UNIVERSITÀ DEGLI STUDI DI MILANO

DOCTORATE COURSE

Molecular and Cellular Biology, XXXIV cycle

Department of Bioscience

IDENTIFICATION AND CHARACTERIZATION OF A NOVEL NFIX PHOSPHORYLATION IN SKELETAL MUSCLE

BIO/17 – Istologia

PhD thesis of

GIADA MURA

Scientific supervisor: **Prof.ssa GRAZIELLA MESSINA**

Doctorate course coordinator: **Prof. ROBERTO MANTOVANI**

A.A.

2020/2021

Table of Contents

List of Figures

List of Tables

Acronyms and Abbreviations

BMD: Becker muscular dystrophy

BMP: Bone morphogenetic protein

CTX: cardiotoxin

DGC: dystrophin-glycoprotein complex

DMD: Duchenne muscular dystrophy

DMEM: Dulbecco's Modified Eagle's Medium

ECM: extracellular matrix

EDTA: Ethylenediaminetetraacetic acid

EGF: epidermal growth factor

ERK: extracellular signal regulated kinase

FA: Formic Acid

FAPs: fibroadipogenic progenitors

FBS: fetal bovine serum

FGF: Fibroblast growth factor

FGFR: fibroblast growth factor receptor

GlcNAc: N-Acetylglucosamine

GSK3: glycogen synthase kinase 3

HA: hemagglutinin

HGF: hepatocyte growth factor

HS: horse serum

HSPCs: hematopoietic stem cells

IAA: Iodoacetamide

IFN-γ: interferon-γ

IGF: insulin-like growth factor

IL: interleukin

IR: insulin receptor

LGMDs: Limb-Girdle Muscular Dystrophies

MAPK: mitogen-activated protein kinase

MD: muscular dystrophy

MRFs: muscle regulatory factors

MRSHSS: Marshall-Smith syndrome

MSCs: mechano-sensitive channels

MTJ: myotendinous junction

mTOR: mammalian target of rapamycin

MuSCs: muscle stem cells

MuSCs-Mb: muscle stem cells-derived myoblasts

MyHC: myosin heavy chain

NaF: sodium fluoride

nanoLC-HRMS: nanoLiquid Chromatography - High Resolution Mass Spectrometry

NFI: nuclear factor I

NMJ: neuromuscular junction

NO: nitric oxide

NOS: nitric oxide synthase

NSCs: neural stem cells

PBS: phosphate buffered saline

PHOSIDA: phosphorylation site database

PI3K: phosphatidylinositol 3-kinase

PMSF: Phenyl methane sulfonyl fluoride

PTMs : post translational modifications

RNA-seq: RNA sequencing

ROS: reactive oxygen species

S6K: S6 kinase

SACs: stretch-activated channels

SCs: satellite cells

SDS: sodium dodecyl sulfate

SDS-PAGE: sodium dodecyl sulfate – polyacrylamide gel electrophoresis

SO: sodium orthovanadate

TBST: Tris-buffered saline with Tween

TCEP: tris(2-carboxyethyl)phosphine)

TEAB: Triethylamonium bicarbonate

TFA: Trifluoroacetic acid

TGFβ: Transforming growth factor β

TNF-α: tissue necrosis factor α

TPA: tissue plasminogen activator

WT: wild type

PART I

Abstract

Nfix is a transcription factor involved in several biological processes spanning from the development of different tissues (brain, muscle, immune system) to cancer. Mutations in the Nfix gene are also at the basis of rare genetic disorders: the Malan syndrome and the Marshall-Smith syndrome.

In skeletal muscle, Nfix drives the transition from embryonic to fetal myogenesis in developmental myoblasts and the skewing from M1 to M2 macrophages. In the adult, Nfix plays a role in the maintenance of the correct timing of skeletal muscle regeneration upon injury regulating the expression of key genes in muscle stem cells. Silencing Nfix in dystrophic mice leads to a morphological and functional amelioration of the dystrophic phenotype by slowing down the degeneration-regeneration cycles, switching muscle fibers towards a slow-twitching phenotype, and by limiting the deposition of fibrotic tissue.

Despite the key roles of this transcription factor, a thorough understanding of its regulatory mechanisms in skeletal muscle and pathological conditions is lacking. Studies focusing on Nfix post-translational modifications (PTMs) are highly underrepresented in the literature. This works aims to fill this knowledge gap, through the characterization of the post-translational modifications of this protein and their biological role.

With this work, we demonstrate that Nfix is phosphorylated in skeletal myoblasts, both in fetal myoblasts during development and in adult muscle cells, *in vitro* and *in vivo*. The presence of at least one phosphorylation site in the C-terminal, transcriptionally active domain of Nfix, phospho-serine 301 (pS301), was confirmed both *in vitro* and *in vivo* in skeletal muscle cells and tissue. Our data are backed by several other high-throughput proteomics data, including data from human muscle biopsies. Furthermore, a search for cancer-relevant variants of S301 on cBioPortal Cancer Genomics database revealed a melanoma patient with a phosphorylation-disruptive serine-to-phenylalanine (S301F) mutation, suggesting a possible role of Nfix phosphorylation in cancer mechanisms. Serine 301 is very well-conserved among Nfix vertebrate orthologs, suggesting a functional relevance. This phosphorylation can be recognized as a slower migrating Nfix species in western blot experiments of cultured myoblasts. The phosphorylation-mediated mobility shift is dependent on the phosphate negative charge, as demonstrated by the S301D (serine-to-aspartate substitution) phosphomimetic Nfix variant. Ser 301 phosphorylation is detected already during fetal myogenesis and its levels remain constant during differentiation of postnatal myoblasts in culture. Data from cell lines obtained from murine *Nfix*-null myoblasts expressing phospho-

deficient (S301A) and phospho-mimetic (S301D) Nfix mutants reveal no evident phenotypical abnormalities, both in proliferation and differentiation.

Strikingly, we found that this phosphorylation is induced upon late activation of muscle stem cells (MuSCs), as it is absent in freshly isolated MuSCs but present in cultured ones. RNA-seq experiments performed on S301A-, S301D-, and Nfix2_WT-expressing myoblasts revealed a common role for Nfix in regulating cell cycle, differentiation, cell metabolism, and cell adhesion processes, all pathways differentially regulated between quiescent and activated MuSCs.

We have also shown that this phosphorylation is sensitive to oxidative stress and proteasome inhibition, as it significantly decreases upon hydrogen peroxide (H_2O_2) and proteasome inhibitors treatment in myoblasts. Intriguingly, the regulation of Nfix phosphorylation by oxidative stress could have interesting impacts in a dystrophic context, where reactive oxygen species such as H_2O_2 are abundant and dysregulated. However, pS301-Nfix levels are unchanged among wild-type and dystrophic *sgca*-null muscles, when detected via western blot.

To conclude, we highlight a possible role for pS301-Nfix in the regulation of MuSCs activation, cell cycle progression, and cell adhesion. Furthermore, Nfix phosphorylation levels could be tuned by the cells' metabolic state via reactive oxygen species.

The importance of this study relies on the clarification of Nfix regulatory mechanisms, a knowledge that could be employed in a dystrophic context to manipulate Nfix transcriptional activity and modulate MuSCs biology.

Furthermore, the dysregulation of Nfix phosphorylation could have a role in Marshall-Smith syndrome pathogenesis. Indeed, the Nfix residues interested by phosphorylation are coded by exons 5 to 8, which are the most affected by mutations in Marshall-Smith patients, that are characterized by de novo heterozygous mutations in the Nfix C-terminal domain, generating mutant proteins that seem to acquire dominant-negative/gain-of-function properties. An impaired interpretation of signaling pathways downstream of Nfix phosphorylation could be a relevant pathogenic mechanism of Marshall-Smith syndrome.

Riassunto

Nfix è un fattore di trascrizione coinvolto in diversi processi biologici che vanno dallo sviluppo di diversi tessuti (cervello, muscoli, sistema immunitario) ai tumori. Mutazioni del gene Nfix sono anche alla base di due malattie genetiche rare: la sindrome di Malan e la sindrome di Marshall-Smith. Nel muscolo scheletrico, Nfix guida la transizione dalla miogenesi embrionale a quella fetale durante lo sviluppo. Nell'adulto, Nfix regola la corretta tempistica di rigenerazione del muscolo scheletrico in seguito a una lesione, regolando l'espressione dei geni chiave nelle cellule staminali muscolari e la conversione dei macrofagi da fenotipo M1 a M2. Il silenziamento di Nfix nei topi distrofici porta a un miglioramento morfologico e funzionale del fenotipo distrofico rallentando i cicli di degenerazione-rigenerazione, indirizzando le fibre muscolari verso un fenotipo a contrazione lenta, e limitando la deposizione di tessuto fibrotico.

Nonostante il ruolo chiave di questo fattore di trascrizione, ad oggi manca una comprensione approfondita dei suoi meccanismi regolatori nel muscolo scheletrico e in condizioni patologiche. Studi incentrati sulle modifiche post-traduzionali di Nfix sono fortemente sottorappresentati in letteratura. Questo lavoro mira a colmare questa lacuna conoscitiva, attraverso la caratterizzazione delle modificazioni post-traduzionali di questa proteina e il loro ruolo biologico.

Con questo lavoro, dimostriamo che Nfix è fosforilato nei mioblasti del muscolo scheletrico, sia nei mioblasti fetali durante lo sviluppo che nelle cellule muscolari adulte, *in vitro* e *in vivo*. La presenza di almeno un sito di fosforilazione nel dominio C-terminale, trascrizionalmente attivo di Nfix, fosfo-serina 301 (pS301), è stata confermata sia *in vitro* che *in vivo* nelle cellule e nei tessuti muscolari scheletrici. I nostri dati sono supportati da numerosi dati di proteomica, inclusi dati provenienti da biopsie muscolari umane. La serina 301 di Nfix è molto ben conservata in diverse specie di vertebrati, suggerendo una importanza funzionale. La carica negativa portata dal gruppo fosfato sulla serina 301 causa una alterata corsa elettroforetica di Nfix e può pertanto essere riconosciuto tramite esperimenti di western blot. La fosforilazione della Ser 301 viene rilevata già durante la miogenesi fetale e i suoi livelli rimangono costanti durante il differenziamento dei mioblasti postnatali in coltura. Dati provenienti da linee cellulari ottenute da mioblasti murini *Nfix*-null trasdotti con mutanti fosfo-deficienti (S301A) e fosfo-mimetici (S301D) di Nfix non rivelano anomalie fenotipiche evidenti, sia in proliferazione che in differenziamento. Questa fosforilazione è invece indotta con l'attivazione delle cellule staminali muscolari (MuSC), poiché è assente nelle MuSC

appena isolate dal tessuto (quiescenti) ma presente in quelle messe in coltura (attivate). Esperimenti di RNA-seq eseguiti su mioblasti che esprimono S301A-, S301D- ed Nfix2_WT hanno rivelato un ruolo di Nfix nella regolazione del ciclo cellulare, differenziamento, metabolismo cellulare e nei processi di adesione cellulare, tutti processi differenzialmente regolati tra MuSC quiescenti e attivate. Abbiamo inoltre dimostrato che questa fosforilazione è sensibile allo stress ossidativo e all'inibizione del proteasoma, poiché diminuisce significativamente in mioblasti trattati con perossido di idrogeno (H_2O_2) e inibitori del proteasoma.

Per concludere, evidenziamo un possibile ruolo per pS301-Nfix nella regolazione dell'attivazione delle cellule staminali del muscolo e nella progressione del ciclo cellulare. La scoperta di questi nuovi meccanismi regolatori di Nfix potrebbe essere impiegata in un contesto distrofico per manipolare l'attività trascrizionale di Nfix e modulare la biologia delle ceullule staminali del muscolo.

Inoltre, la disregolazione della fosforilazione di Nfix potrebbe avere un ruolo nella patogenesi della sindrome di Marshall-Smith. Infatti, i residui Nfix interessati dalla fosforilazione sono codificati dagli esoni da 5 a 8, che sono i più colpiti dalle mutazioni nei pazienti affetti da sindrome di Marshall-Smith, i quali sono caratterizzati da mutazioni eterozigoti nel dominio C-terminale di Nfix, generando proteine mutanti che sembrano acquisire proprietà di dominanti negativi. Un'interpretazione alterata delle vie di segnalazione a valle della fosforilazione di Nfix potrebbe essere un meccanismo rilevante della sindrome di Marshall-Smith.

1 Introduction

1.1 Skeletal Muscle Development

The skeletal muscle is a highly specialized tissue part of the muscoloskeletal system that also consists of nerves (the motor neurons) connected to the muscle by neuromuscular junctions (NMJ), vasculature networks, connective tissue, bones, tendons that attach the muscle to the bones through myotendinous junctions (MTJs), and many other cell types invading the muscle, such as macrophages and FAPs. The basic functional unit of the skeletal muscle is the sarcomere, a structure found in repeated numbers along the myofiber and composed of orderly associated actin and myosin myofibrils. These two are the main contractile proteins that allow the muscle twitching following motor neuron stimulation. Myosins are also a reference used to classify myofibers based on their myosin heavy chain (MyHCs) isoforms expression (slow, type I, or fast, types II) (**Table 1**). Another classification is operated based on myofiber metabolism: oxidative or glycolytic (Schiaffino and Reggiani, 2011).

Early during development, the formation of the skeletal muscle, as many other tissues within the developing embryo, is governed by anterior-posterior, mediolateral and dorsoventral gradients of essential signaling molecules such as BMP, Wnt, FGF, and Notch. Through a complex and finely coordinated intersection of these gradients, several boundaries and segments are formed within the embryo. The body's skeletal muscle patterning is prompted by these signaling pathways and starts on the eighth embryonic day, or E8, in the mouse with the formation of the somites, transient embryonic structures that originate from the paraxial mesoderm, at both sides of the neural tube and notochord, following an anterior-to-posterior developmental gradient. Soon after their formation, somites themselves start to be compartmentalized responding to Wnt, BMP, and Shh signaling gradients from surrounding tissues. It follows the dorsoventral compartmentalization of the somites into a dorsal epithelial dermomyotome and a ventral mesenchymal sclerotome. The skeletal muscle cells arise from the dorsal dermomyotome, which can be further subdivided into a dorsomedial (epaxial lip), a ventrolateral (hypaxial lip), and a central compartment (**Figure 1A**)(Chal and Pourquié, 2017; Tajbakhsh and Buckingham, 2000). Migration, proliferation, and differentiation of the first myogenic progenitors begin from the edges of the dermomyotome, with the activation of the myogenic factor Myf5 in Pax3/7-positive progenitor cells (**Figure 1B**)(Gros et al., 2005; Kahane et al., 2001; Kassar-duchossoy et al., 2005; Relaix et al., 2005). They form the primary myotome with the appearance of the first

postmitotic mononucleated skeletal muscle cells expressing specialized cytoskeletal proteins as slow (*myh7*) and embryonic (*myh3*) myosin heavy chains (MyHCs) (Cinnamon et al., 1999; Denetclaw and Ordahl, 2000; Denetclaw et al., 2001). Afterward, myogenesis proceeds with the progressive migration of more $Pax3^+$ progenitor cells that in part populate and fuse to the primary myotome (Kahane et al., 2001), and in part migrate away from the somite to form muscle masses elsewhere (e.g. limbs, diaphragm), differentiating into slow-$MyHC^+$ myofibers guided by a set of muscle regulatory factors (MRFs), the gatekeepers for entry and progression into the myogenic lineage: Myf5, MyoD, MRF4 (*Myf6*), and myogenin (Davis et al., 1987; Edmondson and Olson, 1989; Rhodes and Konieczny, 1989; Rudnicki et al., 1993; Wright et al., 1989). This represents the primary phase of myogenesis or embryonic myogenesis, which goes from E10.5 to E12.5 in the mouse. During the secondary phase of myogenesis, also called fetal myogenesis (from E14.5 to E17.5), a subset of myogenic precursor cells starts to downregulate Pax3 and to express Pax7 and other specific markers including Nfix and β-enolase (Biressi et al., 2007a; Messina and Cossu, 2009). These cells extensively proliferate and fuse into fast-MyHC-expressing myofibers: fast twitching fibers with a predominantly glycolytic metabolism. Meanwhile (at E16.5), a subset of Pax3/7+ cells originating from the dermomyotome do not express MRFs and migrate to the basal lamina forming the pool of adult muscle stem cells – the satellite cells (MuSCs) (**Figure 1A,B**) (Gros et al., 2005; Kassar-duchossoy et al., 2005; Relaix et al., 2005) (see Chapter 1.2.1).

Embryonic and fetal myoblasts and satellite cells represent distinct classes of myogenic cells, with defined gene expression signatures, as evidenced by several works (Biressi et al., 2007b, 2007a; Cossu and Biressi, 2005; Cossu and Molinaro, 1987; Miller et al., 1999; Stockdale, 1992) although they arise from a common population of muscle progenitors already present at the embryonic stage (Gros et al., 2005; Kassar-duchossoy et al., 2005; Relaix et al., 2005). They differ in proliferation, differentiation, fusion abilities, and sensitivity to myogenesis inhibitors (TGFβ, TPA, BMP-4) (Biressi et al., 2007a; Cossu et al., 1988; Cusella-De Angelis et al., 1994). These differences are exemplified by the unrelated morphologies of primary and secondary myofibers, with the expression of different myosin heavy chain isoforms and muscle enzymes. In particular, embryonic myoblasts proliferate less and are more prone to differentiation but form smaller myotubes with fewer myonuclei characterized by the expression of slow myosin heavy chain isoforms and the lactate dehydrogenase 2 enzyme. In contrast, fetal myoblasts differentiate into larger myotubes with clusters of several myonuclei that express fast myosin heavy chain isoforms and characteristic muscle enzymes such as muscle creatine kinase (*mck*) and βenolase (*eno3*) (Biressi et al., 2007a; Messina et al., 2010).

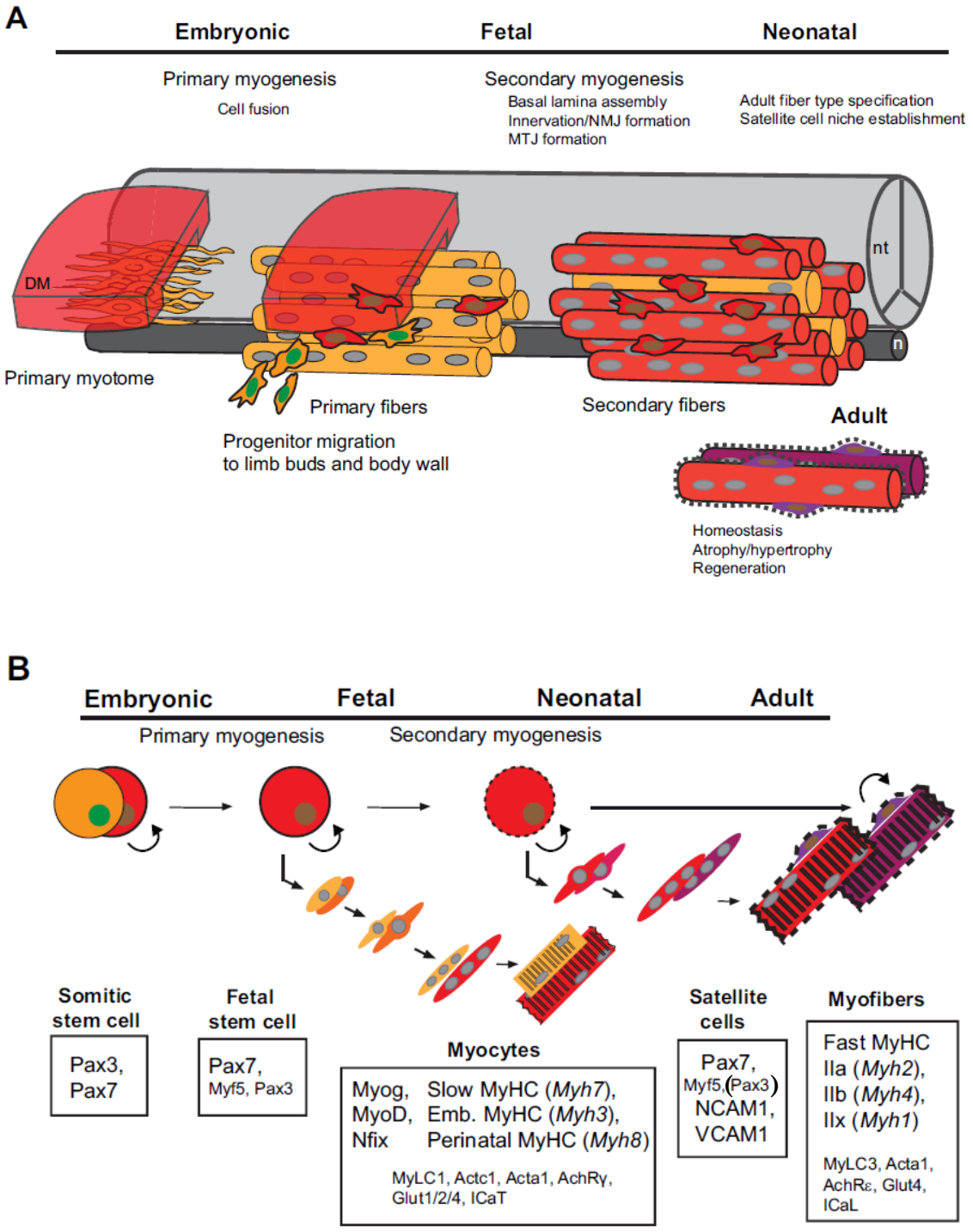


Figure 1. Maturation of somites during mouse embryogenesis

(A) Depiction of somites anterior-posterior maturation. Dermamyotome (DM), neural tube (nt), notochord (n). (B) Schematic representation of the different waves of myogenic progenitors' commitment and differentiation during mouse myogenesis, with the most characteristic genes expressed during the different phases. Image from (Chal and Pourquié, 2017)

1.2 Skeletal Muscle Regeneration

Throughout the life of an individual, the skeletal muscle tissue is continuously subjected to mechanical stress due to its nature as a contractile organ: varying degrees of injury can be experienced by the muscle tissue, from localized and minor damage (such as strains, stretching, or exercise) to more extended and severe insults (such as trauma or degenerative muscular diseases). The coordinated events that are prompted after damage to restore tissue and organ structure and function are collectively categorized under the word *regeneration* (Schiaffino et al., 2008). This process somehow correlates with embryonic muscle development, with many of the genes and regulatory pathways being redeployed during postnatal homeostasis. However, some differences exist between the two processes (Tajbakhsh, 2009).

The regenerative process is characterized by distinct phases: (1) tissue degeneration and inflammation, (2) tissue reconstruction, and (3) tissue remodeling.

Degeneration starts with the dissolution of the myofiber sarcolemma and the formation of a necrotic area that generally remains localized within a few millimeters from the rupture point thanks to the fast formation of a new demarcation sarcolemma that prevents the propagation of the necrosis (**Figure 2A-B**). Inside the gap formed between the two sealed torn fibers, chemotactic factors are released to the extracellular space and a series of inflammatory reactions take place. Specific myeloid populations sequentially invade the necrotic area (**Figure 2B-C**)(Schiaffino et al., 2008). Neutrophils constitute the first leukocyte wave during muscle regeneration (**Figure 2dx**); they prepare the muscle tissue for future remodeling acting on the necrotic area in a seemingly controversial way: they initially contribute to muscle degeneration through reactive oxygen species (ROS) production and releasing proteases that degrade damaged tissue and extracellular matrix (ECM) and then start to remove cellular debris by phagocytosis. Neutrophils also secrete proinflammatory cytokines that promote the subsequent inflammatory waves, whose main protagonists are different subpopulations of phenotypically-distinct macrophages. The earliest to appear (in 1-2 days) is the phagocytic, pro-inflammatory subpopulation, which contributes to the rapid removal of debris (**Figure 2dx**). They derive from circulating monocytes that are recruited and then activated *in situ* by proinflammatory cytokines secreted by neutrophils such as interferon-γ (IFN-γ), tissue necrosis factor α (TNF-α), and interleukin-1 (IL-1) that they also start to express. In the subsequent inflammatory phase, proinflammatory macrophages are readily deactivated by anti-inflammatory cytokines such as IL-4, IL-10, and IL-13, which

instead activate the nonphagocytic, anti-inflammatory (or pro-regenerative) macrophages, which start to express these cytokines themselves (**Figure 2dx**). This shift in macrophage phenotype is obtained *in situ* after phagocytosis of either apoptotic or necrotic myogenic cells, and in turn, it promotes the end of inflammation and the start of tissue repair and remodeling through secretion of factors such as transforming growth factor-β (TGF-β). Different studies suggest a role for macrophages in promoting MuSCs proliferation and differentiation; the proximity of pro-inflammatory macrophages to myogenic cells promotes the proliferation of the latter and their protection from apoptosis. This effect seems to be mediated by TNF-α. Instead, anti-inflammatory macrophages stimulate both the myogenic differentiation and fusion processes (Chazaud et al., 2009; Yang and Hu, 2018). Recently, Shang et al. demonstrated that macrophages communicate with myogenic cells metabolically through the secretion of glutamine and that this sustains a faster regeneration by satellite cells via mTOR pathway (Shang et al., 2020).

Due to the tight linkages between different cell types in the regenerating muscle, a correct appearing time, coordination, and distribution of these cell populations secure the right succession of events that leads to proper muscle regeneration. Perturbations of this balance, as observed in chronic muscle diseases, cause a cascade of catastrophic events and failure of the system (see Chapter 1.3).

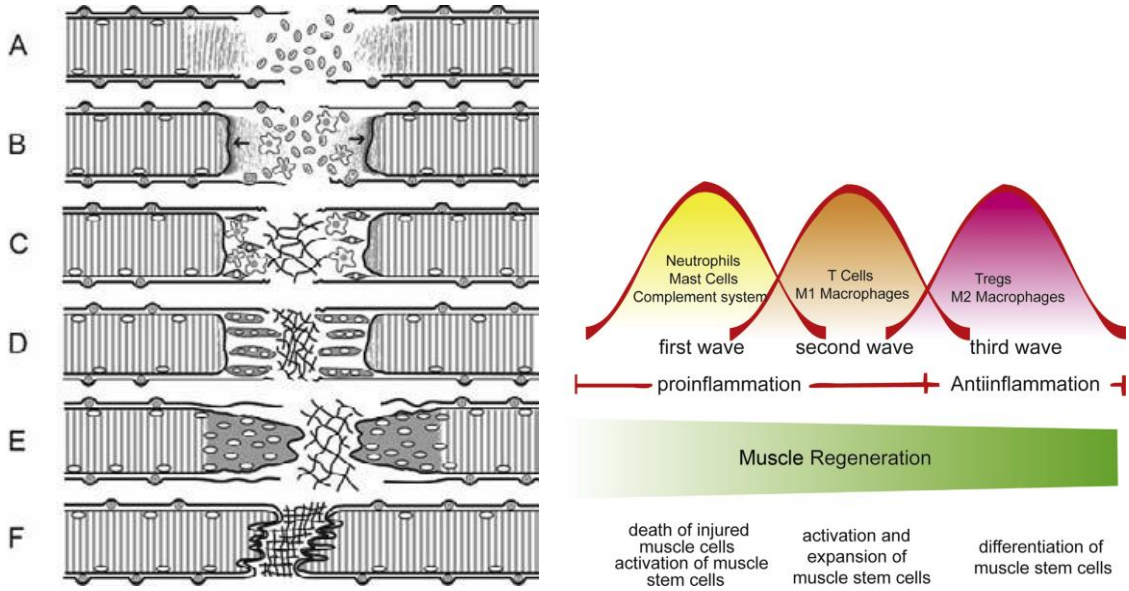


Figure 2. Tissue and cellular rearrangements in regenerating myofibers

(A-F) Schematic representation of the skeletal myofiber repair process upon physiological damage, from scar formation to resolution of inflammation and myofiber reconstruction. Image from (Schiaffino et al., 2008). On the right, highlight of the immune cells waves during muscle regeneration and their cross-talk with myogenic cells. Image from (Yang and Hu, 2018).

Muscle repair begins with the mobilization of committed satellite cells that readily proliferate, differentiate, and fuse to pre-existing myotubes or fuse with one another to form myofibers de novo. Meanwhile, a subset of stem satellite cells proliferates and self-renew generating new myoblasts and concurrently replenishing the MuSCs pool through asymmetric and symmetric divisions (see Chapter 1.2.1). The new myoblasts then differentiate and fuse into myofibers; within 5-6 days the regenerating myotubes invade the connective tissue scar formed in the previously necrotic area (Schiaffino et al., 2008). While homeostatic muscle fibers mainly express slow (type I) and fast (type II) myosin isoforms, during regeneration developmental myosin isoforms, such as embryonic (*myh3*) and neonatal (*myh8*) myosin heavy chains, are transiently re-expressed (**Table 1**). For this reason, they are often used as markers of regenerating fibers (Schiaffino et al., 2015). During this phase, newly forming myofibers can be recognized also by the position of myonuclei that appear at the center of the fiber and are indicative of the ongoing fusion process. These young fibers then increase in size and mature to eventually become indistinguishable from undamaged ones in normal conditions. However, particularly in case of extensive damage or muscular pathologies, it can happen that the myofiber integrity is not fully restored leaving behind scars and abnormal myofiber morphologies (e.g. branched myofibers)(Schiaffino et al., 2008)(see Chapter 1.3). Recently, Roman et al. demonstrated that cellular reconstruction during muscle repair after localized physiological damage can rely on myonuclear migration along the myofiber towards the injury point, independently of satellite cells. Upon injury, myonuclei are attracted to the damaged site by a signaling cascade involving calcium, cdc42, and phosphokinase C and promote sarcomere repair through local delivery of specific mRNAs (Roman et al., 2021).

Protein	Gene	Expression in developing muscle	Expression in adult muscle
Myosin heavy chains			
MyHC-emb	MYH3	Embryonic and fetal muscle	Specialized muscles/ regenerating fibers
MyHC-neo	MYH8	Embryonic and fetal muscle	Specialized muscles/ regenerating fibers
MyHC-slow	MYH7	Embryonic and fetal muscle	Type 1 muscle fibers
MyHC-2A	MYH2	Fetal (human) or early postnatal (mouse) muscle	Type 2A muscle fibers
MyHC-2X	MYH1	Late fetal (human) or early postnatal (mouse) muscle	Type 2X muscle fibers
MyHC-2B	MYH4	Postnatal muscle	Type 2B muscle fibers

Table 1. *MYH* genes expressed in developing and adult mammalian skeletal muscle
Adapted from (Schiaffino et al., 2015)

Remodeling of the injured muscle includes maturation of the myofibers with the formation of fully structured contractile units, formation of new musculotendinous junctions (MTJs), and restoration of the motor innervation with new neuromuscular junctions (NMJs) (Schiaffino et al., 2008).

1.2.1 The Muscle Stem Cells (MuSCs) – Satellite Cells

The adult muscle stem cells, called satellite cells and here abbreviated as MuSCs, in homeostatic conditions are generally quiescent (in G_0 phase) cells found in a niche microenvironment under the myofiber basal lamina (or endomysium), adjacent to the sarcolemma (Mauro, 1961) (**Figure 3**). Limb (hypaxial) and trunk (epaxial) satellite cells originate from the somite (Armand et al., 1983), from a population of progenitor cells that express *Pax7* and *Pax3* but do not upregulate the different MRFs (Gros et al., 2005; Kassar-duchossoy et al., 2005; Relaix et al., 2005) and are incapsulated under the basal lamina when this starts to form, during fetal myogenesis (see Chapter 1.1). Satellite cells sustain the early post-natal growth of the muscle tissue when they account for 30-35% of all sublaminal myonuclei. In late prenatal and neonatal/juvenile stages, the number of myofibers remains constant, but each myofiber grows in size by fusion of MuSCs into both primary and secondary fibers. This maturation process proceeds for about 2-3 weeks after birth in mouse, in which skeletal muscle undergoes robust growth associated with the general growth of the organism (Sambasivan and Tajbakhsh, 2007). In adulthood, they represent only 5-10% of total myonuclei but are still able to support the regeneration of the muscle (Parker et al., 2003).

It is now increasingly clear that MuSCs are a highly heterogeneous group of cells (Collins et al., 2005; Yin et al., 2013), that can be generally identified by molecular markers such as Pax7 (Seale et al., 2000), M-cadherin (Irintchev et al., 1994), CD34 (Beauchamp et al., 2000), and the transmembrane heparan sulfate proteoglycans syndecan 3 and 4 (Cornelison et al., 2001), with the latter two being less specific. Many other markers can be employed to identify different subpopulations of satellite cells, but Pax7 is generally accepted as the universal marker of MuSCs as it is expressed by most of the subpopulations identified so far, both quiescent and actively proliferating (Yin et al., 2013).

