers, four families with two affected brothers, one family with affected maternal uncle and nephew, and one with affected maternal grandfather and grandson.

Mutations. The majority of patients (56/69, 81%) harboured a deletion of one or more *DMD* exons. Ordered by frequency, these were: del 45-48, 14 patients; del 48, 10 patients; del 45-47, 10 patients; del 45-51, 5 patients; del 48-49, 3 patients; del 45-49, del 45-55, del 48-51, del 50-51, del 10-25, 2 patients each; del 3-9, del 10-29, del 11-30, del 34-51, 1 patient each. Two patients (3%) harboured a duplication (dup 13-42 and dup 19-41). Sub-exonic in-frame microdeletions were identified in one patient (c.10099_10101delGAA, p.Glu3386del) and 3 brothers (c.676_678delAAG, p.Lys226del). Two brothers had a frame-shifting sub-exonic microdeletion (c.10587_10588delAG, p.Lys3505AlaFsX8) in exon 74, and a muscle biopsy performed in one of the borthers showed reduced dystrophin immunolabeling with the “DYS1” antibody directed against the rod domain of dystrophin. One patient had a missense mutation (c.478A>C, p.Thr160Pro), and two brothers had a synonymous mutation (c.4299G>T, p.Gly1433Gly) predicted in silico to disrupt an exon splicing enhancer (ESE) site [Desmet et al., 2009]. Splicing out of exon 30 was not confirmed at the RNA level, but the remaining gDNA sequence was normal and WB showed low amounts (10-30%) of dystrophin with slightly reduced molecular weight on WB. The c.4299G>T synonymous variant was not found in the dbSNP, 1000 Genomes, and ExAC databases. Two patients harboured nonsense mutations: c.4980G>A, p.Trp1660* in exon 35, with rescue of 29% dystrophin with slightly reduced molecular weight at WB; and c.3843G>A, p.Trp1281* in exon 28, with rescue of 17% dystrophin with slightly reduced molecular weight. After grouping patients by mutation as explained in Methods, there were 28 patients in the “del 45-x” group, 10 in the “del 48” group, 10 in the “del x-51” group, while the remaining 21 were

grouped as “other”. Average age in the “del 48” and “del x-51” groups was younger than in the “del 45-x” group by 13.0 and 13.3 years respectively (p = 0.08 and 0.02,). Average age by mutation group is summarized in Table I.

Table I. Age and dystrophin quantity by *DMD* mutation.

Mutation group	Individual mutation	Age (years)			Dystrophin (% of control)		
		n	Mean ± SD	median (range)	n	Mean ± SD	median (range)
“del 45-x”	del 45-47	10	38.1 ± 13.0	40.1 (10.7 ~ 55.9)	5	39 ± 27	28 (10 ~ 75)
	del 45-48	14	38.5 ± 17.2	38.0 (9.2 ~ 69.6)	12	57 ± 28	49 (16 ~ 100)
	del 45-49	2	32.1 ± 17.9	32.1 (19.4 ~ 44.7)	1	28 ± NA	28 (28 ~ 28)
	del 45-55	2	29.3 ± 31.7	29.3 (6.9 ~ 51.6)	2	90 ± 14	90 (80 ~ 100)
	"del 45-x" total	28	37.2 ± 16.0	38.3 (6.9 ~ 69.6)	20	54 ± 29	49 (10 ~ 100)
“del 48”	del 48	10	24.2 ± 19.5	14.5 (6.1 ~ 65.8)	7	67 ± 21	58 (40 ~ 100)
“del x-51”	del 34-51	1	50.7 ± NA	NA	1	70 ± NA	NA
	del 45-51	5	16.8 ± 6.8	16.2 (8.6 ~ 25.7)	4	90 ± 12.5	93 (74 ~ 100)
	del 48-51	2	37.8 ± 7.2	37.8 (32.7 ~ 42.9)	2	64 ± 24	64 (47 ~ 81)
	del 50-51	2	14.3 ± 2.1	14.3 (12.9 ~ 15.8)	2	82 ± 18	82 (69 ~ 95)
	"del x-51" total	10	23.9 ± 14.1	18.7 (8.6 ~ 50.7)	9	80 ± 17	81 (47 ~ 100)
“other”	del 3-9	1	13.1 ± NA	NA	1	100 ± NA	NA
	rod domain del*	4	33.7 ± 14.2	29 (22.6 ~ 54.2)	4	100 ± 0	100 (100 ~ 100)
	del 48-49	3	50.2 ± 13.1	54.7 (35.4 ~ 60.3)	3	37 ± 38	30 (3 ~ 78)
	duplications	2	39.9 ± 6.9	39.9 (35 ~ 44.7)	2	30 ± 35	30 (5 ~ 55)
	nonsense	2	28.6 ± 4.8	28.6 (25.2 ~ 32)	2	23 ± 8	23 (17 ~ 29)
	missense	1	37.6 ± NA	NA	1	13 ± NA	NA
	small deletions	6	35.4 ± 18.1	34.9 (17.4 ~ 54.1)	1	39 ± NA	39 (39 ~ 39)
	synonym	2	40.7 ± 8	40.7 (35 ~ 46.3)	2	23 ± 4	23 (20 ~ 26)
	"other" total	21	36.5 ± 14.2	35.0 (16.1 ~ 50.3)	16	51 ± 39	34.5 (3 ~ 100)
Total		69	33.2 ± 16.5	34.9 (6.1~69.6)	52	59 ± 31	57 (3 ~ 100)

SD: standard deviation. **del**: deletion. ***** Including deletions of exons 10-25, 10-29, del 11-30.

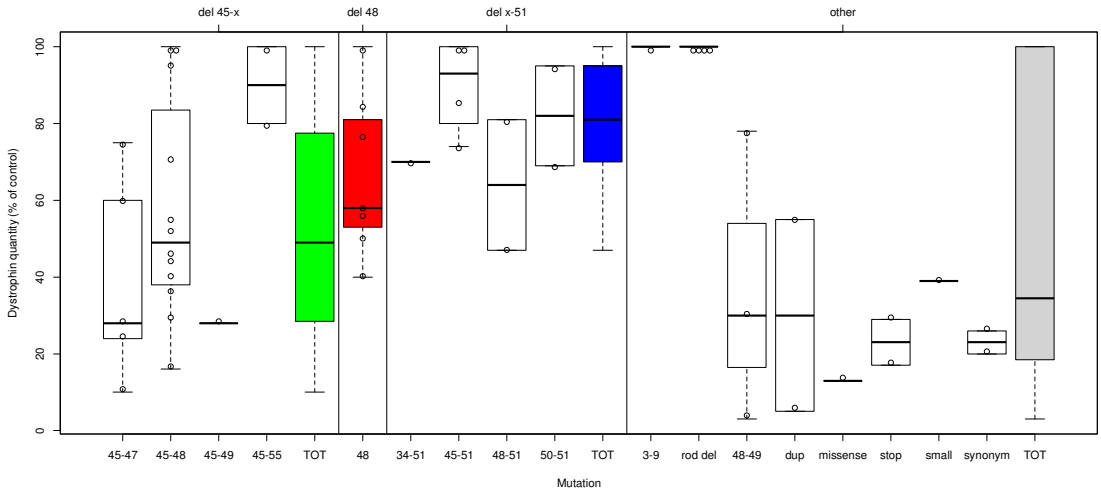

Figure 1. Dystrophin quantity by mutation. Box plot of quantity of dystrophin, expressed as % of control assessed by Western Blot, associated to different kinds of DMD mutations. Data for each mutation group ("del 45-x", "del 48", "del x-51", "other") is presented in a separate rectangle. Data for each individual mutation is represented as empty box plots with circles for individual data points. The rightmost, coloured boxplot in each rectangle represents the total for each mutation group. Axis labels: numbers represent deletions of corresponding exons; TOT: total; rod del: rod domain deletions including deletions of exons 10-25, 10-29, and 11-30; dup: duplications of exons 13-42 and 19-41; stop: nonsense mutations; small: sub-exonic microdeletion.

Dystrophin quantity. Average dystrophin quantity was 59.5 ± 31.3% relative to control. There were differences in dystrophin levels between mutation groups (Figure 1, Table I). In particular, dystrophin levels in the "del x-51" group were higher than in the "del 45-x" group (80.3 ± 17.4% vs. 54.5 ± 29.5%, n = 9 and 20 respectively, p = 0.03). Dystrophin levels associated with exon 48 deletion were in between these two groups (66.6 ± 21.4, n = 7). Mutations showing dystrophin levels similar to control included deletion of exons 45-55 and rod domain deletions (del 10-25, 10-29, and 11-30).

Loss of independent ambulation. At baseline, three/69 patients were using a wheelchair full time: one 60 year old patient carrying the deletion of exon 45-48 (LoA at 59 years); one 47 year old patient carrying the deletion of exon 45-47 (LoA at 41 years); and one 18 year old patient carrying an in-frame microdeletion (LoA at 17 years). Two additional patients lost ambulation during the study: a 22 year old patient carrying a distal out-of-frame microdeletion (with reduced dystrophin on IF); and a 38 year old patient carrying a missense mutation. All LoA events were observed in the "del 45-x" or "other" mutation groups, and none in the "del x-51" or "del 48" groups, although this difference was not statistically significant (log-rank test) due to the small number of events, which also made it impossible to estimate median ages at LoA. Cox regression did not show a significant association of earlier LoA with lower dystrophin levels, although this analysis was scarcely powered, with only 3 patients becoming non-ambulatory and having available dystrophin quantification data. Kaplan-Meier plots of walking ability by age are shown in Figure 2A-B.

Loss of the ability to run. Similar to LoA, all patients who lost the ability to run were in the "del 45-x" (14/28, 50%) or "other" (11/21, 52%) mutation groups. Nine of these could not date the LoR event at a precise age, and were excluded from time-to-event analyses. Median age at LoR in the whole cohort was 31 years (95% confidence interval [CI] 26 ~ ∞), and earlier in the "del 45-x" group (median 26 years, 95% CI 17 ~ 27). The difference in age at LoR between mutation groups was significant (log-rank $p < 0.001$ with 3 degrees of freedom). Cox regression showed a significant association of lower

dystrophin quantity with earlier LoR (HR 0.98, 95% CI 0.90 ~ 0.99, p = 0.03). Kaplan-Meier plots of running ability by age are shown in Figure 2C-D.

