Università degli Studi di Ferrara

DOTTORATO DI RICERCA IN
FARMACOLOGIA E ONCOLOGIA MOLECOLARE

CICLO XXVI

COORDINATORE Prof. Antonio Cuneo

Whole transcriptome expression profiling in collagen VI myopathies KO mice reveals muscle-specific fingerprints and arises the role of circadian clock genes as myopathy biomarkers

Settore Scientifico Disciplinare MED/03

Dottoranda
Dott. ssa Martoni Elena

Tutore
Prof. ssa Ferlini Alessandra

Anni 2011/2013

…ieri è storia, domani è mistero, ma oggi è un dono…

*per questo si chiama "**presente**"…*

A Riccardo ed Enrico

INDEX

SUMMARY

The Col6a1-/- null mouse is a known and well characterized animal model for collagen VI myopathies in humans (Irwin WA, et al. 2003). Inherited mutations within the COL6 A1, 2 and 3 genes cause the autosomal dominant or recessive Bethlem myopathy (BM), the dominant or recessive Ullrich myopathy (UCMD), the autosomal dominant limb-girdle muscular dystrophy (Scacheri PC, et al. 2002) and the recessive rare muscle myosclerosis (Merlini L, et al. 2008).

These congenital muscular dystrophies are heterogeneous phenotypes underlined by a variety of COL6 genes' mutations. The phenotype ranges from mild and late onset myopathies (BM) to very severe forms (UCMD). Defective mitochondrial functions with ultrastructural alterations, increased spontaneous apoptosis and abnormal permeability in the transition pore (PTP) have been characterized both in animal models and in patients (Irwin WA, et al. 2003).

Recently, a further link between COL6 myopathies and mitochondrial dysfunction has been established by the discovery that autophagy is defective in collagen VI muscular dystrophies (Grumati P, et al. 2010). This supports a mechanism of disease pathogenesis that might be due, at least in part, to a defect of autophagy, leading to an accumulation of abnormal mitochondria and sarcoplasmic reticulum. Amelioration/restoration of a proper autophagic flux in mouse models ameliorates these disturbances, however, despite these recent achievements, a complete understanding of COL6-related diseases pathogenesis remains elusive.

Gene expression profiling is a powerful tool for exploring the alteration of normal tissue transcriptional behavior in the course of disease (Shapiro E, et al. 2013 Wang SM, et al. 2013).

In COL6 diseases microarrays are recently employed (Paco S, et al. 2013) also to analyze the transcriptome of UCMD muscle and compare it to healthy muscle and other muscular dystrophies in order to better understand the role of collagen VI in muscle and the mechanism of disease. Bioinformatics analysis reveal the molecular functions and gene networks associated with collagen VI defects; the most significantly regulated pathways were those involved in muscle regeneration, extracellular matrix remodelling and inflammation, opening new insights into disease pathogenesis, biomarkers discovery and novel therapeutic targets.

Therefore, we adopted whole genome expression profiling to study if and how the KO Col6a1-/- mice exhibit peculiar transcriptional patterns, which can address specific pathway alterations and therefore possible transcriptional biomarkers These biomarkers may be useful to better understand the pathogenesis of COL6 myopathies and as exploratory biomarkers of disease progression and drug response.

Our studies , among the involvement of some deregulated pathways, highlighted a misregulation of two genes belonging to the circadian rhythm. CLOCK and EGR1 were found to be consistently upregulated in all KO mice analyzed, though with muscle type-dependent quantitative differences, being tibialis the most affected muscle. The data was technically validated using TaqMan assays and customized Fluidic cards which assayed these and others circadian genes. The involvement of circadian genes in hereditary muscle dystrophy is fully novel, although it is already known as these genes are deeply involved in muscle remodeling, function, performance and aging. This new view opens interesting perspectives on new potential target genes involved in muscle damage pathogenesis.

AIMS

The aim of this work was the identification of pre-clinical, novel genomic or proteomic biomarkers, which may correlate with drug response, diagnosis and disease prognosis of NMDs by determining genomic variations as well as defining specific transcription or protein profiles or disease-related pathways both in human and in suitable animal models.

Biomarkers have been defined as cellular, biochemical, molecular alteration or biological characteristics that are measurable and evaluable in biological material as indicator of normal biological or pathogenic processes and may be used in differential and early diagnosis, in monitoring of disease progression, regression, or therapeutic outcome.

One of the many approaches for biomarkers discovery is gene expression profiling which is a powerful tool for exploring the alteration of normal tissue transcriptional behavior in the course of disease (Shapiro E, et al. 2013; Wang SM, et al. 2013) Great advances in this field have been reached in cancer, and some deregulated transcripts are now considered prognostic/therapeutic biomarkers of neoplastic diseases (Lam SW et al. 2014; Cernei N. et al. 2013).

Bioinformatics is a crucial step to understand and translate gene expression gross data and is carried out both by developing novel tools for data analysis, and by dedicating software to , for instance, whole exome sequencing , expression profiling or RNAsew analysis data storage,.

Using expression profile we identified 47 genes, resulted differentially expressed in COL6 null mice compared to wild type mice; using a customized fluidic card we validated all these genes differentially expressed. Among these genes, some regulators of the circadian rhythm process (Arntl, Atf5, Clock, Dbp, Egr1, Fkbp5, Per1, Per2 and Per3) were found. Some transcriptional

factors important for muscle-related processes (MyoD, MyoG, Mif6, Pax7) were also present confirming the involvement of these pathways in the COL6A1 KO mice pathology

BACKGROUND

Rapidly expanding knowledge of neuromuscular diseases (NMDs) has provided new targets for disease characterization, early diagnosis and drug development whilst presenting many challenges about how to translate this knowledge into clinical practice. This is all the more challenging as initial clinical trials typically run for such a short time that it is difficult for improvement to be effectively measured.

Inherited neuromuscular diseases (NMDs) are chronic degenerative diseases, each of which is individually rare (prevalence < 5\10.000) often associated with severe muscle weakness making difficult the characterization of clinical parameters. They are present in all populations and affect both children and adults. Most NMDs result in chronic long-term disability reduced quality of life and sometimes death may result from cardiac and respiratory muscle involvement and failure. The goal of existing management is to minimize the impact of complications such as joint or spinal deformity and improve cardiac and respiratory function as there currently are no curative treatments for any NMD.

The aim of this work was studying the expression profile in KO COL6 mice in order to increase our knowledge about Collagen VI myopathies pathogenesis in the animal model. This might be also helping in identifying transcriptomic biomarkers of the disease profile, progression and possibly novel therapies monitoring.

Biomarkers have been defined as cellular, biochemical, molecular alteration or biological characteristics that are measurable and evaluable in biological material as indicator of normal biological or pathogenic processes and may be used in differential and early diagnosis, in monitoring of disease progression, regression, or therapeutic outcome (http://www.bio-nmd.eu/forPatients/).

Benefits of biomarkers identification and monitoring can include:

• Blood and urine testing may be able to replace the use of painful and invasive muscle biopsies in the future.

• Diagnosis can happen earlier because testing for biomarkers is quicker and easier than genetic testing.

• Disease progression can be accurately measured allowing better clinical management of symptoms.

• Existing treatments (including drug dosage) can be adjusted to precisely meet the needs of individual patients to ensure they get the maximum benefit.

One of the possible methods for biomarkers discovery is gene expression profiling which is a powerful tool for exploring the alteration of normal tissue transcriptional behavior in the course of disease (Shapiro E, et al. 2013; Wang SM, et al. 2013) Great advances in this field have been reached in cancer, and some deregulated transcripts are now considered prognostic/therapeutic biomarkers of neoplastic diseases (Lam SW et al. 2014; Cernei N. et al. 2013).

Bioinformatics is a crucial step to understand and translate gene expression gross data and is carried both by developing novel tools for data analysis, and by dedicating software to whole exome sequencing analysis data storage, expression profiling and RNAseq analysis.

COLLAGEN VI

COL6 myopathies are well defined genetic diseses which genetic defects are well characterized, (Irwin W. et al. 2003) and representative of many other conditions because they include both early and late onset forms, fast and slow progression and severe and mild symptoms, so any biomarkers found may be more widely applicable.

Collagen VI (ColVI) is an extracellular matrix (ECM) protein forming a microfilamentous network in several organs including skeletal muscle, skin, cornea, lung, blood vessels, intervertebral disks and joints (Keene DR, et al. 1998). It consists of three chains, alpha1(VI), alpha2(VI) and alpha3(VI), encoded by distinct genes (COL6A1, COL6A2, COL6A3, respectively). The alpha1(VI) and alpha2(VI) chains are about 140 kDa, while the alpha3(VI) chain has several alternatively spliced variants with MW of 250−300 kDa (Bonaldo P, et al. 1989; Bonado P, et al. 1990; Chu ML, et al. 1990; Doliana R, et al. 1990;). Each chain contains a short triple helical domain of 335−336 amino acids and two large N− and C−terminal globular ends composed of repeated domains of 200 amino acids sharing similarity with the vWF−A module (Colombatti A, & Bonaldo P. 1991).

ColVI has a complex pathway of intracellular assembly (Colombatti A, et al. 1987; Colombatti A, & Bonaldo P. 1987; Colombatti A, et al. 1995). Equimolar association of the three alpha−chains into a triple helical monomer is followed by formation of dimers and tetramers stabilized by disulfide bonds. After secretion the tetramers form a network of microfilaments with a typical 100−nm periodicity that bridges the surface of cells with the connective tissue (Keene DR, et al. 1988; Bonaldo P, et al. 1990; Kuo HJ, et al. 1997). ColVI makes a wide range of interactions with ECM components (Bonaldo P, et al. 1990; Kuo HJ, et al. 1997; Bidanset DJ, et al. 1992; Sabatelli P, et al. 2001) and cell surface receptors, including integrins and NG2 proteoglycan (Pfaff M, et al. 1993; Burg MA, et al. 1996).

11

ColVI is synthesized and secreted by cells organizing an ECM such as skin fibroblasts and smooth muscle cells, and transcriptional regulation is a key step in its production. In muscle, ColVI is produced both by fibroblasts and myogenic cells, and is a major component of the endomysium (Kuo HJ, et al. 1997) (Fig.1).
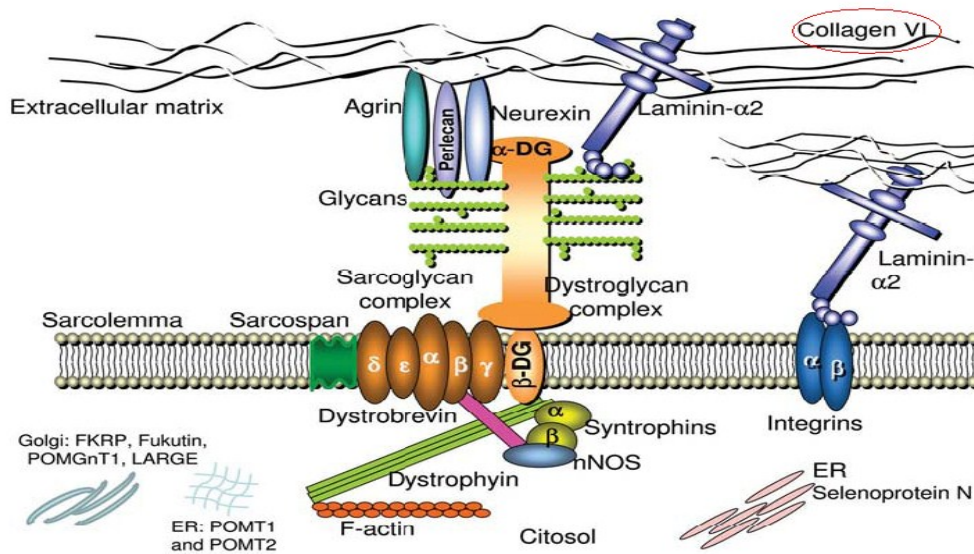

Figure 1. Adapted from Bonaldo P, et al. 1990, Collagen VI location within the extracellular matrix.

Deficiency of ColVI due to mutations in the COL6 genes, causes Collagen type VI-related disorders.

Collagen type VI-related disorders represent a continuum of overlapping phenotypes with Bethlem myopathy at the mild end, Ullrich congenital muscular dystrophy (CMD) at the severe end, and two rare, less well-defined disorders – autosomal dominant limb-girdle muscular dystrophy and autosomal recessive myosclerosis myopathy – in between. Although Bethlem myopathy and Ullrich CMD were defined long before their molecular basis was known, they remain useful for

clarification of prognosis and management (Bertini E. et al. 2011; Mendell JR et al. 2006; Bushby KM, et al. 2014; Lampe AK et al. 1993-2013).

Normal allelic variants

- *COL6A1* comprises 37 exons (35 of which are coding) and produces a single transcript encoding a protein of 1021 amino acids with two C-terminal and one N-terminal vWF type A-like domains.

- *COL6A2* comprises 30 exons (29 of which are coding) and has been shown to produce multiple alternatively spliced mRNAs that differ in the 5'-untranslated region as well as in the 3'-coding and noncoding sequences. It produces at least three α2(VI) protein variants (828-1019 amino acids) with distinct carboxyl termini, which similarly contain two C-terminal and one N-terminal vWF type A-like domain (Saitta B, et al. 1990).

- *COL6A3* comprises 44 exons (43 of which are coding) and encodes the α3(VI) chain, which can vary in size between 2970 and 3176 amino acids. The α3(VI) chain contains two C-terminal vWF type A-like domains, subdomains similar to type III fibronectin repeats and Kunitz protease inhibitors as well as six to ten N-terminal vWF type A-like domains, thus contributing most of the amino-terminal globular domain of the collagen VI heterotrimer. Various N-terminal exons of *COL6A3* are subject to alternative splicing and four variant transcripts encoding proteins with variably sized N-terminal globular domains have been characterized (Stokes DG, et al. 1991, Dziadek M, et al. 2002).

Pathologic allelic variants

- Single amino acid substitutions disrupting the Gly-Xaa-Yaa motif of the highly conserved triple helical domain of any of the three *COL6A* genes (Jöbsis GJ, et al. 1996, Pepe G, et al. 1999a, Scacheri PC, et al. 2002, Lampe AK, et al. 2005, Lucioli S, et al. 2005) constitute the most frequent pathogenic mechanism at 28% of the total pathogenic variants in COL6A1, COL6A2, and COL6A3.

- Mutations that introduce premature termination codons (splice sites and out of frame deletions/insertions) form the second most frequent group at 28%.

- Splice-site mutations that cause exon skipping occur at around the frequency (Lamande SR, et al. 1999, Pepe G, et al. 1999b,Pan TC, et al. 2003, Lampe et al 2005, Lucioli S, et al. 2005), with 27% of the total.

- Other splice-site mutations causing small in-frame deletions or insertions in regions that encode domains flanking the triple helical domain make up 8% of the total of pathogenic variants in *COL6A1*, *COL6A2* and *COL6A,* with large genomic deletions appearing to be rare and occurring at a frequency of around 2% (Vanegas OC, et al. 2002, Lampe AK, et al. 2005, Lucioli S, et al. 2005).

Given the high number of mutations that result in benign amino acid changes described for the three genes encoding the three collagen VI peptide chains, it is difficult to be sure about the pathogenicity of missense mutations other than glycine substitutions within the triple helical domain.