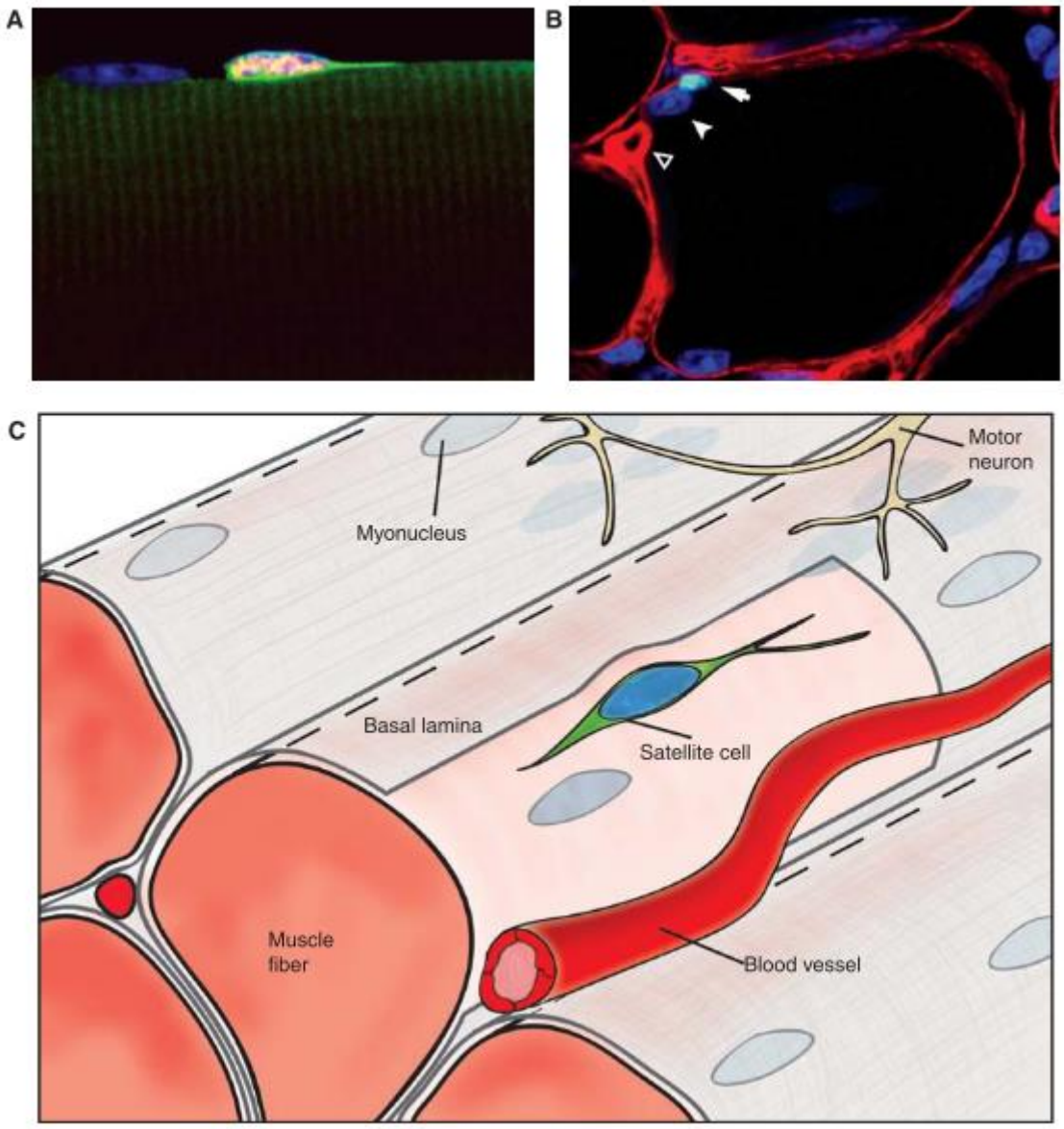


Figure 3. MuSCs and their niche: adult muscle-specific stem cells

(A-B) Satellite cells are found associated with the myofiber (A; cytoplasm in green, Pax7 in red, nucleus in blue) under the basal lamina (B; arrow; Pax7 in green, basal lamina in red) and in proximity to capillaries (B, empty arrowhead) and myonuclei (B; arrowhead; nucleus in blue). (C) schematic representation of the satellite cell and its surrounding niche microenvironment that usually comprehends a tight association with blood vessels and motor neurons. Image from (Bentzinger et al., 2012)

As already discussed, the first step of MuSCs engagement after damage is their activation. The signaling coming from the niche is responsible for the activation state of satellite cells: in homeostatic conditions, they actively maintain the quiescent state, while in response to muscle injury they promote the activation of satellite cells. Cytokines and growth factors released from the injured muscle stimulate membrane receptors of quiescent satellite cells, leading to the activation of different signaling pathways and thus specific gene expression

patterns (Dumont et al., 2015a, 2015b; Yin et al., 2013). Activation is mainly triggered by the signaling pathways downstream to Hepatocyte Growth Factor (HGF) and its interaction with the tyrosine kinase receptor c-Met on MuSCs (see Chapter 1.6.1.1) (Sousa-Victor et al., 2021). Activated satellite cells co-express Pax7, Myf5, and/or Myod and re-enter the cell cycle to perform several rounds of cell divisions before cell cycle re-withdrawal. Nondividing MuSCs-derived myoblasts either differentiate and contribute to muscle repair or re-enter quiescence re-establishing the satellite pool under the basal lamina. The ability of the skeletal muscle system to *self-renew*, that is to repopulate the staminal niche, is a necessary process to support subsequent rounds of muscle regeneration. Several studies have revealed a role for asymmetric division in satellite cell maintenance (Conboy et al., 2007; Kuang et al., 2007; Shinin et al., 2006) (**Figure 4**). However, it is not still proven if self-renewal comes from asymmetrically or symmetrically dividing MuSCs, or both (Kuang et al., 2008). Self-renewing MuSCs downregulate Myod and maintain Pax7 expression while differentiating myoblasts suppress Pax7 and start to express myogenin. A complex interplay of Pax7, Myod, and Myf5 control the fate of MuSCs and myoblasts: cells with a high Pax7/MyoD ratio are kept in their quiescent state; an intermediate Pax7/MyoD but high Myf5/Myod ratio enables proliferation, but not differentiation; low Pax7/MyoD and Myf5/Myod ratios induce satellite cells to differentiate, allowing myogenin expression (Yin et al., 2013). Myod and Pax7 are indeed able to modify the chromatin architecture of myogenic progenitors, orchestrating the changes that promote activation or quiescence (Dall'Agnese et al., 2019; Zhang et al., 2020). The interaction of MuSCs with their niche is fundamental for quiescence maintenance; in part, this is mediated by cell-cell and cell-ECM contacts. Consequently, MuSCs rapidly lose their stemness and differentiate into more committed muscle precursors, termed primary myoblasts, when isolated and expanded in vitro (Fuchs et al., 2004; Fukada et al., 2007; Machado et al., 2017, 2021; Moore and Lemischka, 2006).

Proliferation and differentiation of MuSCs are then coordinated by a complex interplay of signaling molecules coming from the surrounding environment and the communication with other cells, such as fibroblasts, macrophages, endothelial cells, and FAPs. Many of these cells release growth factors that facilitate cell division or differentiation, including HGF, bFGF (basic Fibroblast Growth Factor), EGF (Epidermal Growth Factor), and IGF (Insulin-like Growth Factors). bFGF for example promotes myogenic cell proliferation and is, for this reason, commonly added to the medium of primary myoblasts *in vitro* to stimulate myoblasts expansion (Dumont et al., 2015a; Yin et al., 2013). Some of these pathways are discussed more in detail in Chapter 1.6.1.1.

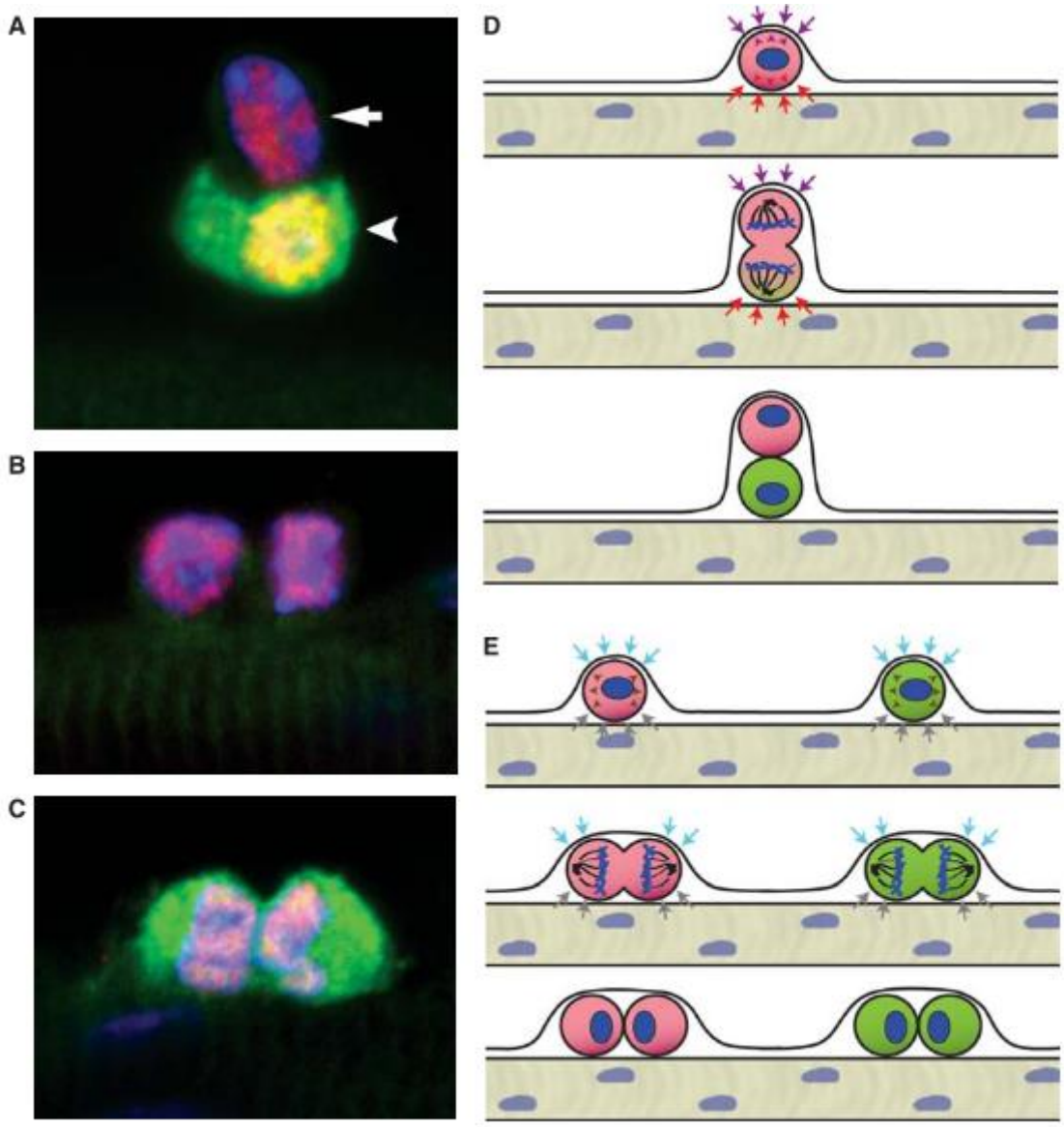


Figure 4. Asymmetric and symmetric cell division of MuSCs

(A, D) asymmetric division with the basal cell remaining in contact with the basal lamina and preserving the stemness (Pax7 in red), while the apical cell, in contact with the myofiber, entering the myogenic commitment program (Myf5 in green). (B, C, E) symmetric divisions for the clonal expansion of either uncommitted satellite cells (B; red cells) or committed myoblasts (C; green cells). Image from (Bentzinger et al., 2012)

1.3 Failure of the System: Muscular Dystrophies

The integrity of the skeletal muscle through repeated contraction cycles during an individual's life is maintained by the mechanisms and players that we discussed above and many others. Fundamental to muscle stability and functionality are the members of the dystrophin-glycoprotein complex (DGC), an intricate system of cytoskeletal and extracellular matrix proteins (**Figure 5**). The complete or even partial absence or dysfunction of one of these proteins often results in a fragile sarcolemma that is damaged upon physiological contraction causing muscle weakness, atrophy, fibrosis, and eventually muscle degeneration, in a group of inherited diseases called muscular dystrophies (MDs) (Allen et al., 2016).

Dystrophin is one of the main components of the complex; it is a long, modular, rod-shaped protein that acts as a bridge between the intra-sarcoplasmic actin cytoskeleton and extracellular matrix components, ultimately anchoring the myofiber to the connective tissue around it. Simultaneously, it serves as a scaffold for many other fundamental modules of the DGC, such as α- and β-dystroglycan, sarcoglycans, α-dystrobrevin, syntrophins, sarcospan, integrin, and biglycan, all with specific structural and signaling properties. The extracellular matrix, and in particular laminin and collagen, is directly contacted by α-dystroglycan, integrin, and biglycan proteins. α-dystroglycan is then connected to dystrophin through its intermembrane subunit β-dystroglycan. Integrin and biglycan interact with the intracellular α-dystrobrevin via the intermembrane sarcospan and sarcoglycans. Dystrophin and α-dystrobrevin are ultimately in contact one each other and both are docking platforms for syntrophins (**Figure 5**)(Allen et al., 2016).

It is thus straightforward that lack of even one component, if not compensated, can cause impairment of the entire system. This is indeed what happens, at various degrees of severity, in muscular dystrophies. In particular, different mutations in the very large, X-linked dystrophin gene (*DMD*) can cause either Duchenne Muscular Dystrophy (DMD) or Becker Muscular Dystrophy (BMD), with the first being the most severe manifestation of all MDs due to complete loss of dystrophin protein. Mutations in any of the sarcoglycan subunits (α, β, γ, δ) result in Limb-Girdle Muscular Dystrophies (LGMDs), with a complete loss of the entire tetramer in the absence of even one of the subunits (Allen et al., 2016). *mdx* (Bulfield et al., 1984) and *sgca*-null (Duclos et al., 1998) mice are the most commonly employed mouse models of MDs, recapitulating the mutations respectively for DMD and LGMD 2D.

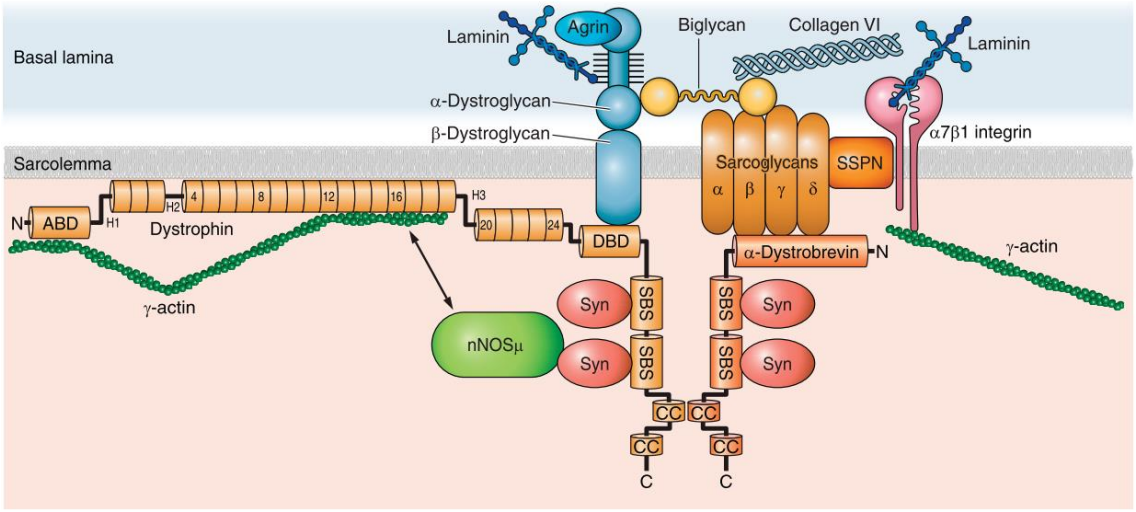


Figure 5. The Dystrophin-Glycoprotein Complex

Schematic representation of the DGC network in skeletal muscle cells. Actin binding domain (ABD, Dystroglycan binding domain (DBD), Syntropin (Syn), Syntropin binding subunit (SBS), sarcospan (SSPN). Image from (Allen et al., 2016)

In spite of the enormous amount of literature and therapy attempts, a cure for MDs (and particularly DMD) is still lacking. The problem is very complex and caused by several intervening factors that will be concisely outlined hereafter.

In absence of an intact DGC, the sarcolemma is fragile and contraction causes its tearing and increased permeability to calcium, also due to dysregulation of the stretch-activated ion channels (SACs or MSCs). Calcium is an important second messenger and this increased sarcoplasmic calcium concentration triggers a series of signaling reactions that ultimately activate cell death pathways and fibrosis (**Figure 6**). As a result, muscle contraction-induced stress generates constant and chronically overlapping cycles of degeneration and regeneration. Eventually, the entire environment is loaded with excessive reactive oxygen species (ROS) coming from different sources; this aggravates membrane permeability through lipid peroxidation and results in increased calcium income (**Figure 6**). ROS accumulation induces also mitochondrial dysfunction and DNA mutations, which contribute to activating cell death pathways (**Figure 6**)(Allen et al., 2016; Starosta and Konieczny, 2021).

Fast fibers are preferentially damaged by muscular dystrophies compared to slow fibers, probably due to the higher mechanical stress to which they are submitted and to their metabolism. Indeed, slow-twitch, oxidative fibers, which contain high levels of protective enzymes against ROS, are more protected from damage-induced oxidative stress and degeneration (Webster et al., 1988). Treatment of dystrophic mice with antioxidants, such as

cyanidin, has proven effective in ameliorating the degenerative signs of the pathology, in part through a muscle fiber-type switch towards type I oxidative fibers (Saclier et al., 2020a).

As already mentioned, the DGC is also a scaffold for signaling molecules; in particular, the dystrophin/α-dystrobrevin/syntrophins network anchors the neuronal Nitric Oxide Synthase (nNOS) enzyme to the sarcolemma. nNOS normally produces nitric oxide (NO) from L-arginine and NO is involved for example in exercise-induced vasodilation. nNOS sarcolemmal localization is impaired in several muscular diseases; the consequent mislocalization of NO from the sarcolemma to the cytoplasm aggravates the dystrophic phenotype (**Figure 6**)(Allen et al., 2016; Starosta and Konieczny, 2021).

NO and ROS levels are also influenced by immune cells that invade the tissue in large numbers in degenerating muscles and create a sustained inflammatory environment that ends up damaging the tissue rather than promoting regeneration (**Figure 6**). The chronic inflammation characteristic of dystrophic muscles exhibits important differences with the acute inflammation typical of muscle regeneration that was outlined in Chapter 1.2. The kinetics of the immune cells' waves is compromised, with the presence of a mixed population of macrophages expressing both TNFα and TGFβ since the early phases of inflammation. TNFα, normally released by M1 in the early phases of inflammation, induces the controlled apoptosis of Fibro-Adipogenic Progenitors (FAPs), which are the main responsible for the deposition of the ECM scaffold useful for regeneration. However, the concomitant presence of both pro-inflammatory and anti-inflammatory macrophages promotes the coexpression of TNFα and TGFβ, and this tips the balance toward FAPs survival and excessive ECM and adipocytes deposition, leading to fibro-fatty scars (**Figure 6**)Figure 6. (Lemos et al., 2015; Muñoz-Cánoves and Serrano, 2015; Theret et al., 2021).

This misregulated environment has a direct effect on MuSCs ability to repair the damage, causing improper activation, proliferation, and differentiation (**Figure 6**). Even in case of successful progression in the myogenic program and reformation of new myofibers, the latter will carry the same DGC aberrations that generated the problem in the first place. Furthermore, recent evidence suggests intrinsic MuSCs dysfunction in absence of dystrophin, specifically affecting their cell polarity and consequently self-renewal ability (Dumont et al., 2015c). A vicious cycle of degeneration/regeneration is thus triggered, leading to progressive replacement of the muscle tissue with adipose and fibrotic tissues (**Figure 6**)(Chang et al., 2016).

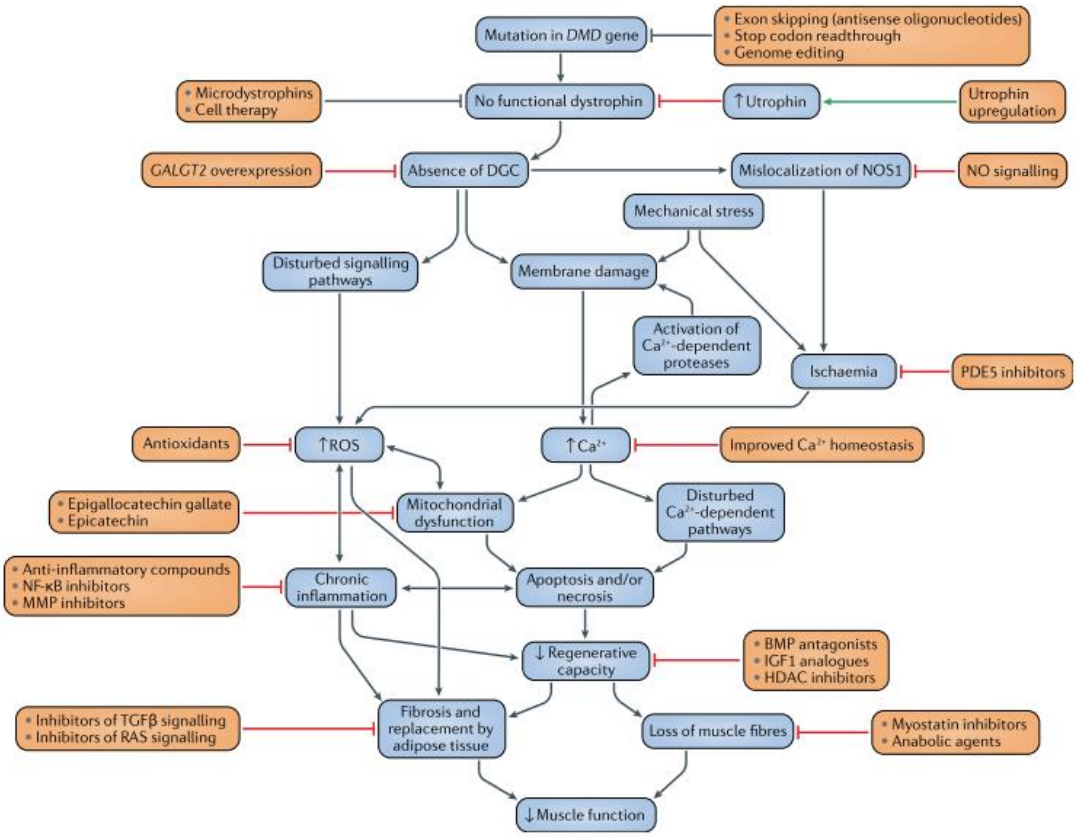


Figure 6. Primary and secondary effects of DGC impairments

Cascade of primary and secondary effects caused by mutations in the *DMD* gene and possible therapeutic approaches. BMP, bone morphogenetic protein; DGC, dystrophin-glycoprotein complex; GALGT2, the gene encoding β1,4 N- acetylgalactosaminyltransferase 2 (also known as B4GALNT2); HDAC, histone deacetylase; IGF1, insulin-like growth factor 1; MMP, matrix metalloproteinase; NF- κB, nuclear factor-κB; NO, nitric oxide; NOS1, nitric oxide synthase; PDE5, phosphodiesterase 5; ROS, reactive oxygen species; TGFβ, transforming growth factor-β. Image from (Verhaart and Aartsma-Rus, 2019).

In DMD patients (approximately 1 in 3,500 males), this cascade of events leads to the appearance of the first symptoms (mainly muscle weakness) between the ages of 2 and 6 years. Between 7 and 12 years, atrophy becomes prominent and most boys are confined to a wheelchair. Mean survival is currently around 25 years, with respiratory failure being the most common cause of death (Allen et al., 2016).

1.4 Nuclear Factor I family

Nuclear Factor I (NFI) proteins are a family of evolutionarily conserved transcription factors found in vertebrates, with single *NFI* gene ancestors found also in *C. elegans* and *D. melanogaster*. Four distinct paralogous genes have been identified in vertebrates: *Nfia*, *Nfib*,

Nfic, and *Nfix* (Apt et al., 1994; Fletcher et al., 1999; Gil et al., 1988; Jones et al., 1987; Santoro et al., 1988). They bind as dimers with high affinity to the dyad symmetric consensus motif 5'-TTGG(A/C)(N5)GCCAA-3', but they can also specifically bind to half-sites (TTGGC or GCCAA), even though less tightly (Gil et al., 1988; Kruse and Sippel, 1994; Nowock et al., 1985). While they have a very similar DNA-binding specificity, diversity is set by multiple genes, splicing isoforms, and homo versus heterodimerization. The first about 200 amino acidic residues at the N-terminal constitute a domain sufficient for DNA-binding and dimerization and is highly conserved among all NFI proteins. In this region, four very well-conserved cysteine residues were identified as necessary for DNA-binding activity (Novak et al., 1992): mutations of three of these Cys residues to several other amino acids abolish DNA-binding activity of NFI. The fourth cysteine is not strictly required for DNA-binding but is an oxidation-sensitive cysteine that may be required for redox modulation of NFI DNA-binding (Bandyopadhyay and Gronostajski, 1994). The C-terminal, proline-rich, portion represents the transcription modulation domain of NFIs. The divergence between *NFI* paralogous genes occurs in their C-terminal part, further accentuated by the presence of splicing sites in this portion of each gene. On the other hand, orthologous *NFI* genes and their splicing variants are highly conserved across vertebrate species (Apt et al., 1994; Chaudhry et al., 1998; Fletcher et al., 1999; Gil et al., 1988; Gronostajski, 2000; Gründer et al., 2003; Jones et al., 1987; Mermod et al., 1989; Qian et al., 1995; Santoro et al., 1988).

All four *NFI* genes play important and distinct roles during mammalian embryogenesis and are involved in different types of cancers and diseases. Their expression patterns during mouse development are specific and sometimes overlapping, and the defects observed in *Nfi* gene-deficient mice suggest common but also unique and nonredundant functions in different organs during development (Chaudhry et al., 1997; Piper et al., 2019; Zenker et al., 2019). Nfia is the first gene of the family to be expressed during mouse embryogenesis, at 9 dpc, mainly restricted to the heart and brain primordia. At 10.5 dpc, Nfib appears in lung buds and Nfic in developing heart and brain. At 11.5 dpc, also Nfix starts to be expressed in the brain, central and proximal mesenchyme, and myotomes. In adult tissues, their expression patterns are in part overlapping and in part unique (**Figure 7**)(Chaudhry et al., 1997). Accordingly to their expression patterns (**Figure 7**), *Nfia*, *Nfib,* and *Nfix* knockout mice display brain anomalies such as megalencephaly, enlarged ventricles and/or hydrocephalus, malformation of the hippocampus, and dysgenesis of the corpus callosum; $Nfia^{-/-}$ mice reveal also defects in kidney and urinary tract; $Nfib^{-/-}$ mice show abnormal lung development with immature lungs and respiratory defects; $Nfix^{-/-}$ mice have defects in bone

maturation and reduced and disorganized musculature (Harris et al., 2015; Piper et al., 2019; Zenker et al., 2019).

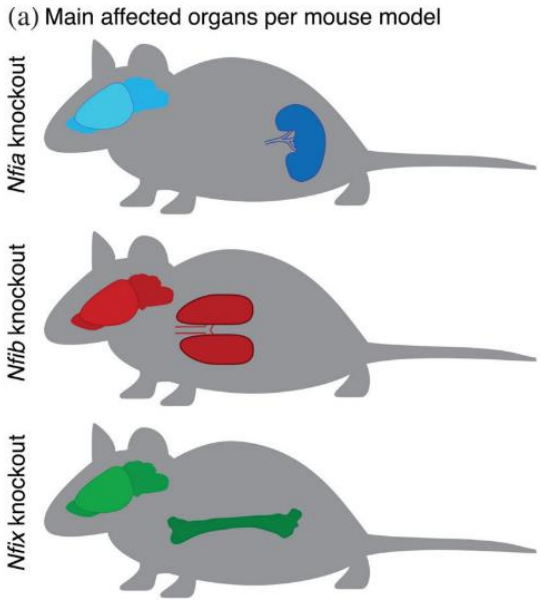


Figure 7. Depiction of the organs mainly affected in Nfi knockout mouse models
Image from (Zenker et al., 2019)

1.4.1 Nfix

Nfix (nuclear factor I/X) is coded by chromosome 8 in *Mus musculus* and its transcription unit encompasses 11-18 exons and encodes four main protein splicing isoforms generated by differential splicing of exons 7 and 9 and the use of alternative transcription initiation sites; they are indicated as Nfix1, Nfix2, Nfix3, and Nfix4. Nfix3 is the longest isoform, retaining both exon 7 and 9, and is mainly expressed in the nervous system. The other isoforms are expressed in a multitude of other tissues, both in the adult and during development. The skeletal muscle mainly expresses the isoform lacking both spliced exons (Nfix2), but also isoform 1 that retains exon 7 (Messina et al., 2010). In *Homo sapiens*, NFIX gene is encoded on the plus strand of chromosome 19 and its structure presents high homology with the murine one. Nfix activity is tissue- and cell-specific, due both to preferential isoforms expression and different interacting partners (Chaudhry et al., 1997; Fletcher et al., 1999; Fraser et al., 2017; Piper et al., 2019).

As revealed by Tabula Muris Atlas (Schaum et al., 2018), Nfix expression is abundant in staminal cells of several tissues in the adult mouse. Accordingly, Nfix activity has been linked

to progenitor quiescence, proliferation, and differentiation in a variety of tissues and cell types.

Recent studies report that NFI proteins can also act as epigenetic regulators in development and cancer as their DNA binding colocalizes with open, active chromatin and active enhancers (Denny et al., 2016; Fane et al., 2017; Hiraike et al., 2017; Martynoga et al., 2013) but also to chromatin boundaries (Pjanic et al., 2013). These are common characteristics for transcription factors active as master switch regulators (Tapscott, 2005).

1.4.1.1 Nfix in Myogenesis, Muscle Regeneration, and Muscular Dystrophy

In skeletal muscle, Nfix is robustly upregulated in the fetal period (Biressi et al., 2007a), where it acts as a transcriptional switch of the embryonic-to-fetal myogenic transition by exerting both activatory and repressive roles. Indeed, its premature expression in embryonic muscles is sufficient to imprint the tissue for the fetal transcriptional program, whereas its ablation confers embryo-like characteristics to the fetal muscles (Messina et al., 2010). In this context, Nfix represses the embryonic, slow-twitching myosin isoform MyHC-I (*myh7*) by a dual mechanism: in concert with Sox6, by facilitating its inhibitory interaction with the *myh7* promoter in fetal myotubes (Taglietti et al., 2016); it also binds to the promoter of *Nfatc-4* and represses its activity in fetal muscles, impeding its induction of the *myh7* promoter (Messina et al., 2010) (**Figure 8**). During the switch to fetal myogenesis, it interacts with Mef2a and Pkcθ; this bridge promotes the phosphorylation of Mef2a by Pkcθ and eventually the transcription of the fetal gene *mck* (muscle creatine kinase). (**Figure 8**) During the transition, Nfix also promotes the transcription of another fetal-specific gene, *eno3* (β-enolase), which is expressed at higher levels in fast-twitch fibers than in slow-twitch fibers and contains NFI recognition motifs in its promoter (Messina et al., 2010) (**Figure 8**). A corresponding role is exerted in zebrafish (*Danio rerio*) by the Nfix ortholog Nfixa, which promotes the fast-twitching program in the zebrafish embryos (Pistocchi et al., 2013).

At the onset of secondary myogenesis, the expression of Nfix is promoted by the AP-1 family transcription factor JunB, whose role is antagonized during primary myogenesis by the RhoA/ROCK signaling axis that inactivates the kinase ERK resulting in lower expression of JunB and consequently of Nfix (Taglietti et al., 2018) (**Figure 8**).

Nfix expression is detected also in adult myonuclei and $Pax7^+$ MuSCs, where it regulates the proper regeneration timing through modulation of Myostatin (Rossi et al., 2016), a secreted factor of the TGF-β superfamily and a critical regulator of prenatal and postnatal myogenesis. Indeed, *Nfix*-null adult mice manifest a reduced muscle weight and cross-sectional area of

myofibers, accompanied by increased expression of MyHC-I, consistent with prenatal observations. Nfix absence in MuSCs causes a delay in their differentiation, as evidenced by a reduced number of myogenin-positive cells in the first 72h hours of differentiation, and this is reflected in a delayed regeneration of CTX-injured Nfix-deficient muscles. This effect is dependent on Myostatin direct repression operated by Nfix in differentiating myotubes and regenerating myofibers (Rossi et al., 2016).

In regenerating muscle, Nfix is also expressed by macrophages, where it supports the conversion to the M2 anti-inflammatory phenotype following phagocytosis (Saclier et al., 2020b). This highlights the manifold importance of Nfix in the process of muscle regeneration. Consistent with Nfix regulation in myoblasts (Taglietti et al., 2018), Nfix expression is regulated by the RhoA-ROCK axis in macrophages too, where this pathway exerts the same inhibitory effect on Nfix (Saclier et al., 2020b).

All the discussed tasks delivered by Nfix in developing and regenerating muscles are the key to interpreting its part in the battle against Muscular Dystrophy (MD). As counterintuitive as it may initially seem, the delayed regeneration prompted by the absence of Nfix becomes a beneficial consequence in the chronically injured dystrophic muscle (Rossi et al., 2017). This effect can be explained in different ways: a slower regeneration rhythm opposes the fast depletion of MuSCs caused by their continuous activation and impaired self-renewal posed by the chronic inflammatory environment of MDs; the conversion to a slow-twitching phenotype and a more oxidative metabolism also preserves the musculature from the contraction-dependent damage and increase in ROS production observed in MDs (Rossi et al., 2017); furthermore, the impaired M1-to-M2 switch in the absence of Nfix limits the deposition of extracellular matrix (ECM) thus reducing fibrosis (Saclier et al., 2022).

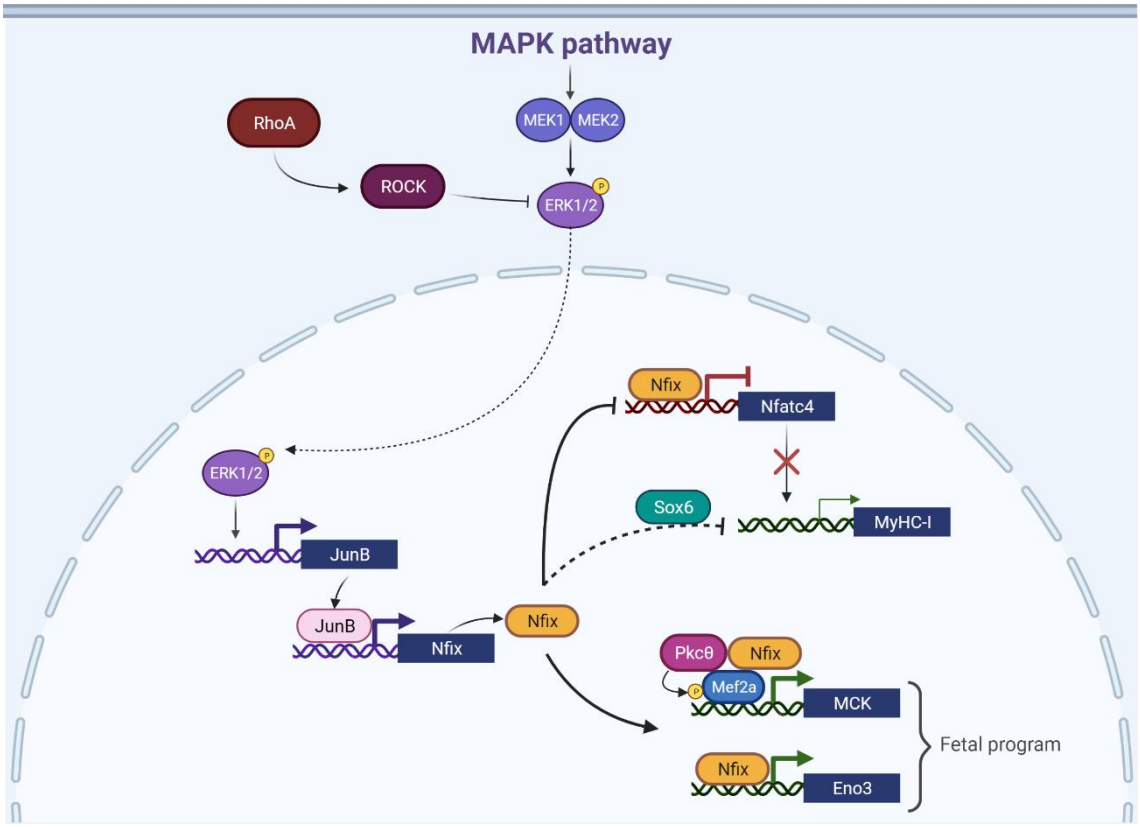


Figure 8. Nfix network in developing muscle

1.4.1.2 Nfix in Other Developing and Adult Tissues

In the central nervous system, Nfix is expressed in neural stem cells (NSCs) populations of several nervous system structures including embryonic and adult cerebellum, hippocampus, and neocortex. Here, it mediates the balance between cell cycle entry and exit, thereby regulating the homeostasis that exists between progenitor pool expansion and tissue regeneration. In particular, Nfix was found to promote a quiescent state in cultured NSCs (Martynoga et al., 2013), induce cell-cycle exit of retinal progenitor cells and generation of late-born retinal cell types (Clark et al., 2019), mediate radial glia differentiation via repression of progenitor self-renewal pathways (Heng et al., 2014), and control timely differentiation of hippocampal progenitors (Harris et al., 2013; Heng et al., 2014) and cerebellar granule neuron precursors (Fraser et al., 2017, 2019; Piper et al., 2011). As a result, a prolonged period of progenitor cell self-renewal during nervous system development is observed in the absence of Nfix, resulting in hydrocephaly (Campbell et al., 2008; Driller et al., 2007). During gliogenesis, Nfix promotes the differentiation of neural progenitor cells into astrocytes, while repressing the oligodendrocyte lineage (Matuzelski et al., 2017; Zhou et al., 2015). Nfix also regulates both proliferation and migration during the development of the murine

subventricular zone neurogenic niche (Evelyn Heng et al., 2015).