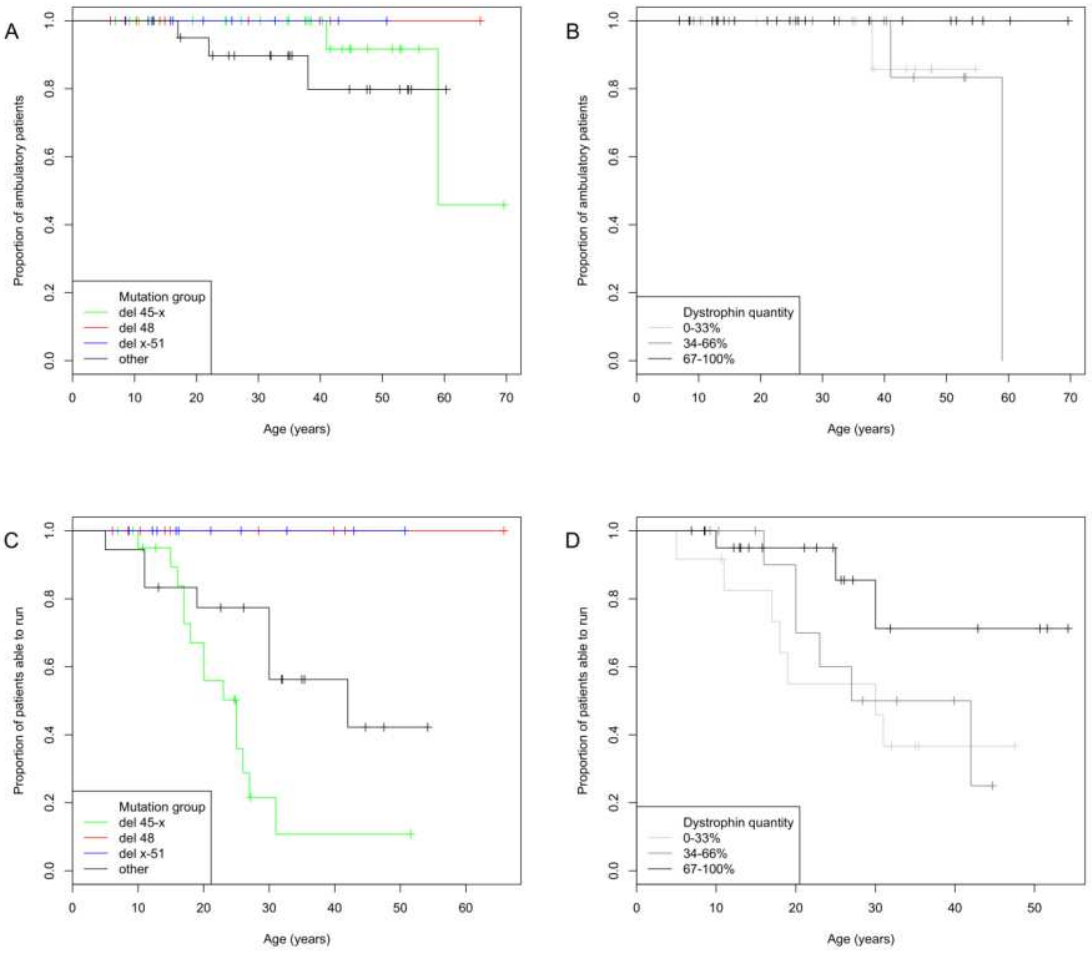

Figure 2. Kaplan-Meier plots of loss of ability to walk and run. The proportion of patients able to walk at increasing ages are represented, grouped by A) mutation group ("del 45-x", "del 48", "del x-51", and "other"), and B) dystrophin quantity (0-33%, 34-66%, 67-100%). The proportion of patients able to run at increasing ages are represented by C) mutation group, and D) dystrophin quantity.

Baseline functional measures. Results of baseline functional evaluations are presented in the following paragraphs and summarized in Table II.

Table II. Baseline functional measures overall and by mutation.

Baseline functional measure	Mutation group	n	Mean value ± SD	Median value (range)	Comparison* between mutation groups
6MWT distance (m)	del 45-x	27	347 ± 167	372 (0 ~ 615)	6MWT differs significantly between mutation groups (p = 0.02)
	del 48	10	459 ± 121	482 (183 ~ 656)	
	del x-51	10	497 ± 100	474 (386 ~ 721)	
	other	21	365 ± 147	413 (0 ~ 595)	
	All BMD	68	391 ± 155	425 (0 ~ 721)	
NSAA score	del 45-x	27	20.9 ± 11.1	20 (2 ~ 34)	NSAA differs significantly between mutation groups (p < 0.001)
	del_48	10	33.9 ± 0.32	34 (33 ~ 34)	
	del_x-51	10	33.7 ± 0.95	34 (31 ~ 34)	
	other	21	23.0 ± 11.2	24 (2 ~ 34)	
	All BMD	68	25.3 ± 10.8	32.5 (2 ~ 34)	
Run/walk 10 m velocity (m/s)	del 45-x	27	1.55 ± 1.05	1.25 (0 ~ 3.33)	Run/walk velocity differs significantly between mutation groups (p < 0.001)
	del 48	9	3.52 ± 1.01	3.33 (1.67 ~ 5)	
	del x-51	10	3.95 ± 1.19	4.17 (2 ~ 5)	
	other	21	1.99 ± 1.41	1.67 (0 ~ 5)	
	All BMD	67	2.31 ± 1.49	3.00 (0 ~ 5)	
Rise from floor velocity (s^ ~ 1)	del 45-x	28	0.19 ± 0.19	0.17 (0 ~ 0.5)	Rise from floor velocity differs significantly between mutation groups (p < 0.001)
	del 48	9	0.52 ± 0.2	0.5 (0.33 ~ 1)	
	del x-51	10	0.62 ± 0.28	0.5 (0.33 ~ 1)	
	other	21	0.27 ± 0.27	0.17 (0 ~ 1)	
	All BMD	68	0.32 ± 0.28	0.33 (0 ~ 1)	
Climb stairs velocity (steps/s)	del 45-x	28	1.20 ± 1.21	0.8 (0 ~ 4)	Cimb stairs velocity differs significantly between mutation groups (p < 0.001)
	del 48	9	2.96 ± 1.25	4 (1.33 ~ 4)	
	del x-51	10	3.13 ± 1.14	4 (1.33 ~ 4)	
	other	21	1.57 ± 1.38	1.33 (0 ~ 4)	
	All BMD	68	1.83 ± 1.46	1.33 (0.00 ~ 4.00)	

6MWT: 6 Minute Walk Test. **NSAA**: North Star Ambulatory Assessment. ***** Kruskal-Wallis rank sum test for differences between mutation groups.

Baseline 6MWT. One patient refused to perform the 6MWT because of recent trauma, and four were assigned a “zero meter” distance for the purpose of cross-sectional baseline analyses (e.g. comparisons within mutation groups). Of these, 3 where using a wheelchair continuously and one could not walk more than a few steps with a wheeled walking frame. Average distance covered in the 6MWT was 391 ± 155 m (414 ± 123 m excluding the four “zero” values). No patient experienced dyspnea, dizziness, palpitations, arrhythmia, or other cardiac or respiratory symptoms during the test. There were no falls and no test interruptions.

Baseline NSAA. NSAA was scored in 68/69 patients (one could not be scored in all items because of recent trauma) and averaged 25.3 ± 10.8, with a range of 2-34 and a heavily right-skewed distribution (median 32.5), due to a “ceiling effect”: 28 patients with substantially normal muscle function scored the maximum (34 points). On the other hand, the minimum observed score of 2 was due all patients having preserved anti-gravity strength of the neck flexor muscles, thus scoring 2 points in the “lift head” item despite inability to stand or perform postural transfers independently.

Baseline TFTs. One patient refused to perform the 10 m walk/run because of recent trauma, and one could not perform TFTs because of a logistic issue. Due to disease progression, 4/67 patients (6%) were unable to walk 10 m, 18/68 (26%) were unable to rise from the floor, and 9/68 (13%) were unable to climb 4 standard steps. These patients were assigned a “zero” velocity in the corresponding activities for the purpose of cross-sectional comparisons. Average baseline velocities were as follows: 10 m run/walk 2.31 ± 1.49 m/s, rise from floor 0.32 ± 0.28 s-1, climb 4 standard steps 1.83 ± 1.46 steps/s. Excluding “zero” velocity values, baseline velocities were 2.45 ± 1.42 m/s, 0.44 ± 0.23 s-1, and 2.11 ± 1.36 steps/s respectively.

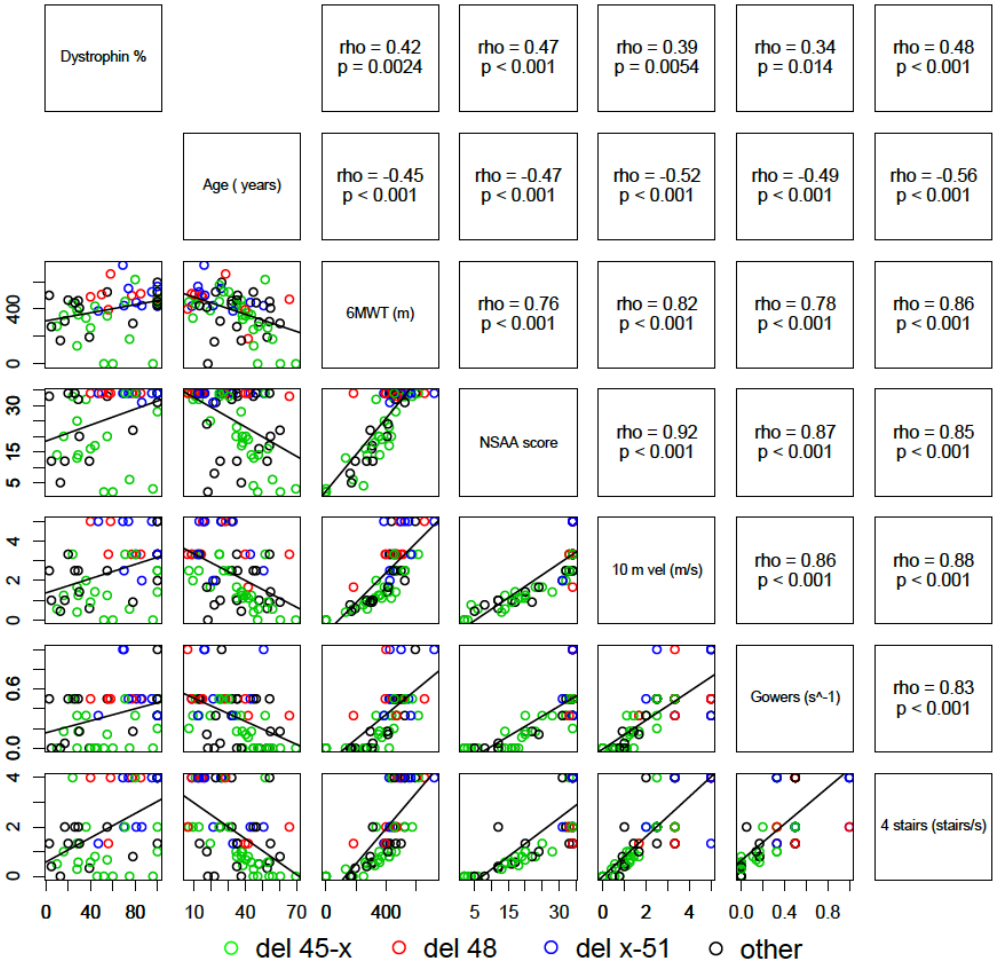

Figure 3. Correlation matrix of dystrophin quantity, age, and functional measures. Panels in the diagonal indicate parameters represented on corresponding columns and rows. Upper panels show correlation parameters (Spearman's ρ and corresponding p-value) between parameters on corresponding the row and column, while lower panels show scatter plots, with data points color-coded for mutation group, and regression lines.

Correlations between baseline functional measures, dystrophin quantity, and age. All functional measures (6MWT, NSAA, and TFTs) were positively correlated with dystrophin quantity assessed by WB, and negatively correlated with age. Furthermore, all functional measures were positively correlated within each other. Scatter plots showing these correlations, and the corresponding parameters (Spearman's ρ, p value), are shown in Figure 3. Due to a NSAA "ceiling effect" in milder mutation groups ("del 48" and "del x-51"), correlations between NSAA and other parameters appeared more linear in the other, more severe groups ("del 45-x", "other"), although still significant in the overall BMD cohort.

Baseline functional differences between mutation groups. For all functional measures at baseline, there were significant differences between mutation groups (Table II). Compared to the "del 45-x" group, patients in the "del 48" and "del x-51" mutation groups walked respectively 112 m and 150 m farther in the 6MWT. Furthermore, while NSAA scores showed an approximately nomal distribution in the "del 45-x group", with a mean ± SD of 20.9 ± 11.1 and median of 20, patients in the "del 48" and "del x-51" groups almost always (90% in both groups) scored a maximum 34/34 in the NSAA, hence the "ceiling effect" observed with this measure, and registered higher velocities in all timed items.