Abnormal gene product

- **Autosomal dominant collagen type VI-related disorders.** Heterozygous single amino acid substitutions disrupting the Gly-Xaa-Yaa motif of the highly conserved triple helical domain of any of the three *COL6A* genes (Jöbsis GJ, et al. 1996, Pepe G, et al. 1999a, Scacheri PC, et al. 2002, Lampe AK, et al. 2005, Lucioli S, et al. 2005), depending on their location, appear to either interfere with intracellular chain assembly or, following successful secretion, cause kinking of the tetramers, thus affecting extracellular microfibril formation (Lamande SR, et al. 2002). Functional haploinsufficiency via a dominant-negative effect has also been reported as the pathogenic mechanism for some missense and splice-site mutations (Lamande SR, et al. 1999). Heterozygous splice mutations (associated with autosomal dominant disease) leading to in-frame exonic deletions as well as in-frame genomic deletions preserve a unique cysteine important for dimer formation, allowing secretion of abnormal tetramers with a consequent dominant-negative effect on microfibrillar assembly (Pan TC, et al. 2003, Baker NL, et al. 2005).

- **Autosomal recessive collagen type VI-related disorders.**

  Most mutations associated with autosomal recessive disease reported to date are protein-truncating nonsense mutations. Some have been shown to result in absence of collagen VI because of nonsense-mediated mRNA decay (Zhang RZ, et al. 2002).

The phenotypes associated with collagen type VI-related disorders, once thought to be distinct entities, were clinically defined long before their molecular basis was discovered. The collagen type VI-related disorders are now recognized to comprise a continuum of overlapping phenotypes with Bethlem myopathy at the mild end, Ullrich congenital muscular dystrophy (CMD) at the severe end, and two less well-defined disorders – autosomal dominant limb-girdle muscular dystrophy and autosomal recessive myosclerosis myopathy – in between. Although these phenotypes are now

recognized to overlap and fall on a continuum, these clinical designations are useful for clarification of prognosis and management.

Bethlem Myopathy (MIM #158810)

The onset of Bethlem myopathy ranges from prenatal to mid-adulthood. Prenatal onset is characterized by decreased fetal movements; neonatal onset with hypotonia or torticollis; early-childhood onset with delayed motor milestones, muscle weakness and contractures; and adult onset (4th-6th decade) with proximal weakness and Achilles tendon or long finger flexor contractures (FIG.2). As some adults are unaware of weakness, age of onset cannot always be established.

The contractures may come and go during childhood, but nearly all affected individuals eventually exhibit flexion contractures of the fingers, wrists, elbows, and ankles. These contractures can become disabling when combined with muscle weakness.

Individuals can have moderate weakness and atrophy of the muscles of the trunk and limbs with proximal muscles being more involved than distal muscles and extensors more than flexors.

As a result of slow but ongoing progression of the condition, more than two-thirds of affected individuals over age 50 years need supportive means (i.e., canes, crutches, or wheelchair) for outdoor mobility (Jöbsis GJ, et al. 1999, Pepe G, et al. 1999b).

Respiratory muscle and especially diaphragmatic involvement necessitating artificial nocturnal respiratory support is part of the clinical spectrum but is rare and appears to be related to severe weakness occurring in later life (Haq RU, et al. 1999). Respiratory failure may supervene prior to loss of ambulation and may be associated with diaphragmatic weakness (Haq RU, et al. 1999).

Cardiac function is usually normal (Mohire MD, et al. 1988, de Visser M, et al. 1992).

Missense mutations in either the triple helical or the vWF−A domains of the three COL6 genes were described for various BM families (Jöbsis GJ, et al. 1996; Scacheri PC, et al. 2002; Sasaki T, et al. 2000). Immunohistochemistry shows apparently normal or mildly reduced levels of ColVI in the endomysium of most BM patients.

Ullrich Congenital Muscular Dystrophy (CMD) (MIM #254090)

In addition to characteristic muscle weakness of early onset, proximal joint contractures, and hyperelasticity of the wrists and ankles, other features observed are congenital hip dislocation, prominent calcanei, and a transient kyphotic deformity at birth (FIG.2).

With time, the distal hyperlaxity can evolve into marked finger flexion contractures and tight Achilles tendons (Furukawa T, & Toyokura Y. 1977; Muntoni F, et al. 2002).

Some affected children acquire the ability to walk independently; however, progression of the disease often results in later loss of ambulation.

Rigidity of the spine is often associated with scoliosis.

Early and severe respiratory involvement may require artificial ventilatory support in the first or second decade of life.

Failure to thrive is common (Demir E, et al. 2002).

Follicular hyperkeratosis over the extensor surfaces of upper and lower limbs and keloid and cigarette paper scar formation are common.

Cardiac involvement has not been documented to date.

Usually, UCMD shows an autosomal recessive inheritance with homozygous or compound heterozygous mutations in the COL6 genes. However, several cases of UCMD with dominant heterozygous mutations were recently reported (Pan TC, et al. 2003; Baker NL, et al. 2005; Angelin A, et al. 2007), and patients with a UCMD phenotype but without mutations in COL6 genes were also described. ColVI appears to be strongly reduced or absent in muscle biopsies from UCMD patients, suggesting that UCMD mutations severely affect the synthesis and secretion of ColVI (Camacho Vanegas O, et al. 2001; Zhang RZ, et al. 2002; Squarzoni S, et al. 2006).
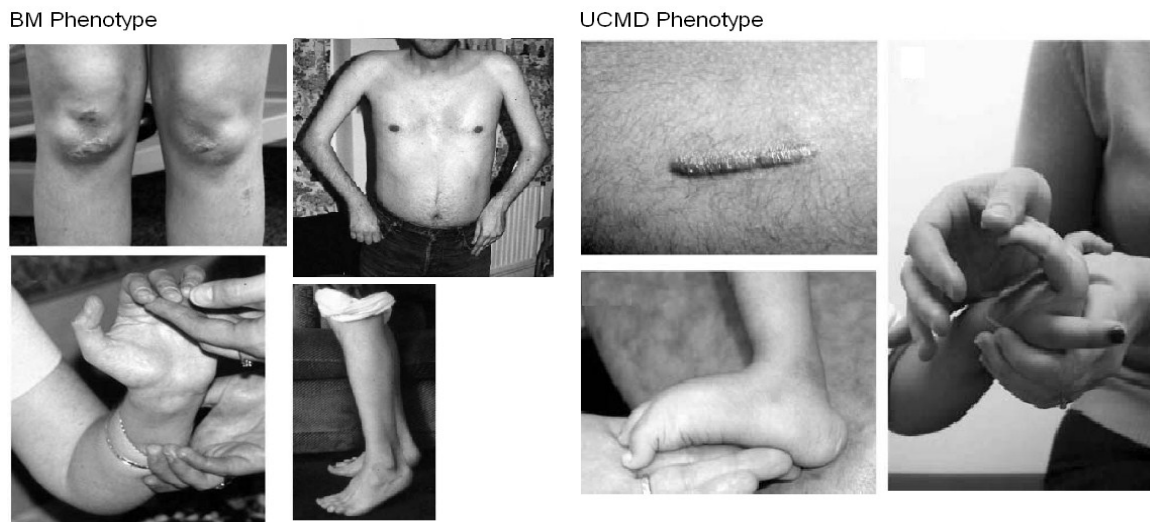

Figure 2 (Lampe AK & Bushby KM. 2005). Pictures from clinical patients with Bethlem Myopathy (BM) phenotype from left to right; Skin follicular hyperkeratosis and distal contractures in elbows, fingers and ankle. Ullrich's Congenital Muscular Dystrophy (UCMD) phenotype from left to right; Keloid scar formation, hyperelasticity in feet and fingers.

About 70 different mutations of the COL6 genes have so far been associated with UCMD and BM (Lampe AK, & Bushby KM. 2005) (see also http://www.dmd.nl/), but the reason why certain COL6 mutations cause BM while others cause the more severe UCMD remains obscure. Based on the

partial overlap between BM and UCMD, it had been proposed that COL6 disorders represent a clinical continuum (Lampe AK, & Bushby KM. 2005; Pepe G, et al. 2002).

Selective receptors directly involved in COL6 binding to muscle fibers have not yet been identified but the cell membrane chondroitin−sulfate proteoglycan NG2 binds COL6 (Tillet E, et al. 2002; Stallcup WB. 2002). NG2 has a highly regulated expression in skeletal muscle, where it is normally detectable in the sarcolemma of postnatal myofibers and gradually declines with age (Petrini S, et al. 2003). NG2 is selectively decreased in the sarcolemma of UCMD patients and Col6a1−/− mice (Petrini S, et al. 2005), suggesting that NG2 may represent an important molecule mediating ColVI−sarcolemma interactions.

Other Phenotypes

The two additional conditions included in the spectrum of collagen VI myopathies are:

- **Autosomal dominant limb-girdle muscular dystrophy (MIM #159000)** caused by mutations in *COL6A1/COL6A2* in three families and *COL6A3* in one family (Scacheri PC, et al. 2002). Although some affected individuals had mild weakness with only limited functional impairment, others had a more severe, dystrophic-like weakness with findings including Gower's maneuver, toe walking, and loss of ambulation. Joint contractures were either absent or much milder than those of typical Bethlem myopathy. Whereas findings of Bethlem myopathy are typically present in infancy, the age at onset in these three families ranged from infancy, to early childhood, to adulthood.

- **Autosomal recessive myosclerosis myopathy (MIM #255600)** caused by mutation of *COL6A2* in two individuals from one family (Merlini L, et al. 2008). Myosclerosis myopathy is characterized by difficulty in walking in early childhood, toe

walking, and progressive contractures of calf muscles. In the early 30s the muscles are slender with a firm "woody" consistency and associated with contractures that restrict range of motion of many joints (FIG.3A-B). By studying an Italian family affected by this rare disorder, we found a homozygous missense mutation in the COL6A2 gene. The mutation is a novel pathogenic variation of the COL6A2 gene and represents the first truncating mutation occurring in homozygosity in the C1 domain of the alpha2(VI) chain. By western blot and immunoprecipitation analysis of cell cultures from the patients, we demonstrated that this mutation causes the synthesis of a truncated alpha2(VI). The mutated chain is still able to form triple helical monomers and dimers with the other two chain, but formation of tetramers is strongly impaired, leading to decreased ColVI secretion. Electron microscopy confirmed the presence of a decreased amount of ColVI microfilaments in the ECM, containing abnormally−shaped tetramers (Merlini L, et al. 2008).
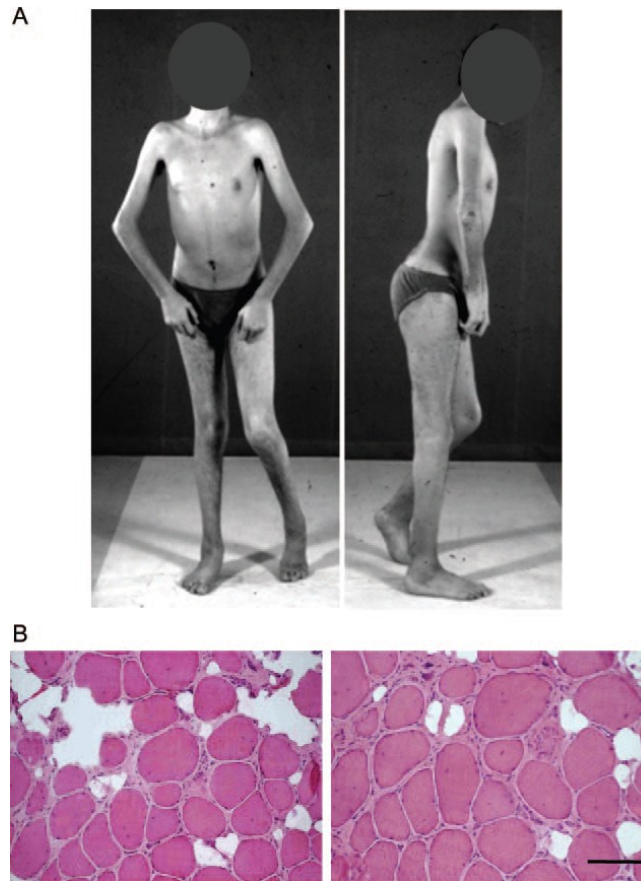
Fig.3.(A) Patient shows diffuse muscle wasting, flexion contractures of both elbows, forward abduction of shoulders, and flexion of fingers. There is fixed flexion of the trunk with lumbar hyperlordosis and marked flexion contractures of the knee and ankle joints in the left lower limb.
(B) Transverse section of muscle biopsy from the index case (P1, left) and the affected sibling (P2, right), showing endomysial fibrosis, increased fatty connective tissue, variability of fiber size, and several internal nuclei. Hematoxylin and eosin staining. Bar, 200 _m.

We have also studied biopsies and myoblasts from patients affected by COL6 muscular dystrophies. UCMD patients display an increased rate of apoptosis in skeletal muscle in vivo and in primary myoblast cultures (Angelin A, et al. 2007). The latter also displayed a measurable fraction of altered mitochondria, with morphological alterations ranging from shape changes to overt swelling; and the presence of a latent mitochondrial dysfunction that could be revealed by the addition of oligomycin, which caused depolarization only in mitochondria from patients (Angelin A, et al. 2007). The mitochondrial defect could be revealed in cultures from UCMD patients irrespective of whether the primary genetic defect was in the COL6A1 or COL6A3 gene, and both in homozygous and

heterozygous mutations (Angelin A, et al. 2007). Remarkably, PTP desensitization could be obtained also with MeAla3EtVal4−cyclosporin (Debio 025), a specific CyP inhibitor that desensitizes the PTP through mitochondrial CyP−D but has no inhibitory effects on calcineurin and therefore no immunosuppressive activity (Hansson MJ, et al. 2004). These findings suggest that PTP opening plays a key role in all COL6 myopathies and open new perspectives for the pharmacological treatment of patients.

Currently the only available animal model of COL6 muscular dystrophies is a mutant mouse with targeted inactivation of the COL6A1 gene in embryonic stem cellsby targeted gene disruption with a vector containing neomyvin resistance cassette in the second exon (Bonaldo P, et al. 1998). Homozygous null (Col6a1−/−) mice completely lack COL6 because in the absence of the alpha1(VI) chain COL6 does not assemble. Col6a1−/− mice are affected by an early onset myopathic disease with muscle weakness and histological changes that are similar to those detected in BM and UCMD patients (Bonaldo P, et al. 1998; Irwin WA, et al 2003). Ultrastructural analysis revealed the presence (in about 30% of the fibers) of marked dilations of the sarcoplasmic reticulum (SR), and of mitochondrial alterations that ranged from tubular cristae, to electron−dense inclusions, to overt swelling (Irwin WA, et al 2003). Remarkably, myofibers with mitochondrial−SR alterations also displayed nuclear features of apoptosis, suggesting a link between organellar changes and increased incidence of cell death (fig.4).