Nfix-null animals show also severe skeletal defects due to delay in endochondral bone ossification and decreased mineralization, with persistent cartilaginous materials and presence of anomalous chondrocyte-like cells in vertebral bodies and femurs. In the latter, the delayed bone development was characterized by an enlarged zone of resting chondrocytes and a reduced zone of proliferating and hypertrophic chondrocytes (Driller et al., 2007).

Nfix depletion affects also hematopoiesis, as it is expressed in adult hematopoietic stem and progenitor cells (HSPCs) where it regulates HSPCs colony-forming potential, survival, and repopulating activity post-transplantation (Holmfeldt et al., 2013). Later during hematopoiesis, Nfix guides B lymphopoiesis and myelopoiesis, influencing both survival of hematopoietic stem cells and lineage commitment, by promoting the myeloid lineage from total bone marrow cells while blocking the B-cell commitment (HALL et al., 2018; O'Connor et al., 2015).

Nfix, together with Nfib, was also shown to be crucial for hair follicle stem cell identity maintenance, acting at the epigenetic level on stem cell-specific super-enhancers. Here, the two transcription factors keep certain chromatin domains open while silencing others, acting as regulators of fate control. Hence, their absence causes unwanted lineages commitment (Adam et al., 2020).

1.4.1.3 Nfix in Cancer and Human Diseases

NFIX is directly involved to various degrees in several types of human diseases, including different kinds of cancer and two genetic disorders namely Marshall-Smith Syndrome and Malan Syndrome.

CANCER

The role of NFIs in cancer is documented as very heterogeneous, with all four family members acting as either tumor suppressors or oncogenes depending on the cancer model. Their involvement in malignancies has been linked to their ability to regulate cell cycle entry and exit that determines the balance between progenitor cells expansion and differentiation. Increasing evidence indicates their contribution to epigenetic control as a bridge between their role in development and cancer progression (Chen et al., 2017; Denny et al., 2016; Fane et al., 2017).

NFIX generally appears to act as a tumor suppressor, with lower expression implying poor

prognosis. However, the picture is more complicated, since in some tumors it can also behave as an oncogene (Fane et al., 2017; Li et al., 2020). Recently, NFIX was shown to promote malignant glioblastoma, a very aggressive brain tumor where NFIX is significantly overexpressed and enhances the cancer cells migration via upregulation of Ezrin (Liu et al., 2020). A similar role is observed in lung cancer, where NFIX was reported to induce cancer cells' proliferation, migration, and metastasis (Rahman et al., 2017). On the contrary, NFIX inhibits the migration of cancer cells in both esophageal squamous cell carcinoma and colorectal cancer (Liu et al., 2017; Mao et al., 2015). In colorectal cancer, miR-647/1914 promote the proliferation and migration of cancer cells by directly targeting NFIX (Liu et al., 2017) while in esophageal squamous cell carcinoma the same effect is delivered by miR-1290 (Mao et al., 2015). NFIX acts as a tumor suppressor also in triple-negative breast cancer, where it likely participates in the suppression of the metastasis-promoting gene ADAM-12 (Ray et al., 2013), and in lung adenocarcinoma (Ge et al., 2018).

MARSHALL-SMITH SYNDROME

The Marshall-Smith syndrome (MRSHSS; MIM# 602535), first described in 1971 (Marshall et al., 1971), is caused by *de novo* heterozygous mutations producing splice-site variants or frameshifts scattered through exons 4-10 of *NFIX* gene (**Figure 9**), leading to proteins with abnormal C-terminus but preserved DNA binding and dimerization domain (Malan et al., 2010; Martinez et al., 2015). Some variants cause *NFIX* mRNA to escape nonsense-mediated decay and the patients retain both WT and mutant transcripts, suggesting a dominant-negative/gain-of-function effect of these mutant NFIX proteins (Malan et al., 2010; Schanze et al., 2014; Zenker et al., 2019). The phenotype is characterized by dysostosis, postnatal failure to thrive, short stature, unusual face, respiratory compromise, intellectual disability, abnormal bone maturation, and moderate to severe developmental delay (Adam et al., 2005; Marshall et al., 1971; Shaw et al., 2010). MRSHSS is a very rare disorder, with less than 60 cases reported in the literature to date (Orpha.net). Currently, there is no specific cure or treatment for MRSHSS, and early mortality (usually by three years of age) is mainly caused by respiratory complications, even though airway support and feeding management increasingly allow survival into adulthood. The respiratory problems are in part caused by a decreased or relaxed muscle tone of the larynx and trachea, making them softer than usual and prone to collapse and obstruct airways (Adam et al., 2005; Diab et al., 2003; Marshall et al., 1971; Mulder et al., 2020; Roodhooft and Acker, 1988; Shaw et al., 2010; Sperli et al., 1993).

MALAN SYNDROME

The Malan syndrome (MS; MIM# 614753; also called Sotos syndrome 2) is an overgrowth disorder caused by heterozygous *de novo* mutations (deletions or point mutations) clustered in exons 2/3 (the DNA-binding/dimerization domain) of *NFIX* gene (**Figure 9**)(Klaassens et al., 2015; Malan et al., 2010; Martinez et al., 2015; Priolo et al., 2012, 2018; Yoneda et al., 2012). In contrast with MRSHSS mutations, Malan syndrome patients show the expression only of the WT allele, while the mutant allele transcript undergoes nonsense-mediated decay, causing haploinsufficiency (Malan et al., 2010). The phenotype is less severe than MRSHSS and is defined as Sotos-like due to the resemblance with Sotos syndrome but lacking the *NSD1* gene mutations. It comprises intellectual disabilities, behavioral anxiety, impaired vision, unusual craniofacial characteristics such as long, triangular face, macrocephaly, prominent forehead, everted lower lip, and prominent chin, and musculo-skeletal anomalies such as abnormal bone maturation, slender habitus, kyphoscoliosis, and pectus excavatum/carinatum (Klaassens et al., 2015; Malan et al., 2010; Priolo et al., 2018; Yoneda et al., 2012).

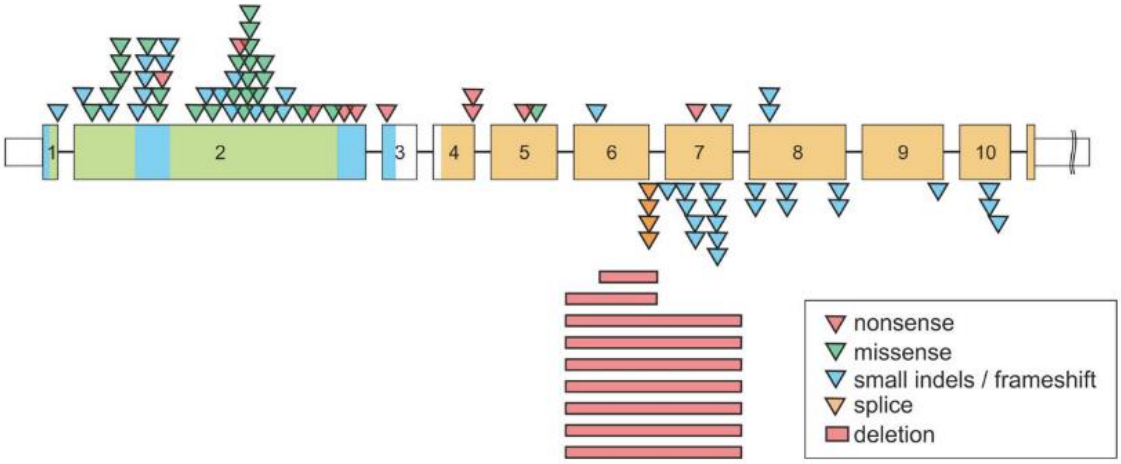


Figure 9. NFIX gene mutations in Malan and Marshall-Smith syndromes
Above the gene are depicted the Malan-specific mutations, while underneath the gene the Marshall-Smith-specific mutations. The DNA binding and dimerization domains are highlighted in blue, in green the MAD homology 1 (MH1) domain, and in orange the C-terminal transactivation/repression CAAT-box transcription factor-nuclear factor I (CTF-NFI) domain. Image from (Martin Zenker et al.).

Overall, despite the involvement of the same gene, Malan syndrome and Marshall-Smith syndrome seem to be two separate entities. This could be explained by the involvement of two different functional domains of Nfix protein and by the dominant-negative properties

of some mutations. Likewise, the twofold nature of Nfix as both oncogene and tumor suppressor and the different effects on cell cycle progression and differentiation during the development of different tissues may suggest a complex, context-dependent regulation of this transcription factor. Interestingly, large-scale proteomic studies gathered in the PhosphoSitePlus database (Hornbeck et al., 2012) evidenced the presence of several post-translational modifications in different domains of Nfix protein, particularly phosphorylation and ubiquitination, which roles on Nfix regulation have not been investigated so far and whose dysregulation could contribute to these diversified phenotypes.

1.5 Post-translational modifications in the NFI Family

One of the main mechanisms through which the cell machinery obtains complexity and diversification of protein function and regulation is represented by post-translational modifications (PTMs). These regulatory devices ensure rapid and dynamic cellular responses to the environmental and intracellular factors.

The presence of terminal GlcNAc (N-Acetylglucosamine) residues on NFI proteins was first identified in 1988 by Jackson et al. (Jackson and Tjian, 1988). In the following years, glycosylation was confirmed for both NFIC (Kane et al., 2002) and NFIB (Mukhopadhyay and Rosen, 2007). In particular, NFIC N-glycosylation is triggered in early involution of the mammary gland (Kane et al., 2002) and NFIB O-linked glycosylation at its C-terminal domain was shown to alter its transcriptional activity (Mukhopadhyay and Rosen, 2007).

Early *in vitro* studies identified differential NFI phosphorylation states between quiescent and proliferating 3T3-L1 adipocytes and this was suggested to be downstream of c-Myc, a key regulatory protein of proliferation and differentiation. In this study, c-Myc-promoted NFI phosphorylation correlated with a lower transcription rate of NFI-dependent promoters, but without affecting their DNA-binding affinity (Yang et al., 1993). Further studies on 3T3-L1 adipocytes identified NFI phosphorylation mediated by insulin as a potential mechanism regulating the downstream gene expression, particularly of the GLUT4 glucose transporter (Cooke and Lane, 1999). However, from these studies, it is not clear which NFI member is responsible for this mechanism and which specific phosphorylation sites are involved. More recently, Nfix serine 265 was identified as a potential target of insulin-induced phosphorylation in adipocytes (Rabiee et al., 2018). However, this regulatory connection was not further investigated.

The only additional data about Nfix PTMs derive from large-scale, high-throughput proteomics data and are available at different databases such as PhosphoSitePlus® www.phosphosite.org (Hornbeck et al., 2012), and PHOSIDA http://www.phosida.com (Gnad et al., 2007). However, most of these modifications (particularly phosphorylation) have not been thoroughly validated and investigated.

1.6 Protein Phosphorylation

Protein phosphorylation is a widely employed and very well-documented posttranslational modification that can alter a broad variety of biological functions and mechanisms. Almost 300,000 phosphorylation sites have now been mapped in the biological proteome (Hornbeck et al., 2012). When it comes to transcription factors, phosphorylation can directly regulate their activity through modulation of cellular localization, protein stability, protein-protein interactions, and DNA binding. Phosphorylation is particularly suitable for transcription factor control due to its rapid kinetic, simplicity, flexibility, and reversibility that allows rapid alteration in their activities in response to extracellular and intracellular stimuli (Cohen, 2002; Hunter, 2012; Whitmarsh and Davis, 2000).

Phosphorylation is accomplished by protein kinases, enzymes able to add phosphate groups (PO_4), with ATP as a phosphoryl donor, to different amino acidic residues, typically the hydroxyamino acids serine, threonine, and tyrosine in eukaryotic cells. It is estimated that there are approximately 500 kinases encoded by the genome, corresponding to about 1.7% of the entire human genome. The reverse reaction is catalyzed by phosphatases, which remove the phosphate groups (Cohen, 2002; Hunter, 2012; Whitmarsh and Davis, 2000). The phosphorylation of a protein is determined by a consensus sequence surrounding the phosphoacceptor site and requires direct binding of the protein kinase to the substrate. Since these consensus sequences are highly common (found in ~80% of all proteins) and limited to a few aminoacids, specificity is often conferred by additional sequences, termed docking domains. One example is the MAP kinases docking domain termed the δ or D domain (Kallunki et al., 1996; Sharrocks et al., 2000).

Phosphorylation sites are commonly found in flexible or unstructured regions of proteins; the phosphoryl group adds negative charges, installs strong hydrogen bond acceptor oxygens, and sharply increases the size of the residue, ultimately perturbing the peptide local structure, altering global protein conformation (Cohen, 2002; Hunter, 2012). The outcome of phosphorylation can have opposing directions depending on the target, kinase, and even

phosphorylation site of the same protein: it can either activate or repress transcriptional activity, promote or inhibit protein degradation, guide shuttling from or to the nucleus, facilitate or impede protein-protein interactions and DNA binding (Whitmarsh and Davis, 2000). Different phosphorylation events on a number of residues within the same transcription factor (multisite phosphorylation) may act either cooperatively or antagonistically to regulate single or multiple functions, integrating stimuli from different signaling pathways.

1.6.1.1 The myogenic kinome: MAPK and Insulin/IGF pathways

Myogenesis, both prenatal and postnatal, is a complex process governed by numerous stimuli; among them, kinase hierarchies and phosphorylation cascades, prompted by the interaction of growth factors with their receptors present in myogenic cells. Some of these circuits will be discussed hereafter.

EXTRACELLULAR SIGNAL-REGULATED KINASE (ERK)

ERK1 and ERK2 (ERK1/2) are two serine/threonine kinases belonging to the MAP kinases whose activity is regulated by growth factors and their binding to their tyrosine kinase receptors. The leading MAP kinase of this signaling pathway is Raf that, once activated by GTP-Ras, phosphorylates and activates MEK, the ERK kinase. ERK activity is critical in myoblasts' proliferation but also in the late differentiation, during the fusion process. Among the mitogens known to activate ERK we find fibroblast growth factor (FGF), hepatocyte growth factor (HGF), and insulin-like growth factor (IGF). During the early phase of MuSCs activation, HGF released from the ECM activates the ERK signaling through interaction with the c-Met receptor expressed by MuSCs and promotes their activation and proliferation. Later on, during myoblasts proliferation, ERK activity is high, stimulated mainly by FGF and their binding to FGF receptors (FGFRs) that activate the Ras/MAPK signaling; with the input of differentiation myocytes progressively lose FGFRs with a consequent decrease in ERK activity. However, this drop in ERK activity, particularly of ERK2, is only temporary, as it rises again to promote myocytes fusion. However, the exact mechanisms of this switch in ERK activity are not well elucidated (Knight and Kothary, 2011; Yin et al., 2013).

p38 MAPK

p38 proteins are a family of serine/threonine kinases that exist in four isoforms: p38α, p38β, p38γ, and p38δ. As for ERK, to which are closely related, their activity is regulated by mitogens. However, their activity is induced during differentiation; specifically, p38α is

essential for primary myoblasts differentiation and this seems to be mediated at least in part by myoblasts cell-cell contact with the involvement of N-cadherin and the GTPase Cdc42, which phosphorylates p38. Another activatory circuit passes through the transforming growth factor β-activated kinase 1 (TAK1) that, upon transforming growth factor (TGF) stimulus, phosphorylates the MAP kinase kinase 3/6 (MKK3/6), which ultimately phosphorylates p38 and promotes myoblasts differentiation (Knight and Kothary, 2011). p38 regulation of differentiation is in part mediated also by reactive oxygen species (ROS), that serve as secondary messengers. Indeed, ROS, which are increased upon differentiation stimuli due to a shift towards oxidative phosphorylation metabolism, induce p38 phosphorylation and promote differentiation and regeneration (L'Honoré et al., 2018). Phosphorylation of p38α/β may also act as a molecular switch to activate the quiescent satellite cells, as it can be detected very early upon myofiber isolation (Jones et al., 2005).

INSULIN/IGF PATHWAY: IGFR, PI3K, AKT, GSK3β, mTOR, S6K

The Insulin/IGF pathway is a fundamental stimulator of myoblasts' proliferation, differentiation, and hypertrophy. IGF and insulin bind to their tyrosine kinase receptors (IGFRs and IRs) stimulating a phosphorylation cascade that starts with the lipid kinase phosphatidylinositol 3-kinase (PI3K). The membrane-bound phosphatidylinositols PI(4)P and PI(4,5)P2 are phosphorylated by PI3K into (PI(3,4)P2) and (PI(3,4,5)P3), this event recruits PDK1, PDK2, and Akt to the membrane, where Akt can be activated by PDK1 and 2. The mammalian target of rapamycin (mTOR) is a major target activated by Akt and its function stimulates protein translation and hypertrophy of muscle fibers through activation of the ribosomal protein S6 kinase 1 (S6K). Another kinase, glycogen synthase kinase 3β (GSK3β), which impedes differentiation and hypertrophy, is deactivated by Akt-dependent phosphorylation. The outcome is a push towards muscle differentiation and hypertrophy. While this represents the canonical pathway, some reports show that the IGF-1 and Akt1 isoforms can be critical also for myoblasts' proliferation in subconfluent conditions, with another isoform, Akt2, antagonizing this function by driving differentiation upon cell-cell contacts (Knight and Kothary, 2011; Yin et al., 2013).

2 Aim of the work

Nfix transcription factor has manifold roles in tissue development and adult homeostasis. In skeletal muscle, Nfix is expressed starting from the 14th day of embryonic development (E14), when the embryo is entering the fetal phase of myogenesis, characterized by a population of myogenic progenitors called fetal myoblasts (Biressi et al., 2007a; Messina et al., 2010). In this context, Nfix imprints the fetal transcriptional program inducing the expression of fetal-specific genes while concurrently repressing genes specific to the previous phase, the embryonic myogenesis (Messina et al., 2010; Taglietti et al., 2016). Once the adult skeletal muscle is formed, Nfix continues to give its contribution in adult muscle stem cells, or MuSCs, where it is expressed in both proliferating and differentiating myoblasts and regulates the timing of muscle regeneration (Rossi et al., 2016). Importantly, dystrophic mice that lack Nfix expression display functional and morphological ameliorations. This effect is in part due to delayed regeneration and in part to the switch to a more oxidative, slow-twitching muscle fibers phenotype (Rossi et al., 2017).

However, as of today, we lack detailed information regarding Nfix upstream and downstream signaling in skeletal muscle. Several studies in many Nfix-regulated systems focused on the transcriptional network of this protein. Instead, very few studies specifically interrogated the post-translational modifications of Nfix protein. Nevertheless, PTMs could to some extent explain the diversified functions of Nfix in different tissues and developmental stages and even the different pathological outcomes of those diseases in which Nfix is involved.

With this study, we aim to address this knowledge gap. Particularly, we question the presence and potential biological role of PTMs on the Nfix protein, with a focus on phosphorylation.

3 Results

3.1 Nfix2 is post-translationally modified in myoblasts

One of the main ways through which the cell machinery obtains complexity and diversification of protein function is represented by post-translational modifications (PTMs). Nfix regulatory mechanisms are poorly characterized in skeletal myoblasts and the presence of PTMs has not been thoroughly investigated. However, some observations suggested us the presence of such modifications.

When analyzed through SDS-PAGE followed by western blot, Nfix appears as multiple electrophoretically distinct polypeptide species in skeletal myoblasts and several other cell types, except for macrophages (**Figure 10A**). These bands may represent different isoforms of Nfix. The Nfix2 isoform, one of the two mainly expressed isoforms in the mouse skeletal muscle tissue, has a predicted molecular weight of 45 kDa. The other isoform expressed in mouse skeletal muscle is isoform 1, which is predicted to be 49 kDa. The isoform 3 is a 54 kDa protein, but it is expressed at very low levels in skeletal myoblasts (Messina et al., 2010). We verified the electrophoretic mobility of the isolated Nfix2 isoform by transfecting C2C12 cells with a plasmid encoding for Nfix2 isoform N-terminally tagged with the hemagglutinin (HA) protein. As observed in **Figure 10B**, the Nfix2 isoform alone represents at least two of these Nfix-specific bands, ruling out the splicing isoforms hypothesis at least for these Nfix species.

To test the presence of PTMs, we compared the SDS-PAGE migration of Nfix expressed by the prokaryotic organism *Escherichia coli* versus myoblasts. *E. coli* is not equipped with the machinery necessary for most eukaryotic post-translational modifications; for this reason, it is a suitable organism for our study. As observed in **Figure 10C**, this experiment reveals that a virtually unmodified Nfix2 protein migrates approximately at its predicted molecular weight of 45 kDa, after cleavage of the GST-tag. Instead, in C2C12 myoblasts all Nfix-specific bands (as verified via shNfix treatment) appear at higher molecular weights (**Figure 10D**). These data point out that Nfix expressed in myoblasts potentially undergoes multiple types of post-translational modifications.

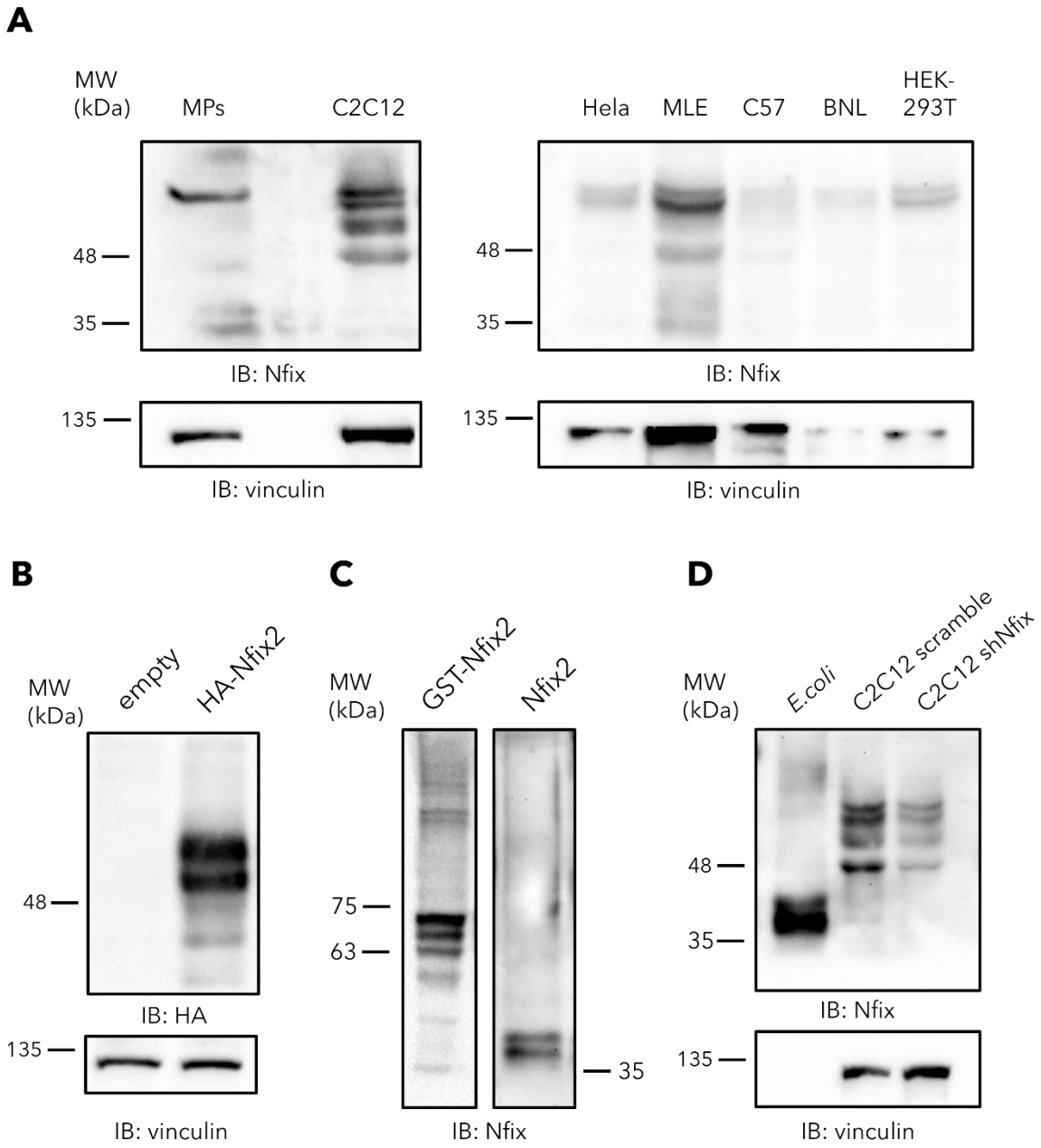


Figure 10. Nfix is post-translationally modified

A) Western blot analyses of different cell types reveal multiple Nfix-positive bands. Macrophages (MPs); Murine myoblast cell line (C2C12); Henrietta Lacks cells (HeLa); Murine Lung Epithelial cell line (MLE); embryonic liver cell line (BNL); Human Embryonic Kidney cell line (HEK-293T). Vinculin is used as a loading control.

B) Western blot showing empty and HA-Nfix2-overexpressing C2C12 immunoblotted with an anti-HA antibody that reveals two Nfix2-specific immunoreactive bands. Vinculin is used as a loading control.

C) Western blot from extracts of GST-Nfix2 expressing *E. coli* before and after cleavage of the GST-tag.

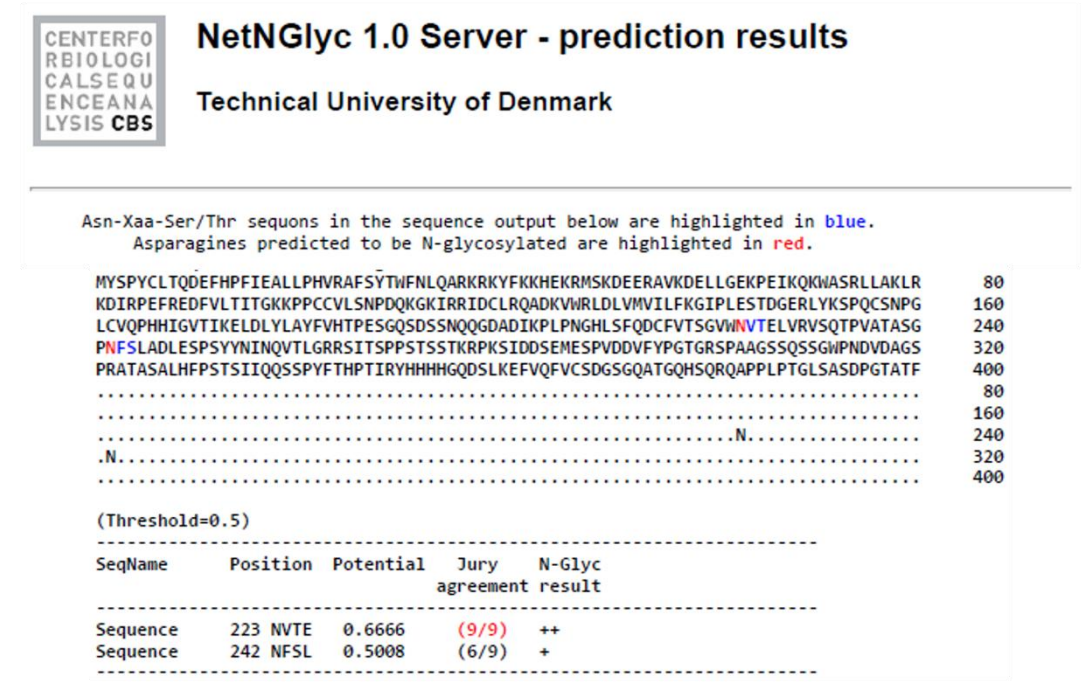
D) Western blot of Nfix2-expressing *E.coli* and wildtype and shNfix C2C12 protein extracts. Vinculin is used as a loading control for C2C12 extracts.

3.2 Nfix is not N-glycosylated in C2C12 myoblasts

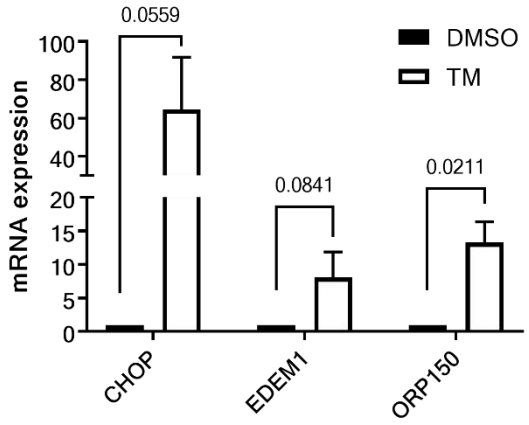
In the plethora of possible PTMs, we investigated N-glycosylation as this is a bulky modification that could explain the important molecular weight shift detected for Nfix. Furthermore, it was one of the few PTMs observed in another NFI member, NFIC (see Chapter 1.5).

Firstly, we performed an *in silico* prediction to evaluate the presence of N-glycosylation sites on the Nfix protein sequence. The analysis highlighted two N-glycosylation motifs on the C-terminal domain of the protein (**Figure 11A**). We thus treated C2C12 myoblasts with the N-glycosylation inhibitor Tunycamicin, verified the treatment efficacy via real-time qPCR of Tunycamicin-responsive genes (**Figure 11B**), and analyzed the protein extracts via western blot. As shown in **Figure 11C**, no detectable changes in Nfix bands' mobility were observed. Hence, we conclude that the mobility shift is not due to the N-glycosylation of the Nfix protein and that Nfix is not N-glycosylated in myoblasts.

A



B



C

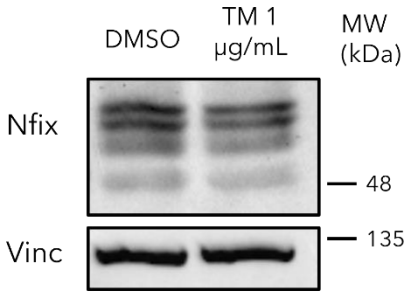


Figure 11. Nfix is not N-glycosylated

A) N-glycosylation in silico prediction for the Nfix2 protein sequence showing two potential N-glycosylation motifs using the NetNGlyc 1.0 prediction server.

B) Quantification of ER-stress marker genes activated by Tunycamycin treatment (1 µg/mL) as a readout of the treatment efficacy. *β-actin* was used as the housekeeping gene.

C) Western blot of protein extracts from control and Tunycamycin (1 µg/mL) treated C2C12 cells showing Nfix-specific signal. Vinculin was used as the loading control.

3.3 Nfix is phosphorylated at Ser 301 in skeletal myoblasts *in vitro*

Phosphorylation is one of the most studied and widely employed PTM in biological systems. It is a versatile and rapidly reversible modification with a broad spectrum of effects and crosstalk with other modifications.

We performed initial research on PTMs literature-curated databases, such as PhosphoSitePlus (www.phosphosite.org), Phospho.ELM (phospho.elm.eu.org), and PHOSIDA (www.phosida.com), which revealed that several whole-cell or whole-tissue proteomics studies identified many phosphorylation sites in the C-terminal part of Nfix. These high throughput data derive from different cell types and tissues and indicate at least six main potentially phosphorylated serines all falling in the transcriptionally active domain of Nfix, spanning from exon 5 to exon 8, with the majority of them falling in exon 6 (**Figure 12A**).

We tested Nfix for phosphorylation taking advantage of the λ-phosphatase, an Mn^{2+}-dependent protein phosphatase with activity toward serine, threonine, and tyrosine phosphorylated residues. The treatment of WT and HA_Nfix2-expressing C2C12 protein extracts with the λ-phosphatase obliterates the slower-migrating bands of Nfix doublets (black arrowheads), while the lower bands (white arrowheads) remain unaffected (**Figure 12B,C**). The effect is specific, as it is abolished by the presence of phosphatase inhibitors in the lysis buffer. In the case of WT myoblasts (**Figure 12B**), the effect is evident for two out of four immunoreactive bands, probably corresponding to the phosphorylated forms of Nfix1 and Nfix2 splicing isoforms (black arrowheads #1 and #2).

Usually, phosphorylation events do not cause a shift in electrophoretic mobility, since single phosphate groups add less than 1 kDa to the protein molecular weight. Rather, when a phosphorylation-dependent electrophoretic mobility shift is observed it is likely due to local context-dependent effects on the flexibility of the peptide chain. When this happens, the detection via polyacrylamide electrophoretic gel offers a rapid and simple method to determine the stoichiometry of phosphorylation (Dephoure et al., 2013).

From this experiment, we conclude that Nfix is phosphorylated in skeletal myoblasts, and this likely causes a conformational reorganization in the protein leading to a variation in its apparent molecular weight. The resultant phosphorylation-dependent mobility shift allows

us to study this modification even in the absence of phosphorylation-specific antibodies.

As already mentioned, our experimental data are backed by several proteomics studies that show the presence of several potentially phosphorylated residues all falling in the C-terminal domain of Nfix. Starting from the proteomic data available at PhosphoSitePlus database (www.phosphosite.org), we selected the six putative phosphorylation sites that were identified by most of the annotated proteomic studies: S265, S268, S280, S288, S301, and S320 (**Figure 12A**). Classically, Ser- and Thr- phosphorylation can be abolished by mutation to alanine. We thus performed site-directed mutagenesis to introduce serine-to-alanine substitutions in the plasmid encoding for HA-Nfix2 (pLenti_HA-Nfix2). The expression of these Nfix2 variants in primary myoblasts and their detection through western blot revealed that the slower-migrating Nfix form (**Figure 12D, black arrowhead**) is phosphorylated on serine 301. Indeed, the S301A mutant Nfix2 species presents a single immunoreactive band (**Figure 12D, white arrowhead**) that corresponds to the one that remains after phosphatase treatment (**Figure 12C, white arrowhead**). Accordingly, treatment of protein lysates from HA-Nfix2_S301A myoblasts with λ-phosphatase does not lead to any further changes in Nfix electrophoretic mobility, compared to HA-Nfix2_WT controls (**Figure 12E**).

Ser-phosphorylation can be mimicked (with some limitations) by mutations to aspartate or glutamate. These substitutions are of similar volume to phosphoserine residues and provide a single negative charge (versus the -1.5 formal charges of a phosphoresidue) (Pérez-Mejías et al., 2020).

We decided to introduce a serine-to-aspartate substitution (S301D) in the pLenti_HA-Nfix2 plasmid using the same site-directed mutagenesis settings used to obtain the phosphodeficient (S301A) mutants. We then transfected murine MuSCs-derived myoblasts with WT, S301A, and S301D Nfix2 plasmids to compare the different Ser 301 variants of Nfix2. Strikingly, the introduction of the negatively charged amino acid aspartate in place of the phosphorylable serine 301, even without the addition of a phosphate group per se, successfully mimics the phosphorylation-dependent mobility shift of pS301 Nfix. Indeed, as observed in **Figure 12F**, myoblasts transfected with the pLenti_HA-Nfix2_S301D display a single, higher band with respect to the HA-Nfix2_S301A variant, roughly corresponding to the higher band of the HA_Nfix2_WT doublet, although a bit lower. We conclude that the introduction of a negative charge on position 301 of the Nfix aminoacidic sequence is necessary and sufficient to induce the Nfix phosphorylation-dependent mobility shift. Thus, the Nfix2_S301D mutant could represent a valid phosphomimetic Nfix variant to explore

the biological function of serine 301 phosphorylation in myoblasts' biology, together with its Nfix2_S301A counterpart.