Functional changes after 1 year. Fifty-four patients completed the 6MWT at baseline and after 1 year, showing stability of this measure: average change was 3 ± 66 m. Conversely, there was a significant decrease of -0.9 ± 1.6 NSAA score points in 57 patients who completed this evaluation after 1 year ($p < 0.001$). TFTs were stable after 1 year, with non-significant changes of -0.03 ± 0.58 m/s in run/walk velocity, 0.06 ± 0.25 s-1 in rise from floor velocity, and -0.13 ± 0.6 steps/s in climb stairs velocity. Two patients who lost ambulation during follow-up (see paragraph about ambulatory status and disease milestones) were assigned "zero" values for 6MWT distance and 10 m walk velocity at the 1-year evaluation. At baseline, their 6MWT distances were 161 and 170 m, and their 10 m walk speeds were 0.77 and 0.45 m/s. No patient lost the ability to rise

from the floor during follow up, while two (47 year old patient carrying a deletion of exons 45-47, and a 53 year old patient carrying a small sub-exonic deletion) lost the ability to climb 4 standard steps after 1 year. Grouping of functional changes by mutation groups (Figure 4) showed that patients in the “del 45-x” group (n = 27 with complete longitudinal data) “lost” -12 ± 31 m in average 6MWT distance after 1 year (p = 0.059). The same patients presented a significant within-group decrease of -1.3 ± 1.7 NSAA score points at 1 year (p = 0.001). We did not observe any significant velocity changes in any of the TFTs after 1 year. Functional changes at 1 year are summarized by mutation group in Table III. There was a trend towards a correlation between dystrophin quantity and 1-year changes of 6MWT ($\rho = 0.3$, $p = 0.055$) and NSAA ($\rho = 0.025$, $p = 0.09$), while no correlation between dystrophin quantity and timed item changes was observed (Figure 5). Functional changes at 1 year by dystrophin quantity classes are summarized in Table IV.

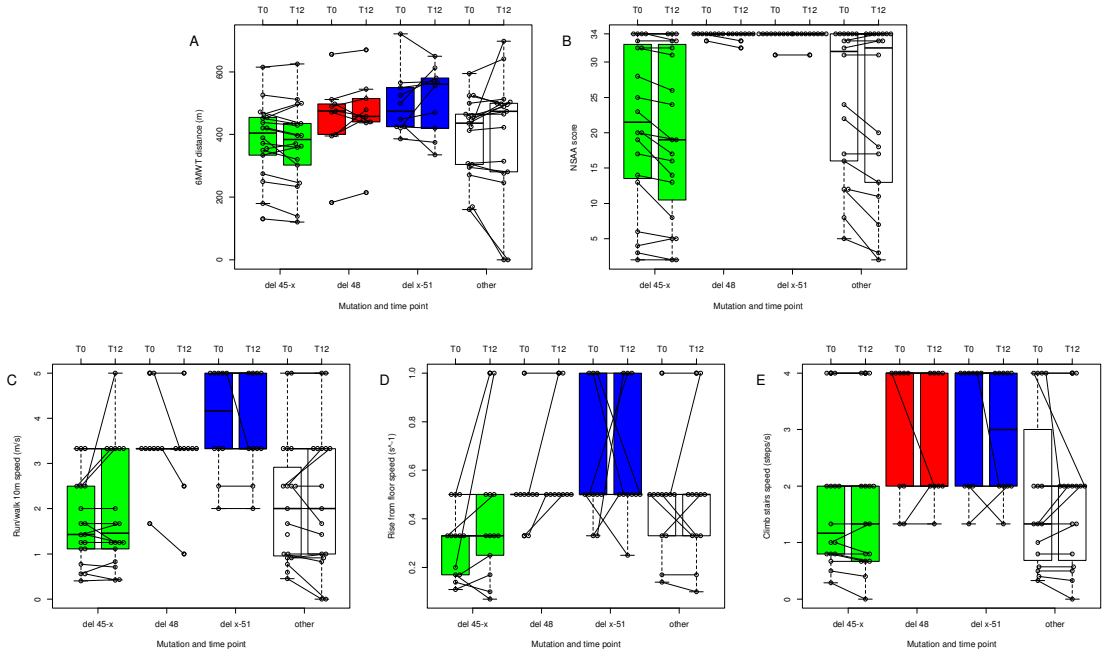

Figure 4. Functional changes after 1 year in different BMD mutation groups. Box plots showing baseline and 1-year values of A) 6MWT distance, B) NSAA score, C) 10 m run/walk velocity, D) rise from floor velocity, and E) climb 4 standard steps velocity. Boxes are color-coded for mutation group, and trajectories of each individual patient are illustrated by dots connected by segments.

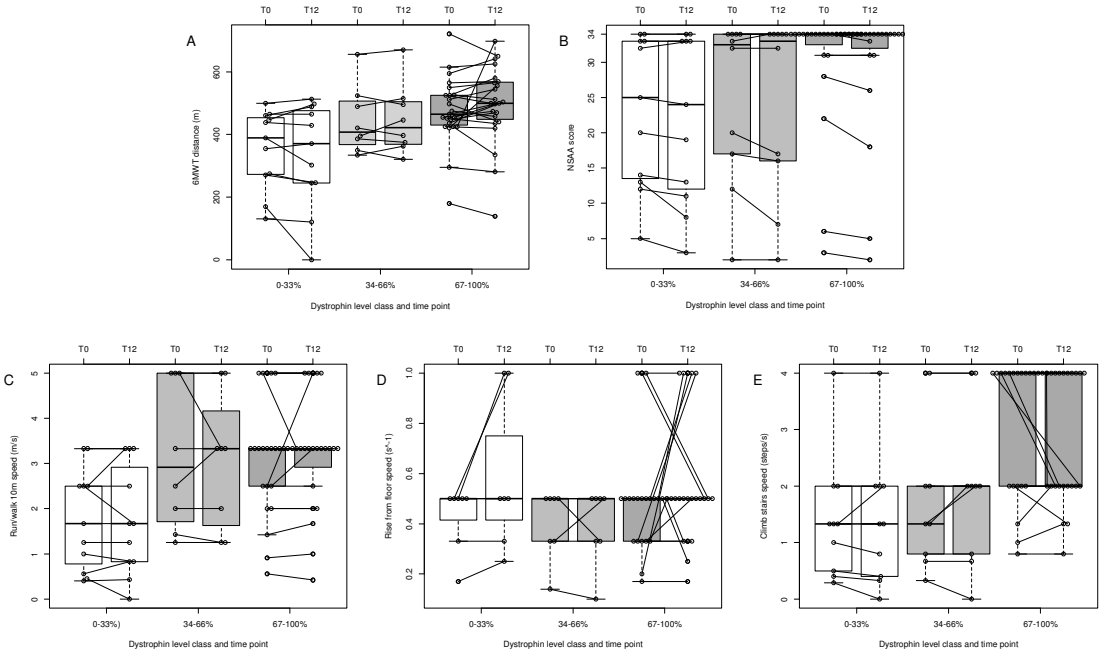

Figure 5. Functional changes after 1 year in different dystrophin quantity level groups. Box plots showing baseline and 1-year values of A) 6MWT distance, B) NSAA score, C) 10 m run/walk velocity, D) rise from floor velocity, and E) climb 4 standard steps velocity. Boxes are color-coded for dystrophin levels of 0-33%, 34-66%, and 67-100% relative to control, and trajectories of each individual patient are illustrated by dots connected by segments.

Table III. Functional changes after 1 year, overall and by mutation.

Functional measure change at 1 year	Mutation class	n	Mean change ± SD	Median change (range)
6MWT distance (m)	del 45-x	27	-12 ± 31[§]	-15 (-87 ~ 45)
	del 48	9	14 ± 43.2	79 (-58 ~ 79)
	del x-51	10	17 ± 71	135 (-90 ~ 135)
	other	17	4 ± 95	7 (-170 ~ 273)
	All BMD	54	3 ± 66	-2.5 (-170 ~ 273)
NSAA score	del 45-x	20	-1.3 ± 1.7**	-1 (-5 ~ 1)
	del_48	9	-0.3 ± 0.5	0 (-1 ~ 0)
	del_x-51	10	0 ± 0	0 (0 ~ 0)
	other	18	-1.3 ± 2.2	0 (-6 ~ 1)
	All BMD	57	-0.9 ± 1.6***	0 (-6 ~ 1)
Run/walk 10 m velocity (m/s)	del 45-x	18	0.22 ± 0.65	0 (-0.42 ~ 2.5)
	del 48	9	-0.36 ± 0.6	0 (-1.67 ~ 0)
	del x-51	10	-0.17 ± 0.53	0 (-1.67 ~ 0)
	other	17	-0.04 ± 0.43	0 (-0.83 ~ 0.83)
	All BMD	54	-0.03 ± 0.58	0 (-1.67 ~ 2.5)
Rise from floor velocity (s^ ~ 1)	del 45-x	13	0.11 ± 0.25	0 (-0.1 ~ 0.8)
	del 48	9	0.1 ± 0.17	0.5 (0 ~ 0.5)
	del x-51	10	0.01 ± 0.38	0.67 (-0.5 ~ 0.67)
	other	12	0.024 ± 0.17	0 (-0.17 ~ 0.5)
	All BMD	44	0.06 ± 0.25	0 (-0.5 ~ 0.8)
Climb stairs velocity (steps/s)	del 45-x	16	-0.03 ± 0.16	0 (-0.29 ~ 0.33)
	del 48	9	-0.15 ± 0.73	0 (-2 ~ 0.67)
	del x-51	10	-0.2 ± 0.71	0 (-2 ~ 0.67)
	other	16	-0.17 ± 0.76	0 (-2 ~ 0.67)
	All BMD	51	-0.13 ± 0.6	0 (-2 ~ 0.67)

6MWT: 6 Minute Walk Test. **NSAA**: Norths Star Ambulatory Assessment. **§** Functional change close to statistical significance (p = 0.059). ****** Statistically significant functional change (p = 0.001). *******Statistically significant functional change (p < 0.001).

Table IV. Functional changes after 1 year, by dystrophin quantity.

Functional change at 1 year	Dystrophin (% of control)	n	Mean change ± SD	Median change (range)	Correlation between dystrophin % and functional change
6MWT distance (m)	0-33	11	-21 ± 62	-9 (-170 ~ 52)	ρ = 0.3, p = 0.055
	34-66	8	4 ± 31	1.5 (-29 ~ 51)	
	67-100	23	18 ± 75	10 (-90 ~ 273)	
NSAA score	0-33	11	-0.91 ± 1.58	-1 (-5 ~ 1)	ρ = 0.025, p = 0.09
	34-66	10	-0.8 ± 1.82	0 (-5 ~ 1)	
	67-100	24	-0.38 ± 0.93	0 (-4 ~ 0)	
Run/walk 10 m velocity (m/s)	0-33	11	-0.03 ± 0.41	0 (-0.83 ~ 0.83)	ρ = 0.22, p = n.s.
	34-66	8	-0.13 ± 0.7	0 (-1.67 ~ 0.83)	
	67-100	23	0.09 ± 0.67	0 (-1.67 ~ 2.5)	
Rise from floor velocity (s^-1)	0-33	7	0.16 ± 0.24	0 (0 ~ 0.5)	ρ = 0.07, p = n.s.
	34-66	7	-0.01 ± 0.1	0 (-0.17 ~ 0.17)	
	67-100	22	0.07 ± 0.32	0 (-0.5 ~ 0.8)	
Climb stairs velocity (steps/s)	0-33	10	0 ± 0.26	0 (-0.29 ~ 0.67)	ρ = 0.06, p = n.s.
	34-66	9	0.11 ± 0.33	0 (-0.33 ~ 0.67)	
	67-100	22	-0.35 ± 0.83	0 (-2 ~ 0.67)	

6MWT: 6 Minute Walk Test. **NSAA**: North Star Ambulatory Assessment.