However, has been demostrated that an other pathway correlated with abnormal mitochondria and involved in myopathy is an autophagic pathway; indeed a major autophagic effector protein, beclin-1, is reduced in Col6a1−/− muscles. Thus, beclin-1 seems to be necessary and sufficient for basal autophagy and for its correct induction (Grumati P, et al. 2010).
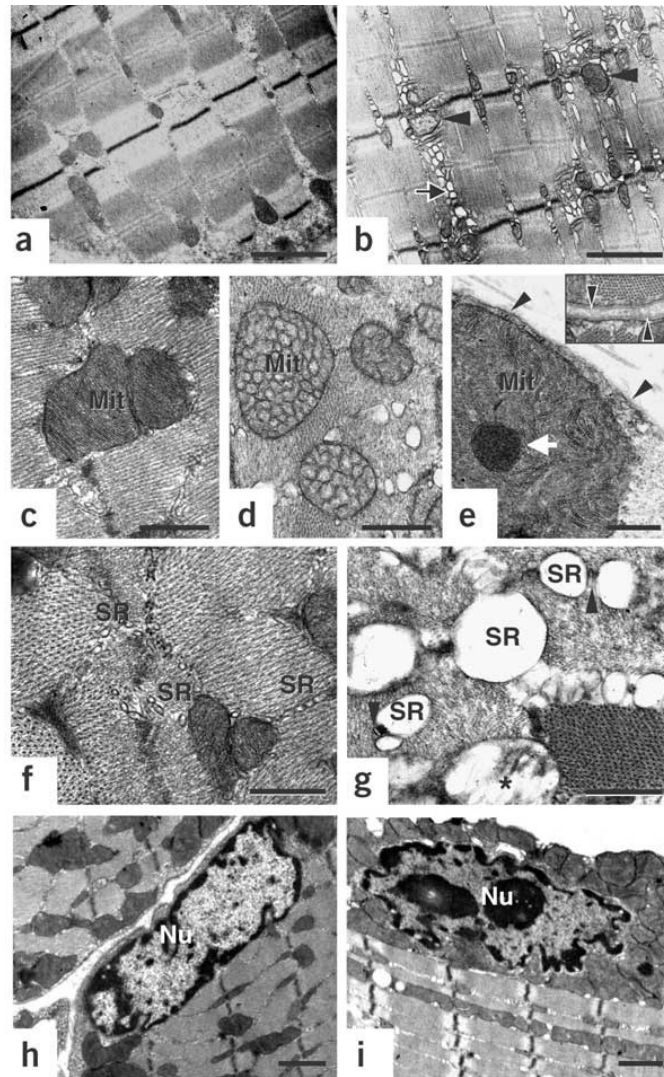
Fig. 4: Ultrastructural defects in collagen VI−deficient muscles (Irwin WA, et al. 2003). Electron micrographs of diaphragm and FDB muscles from wild-type (**a**,**c**,**f**,**h**) and *Col6a1⁻/⁻* (**b**,**d**,**e**,**g**,**i**) mice. (**a**,**b**) Low-power view of FDB longitudinal sections showing alterations in SR (arrows) and mitochondria (arrowheads) in *Col6a1⁻/⁻* fibers. Myofibrillar network is comparable in *Col6a1⁻/⁻* and wild-type mice. (**c−e**) High-power view of diaphragm transverse sections showing abnormal mitochondria (Mit) with altered cristae and dense bodies (white arrow) in *Col6a1⁻/⁻* fibers. Sarcolemma and basal lamina appear normal (arrowheads and inset in **e**). (**f**,**g**) High-power view of the triadic system showing dilation of the terminal cisternae of SR and normal T-tubules (arrowheads) in *Col6a1⁻/⁻* fibers. Swelling is visible in some *Col6a1⁻/⁻* mitochondria (asterisk). (**h**,**i**) Peripheral chromatin condensation and irregular shape characteristic of apoptosis are detected in *Col6a1⁻/⁻* myonuclei (Nu).

MAIN PUBLICATIONS DURING PHD

During my PhD I have contributed to the definition of col6-related disorders, which are a heterogeneous group of congenital myopathies, in order to increase both our knowledge about Collagen VI myopathies pathogenesis and the identification of specific transcription or protein profiles or disease-related pathways both in human and in suitable animal models.

First step of this huge work has been the identification of the suitability of peripheral blood macrophages as a reliable, minimally invasive tool for supplementing or replacing muscle/skin biopsies in the diagnosis and monitoring of collagen VI-related myopathies (Gualandi F, et al. 2011); In the dominant BM patient, no collagen VI alterations were detectable in macrophages or muscle biopsy but in the remaining patients, the protein defect caused by the selected mutations, as well as the transcriptional abnormalities, were readily detectable in macrophages, at levels comparable to those observed in muscle biopsy samples and cultured skin fibroblasts.

Interestingly two additional collagen VI chains were identified in humans, the α5 and α6 chains (Sabatelli P, et al. 2012): results showed a restricted and differential distribution of the novel α6 and α5 chains in skeletal muscle when compared to the widely distributed, homologous α3 chain, suggesting that these new chains may play specific roles in specialized ECM structures. While the α5 chain may have a specialized function in tissue areas subjected to tensile stress, the α6 chain appears implicated in ECM remodeling during muscle fibrosis.

A case report has also been described (Martoni E, et al. 2013) focusing the attention on mutations within the C-terminal region of the COL6A1 gene where they were only detected in Ullrich/Bethlem patients on extremely rare occasions. Two Brazilian brothers with a classic Ullrich

phenotype and compound heterozygous for two truncating mutations in COL6A1 gene, were expected to result in the loss of the α1(VI) chain C2 subdomain. Despite the reduction in COL6A1 RNA level due to nonsense RNA decay, three truncated alpha1 (VI) chains were produced as protein variants encoded by different out-of-frame transcripts. Collagen VI matrix was severely decreased and intracellular protein retention evident. The altered deposition of the fibronectin network highlighted abnormal interactions of the mutated collagen VI, lacking the α1(VI) C2 domain, within the extracellular matrix.

More recently we have identified a new form of extracellular matrix-related myopathy (Hicks D, et al. 2014); it was already known that should be additional causal genes to identify as in ~50% of BM cases no mutations in the COL6 genes are identified. In a cohort of -24 patients with a BM-like phenotype, we first sequenced 12 candidate genes based on their function, including genes for known binding partners of collagen VI, and those enzymes involved in its correct post-translational modification, assembly and secretion. Proceeding to whole-exome sequencing (WES), we identified mutations in the COL12A1 gene, a member of the FACIT collagens (fibril-associated collagens with interrupted triple helices) in five individuals from two families. Both families showed dominant inheritance with a clinical phenotype resembling classical BM. Abnormality at the protein level was confirmed in both families by the intracellular retention of collagen XII in patient dermal fibroblastsWe conclude that the spectrum of causative genes in extracellular matrix (ECM)-related myopathies be extended to include COL12A1.

METHODS

Samples selection

Four mice of six months of age for each condition (wild type and knock out) were selected and total RNA was extracted and pooled per muscle type: diaphragm, gastrocnemius and tibialis.

RNA extraction

Total RNAs was extracted from tibialis, diaphragm and gastrocnemius muscle tissue of WT and KO mice by using TRIzol Reagent (Invitrogen). Tissues were homogenized in 1 ml of TRIzol and incubated for 5' at room temperature to permit the complete dissociation of nucleoproteins complexes. 200 µl of chloroform were added and tubes were incubated for 3' at room temperature. After centrifugation (12000 g x 15' at 4°C) the mixture separated in a lower red, phenol-chloroform phase, an interphase and a colorless upper aqueous phase containing RNA. RNA from each muscle was purified with TRIzol Plus RNA Purification Kit (Invitrogen) and quantified by means of a Nanodrop (Thermo scientific) spectrophotometer and evaluated for the integrity with a 2100 Bioanalyser (Agilent) prior to be pooled in the same amount for muscles of the same kind.

Whole genome expression microarray analysis

To analyse the transcriptional profile in the mouse samples available, we selected the Whole Mouse Genome gene expression kit composed by four arrays of 44000 probes (Agilent Technologies, G4122F).

Sample labeling and hybridization were performed in triplicate according to the protocols provided by Agilent (One-Color Microarray-Based Gene Expression Analysis version 5.0.1). The array was analyzed using the Agilent scanner and Feature Extraction software (v9.1). Quality of the results was verified via Agilent Quality Controls (Spike-in), consisting of a mixture of 10 in vitro-synthesized, polyadenylated transcripts derived from the Adenovirus E1A gene, premixed at concentrations spanning six logs and differing by one-log or half-log increments.

Data analysis was performed using GeneSpring GX v.11. Data were normalized and log transformed. Unreliable "flagged" probes were excluded from further analysis. 2wayANOVA were performed for WT vs KO muscles with $p<0.05$ and Benjamini-Hochberg correction for multiple testing followed by fold change analysis (cut-off >3.0). Gene Ontology ($p<0.01$) was performed to group the differentially expressed genes into three main domains: cellular component, molecular function and biological process.

Microarray data analysis validation by fluidic cards

Following the pathway's bioinformatics analysis, 46 genes have been selected which resulted both differentially expressed between wild type and KO mice, in addition we analyzed Beclin-1 as candidate gene from literature data. The 47 selected genes predominantly include circadian pathway and muscle regeneration, but also known genes involved in autophagy and apoptosis. These genes were used to design customized TaqMan systems based on the Applied Biosystems database in order to set up a TaqMan low density array (TLDA or Fluidic card). The Fluidcard was used to validate the results obtained using the microarray platform. One gene (Mip) was selected as negative control for its absence of expression in the microarray data.

The fluidic card format 48, able to analyze 48 genes (47 test and the 18S housekeeping gene) simultaneously in eight samples, was run five times for each sample.

A total of 150 ng for each pool was retrotranscribed by means of a High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems), according to the manufacturer's instructions to be loaded in a single port of the TLDA and run on an ABI 7900HT System (Applied Biosystem) for 2 min at 50 °C, 10 min at 95 °C, followed by 40 cycles for 15 s at 97 °C and 1 min at 60 °C.

Data analysis was performed following the $\Delta\Delta$CT Method (Applied Biosystems User Bulletin #2) and using the 18S gene as calibrator and the WT samples as controls in respect to the KO samples. The mean 2-DDCT obtained has been Log transformed to be compared with the normalized fold change microarray data (raw data table).

RESULTS

To analyze the transcriptional profile in the KO Col6a1$^{-/-}$ mice as compared to WT mice we selected the Whole Mouse Genome gene expression kit (Agilent Technologies, G4122F) which is composed of four arrays and a total of 44000 probes. Gene expression profiles were analyzed in three muscle types; diaphragm, tibialis and gastrocnemius.

Data processing

Only 6-months-old mice were included in analysis and intensity data with local background adjustment by spatial detrending was extracted from Agilent files. This resulted in 45018 probes and 48 samples. Synonymous probes (which match to the same gene) were averaged. The final dataset consisted of 28052 probes. Additionally, data was log-transformed.

Quantile between-array normalization was utilized to make distributions of probes equal for each array. Box plot for probe distribution for each array before and after normalization is shown in fig 5.
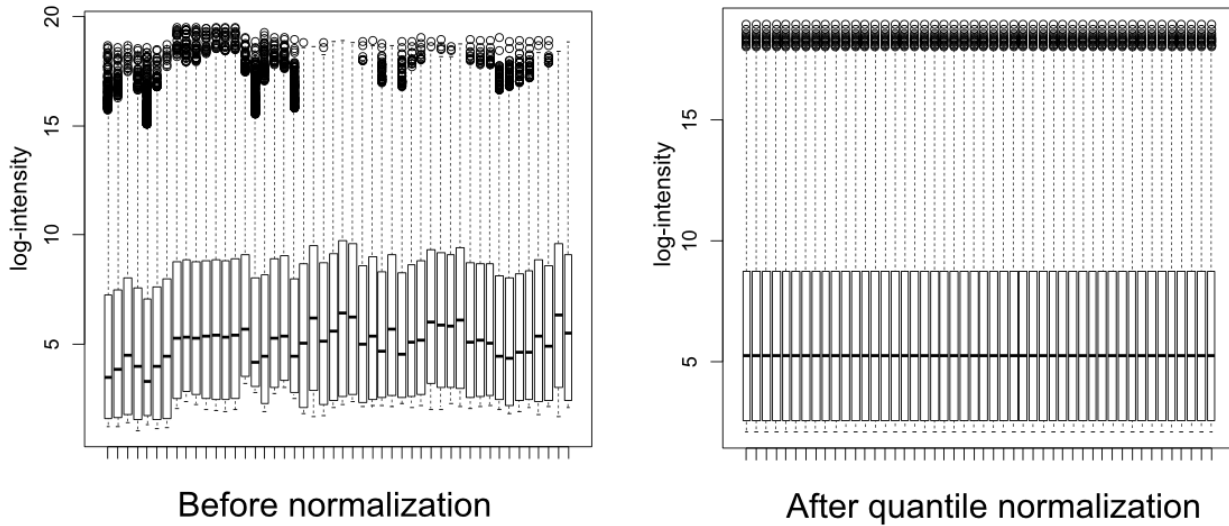
29

Fig 5: Data normalization

To perform multi-way comparisons in complex experiment design we utilized *limma* package (Wettenhall JM, & Smyth GK. 2004). This package fits separate linear model for each gene and uses moderated *t*-statistics to estimate differences between arrays. Empirical Bayes analysis is used to improve statistical power in small sample sizes. For each gene we fitted linear model with zero intercept parameterization because of the factorial design of the experiment. Corresponding p-values were corrected for multiple testing using Benjamini & Hochberg approach (Benjamini & Hochberg, 1995).

All data analysis and graphics were performed using free R statistical language (www.r-project.org) and set of bioinformatics-related packages, Bioconductor (www/bioconductor.org), Matlab and Ariadne Pathway Studio (ResNet8).

Samples clustering

A gene cluster analysis of the expression profiles in each muscle type samples in both the KO and WT mice was performed. We utilized complete linkage hierarchical clustering coupled with Spearman correlation as a distance measure. The resulting dendrogram is shown in figure 6. The

expression profiles of KO and WT gastrocnemius were most similar and thus clustered together. Those of diaphragm were also fairly well correlated and clustered together, although they were significantly different than the patterns seen in gastrocnemius. Interestingly, the expression profiles of KO vs. WT tibialis were most varied; with the profile of WT clustering with those gastrocnemius and that of the KO tibialis clustering with diaphragm.
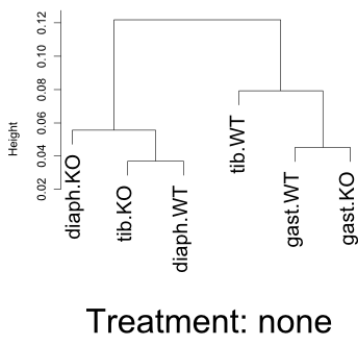
Fig 6: Data clusterization

Significant genes selection

To determine those genes that were significantly differentially expressed (p-value < 0.05, limma package) we analyzed each muscle type separately and in "pooled" tissue (averaged expression change over all muscle types).

The tissue most affected by col6 knockout is tibialis with 11170 differentially expressed genes. Diaphragm is second with 6790 differentially expressed genes and the least affected is gastrocnemius (3759 genes) .

KO vs WT effect

We applied "Find Go groups enriched with selected entities: cellular component" tool to genes with

p-value <0.05 and chose top 10 significant (p-value <0.05) results by overlap for each muscle and

for pooled dataset to ascribe differentially expressed genes to cellular components (tab. 1).