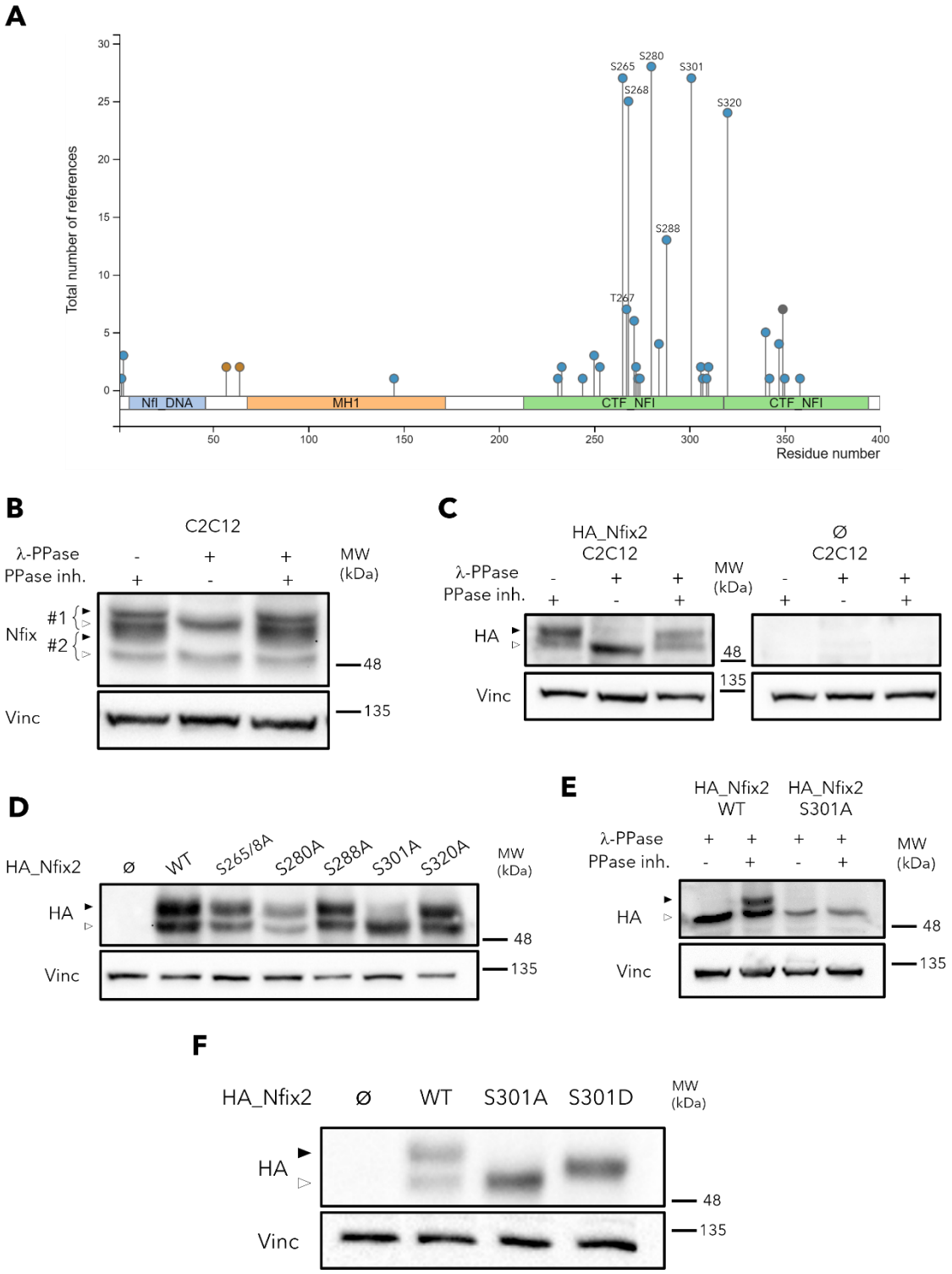


Figure 12. Nfix is phosphorylated at Ser 301 in myoblasts *in vitro*

A) PhosphositePlus graph depicting PTMs sites on the Nfix protein sequence identified by high-throughput studies. The letters and numbers indicate the amino acid its relative position in the protein sequence of Nfix of the phosphorylation sites with the highest number of references.

B) Western blot of wildtype C2C12 protein extracts treated with λ-Phosphatase (λ-PPase) alone or in combination with phosphatase inhibitors (PPase inh.). Phosphorylated Nfix bands are indicated with black arrowheads; white arrowheads indicate de-phosphorylated Nfix. #1 and #2 indicate the two phosphorylated/unphosphorylated doublets presumably belonging to isoform 1 and 2 of Nfix. Vinculin is used as a loading control.
C) Western blot of HA-Nfix2-overexpressing and control C2C12 protein extracts treated with λ-Phosphatase (λ-PPase) alone or in combination with phosphatase inhibitors (PPase inh.). Phosphorylated and de-phosphorylated Nfix2 are indicated with black and white arrowheads, respectively. Vinculin is used as loading control.
D) Western blot of MuSCs-derived myoblasts transfected with different serine-to-alanine HA-Nfix2 mutants. pS301-Nfix is indicated with the black arrowhead, while not phosphorylated S301 is indicated with the white arrowhead. Vinculin is used as loading control.
E) Western blot of HA-Nfix2 and HA_Nfix2_S301A expressing MuSCs-derived myoblasts protein extracts treated with λ-Phosphatase (λ-PPase) alone or in combination with phosphatase inhibitors (PPase inh.). . pS301-Nfix is indicated with the black arrowhead, while not phosphorylated S301 is indicated with the white arrowhead. Vinculin is used as loading control.
F) Western blot of control, HA-Nfix2, HA_Nfix2_S301A, and HA_Nfix2_S301D expressing MuSCs-derived myoblasts. pS301-Nfix is indicated with the black arrowhead, while not phosphorylated S301 is indicated with the white arrowhead. Vinculin is used as loading control.

BLAST alignment evidenced that Ser 301 is conserved across different vertebrate species, including zebrafish (**Figure 13A**), in which the ortholog nfixa has conserved roles during myogenesis (Pistocchi et al., 2013).

Interestingly, a phosphorylation-disruptive serine-to-phenylalanine mutation of this residue (S301F) was identified in a melanoma patient and is reported on the cBioPortal Cancer Genomics database (https://www.cbioportal.org/)(**Figure 13B**). This could indicate a possible implication of Ser 301 phosphorylation in cancer mechanisms.

As can be observed in **Figure 13C**, the Ser 301 residue is predicted to fall in a highly disordered region, and the theoretical modeling artificial intelligence system implemented by the AlphaFold Protein Structure Database (https://alphafold.ebi.ac.uk/; (Tunyasuvunakool et al., 2021)) is not able to confidently predict the structure of this region of Nfix.

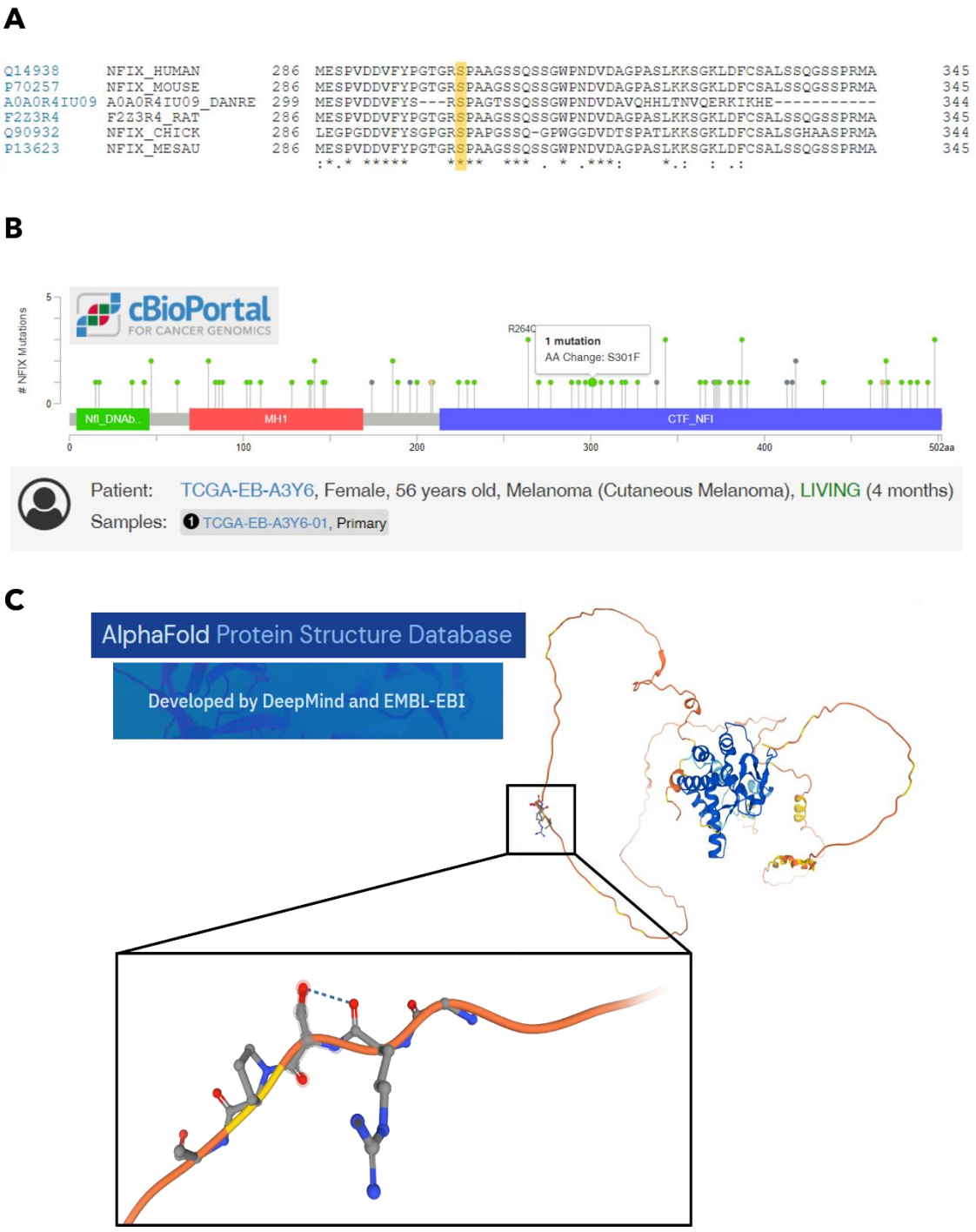


Figure 13. Nfix S301 is conserved across vertebrates and falls in a highly disordered region

A) BLAST alignment of the Nfix sequence of different vertebrate species (Human, Mouse, Zebrafish, Rat, Chicken, Golden hamster) showing the conservation of Ser 301 amino acid.

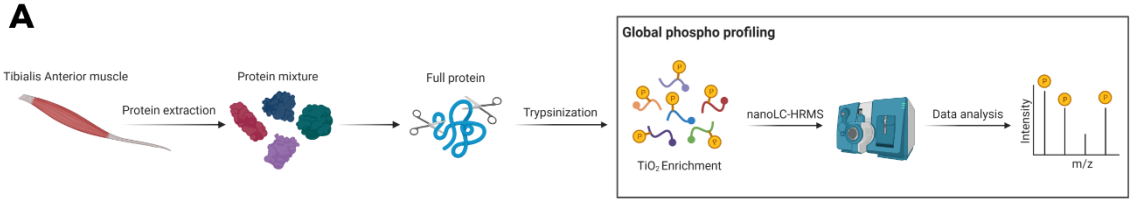
B) cBioPortal image that shows a Nfix S301F mutation in a Cutaneous Melanoma patient.

C) Image depicting the computationally predicted Nfix structure from the AlphaFold Protein Structure Database.

3.4 Nfix is phosphorylated at Ser 301 in skeletal muscle *in vivo*

To confirm the occurrence of S301 phosphorylation also *in vivo*, we performed a phosphoproteomics analysis on *Tibialis Anterior* (TA) samples of adult CD1 WT mice (**Figure 14A**). The mass spectrometry data identified the S301 phosphorylation (**Figure 14B,C**) with 99-100% confidence, thus validating the *in vitro* studies. The phosphoproteomics also identified phosphorylation on serine 265 and threonine 267, even though with lower confidence (data not shown). Furthermore, the presence of Nfix pS301 in murine WT muscles can be distinguished via western blot analysis thanks to the characteristic phosphorylation-dependent mobility shift (**Figure 14D, black arrowheads**), as already described in cultured myoblasts. Also in this case, we observe two doublets (#1 and #2) that probably belong to isoforms 1 and 2 of Nfix.

A

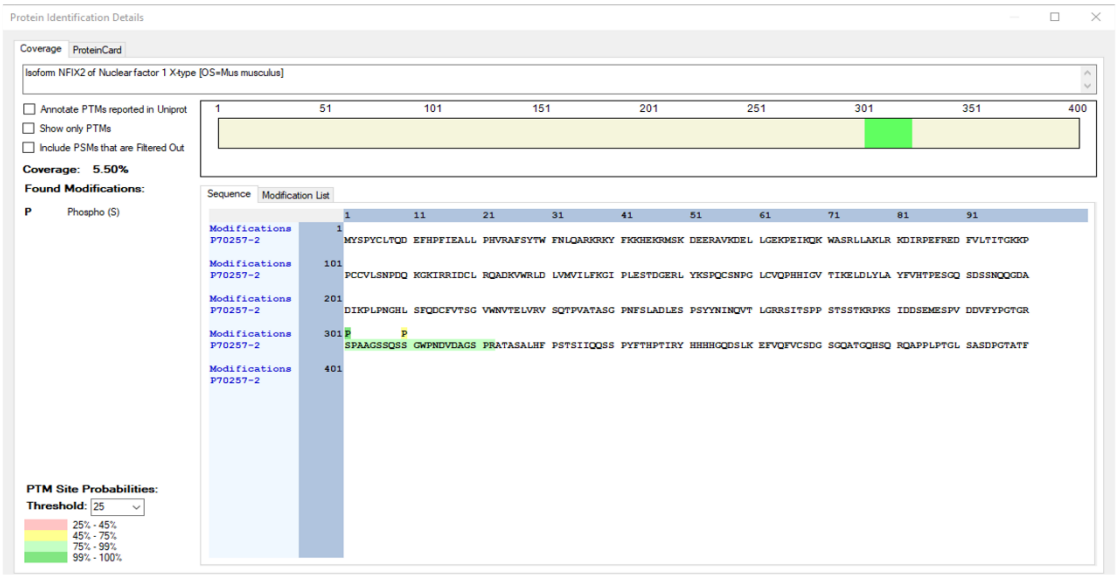


B

Master	Accession	Description	Coverage [%]	# Peptides	# PSMs	# Unique Peptides	# AAs	MW [kDa]	calc. pI	Score Sequest HT: Sequest HT
Master Protein	P70257-2	**Isoform NFIX2 of Nuclear factor 1 X-type [OS=Mus musculus]**	6	1	2	1	400	44.6	7.94	6.83

Confidence	Sequence	Modifications	Qvality PEP	Qvality q-value	Master Protein Accessions	Theo. MH+ [Da]	Confidence	m/z [Da]
High	**S**PAAGSSQSSGWPNDVDAGSPR	**1xPhospho [S1(100)]**	5.8452E-09	0.000230791	P70257-2	2209.91459	High	1105.46342

C



D

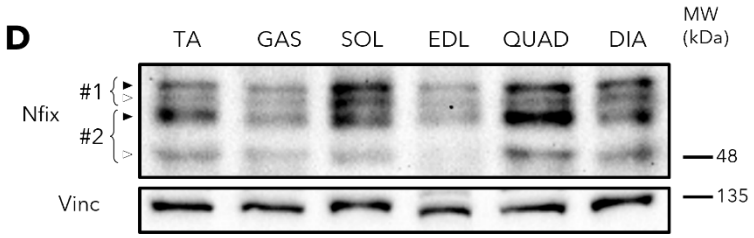


Figure 14. Nfix is phosphorylated at Ser 301 in skeletal muscle *in vivo*

A) Schematic representation of phosphoproteomic analysis workflow.

B) Extract from the phosphoproteomic results table (UNITECH OMICs) showing the identified Nfix2 S301 phosphorylation in the mouse *Tibialis anterior* sample.

C) Screenshot from the mass spectrometry analysis software highlighting the identification of the Ser 301 phosphorylation site on the Nfix2 protein sequence (UNITECH OMICs).

D) Western blot analysis of wild-type murine muscle protein extracts showing the characteristic mobility shift of Nfix Ser 301 phosphorylation (black arrowheads), for both Nfix muscle isoforms (#1 and #2).

Vinculin is used as a loading control. TA=Tibialis anterior; GAS= Gastrocnemius; SOL=Soleus; EDL= Extensor digitorum longus; QUAD=Quadriceps; DIA=Diaphragm

3.5 Nfix S301 phosphorylation is present during development and in both proliferating and differentiating postnatal myogenic cells

Nfix during development is expressed during fetal myogenesis. At E16.5 it is highly expressed in myoblasts. We thus verified the phosphorylation status of Nfix in fetal myoblasts and we found that Nfix is phosphorylated at Ser 301 also in these cells (**Figure 15A, black arrowhead**). These data support a conserved role of Nfix Ser 301 phosphorylation also during prenatal development.

To verify whether Nfix S301 phosphorylation was differentially regulated between proliferation and differentiation of postnatal MuSCs, we induced the differentiation of isolated murine MuSCs expanded in vitro. We quantified the levels of phosphorylated Nfix (pS301-Nfix, black arrowhead in **Figure 15B**) relative to the non-phosphorylated Nfix (S301-Nfix, white arrowhead in **Figure 15B**) via western blot analysis. As shown and quantified in **Figure 15B** and **C**, respectively, phosphorylation of serine 301 just slightly decreases in a not-statistically significant way at 2 and 3 days in differentiation medium.

We conclude that Nfix S301 phosphorylation levels are mainly independent of the differentiation status of myoblasts *in vitro*, as they remain constant at least for the first 3 days of differentiation *in vitro*.

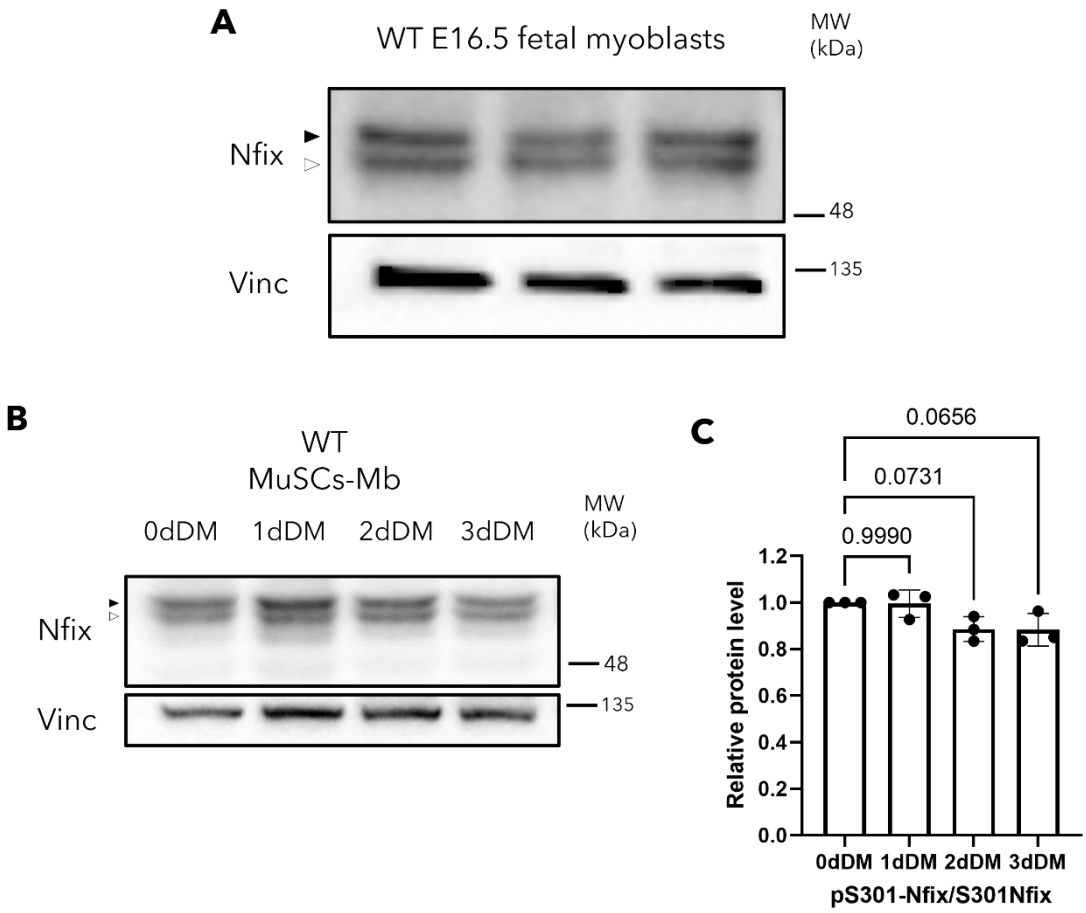


Figure 15. Nfix S301 phosphorylation is present during prenatal myogenesis and in both proliferating and differentiated muscle cells

A) Western blot showing the presence of pS301-Nfix also in fetal myoblasts. The black arrowhead indicates pS301-Nfix; the white arrowhead indicated not phosphorylated S301-Nfix. Vinculin is used as a loading control.

B) Representative western blot showing Nfix immunoblot from proliferating and differentiating MuSC-derived myoblasts (MuSC-Mb). The black arrowhead indicates pS301-Nfix; the white arrowhead indicates not phosphorylated S301-Nfix. Vinculin is used as a loading control.

C) Quantification of the levels of pS301-Nfix over the not-phosphorylated S301-Nfix of proliferating versus differentiating MuSC-derived myoblasts. pS301-Nfix is indicated by the black arrowhead while S301-Nfix is indicated by the white arrowhead. n = 3

3.6 Nfix Ser 301 phosphorylation is induced upon MuSCs late activation

From the literature, we know that Nfix is involved also in quiescence maintenance of neural stem cells (Martynoga et al., 2013). Furthermore, early studies demonstrated different NFI phosphorylation status exists between quiescent and actively proliferating 3T3-L1 adipocytes (Yang et al., 1993).

We thus decided to verify whether Nfix was phosphorylated in freshly-isolated, uncultured myoblasts. As can be observed from **Figure 16A**, while cultured MuSCs display the characteristic mobility shift due to pS301-Nfix (**Figure 16A, black arrowhead**), MuSCs lysed directly after isolation, without seeding, display mainly the nonphosphorylated Nfix bands of the two main Nfix isoforms expressed by the muscle (**Figure 16B, white arrowheads #1 and #2**). Furthermore, total Nfix protein levels (both phosphorylated and not-phosphorylated) significantly increase in cultured MuSCs-derived myoblasts with respect to freshly isolated MuSCs (**Figure 16C**).

This experiment confirms that quiescent/early activated MuSCs express a Nfix form that is not phosphorylated on Ser 301, and that both Nfix total protein levels and Nfix Ser 301 phosphorylation are induced in committed/proliferating MuSCs-derived myoblasts. Therefore, pS301-Nfix could be involved in processes of late MuSCs activation that take place in actively cycling myoblasts.

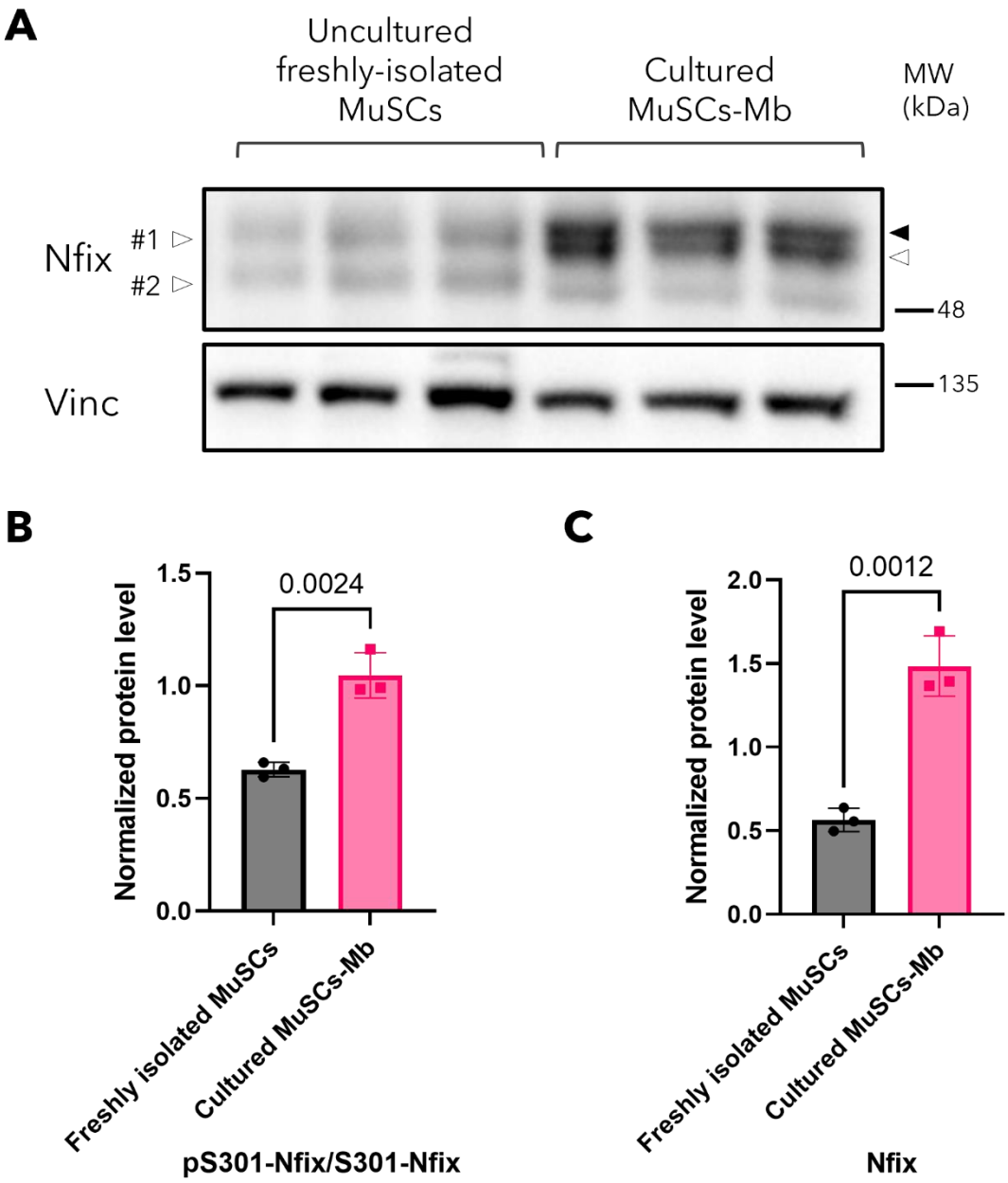


Figure 16. Nfix Ser 301 phosphorylation is induced in committed/proliferating MuSCs-derived myoblasts

A) Western blot comparing freshly isolated MuSCs that were lysed directly after isolation with MuSCs isolated and seeded in culture dishes with growth medium for 62h before lysis and protein extraction. The black arrowhead indicates the phosphorylated Nifx (pS301-Nfix), while the #1 and #2 white arrowheads indicate the nonphosphorylated Nfix (S301-Nfix), presumably of Nfix1 and 2 isoforms. The membrane was immunoblotted with anti-Nfix antibody. Vinculin was used as a loading control.

B) Quantification of pS301-Nfix (black arrowhead) on nonphosphorylated S301 Nfix (white arrowhead) of the western blot in figure A.

C) Quantification of total Nfix (both black and white arrowheads) of the western blot in figure A.

3.7 Nfix S301 phosphorylation is not downstream to AKT or mTOR in myoblasts *in vitro*

Early studies on NFIs highlighted the involvement of the insulin pathway in the regulation of NFI transcriptional activity via phosphorylation (Cooke and Lane, 1999). Furthermore, the PI3K-AKT-mTOR pathways downstream of Insulin/IGF-1 signaling are involved in the activation of MuSCs (Zhou et al., 2022).

We thus decided to verify whether Nfix S301 phosphorylation could be a target of the insulin pathway. As a model to study this possibility, we used myoblasts derived from adult (5-8 weeks old) postnatal murine WT MuSCs and we treated them with two well-characterized inhibitors of the insulin/IGF1 pathway: LY294002, an inhibitor of PI3K, and PP242, a selective inhibitor of mTORC1 and 2. We treated the cells for 4h, 8h, and 16h and verified the efficacy of the treatment by looking at the levels of known phosphorylated targets of PI3K and mTOR. In particular, we used pAKT as a readout of PI3K inhibition and pS6 for mTORC1/2 inhibition. As shown in **Figure 17A,D** and quantified in **Figure 17B,E**, both treatments were effective at all time points. However, no significant changes in Nfix pS301 levels could be detected (**Figure 17C,F**). We conclude that this particular phosphorylated residue is not a target of the PI3K or mTOR axes.

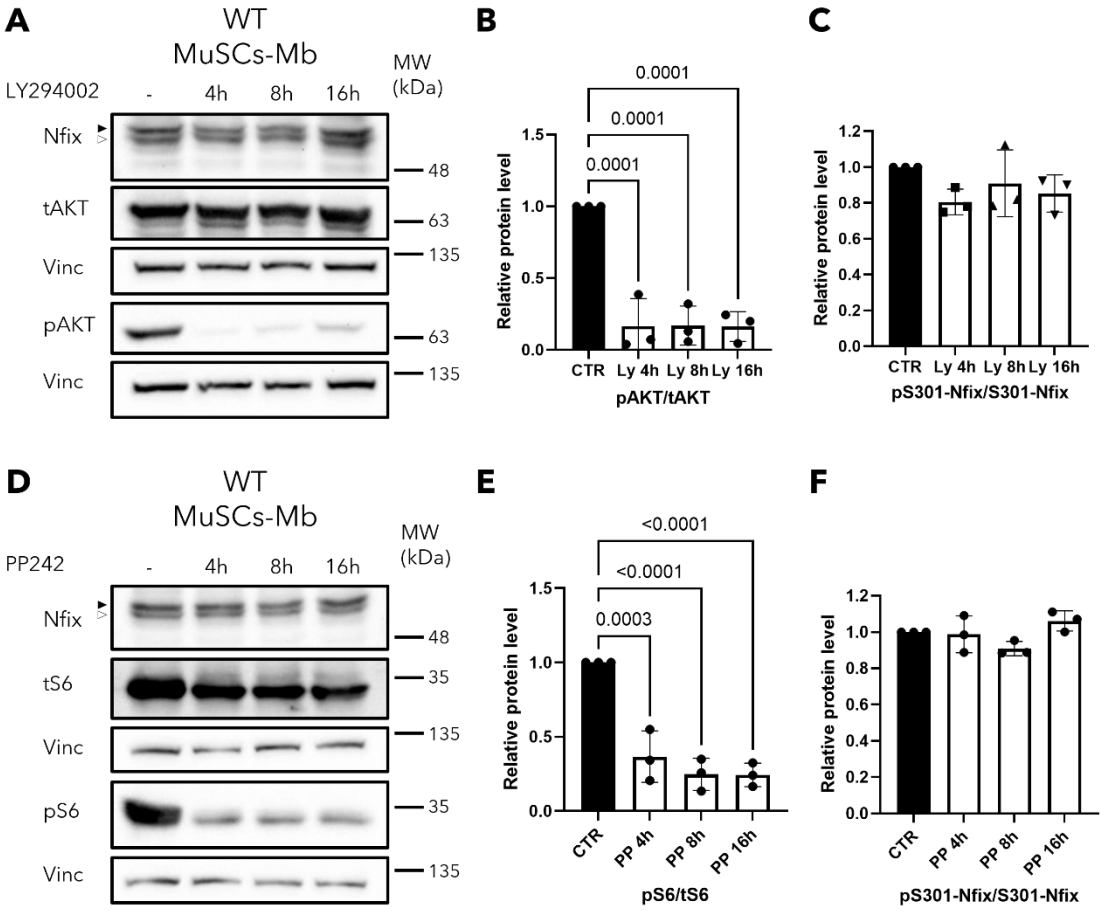


Figure 17. Nfix S301 phosphorylation is not downstream to AKT or mTOR

A) Representative western blot of MuSCs-derived myoblasts treated with the PI3K inhibitor LY294002 (10 μM) or DMSO for 4h, 8h, 16h. pAKT and tAKT are used as a readout of treatment efficacy and vinculin is the loading control. The black arrowhead indicates pS301-Nfix band, while the white arrowhead indicates the S301-Nfix not-phosphorylated band.

B) Quantification of the LY294002 treatment efficacy via pAKT/tAKT ratio. n = 3

C) Quantification of the S301 phosphorylation levels (pS301-Nfix over S301-Nfix) upon LY294002 treatment. n = 3

D) Representative western blot of MuSCs-derived myoblasts treated with the mTORC inhibitor PP242 (500 nM) or DMSO for 4h, 8h, 16h. pS6 and tS6 are used as a readout of treatment efficacy and vinculin is the loading control. The black arrowhead indicates pS301-Nfix band, while the white arrowhead indicates the S301-Nfix not-phosphorylated band.

E) Quantification of the PP242 treatment efficacy via pS6/tS6 ratio. n = 3

F) Quantification of the S301 phosphorylation levels (pS301-Nfix over S301-Nfix) upon PP242 treatment. n = 3

3.8 Proteasome inhibition affects Nfix S301 phosphorylation

In some cases, phosphorylation affects the stability of the protein, usually by promoting protein degradation. We thus treated myoblasts with different proteasome inhibitors to test whether this was the case for Nfix. We used C2C12 cells as a model of post-natal myoblasts and isolated fetal myoblasts as a pre-natal myoblast model expressing Nfix.

Contrarily to what was expected, we documented a specific and significant decrease in pS301-Nfix levels (black arrowheads) upon proteasome inhibition, both in fetal myoblasts (**Figure 18A-D**) and in C2C12 (**Figure 18E-F**), equivalently at different time points, concentrations and with different proteasome inhibitors. We can conclude that, for a yet unknown reason, inhibition of proteasome function leads to a specific drop of Nfix Ser 301 phosphorylation in myoblasts, independently of the drug used.