Power calculation for a hypothetical BMD clinical trial. As the NSAA was able to detect a significant functional change in one year, we could perform a power calculation in order to answer the question: how many BMD patients per study arm would be needed, in order to show the effectiveness of a 1-year course of treatment in slowing NSAA decrease? As the NSAA decrease was mostly due to patients in the “del 45-x” group (typical BMD phenotype), and was significant in this group in itself, the power calculation is based on this population alone, which showed a mean ± SD NSAA change of -1.3 ± 1.7. Assuming a type I error $\alpha = 0.05$ and power $(1 - \beta) = 0.8$, it would take approximately 15 patients per study arm to identify the effect of an intervention able to arrest disease progression (1.3 point difference in NSAA change). This figure would correspond to approximately 20, 40, and 65 patients per arm for NSAA change

differences of 1, 0.75, and 0.5 respectively (Figure 6). Inclusion criteria in this hypothetical trial should be limited to patients with “typical” BMD, based on mutation or baseline functional status, as it is clear from longitudinal NSAA data (Table III, Figure 5) in the “del x-51” and “del 48” group that patients with mild BMD have a “ceiling” effect with NSAA and stable scores after 1 year.

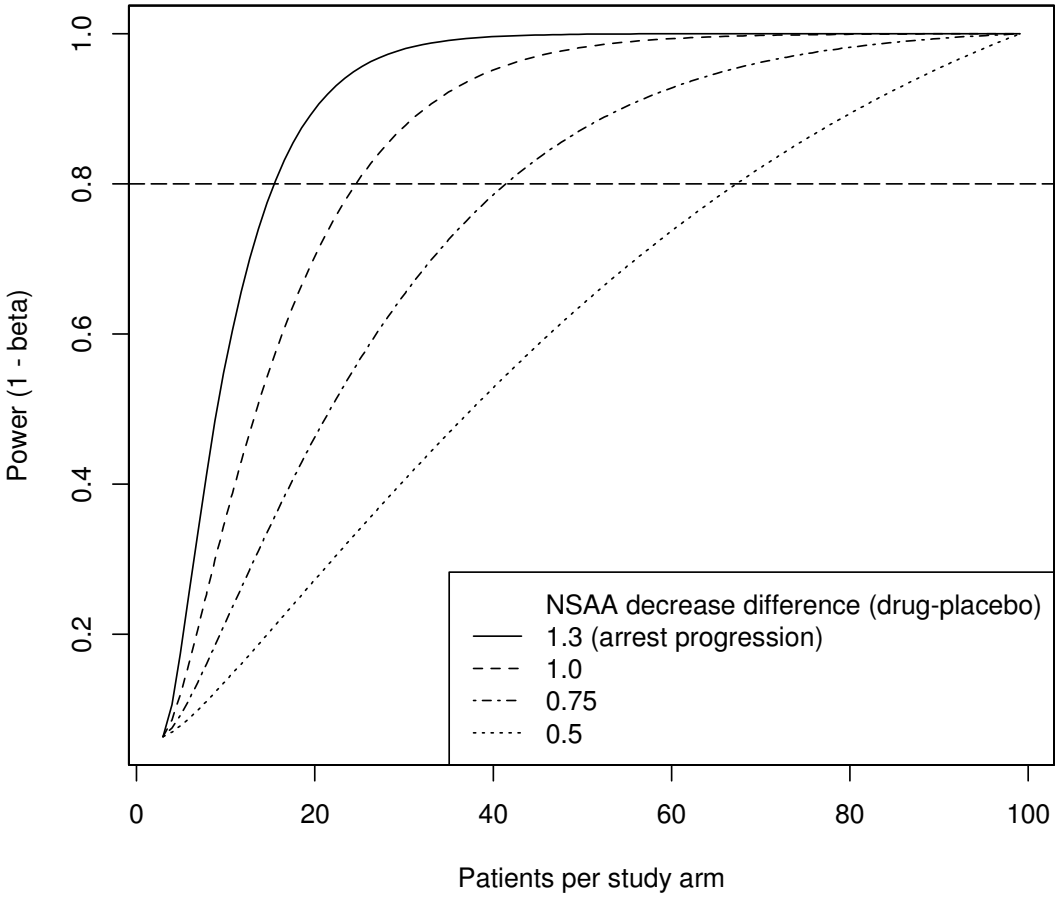

Figure 6. Power calculation for a hypothetical 1-year trial in typical BMD, using NSAA as an outcome measure. This power calculation assumes selection of BMD patients with a typical BMD phenotype based on functional or genetic criteria, and exclusion of patients with mild/asymptomatic or DMD-like phenotypes; an accepted type I error rate of $\alpha = 0.05$, and a required statistical power $(1 - \beta) = 0.8$, and a blinded placebo-controlled design. The number of patients per study arm required for adequate power increases, as the hypothetical difference in 1-year NSAA change between treatment and placebo decreases.

Discussion

The patient population recruited in our study recapitulates the usually described BMD mutation distribution (81% deletions), and the phenotypic variability commonly observed in BMD [Darras et al., 2000]. The most severe cases showed progressive muscle wasting and LoA in early adulthood, while milder cases presented normal muscle strength and function [Angelini et al., 1994a].

We confirm that deletions ending on exon 51 are associated with relatively high dystrophin levels, averaging ~80%, very similar to levels measured in the same mutation group by a different technique quantitative immunohistochemistry by other laboratories [Anthony et al., 2011]. Although the two dystrophin quantification techniques are not directly comparable, both our data and previous report point to relatively abundant dystrophin in this mutation group. As expected, these patients displayed normal or close-to-normal muscle strength and function, and did not present any functional deterioration over the course of one year. The functional picture was similar in a group of patients carrying the isolated deletion of exon 48, despite slightly lower dystrophin levels, averaging ~65%. Patients in both the “del x-51” and “del 48” groups did not suffer from loss of the ability to walk or run, even in the adult age, although this observation was somewhat limited by younger average age in these groups.

Patients in the “del 45-x” group, on the other hand, were confirmed to display a typical BMD phenotype [Bushby et al., 1993; Anthony et al., 2014a; van den Bergen et al., 2014b], with progressive muscle wasting and weakness resulting in loss of motor function, as shown by loss of the ability to run and rarely to walk, and by reduced 6MWT distance, NSAA scores, and TFT velocities. Dystrophin WB quantitative data paralleled functional data in this group, showing dystrophin levels averaging 40~50%. This was again close to levels measured by quantitative immunohistochemistry in a similar BMD subpopulation [Anthony et al., 2014a]. The exception within the “del 45-x” group was represented by two patients carrying the deletion of exons 45-55, who had higher dystrophin levels (~90%) and a milder phenotype (both NSAA = 34, 6MWT > 400 m), as

previously reported for this mutation [Ferreiro et al., 2009]. This is probably due to the the fact that the 45-55 deletion spans exons 50 and 51, encoding the hinge III domain, whose absence from internally deleted dystrophin seems to stabilize the protein [Carsana et al., 2005]. Subsequently, while these patients have been included in the "del 45-x" group based on DMD exon skipping models, they bear in fact more similarities with the "del x-51" group from the standpoint of genotype-phenotype correlation.

Taken together, these observations confirm that skipping of exon 51 could be expected to be more effective than skipping of exon 45 for the treatment of DMD [Anthony et al., 2011; Anthony et al., 2014a]. Theoretically, AON cocktails promoting the skipping of multiple exons may be expected to be more effective than single-exon skipping AONs for DMD patients with "hotspot" deletions, if multi-exon skipping caused the exclusion, rather than inclusion, of exons 50 and 51 from the resulting spliced transcript. Multi-exon skipping of exons 45-55 has been proposed as a potential therapy for up to 40-45% DMD patients, although issues related to low efficacy and toxicity of AON mixtures complicates the translation of this approach to human trials [Aoki et al., 2013]. If these technical hurdles were overcome, this approach could be considered also for BMD patients with relatively common deletions such as those of exons 45-47, 45-48, and 45-49, for whom skipping of exons up to 55 (or 51) could be expected to convert typical BMD into mild or asymptomatic BMD.

Some individual patients with rare mutations in our cohort offer further interesting insights into genoype-phenotype correlations. A 13 year old patient with a deletion of exons 3-9, which ablates most of the N-terminal actin binding domain, nevertheless presented a mild clinical picture with a NSAA score of 34/34. Dystrophin function might be preserved because of the existence of secondary actin-binding sites within the rod domain of dystrophin [Amann et al., 1998]. The novel missense mutation p.Thr160Pro was associated to a relatively severe phenotype (LoA at 38 years in a patient who previously underwent a heart transplant). This mutation introduces a proline, probably disrupting secondary structure, within the functionally critical actin-binding domains encoded by exon 6. A synonym nucleotide substitution (c.4299G>T, p.Gly1433Gly),

which likely causes the exclusion of exon 30 from the mature mRNA by disrupting an ESE, caused mild dystrophinopathy (NSAA 34/34, 6MWT > 400 m) in two brothers aged 35 and 47 years who had reduced (10~30%) dystrophin with slightly reduced molecular weight at WB. These cases represent a rare kind of pathogenetic mutation, and also display well preserved muscle strength and function despite relatively low dystrophin quantity ascertained by WB. Two patients in our cohort presented nonsense mutations. One, carrying the mutation c.4980G>A, p.Trp1660* in exon 35, presented a relatively mild BMD phenotype (32 year old, NSAA 32/34, 6MWT > 400 m) and 29% dystrophin at WB, while the other, carrying the mutation c.3843G>A, p.Trp1281* in exon 28, suffered from substantial weakness (25 years, NSAA 12/34, 6MWT 312 m) and had 17% dystrophin. A BMD phenotype is sometimes observed with nonsense mutations situated within in-frame exons of the rod domain, especially if the corresponding exon is predicted to have weak exon recognition signals, i.e. low splice acceptor site strength and low ESE density, leading to endogenous exon skipping [Flanigan et al., 2011]. Interestingly, exons 28 and 35 are actually predicted in the cited paper by Flanigan et al. to have relatively strong splice acceptor sites and average-to-high ESE density, so that other unidentified factors may be at play in these mutations, which determine alternative splicing to occur. These nonsense mutations have not been observed in other patients in the Leiden database [Aartsma-Rus et al., 2006]. Two brothers had an out-of-frame microdeletion in exon 74 (c.10587_10588delAG, p.Lys3505AlaFsX8), and were included in this study because of the presence of reduced dystrophin identified by IHC with the DYS-1 antibody directed towards the rod domain, and a BMD phenotype (one brother lost ambulation at the age of 22 years, and the other is ambulatory at 17 years). The distal situation of this frameshifting mutation may determine the escape of part of the transcript from nonsense-mediated decay, leading to reduced amounts of distally truncated dystrophin. Unfortunately, no muscle tissue was available to quantify dystrophin by WB. Finally, two patients with large duplications involving exons 13-42 and 19-41, which gave rise to abnormally large dystrophin proteins (up to 600 kDa), presented symptomatic BMD, suggesting altered structure and function of these

aberrant proteins. The patients carrying the larger duplication (13-42) was previously described as displaying a relatively mild phenotype [Angelini et al., 1990], but 25 years from the original report, his weakness has progressed (NSAA score of 12 at 35 years of age); conversely, the patient with the relatively smaller mutation (19-41) is more mildly affected (NSAA 33 at 47 years of age). Further studies would be required to establish if sheer protein size, or subverted phasing of protein domains, are dictating phenotype in these patients.