% of genes assosiated with Cellular Component	Muscle Type			
	Tibialis	Diaphragm	Gastrocnemi	All
membrane	26%	24%	23%	23%
cytoplasm	24%	22%	20%	22%
nucleus	23%	22%	18%	21%
plasma membrane	12%	12%	11%	-
cytosol	8%	8%	6%	5%
mitochondrion	8%	8%	6%	7%
extracellular region	8%	8%	8%	7%
endoplasmic reticulum	6%	5%	5%	4%
integral to plasma membrane	5%	4%	4%	-
Golgi apparatus	4%	4%	4%	3%
extracellular space	4%	4%	4%	-
cytoskeleton	4%	3%	3%	4%
proteinaceous extracellular	1%	1%	2%	3%

Tab 1. differentially expressed genes associated to cellular components

Muscles comparison

Genes associated with muscle development and function displayed altered expression in Col6-/-

mice. These changes affected multiple aspects of muscle biology including structure, metabolism,

transcription and regulation and varied between the tested muscle types. 60% (26 out of 44) of

muscle-specific genes change their expression in an opposite direction when comparing tibialis to

diaphragm, 23% between tibialis and gastrocnemius (7 out of 30) and 67% between gastrocnemius

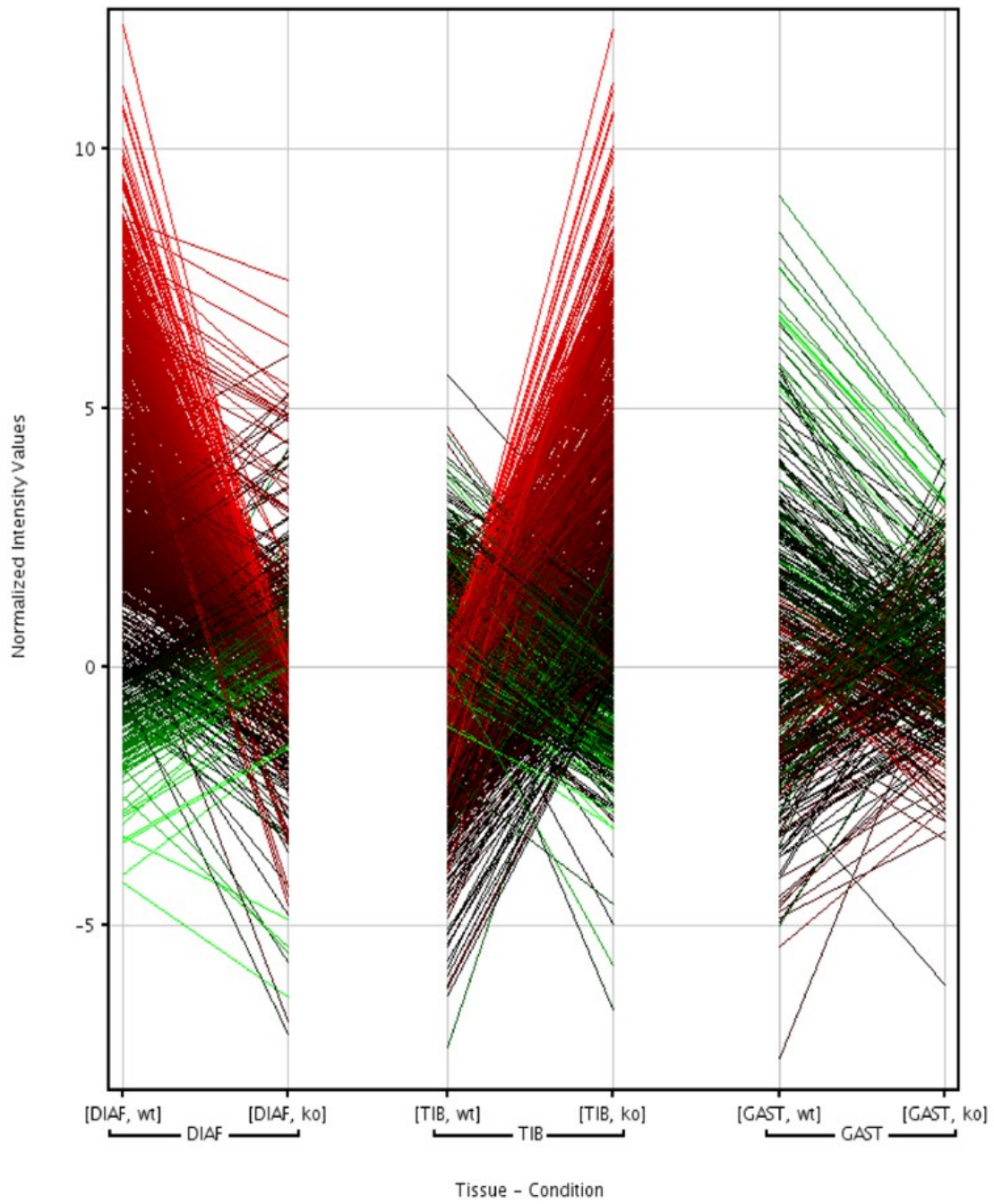and diaphragm (12 of 18) (each set filtered by pvalue KO vs WT mice 0.05) (Fig.7).

Fig. 7. each set filtered by p-value KO vs WT mice 0.05 in diaphragm, tibialis and gastrocnemius muscles

Below is the cluster analysis of non-filtered muscle-related genes in different muscle types (fig. 8).

Fig. 8: Muscle-related genes KO vs WT

Comparing to other muscle types Tibialis displayed differential expression of variant myosin, troponin and calsequestrin isoforms (predominantly cardiac and slow), and increased expression of PAX3. Genes, differentially expressed in diaphragm and tibialis, are almost all members of the dystrophin-glycan complex (dystrophin, dystrobrevin, dystroglycan 1, syntrophin, α-, β-, γ-sarcoglycans) and several transcriptional factors important for muscle-related processes, such as MYOD1, MYOG, MYF6, NFATC2, PAX7. In diaphragm embryonic isoforms of myosin are up-regulated, which can be regarded as a sign of regeneration as well as up-regulated PAX7, a marker of satellite cells activity.

Differentially expressed genes

The majority of gene expression changes in KO vs. WT mice differ for the different muscle types. Direct inspection of all genes displaying significant changes in expression level with the same sign in all tissues identified a relatively small subset of only 479 genes. This small number of consistent changes in expression corresponds to the high heterogeneity of effects that the col6 mutation has in the different muscle groups. Gene set enrichment analysis of these 479 genes was employed to identify those gene ontology (GO) groups with the most significant ($p < 0.05$) level of consistent change. The top 10 GO groups are listed in Table 2.

479 genes have the same sign effect in KO vs WT with p-value <0.05 in each muscle	Up	Down
inflammatory response	18	1
positive regulation of cell proliferation	10	9
regulation of transcription	33	12
hydrolase activity	21	22
transferase activity	19	21
modification-dependent protein catabolic process	8	7
GTPase activity	7	6
cholesterol metabolic process	5	1
positive regulation of mast cell cytokine production	3	0
positive regulation of integrin activatioin	1	1

Tab.2: The top 10 GO groups

Unfortunately, the majority of the GO groups identified are somewhat vague and not specific enough to yield insight regarding the effects of the col6 KO. Thus, we further probed the composite gene expression profile of the "pooled" data from all muscle groups, which indicated 334 relevant genes. Gene set enrichment analysis of these 334 genes identified a different subset of GO groups (Table 3).

334 genes are significantly different for pooled muscles in KO vs	Up	Down
rhythmic process	0	6
nuclear localization sequence binding	0	3
retinol metabolic process	0	3
regulation of transcription	17	9
ATP binding	10	14
Wnt-protein binding	2	1
response to cadmium ion	1	2
thyroid hormone generation	2	0
transport	7	11
apoptosis	4	5

Tab. 3: different subset of GO groups from "pooled" data

The intersection of two lists above has 94 genes with the following GO groups enriched (tab. 4):

Intersected 94 KO vs WT genes	Up	Down
transferase activity	7	8
negative regulation of signal transduction	2	1
negative regulation of ubiquitin-protein ligase activity	0	3
response to lead ion	1	1
transcription factor activity	5	4

Tab. 4: intersection of the two lists

It is clear that if data from different muscles are pooled, there are not so many inflammatory genes that have significant expression change between KO and WT mice. This leaded us to conclusion that inflammatory genes are most variable between different muscle types, so they are probably not suitable biomarkers.

Interestingly, one of the top 10 GO groups selected was "rhythmic process" which is a set of genes that are involved in the regulation of circadian rhythms.

Significant regulators

For further identifying key regulators that might affect or be effected in col6 disease, we employed sub-network enrichment analysis (SNEA) using Pathway Studio (Elsevier). SNEA utilizes literature (ResNet) and experimentally determined data to define the relationships between all molecular entities within the cell. In this way key pathways, either upstream or downstream, of identified gene expression changes can be elucidated and thus provide insight into potential key regulators of an altered cellular process. Again we performed this analysis on both the expression profiles of each muscle group individually as well as in the "pooled" dataset. Furthermore, we explored the intersection of these two datasets. For this initial SNEA experiment we used the Fisher Exact test as a measure of statistical significance. Identified key regulators are presented in figure 9.
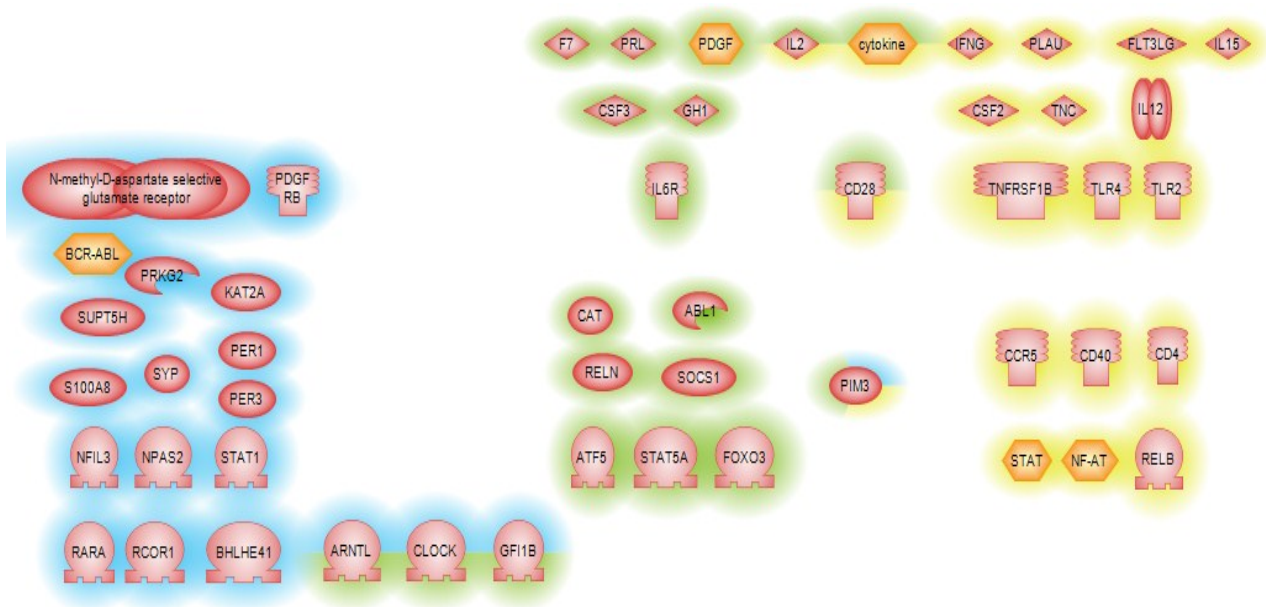

Fig. 9. Regulators of KO vs WT significant genes: each muscle group individually (blue highlight; "pooled" dataset yellow highlighted and the intersection red highlighted)

Genes with similar effect, but different base levels in different muscle groups are primarily regulated by inflammatory cytokines and receptors. In the "pooled" dataset a large number of key regulators for the most stable group of differentially expressed genes are regulators of cellular circadian rhythm (ARNTL, CLOCK, PER1, PER3) in good agreement the prior GO analysis. Unsurprisingly, identified regulators for the intersection of the two datasets shows features of the both previous groups, with the following intersecting genes: CLOCK, ARNTL, GFI1B, IL2, cytokine, CD28, PIM3. In total 47 key regulators were selected.

To verify the selected key regulators, we used additional statistical procedures within the Pathway Studio software (sub-network enrichment analysis with Mann-Whitney test, p-value 0.05, top 100 regulators) to verify potential regulators of the observed expression changes. All log ratio values KO vs. WT in the 3 muscle groups and the "pooled" dataset were used. Data was filtered to accept only gene expression changes with p-values <0.05 and used SNEA afterwards. This procedure gave us 8 different ranking lists (Table 5). Regulators, found in at least 4 of these ranking lists are presented below. Only two of the SNEA regulators were found more than 4 times – GH1 and, again, CLOCK. Noteworthy, none of the most common regulators is significant for gastrocnemius muscle.

Rank value	All	All (diff. exp pval<0.05)	Diaphragm	Diaphragm (diff. exp pval<0.05)	Gastrocnemius	Gastrocnemius (diff. exp pval<0.05)	Tibialis	Tibialis (diff. exp pval<0.05)	Total Sum of #
GH1	8		15	5			3	3	34
CLOCK		2	26	21			19	28	96
HNF4A			1	1			1	1	4
HNF1A			2	2			2	2	8
NR1H4			5	6			4	4	19
NR1I3			3	10			8	5	26
PPARA			7	8			5	11	31
FOXA2			4	4			16	8	32
CEBPA			9	7			9	10	35
RXRA			14	11			13	9	47
retinoid-X receptor			21	18			6	6	51
NR1I2			8	17			14	14	53
IL6			22	15			12	7	56
HNF1B			16	13			15	15	59
GCG			13	16			17	20	66
HNF4G			10	14			26	22	72
FOXM1			6	3			44	23	76
LXR			19	31			10	17	77
NR1H3	9		53				11	18	91
CEBPB			30	30			18	13	91
NR			23	20			33	16	92
ONECUT1			11	12			35	35	93
nuclear hormone receptor			25	24			20	27	96
NR0B2			20	26			30	32	108
FOXA3			31	29			24	25	109
fibrinogen			28	36			29	21	114
SREBF1			48	49			7	12	116
NR5A2			38	35			23	24	120
NR3C1			36	38			28	19	121
THRB			39	27			49	31	146
F9			32	37			38	41	148
INS			51	25			37	40	153
FABP1			40	63			96	43	242
DBP			58	50			61	80	249
FGF21			45	52			64	93	254

Table 5: ranking list

Among the possible pathways of expression changes it seems that circadian rhythm is one of the most affected circuit (significant GO group, CLOCK and ARNTL as regulators found by several statistical procedures). We used corresponding sub-networks found above to construct a possible circadian pathway (Fig. 10).
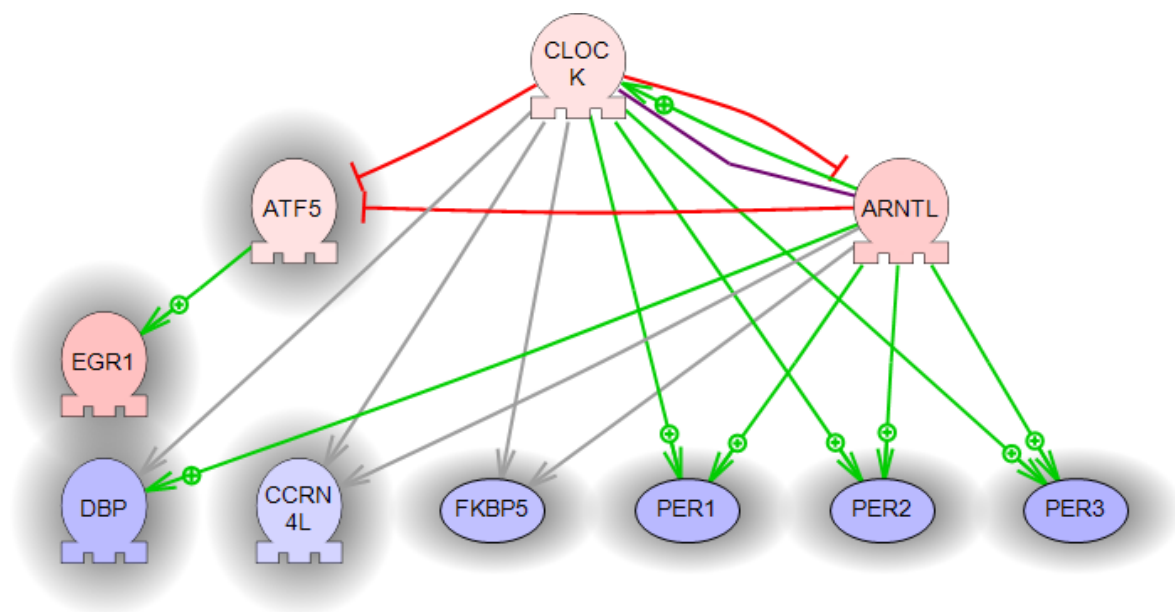

Fig 10: CLOCK ARNTL pathway. Highlighted - differentially expressed genes KO vs WT in pooled tissue sample

Microarray data analysis validation (fluidic card)