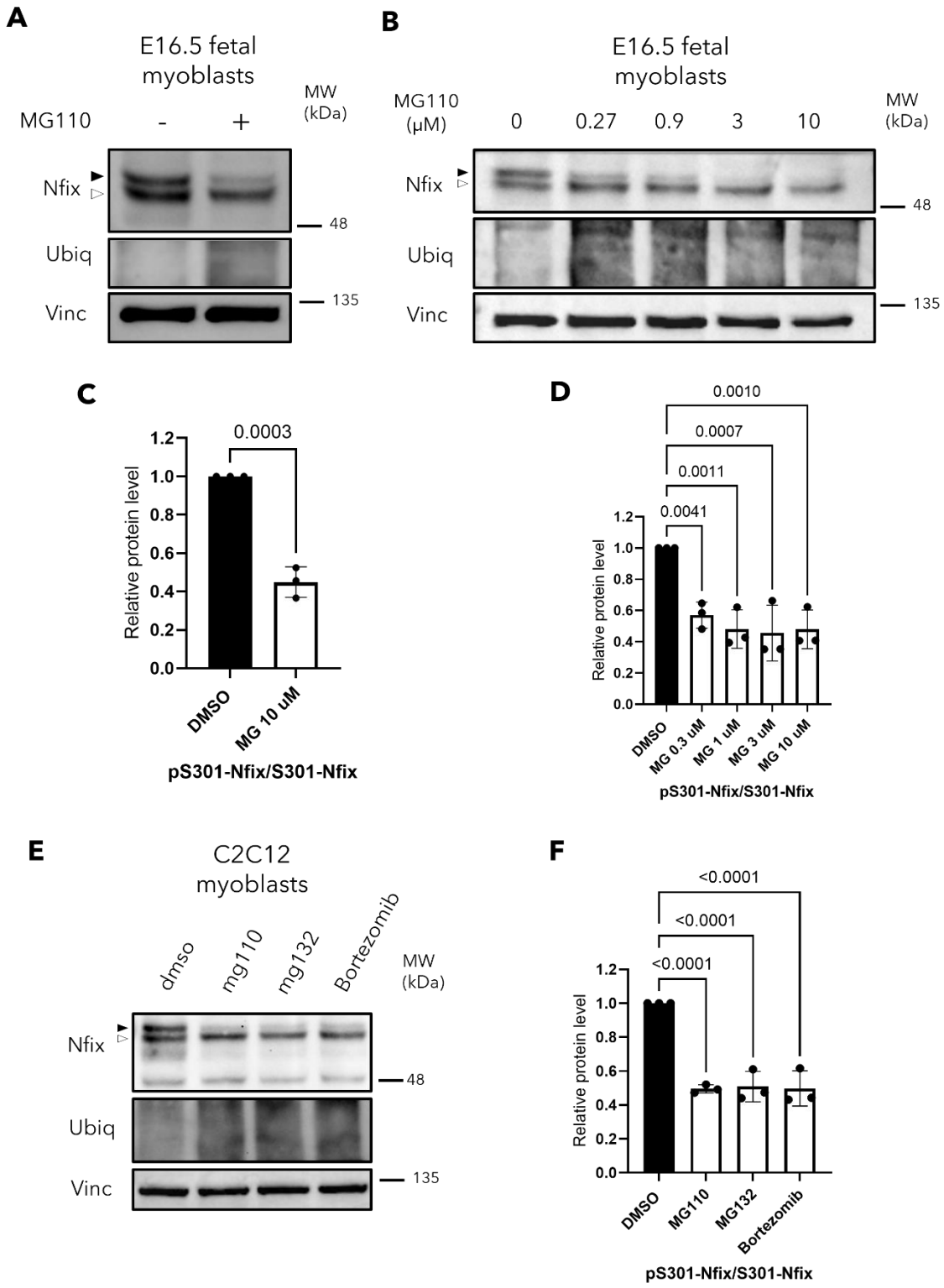


Figure 18. pS301 Nfix levels decrease upon proteasome inhibition

A) Representative western blot of primary fetal myoblasts treated with the proteasome inhibitor MG110 (10 μM) for 4h. Treatment efficacy is verified via ubiquitinated lysines (Ubiq) increase. Vinculin is used as a loading control. The black arrowhead indicates pS301-Nfix band, while the white arrowhead indicates the S301-Nfix not-phosphorylated band.

B) Representative western blot of primary fetal myoblasts treated with different concentrations of the proteasome inhibitor MG110 (0, 0.3, 1, 3, 10 μM) for 12h. Treatment efficacy is verified via ubiquitinated lysines (Ubiq) increase. Vinculin is used as a loading control. The black arrowhead indicates pS301-Nfix band, while the white arrowhead indicates the S301-Nfix not-phosphorylated band.
C) Quantification of pS301-Nfix (black arrowhead) over S301-Nfix (white arrowhead) (pS301-Nfix/S301-Nfix) levels in primary fetal myoblasts treated with the proteasome inhibitor MG110 (10 μM) for 4h. n = 3
D) Quantification of pS301-Nfix (black arrowhead) over S301-Nfix (white arrowhead) (pS301-Nfix/S301-Nfix) levels in primary fetal myoblasts treated with different concentrations of the proteasome inhibitor MG110 for 12h. n = 3
E) Representative western blot of C2C12 myoblasts treated with different proteasome inhibitors (MG110 10 μM; MG132 10 μM, Bortezomib 1uM) for 4h. Treatment efficacy is verified via ubiquitinated lysines (Ubiquitin) increase. The black arrowhead indicates pS301-Nfix band, while the white arrowhead indicates the S301-Nfix not-phosphorylated band. Vinculin is used as a loading control.
F) Quantification of pS301-Nfix (black arrowhead) over S301-Nfix (white arrowhead) (pS301-Nfix/S301-Nfix) levels in C2C12 myoblasts treated with different proteasome inhibitors for 4h. n = 3

3.9 Nfix S301 phosphorylation is sensitive to oxidative stress

The levels of ROS and oxidative stress in cells are known factors influencing protein phosphorylation and signaling cascades. We thus tested whether Nfix phosphorylation is influenced by oxidative stress by treating wildtype post-natal MuSC-derived myoblasts with hydrogen peroxide (H_2O_2) at two different concentrations for 4h. Strikingly, Nfix pS301 (black arrowheads) levels significantly decrease upon H_2O_2 treatment at both concentrations, with no significant effects on the total Nfix levels (**Figure 19A-C**). This experiment shows that Nfix S301 phosphorylation is oxidation-sensitive.

In dystrophic muscles, the chronic inflammatory environment increases local ROS levels and triggers pathological oxidative stress responses. We thus tested whether Nfix pS301 levels were influenced by this context in a dystrophic mouse model: the *sgca*-null mouse. Muscles from adult *sgca*-null mice were isolated and protein extracts were analyzed via western blot. The levels of pS301-Nfix relative to S301-Nfix were quantified for both the Nfix isoforms (arrowheads #1 and #2) that are evident in western blot analyses from muscle sections, taking advantage of the phosphorylation-dependent mobility shift of pS301-Nfix (black arrowheads) (**Figure 19D**). However, no detectable changes in pS301-Nfix levels could be quantified in dystrophic versus wild-type muscles (**Figure 19E-F**).

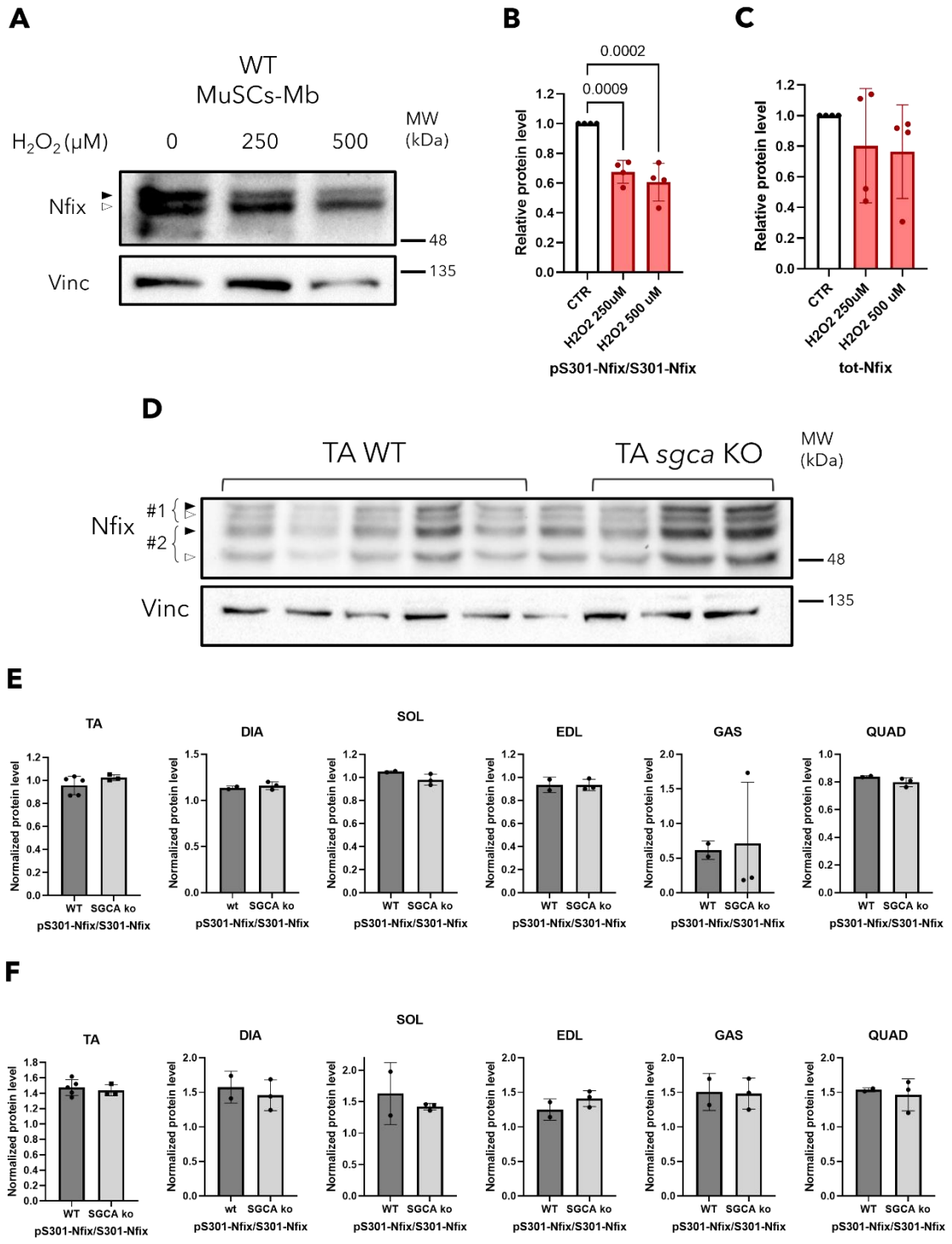


Figure 19. Nfix S301 phosphorylation is sensitive to oxidative stress

A) Representative western blot of MuSC-derived myoblasts treated with hydrogen peroxide (H_2O_2) for 4h. Vinculin is used as a loading control. The black arrowhead indicates pS301-Nfix band, while the white arrowhead indicates the S301-Nfix not-phosphorylated band.

B) Quantification of pS301-Nfix (black arrowhead) over S301-Nfix (white arrowhead) levels in MuSC-derived myoblasts treated with hydrogen peroxide (H_2O_2) for 4h. n = 3

C) Quantification of total Nfix levels (both black and white arrowheads) in MuSC-derived myoblasts treated with hydrogen peroxide (H_2O_2) for 4h. n = 3
D) Western blot of Tibialis anterior (TA) muscle protein extracts from wildtype and *sgca*-null mice. The pS301-Nfix band and S301-Nfix not-phosphorylated band are indicated by black and white arrowheads, respectively. #1 and #2 indicate the band couples that presumably correspond to two different Nfix isoforms. Vinculin is used as a loading control.
E) Quantification of pS301-Nfix (black arrowhead #1) over S301-Nfix levels (white arrowhead #1) in muscles from wildtype and *sgca*-null mice. For TA muscles: wildtype n = 5; *sgca*-null n = 3; for all other muscles: wildtype n = 2; *sgca*-null n = 3
F) Quantification of pS301-Nfix (black arrowhead #2) over S301-Nfix levels (white arrowhead #2) in muscles from wildtype and *sgca*-null mice. For TA muscles: wildtype n = 5; *sgca*-null n = 3; for all other muscles: wildtype n = 2; *sgca*-null n = 3

3.10 Effects of Nfix S301 phosphorylation mutants expression on myoblasts

To study the biological effect of Nfix phosphorylation mutants in a muscular system without the interference of endogenous Nfix, we established *Nfix*-null muscle cell lines from *Nfix*-null juvenile (P13) mice (**Figure 20A**). We verified the absence of Nfix protein expression in these cells compared to their WT counterparts (**Figure 20B**). Contextually, we confirmed that Nfix phosphorylation is maintained also in juvenile WT MuSCs-derived myoblasts: Nfix appears as multiple specific bands (black and white arrowheads), which are at least in part due to phosphorylation (black arrowheads), as verified by λ-phosphatase treatment (**Figure 20C**). We can again observe two distinct couples of bands (#1 and #2), that we assign to isforoms 1 and 2 of Nfix, both having a phosphorylated (black arrowhead) and a not-phosphorylated band (white arrowhead) (**Figure 20B, C**).

Starting from the $Nfix^{-/-}$ myoblasts (here referred to as $Nfix^{-/-}$ MuSC-Mb), we obtained cell lines stably expressing the Nfix2 variants (Nfix2_WT, Nfix2_S301A, and Nfix2_S301D) (**Figure 20D**). As verified through real-time qPCR and western blot analyses, all transduced MuSC-Mb lines, except for $Nfix^{-/-}$ MuSC-Mb controls, correctly express Nfix at comparable levels (**Figure 20E-F**). It is thus likely that both phosphorylation-site mutants undergo normal folding and do not cause nonspecific damage to the protein, since proteins that cannot fold correctly are normally destroyed (Dephoure et al., 2013).

Furthermore, immunofluorecence staining against the HA tag confirms that both S301

variants produce a nuclear Nfix protein, as the Nfix2_WT (**Figure 21**). We conclude that this particular phosphorylation site does not influence the nuclear localization of the protein, further confirming that the two mutations do not compromise the key protein functions.

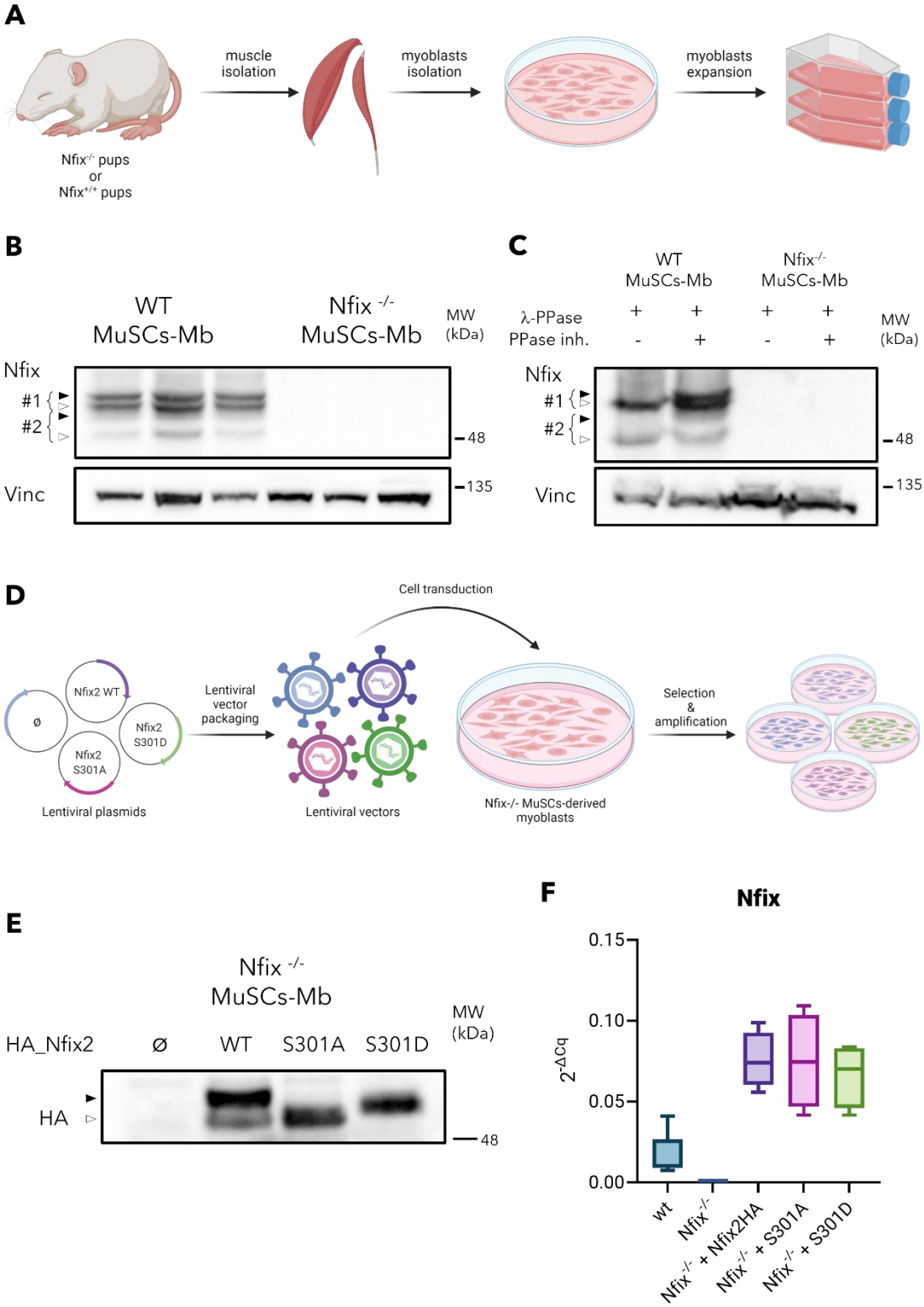


Figure 20. Generation of Nfix S301 phosphorylation mutant cell lines from *Nfix*-null MuSCs-derived myoblasts

A) Schematic representation of the generation of *Nfix*-null and wildtype MuSCs-derived myoblasts lines.

B) Western blot to compare Nfix expression between wildtype and Nfix-null MuSCs-Mb lines. Vinculin is used as loading control. Nfix specific bands are indicated by arrowheads: black for the pS301 phosphorylated form and white for the S301 not-phosphorylated species. #1 and #2 presumably correspoding to isoforms 1 and 2 of Nfix.

C) Western blot of wildtype and *Nfix*-null protein extracts treated with λ-Phosphatase (λ-PPase) alone or in combination with phosphatase inhibitors (PPase inh.). Vinculin is used as loading control. Phosphorylated Nfix (pS301) is indicated with the black arrowheads, and the not-phosphorylated bands (S301) are indicated by the white arrowheads. #1 and #2 presumably correspoding to isoforms 1 and 2 of Nfix.

D) Schematic representation of the generation of *Nfix*-null MuSCs-derived myoblasts lines stably expressing the HA-Nfix2 S301 variants (HA-Nfix2_WT, HA-Nfix2_S301A, HA-Nfix2_S301D).

E) Representative western blot of *Nfix*-null MuSCs-derived myoblasts lines stably expressing the HA-Nfix2 S301 variants. The membrane was immunoblotted with a primary antibody against HA. Vinculin is used as a loading control.

F) RT-qPCR quantification of Nfix expression. β-actin is used as housekeeping gene. n = 3

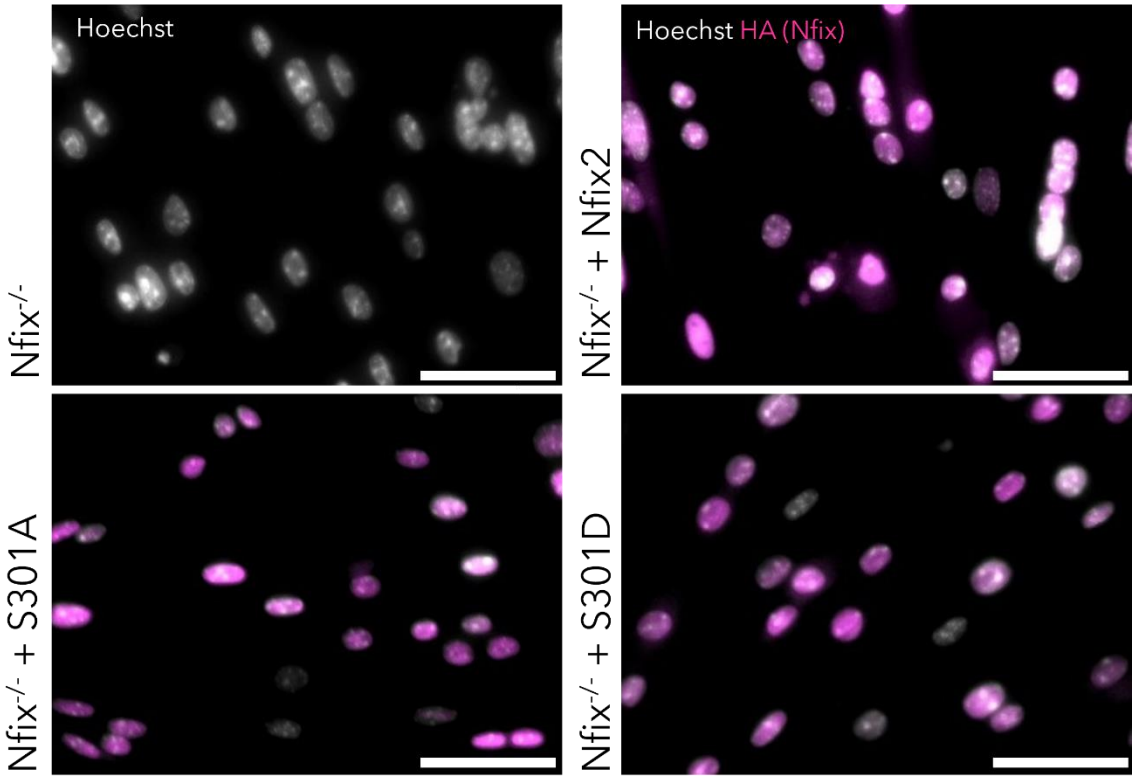


Figure 21. Nfix S301 phosphorylation mutants correctly retain nuclear localization

Immunofluorescence of MuSC-derived myoblasts lines from Nfix-null mice transduced with HA-Nfix2, HA-Nfix2_S301A, HA-Nfix2_S301D and control lentiviral vectors. A primary antibody against HA was used to stain Nfix. Hoechst was used to counterstain the nuclei. Scale bar 50 μm.

We then assessed the myogenic properties of these cell lines looking at the expression of different myogenic markers, both in proliferation and differentiation. All cell lines maintain the myoblast identity, as verified via Pax7, Myod, and Myog markers, without significant differences in their expression levels among the different Nfix2 variants (**Figure 22A,E** and **Figure23A-F**), except for Pax7+ cells, that statistically increase in S301A mutant cells versus *Nfix*-null myoblasts (**Figure 22C**). However, this is not true for Pax7 protein levels (**Figure 23A,F**). All myoblasts correctly differentiate into MyHC-positive myotubes (**Figure 22D** and **Figure23B,D**). Differentiation and fusion indexes were both unchanged (**Figure 22F-G**). Furthermore, all myoblast cell lines have comparable numbers of Ki67+ cells, which marks cycling cells from those found in G0 (**Figure 22B**).

We conclude that phosphorylation of Nfix S301 does not influence the expression of these markers or the general myoblasts' differentiation process into myotubes.

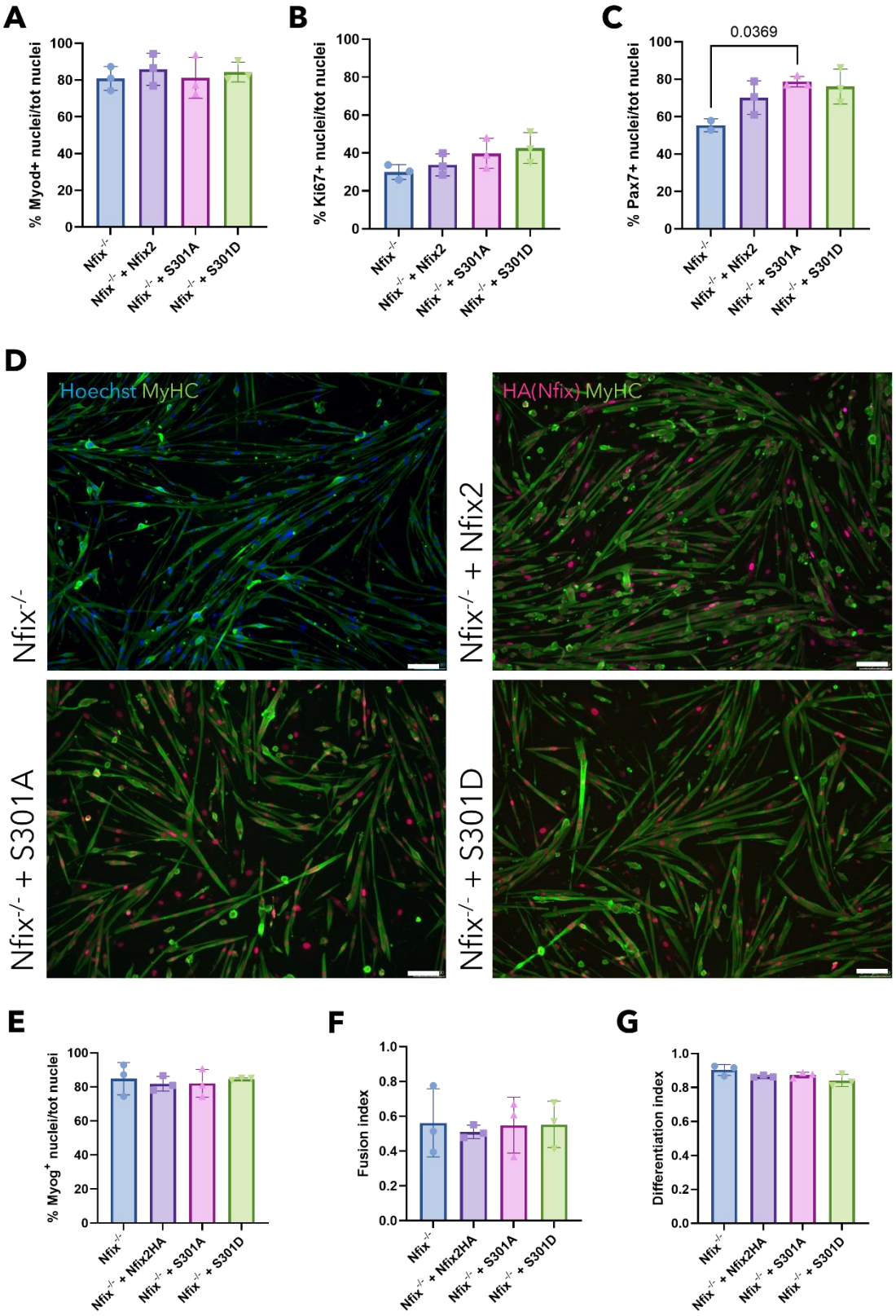


Figure 22. Characterization of Nfix S301 phosphorylation mutants by immunofluorescence staining

A) Quantification of Myod+ nuclei in proliferating Nfix-null and Nfix2-expressing MuSCs-Mb lines. n= 3

B) Quantification of Ki67+ nuclei in proliferating Nfix-null and Nfix2-expressing MuSCs-Mb lines. n= 3

C) Quantification of Pax7+ nuclei in proliferating Nfix-null and Nfix2-expressing MuSCs-Mb lines. *Nfix*-null MuSCs-Mb n= 2, Nfix2-expressing MuSCs-Mb n = 3.

D) Representative immunofluorescence images of differentiating MuSCs-derived myoblasts stable lines after 3 days in differentiation medium. Staining against myosin heavy chains (MyHC), HA (Nfix), and Hoechst (nuclei). Scale bar 100 µm.

E) Quantification of Myog+ nuclei in Nfix-null and Nfix2-expressing MuSCs-Mb lines after 3 days in differentiation medium. n= 3

F) Quantification of fusion index (myotubes with ≥3 nuclei) in Nfix-null and Nfix2-expressing MuSCs-Mb lines after 3 days in differentiation medium. n= 3

G) Quantification of differentiation index in Nfix-null and Nfix2-expressing MuSCs-Mb lines after 3 days in differentiation medium. n= 3

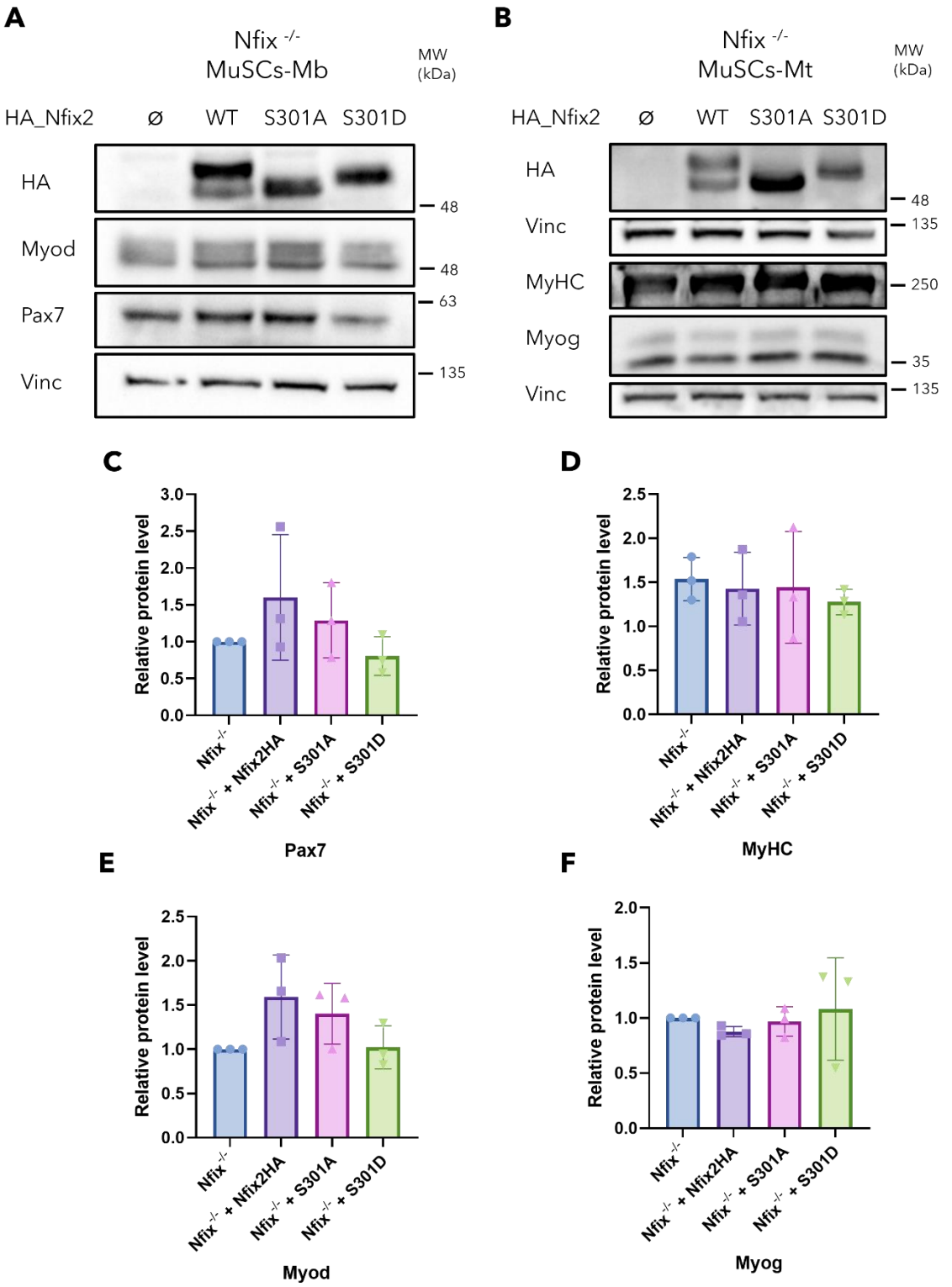


Figure 23. Nfix S301 phosphorylation mutant myoblasts retain expression of key myogenic factors

A) Representative western blot of proliferating MuSCs-Mb stable lines.

B) Representative western blot of differentiated MuSCs-derived myoblasts stable lines after 3 days in differentiation medium.

C) Quantification of the expression of Pax7 in proliferating MuSCs-Mb stable lines. n = 3

D) Quantification of the expression of MyHC in differentiated MuSCs-derived myoblasts stable lines after 3 days in differentiation medium. n = 3
E) Quantification of the expression of Myod in proliferating MuSCs-Mb stable lines. n = 3
F) Quantification of the expression of Myog in differentiated MuSCs-derived myoblasts stable lines after 3 days in differentiation medium. n = 3

3.11 RNA-seq analysis of Nfix phosphorylation mutants

The Ser 301 residue is situated in the transcriptionally active domain of Nfix. Thereby, an implication of phosphorylation in Nfix transcriptional activity is possible. We thus decided to perform an RNA-seq analysis to compare the transcriptome changes between the different MuSC-Mb cell lines obtained and described in the previous results. Since S301 phosphorylation is present in both proliferating and differentiating myoblasts, we performed the analysis starting with proliferating myoblasts. Group names as indicated in the RNA-seq analysis figures, with their description, are indicated in **Figure 24A**.

All myoblast lines display comparable numbers of differentially expressed genes (DEGs) (**Figure 24B-D**). Since we were interested in the identification of possible biological processes differentially regulated by the two Nfix forms, we performed an enrichment analysis of significantly enriched Gene Ontology (GO) terms of DEGs in Nfix-expressing myoblasts compared to *Nfix*-null myoblasts using DAVID Functional Annotation Tool. This tool clusters redundant annotation terms to give a more comprehensive result.

When considering upregulated genes between Nfix-expressing and Nfix-null groups, the analysis reveals a prevalence of cell cycle, cell division, and mitosis-related GO terms, particularly in Nfix2_WT and S301A myoblasts (**Figure 25B-D**). Indeed, we can observe how these gene sets are particularly overrepresented in S301A mutants, with both a higher gene count and statistical significance for each GO term, while in S301D mutants these terms are much less represented and do not reach the significance (compare **Figure 25C** with **Figure 25B** and **D**). The other two gene sets present in all three conditions involve angiogenesis and cell adhesion. Interestingly, cell adhesion is the only significantly enriched GO term in S301D mutants (**Figure 25D**). Both phosphorylation mutants display enrichment, even though not significant, in lipid metabolism-related genes (**Figure 25B,C**).

We noticed that nonphosphorylated S301A mutants display upregulation of *Rgs2 (regulator of G-protein signaling 2)* and *Bmp2* (*bone morphogenetic protein 2*) genes, which is not observed in Nfix2_WT or S301D mutants vs. KO comparison. These are two of the genes that were

found to be upregulated in freshly-isolated, uncultured MuSCs (Fukada et al., 2007).

When considering downregulated genes between Nfix-expressing and Nfix-null groups, the most significantly enriched GO term in all conditions is "Differentiation" (**Figure 26A-D**). In the three myoblast lines, we can find terms related to both neurogenesis and myogenesis, with different degrees of enrichment (**Figure 26A-D**). Furthermore, we can also find terms associated with Wnt and Notch signaling pathways, transcriptional regulation, and ion (sodium and potassium) transport (**Figure 26A-D**). No striking differences in downregulated gene-set enrichment are observed between phosphorylation mutant samples.