As proposed by other authors [Angelini et al., 1994a; Angelini et al., 1996; Comi et al., 1994; Anthony et al., 2011; van den Bergen et al., 2014b], we did identify a moderate correlation of dystrophin quantity with phenotype severity in the whole cohort (Figure 3). It is interesting to speculate on which molecular mechanisms stabilize internally deleted DMD transcripts and/or dystrophin proteins, and thus modulate the BMD phenotype. We did not observe a clear correlation between dystrophin quantity and phenotype severity within the most numerous mutation groups (deletions of exons 45-47 and 45-48, see the “del 45-x” data series in Figure 3), which was similar to findings in a Dutch cohort of BMD patients [van den Bergen et al., 2014b]. Observing a correlation within a homogeneous mutation group would eliminate the confounding effects due different mutations, and suggest more clearly that disease severity might be actually dictated by inter-individual variability in dystrophin quantity. Although this analysis is hindered by low patient numbers, challenges in accurate dystrophin quantification [Anthony et al., 2014b], and sampling variability of dystrophin in human muscle biopsies, there does not seem to exist such a linear correlation within homogeneous mutation groups. On the other hand, parallel differences in both dystrophin quantity and phenotype severity across different mutation groups, which are commonly and consistently observed, suggest that the physical properties of internally deleted dystrophin proteins dictate disease severity, by causing varying degrees of downstream pathogenetic phenomena, such as altered costamere resistance to mechanical stress, membrane hyperpermeability, and disrupted recruitment of, and interaction with dystrophin-associated glycoproteins and other relevant factors (e.g. nNOS, syntrophin,

dystrobrevin). It has recently been shown that the inflammatory milieu characterizing active dystrophic pathology, and specifically the activation of the TNFα - NF-κB signaling pathway, upregulate several microRNAs which target the DMD 3' untranslated region, thus inhibiting dystrophin translation [Fiorillo et al., 2015]. Based on this model, quantitative dystrophin reduction might be a consequence of qualitative dystrophin alterations in BMD, and in turn sustain a vicious cycle of exacerbated pathology. As a corollary, the use of anti-inflammatory agents might be beneficial in BMD with active dystrophic pathology, if compounds were available that had an improved benefit-to-side-effect ratio than glucocorticoids, which are currently considered too harmful for long-term treatments in typical BMD.

One of the main aims of this study was to test the adequacy of the 6MWT as an outcome measure in a BMD population. Despite a wide variability, the correlations with dystrophin quantity, age, and all other functional measures support its clinical meaningfulness. Patients with a mild or asymptomatic clinical picture (34/34 NSAA score) always walked more than 400 m, which is considered a threshold for exclusion from some DMD clinical trials because of low probability of functional deterioration in the following one to three years [Pane et al., 2014a]. The only exception was one patient suffering from a schizophrenic disorder, who could collaborate fully to study procedures, but had a slow gait with mixed hysterical and iatrogenic dystonic alterations, and covered only 183 m in the 6MWT despite a 34/34 score in the NSAA. Within the "del 45-x" mutation group, there was a strong inverse correlation of 6MWT distance with age, similar to findings reported in other cohorts with a composite muscle strength quantification [van den Bergen et al., 2014b].

The NSAA, on the other hand, showed an obvious "ceiling effect" in mild or asymptomatic patients, especially those in the "del 48" and "del x-51" mutation groups. For these patients, NSAA does not appear to be a feasible clinical endpoint. On the contrary, for typical BMD patients (e.g. the "del 45-x" group) there was an even distribution of patients among scale scores, a tight inverse correlation with age, and a strong correlation with 6MWT and timed items, suggesting a clinically meaningful

outcome measures. Similar considerations apply to TFTs. Asymptomatic patients were very fast in the proposed tasks, so that the accuracy of time measurements sometimes became an issue, but there was a gradual decrease of velocities with age and good correlations with other measures in typical BMD.

The 6MWT distance was stable after 1 year in the overall population, suggesting that the observation period was not long enough to detect a deterioration of walking function in BMD. Breakdown of 6MWT changes by mutation group revealed some apparent “improvements” in a few young, mildly affected or asymptomatic patients, probably due to a limitation of our study which did not include a baseline test-retest, causing some “learning effect”. Importantly, however, a trend towards a distance decrease (-12 m) was observed in the typical BMD population (Figure 4, panel A), hinting at a possible detection of significant changes with longer observation periods, and supporting the clinical meaningfulness of longitudinal 6MWT in BMD.

NSAA changes, although not informative for mild patients (“del 48”, “del x-51”) appeared even more sensitive in detecting disease progression, as there was a statistically significant decrease of -0.9 points in the whole cohort (Figure 4, panel B). This change was driven by a decrease of -1.3 points in the “del 45-x” group (Figure 4, panel B), which was in itself statistically significant. In this group, most patients (13/18) decreased of one or more point, and only one increased one point. The design of the NSAA scale renders the clinical meaning of a 1-point decrease immediate and specific by comparison with baseline values, e.g. the patient has lost the ability to rise from the floor, or needs handrail support to climb a step. The observation of a statistically significant decrease of NSAA in a group of patients with typical BMD has several important implications. First, the selection of patients based on functional characteristics appears paramount in order to be able to measure differences in disease progression due to the intervention under study; a baseline NSAA score of 10 to 32 might be an example of such a criterion. Second, we were able to provide a power calculation for a hypothetical BMD clinical trial, which should be considered preliminary and warrants collection of natural history data in larger BMD populations for longer

time periods, but nevertheless represents a step towards clinical trial readiness in BMD. Third, more in-depth analyses of the patterns of loss of function underlying decreased NSAA score are warranted to improve our understanding of BMD natural history (e.g. chronological order in the loss of differenct motor functions) and possibly adapt the NSAA scale, which was tailored specifically on DMD patterns of functional deterioration by Rasch analysis [Mayhew et al., 2011].

Longitudinal changes of TFTs did not appear to be sensitive in describing disease progression. Even more than with 6MWT, we observed noise due to learning effect and inconsistent performances, so that longer observation times are probably needed to accomplish meaningful results with these measures.

We acknowledge several limitations to this study, including the lack of test - retest at baseline, the unavailability of WB dystrophin quantification in all patients, and the loss of follow-up at 1 year in some patients. We plan to continue the collection of natural history data and expand the cohort to overcome these limitations.

In conclusion, we confirm genotype-phenotype correlations that have been previously established for BMD: deletions bordering exon 45 (but not including exons 50-51) cause a typical BMD phenotype, while deletions bordering exon 51 and isolated exon 48 deletions cause mild or no muscle wasting and weakness. In the perspective of exon skipping outcomes for DMD caused by “hotspot” (exon 45-53) deletions, we also confirm an outlook of potential better outcomes for exon skipping treatments resulting in the absence of the hinge III domain (exons 50-51) from the rescued, internally deleted dystrophin protein. Based on the novel longitudinal natural history data presented here, it appears crucial to select BMD patients with measurable muscle weakness and dysfunction at baseline, in order to observe meaningful functional changes during a 1-year time frame. NSAA, and to a lesser extent 6MWT, hold promise as feasible and clinically meaningful outcome measures for BMD clinical trials.

Conclusions

The development of the aims of this thesis has led to relevant results, which help deciphering several different and interconnected aspects of the clinical variability observed in Duchenne and Becker muscular dystrophy.

1. Our study of the known *DMD* modifiers in the CINRG-DNHS [Bello et al., 2015a], in addition to providing independent validation of the genetic associations described in the original cohorts, offered insights into the dynamics of disease modifying mechanisms. In particular, it became apparent that *SPP1* genotype may be a pharmacodynamic biomarker of glucocorticoid treatment response in DMD, rather than a modifier of disease progression by itself; and we highlighted the relevance of controlling for population stratification in DMD cohorts recruited in clinical studies, as this can be a powerful confounder in the interpretation of results. These findings have relevant implications for patient care, trial design, and interpretation of the data of natural history studies and interventional trials.
2. Genotype-phenotype correlations in the CINRG-DNHS confirmed that deletions amenable to exon 44 skipping often present a mild phenotype, and that in DMD the selection of placebo groups in trials of mutation-specific drugs should probably be limited to patients with similar or comparable mutations from the standpoint of natural history. Similarly, trials of non-mutation specific drugs should make an effort to balance mutations with “atypical” disease courses equally between treatment and placebo groups.
3. Treatment with deflazacort may have a greater efficacy than treatment with prednisone in DMD [Bello et al., 2015b], a hypothesis generated by the CINRG-DNHS observational data, and currently being tested by the phase 3 “FOR-DMD” trial for Finding the Optimum glucocorticoid Regimen in DMD.
4. We performed the first GWAS aiming to discover genetic modifiers in a rare neuromuscular disease. This GWAS and the subsequent validation studies included a

total of 769 DMD patients from 4 continents, and demonstrated that *CD40* is a novel modifier locus of DMD, highlighting the relevance of cell-mediated inflammatory mechanisms in DMD pathogenesis, and representing an interesting therapeutic target.

5. Our longitudinal natural history study of BMD identifies NSAA and 6MWT as feasible, clinically meaningful outcome measures for clinical trials in BMD, and provides the bases to power these studies appropriately and design inclusion criteria. Furthermore, we were able to predict better efficacy of exon 51 than 45 skipping in DMD.

Future perspectives include: functional studies to elucidate through which molecular mechanisms CD40 signaling modifies dystropathology, and fine-tune potential interventions aimed at modulating this pathway in “in vivo” models; continued collection of DMD and BMD natural history data to expand genetic modifier studies with enlarged populations; discovery studies of genetic modifiers in other neuromuscular diseases.

Ultimately, the overarching aim of our research in the field of genetic modifiers of neuromuscular diseases will be to exploit the growing potential of genotyping and next generation sequencing, in order to not only identify the pathogenetic mutation for each patient, but also interpret the role of the genetic background in dictating the phenotype, the clinical course, and the response to treatments; and use this information to provide improved care.

Acknowledgements

I am most grateful to my Supervisor at the University of Padova, Prof. Elena Pegoraro, and my Supervisor at the Research Center for Genetic Medicine, Children's National Medical Center, Dr. Eric P. Hoffman, for their joint scientific teaching and guidance, and their continuing professional and human support.

I also wish to thank:

- All the laboratory members and staff at the Research Center for Genetic Medicine, Children's National Medical Center, Washington, DC, USA, particularly but not exclusively: Eric Hoffman, Jaya Punetha, Peter Nghiem, Kitipong Uaesoontrachoon, Akanchha Kesari, Mamta Giri, Heather Gordish-Dressman, Gina Many, Whitney Barfield, Sebahattin Cirak, Alyson Fiorillo, Chris Heier, Jelena Perovanovic, Jesse Damsker, Blythe Dillingham, Sherry Dadgar, Joe Devaney, Yetrib Hathout, Haeri Seol, Kristy Brown, Javad Nazarian, Mojca Štampar, Gary Cunningham, Mary Pichaske, Lisa Sheehey;
- All the CINRG Study Chairs, Coordinating Center members and staff: Craig McDonald, Avital Cnaan, Tina Duong, Erik Henricson, Paula Clemens, Lauren Morgenroth, Adrienne Arrieta, Zoë Sund, Mohammad Ahmed; and all the CINRG Investigators (www.cinrgresearch.org);
- All the Neuromuscular laboratory and clinic members at the University of Padova, particularly but not exclusively: Elena Pegoraro, Andrea Barp, Sara Vianello, Claudio Semplicini, Gianni Sorarù, Giorgia Querin, Michelangelo Cao, Alessandra Gaiani, Chiara Ferrati, Paola Campadello, Marina Fanin, Cinzia Bertolin, Bruno Gavassini, Boris Pantic, Fabio Busato;
- Drs. Stanley Nelson and Richard Wang at the UCLA Human Genetics Institute for generating the Exome Chip data;
- Dr. Suzanne Leal (Baylor University) for admitting me into, and teaching in the Advanced Genome Mapping Course in Rockefeller University, New York City, NY, USA.