To validate the 47 selected key regulators, we again explored the differential expression of these proteins using Taqman low-density array (TLDA) cards (Fig. 11).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A	Abl1	Arntl	Atf5	Becn1	Mip	Ccr5	Ccrn4l	Cd40	Clock	Col6a1	18S	Csf2	Csf3	Dbp	Egr1	F7	Fkbp5	Flt3l	Foxo3	Gfi1b	Gh	Ifng	Il15	Il2
B	Il6ra	Kat2a	Myf6	Myod1	Myog	Pax7	Per1	Per2	Per3	Pim3	Plau	Prkg2	Prl	Relb	Reln	Socs1	Stat1	Stat5a	Tlr2	Tlr4	Tnc	Tnfrsf1b	Nfatc2	Pax3
C	Abl1	Arntl	Atf5	Becn1	Mip	Ccr5	Ccrn4l	Cd40	Clock	Col6a1	18S	Csf2	Csf3	Dbp	Egr1	F7	Fkbp5	Flt3l	Foxo3	Gfi1b	Gh	Ifng	Il15	Il2
D	Il6ra	Kat2a	Myf6	Myod1	Myog	Pax7	Per1	Per2	Per3	Pim3	Plau	Prkg2	Prl	Relb	Reln	Socs1	Stat1	Stat5a	Tlr2	Tlr4	Tnc	Tnfrsf1b	Nfatc2	Pax3
E	Abl1	Arntl	Atf5	Becn1	Mip	Ccr5	Ccrn4l	Cd40	Clock	Col6a1	18S	Csf2	Csf3	Dbp	Egr1	F7	Fkbp5	Flt3l	Foxo3	Gfi1b	Gh	Ifng	Il15	Il2
F	Il6ra	Kat2a	Myf6	Myod1	Myog	Pax7	Per1	Per2	Per3	Pim3	Plau	Prkg2	Prl	Relb	Reln	Socs1	Stat1	Stat5a	Tlr2	Tlr4	Tnc	Tnfrsf1b	Nfatc2	Pax3
G	Abl1	Arntl	Atf5	Becn1	Mip	Ccr5	Ccrn4l	Cd40	Clock	Col6a1	18S	Csf2	Csf3	Dbp	Egr1	F7	Fkbp5	Flt3l	Foxo3	Gfi1b	Gh	Ifng	Il15	Il2
H	Il6ra	Kat2a	Myf6	Myod1	Myog	Pax7	Per1	Per2	Per3	Pim3	Plau	Prkg2	Prl	Relb	Reln	Socs1	Stat1	Stat5a	Tlr2	Tlr4	Tnc	Tnfrsf1b	Nfatc2	Pax3
I	Abl1	Arntl	Atf5	Becn1	Mip	Ccr5	Ccrn4l	Cd40	Clock	Col6a1	18S	Csf2	Csf3	Dbp	Egr1	F7	Fkbp5	Flt3l	Foxo3	Gfi1b	Gh	Ifng	Il15	Il2
J	Il6ra	Kat2a	Myf6	Myod1	Myog	Pax7	Per1	Per2	Per3	Pim3	Plau	Prkg2	Prl	Relb	Reln	Socs1	Stat1	Stat5a	Tlr2	Tlr4	Tnc	Tnfrsf1b	Nfatc2	Pax3
K	Abl1	Arntl	Atf5	Becn1	Mip	Ccr5	Ccrn4l	Cd40	Clock	Col6a1	18S	Csf2	Csf3	Dbp	Egr1	F7	Fkbp5	Flt3l	Foxo3	Gfi1b	Gh	Ifng	Il15	Il2
L	Il6ra	Kat2a	Myf6	Myod1	Myog	Pax7	Per1	Per2	Per3	Pim3	Plau	Prkg2	Prl	Relb	Reln	Socs1	Stat1	Stat5a	Tlr2	Tlr4	Tnc	Tnfrsf1b	Nfatc2	Pax3
M	Abl1	Arntl	Atf5	Becn1	Mip	Ccr5	Ccrn4l	Cd40	Clock	Col6a1	18S	Csf2	Csf3	Dbp	Egr1	F7	Fkbp5	Flt3l	Foxo3	Gfi1b	Gh	Ifng	Il15	Il2
N	Il6ra	Kat2a	Myf6	Myod1	Myog	Pax7	Per1	Per2	Per3	Pim3	Plau	Prkg2	Prl	Relb	Reln	Socs1	Stat1	Stat5a	Tlr2	Tlr4	Tnc	Tnfrsf1b	Nfatc2	Pax3
O	Abl1	Arntl	Atf5	Becn1	Mip	Ccr5	Ccrn4l	Cd40	Clock	Col6a1	18S	Csf2	Csf3	Dbp	Egr1	F7	Fkbp5	Flt3l	Foxo3	Gfi1b	Gh	Ifng	Il15	Il2
P	Il6ra	Kat2a	Myf6	Myod1	Myog	Pax7	Per1	Per2	Per3	Pim3	Plau	Prkg2	Prl	Relb	Reln	Socs1	Stat1	Stat5a	Tlr2	Tlr4	Tnc	Tnfrsf1b	Nfatc2	Pax3

Fig 11: Fluidic card design

Comparison of the data obtained from both the microarray experiments and those with the TLDA cards resulted in the same sign and trend of fold change in 37 out of 44 genes (84%) in the diaphragm and tibialis tissues, whereas in the gastrocnemious we found 32 out of 44 (75%) of genes with these characteristics (fig.12)
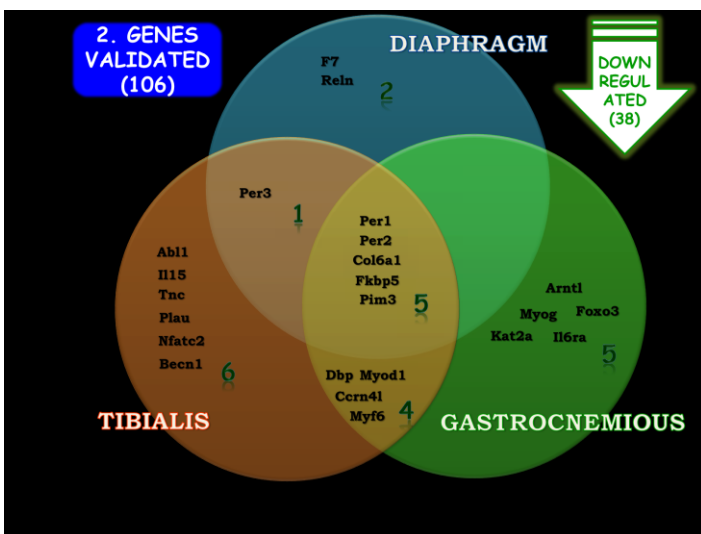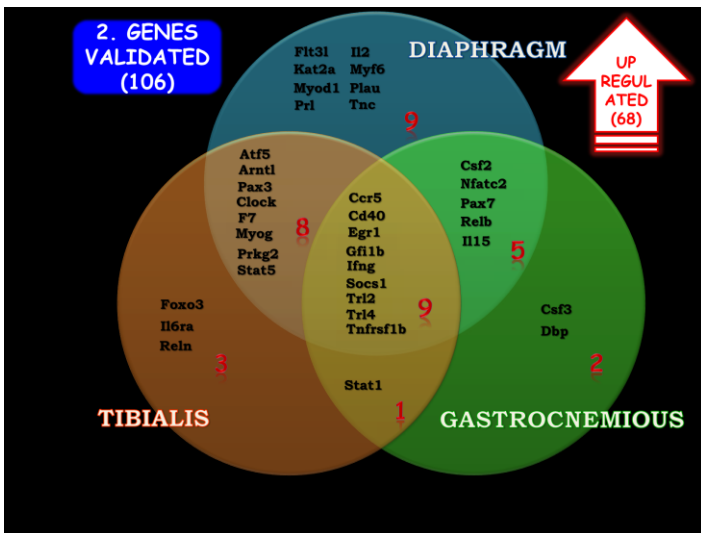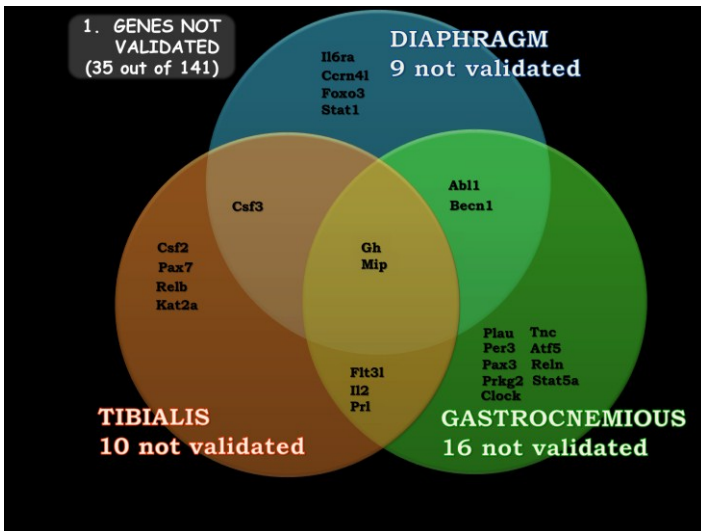
Fig 12. representation of non validated/validated genes in the three tissues.

However, a number of genes (Ccr5, Gh1, Ifng, Il6ra, Il2, Prl and Socs1) for which a significant fold-change was detected by the microarray, were poorly or not at all detected by the TLDA card. The Csf3 gene was only detected by the TLDA in tibialis muscle.

Up-regulated and down-regulated genes as detected by the TLDA cards in KO vs. WT mice are presented in figure 12. Negative control genes, Mip and Col6a1, were undetectable or down-regulated, respectively, for both methods in all three muscle types representing a true negative control. The TLDA and microarray data were in good agreement for regulators of the rhythmic process (Arntl, Atf5, Clock, Dbp, Egr1, Fkbp5, Per1, Per2 and Per3) and transcriptional factors important for muscle-related processes (MyoD, MyoG, Mif6, Pax7) in both diaphragm and tibialis and to a lesser extent in gastrocnemious confirming the involvement of these pathways in the Col6a1 KO mice (fig. 12).

We confirmed the down regulation of Beclin-1, as reported by Grumati P. et al., in KO vs WT mice in all 3 muscle tissues analyzed, even though this deregulation does not appear significantly relevant.

DISCUSSION

Collagen VI myopathies are a heterogeneous group of inherited diseases caused by mutations in *COL6A* genes. It has been previously identified a latent mitochondrial dysfunction in myoblasts from patients which displayed altered mitochondria and increased occurrence of spontaneous apoptosis.

The Col6a1-/- null mouse is a known and well characterized animal model for collagen VI myopathies in humans (Irwin WA, et al. 2003). Inherited mutations within the COL6 A1, 2 and 3 genes cause the autosomal dominant or recessive Bethlem myopathy (BM), the dominant or recessive Ullrich myopathy (UCMD), the autosomal dominant limb-girdle muscular dystrophy (Scacheri PC, et al. 2002) and the recessive rare muscle myosclerosis (Merlini L, et al. 2008).

These congenital muscular dystrophies are heterogeneous phenotypes underlined by a variety of COL6 genes' mutations. The phenotype ranges from mild and late onset myopathies (BM) to very severe forms (UCMD). Defective mitochondrial functions with ultrastructural alterations, increased spontaneous apoptosis and abnormal permeability in the transition pore (PTP) have been characterized both in animal models and in patients (Irwin WA, et al. 2003). The PTP abnormalities are reverted by inhibition of the pore via Cyclosporine A (CsA) treatment. Proof of principle of these effects of CsA have been demonstrated *in vitro* as well as in animal models and has subsequently been translated into a pivotal clinical trial (Angelin A, et al. 2007).

Recently, a further link between COL6 myopathies and mitochondrial dysfunction has been established by the discovery that autophagy is defective in collagen VI muscular dystrophies (Grumati P, et al. 2010). This supports a mechanism of disease pathogenesis that might be due, at least in part, to a defect of autophagy, leading to an accumulation of abnormal mitochondria and

sarcoplasmic reticulum. Amelioration/restoration of a proper autophagic flux in mouse models ameliorates these disturbances, however, despite these recent achievements, a complete understanding of COL6 disease pathogenesis remains elusive.

However in our data, beclin-1, a major autophagic effector protein, does not appear to be significantly deregulated suggesting the involvement of other candidated pathways in the pathophysiological mechanism of Col6 related myopathies.

Skeletal muscle biopsy is a choice method to monitor disease progression and the effects of treatments; however, it is an invasive and painful method and requires surgical expertise. Cultured skin fibroblasts are currently used to evaluate the expression pattern of collagen VI in UCMD/BM patients. However, cultured fibroblasts are not useful for functional studies since they do not display the characteristic mitochondria alterations.

Gene expression profiling is a powerful tool for exploring the alteration of normal tissue transcriptional behavior in the course of disease (Shapiro E, et al. 2013 Wang SM, et al. 2013).

Great advances in this field have been reached in cancer, and some deregulated transcripts are now considered prognostic/therapeutic biomarkers of neoplastic diseases (Lam SW et al. 2014. Cernei N. et al. 2013).

Animal models do represent an established model to study expression profile in human diseases, including muscular dystrophies and muscle pathologies ('t Hoen PA, et al. 2006; Marotta M, et al. 2009).

In COL6 diseases microarrays are recently employed (Paco S, et al. 2013) also to analyze the transcriptome of UCMD muscle and compare it to healthy muscle and other muscular dystrophies in order to better understand the role of collagen VI in muscle and the mechanism of disease.

Bioinformatics analysis reveal the molecular functions and gene networks associated with collagen VI defects; the most significantly regulated pathways were those involved in muscle regeneration, extracellular matrix remodelling and inflammation, opening new insights into disease pathogenesis, biomarkers discovery and novel therapeutic targets.

We adopted whole genome expression profiling to study if and how the KO Col6a1-/- mice exhibit peculiar transcriptional patterns which can address specific pathway alterations and therefore possible transcriptional biomarkers. These biomarkers may be useful to better understand the pathogenesis of COL6 myopathies and as exploratory biomarkers of disease progression and drug response.

To analyze the transcriptional profile in the KO mice compared to WT we used the whole mouse genome expression kit (Agilent technologies G4122F) in three different muscular district: tibialis, diaphragm and gastrocnemius of mice of six months of age. We differentiated our analysis studying those genes that were significantly differentially expressed in each muscle type both separately and in "pooled" tissues (all tissues taken together as one) and among the three muscles analyzed, tibialis and diaphragm seemed to be the most affected whereas gastrocnemius showed only few differences.

We identified changes in genes associated with muscle function including structure, metabolism, transcription and regulation and varying between the tested muscles.

The number of genes significantly changed in all tissues analyzed both separately and "pooled" together with the same sign of effect is relatively small, 479 genes and 334 genes respectively; gene set enrichment of these genes was employed to identify a list of top GO groups, one for tissues analyzed separately and one for "pooled" muscles. The intersection of these two GO groups allowed us to conclude that inflammatory genes (found at the top of GO group of genes analyzed separately as inflammatory response) are probably not suitable transcriptomic biomarkers; they

disappear in prioritization after the GO intersection. On the contrary one of the top GO groups, maintained after intersection, was "rhythmic process" which is a set of genes involved in circadian rhythm regulation.

In order to further refine the transcriptomic signature, we employed a sub-network enrichment analysis again on both separate tissues and "pooled" ones. Furthermore we explored again the intersection of the two dataset. In total 47 key regulators were selected.

To verify the selected key regulators we used an additional statistical procedures, data was filtered several times giving us 8 ranking lists; regulators found at least in four of these ranking lists have been taken and just two regulators were found more than four times: CLOCK and GH1.

Among the possible mechanism of expression it seems that circadian rhythm is one of the most affected (significant GO groups and CLOCK and ARNTL s regulators found by several statistical procedures).