A

Group Name	Description
KO	Nfix-null MuSCs-Mb
Nfix2_WT	Nfix-null MuSCs-Mb stably expressing the wildtype Nfix2 isoform (Nfix2_WT)
Nfix2_S301A	Nfix-null MuSCs-Mb stably expressing the Nfix2_S301A phosphomutant
Nfix2_S301D	Nfix-null MuSCs-Mb stably expressing the Nfix2_S301D phosphomimetic

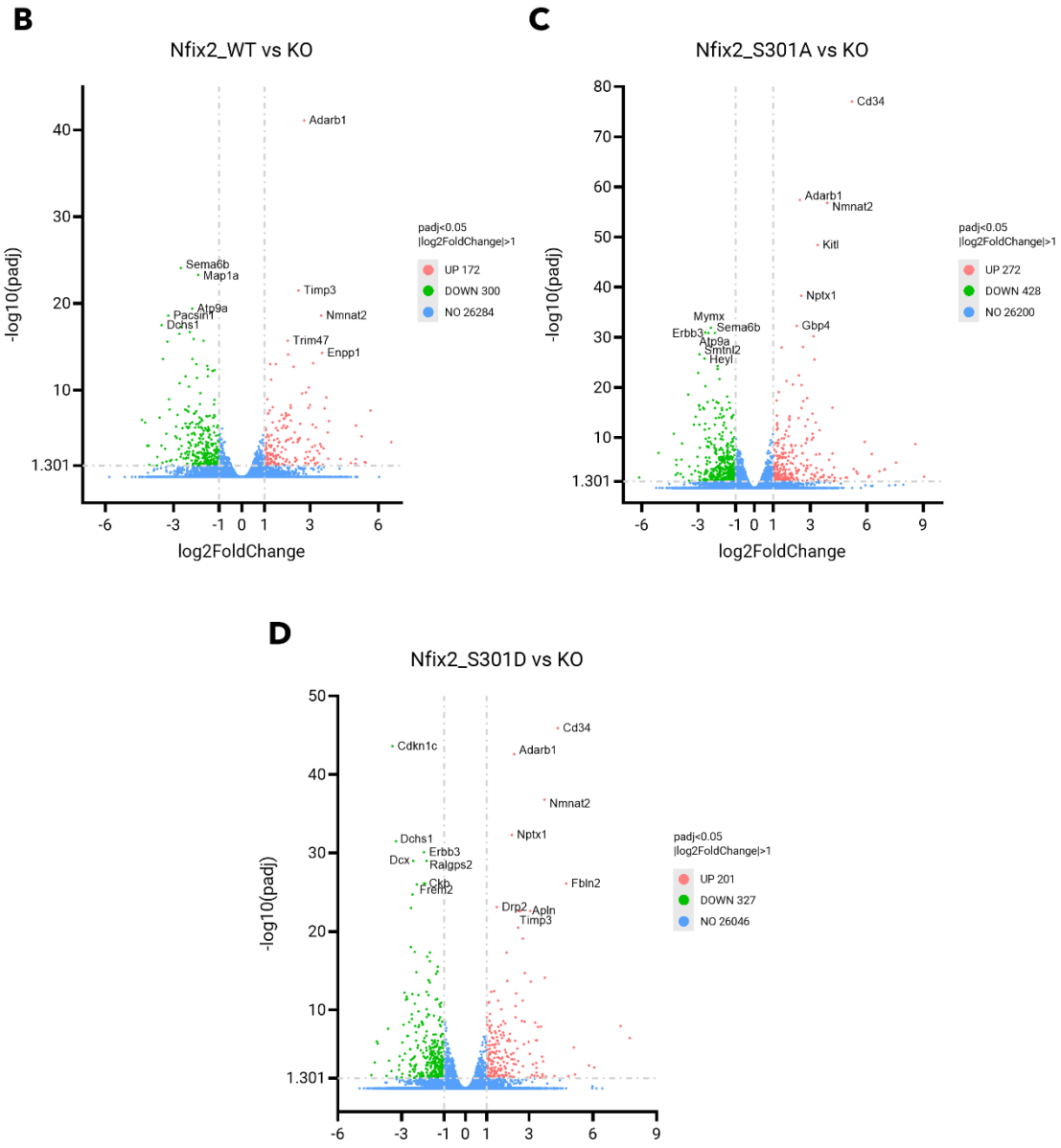


Figure 24. Effect of S301 Nfix2 variants' expression on the number of differentially upregulated and downregulated genes in myoblasts

A) Table with RNA-seq group names and relative description

B) Volcano plot of differentially expressed genes between Nfix2_WT-expressing *Nfix*-null myoblasts versus *Nfix*-null myoblasts with the most differentially expressed genes outlined.

C) Volcano plot of differentially expressed genes between Nfix2_S301D-expressing *Nfix*-null myoblasts versus *Nfix*-null myoblasts with the most differentially expressed genes outlined.

D) Volcano plot of differentially expressed genes between Nfix2_S301A-expressing *Nfix*-null myoblasts versus *Nfix*-null myoblasts with the most differentially expressed genes outlined.

A Functional Annotation Chart: Nfix2_WT vs KO_up

Category	Term	RT	Genes	Count	%	P-Value	FDR
UP_KW_BIOLOGICAL_PROCESS	Mitosis	RT		25	4,0	1,8E-6	2,1E-4
UP_KW_BIOLOGICAL_PROCESS	Cell cycle	RT		39	6,2	3,4E-5	2,0E-3
UP_KW_BIOLOGICAL_PROCESS	Cell division	RT		27	4,3	6,2E-5	2,4E-3
UP_KW_BIOLOGICAL_PROCESS	Angiogenesis	RT		13	2,1	5,4E-4	1,6E-2
UP_KW_BIOLOGICAL_PROCESS	Cell adhesion	RT		26	4,1	2,4E-3	5,7E-2
UP_KW_BIOLOGICAL_PROCESS	Apoptosis	RT		27	4,3	3,5E-3	6,8E-2
UP_KW_BIOLOGICAL_PROCESS	Stress response	RT		7	1,1	2,9E-2	4,9E-1
UP_KW_BIOLOGICAL_PROCESS	rRNA processing	RT		7	1,1	5,4E-2	7,6E-1
UP_KW_BIOLOGICAL_PROCESS	Differentiation	RT		30	4,8	5,8E-2	7,6E-1

B Functional Annotation Chart: Nfix2_S301A vs KO_up

Category	Term	RT	Genes	Count	%	P-Value	FDR
UP_KW_BIOLOGICAL_PROCESS	Mitosis	RT		46	5,3	6,2E-16	7,4E-14
UP_KW_BIOLOGICAL_PROCESS	Cell division	RT		50	5,8	6,0E-13	3,6E-11
UP_KW_BIOLOGICAL_PROCESS	Cell cycle	RT		65	7,5	2,1E-11	8,3E-10
UP_KW_BIOLOGICAL_PROCESS	Angiogenesis	RT		17	2,0	9,6E-5	2,9E-3
UP_KW_BIOLOGICAL_PROCESS	Chromosome partition	RT		8	0,9	2,0E-3	4,6E-2
UP_KW_BIOLOGICAL_PROCESS	Cell adhesion	RT		33	3,8	2,3E-3	4,6E-2
UP_KW_BIOLOGICAL_PROCESS	Apoptosis	RT		33	3,8	7,5E-3	1,3E-1
UP_KW_BIOLOGICAL_PROCESS	One-carbon metabolism	RT		4	0,5	3,4E-2	4,1E-1
UP_KW_BIOLOGICAL_PROCESS	Blood coagulation	RT		6	0,7	3,4E-2	4,1E-1
UP_KW_BIOLOGICAL_PROCESS	Hemostasis	RT		6	0,7	3,4E-2	4,1E-1
UP_KW_BIOLOGICAL_PROCESS	DNA replication	RT		9	1,0	4,2E-2	4,6E-1
UP_KW_BIOLOGICAL_PROCESS	Lipid metabolism	RT		37	4,3	6,8E-2	6,5E-1
UP_KW_BIOLOGICAL_PROCESS	Prostaglandin biosynthesis	RT		3	0,3	7,0E-2	6,5E-1
UP_KW_BIOLOGICAL_PROCESS	Lipid transport	RT		10	1,2	8,3E-2	7,1E-1
UP_KW_BIOLOGICAL_PROCESS	Prostaglandin metabolism	RT		3	0,3	9,5E-2	7,6E-1

C Functional Annotation Chart: Nfix2_S301D vs KO_up

Category	Term	RT	Genes	Count	%	P-Value	FDR
UP_KW_BIOLOGICAL_PROCESS	Cell adhesion	RT		29	4,5	1,3E-4	1,6E-2
UP_KW_BIOLOGICAL_PROCESS	Angiogenesis	RT		12	1,9	1,4E-3	8,7E-2
UP_KW_BIOLOGICAL_PROCESS	Mitosis	RT		18	2,8	2,2E-3	9,0E-2
UP_KW_BIOLOGICAL_PROCESS	Lipid metabolism	RT		32	4,9	7,0E-3	2,1E-1
UP_KW_BIOLOGICAL_PROCESS	Prostaglandin biosynthesis	RT		3	0,5	3,7E-2	6,1E-1
UP_KW_BIOLOGICAL_PROCESS	Chromosome partition	RT		5	0,8	3,8E-2	6,1E-1
UP_KW_BIOLOGICAL_PROCESS	Inflammatory response	RT		11	1,7	4,4E-2	6,1E-1
UP_KW_BIOLOGICAL_PROCESS	Cell division	RT		18	2,8	4,7E-2	6,1E-1
UP_KW_BIOLOGICAL_PROCESS	Cell cycle	RT		27	4,2	4,8E-2	6,1E-1
UP_KW_BIOLOGICAL_PROCESS	Prostaglandin metabolism	RT		3	0,5	5,1E-2	6,1E-1
UP_KW_BIOLOGICAL_PROCESS	Amino-acid transport	RT		5	0,8	5,5E-2	6,1E-1
UP_KW_BIOLOGICAL_PROCESS	Fatty acid metabolism	RT		9	1,4	6,8E-2	6,9E-1
UP_KW_BIOLOGICAL_PROCESS	Lipid transport	RT		8	1,2	7,8E-2	6,9E-1
UP_KW_BIOLOGICAL_PROCESS	Apoptosis	RT		21	3,2	7,9E-2	6,9E-1
UP_KW_BIOLOGICAL_PROCESS	Growth arrest	RT		3	0,5	9,0E-2	7,3E-1

Figure 25. Comparison of upregulated GO terms among the different Nfix-expressing myoblast lines versus *Nfix*-null myoblasts

A) Gene Ontology enrichment of upregulated genes relative to the Nfix2_WT-expressing *Nfix*-null myoblasts versus *Nfix*-null myoblasts comparison.

B) Gene Ontology enrichment of upregulated genes relative to the Nfix2_S301A-expressing *Nfix*-null myoblasts versus *Nfix*-null myoblasts comparison.

C) Gene Ontology enrichment of upregulated genes relative to the Nfix2_S301D-expressing *Nfix*-null myoblasts versus *Nfix*-null myoblasts comparison.

A Functional Annotation Chart: Nfix2_WT vs KO_down

Category	Term	RT	Genes	Count	%	P-Value	FDR
UP_KW_BIOLOGICAL_PROCESS	Differentiation	RT		59	8,0	5,8E-11	5,5E-9
UP_KW_BIOLOGICAL_PROCESS	Wnt signaling pathway	RT		19	2,6	3,1E-5	1,5E-3
UP_KW_BIOLOGICAL_PROCESS	Cell adhesion	RT		31	4,2	1,8E-4	4,9E-3
UP_KW_BIOLOGICAL_PROCESS	Neurogenesis	RT		22	3,0	2,0E-4	4,9E-3
UP_KW_BIOLOGICAL_PROCESS	Notch signaling pathway	RT		8	1,1	6,4E-4	1,2E-2
UP_KW_BIOLOGICAL_PROCESS	Myogenesis	RT		7	0,9	1,7E-3	2,7E-2
UP_KW_BIOLOGICAL_PROCESS	Transcription regulation	RT		75	10,2	2,6E-2	3,2E-1
UP_KW_BIOLOGICAL_PROCESS	Angiogenesis	RT		10	1,4	2,7E-2	3,2E-1
UP_KW_BIOLOGICAL_PROCESS	Transcription	RT		76	10,3	3,4E-2	3,6E-1
UP_KW_BIOLOGICAL_PROCESS	Inflammatory response	RT		12	1,6	4,1E-2	3,9E-1
UP_KW_BIOLOGICAL_PROCESS	Potassium transport	RT		8	1,1	7,5E-2	6,5E-1
UP_KW_BIOLOGICAL_PROCESS	Chondrogenesis	RT		3	0,4	1,0E-1	7,9E-1

B Functional Annotation Chart: Nfix2_S301A vs KO_down

Category	Term	RT	Genes	Count	%	P-Value	FDR
UP_KW_BIOLOGICAL_PROCESS	Differentiation	RT		59	6,9	1,9E-9	1,9E-7
UP_KW_BIOLOGICAL_PROCESS	Neurogenesis	RT		23	2,7	2,6E-4	1,3E-2
UP_KW_BIOLOGICAL_PROCESS	Cell adhesion	RT		31	3,6	7,7E-4	2,3E-2
UP_KW_BIOLOGICAL_PROCESS	Transcription regulation	RT		90	10,6	9,0E-4	2,3E-2
UP_KW_BIOLOGICAL_PROCESS	Transcription	RT		90	10,6	2,2E-3	4,4E-2
UP_KW_BIOLOGICAL_PROCESS	Myogenesis	RT		7	0,8	2,7E-3	4,4E-2
UP_KW_BIOLOGICAL_PROCESS	Wnt signaling pathway	RT		15	1,8	6,0E-3	8,5E-2
UP_KW_BIOLOGICAL_PROCESS	Ion transport	RT		34	4,0	8,1E-3	1,0E-1
UP_KW_BIOLOGICAL_PROCESS	Potassium transport	RT		10	1,2	1,9E-2	2,1E-1
UP_KW_BIOLOGICAL_PROCESS	Notch signaling pathway	RT		6	0,7	2,2E-2	2,1E-1
UP_KW_BIOLOGICAL_PROCESS	Growth arrest	RT		4	0,5	2,3E-2	2,1E-1
UP_KW_BIOLOGICAL_PROCESS	Angiogenesis	RT		10	1,2	4,3E-2	3,6E-1
UP_KW_BIOLOGICAL_PROCESS	Biological rhythms	RT		10	1,2	7,4E-2	5,7E-1

C Functional Annotation Chart: Nfix2_S301D vs KO_down

Category	Term	RT	Genes	Count	%	P-Value	FDR
UP_KW_BIOLOGICAL_PROCESS	Differentiation	RT		48	6,7	4,6E-7	4,2E-5
UP_KW_BIOLOGICAL_PROCESS	Neurogenesis	RT		25	3,5	2,9E-6	1,3E-4
UP_KW_BIOLOGICAL_PROCESS	Transcription regulation	RT		84	11,7	9,3E-5	2,8E-3
UP_KW_BIOLOGICAL_PROCESS	Transcription	RT		85	11,8	1,5E-4	3,4E-3
UP_KW_BIOLOGICAL_PROCESS	Notch signaling pathway	RT		8	1,1	4,4E-4	8,1E-3
UP_KW_BIOLOGICAL_PROCESS	Potassium transport	RT		10	1,4	7,8E-3	1,2E-1
UP_KW_BIOLOGICAL_PROCESS	Cell adhesion	RT		24	3,3	1,2E-2	1,6E-1
UP_KW_BIOLOGICAL_PROCESS	Ion transport	RT		29	4,0	1,7E-2	2,0E-1
UP_KW_BIOLOGICAL_PROCESS	Wnt signaling pathway	RT		12	1,7	2,6E-2	2,6E-1
UP_KW_BIOLOGICAL_PROCESS	Myogenesis	RT		5	0,7	3,3E-2	3,0E-1
UP_KW_BIOLOGICAL_PROCESS	Sodium transport	RT		8	1,1	5,0E-2	4,1E-1
UP_KW_BIOLOGICAL_PROCESS	Biological rhythms	RT		9	1,3	7,8E-2	5,9E-1
UP_KW_BIOLOGICAL_PROCESS	Growth arrest	RT		3	0,4	9,9E-2	6,9E-1

Figure 26. Comparison of downregulated GO terms among the different Nfix-expressing myoblast lines versus *Nfix*-null myoblasts

A) Gene Ontology enrichment of downregulated genes relative to the Nfix2_WT-expressing *Nfix*-null myoblasts versus *Nfix*-null myoblasts comparison.

B) Gene Ontology enrichment of downregulated genes relative to the Nfix2_S301A-expressing *Nfix*-null myoblasts versus *Nfix*-null myoblasts comparison.

C) Gene Ontology enrichment of downregulated genes relative to the Nfix2_S301D-expressing *Nfix*-null myoblasts versus *Nfix*-null myoblasts comparison.

4 Discussion

NFI transcription factors have been known as important players in DNA replication and transcription since late 1900s (Gil et al., 1988; Jones et al., 1987; Kruse and Sippel, 1994; Mermod et al., 1989; Santoro et al., 1988), with several studies demonstrating their central role in developmental processes across several tissues (Chaudhry et al., 1997; Driller et al., 2007; Fane et al., 2017; Gil et al., 1988; Gronostajski, 2000; Piper et al., 2011). More recently, Nfix involvement in pathological processes was revealed (Piper et al., 2019). For example, Nfix gene mutations have been identified as the genetic cause of two rare syndromes, the Malan and Marshall-Smith syndromes (Klaassens et al., 2015; Malan et al., 2010; Martinez et al., 2015; Mulder et al., 2020; Priolo et al., 2018; Yoneda et al., 2012). Differential levels of Nfix expression can be either a positive or negative prognostic marker in different types of cancer, influencing proliferation and migration of cancer cells (Fane et al., 2017; Li et al., 2020; Liu et al., 2020; Rahman et al., 2017; Liu et al., 2017; Mao et al., 2015; Ray et al., 2013; Ge et al., 2018). Finally, Nfix levels in dystrophic mice influence the severity of the degenerative process typical of this ominous disease (Rossi et al., 2017). Thus, it is not surprising that the research activity focused on Nfix activity, both regarding developmental and pathological mechanisms, is constantly increasing. So far, most Nfix-centered studies mainly investigated those pathways disrupted by total or partial genetic ablation of Nfix expression in different tissues. Thanks to these studies, we now know that Nfix is central to maintaining the correct balance between quiescence, proliferation, and differentiation of progenitor and stem cells in developing and adult tissues, through transcriptional and epigenetic control of gene expression (Adam et al., 2020; Campbell et al., 2008; Clark et al., 2019; Driller et al., 2007; Heng et al., 2014; Holmfeldt et al., 2013; Martynoga et al., 2013; O'Connor et al., 2015; Piper et al., 2011; Pjanic et al., 2013; Zhou et al., 2015). These effects partly explain the outcomes observed in pathological conditions that involve Nfix. However, in some cases, such as the heterozygous Nfix mutations that cause Marshall-Smith syndrome, we deal with complex biological consequences that cannot be explained by the total absence of Nfix protein, rather by uncharacterized effects of mutant Nfix protein expression (Malan et al., 2010; Martinez et al., 2015; Mulder et al., 2020). Furthermore, unique and sometimes opposite roles of Nfix can be distinguished between developing and adult stem cells of different tissues (Harris et al., 2015). Indeed, the activity of many transcription factors is not only regulated by changes in their expression levels, which is a rather slow regulatory mechanism. A large and important branch of transcriptional factors' regulation is played by post-translational modifications (PTMs). However, as of today, this aspect is poorly

characterized for Nfix.

With this study, we demonstrate that Nfix is post-translationally modified and in particular it is subjected to phosphorylation. We were able to experimentally confirm at least one phosphorylated serine residue, serine 301, which is phosphorylated in skeletal muscle cells, both *in vitro* and *in vivo*, as well as in other cell types. Ser 301 is also well conserved among Nfix vertebrate orthologs, suggesting functional importance. In cultured prenatal and postnatal myoblasts, and in skeletal muscle tissue, we can detect both Ser 301 phosphorylated and nonphosphorylated forms. The detection is facilitated by a characteristic mobility shift in SDS-PAGE experiments, which is mediated by the negative charge carried by the phosphate group. This effect, besides being experimentally useful, could indirectly be linked to structural differences between the phosphorylated and nonphosphorylated Nfix species. Indeed, Ser 301 is homed in a region of Nfix predicted to be highly disordered, thus the addition of the phosphate negative charges could influence the structure of this region, and potentially regulate interaction with partner proteins. The impediment or forced simulation of S301 phosphorylation via site-directed mutagenesis (S301A and S301D mutations, respectively) does not affect either the protein expression or its nuclear localization in myoblasts, indicating that both forms retain key functions.

Strikingly, Nfix phosphorylation at Ser 301 is absent in freshly-isolated uncultured MuSCs. Instead, in cultured myoblasts and whole muscle lysates, we always observe both phosphorylated and nonphosphorylated Nfix. Furthermore, a general and significant increase in Nfix protein levels is observed in cultured versus freshly isolated MuSCs. It is important to notice that cultured MuSCs undergo activation upon isolation from the tissue (Machado et al., 2017, 2021; van Velthoven et al., 2017), and further changes are induced upon MuSCs seeding in culture dishes (Fukada et al., 2007). This is due, at least in part, to activation of stress signals that irreversibly change the transcriptome profile of cultured and activated MuSCs (Fukada et al., 2007; Machado et al., 2017, 2021; van Velthoven et al., 2017), and to growth factors that activate signaling pathways, mediated also by phosphorylation cascades (Baghdadi et al., 2018; Evano and Tajbakhsh, 2018; Jones et al., 2005; L'Honoré et al., 2018; Relaix et al., 2021; Zhou et al., 2022). Several groups demonstrated roles for Nfix in the maintenance of neural stem cells quiescence (Martynoga et al., 2013), proliferation (Fraser et al., 2019), migration (Evelyn Heng et al., 2015), and differentiation (Clark et al., 2019; Heng et al., 2014; Piper et al., 2011). Therefore, our observations suggest that pS301-Nfix may be implicated in the MuSCs late activation processes. Interestingly, we recently demonstrated that Nfix expression during prenatal myogenesis is regulated by JunB (Taglietti

et al., 2018), a member of the AP-1 family induced upon MuSCs activation (Machado et al., 2017, 2021; Pietrosemoli et al., 2017). Since Nfix appears phosphorylated both during prenatal myogenesis, in fetal myoblasts, and in proliferating and differentiating postnatal myoblasts in culture and in skeletal muscles *in vivo*, pS301-Nfix may fulfill a common function that is required in all these situations but not in quiescent/early activated MuSCs. From our experiments, we cannot conclude whether all cultured MuSCs co-express both phosphorylated and nonphosphorylated Nfix or rather if we are dealing with a mix of two cell populations, one mainly expressing the phosphorylated Nfix and the other one the nonphosphorylated Nfix. Antibodies specifically recognizing pS301-Nfix could help us answer this question, but unfortunately, they are not yet available.

Unfortunately, RNA-seq data failed to highlight pathways and genes specifically upregulated or downregulated only by the phosphomimetic Nfix mutant. This could be explained by the fact that the serine-to-aspartate substitution is not sufficient to mimic the transcriptional activity of pNfix, as the phosphate group could mediate important functional properties that cannot be recapitulated by the aspartate. Furthermore, the cultured conditions of the myoblasts we used for the RNA-seq analysis could have masked potential transcriptome differences that may occur in the passage from freshly isolated to cultured MuSCs, when we observe the induction of Ser 301 Nfix phosphorylation. Nevertheless, we found that, in general, Nfix expression in cultured MuSCs-derived myoblasts potentially regulates the expression of genes involved in the control of cell cycle, cell adhesion, differentiation, Notch and Wnt signaling, and metabolism: all pathways differentially regulated between quiescent/early activated MuSCs and actively cycling myoblasts (Evano and Tajbakhsh, 2018; Relaix et al., 2021; Zhou et al., 2022). More precisely, we found that S301A mutant myoblasts display the highest number of upregulated genes significantly enriched in cell cycle, cell division, and mitosis processes. In particular, S301A, but not Nfix2_WT or S301D, myoblasts significantly upregulate *Rgs2* and *Bmp2* genes, that are associated with quiescent/early activated MuSCs (Fukada et al., 2007).

Early studies revealed a possible involvement of the insulin pathway in the regulation of NFI phosphorylation (Cooke and Lane, 1999). Furthermore, PI3K-Akt-mTOR pathways have been implicated in the regulation of quiescence versus activation of MuSCs; indeed, PI3K-Akt-mTORC1 signaling is inactive under quiescence but rapidly induced upon activation (Zhou et al., 2022). However, we excluded the AKT and mTOR signaling cascades as upstream regulators of Nfix Ser 301 phosphorylation in cultured MuSCs. Another signaling cascade involved in MuSCs activation and proliferation is guided by the MAPK ERK1/2

(Machado et al., 2021; Zhou et al., 2022). We recently demonstrated that Nfix protein levels are influenced by pERK levels (Taglietti et al., 2018). However, retrospectively looking at these data, we can observe that inhibition of pERK causes a decrease of both phosphorylated and nonphosphorylated Nfix. Thus, pS301-Nfix levels are not directly regulated by pERK either. Therefore, the kinase and signaling pathways responsible for this phosphorylation remain to be discovered.

Of interest, NFI phosphorylation state was found different between quiescent and actively growing 3T3-L1 adipocytes, an effect dependent at least in part on c-Myc, a known factor involved in cell proliferation and differentiation (Yang et al., 1993). Importantly, c-Myc was identified as one of the early induced genes involved in the activation of freshly-isolated MuSCs (Machado et al., 2017, 2021). Functional analyses demonstrated that NFIs were able to alter the interaction of reconstituted nucleosomal cores with DNA in vitro in a growth factor-dependent manner (Alevizopoulos et al., 1995), possibly linking the epigenetic activity of Nfix with its phosphorylation status.

Reactive oxygen species (ROS) are partially reduced forms of molecular oxygen that are formed as by-products of aerobic metabolism in all existing organisms. Excessive ROS production is harmful and is observed in several conditions, such as neurodegeneration, aging, cancer, and muscular dystrophies. However, some ROS, such as H_2O_2 and the NO radical, at small and local physiological levels can act as signaling molecules. Cysteine (Cys) and methionine (Met) are two sulfur-containing amino acids susceptible to oxidation, both in their free form and protein-bound state. In the latter case, oxidation represents a reversible protein PTM and a sophisticated method to regulate protein function based on their redox state (Drazic and Winter, 2014; Kim et al., 2014). The redox status of the cells may influence protein phosphorylation by different means: it may impact the kinase and phosphatase activity or directly regulate the interplay between the target protein oxidation and phosphorylation. In particular, oxidation of methionines was documented to interface with surrounding phosphorylation sites. Interestingly, one of the very few methionine residues in the Nfix amino acidic sequence is found near serine 301: methionine 286. Proteomic studies derived from H_2O_2-stressed Jurkat cells highlighted a significant overrepresentation of acidic amino acids (glutamate and aspartate) in the neighborhood of oxidized methionines in proteins (Aledo and Aledo, 2020). Strikingly, Met 286 of Nfix is surrounded on both sites by the negatively charged amino acid glutamate. Interestingly, we found that S301-Nfix phosphorylation is negatively affected by hydrogen peroxide (H_2O_2) treatment *in vitro*. Thus, an interplay between Ser 301 phosphorylation and Met 286 oxidation could exist, and Nfix

may act as a redox sensor via differential phosphorylation of Ser 301. Yet, pS301-Nfix levels do not change between wild-type and *sgca*-null dystrophic muscles when analyzed via western blot. However, as muscle tissue lysates contain proteins from several cell types, the effect may be diluted and difficult to observe in such a diversified sample. For example, we show that macrophages display a single Nfix-immunoreactive band, thus Nfix might not be phosphorylated in macrophages. This could confound the results in whole dystrophic muscle lysates, where macrophages are present in higher numbers. Alternatively, it is possible that ROS levels in dystrophic muscles from *sgca*-null mice do not reach the concentration needed to observe the effect on Nfix phosphorylation, or that compensatory mechanisms that prevent this effect occur *in vivo*. ROS have also been characterized as secondary messengers linking MuSCs metabolic state with their differentiation propensity (L'Honoré et al., 2018; Relaix et al., 2021). MuSCs in different phases of their myogenic program (quiescence vs. activation vs. proliferation vs. differentiation) rely on different energy sources. Indeed, there is a progressive metabolic shift from fatty acid oxidation (FAO) and oxidative phosphorylation (OXPHOS) of quiescent MuSCs, to glycolysis of proliferating myoblasts, to OXPHOS again during differentiation (Relaix et al., 2021). Interestingly, RNAseq data highlighted a possible role for Nfix in the regulation of myoblasts metabolism, as Nfix phosphomutants appear to positively regulate lipid metabolism-related genes. Since Nfix appears differentially phosphorylated between early and late activated MuSCs, the effect of hydrogen peroxide treatment on Nfix phosphorylation could have an explanation linked to the myoblasts' activation and metabolic state control.

We also discovered that treatment with proteasome inhibitors causes a drop in pS301-Nfix levels. The main effect of proteasome inhibition is the prevention of protein degradation, but several secondary consequences can be caused as a result of this primary effect. A common aspecific outcome is the mitotic arrest of treated cells since cell cycle progression highly relies on protein degradation (Varetti et al., 2011; Visconti et al., 2010; Zeng et al., 2010). As discussed previously, Nfix is involved in cell cycle progression and proliferation in several cell types, and recent findings from Clark et al. revealed that NFI factors promote cell-cycle exit during retinal development (Clark et al., 2019). Our RNA-seq data obtained from Nfix-expressing and *Nfix*-null myoblasts confirm the involvement of Nfix in the control of cell cycle progression, and in particular in the regulation of mitotic processes. Interestingly, S301A mutant myoblasts have a more evident significant enrichment of cell cycle, cell division, and mitosis-related gene sets. It remains to be elucidated if the significant decrease in Nfix phosphorylation that we observe upon proteasome inhibition may be related

to specific and nonoverlapping control of different aspects of cell cycle progression by phosphorylated and nonphosphorylated Nfix.

Ser 301 is highly conserved among Nfix orthologs of different vertebrate species and is present in all Nfix splicing isoforms, as it is homed in the exon 6 of Nfix, which is not subjected to alternative splicing. In this work, we focused on the murine Nfix2 isoform in skeletal muscle cells, but considering the transversal presence of this phosphorylation site, our findings may be translated also to different systems. Furthermore, several other Nfix phosphorylation sites emerge from proteomic studies from different cells and tissues, opening new and unexplored questions on their biological role and implications in pathological conditions. Potentially, other phosphorylation sites on the Nfix protein sequence could contribute to finely and timely regulate the processes described here. Of interest, one of these sites, Ser 265, was detected as phosphorylated in our phosphoproteomic analysis of murine tibialis anterior muscles.

Dysregulation of Nfix phosphorylation could be implicated in the pathogenesis of Malan and Marshall-Smith syndromes, two rare genetic diseases with heterozygous mutations on the Nfix gene (Malan et al., 2010; Marshall et al., 1971; Martinez et al., 2015; Rai et al., 2018; Yoneda et al., 2012). Particularly, Marshall-Smith syndrome mutations involve the same exons interested by most of the phosphorylation sites, including Ser 301: exons 6-8 (Zenker et al., 2019).

To conclude, our work, despite not conclusive and still open, highlights the importance of studying Nfix PTMs, and in particular phosphorylation, as understanding the molecular pathways involved could give transversal answers to Nfix highly diversified roles in different contexts, spanning from developmental and adult stem cells biology to cancer and genetic syndromes.

5 Future Perspectives

As a continuation of this research work, our main goal would be to define the biological role of Nfix phosphorylation in MuSCs. To this aim, future studies will try to address the specific relevance of Ser 301 phosphorylation in the context of MuSCs activation, via RNA-seq experiments comparing the quiescent, freshly-isolated MuSCs with the actively cycling MuSCs, in wildtype versus Nfix-null contexts. This experiment would allow us to distinguish the different contributions of non-phosphorylated (in quiescent MuSCs) and phosphorylated (in actively cycling MuSCs) endogenous Nfix. With this type of experiment, we could observe

the contributions of both Nfix isoforms expressed by the muscle at physiological levels, contrary to our current RNA-seq experiment in which we overexpressed only isoform 2 in a *Nfix*-null background. Since also isoform 1 of Nfix is normally expressed in MuSCs, and since Ser 301 is present also in this isoform, it is possible that with the experimental setting used in this study we are missing some important information on the function of S301 phosphorylation of Nfix1 isoform. In the future, silencing strategies could be employed to distinguish between specific roles of Nfix1 and Nfix2 isoforms in skeletal muscle, by targeting the exon 7, which is present in isoform 1 but absent in isoform 2.

In addition, experiments of in vivo muscle injury and subsequent isolation of MuSCs from injured and uninjured mice will be performed. This would reveal if Nfix phosphorylation is induced also in activated and cycling MuSCs *in vivo*, as observed in MuSCs *in vitro*. Furthermore, MuSCs will be isolated from *sgca*-null mice to remove the contribution of the other cell types present in the muscle that may have masked the different phosphorylation status of Nfix in dystrophic versus wild type MuSCs.

Finally, we plan to perform transfection experiments of the Nfix1 and Nfix2 S301 mutants in single fibers preparatives from *Nfix*-null mice and/or electroporation of Nfix plasmids in muscles of *Nfix*-null animals. With these experiments, we aim to decipher the downstream effects caused by the manipulation of Nfix S301 phosphorylation in the MuSCs activation process.

6 Materials and methods

6.1 Animal experimentation

For this work, I used the following mouse lines: the CD1 and C57BL/6 wild-type mice (Charles River), the *Nfix*-null mouse model obtained from Prof. Richard M. Gronostajski (University of Buffalo) (Driller et al., 2007), and the *sgca*-null dystrophic mouse model (Duclos et al., 1998).

Adult 5-weeks old wildtype CD1 mice were used to obtain muscles for the phosphoproteomic and western blot analyses. C57BL/6 wild-type and *Nfix*-null 13-days old (P13) animals were used to obtain secondary MuSCs-derived myoblasts. Adult 8-weeks old *sgca*-null and control C57BL/6 wild-type mice were used to obtain muscles samples. Adult 8-weeks old C57BL/6 wild-type mice were used to obtain freshly-isolated MuSCs and cultured MuSCs-myoblasts.

Female CD1 mice were mated with males (2:1) and examined every morning for copulatory plugs. The day on which a vaginal plug was seen was designated as 0.5 gestation day (E0.5). All the female mice used for the experiments were at least 7 weeks old. Fetal myoblasts were isolated from E16.5 fetuses.

All the mice were kept under pathogen-free conditions with a 12-h/12-h light/dark cycle. All the procedures on animals are conformed to Italian law (D. Lgs n. 2014/26, as the implementation of 2010/63/UE) and were approved by the University of Milan Animal Welfare Body and by the Italian Ministry of Health.

6.2 Muscle isolation

For the isolation of muscles from adult mice for phosphoproteomic and western blot analyses, mice were sacrificed via cervical dislocation, muscles were isolated, transferred in 2mL tubes, and immediately immersed in liquid nitrogen. Samples were then stored at -80°C upon use.Protein extraction and preparation for phosphoproteomics

6.3 Cell culture

6.3.1 C2C12

The C2C12 cell line, originally derived from *Mus musculus* myoblasts, was maintained in culture conditions at 37.0°C, 5% CO_2. The base medium for this cell line is Dulbecco's

Modified Eagle's Medium (DMEM), supplemented with 10% Fetal Bovine Serum (FBS), 1% of Penicillin/streptomycin, and 1% Glutamine (referred to as growth medium, GM). Cell growth was monitored daily, and cells were stopped and detached with trypsin (Sigma T4174) once they reached the desired confluence. The seeding density required for proliferation analysis is around $5.0x10^3$ viable cells/cm^2. To proceed with the molecular analysis (protein extraction) the cells were briefly washed with cold PBS 1X (EuroClone ECB4004L) and frozen at -20°C.

6.3.2 Fetal myoblasts

Fetal myoblasts were isolated from wild-type (CD1) E16.5 fetuses. After mother sacrifice, fetuses were extracted from the placenta, and the vitelline yolk was removed. The heads were immediately separated from the body, and the fetuses were dissected under a stereomicroscope in a Ø=150mm petri with PBS 1X (EuroClone). All internal organs and the skin were completely removed, leaving only the skeleton with the muscles attached to it. The dissected fetuses were then mechanically and enzymatically digested in 50-ml tubes with 8 ml of Digestion Solution every two fetuses, containing 0.15 mg/mL of Collagenase Type I (Sigma), 1.5 mg/mL of Dispases (Gibco) and 0.1 mg/mL of DNases (Roche) in PBS 1X at 37°C under strong shaking for 30 minutes. Multiple (2-3) digestions rounds were performed until complete dissociation was reached. In our conditions, 2 or 3 cycles are usually sufficient. After every digestion cycle, the cells in suspension were gathered into a new tube with an FBS-containing medium (see **Table 1**) to inactivate the enzymes and cells were centrifuged at 900 rcf for 5 minutes at 4°C, the supernatants discarded, and the cellular pellet was resuspended into 2ml of erythrocytes lysis buffer (see **Table 2**) for 2' in ice. The cell suspension was diluted with 5ml of PBS 1X and centrifuged at 900 rcf at 4°C. The cells were resuspended in 20% HS (see **Table 2**) then filtered using a 40-µm filter. Finally, the cell suspension was pre-plated on Ø=150mm plastic petri dishes (about one petri for 2-3 fetuses) in a total volume of 15 ml of Growth Medium (see **Table 2**) for 1h at 37°C to deplete the cell suspension of fibroblasts. The enriched fetal myoblast preparative is finally plated on collagen-coated (Sigma C9791) petri at a density of 45.000-50.000 cells/cm^2. After 24 hours, the fetal myoblasts should be adherent, and differentiation can be induced by changing the medium with a differentiation medium (see **Table 2**) and maintaining the cells in culture for 48-72h.