Some of the text which constitutes the Introduction to this thesis is adapted from a book chapter, of which I am the primary author: Bello L, Hoffman EP, Pegoraro E. Dystrophinopathies. In: Angelini C, editor. Muscular Dystrophy: causes and management. Hauppauge NY: Nova Science Publisher, 2013. 69-96. © 2013 Nova Science Publisher, Inc.

The work presented in Aim 1 of this thesis is published as Bello L, Kesari A, Gordish-Dressman H, Cnaan A, Morgenroth LP, Punetha J, Duong T, Henricson EK, Pegoraro E, McDonald CM, Hoffman EP; CINRG Investigators. Genetic modifiers of ambulation in the Cooperative International Neuromuscular Research Group Duchenne Natural History Study. Ann Neurol. 2015 Apr;77(4):684-96, © 2015 The Authors Annals of Neurology published by Wiley Periodicals, Inc. on behalf of American Neurological Association. This is an open access article under the terms of the Creative Commons Attribution NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes. This study was funded by grants from the National Institutes of Health (#U54HD053177, #R24HD050846, #UL1RR031988, #UL1RR024992, #U54RR026139, #G12RR003051, #1R01AR061875, #R01AR062380), the U.S. Department of Education/NIDRR (#H133B031118, #H133B090001), the U.S. Department of Defense (#W81XWH-09-1-0592).

The work presented in Aim 2 of this thesis is currently submitted for publication as a research manuscript by Bello L, Morgenroth LP, Gordish-Dressman H, Hoffman EP, McDonald CM, Cirak S; CINRG Investigators. This study was funded by grants from the U.S. Department of Education/NIDRR (#H133B031118, #H133B090001); U.S. Department of Defense (#W81XWH-12-1-0417); National Institutes of Health/NIAMS (#R01AR061875); German Research Foundation (DFG) and MDA Developmental Grant awarded to Sebahattin Cirak.

The work presented in Aim 3 of this thesis is published as Bello L, Gordish-Dressman H, Morgenroth LP, Henricson EK, Duong T, Hoffman EP, Cnaan A, McDonald CM; CINRG Investigators. Prednisone/prednisolone and deflazacort regimens in the CINRG

Duchenne Natural History Study. Neurology. 2015 Sep 22;85(12):1048-55, © 2015 American Academy of Neurology. This study was funded by grants from the National Institutes of Health (#U54HD053177, #R24HD050846, #UL1RR031988, #UL1RR024992, #U54RR026139, #G12RR003051, #1R01AR061875, #RO1AR062380), the U.S. Department of Education/NIDRR (#H133B031118, #H133B090001), the U.S. Department of Defense (#W81XWH-09-1-0592).

The work presented in Aim 4 of this thesis is currently under preparation for submission as a research manuscript by Bello L, Flanigan KM, Weiss RB, UDP Investigators, Spitali P, Aartsma-Rus A, Muntoni F, Zarahieva I, Ferlini A, Tuffery-Giraud S, Lochmüller H, Straub V, Bio-NMD Investogators, Barp A, Vianello S, Pegoraro E, Punetha J, Gordish-Dressman H, Giri M, McDonald CM, Hoffman EP, CINRG Investigators. This study was funded by grants from the National Institutes of Health (#U54HD053177, #R24HD050846, #UL1RR031988, #UL1RR024992, #U54RR026139, #G12RR003051, #1R01AR061875, #RO1AR062380), the U.S. Department of Education/NIDRR (#H133B031118, #H133B090001), the U.S. Department of Defense (#W81XWH-09-1-0592), the AFM grant 15092, the EU funded projects BIO-NMD (grant EU 241665), Neuromics (No. 305121) and RD-Connect (No. 305444), the Medical Research Council (MRC) Centre for Neuromuscular Diseases Biobanks (Newcastle and London) which are part of EuroBioBank., and the National Institute for Health Research Biomedical Research Centre at Great Ormond Street Hospital for Children NHS Foundation Trust and University College London.

The work presented in Aim 5 of this thesis is currently submitted for publication as a research manuscript by Bello L, Campadello P, Barp A, Fanin M, Semplicini S, Sorarù G, Angelini C, Pegoraro E. This study was funded by Telethon Genetic BioBank (GTB12001D) and the Eurobiobank network.

References

Aartsma-Rus A, et al. Entries in the Leiden Duchenne muscular dystrophy mutation database: an overview of mutation types and paradoxical cases that confirm the reading-frame rule. Muscle Nerve. 2006;34:135-144.

Aartsma-Rus A. Overview on DMD exon skipping. Methods Mol Biol. 2012;867:97-116.

Amann KJ, et al. A cluster of basic repeats in the dystrophin rod domain binds F-actin through an electrostatic interaction. J Biol Chem. 1998;273:28419-28423.

Anderson JE, et al. Deflazacort increases laminin expression and myogenic repair, and induces early persistent functional gain in mdx mouse muscular dystrophy. Cell Transplant. 2000;9:551-564.

Angelini C, et al. Enormous dystrophin in a patient with Becker muscular dystrophy. Neurology. 1990;40:808-812.

Angelini C, et al. Clinical-molecular correlation in 104 mild X-linked muscular dystrophy patients: characterization of sub-clinical phenotypes. Neuromuscul Disord. 1994a;4:349-358.

Angelini C, et al. Deflazacort in Duchenne dystrophy: study of long-term effect. Muscle Nerve. 1994b;17:386-391.

Angelini C, et al. Prognostic factors in mild dystrophinopathies. J Neurol Sci. 1996;142:70-78.

Anthony K, et al. Dystrophin quantification and clinical correlations in Becker muscular dystrophy: implications for clinical trials. Brain. 2011;134: 3547-59.

Anthony K, et al. Biochemical characterization of patients with in-frame or out-of-frame DMD deletions pertinent to exon 44 or 45 skipping. JAMA Neurol. 2014a;71:32-40.

Anthony K, et al. Dystrophin quantification: Biological and translational research implications. Neurology. 2014b;83:2062-2069.

Aoki Y, et al. Development of multiexon skipping antisense oligonucleotide therapy for Duchenne muscular dystrophy. Biomed Res Int. 2013;2013:402369.

Arechavala-Gomeza V, et al. Antisense oligonucleotide-mediated exon skipping for Duchenne muscular dystrophy: progress and challenges. Curr Gene Ther. 2012;12:152-160.

Barfield WL, et al. Eccentric muscle challenge shows osteopontin polymorphism modulation of muscle damage. Hum Mol Genetg. 2014;43:4043-4050.

Barp A, et al. Genetic modifiers of Duchenne muscular dystrophy and dilated cardiomyopathy. PLoS One. 2015;10:e0141240.

Bello L, et al. Importance of SPP1 genotype as a covariate in clinical trials in Duchenne muscular dystrophy. Neurology. 2012;79:159-162.

Bello L, et al. Genetic modifiers of ambulation in the Cooperative International Neuromuscular Research Group Duchenne natural history study. Ann Neurol. 2015a;77:684-696.

Bello L, et al. Prednisone/prednisolone and deflazacort regimens in the CINRG Duchenne Natural History Study. Neurology. 2015b;85:1-8.

Biggar WD, et al. Deflazacort treatment of Duchenne muscular dystrophy. J Pediatr. 2001;138:45-50.

Blanco-Kelly F, et al. CD40: novel association with Crohn's disease and replication in multiple sclerosis susceptibility. PLoS One. 2010;5:e11520.

Bonifati MD, et al. A multicenter, double-blind, randomized trial of deflazacort versus prednisone in Duchenne muscular dystrophy. Muscle Nerve. 2000;23:1344-1347.

Bortolini ER, Zatz M. Duchenne muscular dystrophy: comparison among different racial groups. Am J Med Genet. 1987;28:925-929.

Brooke MH, et al. Clinical investigation of Duchenne muscular dystrophy. Interesting results in a trial of prednisone. Arch Neurol. 1987;44:812-817.

Bush A, Dubowitz V. Fatal rhabdomyolysis complicating general anaesthesia in a child with Becker muscular dystrophy. Neuromuscul Disord. 1991;1:201-204.

Bushby KM, et al. Identification of a mutation in the promoter region of the dystrophin gene in a patient with atypical Becker muscular dystrophy. Hum Genet. 1991;88:195-199.

Bushby KM, et al. The clinical, genetic and dystrophin characteristics of Becker muscular dystrophy. II. Correlation of phenotype with genetic and protein abnormalities. J Neurol. 1993;240:105-12.

Bushby K, et al. Diagnosis and management of Duchenne muscular dystrophy, part 1: diagnosis, and pharmacological and psychosocial management. Lancet Neurol. 2010a;9:77-93.

Bushby K, et al. Diagnosis and management of Duchenne muscular dystrophy, part 2: implementation of multidisciplinary care. Lancet Neurol. 2010b;9:177-189.

Bushby K, et al. Ataluren treatment of patients with nonsense mutation dystrophinopathy. Muscle Nerve. 2014;50:477-487.

Carsana A, et al. Analysis of dystrophin gene deletions indicates that the hinge III region of the protein correlates with disease severity. Ann Hum Genet. 2005;69:253-9.

Cascabulho CM, et al. Defective T-lymphocyte migration to muscles in dystrophin-deficient mice. Am J Pathol. 2012;181:593-604.

Ceco E, McNally EM. Modifying muscular dystrophy through transforming growth factor-β. FEBS J. 2013;280:4198-4209.

Ceco E, et al. Targeting latent TGFβ release in muscular dystrophy. Sci Transl Med. 2014;6:259ra144.

Chen F, et al. CD40 gene polymorphisms confer risk of Behcet's disease but not of Vogt-Koyanagi-Harada syndrome in a Han Chinese population. Rheumatology. 2012;51:47-51.

Chen JM, et al. The association of CD40 polymorphisms with CD40 serum levels and risk of systemic lupus erythematosus. BMC Genet. 2015;16:121.

Chen YW, et al. Early onset of inflammation and later involvement of TGFbeta in Duchenne muscular dystrophy. Neurology. 2005;65:826-834

Cirak S, et al. Exon skipping and dystrophin restoration in patients with Duchenne muscular dystrophy after systemic phosphorodiamidate morpholino oligomer treatment: an open-label, phase 2, dose-escalation study. Lancet. 2011;378:595-605.

Comi GP, et al. Clinical variability in Becker muscular dystrophy. Genetic, biochemical and immunohistochemical correlates. Brain. 1994;117:1-14.

Dadgar S, et al. Asynchronous remodeling is a driver of failed regeneration in Duchenne muscular dystrophy. J Cell Biol. 2014;207:139-158.

Darras BT, et al. Dystrophinopathies. 2000 Sep 05 [updated. 2014 Nov 26]. In: Pagon RA et al, editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2015. Available from http://www.ncbi.nlm.nih.gov/books/NBK1119/

del Río-Espínola A, et al. CD40-1C>T polymorphism (rs1883832) is associated with brain vessel reocclusion after fibrinolysis in ischemic stroke. Pharmacogenomics. 2010;11:763-772.