To validate the 47 selected key regulators we explored again the differential expression of these transcripts using low-density array cards (fluidic cards); fluidic cards and microarray data were in good agreement for:

- Regulators of the circadian rhythm process (Arntl, Atf5, Clock, Dbp, Egr1, Fkbp5, Per1, Per2 and Per3).

- Transcription factors important for muscle related process (MyoD, MyoG, Mif6, Pax7).

All these genes were deregulated both in diaphragm and in tibialis and lower in gastrocnemius.

Interestingly, in null COLVI mice the expression profile was slightly moved toward the up-regulation respect to the down-regulation suggesting a possible gain of function of the involved genes rather than a loss of function.

In summary our studies revealed a clear involvement of two genes known to be controlling the circadian rhythm, CLOCK and EGR1 found to be consistently upregulated in all KO mice analyzed, although with muscle-type dependent quantitative differences. Tibialis is the muscle most affected. The graphs below show the gene expression microarray data analysis (blue) and the fluidic cards validation (red) for each muscle (Fig.13).
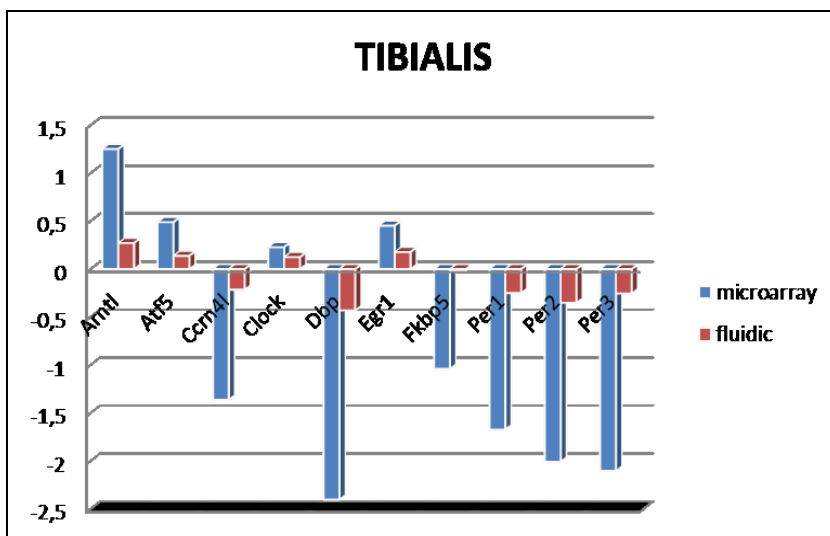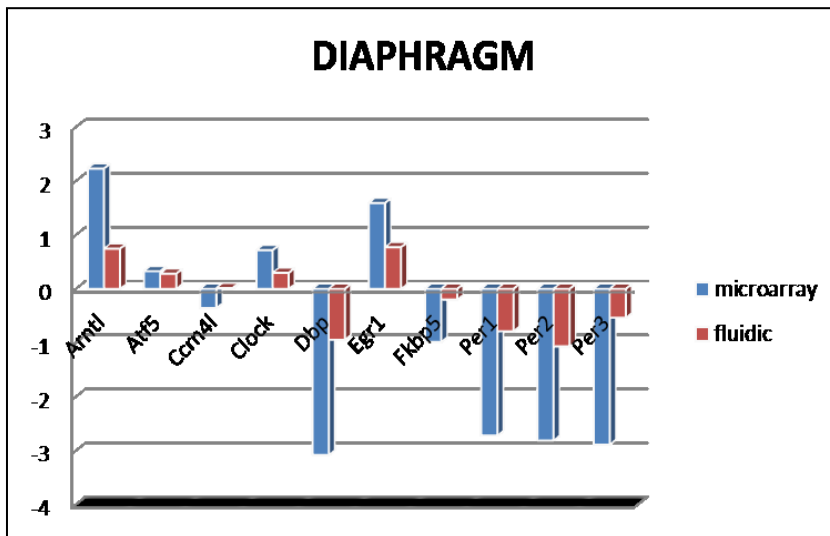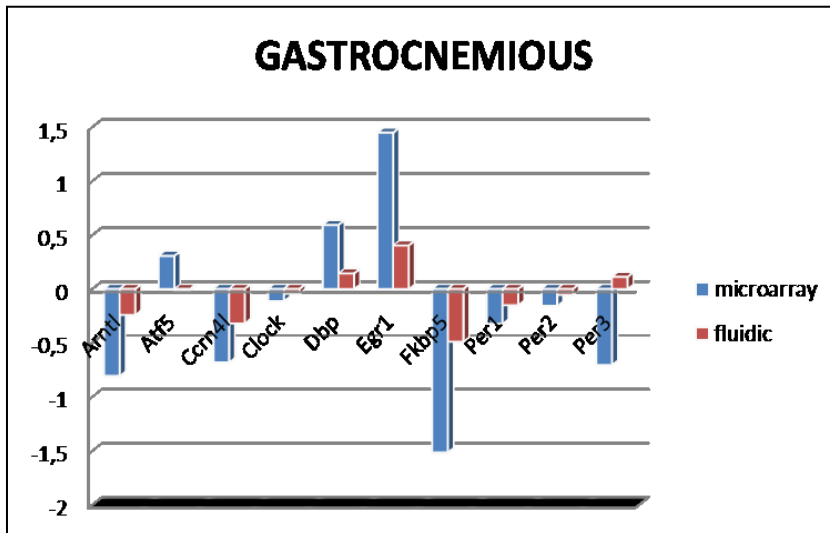
Fig 13. gene expression microarray data analysis (blue) and the fluidic cards validation (red) for each muscle in circadian pathway

The involvement of circadian genes in hereditary muscle dystrophy is fully novel, although it is already known that these genes are deeply involved in muscle remodeling, function, performance and aging (Andrews JL, et al. 2010; McCarthy JJ, et al. 2007). This new view opens interesting perspectives on new potential target genes involved in muscle damage pathogenesis and potentially a cure.

Circadian rhythms are approximate 24 hours (24-h) behavioral and physiological cycles that function to prepare an organism for daily environmental changes.

The basic clock mechanism is a network of transcriptional-translational feedback loops that drive rhythmic expression of genes over a 24-h period. The cycles are endogenously generated self-sustaining rhythms but can be influenced, in mammals, by environmental stimuli such as light and feeding (Ko CH. Et al, 2006, Levi F. et al. 2007, Shibler U. 2005). The temporal coordination between endogenous cellular rhythms and the environment has been experimentally shown to provide an adaptive advantage by enhancing an organism's ability to anticipate daily changes in

light, temperature, and humidity (Woelfle MA. Et al, 2004). The molecular mechanism regulating circadian rhythms is described as a network composed of transcriptional-translational feedback loops and is referred to as the core clock (Lowrey PL. et al, 2004, Sangoram AM. et al, 1998). The positive loop of the core clock is formed by two members *Clock* (circadian locomotor output control kaput), and *Bmal1* (brain muscle arnt-like 1). The CLOCK:BMAL1 heterodimer activates transcription of additional core clock genes *Period* (*Per1*, *Per2,* and *Per3*) and *Cryptochrome* (*Cry1* and *Cry2*). The CRY and PER proteins constitute the negative loop of the core clock by forming a multimeric complex that inhibits CLOCK:BMAL1 activity upon translocation to the nucleus. The orphan nuclear receptors *Rev-erb_*, *Rev-erb_,* and *Rora* (RAR-related orphan receptor alpha) are the remaining core clock genes and are interwoven into the core clock mechanism via their action in repressing and activating *Bmal1* gene transcription, respectively (Akashi M, et al. 2005, Lowrey PL. et al, 2004, Sato TK, et al. 2004, Ueda HR. et al, 2002) (Fig 14).
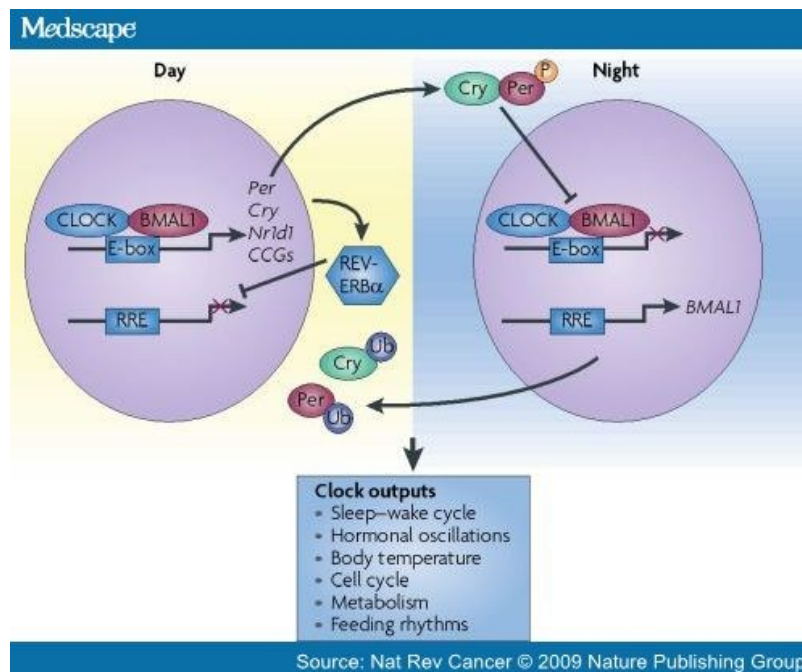

Fig. 14: schematic representation of circadian rhythm

We have provided, at least in part, the evidence that the core regulators of the circadian rhythm are deregulated in Col6a1$^{-/-}$ mice versus WT, highlighting a possible link between NMDs and circadian oscillators to be further and deeper clarified.

Expression profile has already been employed to determine the circadian transcriptome of adult mouse skeletal muscle. (McCarthy JJ, et al. 2007). Identifying genes associated with regulation of transcription, protein synthesis and degradation, and lipid metabolism illustrating the fundamental role that the core clock plays in maintaining cellular homeostasis. Few genes were identified as having a circadian pattern of expression in adult skeletal muscle: among of these Arntl 1 (Bmal1) Period 2 (Per2) transcript levels oscillate over 48h with a circadian period in adult skeletal muscle (Fig.15).
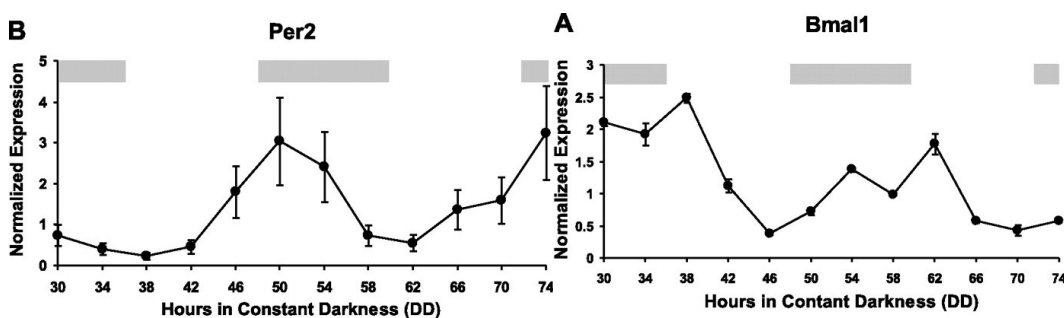

Fig 15. Arntl 1 (Bmal1) Period 2 (Per2) transcript levels oscillate over 48h with a circadian period in adult skeletal muscle (McCarthy J et al, 2007)

Clock mutant shown also an altered expression of circadian genes and MyoD, a known transcription factor key regulator of muscle differentiation; the identification of *Myod1* as a circadian gene was an exciting finding because it expands the potential function of *Myod1* beyond its known role in myogenesis to include an active role in the daily maintenance of adult skeletal muscle.

The circadian expression of *Myod1* transcript may be conserved between mouse and human as suggested by a recent study that found daily variation in *Myod1* expression in human skeletal muscle (Vissing K, et al. 2005). The circadian function of *Myod1* remains to be determined, but the possible conservation of *Myod1* circadian expression between species indicates an important role in skeletal muscle.

It is important to note that we have also found MyoD as a deregulated gene into GO group of transcriptional factors important for muscle related processes among those validated by fluidic cards.

MyoD is a master regulator of myogenesis and exhibits a circadian rhythm in its protein and RNA levels suggesting a possible role in the daily maintenance of muscle phenotype and function.

Another work (Andrews JL, et al. 2010) reports that MyoD is a direct target of the circadian transcriptional activators CLOCK and BMAL1, which bind in a rhythmic manner to the core enhancer of the MyoD promoter (all of them resulted deregulated in our findings). Skeletal muscle of ClockΔ19 and Bmal1−/− mutant mice exhibited ∼30% reductions in normalized maximal force miming NMDs (Fig. 16).
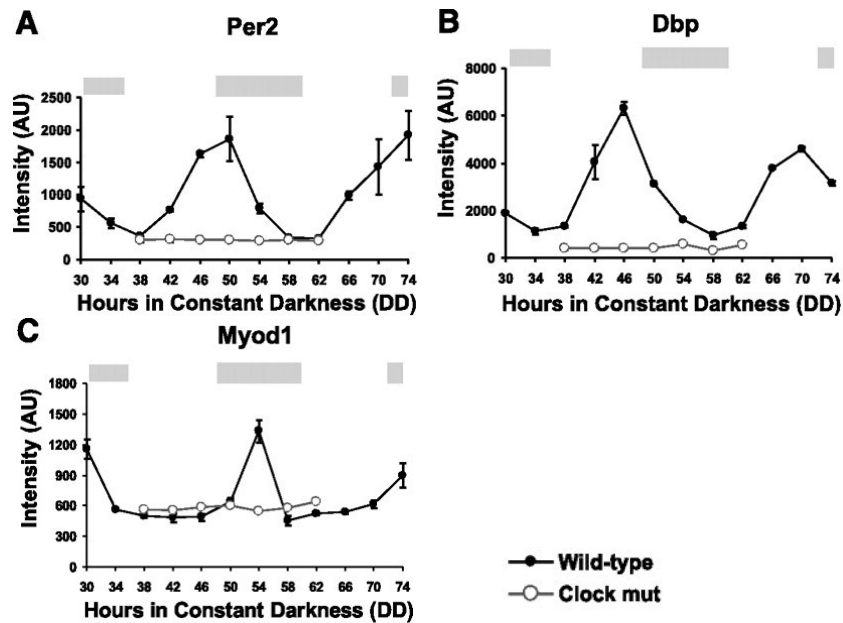
Fig. 16: Altered expression of circadian genes in *Clock* mutant. Approximately 30% of the circadian genes in the Clock mutant showed a loss in cycling. Graphs comparing wild-type vs. *Clock* mutant expression (*A–C*) (Andrews J L et al. PNAS 2010;107:19090-19095)

A similar reduction in force was observed at the single-fiber level. Electron microscopy (EM) showed that the myofilament architecture was disrupted in skeletal muscle of ClockΔ19, Bmal1−/−, and MyoD−/− mice. The alteration in myofilament organization was associated with decreased expression of actin, myosins, titin, and several MyoD target genes (Fig.17).
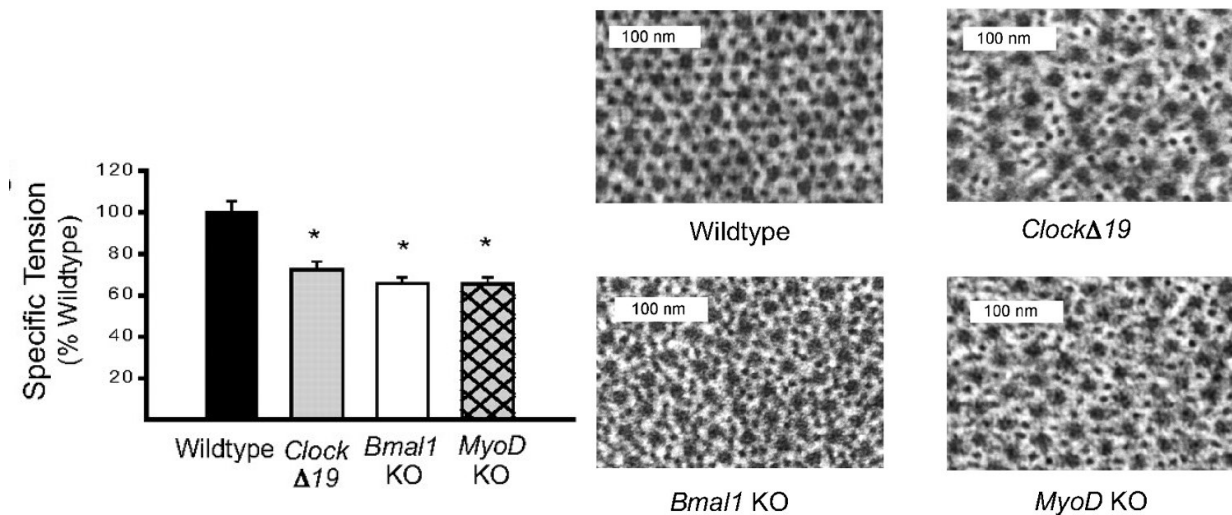
Fig 17. Decreased whole-muscle function and myofilament structure in ClockΔ19, Bmal−/−, and MyoD−/− mice.