For protein or RNA extraction studies, cells were briefly washed with cold PBS 1X and frozen at -20°C.

Table 2. Fetal myoblasts media and buffer composition.

Growth medium	Differentiation medium	Erythrocyte lysis solution
20% Horse Serum (HS)	2% HS	NH_4Cl (0.15M)
1% Penicillin/Streptomycin (P/S) (Euroclone, ECB3001D)	1% P/S	$KHCO_3$ (10mM)
1% Glutamine	1% Glutamine	EDTA tetrasodium salt hydrate (0.1mM)
0.1% Gentamycin	0.1% Gentamycin	
To volume with DMEM high glucose (EuroClone)	To volume with DMEM high glucose	To volume in PBS 1X

All media were filtered and stored at 4°C upon use. Glutamine was re-added every two weeks.

6.3.3 Primary myoblasts

6.3.3.1 Standard protocol

Postnatal myoblasts were obtained from postnatal day 13 (P13) *Nfix*-null and wt CD1 mouse pups or from 5-weeks old wt CD1 mice. All limbs, pectoralis, and diaphragm muscles were collected from each mouse under sterile conditions and placed into DMEM, 50% FBS, 1%P/S medium. Muscles were cleaned in a culture dish containing DMEM/F12, 1% P/S under a stereomicroscope in a horizontal laminar flow hood to remove all non-muscular tissues in sterile conditions. Muscles were mechanically and enzymatically digested for 10 min at 37°C under strong agitation with digestion medium (see **Table 3**) and dissociated cells were collected in growth medium (see **Table 3**). After 3-4 digestion rounds, cells were centrifuged at 300g for 7 minutes, the supernatant was discarded, and the pellet was resuspended in proliferation medium (see **Table 3**). Cells were then filtered using the 40-µm filter and subjected to negative selection to remove non-myogenic cells using the Miltenyi MACS® Columns and Satellite cells Isolation Kit (Miltenyi Biotec # 130-104-268), according to manufacturer instructions. The isolated cells were then pre-plated on Ø=150mm plastic dishes for 1h at 37°C to allow residual fibroblast adhesion. The final estimated percentage of residual fibroblasts in culture is less than 5%. The non-adherent cells were collected and centrifuged at 300g for 7 minutes. The recovered myoblasts were plated at low density (500 cells/cm^2) in culture plates coated with 0.02% gelatin (Sigma G8150) with proliferation medium to allow expansion. The proliferation medium was completely replaced every 24h/48h.

Table 3. SC-derived myoblasts media and buffer composition

Proliferation medium	Differentiation medium	Digestion medium
DMEM/F12 (Gibco 31331-028)	DMEM/F12	DMEM/F12
20% FBS	2% HS	0.14% Protease (Sigma P8811)
2,5 ng/ml bFGF (peprotech)	1% P/S	1% P/S
1% P/S		

6.3.3.2 Freshly isolated versus cultured MuSCs protocol

For freshly isolated MuSCs, a modified version of the T3 protocol from (Machado et al., 2017) was performed. Briefly, wild type 8-weeks-old mice were sacrificed by cervical dislocation and all limbs and pectoralis muscles were dissected and cleaned from fat and other associated tissue. Muscles were then immediately chopped with scissors for up to 10min until a homogeneous, paste-like slurry was formed. The muscle preparation was then transferred into a 50ml tube with 30 ml of ice-cold DMEM. Once all samples were collected, muscle slurries were centrifuged at 4°C for 5min/600g and the supernatant discarded. The cell pellet was resuspended in 20ml of warm (37°C) digestion solution (Dispase II, 3U/ml, Collagenase A 0.5U/ml, 0.2% BSA, 100 U/ml of penicillin, and 100 mg/ml of streptomycin) and incubated for 2h at 37°C in a shaking water bath. The digested muscle was spun at 4°C for 5min/600g and the pellet was resuspended with 1ml of ice-cold DMEM with a P1000 pipette. Additional 19 ml of ice-cold DMEM were added. The digestion was mixed gently by inverting the tube 5-10 times, then filtered through a 70µm cell strainer into a new 50ml tube, and spin at 50g for 5min at 4°C. The supernatant was filtered again through a 40µm cell strainer and spun at 4°C for 5min/600g. The pellet was washed twice with ice-cold PBS and then subjected to negative selection to remove non-myogenic cells using the Miltenyi MACS® Columns and Satellite cells Isolation Kit (Miltenyi Biotec # 130-104-268) as described previously. The final estimated percentage of residual fibroblasts is around 5%. Freshly isolated MuSCs were immediately lysed in RIPA buffer as described in section 6.8. For cultured MuSCs-derived myoblasts, MuSCs were plated in a 35-cm dish coated with gelatin 0.02% in proliferation medium (**Table 3.** SC-derived myoblasts media and buffer composition. The proliferation medium was completely replaced every 24h and cells were allowed to attach and proliferate for 62h before lysis with RIPA buffer.

6.4 Treatments

Cells were treated with the following molecules:

Table 4. List of inhibitors/reagent used in this study

Inhibitor/reagent type	Inhibitor name	Manufacturer	Concentration
Glycosylation inhibitors	Tunicamycin	Sigma-Aldrich	1 μg/mL
Kinase inhibitors	LY294002 (PI3K inh.)	Sigma-Aldrich	10 μM
	PP242 (mTORC inh.)		500 nM
Proteasome inhibitors	mg132	Sigma-Aldrich	10 μM
	mg110	Sigma-Aldrich	0.3, 1, 3, 10 μM
	Bortezomib	EMD Millipore	1 μM
Oxidative stress	H_2O_2	Sigma-Aldrich	250 nM, 500 nM

6.5 Site-directed mutagenesis

Point mutations were introduced on Nfix2 protein sequence using the QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies, cat. 200524). The plasmid pLenti.HA-Nfix2 (**Figure 27**) was used as dsDNA template. For each mutation site, a couple of specific complementary primers, both containing the desired mutation, were designed, using the web-based QuickChange Primer Design Program. The sequence of the designed primers is specified in the table below (**Table 5**), with mutated bases highlighted in capital bold letters.

Table 5. List of primers used for mutagenesis experiments

Primer name	Primer sequence
Nfix2_S265/268A_FOR	ccctgggaaggcgg**G**ccatcacc**G**cccctccttccacc
Nfix2_S265/268A_REV	ggtggaaggaggggg**C**ggtgatgg**C**ccgccttcccaggg
Nfix2_S280A_FOR	cagcaccaagcgccccaag**G**ccatcgacgacagtgagatg
Nfix2_S280A_REV	catctcactgtcgtcgatgg**C**cttggggcgcttggtgctg
Nfix2_S288A_FOR	cgacgacagtgagatggag**GC**tccagtagatgatgtgttc
Nfix2_S288A_REV	gaacacatcatctactgga**GC**ctccatctcactgtcgtcg
Nfix2_S301A_FOR	cctgggacaggccgc**G**ctccggccgctggcag
Nfix2_S301A_REV	ctgccagcggccggag**C**gcggcctgtcccagg
Nfix2_S320A_FOR	gacgtggatgcaggg**GC**cccccgggccacgg
Nfix2_S320A_REV	ccgtggcccgggggg**GC**ccctgcatccacgtc
Nfix2_S301D_FOR	cctgggacaggccgc**GA**tccggccgctggcag
Nfix2_S301D_REV	ctgccagcggccgga**TC**gcggcctgtcccagg

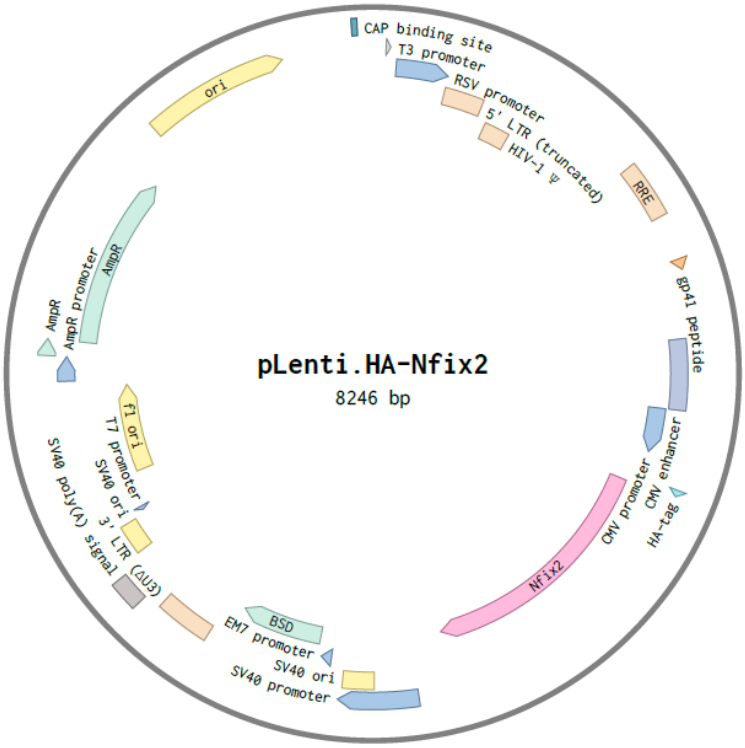


Figure 27. Map of pLenti_HA-Nfix2 plasmid

Sample reaction:

- µl of 10× reaction buffer
- 0.5 µl (50 ng) of dsDNA template (100 ng/µl)
- 1.25 µl (125 ng) of oligonucleotide primer FOR
- 1.25 µl (125 ng) of oligonucleotide primer REV
- 1 µl of dNTP mix
- 40 µl ddH2O to a final volume of 50 µl
- 1 µl *PfuUltra* High Fidelity DNA polymerase (2.5 U/µl)

A rolling circle amplification was performed to introduce the desired mutations using the cycling parameters outlined here:

Table 6. Thermal protocol for mutagenesis

Segment	Cycles	Temperature	Time
Pre-denaturation	1	95°C	30 seconds
Denaturation	16	95°C	30 seconds
Annealing		55°C	1 minute
Elongation		68°C	10 minutes
Final elongation	1	68°C	12 minutes
End	1	4°C	∞

Following the thermal protocol, 1 µl of the Dpn I restriction enzyme (10 U/µl) was added directly to each amplification reaction and incubated at 37°C for 1h. Dpn I is a specific endonuclease used to digest the parental methylated DNA template and to select mutation-containing *in vitro* synthesized DNA.

Finally, the nicked vector DNA containing the desired mutations was transformed into XL1-Blue supercompetent cells. Briefly, 50 µl of the supercompetent cells were aliquoted to a prechilled 14-ml BD Falcon polypropylene round-bottom tube and 3 µl of the Dpn I-treated DNA from each reaction was transferred to separate aliquots of the supercompetent cells. The transformation reactions were swirled gently to mix and incubated on ice for 30 minutes. Then they were pulse-heated for 45 seconds at 42°C and placed on ice for 2 minutes. 0.5 ml of LB broth preheated to 37°C was added and the transformation reactions were incubated at 37°C for 1 hour with shaking at 190 rpm. For each transformation reaction, cells were spread on LB–ampicillin agar plates and incubated at 37°C for >16 hours. Colonies were picked and let grow in LB with Ampicillin (50 µg/ml) overnight. A portion was stocked in 30% glycerol, while the rest was used to isolate the plasmid DNA.

6.5.1 Mini- and Maxi-prep

For mini-prep, single colonies were collected and inoculated in 15mL falcon containing 3 mL of LB and Ampicillin (50 µg/ml) and kept on shaking O/N at 37°C.

To isolate plasmid DNA, PureYieldTM Plasmid Mini-prep System (PROMEGA) was used. Briefly, the cells were pelleted and resuspended in 600 µl of SALF water and lysed with 100 µl of Cell Lysis Buffer. Cold (4-8°C) Neutralization Solution was added and mixed thoroughly by inverting. The samples were centrifugated at maximum speed for 3 minutes and the supernatant was transferred to a PureYeld TM Minicolumn placed into a Collection Tube and centrifugated at maximum speed for 15 seconds. Two subsequent washes with Endotoxin Removal Wash (ERB) and Column Wash Solution (CWC) were performed. Finally, the plasmid DNA was eluted, quantified, and stored at -20°C.

For the maxi-prep, the transformed cells stocked in glycerol were grown in 300 ml of LB and Ampicillin overnight at 37°C in agitation (180-190 rpm). To obtain plasmid purification, the NucleoBond Xtra Maxi EF Kit (Macherey-Nagel) was used. The bacterial culture was centrifuged at 5000 x g for 15 minutes at 4°C. The supernatant was discarded, and the cell pellet was resuspended and lysed. The precipitate was removed by centrifugation at 5000 x g

for at least 10 minutes. The cleared lysate was loaded on the pre-equilibrated NucleoBond® Xtra Column together with the inserted column filter. The column was then washed thrice, and the plasmid DNA was then eluted and precipitated with isopropanol. Finally, 70 % EtOH was added to wash the DNA pellet and the samples were centrifugated at 5000xg for 5 minutes. Ethanol was removed completely from the tube, allowing the pellet to dry at room temperature for 15 minutes. As the last step, the DNA pellet was dissolved in an appropriate volume of S.A.L.F. water and stored at -20°C.

The samples were quantified through NanoDrop Spectrophotometer (Implen), and the plasmid integrity was confirmed by 7% agarose gel electrophoresis.

6.5.2 Sequencing

Sequencing was carried out by using the TubeSeq Service from Eurofins Genomics. 15 µl of plasmid DNA (50-100 ng/µl) and 2 µl of sequencing primers (10 pg/µl) (**Table 7**) were collected in each Eppendorf signed by a specific TubSeq Label and sent to Eurofins Genomics to be sequenced. The sequencing chromatograms were analyzed with SnapGene.

Table 7. Sequencing primers

Primer name	Primer sequence
Nfix2_seq1_FOR	ggagtcacaatcaaagaact
Nfix2_seq2 (CMV)_FOR	cgcaaatgggcggtaggcgtg

6.6 Cells Transfection

Cells were transfected at 70-80% confluency using the Lipofectamine™ 3000 Transfection Reagent protocol (Invitrogen, L3000001). The medium was changed after 16h. Myoblasts were harvested 48h after transfection.

6.7 Lentiviral production and transduction

Viral particles were prepared in HEK293T cells by using Calcium Phosphate Transfection method. The packaging plasmids pΔ8.4 (16.25 µg), pVSV-G (9 µg) and pRSV-REV (6.25 µg) were co-transfected with the following lenti-plasmids: pLenti_HA-Nfix2, pLenti_HA-Nfix2_S301A, pLenti_HA-Nfix2_S301D, and the pLenti_empty as control. Culture medium containing the viral particles was collected 48h after transfection, filtered with a 0.45 µm filter, and stored as supernatants at -80°C until use.

Cells were transduced with the supernatants with the addition of polybrene (8 μg/mL) for 16h before changing the medium. After an additional 24h, Blasticidin (GeneSpin STS-BLA20) (10 μg/mL) was added for selection.

6.8 Protein extraction for western blot

Protein extracts were obtained from cultured myoblasts lysed for 30' on ice with RIPA buffer (10mM Tris-HCl pH 8.0, 1mM EDTA, 1% Triton-X, 0.1% sodium deoxycholate, 0.1% SDS, 150mM NaCl in deionized water), while total protein extracts from adult muscles were obtained from tissues homogenized and incubated for 30' on ice in Tissue Lysis buffer (50mM Tris-HCl, 1mM EDTA, 1% Triton-X, 150mM NaCl). Both RIPA and Tissue Lysis buffers were supplemented with protease and phosphatases inhibitors (Sodium Orthovanadate, SO, 1mM; Phenyl methane sulfonyl fluoride, PMSF, 1mM; Sodium fluoride, NaF, 50mM; 1X cOmplete Mini Protease Inhibitors Cocktail). Then the tubes were centrifugated at 11000 rcf for 10 minutes at 4°C and the supernatants were collected in new tubes and quantified through the DC Protein Assay (Bio-Rad). The absorbance of each sample was determined spectrophotometrically at 750 nm with the Kaleido software of the Ensight plate reader and the concentration of the samples was calculated using concentration standards.

6.8.1 Phosphatase Treatment

Treatments with λ phosphatase (λ -PPase, sc-200312 Santa Cruz Biotechnology) were performed on protein extracted from C2C12 myoblasts. One set of cells were lysed using RIPA lysis buffer with phosphatase and protease inhibitors (NaF, SO; PMSF and EDTA-free protease inhibitors). Another set of cells was lysed with a RIPA buffer supplemented with protease inhibitors (PMSF and EDTA-free protease inhibitors) but without phosphatase inhibitors. After protein concentration determination, 50 μg of protein per sample were incubated for 30 minutes at 30 °C with 1 μl (400U) λ phosphatase and its buffers with or without phosphatase inhibitors.

	Untreated	Treated	Rescue
λ Phosphatase (400U)	-	+	+
Phosphatase inhibitors	+	-	+
1x λ-PPase and 1x $MnCl_2$ buffers	-	+	+

6.8.2 Western blot

35-50 µg of protein extracts were denatured at 95°C for 5 minutes with SDS Page Loading sample buffer (100mM Tris pH 6.8, 4% SDS, 0.2% Bromophenol blue, 20% Glycerol, and 10mM dithiothreitol) and loaded on 8-10% SDS polyacrylamide gel. After electrophoresis, proteins were blotted to a nitrocellulose membrane (Whatman, Protran Nitrocellulose Transfer Membrane) at 110 mV, 250 mA for 1h-2h. The membranes were blocked in 5% Milk TBST 1X (Tris-buffered saline solution plus 0.02% Tween20) for 1 hour to saturate aspecific sites. Primary antibodies were dissolved in 5% Milk TBST 1X or Signal Boost and were incubated O/N at 4°C under gentle agitation. **Table 8** lists the primary antibodies used for this study. The day after, blots were washed and incubated for 45 minutes at RT with secondary antibodies diluted in 5% Milk TBST 1X (1:10000, IgG-HRP, Bio-Rad) and washed again. Bands were revealed using ECL detection reagent (ThermoFisher) and images were acquired using the ChemiDoc MP System (Biorad). Finally, images were quantified using the software ImageJ.

Table 8. List of western blot primary antibodies used in this study

Antibody	Host	Concentration	Type	Buffer
anti-Nfix	Rabbit	1:5000	Active Motif	Signal Boost
anti-HA.11	Mouse	1:1000	MMS-101P-500 *Covance*	Signal Boost
anti-Vinculin	Mouse	1:2500	D4505 Sigma-Aldrich	Milk 5%
anti-Ubiquitin	Rabbit	1:1000	Santa Cruz	Milk 5%
anti-pAKT	Rabbit	1:1000	Cell Signaling Technology	Signal Boost
anti-tAKT	Rabbit	1:1000	Cell Signaling Technology	Signal Boost
anti-pS6	Rabbit	1:1000	Cell Signaling Technology	Signal Boost
anti-S6	Rabbit	1:1000	Cell Signaling Technology	Signal Boost
anti-MyHCs	Mouse	1:5	DSHB – MF20	Milk 5%
anti-Myod	Mouse	1:1000		Signal Boost
anti-Pax7	Mouse	1:5	DSHB	Milk 5%
anti-Myog	Mouse	1:5	DSHB – F5D	Milk 5%

6.9 Protein extraction and preparation for phosphoproteomics

Tibialis Anterior (TA) muscle biopsy specimens were dissected from 5-weeks-old wildtype CD1 mice and rapidly transferred to liquid nitrogen and stored at –80 °C until analyzed. Muscle samples were then digested with a urea solution containing phosphatase inhibitors (8

M urea, 50 mM TrisHCl, 1 mM sodium ortho-vanadate, 5 mM sodium fluoride, PMSF, 1x protease inhibitor cocktail, 25mM NEM, pH = 8.0) and homogenized in liquid nitrogen with ceramic mortar and pestle. Next, sonication on ice was performed (Intensity: high; Pulses: 30" ON, 30" OFF; Time: 10'-15'). Samples were then centrifuged (20 min, 16,000xg, 4 °C) and supernatants were retrieved and stored at -80°C.

The TA protein samples were prepared for phosphoproteomic analysis through TiO_2 phosphopeptide enrichment. 3 mg of protein sample were diluted with 100 mM of TAEB (300 µl final volume), equilibrated to pH=8 and incubated with 15 µl of TCEP 200 mM for 1h at 55°C. Next, samples were treated with 15 µL of IAA 375 mM for 30' at RT in the dark. Digestion was performed with 45 µL of trypsin 1 µg/µL (enzyme:protein ratio 1:66) at 37°C overnight. The reaction was stopped with the addition of 1 µL TFA 100%. The samples were concentrated with a speedvac and reconstituted in 300 µL di TFA 0.1%. The proteins were then desalted using Pierce™ Peptide Desalting Spin Columns (cat n. 89852) and concentrated again with a speedvac. Phosphopeptides were enriched using the High-Select™ TiO_2 Phosphopeptide Enrichment Kit (cat n. A32993). Concentrated samples were reconstituted in FA 0.1% and analyzed in nanoLC-HRMS (nanoLiquid Chromatography - High Resolution Mass Spectrometry).

6.9.1 Mass Spectrometry analysis

All samples have been analyzed at UNITECH OMICs (University of Milano, Italy) using: Dionex Ultimate 3000 nano-LC system (Sunnyvale CA, USA) connected to Orbitrap Fusion™ Tribrid™ Mass Spectrometer (Thermo Scientific, Bremen, Germany) equipped with nano electrospray ion source. Peptide mixtures were pre-concentrated onto a Acclaim PepMap 100 – 100µm x 2cm C18 (Thermo Scientific) and separated on EASY-Spray column ES802A, packed with Thermo Scientific Acclaim PepMap RSLC C18, 3 µm, 100 Å using mobile phase A (0.1 % formic acid in water) and mobile phase B (0.1% formic acid in acetonitrile 20/80, v/v) at a flow rate of 0.300 µL/min. The temperature was set to 35°C and the samples were injected in duplicates. The sample injection volume is 2 µL.

6.10 RNA isolation

Total RNA was isolated using the TRIzol™ Reagent protocol (Invitrogen). After homogenizing the sample in 1mL TRIzol™ Reagent under a chemical hood, chloroform (1/5 of the total volume = 0.2 mL/1mL of TRIzol™ Reagent) was added, and the tubes were vigorously shaken by hand to allow the solubilization of the RNA in it. After 2-3

minutes of incubation, samples were centrifuged at 12000xg for 15' at 4°C. The clear upper aqueous phase (containing RNA), was recovered and the nucleic acids were precipitated with isopropanol and washed with 70% ethanol. Finally, the reconstituted RNA was quantified with a NanoDrop Spectrophotometer (Implen), and stored at -80°C.

6.10.1 Real-time qPCR

Once RNA was quantified, 0.5 µg of total RNA was retrotranscribed with 5X iScript™ Reverse Transcription Supermix for RT-qPCR (Bio-Rad) with the following thermal protocol:

- 25°C for 5'
- 42°C for 30'
- 85°C for 5'
- The cycles end and the machine reaches 16°C.

The cDNAs obtained were diluted 1:10 in S.A.L.F. water and stored at -20°C.

5 µl of diluted cDNA was used for the Real-Time qPCR, which was performed by using SYBR Green Supermix (Biorad) and the primers listed in **Table 9**. The thermic protocol used was the following:

- Pre-denaturation at 98°C for 30 seconds
- Denaturation at 95°C for 30 seconds x 35-41 cycles.
- Annealing at 61°C for 30 seconds x 35-41 cycles
- Elongation at 72 °C for 10 seconds x 35-41 cycles.
- Melting curve 65-95 °C

β-actin was always present as the reference gene used to normalize the relative mRNA expression level of the genes of interest. Expression values (2^{-ddCt}) were calculated using Excel.

Table 9. List of primers for real-time qPCR used in this study

Target name	Primer sequence
β-actin_FOR	ctctggctcctagcaccatgaaga
β-actin_REV	gtaaaacgcagctcagtaacagtccg
Nfix_FOR	ctggcttactttgtccacactc
Nfix_REV	ccagctctgtcacattccagac
CHOP_FOR	cccaggaggaagaggaggaa
CHOP_REV	cttcatgcgttgcttcccag
EDEM1_FOR	cggtgctctggttggtctt

EDEM1_REV	aagccgaagctgagcggaaa
ORP150_FOR	catcaatgcggacgaagctg
ORP150_REV	ccgcacaacaaatggcttca

6.10.2 RNA-seq analysis

RNA for RNA-seq analysis was extracted as discussed previously. RNA integrity was assessed by the RNA integrity number (RIN) algorithm, using the Agilent 2100 Bioanalyzer and RIN number for all RNA samples was 10. RNA integrity analysis, library preparation, sequencing, and bioinformatic analysis were performed by Novogene Co.,Ltd. Functional annotation was performed using DAVID (https://david.ncifcrf.gov/).

6.11 Immunofluorescence

Cells were washed twice with PBS 1x and fixed for 15 minutes in 4%PFA. After three PBS washes, the membranes were permeabilized with a solution of 0.2% Triton/1% BSA in PBS for 30 minutes and then incubated for 1h with 10% goat serum as a blocking agent. Primary antibodies (**Table 10**) were added to the cells and incubated overnight at 4°C. The day after, the samples were washed with 0.2% Triton/1% BSA in PBS and labeled with fluorescent secondary antibodies for 45 minutes at RT in the dark. Three final washes were performed with 0.2% Triton in PBS prior to mounting and observation. Images were acquired with a Leica-DMI6000B fluorescence microscope equipped with a Leica DFC365FX digital camera. The merges and analyses were performed using the software ImageJ.

Table 10. List of IF primary antibodies used in this study

Antibody	Host	Concentration	Type	Buffer
anti-HA	Rabbit	1:800	Cell-Signaling #3724S	10% Goat Serum
anti-Ki67	Mouse	1:50	BD Biosciences #550609	10% Goat Serum
anti-Pax7	Mouse	Pure	Developmental Studies Hybridoma Bank (DSHB)	Pure
anti-Myod	Mouse	1:50	Santa Cruz Biotech #sc-377460	10% Goat Serum
anti-Myog	Mouse	Pure	DSHB – F5D	Pure
anti-MyHC	Mouse	Pure	DSHB – MF20	Pure

6.12 Statistical analyses

All data are expressed as mean +/- SD. Graphs are obtained using GraphPad Prism 9 (GraphPad Software, San Diego, CA, USA) and analyzed with one-way ANOVA test with Dunnett's correction or two-side Student t-test. * $p<0.01$; ** $p<0.001$; *** $p<0.0001$; confidence interval 95%; alpha level 0.05

7 Acknowledgments

Part of this work was carried out in OMICs, an advanced mass spectrometry platform established by the Università degli Studi di Milano. The RNA-sequencing and bioinformatic analyses were performed by Novogene. The Nfix expression and purification experiments in *E. coli* were performed by Dr. Michela Lapi under Prof. Marco Nardini supervision at Università degli Studi di Milano.

I thank Prof. Stefano Biffo and Dr. Nicola Manfrini for their kind help with the insulin pathway treatment reagents and the precious advice. I thank Dr. Silvia Torchio for the helpful discussion during these years and the reagents exchange.

8 Bibliography

Adam, M.P., Hennekam, R.C.M., Keppen, L.D., Bull, M.J., Clericuzio, C.L., Burke, L.W., Ormond, K.E., and Hoyme, H.E. (2005). Marshall-Smith syndrome: Natural history and evidence of an osteochondrodysplasia with connective tissue abnormalities. Am. J. Med. Genet. *137 A*, 117–124.

Adam, R.C., Yang, H., Ge, Y., Infarinato, N.R., Gur-Cohen, S., Miao, Y., Wang, P., Zhao, Y., Lu, C.P., Kim, J.E., et al. (2020). NFI transcription factors provide chromatin access to maintain stem cell identity while preventing unintended lineage fate choices (Springer US).

Aledo, J.C., and Aledo, P. (2020). Susceptibility of protein methionine oxidation in response to hydrogen peroxide treatment–ex vivo versus in vitro: A computational insight. Antioxidants *9*, 1–24.

Alevizopoulos, A., Dussere, Y., Tsai-Pflugfelder, M., von der Weid, T., Wahli, W., and Mermod, N. (1995). A proline-rich TGF responsive transcriptional activator interacts with histone H3. Genes Dev. 3051–3066.

Allen, D.G., Whitehead, N.P., and Froehner, S.C. (2016). Absence of dystrophin disrupts skeletal muscle signaling: roles of Ca 2+ , reactive oxygen species, and nitric oxide in the development of muscular dystrophy. Physiol. Rev. *96*, 253–306.

Apt, D., Lui, Y., and Bernard, H. (1994). Cloning and functional analysis of spliced isoforms of human nuclear factor I-X: interference with transcriptional activation by NFI/CTF in a cell type specific manner. Nucleic Acids Res. *22*, 3825–3833.

Armand, O., Boutineau, A.M., Mauger, A., Pautou, M.P., and Kieny, M. (1983). Origin of satellite cells in avian skeletal muscles. Arch. Anat. Microsc. Morphol. Exp. *72*, 163–181.

Baghdadi, M.B., Castel, D., Machado, L., Fukada, S., Birk, D.E., Relaix, F., Tajbakhsh, S., and Mourikis, P. (2018). Reciprocal signalling by Notch–Collagen V–CALCR retains muscle stem cells in their niche. Nature *557*, 714–718.

Bandyopadhyay, S., and Gronostajski, R.M. (1994). Identification of a conserved oxidation-sensitive cysteine residue in the NFI family of DNA-binding proteins. J. Biol. Chem. *269*, 29949–29955.

Beauchamp, J.R., Heslop, L., Yu, D.S.W., Tajbakhsh, S., Kelly, R.G., Wernig, A., Buckingham, M.E., Partridge, T.A., and Zammit, P.S. (2000). Expression of CD34 and Myf5

defines the majority of quiescent adult skeletal muscle satellite cells. J. Cell Biol. *151*, 1221–1233.

Bentzinger, C.F., Wang, Y.X., and Rudnicki, M.A. (2012). Building Muscle : Molecular Regulation of Myogenesis. Cold Spring Harb Perspect Biol *4*, 1–16.

Biressi, S., Tagliafico, E., Lamorte, G., Monteverde, S., Tenedini, E., Roncaglia, E., Ferrari, S.S., Ferrari, S.S., Cusella-De Angelis, M.G., Tajbakhsh, S., et al. (2007a). Intrinsic phenotypic diversity of embryonic and fetal myoblasts is revealed by genome-wide gene expression analysis on purified cells. Dev. Biol. *304*, 633–651.

Biressi, S., Molinaro, M., and Cossu, G. (2007b). Cellular heterogeneity during vertebrate skeletal muscle development. Dev. Biol. *308*, 281–293.

Bulfield, G., Siller, W.G., Wight, P.A., and Moore, K.J. (1984). X chromosome-linked muscular dystrophy (mdx) in the mouse. Proc. Natl. Acad. Sci. *81*, 1189–1192.

Campbell, C.E., Piper, M., Plachez, C., Yeh, Y.-T.T., Baizer, J.S., Osinski, J.M., Litwack, E.D., Richards, L.J., and Gronostajski, R.M. (2008). The transcription factor Nfix is essential for normal brain development. BMC Dev. Biol. *8*, 52.

Chal, J., and Pourquié, O. (2017). Making muscle: Skeletal myogenesis in vivo and in vitro. Dev. *144*, 2104–2122.

Chang, N.C., Chevalier, F.P., and Rudnicki, M.A. (2016). Satellite Cells in Muscular Dystrophy - Lost in Polarity. Trends Mol. Med. *22*, 479–496.

Chaudhry, A.Z., Lyons, G.E., and Gronostajski, R.M. (1997). Expression patterns of the four nuclear factor I genes during mouse embryogenesis indicate a potential role in development. Dev. Dyn. *208*, 313–325.

Chaudhry, A.Z., Vitullo, A.D., and Gronostajski, R.M. (1998). Nuclear factor I (NFI) isoforms differentially activate simple versus complex NFI-responsive promoters. J. Biol. Chem. *273*, 18538–18546.

Chazaud, B., Brigitte, M., Yacoub-Youssef, H., Arnold, L., Gherardi, R., Sonnet, C., Lafuste, P., and Chretien, F. (2009). Dual and beneficial roles of macrophages during skeletal muscle regeneration. Exerc. Sport Sci. Rev. *37*, 18–22.

Chen, K.S., Lim, J.W.C., Richards, L.J., and Bunt, J. (2017). The convergent roles of the nuclear factor I transcription factors in development and cancer. Cancer Lett. *410*, 124–138.

Cinnamon, Y., Kahane, N., and Kalcheim, C. (1999). Characterization of the early development of specific hypaxial muscles from the ventrolateral myotome. Development *126*, 4305–4315.

Clark, B.S., Stein-O'Brien, G.L., Shiau, F., Cannon, G.H., Davis-Marcisak, E., Sherman, T., Santiago, C.P., Hoang, T. V., Rajaii, F., James-Esposito, R.E., et al. (2019). Single-Cell RNA-Seq Analysis of Retinal Development Identifies NFI Factors as Regulating Mitotic Exit and Late-Born Cell Specification. Neuron *102*, 1111-1126.e5.

Cohen, P. (2002). The origins of protein phosphorylation. Nat. Cell Biol. *4*.

Collins, C.A., Olsen, I., Zammit, P.S., Heslop, L., Petrie, A., Partridge, T.A., and Morgan, J.E. (2005). Stem cell function, self-renewal, and behavioral heterogeneity of cells from the adult muscle satellite cell niche. Cell *122*, 289–301.

Conboy, M.J., Karasov, A.O., and Rando, T.A. (2007). High incidence of non-random template strand segregation and asymmetric fate determination in dividing stem cells and their progeny. PLoS Biol. *5*, 1120–1126.

Cooke, D.W., and Lane, M.D. (1999). The Transcription Factor Nuclear Factor I Mediates Repression of the GLUT4 Promoter by Insulin. J. Biol. Chem. *274*, 12917–12924.

Cornelison, D.D.W., Filla, M.S., Stanley, H.M., Rapraeger, A.C., and Olwin, B.B. (2001). Syndecan-3 and syndecan-4 specifically mark skeletal muscle satellite cells and are implicated in satellite cell maintenance and muscle regeneration. Dev. Biol. *239*, 79–94.

Cossu, G., and Biressi, S. (2005). Satellite cells, myoblasts and other occasional myogenic progenitors: Possible origin, phenotypic features and role in muscle regeneration. Semin. Cell Dev. Biol. *16*, 623–631.

Cossu, G., and Molinaro, M. (1987). Cell heterogeneity in the myogenic lineage. Curr. Top. Dev. Biol. *23*, 185–208.

Cossu, G., Ranaldi, G., Senni, M.I., Molinaro, M., and Vivarelli, E. (1988). "Early" mammalian myoblasts are resistant to phorbol ester-induced block of differentiation. Development *102*, 65–69.

Cusella-De Angelis, M.G., Molinari, S., Le Donne, A., Coletta, M., Vivarelli, E., Bouche, M., Molinaro, M., Ferrari, S., and Cossu, G. (1994). Differential response of embryonic and fetal myoblasts to TGF beta: a possible regulatory mechanism of skeletal muscle histogenesis.

Development *120*, 925–933.

Dall'Agnese, A., Caputo, L., Nicoletti, C., di Iulio, J., Schmitt, A., Gatto, S., Diao, Y., Ye, Z., Forcato, M., Perera, R., et al. (2019). Transcription Factor-Directed Re-wiring of Chromatin Architecture for Somatic Cell Nuclear Reprogramming toward trans-Differentiation. Mol. Cell *76*, 453-472.e8.