DeSilva S, et al. Prednisone treatment in Duchenne muscular dystrophy. Long-term benefit. Arch Neurol. 1987;44:818-822.

Desmet FO, et al. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res. 2009;37:e67.

Doriguzzi C, et al. Exercise intolerance and recurrent myoglobinuria as the only expression of Xp21 Becker type muscular dystrophy. J Neurol. 1993;240:269-271.

Drachman DB, et al. Prednisone in Duchenne muscular dystrophy. Lancet. 1974;2:1409-1412.

Dubowitz V. Steroids in Duchenne dystrophy. Neuromuscul Disord. 2013a;23:527-528.

Dubowitz V. Response. Neuromuscul Disord. 2013b;23:697.

Dupont WD, Plummer WD. Power and sample size calculations: a review and computer program. Controlled Clinical Trials. 1990;11:116-28.

Dwianingsih EK, et al. A novel splicing silencer generated by DMD exon 45 deletion junction could explain upstream exon 44 skipping that modifies dystrophinopathy. J Hum Genet. 2014;59:423-429.

Emery AE. The muscular dystrophies. Lancet. 2002;359:687-695.

Emond MJ, et al. Exome sequencing of extreme phenotypes identifies DCTN4 as a modifier of chronic Pseudomonas aeruginosa infection in cystic fibrosis. Nat Genet. 2012;44:886-889.

Enoch MA, et al. Using ancestry-informative markers to define populations and detect population stratification. J Psychopharmacol. 2006;20:19-26.

Escolar DM, et al. Randomized, blinded trial of weekend vs daily prednisone in Duchenne muscular dystrophy. Neurology. 2011;77:444-452.

Farini A, et al. T and B lymphocyte depletion has a marked effect on the fibrosis of dystrophic skeletal muscles in the scid/mdx mouse. J Pathol. 2007;213:229-238.

Fenichel GM, et al. Long-term benefit from prednisone therapy in Duchenne muscular dystrophy. Neurology. 1991;41:1874-1877.

Ferlini A, et al. X-linked dilated cardiomyopathy and the dystrophin gene. Neuromuscul Disord. 1999;9:339-346.

Ferreiro V, et al. Asymptomatic Becker muscular dystrophy in a family with a multiexon deletion. Muscle Nerve. 2009;39:239-243.

Field J, et al. The MS risk allele of CD40 Is Associated with Reduced Cell-Membrane Bound Expression in Antigen Presenting Cells: Implications for Gene Function. PLoS One. 2015;10:e0127080.

Fiorillo AA, et al. TNF-α-Induced microRNAs Control Dystrophin Expression in Becker Muscular Dystrophy. Cell Rep. 2015;12:1678-90.

Fisher I, et al. Prednisolone-induced changes in dystrophic skeletal muscle. FASEB J. 2005;19:834-836.

Flanigan KM, et al. DMD Trp3X nonsense mutation associated with a founder effect in North American families with mild Becker muscular dystrophy. Neuromuscul Disord. 2009;19:743-748.

Flanigan KM, et al. Nonsense mutation-associated Becker muscular dystrophy: interplay between exon definition and splicing regulatory elements within the DMD gene. Hum Mutat. 2011;32:299-308.

Flanigan KM, et al. LTBP4 genotype predicts age of ambulatory loss in Duchenne muscular dystrophy. Ann Neurol. 2013;73:481-488.

Flanigan KM, et al. Pharmacokinetics and safety of single doses of drisapersen in non-ambulant subjects with Duchenne muscular dystrophy: results of a double-blind randomized clinical trial. Neuromuscul Disord. 2014;24:16-24.

Fox DJ, et al. Trends With Corticosteroid Use in Males With Duchenne Muscular Dystrophy Born 1982-2001. J Child Neurol. 2015;30:21-26.

Gandhi KS, et al. The multiple sclerosis whole blood mRNA transcriptome and genetic associations indicate dysregulation of specific T cell pathways in pathogenesis. Hum Mol Genet. 2010;19:2134-2143.

García-Bermúdez M, et al. Study of association of CD40-CD154 gene polymorphisms with disease susceptibility and cardiovascular risk in Spanish rheumatoid arthritis patients. PLoS One. 2012;7:e49214.

Genin E, et al. Identifying modifier genes of monogenic disease: strategies and difficulties. Hum Genet. 2008;124:357-368.

Giacopelli F, et al. Polymorphisms in the osteopontin promoter affect its transcriptional activity. Physiol Genomics. 2004;20:87-96.

Goemans NM, et al. Systemic administration of PRO051 in Duchenne's muscular dystrophy. N Engl J Med. 2011;364:1513-1522.

Gospe SM Jr, et al. Familial X-linked myalgia and cramps: a nonprogressive myopathy associated with a deletion in the dystrophin gene. Neurology. 1989;39:1277-1280.

Griggs RC, et al. Prednisone in Duchenne dystrophy. A randomized, controlled trial defining the time course and dose response. Clinical Investigation of Duchenne Dystrophy Group. Arch Neurol. 1991;48:383-388.

Griggs RC, et al. Corticosteroids in Duchenne muscular dystrophy: major variations in practice. Muscle Nerve. 2013;48:27-31.

Gualandi F, et al. Intronic breakpoint definition and transcription analysis in DMD/BMD patients with deletion/duplication at the 5' mutation hot spot of the dystrophin gene. Gene. 2006;370:26-33.

Guglieri M, Bushby K. Molecular treatments in Duchenne muscular dystrophy. Curr Opin Pharmacol. 2010;10:331-337.

Gussoni E, et al. Specific T cell receptor gene rearrangements at the site of muscle degeneration in Duchenne muscular dystrophy. J Immunol. 1994;153:4798-4805.

Haas M, et al. European Medicines Agency review of ataluren for the treatment of ambulant patients aged 5 years and older with Duchenne muscular dystrophy resulting from a nonsense mutation in the dystrophin gene. Neuromuscul Disord. 2015;25:5-13.

Heier CR, et al. VBP15, a novel anti-inflammatory and membrane-stabilizer, improves muscular dystrophy without side effects. EMBO Mol Med. 2013;5:1569-1585.

Henricson EK, et al. The cooperative international neuromuscular research group Duchenne natural history study: glucocorticoid treatment preserves clinically meaningful functional milestones and reduces rate of disease progression as measured by manual muscle testing and other commonly used clinical trial outcome measures. Muscle Nerve. 2013;48:55-67.

Heydemann A, et al. Latent TGF-beta-binding protein 4 modifies muscular dystrophy in mice. J Clin Invest. 2009;119:3703-3712.

Hoffman EP, et al. Dystrophin: the protein product of the Duchenne muscular dystrophy locus. Cell. 1987;51:919-928.

Hoffman EP, et al. Characterization of dystrophin in muscle-biopsy specimens from patients with Duchenne's or Becker's muscular dystrophy. N Engl J Med. 1988;318: 1363-1368.

Hoffman EP, et al. Improved diagnosis of Becker muscular dystrophy by dystrophin testing. Neurology. 1989;39:1011-1017.

Hoffman EP, et al. Is the carboxyl-terminus of dystrophin required for membrane association? A novel, severe case of Duchenne muscular dystrophy. Ann Neurol. 1991;30:605-610.

Hoffman EP, et al. Restoring dystrophin expression in duchenne muscular dystrophy muscle progress in exon skipping and stop codon read through. Am J Pathol. 2011;179:12-22.

Hoffman EP, et al. Novel approaches to corticosteroid treatment in Duchenne muscular dystrophy. Phys Med Rehabil Clin N Am. 2012;23:821-828.

Hoffman EP, et al. Alterations in osteopontin modify muscle size in females in both humans and mice. Med Sci Sports Exerc. 2013;45:1060-1068.

Hoffman EP, McNally EM. Exon-skipping therapy: a roadblock, detour, or bump in the road? Sci Transl Med. 2014;6:230fs14.

Holtzer C, et al. Disparities in the diagnostic process of Duchenne and Becker muscular dystrophy. Genet Med. 2011;13:942-947.

Humbertclaude V, et al. Motor and respiratory heterogeneity in Duchenne patients: implication for clinical trials. Eur J Paediatr Neurol. 2012;16:149-160.

İnal EE, et al. Associations of rs4810485 and rs1883832 polymorphisms of CD40 gene with susceptibility and clinical findings of Behçet's disease. Rheumatol Int. 2015;35:837-843.

Jacobs SC, et al. Prednisone can protect against exercise-induced muscle damage. J Neurol. 1996;243:410-416.

Jacobson EM, et al. A Graves' disease-associated Kozak sequence single-nucleotide polymorphism enhances the efficiency of CD40 gene translation: a case for translational pathophysiology. Endocrinology. 2005;146:2684-2691.

Jiang DK, et al. Genetic variants in five novel loci including CFB and CD40 predispose to chronic hepatitis B. Hepatology. 2015;62118-621128.

Joo YB, et al. Association of genetic polymorphisms in CD40 with susceptibility to SLE in the Korean population. Rheumatology. 2013;52:623-630.

Kaspar RW, et al. Analysis of dystrophin deletion mutations predicts age of cardiomyopathy onset in Becker muscular dystrophy. Circ Cardiovasc Genet. 2009;2:544-51.

Kenneson A, et al. Trends and racial disparities in muscular dystrophy deaths in the United States, 1983-1998: an analysis of multiple cause mortality data. Am J Med Genet. 2006;140:2289-2297.

Kesari A, et al. Integrated DNA, cDNA, and protein studies in Becker muscular dystrophy show high exception to the reading frame rule. Hum Mutat. 2008;29:728-737.

Kim S, et al. Corticosteroid treatments in males with Duchenne muscular dystrophy: treatment duration and time to loss of ambulation. J Child Neurol. 2015;30:1275-1280.

Kinali M, et al. Local restoration of dystrophin expression with the morpholino oligomer AVI-4658 in Duchenne muscular dystrophy: a single-blind, placebo-controlled, dose-escalation, proof-of-concept study. Lancet Neurol. 2009;8:918-928.

Kissel JT, et al. Mononuclear cell analysis of muscle biopsies in prednisone-treated and untreated Duchenne muscular dystrophy. Neurology. 1991;41:667-672.

Li M, et al. CD40 C/T-1 polymorphism plays different roles in Graves' disease and Hashimoto's thyroiditis: a meta-analysis. Endocr J. 2012;59:1041-1050.

Lu QL, et al. What Can We Learn From Clinical Trials of Exon Skipping for DMD? Mol Ther Nucleic Acids. 2014;3:e152.

Manzur AY, et al. Glucocorticoid corticosteroids for Duchenne muscular dystrophy. Cochrane Database Systematic Rev. 2008;1:CD003725.

Mayhew A, et al. Moving towards meaningful measurement: Rasch analysis of the North Star Ambulatory Assessment in Duchenne muscular dystrophy. Dev Med Child Neurol. 2011;53:535-542.

Mazzone E, et al. Reliability of the North Star Ambulatory Assessment in a multicentric setting. Neuromuscul Disord. 2009;19:458-461.

Mazzone E, et al. North Star Ambulatory Assessment, 6-minute walk test and timed items in ambulant boys with Duchenne muscular dystrophy. Neuromuscul Disord. 2010;20:712-716.

Mazzone E, et al. Functional changes in Duchenne muscular dystrophy: a 12-month longitudinal cohort study. Neurology. 2011;77:250-256.

Mazzone E, et al. 24 month longitudinal data in ambulant boys with Duchenne muscular dystrophy. PLoS One. 2013;8:e52512.