It is important to highlight that similar reduction in specific tension, and these deficits are very comparable to those observed in Duchenne muscular dystrophy and others NMDs such as collagen VI myopathies as well as those seen in aging (Gonzalez E, et al. 2000; Lynch GS, Et al. 2001).

This suggests that the abnormalities observed in skeletal-muscle function might also be due to alterations in molecular clock function. The deficits recorded for whole-muscle tissue were also present at the single-cell level. These findings are consistent with the concept that proper function of the molecular clock in skeletal muscle fiber/cell is critical for maintenance of muscle function.

EM analysis also demonstrated that muscle from both ClockΔ19 and Bmal1−/− mice had a 40% reduction in mitochondrial volume. The remaining mitochondria in these mutant mice displayed aberrant morphology characterized by swelling and disrupted cristae and increased uncoupling of respiration (Fig. 18).
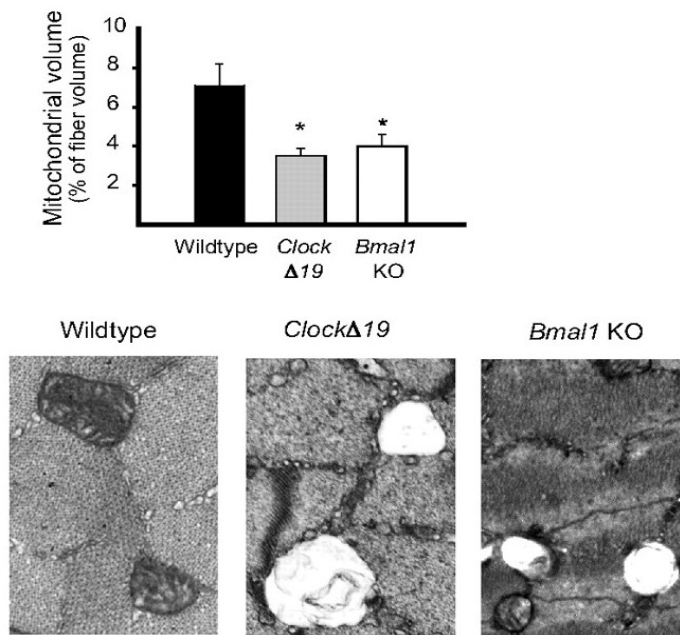
Fig 18. Decreased mitochondrial volume and respiratory function in muscle of ClockΔ19 and Bmal1−/− mice.

This result recapitulates the mitochondrial dysfunction found in mice with collagen VI deficiency (Irwin WA. et al. 2003) where authors have shown that Col6a1$^{-/-}$ muscles have a loss of contractile strength associated with ultrastructural alterations of sarcoplasmic reticulum (SR) and mitochondria with abnormal cristae with tubular shape and altered matrix density associated with the presence of dense bodies.

Taken together these results demonstrate that disruption of CLOCK or BMAL1 leads to structural and functional alterations at the cellular level in skeletal muscle. The identification of MyoD as a clock-controlled gene provides a mechanism by which the circadian clock may generate a muscle-specific circadian transcriptome in an adaptive role for the daily maintenance of adult skeletal muscle.

CONCLUSIONS

In conclusion evidence is growing that links inflammation, cancer, cardiovascular disease, sleep disorders, and metabolic disease with altered expression of the molecular clock in peripheral tissues (Duez H, et al. 2008. Maywood ES, et al. 2006. Takahashi JS, 2008). Thus, the significant force and metabolic deficits in the muscle of ClockΔ19 and Bmal1−/− mice, abnormal sarcoplasmic reticulum (SR) and mitochondria seen in Col6a1$^{-/-}$ mice, and pathways involved in muscle regeneration, extracellular matrix remodeling and inflammation, can open the possibility that the profound peripheral weakness and fatigue seen in chronic diseases may also be the result from disruption of proper circadian factor expression in muscle. Therefore, circadian oscillators and proteins might represent molecular biomarkers measuring the muscle damage and impairment in COL6 myopathies and other NMDs. Their strong link with environment opens also new avenues for novel therapeutic options.

REFERENCES

1. **Akashi M**, Takumi T. The orphan nuclear receptor RPRalpha regulates circadian transcription of the mammalian core-clock Bmal1. *Nat Struct Mol Biol* 12: 441-448, 2005

2. **Andrews JL**, Zhang X, McCarthy JJ, McDearmon EL, Hornberger TA, Russell B, Campbell KS, Arbogast S, Reid MB, Walker JR, Hogenesch JB, Takahashi JS, Esser KA. CLOCK and BMAL1 regulate MyoD and are necessary for maintenance of skeletal muscle phenotype and function. *Proc Natl Acad Sci U S A*. 2;107(44):19090-5, 2010

3. **Angelin A**, Tiepolo T, Sabatelli P, Grumati P, Bergamin N, Golfieri C, Mattioli E, Gualandi F, Ferlini A, Merlini L, Maraldi NM, Bonaldo P &Bernardi P. Mitochondrial dysfunction in the pathogenesis of Ullrich congenital muscular dystrophy and prospective therapy with cyclosporins. *Proc Natl Acad Sci USA* 104, 991−996, 2007

4. **Baker NL**, Morgelin M, Peat R, Goemans N, North KN, Bateman JF &Lamande SR, Dominant collagen VI mutations are a common cause of Ullrich congenital muscular dystrophy. *Hum Mol Genet* 14, 279−293, 2005

5. **Bertini E**, D'Amico A, Gualandi F, Petrini S. Congenital muscular dystrophies: a brief review. *Semin Pediatr Neurol*. 18(4):277-88, 2011

6. **Bidanset DJ**, Guidry C, Rosenberg LC, Choi HU, Timpl R &Hook M. Binding of the proteoglycan decorin to collagen type VI. *J Biol Chem* 267, 5250−5256, 1992

7. **Bonaldo P**, Braghetta P, Zanetti M, Piccolo S, Volpin D &Bressan GM. Collagen VI deficiency induces early onset myopathy in the mouse: an animal model for Bethlem myopathy. *Hum Mol Genet* 7, 2135−2140, 1998

8. **Bonaldo P**, Russo V, Bucciotti F, Doliana R &Colombatti A. Structural and functional features of the alpha 3 chain indicate a bridging role for chicken collagen VI in connective tissues. *Biochemistry* 29, 1245−1254, 1990

9. **Bonaldo P**, Russo V, Bucciotti F, Bressan GM &Colombatti A. Alpha 1 chain of chick type VI collagen. The complete cDNA sequence reveals a hybrid molecule made of one short collagen and three von Willebrand factor type A−like domains. *J Biol Chem* 264, 5575−5580, 1989

10. **Burg MA**, Tillet E, Timpl R &Stallcup WB. Binding of the NG2 proteoglycan to type VI collagen and other extracellular matrix molecules. *J Biol Chem* 271, 26110−26116, 1996

11. **Bushby KM**, Collins D. Collagen Type VI Myopathies. *Adv Exp Med Biol*. 802:185-99, 2014

12. **Chu ML**, Zhang RZ, Pan TC, Stokes D, Conway D, Kuo HJ, Glanville R, Mayer U, Mann K, Deutzmann R &Timpl R. Mosaic structure of globular domains in the human type VI collagen alpha 3 chain: similarity to von Willebrand factor, fibronectin, actin, salivary proteins and aprotinin type protease inhibitors. *EMBO J* 9, 385−393, 1990

13. **Cernei N**, Heger Z, Gumulec J, Zitka O, Masarik M, Babula P, Eckschlager T, Stiborova M, Kizek R, Adam V. Sarcosine as a potential prostate cancer biomarker--a review. *Int J Mol Sci*. 4;14(7):13893-908, 2013

14. **Camacho Vanegas O**, Bertini E, Zhang RZ, Petrini S, Minosse C, Sabatelli P, Giusti B, Chu ML &Pepe G. Ullrich scleroatonic muscular dystrophy is caused by recessive mutations in collagen type VI. *Proc Natl Acad Sci USA* 98, 7516−7521. 2001

15. **Colombatti A &Bonaldo P**. The superfamily of proteins with von Willebrand factor type A−like domains: one theme common to components of extracellular matrix, hemostasis, cellular adhesion, and defense mechanisms. *Blood* 77, 2305−2315, 1991

16. **Colombatti A**, Bonaldo P, Ainger K, Bressan GM &Volpin D. Biosynthesis of chick type VI collagen. I. Intracellular assembly and molecular structure. *J Biol Chem* 262, 14454−14460, 1987

17. **Colombatti A &Bonaldo P**. Biosynthesis of chick type VI collagen. II. Processing and secretion in fibroblasts and smooth muscle cells. *J Biol Chem* 262, 14461−14466, 1987

18. **Colombatti A**, Mucignat MT &Bonaldo P. Secretion and matrix assembly of recombinant type VI collagen. *J Biol Chem* 270, 13105−13111, 1995

19. **Demir E**, Sabatelli P, Allamand V, Ferreiro A, Moghadaszadeh B, Makrelouf M, Topaloglu H, Echenne B, Merlini L &Guicheney P. Mutations in COL6A3 cause severe and mild phenotypes of Ullrich congenital muscular dystrophy. *Am J Hum Genet* 70, 1446−1458, 2002

20. **de Visser M**, de Voogt WG, la Riviere GV. The heart in Becker muscular dystrophy, facioscapulohumeral dystrophy, and Bethlem myopathy. *Muscle Nerve.* 15:591–6, 1992

21. **Doliana R**, Bonaldo P &Colombatti A. Multiple forms of chicken alpha 3(VI) collagen chain generated by alternative splicing in type A repeated domains. *J Cell Biol* 111, 2197−2205, 1990

22. **Duez H**, Staels B. The nuclear receptors Rev-erbs and RORs integrate circadian rhythms and metabolism. *Diab Vasc Dis Res* 5: 82-88, 2008

23. **Dziadek M**, Kazenwadel JS, Hendrey JA, Pan TC, Zhang RZ, Chu ML. Alternative splicing of transcripts for the alpha 3 chain of mouse collagen VI: identification of an abundant isoform lacking domains N7-N10 in mouse and human. *Matrix Biol.* 21(3):227-41, 2002

24. **Furukawa T**, Toyokura Y. Congenital, hypotonic-sclerotic muscular dystrophy. *J Med Genet.* 14:426–9, 1977

25. **Gonzalez E**, Messi ML, Delbono O. The specific force of single intact extensor digito rum longus and soleus mouse muscle fibers declines with aging. *J Membr Biol* 178: 175-183, 2000

26. **Grumati P**, Coletto L, Sabatelli P, Cescon M, Angelin A. et al. Autophagy is defective in collagen VI muscular dystrophies, and its reactivation rescues myofiber degeneration. *Nat Med* 16: 1313-1320, 2010

27. **Gualandi F**, Curci R, Sabatelli P, Martoni E, Bovolenta M, Maraldi MN, Merlini L, Ferlini A. Macrophages: a minimally invasive tool for monitoring collagen VI myopathies. *Muscle Nerve* 44 (1):80-4, 2011

28. **Hansson MJ**, Mattiasson G, Mansson R, Karlsson J, Keep MF, Waldmeier P, Ruegg UT, Dumont JM, Besseghir K &Elmer E. The nonimmunosuppressive cyclosporin analogs NIM811 and UNIL025 display nanomolar potencies on permeability transition in brain−derived mitochondria. *J Bioenerg Biomembr* 36, 407−413, 2004

29. **Hicks D,** Farsani GT, Laval S, Collins J, Sarkozy A, Martoni E, Shah A, Zou Y, Koch M, Bönnemann CG, Roberts M, Lochmüller H, Bushby K, Straub V. Mutations in the Collagen XII gene define a new form of extracellular matrix-related myopathy. *Hum Mol Genet* Jan 12. [Epub ahead of print], 2014

30. **Haq RU**, Speer MC, Chu ML, Tandan R. Respiratory muscle involvement in Bethlem myopathy. *Neurology*. 52:174–6, 1999

31. **Irwin WA**, Bergamin N, Sabatelli P, Reggiani C, Megighian A, Merlini L, Braghetta P, Columbaro M, Volpin D, Bressan GM, Bernardi P &Bonaldo P. Mitochondrial dysfunction and apoptosis in myopathic mice with collagen VI deficiency. *Nat Genet* 35, 267−271, 2003

32. **Jöbsis GJ**, Boers JM, Barth PG, de Visser M. Bethlem myopathy: a slowly progressive congenital muscular dystrophy with contractures. *Brain*. 122 (Pt 4):649-55, 1999

33. **Jöbsis GJ**, Keizers H, Vreijling JP, de Visser M, Speer MC, Wolterman RA, Baas F, Bolhuis PA. Type VI collagen mutations in Bethlem myopathy, an autosomal dominant myopathy with contractures. *Nat Genet*. 14(1):113-5, 1996

34. **Keene DR**, Engvall E &Glanville RW. Ultrastructure of type VI collagen in human skin and cartilage suggests an anchoring function for this filamentous network. *J Cell Biol* 107, 1995−2006, 1988

35. Klipper-Aurbach Y, Wasserman M, Braunspiegel-Weintrob N, Borstein D, Peleg S, Assa S, Karp M, **Benjamini** Y, **Hochberg** Y, Laron Z. Mathematical formulae for the prediction of the residual beta cell function during the first two years of disease in children and adolescents with insulin-dependent diabetes mellitus. *Med Hypotheses*. 45(5):486-90, 1995

36. **Ko CH**, Takahashi JS. Molecular components of the mammalian circadian clock. *Hum Mol Genet* 15, *Spec No* 2: R271-R277, 2006.

37. **Kuo HJ**, Maslen CL, Keene DR &Glanville RW. Type VI collagen anchors endothelial basement membranes by interacting with type IV collagen. *J Biol Chem* 272, 26522−26529, 1997

38. **Lam SW**, Jimenez CR, Boven E. Breast cancer classification by proteomic technologies: current state of knowledge. *Cancer Treat Rev.* 40(1):129-38, 2014

39. **Lamandé SR**, Mörgelin M, Selan C, Jöbsis GJ, Baas F, Bateman JF. Kinked collagen VI tetramers and reduced microfibril formation as a result of Bethlem myopathy and introduced triple helical glycine mutations. *J Biol Chem*. 18;277(3):1949-56, 2002

40. **Lamandé SR**, Shields KA, Kornberg AJ, Shield LK, Bateman JF. Bethlem myopathy and engineered collagen VI triple helical deletions prevent intracellular multimer assembly and protein secretion. *J Biol Chem*. 274(31):21817-22, 1999

41. **Lampe AK &Bushby KM**. Collagen VI related muscle disorders. *J Med Genet* 42, 673−685, 2005

42. **Lampe AK,** Flanigan KM, Bushby KM, Hicks D, Pagon RA, Adam MP, Bird TD, Dolan CR, Fong CT, Stephens K. Collagen Type VI-Related Disorders. *GeneReviews™* [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2013.