Davis, R.L., Weintraub, H., and Lassar, A.B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell *51*, 987–1000.

Denetclaw, W.F., and Ordahl, C.P. (2000). The growth of the dermomyotome and formation of early myotome lineages in thoracolumbar somites of chicken embryos. Development *127*, 893–905.

Denetclaw, J., Berdougo, E., Venters, S.J., and Ordahl, C.P. (2001). Morphogenetic cell movements in the middle region of the dermomyotome dorsomedial lip associated with patterning and growth of the primary epaxial myotome. Development *128*, 1745–1755.

Denny, S.K., Yang, D., Chuang, C.H., Brady, J.J., Lim, J.S.S., Grüner, B.M., Chiou, S.H., Schep, A.N., Baral, J., Hamard, C., et al. (2016). Nfib Promotes Metastasis through a Widespread Increase in Chromatin Accessibility. Cell *166*, 328–342.

Dephoure, N., Gould, K.L., Gygi, S.P., and Kellogg, D.R. (2013). Mapping and analysis of phosphorylation sites: A quick guide for cell biologists. Mol. Biol. Cell *24*, 535–542.

Diab, M., Raff, M., and Gunther, D.F. (2003). Clinical Report Osseous Fragility in Marshall – Smith Syndrome. Am. J. Med. Genet. *119A*, 218–222.

Drazic, A., and Winter, J. (2014). The physiological role of reversible methionine oxidation. Biochim. Biophys. Acta - Proteins Proteomics *1844*, 1367–1382.

Driller, K., Pagenstecher, A., Uhl, M., Omran, H., Berlis, A., Gründer, A., and Sippel, A.E. (2007). Nuclear Factor I X Deficiency Causes Brain Malformation and Severe Skeletal Defects. Mol. Cell. Biol. *27*, 3855–3867.

Duclos, F., Straub, V., Moore, S.A., Venzke, D.P., Hrstka, R.F., Crosbie, R.H., Durbeej, M., Lebakken, C.S., Ettinger, A.J., van der Meulen, J., et al. (1998). Progressive muscular dystrophy in alpha-sarcoglycan-deficient mice. J. Cell Biol. *142*, 1461–1471.

Dumont, N.A., Wang, Y.X., and Rudnicki, M.A. (2015a). Intrinsic and extrinsic mechanisms regulating satellite cell function. Development *142*, 1572–1581.

Dumont, N.A., Bentzinger, C.F., Sincennes, M.-C.C., and Rudnicki, M.A. (2015b). Satellite Cells and Skeletal Muscle Regeneration. Compr. Physiol. *5*, 1027–1059.

Dumont, N.A., Wang, Y.X., Von Maltzahn, J., Pasut, A., Bentzinger, C.F., Brun, C.E., and Rudnicki, M.A. (2015c). Dystrophin expression in muscle stem cells regulates their polarity and asymmetric division. Nat. Med. *21*, 1455–1463.

Edmondson, D.G., and Olson, E.N. (1989). A gene with homology to the myc imilarity region of MyoDl is expressed uring myogenesis and is sufficient to activate the muscle differentiation program.

Evano, B., and Tajbakhsh, S. (2018). Skeletal muscle stem cells in comfort and stress. Npj Regen. Med. *3*, 1–13.

Evelyn Heng, Y.H., Zhou, B., Harris, L., Harvey, T., Smith, A., Horne, E., Martynoga, B., Andersen, J., Achimastou, A., Cato, K., et al. (2015). NFIX regulates proliferation and migration within the murine SVZ neurogenic niche. Cereb. Cortex *25*, 3758–3778.

Fane, M., Harris, L., Smith, A.G., and Piper, M. (2017). Nuclear factor one transcription factors as epigenetic regulators in cancer. Int. J. Cancer *140*, 2634–2641.

Fletcher, C.F., Jenkins, N.A., Copeland, N.G., Chaudhry, A.Z., and Gronostajski, R.M. (1999). Exon structure of the nuclear factor I DNA-binding domain from C. elegans to mammals. Mamm. Genome *10*, 390–396.

Fraser, J., Essebier, A., Gronostajski, R.M., Boden, M., Wainwright, B.J., Harvey, T.J., and Piper, M. (2017). Cell-type-specific expression of NFIX in the developing and adult cerebellum. Brain Struct. Funct. *222*, 2251–2270.

Fraser, J., Essebier, A., Brown, A.S., Davila, R.A., Sengar, A.S., Tu, Y.S., Ensbey, K.S., Day, B.W., Scott, M.P., Gronostajski, R.M., et al. (2019). Granule neuron precursor cell proliferation is regulated by NFIX and intersectin 1 during postnatal cerebellar development. Brain Struct. Funct. *224*, 811–827.

Fuchs, E., Tumbar, T., and Guasch, G. (2004). Socializing with the neighbors: Stem cells and their niche. Cell *116*, 769–778.

Fukada, S., Uezumi, A., Ikemoto, M., Masuda, S., Segawa, M., Tanimura, N., Yamamoto, H., Miyagoe-Suzuki, Y., and Takeda, S. (2007). Molecular Signature of Quiescent Satellite Cells in Adult Skeletal Muscle. Stem Cells *25*, 2448–2459.

Ge, J., Dong, H., Yang, Y., Liu, B., Zheng, M., Cheng, Q., Peng, L., and Li, J. (2018). NFIX downregulation independently predicts poor prognosis in lung adenocarcinoma, but not in squamous cell carcinoma. Futur. Oncol. *14*, 3135–3144.

Gil, G., Smith, J.R., Goldstein, J.L., Slaughter, C.A., Orth, K., Brown, M.S., and Osborne, T.F. (1988). Multiple genes encode nuclear factor 1-like proteins that bind to the promoter for 3-hydroxy-3-methylglutaryl-coenzyme A reductase. Proc Natl Acad Sci U S A *85*, 8963–8967.

Gnad, F., Ren, S., Cox, J., Olsen, J. V., Macek, B., Oroshi, M., and Mann, M. (2007). PHOSIDA (phosphorylation site database): Management, structural and evolutionary investigation, and prediction of phosphosites. Genome Biol. *8*.

Gronostajski, R.M. (2000). Roles of the NFI/CTF gene family in transcription and development. Gene *249*, 31–45.

Gros, J., Manceau, M., Thomé, V., and Marcelle, C. (2005). A common somitic origin for embryonic muscle progenitors and satellite cells. Nature *435*, 954–958.

Gründer, A., Qian, F., Ebel, T.T., Mincheva, A., Lichter, P., Kruse, U., and Sippel, A.E. (2003). Genomic organization, splice products and mouse chromosomal localization of genes for transcription factor Nuclear Factor One. Gene *304*, 171–181.

HALL, T., WALKER, M., GANUZA, M., HOLMFELDT, P., BORDAS, M., KANG, G., PALMER, L.E., BI, W., FINKELSTEIN, D., and MCKINNEY-FREEMAN, S. (2018). Nfix Promotes Survival of Immature Hematopoietic Cells via Regulation of c-Mpl TRENT. Stem Cells *36*, 943–950.

Harris, L., Dixon, C., Cato, K., Heng, Y.H.E., Kurniawan, N.D., Ullmann, J.F.P., Janke, A.L., Gronostajski, R.M., Richards, L.J., Burne, T.H.J., et al. (2013). Heterozygosity for Nuclear Factor One X Affects Hippocampal-Dependent Behaviour in Mice. PLoS One *8*.

Harris, L., Genovesi, L.A., Gronostajski, R.M., Wainwright, B.J., and Piper, M. (2015). Nuclear factor one transcription factors: Divergent functions in developmental versus adult stem cell populations. Dev. Dyn. *244*, 227–238.

Heng, Y.H.E., McLeay, R.C., Harvey, T.J., Smith, A.G., Barry, G., Cato, K., Plachez, C., Little, E., Mason, S., Dixon, C., et al. (2014). NFIX regulates neural progenitor cell differentiation during hippocampal morphogenesis. Cereb. Cortex *24*, 261–279.

Hiraike, Y., Waki, H., Yu, J., Nakamura, M., Miyake, K., Nagano, G., Nakaki, R., Suzuki, K., Kobayashi, H., Yamamoto, S., et al. (2017). NFIA co-localizes with PPARγ and transcriptionally controls the brown fat gene program. Nat. Cell Biol. *19*, 1081–1092.

Holmfeldt, P., Pardieck, J., Saulsberry, A.C., Nandakumar, S.K., Finkelstein, D., Gray, J.T., Persons, D.A., and McKinney-Freeman, S. (2013). Nfix is a novel regulator of murine hematopoietic stem and progenitor cell survival. Blood *122*, 2987–2996.

Hornbeck, P. V., Kornhauser, J.M., Tkachev, S., Zhang, B., Skrzypek, E., Murray, B., Latham, V., and Sullivan, M. (2012). PhosphoSitePlus: A comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. Nucleic Acids Res. *40*, 261–270.

Hunter, T. (2012). Why nature chose phosphate to modify proteins. Philos. Trans. R. Soc. B Biol. Sci. *367*, 2513–2516.

Irintchev, A., Zeschnigk, M., Starzinski-Powitz, A., and Wernig, A. (1994). Expression pattern of M-cadherin in normal, denervated, and regenerating mouse muscles. Dev. Dyn. *199*, 326–337.

Jackson, S.P., and Tjian, R. (1988). O-glycosylation of eukaryotic transcription factors: Implications for mechanisms of transcriptional regulation. Cell *55*, 125–133.

Jones, K.A., Kadonaga, J.T., Rosenfeld, P.J., Kelly, T.J., and Tjian, R. (1987). A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. Cell *48*, 79–89.

Jones, N.C., Tyner, K.J., Nibarger, L., Stanley, H.M., Cornelison, D.D.W., Fedorov, Y. V., and Olwin, B.B. (2005). The p38α/β MAPK functions as a molecular switch to activate the quiescent satellite cell. J. Cell Biol. *169*, 105–116.

Kahane, N., Cinnamon, Y., Bachelet, I., and Kalcheim, C. (2001). The third wave of myotome colonization by mitotically competent progenitors: regulating the balance between differentiation and proliferation during muscle development. Development *128*, 2187–2198.

Kallunki, T., Deng, T., Hibi, M., and Karin, M. (1996). c-Jun can recruit JNK to phosphorylate dimerization partners via specific docking interactions. Cell *87*, 929–939.

Kane, R., Murtagh, J., Finlay, D., Marti, A., Jaggi, R., Blatchford, D., Wilde, C., and Martin, F. (2002). Transcription factor NFIC undergoes N-glycosylation during early mammary gland involution. J. Biol. Chem. *277*, 25893–25903.

Kassar-duchossoy, L., Giacone, E., Gayraud-Morel, B., Jory, A., Gomès, D., and Tajbakhsh, S. (2005). Pax3/Pax7 mark a novel population of primitive myogenic cells during development. Genes Dev. *19*, 1426–1431.

Kim, G., Weiss, S.J., and Levine, R.L. (2014). Methionine oxidation and reduction in proteins. Biochim. Biophys. Acta - Gen. Subj. *1840*, 901–905.

Klaassens, M., Morrogh, D., Rosser, E.M., Jaffer, F., Vreeburg, M., Bok, L.A., Segboer, T., Van Belzen, M., Quinlivan, R.M., Kumar, A., et al. (2015). Malan syndrome: Sotos-like overgrowth with de novo NFIX sequence variants and deletions in six new patients and a review of the literature. Eur. J. Hum. Genet. *23*, 610–615.

Knight, J.D.R., and Kothary, R. (2011). The myogenic kinome: Protein kinases critical to mammalian skeletal myogenesis. Skelet. Muscle *1*, 1–18.

Kruse, U., and Sippel, A.E. (1994). Transcription factor nuclear factor I proteins form stable homo- and heterodimers. FEBS Lett. *348*, 46–50.

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

Kuang, S., Gillespie, M.A., and Rudnicki, M.A. (2008). Niche Regulation of Muscle Satellite Cell Self-Renewal and Differentiation. Cell Stem Cell *2*, 22–31.

L'Honoré, A., Commère, P.H., Negroni, E., Pallafacchina, G., Friguet, B., Drouin, J., Buckingham, M., and Montarras, D. (2018). The role of Pitx2 and Pitx3 in muscle 1 stem cells gives new insights into P38α MAP kinase and redox regulation of muscle regeneration. Elife *7*.

Lemos, D.R., Babaeijandaghi, F., Low, M., Chang, C.K., Lee, S.T., Fiore, D., Zhang, R.H., Natarajan, A., Nedospasov, S.A., and Rossi, F.M.V. (2015). Nilotinib reduces muscle fibrosis in chronic muscle injury by promoting TNF-mediated apoptosis of fibro/adipogenic progenitors. Nat. Med. *21*, 786–794.

Li, Y., Sun, C., Tan, Y., Li, L., Zhang, H., Liang, Y., Zeng, J., and Zou, H. (2020). Transcription levels and prognostic significance of the NFI family members in human cancers. PeerJ *2020*, 1–34.

Liu, S., Qu, D., Li, W., He, C., Li, S., Wu, G., Zhao, Q., Shen, L., Zhang, J., and Zheng, J. (2017). MiR-647 and miR-1914 promote cancer progression equivalently by downregulating

nuclear factor IX in colorectal cancer. Mol. Med. Rep. *16*, 8189–8199.

Liu, Z., Ge, R., Zhou, J., Yang, X., Cheng, K.K. yip, Tao, J., Wu, D., and Mao, J. (2020). Nuclear factor IX promotes glioblastoma development through transcriptional activation of Ezrin. Oncogenesis *9*.

Machado, L., Esteves de Lima, J., Fabre, O., Proux, C., Legendre, R., Szegedi, A., Varet, H., Ingerslev, L.R., Barrès, R., Relaix, F., et al. (2017). In Situ Fixation Redefines Quiescence and Early Activation of Skeletal Muscle Stem Cells. Cell Rep. *21*, 1982–1993.

Machado, L., Geara, P., Camps, J., Dos Santos, M., Teixeira-Clerc, F., Van Herck, J., Varet, H., Legendre, R., Pawlotsky, J.M., Sampaolesi, M., et al. (2021). Tissue damage induces a conserved stress response that initiates quiescent muscle stem cell activation. Cell Stem Cell *28*, 1125-1135.e7.

Malan, V., Rajan, D., Thomas, S., Shaw, A.C., Louis Dit Picard, H., Layet, V., Till, M., Van Haeringen, A., Mortier, G., Nampoothiri, S., et al. (2010). Distinct effects of allelic NFIX mutations on nonsense-mediated mRNA decay engender either a sotos-like or a Marshall-Smith Syndrome. Am. J. Hum. Genet. *87*, 189–198.

Mao, Y., Liu, J., Zhang, D., and Li, B. (2015). MiR-1290 promotes cancer progression by targeting nuclear factor I/X(NFIX) in esophageal squamous cell carcinoma (ESCC). Biomed. Pharmacother. *76*, 82–93.

Marshall, R.E., Graham, C.B., Scott, C.R., and Smith, D.W. (1971). Syndrome of accelerated skeletal maturation and relative failure to thrive: A newly recognized clinical growth disorder. J. Pediatr. *78*, 95–101.

Martinez, F., Marín-Reina, P., Sanchis-Calvo, A., Perez-Aytés, A., Oltra, S., Roselló, M., Mayo, S., Monfort, S., Pantoja, J., and Orellana, C. (2015). Novel mutations of NFIX gene causing Marshall-Smith syndrome or Sotos-like syndrome: One gene, two phenotypes. Pediatr. Res. *78*, 533–539.

Martynoga, B., Mateo, J.L., Zhou, B., Andersen, J., Achimastou, A., Urbán, N., van den Berg, D., Georgopoulou, D., Hadjur, S., Wittbrodt, J., et al. (2013). Epigenomic enhancer annotation reveals a key role for NFIX in neural stem cell quiescence. Genes Dev. *27*, 1769–1786.

Matuzelski, E., Bunt, J., Harkins, D., Lim, J.W.C., Gronostajski, R.M., Richards, L.J., Harris, L., and Piper, M. (2017). Transcriptional regulation of Nfix by NFIB drives astrocytic

maturation within the developing spinal cord. Dev. Biol. *432*, 286–297.

Mauro, A. (1961). Satellite cell of skeletal muscle fibers. J. Biophys. Biochem. Cytol. *9*, 493–495.

Mermod, N., O'Neill, E.A., Kelly, T.J., and Tjian, R. (1989). The proline-rich transcriptional activator of CTF/NF-I is distinct from the replication and DNA binding domain. Cell *58*, 741–753.

Messina, G., and Cossu, G. (2009). The origin of embryonic and fetal myoblasts: a role of Pax3 and Pax7. Genes Dev. *23*, 902–905.

Messina, G., Biressi, S., Monteverde, S., Magli, A., Cassano, M., Perani, L., Roncaglia, E., Tagliafico, E., Starnes, L., Campbell, C.E., et al. (2010). Nfix Regulates Fetal-Specific Transcription in Developing Skeletal Muscle. Cell *140*, 554–566.

Miller, J.B., Schaefer, L., and Dominov, J.A. (1999). Seeking muscle stem cells. Curr Top Dev Biol *43*, 191–219.

Moore, K.A., and Lemischka, I.R. (2006). Stem cells and their niches. Science (80-.). *311*, 1880–1885.

Mukhopadhyay, S.S., and Rosen, J.M. (2007). The C-terminal domain of the nuclear factor I-B2 isoform is glycosylated and transactivates the WAP gene in the JEG-3 cells. Biochem. Biophys. Res. Commun. *358*, 770–776.

Mulder, P.A., C, I.D., Landlust, A.M., Priolo, M., Menke, L.A., Acero, I.H., Alkuraya, F.S., Arias, P., Bernardini, L., Bijlsma, E.K., et al. (2020). Development , behaviour and sensory processing in Marshall – Smith syndrome and Malan syndrome : phenotype comparison in two related syndromes. J. Intellect. Disabil. Res. *64*, 956–969.

Muñoz-Cánoves, P., and Serrano, A.L. (2015). Macrophages decide between regeneration and fibrosis in muscle. Trends Endocrinol. Metab. *26*, 449–450.

Novak, A., Goyal, N., and Gronostajski, R.M. (1992). Four conserved cysteine residues are required for the DNA binding activity of nuclear factor I. J. Biol. Chem. *267*, 12986–12990.

Nowock, J., Borgmeyer, U., Puschel, A.W., Rupp, R.A.W., and Sippel, A.E. (1985). The TGGCA protein binds to the MMTV-LTR, the adenovirus origin of replication, and the BK virus enhancer. Nucleic Acids Res. *13*, 2045–2061.

O'Connor, C., Campos, J., Osinski, J.M., Gronostajski, R.M., Michie, A.M., and Keeshan, K.

(2015). Nfix expression critically modulates early B lymphopoiesis and myelopoiesis. PLoS One *10*, 1–15.

Pérez-Mejías, G., Velázquez-Cruz, A., Guerra-Castellano, A., Baños-Jaime, B., Díaz-Quintana, A., González-Arzola, K., Ángel De la Rosa, M., and Díaz-Moreno, I. (2020). Exploring protein phosphorylation by combining computational approaches and biochemical methods. Comput. Struct. Biotechnol. J. *18*, 1852–1863.

Pietrosemoli, N., Mella, S., Yennek, S., Baghdadi, M.B., Sakai, H., Sambasivan, R., Pala, F., Di Girolamo, D., and Tajbakhsh, S. (2017). Comparison of multiple transcriptomes exposes unified and divergent features of quiescent and activated skeletal muscle stem cells. Skelet. Muscle *7*, 1–15.

Piper, M., Harris, L., Barry, G., Heng, Y.H.E., Plachez, C., Gronostajski, R.M., and Richards, L.J. (2011). Nuclear factor one X regulates the development of multiple cellular populations in the postnatal cerebellum. J. Comp. Neurol. *519*, 3532–3548.

Piper, M., Gronostajski, R., and Messina, G. (2019). Nuclear Factor One X in Development and Disease. Trends Cell Biol. *29*, 20–30.

Pistocchi, A., Gaudenzi, G., Foglia, E., Monteverde, S., Moreno-Fortuny, A., Pianca, A., Cossu, G., Cotelli, F., and Messina, G. (2013). Conserved and divergent functions of Nfix in skeletal muscle development during vertebrate evolution. Development *140*, 2443–2443.

Pjanic, M., Schmid, C.D., Gaussin, A., Ambrosini, G., Adamcik, J., Pjanic, P., Plasari, G., Kerschgens, J., Dietler, G., Bucher, P., et al. (2013). Nuclear Factor I genomic binding associates with chromatin boundaries. BMC Genomics *14*, 1–18.

Priolo, M., Grosso, E., Mammì, C., Labate, C., Naretto, V.G., Vacalebre, C., Caridi, P., and Laganà, C. (2012). A peculiar mutation in the DNA-binding/dimerization domain of NFIX causes Sotos-like overgrowth syndrome: A new case. Gene *511*, 103–105.

Priolo, M., Schanze, D., Tatton-Brown, K., Mulder, P.A., Tenorio, J., Kooblall, K., Acero, I.H., Alkuraya, F.S., Arias, P., Bernardini, L., et al. (2018). Further delineation of Malan syndrome. Hum. Mutat. *39*, 1226–1237.

Qian, F., Kruse, U., Lichter, P., and Sippel, A.E. (1995). Chromosomal Localization of the four genes (NFIA, B, C, and X) for the human transcription factor nuclear factor I by FISH. Genomics *28*, 66–73.

Rabiee, A., Krüger, M., Ardenkjær-Larsen, J., Kahn, C.R., and Emanuelli, B. (2018). Distinct signalling properties of insulin receptor substrate (IRS)-1 and IRS-2 in mediating insulin/IGF-1 action. Cell. Signal. *47*, 1–15.

Rahman, N.I.A., Abdul Murad, N.A., Mollah, M.M., Jamal, R., and Harun, R. (2017). NFIX as a master regulator for lung cancer progression. Front. Pharmacol. *8*, 1–12.

Rai, A., Narayanan Lakshmi, D., and Padhke, S.R. (2018). Malan syndrome : Extension of genotype and phenotype spectrum. Am. J. Med. Genet. 2896–2900.

Ray, B.K., Dhar, S., Henry, C., Rich, A., and Ray, A. (2013). Epigenetic regulation by Z-DNA silencer function controls cancer-associated ADAM-12 expression in breast cancer: Cross-talk between MeCP2 and NF1 transcription factor family. Cancer Res. *73*, 736–744.

Relaix, F., Rocancourt, D., Mansouri, A., and Buckingham, M. (2005). A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. Nature *435*, 948–953.

Relaix, F., Bencze, M., Borok, M.J., Der Vartanian, A., Gattazzo, F., Mademtzoglou, D., Perez-Diaz, S., Prola, A., Reyes-Fernandez, P.C., Rotini, A., et al. (2021). Perspectives on skeletal muscle stem cells. Nat. Commun. *12*, 1–11.

Rhodes, S.J., and Konieczny, S.F. (1989). Identification of MRF4: a new member of the muscle regulatory factor gene family. Genes Dev. *3*, 2050–2061.

Roman, W., Pinheiro, H., Pimentel, M.R., Segalés, J., Oliveira, L.M., García-domínguez, E., Gómez-cabrera, M.C., Serrano, A.L., Gomes, E.R., and Muñoz-cánoves, P. (2021). Muscle repair after physiological damage relies on nuclear migration for cellular reconstruction. Science (80-.). *374*, 355–359.

Roodhooft, A.M., and Acker, K.J. Van (1988). Marshall-Smith Syndrome : New Aspects. Neuropediatrics *19*, 179–182.

Rossi, G., Antonini, S., Bonfanti, C., Monteverde, S., Vezzali, C., Tajbakhsh, S., Cossu, G., and Messina, G. (2016). Nfix Regulates Temporal Progression of Muscle Regeneration through Modulation of Myostatin Expression. Cell Rep. *14*, 2238–2249.

Rossi, G., Bonfanti, C., Antonini, S., Bastoni, M., Monteverde, S., Innocenzi, A., Saclier, M., Taglietti, V., and Messina, G. (2017). Silencing Nfix rescues muscular dystrophy by delaying muscle regeneration. Nat. Commun. *8*, 1–12.

Rudnicki, M.A., Schnegelsberg, P.N.J., Stead, R.H., Braun, T., Arnold, H.H., and Jaenisch,

R. (1993). MyoD or Myf-5 is required for the formation of skeletal muscle. Cell *75*, 1351–1359.

Saclier, M., Bonfanti, C., Antonini, S., Angelini, G., Mura, G., Zanaglio, F., Taglietti, V., Romanello, V., Sandri, M., Tonelli, C., et al. (2020a). Nutritional intervention with cyanidin hinders the progression of muscular dystrophy. Cell Death Dis. *11*.

Saclier, M., Lapi, M., Bonfanti, C., Rossi, G., Antonini, S., and Messina, G. (2020b). The Transcription Factor Nfix Requires RhoA-ROCK1 Dependent Phagocytosis to Mediate Macrophage Skewing during Skeletal Muscle Regeneration. Cells *9*, 1–18.

Saclier, M., Angelini, G., Bonfanti, C., Mura, G., Temponi, G., and Messina, G. (2022). Selective ablation of Nfix in macrophages attenuates muscular dystrophy by inhibiting fibro-adipogenic progenitor-dependent fibrosis. J. Pathol.

Sambasivan, R., and Tajbakhsh, S. (2007). Skeletal muscle stem cell birth and properties. Semin. Cell Dev. Biol. *18*, 870–882.

Santoro, C., Mermod, N., Andrews, P.C., and Tjian, R. (1988). A family of human CCAAT-box-binding proteins active in transcription and DNA replication: cloning and expression of multiple cDNAs. Nature *334*, 218–224.

Schanze, D., Neubauer, D., Cormier-Daire, V., Delrue, M.A., Dieux-Coeslier, A., Hasegawa, T., Holmberg, E.E., Koenig, R., Krueger, G., Schanze, I., et al. (2014). Deletions in the 3' part of the NFIX gene including a recurrent alu-mediated deletion of exon 6 and 7 account for previously unexplained cases of marshall-smith syndrome. Hum. Mutat. *35*, 1092–1100.

Schaum, N., Karkanias, J., Neff, N.F., May, A.P., Quake, S.R., Wyss-Coray, T., Darmanis, S., Batson, J., Botvinnik, O., Chen, M.B., et al. (2018). Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. Nature *562*, 367–372.

Schiaffino, S., and Reggiani, C. (2011). Fiber Types in Mammalian Skeletal Muscles. Physiol. Rev. *91*, 1447–1531.

Schiaffino, S., Partridge, T., Stockdale, F.E., Buckingham, M., Montarras, D., Zammit, P.S., Messina, G., Biressi, S., Cossu, G., Zhao, P., et al. (2008). Skeletal Muscle Repair and Regeneration (Springer).

Schiaffino, S., Rossi, A.C., Smerdu, V., Leinwand, L.A., and Reggiani, C. (2015). Developmental myosins: Expression patterns and functional significance. Skelet. Muscle *5*,

1–14.

Seale, P., Sabourin, L.A., Girgis-Gabardo, A., Mansouri, A., Gruss, P., and Rudnicki, M.A. (2000). Pax7 Is Required for the Specification of Myogenic Satellite Cells. Cell *102*, 777–786.

Shang, M., Cappellesso, F., Amorim, R., Serneels, J., Virga, F., Eelen, G., Carobbio, S., Rincon, M.Y., Maechler, P., De Bock, K., et al. (2020). Macrophage-derived glutamine boosts satellite cells and muscle regeneration. Nature *587*, 626–631.

Sharrocks, A.D., Yang, S.H., and Galanis, A. (2000). Docking domains and substrate-specificity determination for MAP kinases. Trends Biochem. Sci. *25*, 448–453.

Shaw, A.C., Van Balkom, I.D.C., Bauer, M., Cole, T.R.P., Delrue, M.A., Van Haeringen, A., Holmberg, E., Knight, S.J.L., Mortier, G., Nampoothiri, S., et al. (2010). Phenotype and natural history in Marshall-Smith syndrome. Am. J. Med. Genet. Part A *152*, 2714–2726.

Shinin, V., Gayraud-Morel, B., Gomès, D., and Tajbakhsh, S. (2006). Asymmetric division and cosegregation of template DNA strands in adult muscle satellite cells. Nat. Cell Biol. *8*, 677–682.

Sousa-Victor, P., García-Prat, L., and Muñoz-Cánoves, P. (2021). Control of satellite cell function in muscle regeneration and its disruption in ageing. Nat. Rev. Mol. Cell Biol. *0123456789*.

Sperli, D., Concolino, D., Barbato, C., Strisciuglio, P., and Andria, G. (1993). Long survival of a patient with Marshall-Smith syndrome without respiratory complications. J Med Genet 877–879.

Starosta, A., and Konieczny, P. (2021). Therapeutic aspects of cell signaling and communication in Duchenne muscular dystrophy. Cell. Mol. Life Sci.

Stockdale, F.E. (1992). Myogenic cell lineages. Dev. Biol. *154*, 284–298.

Taglietti, V., Maroli, G., Cermenati, S., Monteverde, S., Ferrante, A., Rossi, G., Cossu, G., Beltrame, M., and Messina, G. (2016). Nfix Induces a Switch in Sox6 Transcriptional Activity to Regulate MyHC-I Expression in Fetal Muscle. Cell Rep. *17*, 2354–2366.

Taglietti, V., Angelini, G., Mura, G., Bonfanti, C., Caruso, E., Monteverde, S., Carrou, G. Le, Tajbakhsh, S., Relaix, F., and Messina, G. (2018). Rhoa and erk signalling regulate the expression of the transcription factor nfix in myogenic cells. Development *145*, dev.163956.

Tajbakhsh, S. (2009). Skeletal muscle stem cells in developmental versus regenerative

myogenesis. J. Intern. Med. *266*, 372–389.

Tajbakhsh, S., and Buckingham, M. (2000). The Birth of Muscle Progenitor Cells in the Mouse: Spatiotemporal Considerations. In Current Topics in Developmental Biology, pp. 225–268.

Tapscott, S.J. (2005). The circuitry of a master switch: Myod and the regulation of skeletal muscle gene transcription. Development *132*, 2685–2695.

Theret, M., Saclier, M., Messina, G., and Rossi, F.M.V. (2021). Macrophages in Skeletal Muscle Dystrophies, An Entangled Partner. J. Neuromuscul. Dis. *9*, 1–23.

Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Žídek, A., Bridgland, A., Cowie, A., Meyer, C., Laydon, A., et al. (2021). Highly accurate protein structure prediction for the human proteome. Nature *596*, 590–596.

Varetti, G., Guida, C., Santaguida, S., Chiroli, E., and Musacchio, A. (2011). Homeostatic control of mitotic arrest. Mol. Cell *44*, 710–720.

van Velthoven, C.T.J., de Morree, A., Egner, I.M., Brett, J.O., and Rando, T.A. (2017). Transcriptional Profiling of Quiescent Muscle Stem Cells In Vivo. Cell Rep. *21*, 1994–2004.

Verhaart, I.E.C., and Aartsma-Rus, A. (2019). Therapeutic developments for Duchenne muscular dystrophy. Nat. Rev. Neurol. *15*, 373–386.

Visconti, R., Palazzo, L., and Grieco, D. (2010). Requirement for proteolysis in spindle assembly checkpoint silencing. Cell Cycle *9*, 564–569.

Webster, C., Silberstein, L., Hays, A.P., and Blau, H.M. (1988). Fast muscle fibers are preferentially affected in Duchenne muscular dystrophy. Cell *52*, 503–513.

Whitmarsh, A.J., and Davis, R.J. (2000). Regulation of transcription factor function by phosphorylation. Cell. Mol. Life Sci. *57*, 1172–1183.

Wright, W.E., Sassoon, D.A., and Lin, V.K. (1989). Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD. Cell *56*, 607–617.

Yang, W., and Hu, P. (2018). Skeletal muscle regeneration is modulated by inflammation. J. Orthop. Transl. *13*, 25–32.

Yang, B.S., Gilbert, J.D., and Freytag, S.O. (1993). Overexpression of Myc suppresses CCAAT transcription factor/nuclear factor 1-dependent promoters in vivo. Mol. Cell. Biol.

13, 3093–3102.

Yin, H., Price, F., and Rudnicki, M.A. (2013). Satellite Cells and the Muscle Stem Cell Niche. Physiol Rev *93*, 23–67.

Yoneda, Y., Saitsu, H., Touyama, M., Makita, Y., Miyamoto, A., Hamada, K., Kurotaki, N., Tomita, H., Nishiyama, K., Tsurusaki, Y., et al. (2012). Missense mutations in the DNA-binding/dimerization domain of NFIX cause Sotos-like features. J. Hum. Genet. *57*, 207–211.

Zeng, X., Sigoillot, F., Gaur, S., Choi, S., Pfaff, K.L., Oh, D.C., Hathaway, N., Dimova, N., Cuny, G.D., and King, R.W. (2010). Pharmacologic inhibition of the anaphase-promoting complex induces a spindle checkpoint-dependent mitotic arrest in the absence of spindle damage. Cancer Cell *18*, 382–395.

Zenker, M., Bunt, J., Schanze, I., Schanze, D., Piper, M., Priolo, M., Gerkes, E.H., Gronostajski, R.M., Richards, L.J., Vogt, J., et al. (2019). Variants in nuclear factor I genes influence growth and development. Am. J. Med. Genet. Part C Semin. Med. Genet. *181*, 611–626.

Zhang, N., Mendieta-Esteban, J., Magli, A., Lilja, K.C., Perlingeiro, R.C.R., Marti-Renom, M.A., Tsirigos, A., and Dynlacht, B.D. (2020). Muscle progenitor specification and myogenic differentiation are associated with changes in chromatin topology. Nat. Commun. *11*, 1–18.

Zhou, B., Osinski, J.M., Mateo, J.L., Martynoga, B., Sim, F.J., Campbell, C.E., Guillemot, F., Piper, M., and Gronostajski, R.M. (2015). Loss of NFIX Transcription Factor Biases Postnatal Neural Stem/Progenitor Cells Toward Oligodendrogenesis. Stem Cells Dev. *24*, 2114–2126.

Zhou, S., Han, L., and Wu, Z. (2022). A Long Journey before Cycling: Regulation of Quiescence Exit in Adult Muscle Satellite Cells. Int. J. Mol. Sci. *23*, 1748.

PART II

List of scientific publications

Saclier M, Angelini G, Bonfanti C, **Mura G**, Temponi G, and Messina G; Selective ablation of Nfix in macrophages attenuates muscular dystrophy by inhibiting fibro-adipogenic progenitor-dependent fibrosis; J Pathol, 2022 Mar 17; doi: 10.1002/path.5895

Angelini G*, **Mura G***, Messina G; Therapeutic approaches to preserve the musculature in Duchenne Muscular Dystrophy: the importance of the secondary therapies; Experimental Cell Research, 2022 Jan 15 ;410(2):112968. doi: 10.1016/j.yexcr.2021.112968

***These authors contributed equally**

Saclier M, Bonfanti C, Antonini S, Angelini G, **Mura G**, Zanaglio F, Taglietti V, Romanello V, Sandri M, Tonelli C, Petroni K, Cassano M, Messina G; Nutritional intervention with cyanidin hinders the progression of muscular dystrophy; Cell Death & Disease, 2020 Feb 18;11(2):127. doi: 10.1038/s41419-020-2332-4

Taglietti V, Angelini G, **Mura G**, Bonfanti C, Caruso E, Monteverde S, Le Carrou G, Tajbakhsh S, Relaix F, Messina G; RhoA and ERK signalling regulate the expression of the transcription factor Nfix in myogenic cells; Development, 2018 Oct 29;145(21): dev163956. doi: 10.1242/dev.163956