McAdam LC, et al. The Canadian experience with long-term deflazacort treatment in Duchenne muscular dystrophy. Acta Myol. 2012;31:16-20.

McDonald CM, et al. The 6-minute walk test as a new outcome measure in Duchenne muscular dystrophy. Muscle Nerve. 2010;41:500-510.

McDonald CM, et al. The cooperative international neuromuscular research group Duchenne natural history study--a longitudinal investigation in the era of glucocorticoid therapy: design of protocol and the methods used. Muscle Nerve. 2013;48:32-54.

Melacini P, et al. Myocardial involvement is very frequent among patients affected with subclinical Becker's muscular dystrophy. Circulation. 1996;94:3168-3175.

Mendell JR, et al. Randomized, double-blind six-month trial of prednisone in Duchenne's muscular dystrophy. N Engl J Med. 1989;320:1592-1597.

Mendell JR, et al. Eteplirsen for the treatment of Duchenne muscular dystrophy. Ann Neurol. 2013;74:637-647.

Mendell JR, et al. Longitudinal effect of eteplirsen vs. historical control on ambulation in DMD. Ann Neurol. 2015 Nov 17. doi: 10.1002/ana.24555. [Epub ahead of print]

Monaco AP, et al. An explanation for the phenotypic differences between patients bearing partial deletions of the DMD locus. Genomics. 1988;2:90-95.

Morrison J, et al. T-cell-dependent fibrosis in the mdx dystrophic mouse. Lab Invest. 2000;80:881-891.

Morrison J, et al. Effects of T-lymphocyte depletion on muscle fibrosis in the mdx mouse. Am J Pathol. 2005;166:1701-1710.

Morrone A, et al. Asymptomatic dystrophinopathy. Am J Med Genet. 1997;69:261-267.

Moser H, Emery AE. The manifesting carrier in Duchenne muscular dystrophy. Clin. Genet. 1974;5:271-284.

Mostacciuolo ML, et al. Population data on benign and severe forms of X-linked muscular dystrophy. Hum Genet. 1987;75:217-220.

Moxley RT 3rd, et al. Practice parameter: corticosteroid treatment of Duchenne dystrophy: report of the Quality Standards Subcommittee of the American Academy of Neurology and the Practice Committee of the Child Neurology Society. Neurology. 2005;64:13-20.

Muntoni F, et al. Deletions in the 5' region of dystrophin and resulting phenotypes. J Med Genet. 1994;31:843-847.

Neale BM, Purcell S. The positives, protocols, and perils of genome-wide association. Am J Med Genet B Neuropsychiatr Genet. 2008;147B:1288-1294.

Nelson SF, Griggs RC. Predicting the severity of Duchenne muscular dystrophy: implications for treatment. Neurology. 2011;76:208-209.

Nieters A, et al. A functional TNFRSF5 polymorphism and risk of non-Hodgkin lymphoma, a pooled analysis. Int J Cancer. 2011;128:1481-1485.

Onouchi Y, et al. A genome-wide association study identifies three new risk loci for Kawasaki disease. Nat Genet. 2012;44:517-521.

Orozco G, et al. Association of CD40 with rheumatoid arthritis confirmed in a large UK case-control study. Ann Rheum Dis. 2010;69:813-816.

Panach L, et al. The role of CD40 and CD40L in bone mineral density and in osteoporosis risk: A genetic and functional study. Bone. 2015;83:94-103.

Pane M, et al. Long term natural history data in ambulant boys with Duchenne muscular dystrophy: 36-month changes. PLoS One. 2014a;9:e108205.

Pane M, et al. 6 Minute walk test in Duchenne MD patients with different mutations: 12 month changes. PLoS One. 2014b;9:e83400.

Pegoraro E, et al. Genetic and biochemical normalization in female carriers of Duchenne muscular dystrophy: evidence for failure of dystrophin production in dystrophin-competent myonuclei. Neurology. 1995;45:677-690.

Pegoraro E, et al. SPP1 genotype is a determinant of disease severity in Duchenne muscular dystrophy. Neurology. 2011;76:219-226.

Pineda B, et al. A C >T polymorphism located at position -1 of the Kozak sequence of CD40 gene is associated with low bone mass in Spanish postmenopausal women. Osteoporos Int. 2008;19:1147-1152.

Pinto-Mariz F, et al. Differential integrin expression by T lymphocytes: potential role in DMD muscle damage. J Neuroimmunol. 2010;223:128-130.

Piva L, et al. TGFBR2 but not SPP1 genotype modulates osteopontin expression in Duchenne muscular dystrophy muscle. J Pathol. 2012;228:251-259.

Prior TW, et al. Dystrophin expression in a Duchenne muscular dystrophy patient with a frame shift deletion. Neurology. 1997;48:486-488.

Purcell S, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559-575.

Raychaudhuri S, et al. Common variants at CD40 and other loci confer risk of rheumatoid arthritis. Nat Genet. 2008;40:1216-1223.

Reitter B. Deflazacort vs. prednisone in Duchenne muscular dystrophy: trends of an ongoing study. Brain Dev. 1995;S17:39-43.

Rhen T, Cidlowski JA. Antiinflammatory action of glucocorticoids--new mechanisms for old drugs. N Engl J Med. 2005;353:1711-1723.

Ricotti V, et al. Long-term benefits and adverse effects of intermittent versus daily glucocorticoids in boys with Duchenne muscular dystrophy. J Neurol Neurosurg Psychiatry. 2013a;84:698-705.

Ricotti V, et al. Steroids in Duchenne muscular dystrophy. Neuromusc Disord. 2013b;23:696-697.

Rodríguez-Rodríguez L, et al. Influence of CD40 rs1883832 polymorphism in susceptibility to and clinical manifestations of biopsy-proven giant cell arteritis. J Rheumatol. 2010;37:2076-2080.

Rosenberg AS, et al. Immune-mediated pathology in Duchenne muscular dystrophy. Sci Transl Med. 2015;7:299rv4.

Scott E, et al. Development of a functional assessment scale for ambulatory boys with Duchenne muscular dystrophy. Physiother Res Int. 2012;17:101-9.

Sejerson T, et al. Standards of care for Duchenne muscular dystrophy: brief TREAT-NMD recommendations. Adv Exp Med Biol. 2009;652:13-21.

Servais L, et al. Non-ambulant duchenne patients theoretically treatable by exon 53 skipping have severe phenotype. J Neuromusc Dis. 2015;2:269-279.

Shuang C, et al. Association of CD40 gene polymorphisms with sporadic breast cancer in Chinese Han women of Northeast China. PLoS One. 2011;6:e23762.

Skibola CF, et al. A functional TNFRSF5 gene variant is associated with risk of lymphoma. Blood. 2008;111:4348-4354.

Sklar RM, Brown RH, Jr. Methylprednisolone increases dystrophin levels by inhibiting myotube death during myogenesis of normal human muscle in vitro. J Neurol Sci. 1991;101:73-81.

Sokolova EA, et al. Association of SNPs of CD40 gene with multiple sclerosis in Russians. PLoS One. 2013;8:e61032.

Stark AE. Determinants of the incidence of Duchenne muscular dystrophy. Ann Transl Med. 2015;3:287.

Sugiura T, et al. Increased CD40 expression on muscle cells of polymyositis and dermatomyositis: role of CD40-CD40 ligand interaction in IL-6, IL-8, IL-15, and monocyte chemoattractant protein-1 production. J Immunol. 2000;164:6593-6600.

Tanizawa K, et al. A CD40 single-nucleotide polymorphism affects the lymphocyte profiles in the bronchoalveolar lavage of Japanese patients with sarcoidosis. Tissue Antigens. 2011;78:442-445.

Tian C, et al. Accounting for ancestry: population substructure and genome-wide association studies. Hum Mol Genet. 2008;17:R143-50.

Turner, SD. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. biorXiv. 2014; doi: 10.1101/005165.

Uaesoontrachoon K, et al. Osteopontin deficiency delays inflammatory infiltration and the onset of muscle regeneration in a mouse model of muscle injury. Dis Model Mech. 2013;6:197-205.

van den Bergen JC, et al. Prolonged ambulation in duchenne patients with a mutation amenable to exon 44 skipping. J Neuromusc Dis. 2014a;1:91-94.

van den Bergen JC, et al. Dystrophin levels and clinical severity in Becker muscular dystrophy patients. J Neurol Neurosurg Psychiatry. 2014b;85: 747-53.

van den Bergen JC, et al. Validation of genetic modifiers for Duchenne muscular dystrophy: a multicentre study assessing SPP1 and LTBP4 variants. J Neurol Neurosurg Psychiatry. 2015;86:1060-1065.

van der Linden MP, et al. Association of a single-nucleotide polymorphism in CD40 with the rate of joint destruction in rheumatoid arthritis. Arthritis Rheum. 2009;60:2242-2247.

van Deutekom JC, et al. Local dystrophin restoration with antisense oligonucleotide PRO051. N Engl J Med. 2007;357:2677-2686.

van Essen AJ, et al. Birth and population prevalence of Duchenne muscular dystrophy in the Netherlands. Hum Genet. 1992;88:258-266.

van Putten M, et al. Low dystrophin levels increase survival and improve muscle pathology and function in dystrophin/utrophin double-knockout mice. FASEB J. 2013;27:2484-2495.

Vetrone SA, et al. Osteopontin promotes fibrosis in dystrophic mouse muscle by modulating immune cell subsets and intramuscular TGF-beta. J Clin Invest. 2009;119:1583-1594.

Villalta SA, et al. Regulatory T cells suppress muscle inflammation and injury in muscular dystrophy. Sci Transl Med. 2014;6:258ra142.

Vo AH, McNally EM. Modifier genes and their effect on Duchenne muscular dystrophy. Curr Opin Neurol. 2015;28:528-534.

Voit T, et al. Safety and efficacy of drisapersen for the treatment of Duchenne muscular dystrophy (DEMAND II): an exploratory, randomised, placebo-controlled phase 2 study. Lancet Neurol. 2014;13:987-96.

Wang M, et al. The CD40 gene polymorphism rs1883832 is associated with risk of acute coronary syndrome in a Chinese case-control study. DNA Cell Biol. 2011;30:173-178.

Wein N, et al. Translation from a DMD exon 5 IRES results in a functional dystrophin isoform that attenuates dystrophinopathy in humans and mice. Nat Med. 2014;20:992-1000.

White SJ, den Dunnen JT. Copy number variation in the genome; the human DMD gene as an example. Cytogenet Genome Res. 2006;115:240-6.

Wills AM, et al. A large-scale international meta-analysis of paraoxonase gene polymorphisms in sporadic ALS. Neurology. 2009;73:16-24.

Winnard AV, et al. Frameshift deletions of exons 3-7 and revertant fibers in Duchenne muscular dystrophy: mechanisms of dystrophin production. Am J Hum Genet. 1995;56:158-166.

Wu CJ, et al. Association of CD40 polymorphisms and haplotype with risk of systemic lupus erythematosus. Rheumatol Int. 2015 Aug 21. doi: 10.1007/s00296-015-3345-7. [Epub ahead of print]

Yang J, et al. CD40 C/T(-1) and CTLA-4 A/G(49) SNPs are associated with autoimmune thyroid diseases in the Chinese population. Endocrine. 2012;41:111-115.

Yun Y, et al. The SNP rs1883832 in CD40 gene and risk of atherosclerosis in Chinese population: a meta-analysis. PLoS One. 2014;9:e97289.

Zatz M, et al. Milder course in Duchenne patients with nonsense mutations and no muscle dystrophin. Neuromuscul Disord. 2014;24:986-989.