43. **Levi F**, Shibler U. Circadian rhythm: mechanisms and therapeutic implications. *Annu Rev Pharmacol Toxicol* 47: 593-628, 2007

44. **Lowrey PL**, Takahashi JS. Mammalian circadian biology: elucidating genome-wide levels of temporal organization. *Annu Rev Genomics Hum Genet* 5:407-441, 2004

45. **Lucioli** S, Giusti B, Mercuri E, Vanegas OC, Lucarini L, Pietroni V, Urtizberea A, Ben Yaou R, de Visser M, van der Kooi AJ, Bönnemann C, Iannaccone ST, Merlini L, Bushby

K, Muntoni F, Bertini E, Chu ML, Pepe G. Detection of common and private mutations in the COL6A1 gene of patients with Bethlem myopathy. *Neurology*. 64(11):1931-7, 2005

46. **Lynch GS,** Hinkle RT, Chamberlain JS, Brooks SV, Faulkner JA. Force and power output of fast and slow skeletal muscles from mdx mice 6-28 months old. *J Physiol* 535: 591-600, 2001

47. **Marotta M**, Ruiz-Roig C, Sarria Y, Peiro JL, Nuñez F, Ceron J, Munell F, Roig-Quilis M. Muscle genome-wide expression profiling during disease evolution in mdx mice. *Physiol Genomics*. 37(2):119-32, 2009

48. **Martoni E**, Petrini S, Trabanelli C, Sabatelli P, Urciuolo A, Selvatici R, D'Amico A, Falzarano S, Bertini E, Bonaldo P, Ferlini A, Gualandi F. Characterization os a rare case of Ullrich Congenital Muscular Dystrophy due to truncating mutations within the COL6A1 gene C-Terminal domain: a case report. *BMC Ned Genet* 5:14-59, 2013

49. **Maywood ES**, O'Neill J, Wong GK, Reddy AB, Hasting MH. Circadian timing in health and disease. *Prog Brain Res* 153: 253-269, 2006

50. **McCarthy JJ**, Andrews JL, McDearmon EL, Campbell KS, Barber BK, Miller BH, Walker JR, Hogenesch JB, Takahashi JS, Esser KA. Identification of the circadian transcriptome in adult mouse skeletal muscle. *Physiol Genomics* 31: 86-95, 2007

51. **Mendell JR**, Boué DR, Martin PT. The congenital muscular dystrophies: recent advances and molecular insights. *Pediatr Dev Pathol.* 9(6):427-43, 2006

52. **Merlini L**, Martoni E, Grumati P, Sabatelli P, Squarzoni S, Urciuolo A, Ferlini A, Gualandi F, Bonaldo P. Autosomal recessive myosclerosis myopathy is a collagen VI disorder. *Neurology*. 71:1245–53, 2008

53. **Mohire MD**, Tandan R, Fries TJ, Little BW, Pendlebury WW, Bradley WG. Early-onset benign autosomal dominant limb-girdle myopathy with contractures (Bethlem myopathy). *Neurology*. 38:573–80, 1998

54. **Muntoni F**, Bertini E, Bonnemann C, Brockington M, Brown S, Bushby K, Fiszman M, Korner C, Mercuri E, Merlini L, Hewitt J, Quijano-Roy S, Romero N, Squarzoni S, Sewry CA, Straub V, Topaloglu H, Haliloglu G, Voit T, Wewer U, Guicheney P (2002) 98th ENMC International Workshop on Congenital Muscular Dystrophy (CMD), 7th Workshop of the International Consortium on CMD, 2nd Workshop of the MYO CLUSTER project GENRE. 26-28th October, 2001, Naarden, The Netherlands. *Neuromuscul Disord* 12:889-96

55. **Paco S**, Kalko SG, Jou C, Rodriguez MA, Corbera J, Muntoni F, Feng L, Rivas E, Torner F, Gualandi F, Gomez-Foix AM, Ferrer A, Ortez C, Nascimento A, Colomer J, Jimenez-Mallebrera C. Gene expression profiling identifies molecular pathways associated with collagen VI deficiency and provides novel therapeutic targets. *PLoS One*. 8(10):e77430, 2013

56. **Pan TC**, Zhang RZ, Sudano DG, Marie SK, Bonnemann CG &Chu ML. New molecular mechanism for Ullrich congenital muscular dystrophy: a heterozygous in−frame deletion in the COL6A1 gene causes a severe phenotype. *Am J Hum Genet* 73, 355−369. 2003

57. **Pepe G**, Bertini E, Bonaldo P, Bushby K, Giusti B, de Visser M, Guicheney P, Lattanzi G, Merlini L, Muntoni F, Nishino I, Nonaka I, Yaou RB, Sabatelli P, Sewry C, Topaloglu H &van der Kooi A. Bethlem myopathy (BETHLEM) and Ullrich scleroatonic muscular dystrophy: 100th ENMC international workshop, 23−24 November 2001, Naarden, The Netherlands. *Neuromuscul Disord* 12, 984−993, 2002

58. **Pepe G**, Bertini E, Giusti B, Brunelli T, Comeglio P, Saitta B, Merlini L, Chu ML, Federici G, Abbate R. A novel de novo mutation in the triple helix of the COL6A3 gene in a two-generation Italian family affected by Bethlem myopathy. A diagnostic approach in the mutations' screening of type VI collagen. *Neuromuscul Disord*. 9(4):264-71, 1999

59. **Pepe G**, Giusti B, Bertini E, Brunelli T, Saitta B, Comeglio P, Bolognese A, Merlini L, Federici G, Abbate R, Chu ML. A heterozygous splice site mutation in COL6A1 leading to an in-frame deletion of the alpha1(VI) collagen chain in an italian family affected by bethlem myopathy. *Biochem Biophys Res Commun*. 19;258(3):802-7, 1999

60. **Petrini S**, Tessa A, Carrozzo R, Verardo M, Pierini R, Rizza T &Bertini E. Human melanoma/NG2 chondroitin sulfate proteoglycan is expressed in the sarcolemma of

postnatal human skeletal myofibers. Abnormal expression in merosin−negative and Duchenne muscular dystrophies. *Mol Cell Neurosci* 23, 219−231, 2003

61. **Petrini S**, Tessa A, Stallcup WB, Sabatelli P, Pescatori M, Giusti B, Carrozzo R, Verardo M, Bergamin N, Columbaro M, Bernardini C, Merlini L, Pepe G, Bonaldo P &Bertini E. Altered expression of the MCSP/NG2 chondroitin sulfate proteoglycan in collagen VI deficiency. *Mol Cell Neurosci* 30, 408−417, 2005

62. **Pfaff M**, Aumailley M, Specks U, Knolle J, Zerwes HG &Timpl R. Integrin and Arg−Gly−Asp dependence of cell adhesion to the native and unfolded triple helix of collagen type VI. *Exp Cell Res* 206, 167−176, 1993

63. **Sabatelli P**, Bonaldo P, Lattanzi G, Braghetta P, Bergamin N, Capanni C, Mattioli E, Columbaro M, Ognibene A, Pepe G, Bertini E, Merlini L, Maraldi NM &Squarzoni S. Collagen VI deficiency affects the organization of fibronectin in the extracellular matrix of cultured fibroblasts. *Matrix Biol* 20, 475−486, 2001

64. **Sabatelli P,** Gualandi F, Gara SK, Grumati P, Zamparelli A, Martoni E, Pellegrini C, Merlini L, Ferlini A, Bonaldo P, Maraldi NM, Paulsson M, Squarzoni S, Wagener R. Expression of Collagen VI α5 and α6 chains in human muscle and in Duchenne muscular dystrophy-related muscle fibrosis. *Matrix Biol* 31(3):187-96, 2012

65. **Saitta** B, Stokes DG, Vissing H, Timpl R, Chu ML. Alternative splicing of the human alpha 2(VI) collagen gene generates multiple mRNA transcripts which predict three protein variants with distinct carboxyl termini. *J Biol Chem*. 15;265(11):6473-80, 1990

66. **Sangoram AM**, Saez L, Antoch MP, Gekakis N, Staknis D, Whiteley A, Fruechte EM, Vitaterna MH, Shimomura K, King DP, Young MW, Weitz CJ, Takahashi JS. Mammalian circadian autoregulatory loop: a timeless ortholog and mPer1 interact and negatively regulate CLOCK – BMAL1 – induced transcription. *Neuron* 21: 1101-1113, 1998.

67. **Sasaki T**, Hohenester E, Zhang RZ, Gotta S, Speer MC, Tandan R, Timpl R &Chu ML. A Bethlem myopathy Gly to Glu mutation in the von Willebrand factor A domain N2 of the collagen alpha3(VI) chain interferes with protein folding. *FASEB J* 14, 761−768, 2000

68. **Sato TK**, Panda S, Miraglia LJ, Reyes TM, Rudic RD, McNamara P, Naik KA, FitzGerald GA, Kay SA, Hogenesch JB. A functional genomic strategy reveals Rora as a component of the mammalian circadian clock. *Neuron* 43: 527-537, 2004

69. **Scacheri PC**, Gillanders EM, Subramony SH, Vedanarayanan V, Crowe CA, Thakore N, Bingler M, Hoffman EP Novel mutations in collagen VI genes: expansion of the Bethlem myopathy phenotype. *Neurology*. 26;58(4):593-602, 2002

70. **Shapiro E**, Biezuner T, Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet.*14(9):618-30, 2013

71. **Shibler U**. The dailt rhythm of genes and organs. Biological clocks and circadian timing in cells. *EMBO Rep* 6, *Spec No*: S9-13, 2005

72. **Squarzoni S**, Sabatelli P, Bergamin N, Guicheney P, Demir E, Merlini L, Lattanzi G, Ognibene A, Capanni C, Mattioli E, Columbaro M, Bonaldo P &Maraldi NM. Ultrastructural defects of collagen VI filaments in an Ullrich syndrome patient with loss of the alpha3(VI) N10−N7 domains. *J Cell Physiol* 206, 160−166, 2006

73. **Stallcup WB**. The NG2 proteoglycan: past insights and future prospects. *J Neurocytol* 31: 423−435, 2002

74. **Stokes** DG, Saitta B, Timpl R, Chu ML. Human alpha 3(VI) collagen gene. Characterization of exons coding for the amino-terminal globular domain and alternative splicing in normal and tumor cells. *J Biol Chem*. 5;266(13):8626-33, 1991

75. **Takahashi JS**, Hong HK, Ko CH, McDearmon EL. The genetics of mammalian circadian order and disorder: implications for physiology and disease. *Nat Rev Genet* 9: 764-775, 2008

76. **'t Hoen** PA, van der Wees CG, Aartsma-Rus A, Turk R, Goyenvalle A, Danos O, Garcia L, van Ommen GJ, den Dunnen JT, van Deutekom JC. Gene expression profiling to monitor therapeutic and adverse effects of antisense therapies for Duchenne muscular dystrophy. *Pharmacogenomics*. 7(3):281-97, 2006

77. **Tillet E**, Gential B, Garrone R &Stallcup WB. NG2 proteoglycan mediates beta1 integrin−independent cell adhesion and spreading on collagen VI. *J Cell Biochem* 86, 726−736, 2002

78. **Ueda HR**, Chen W, Adachi A, Wakamatsu H, Hayashi S, Takasugi T, Nagano M, Nakahama K, Suzuky Y, Sugano S, Iino M, Shigeyoshi Y, Hashimoto S. A transcription factor response element for gene expression during circadian night. *Nature* 418: 534-539, 2002

79. **Vanegas** OC, Zhang RZ, Sabatelli P, Lattanzi G, Bencivenga P, Giusti B, Columbaro M, Chu ML, Merlini L, Pepe G. Novel COL6A1 splicing mutation in a family affected by mild Bethlem myopathy. *Muscle Nerve*. 25(4):513-9, 2002

80. **Wang SM**, Zhao G, Peng L, Yang W. Exome sequencing reveals new insights into the progression of abdominal aortic aneurysm. E*ur Rev Med Pharmacol Sci*. 17(17):2401-9, 2013

81. **Wettenhall JM**, Smyth GK. limmaGUI: a graphical user interface for linear modeling of microarray data. *Bioinformatics*. 12;20(18):3705-6, 2004

82. **Vissing K**, Andersen HL, Schjerling P. Are exercise-induced genes induced by exercise? *FASEB J* 19: 94 − 96, 2005

83. **Woelfle MA**, Ouyang Y. Phanvijhitsiri K. Johnson CH. The adaptive value of circadian clocks: an experimental assessment in cyanobacteria. *Curr Biol* 14: 1481-1486, 2004

84. **Zhang RZ**, Sabatelli P, Pan TC, Squarzoni S, Mattioli E, Bertini E, Pepe G &Chu ML. Effects on collagen VI mRNA stability and microfibrillar assembly of three COL6A2 mutations in two families with Ullrich congenital muscular dystrophy. *J Biol Chem* 277, 43557−43564, 2002

ACKNOWLEDGMENTS

Giunti a questo nuovo ed ultimo traguardo non mi resta che ringraziare tutti coloro che lo hanno reso possibile e mi hanno accompagnato durante tutto il percorso del dottorato di ricerca.

Innanzitutto ringrazio il mio tutor, la Prof.ssa Ferlini che dopo la tesi di laurea e la tesi di specializzazione non ha mancato di seguirmi anche durante il cammino del dottorato regalandomi, nel tragitto, la possibilità di una meravigliosa esperienza a Newcastle che sarà per sempre un grande bagaglio di crescita.

Ringrazio Francesca Gualandi che ha saputo insegnarmi quasi tutto quello che ho imparato, insieme a Cecilia Trabanelli, Paola Rimessi, Anna Venturoli, Anna Ravani, Marina Taddei e tutta la di genetica medica perchè mi hanno sempre supportato negli esperimenti di laboratorio e nella teoria.

Ringrazio il mio collega Matteo Bovolenta e le mie colleghe e amiche Chiara Scotton, Marcella Neri, Marina Fabris, Chiara Passarelli, Sofia Falzarano, Simona Brioschi, Elena Bassi, Annarita Armaroli e Francesca di Raimo per gli anni bellissimi e intensi trascorsi assieme.

Un grazie specialissimo e unico va alla mia famiglia: al mio piccolino Riccardo che è diventato il centro della mia vita per la persona migliore che mi ha fatto diventare giorno dopo giorno, perché tornare un po' bambini è davvero bellissimo. A Enrico che è la mia roccia, che mi ha sempre supportato e tanto sopportato nei periodi più difficili senza mai farmi perdere d'animo.

Ai miei genitori, ai miei zii e ai miei suoceri, che per quanto brontolo mi hanno aiutato come nessuno potrebbe, mostrandosi sempre così disponibili con il nipotino e con i miei mille orari sempre diversi. Senza di voi non ce l'avrei davvero mai fatta!

A mio fratello Andrea e a mia nonna Marina per essermi sempre vicino anche se purtroppo ci vediamo poco.

Alle mie amiche di sempre, a Lara ed Elena per avere la fortuna di avere amiche Vere come lo siete voi, e per esserci davvero sempre state nei momenti spensierati e soprattutto in quelli più difficili.

A Cristian, inseparabile amico d'infanzia, per tutto ciò che abbiamo trascorso insieme e che ricordo sempre con un nodo alla gola, e per avere dato una svolta alla mia vita.

Ora inizia per me un nuovo percorso, ma non dimenticherò mai questi anni bellissimi, grazie a tutti voi.