Molecular mechanisms in muscular dystrophy : a gene expression profiling study

Turk, R.

Molecular Mechanisms In Muscular Dystrophy

A Gene Expression Profiling Study

Molecular Mechanisms In Muscular Dystrophy

A Gene Expression Profiling Study

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van de Rector Magnificus Dr. D.D.Breimer,
hoogleraar in de faculteit der Wiskunde en
Natuurwetenschappen en die der Geneeskunde,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 27 september 2006
klokke 15.00 uur

door

Rolf Turk

Geboren te Leiden in 1975

Promotiecommissie

Promotor	Prof. Dr. G.J.B. van Ommen
Co-promotores	Dr. J.T. den Dunnen Dr. P.A.C. ‘t Hoen
Referent	Prof. Dr. R.M.W. Hofstra (Rijksuniversiteit Groningen)
Overige leden	Prof. Dr. M. Koenig (Université Louis Pasteur de Strasbourg)

An experiment is a question which science poses to Nature, and a measurement is the recording of Nature's answer.

Max Planck

Aan Maaike, Gerard en Annet

Printed by: Drukkerij Duineveld

ISBN-10: 90-9021042-3
ISBN-13: 978-90-9021042-1

Turk, Rolf
Molecular mechanisms in muscular dystrophy. A gene expression profiling study.
Thesis, Leiden University Medical Center
September 27, 2006

Molecular Mechanisms In Muscular Dystrophies

Preface

The muscular dystrophies are a group of neuromuscular disorders characterized by progressive muscle weakness and wasting. In the past two decades, the genetic causes of individual muscular dystrophies have been elucidated, which increased successful diagnosis and subsequent classification of the various muscular dystrophies. Although the underlying genetic defects of a large number of muscular dystrophies are now known, the molecular mechanisms resulting in the devastating effects of the disease are not yet clear. Furthermore, the muscular dystrophies differ in clinical presentation and severity. The processes responsible for this divergence are largely unknown as well.

In this thesis, gene expression profiling has been applied to study the molecular and cellular mechanisms and subsequent biological processes that play a role in muscular dystrophy. To this extent, we have determined gene expression levels in muscle tissue from different mouse models for muscular dystrophy. To characterize the processes associated with regeneration, we have compared gene expression levels in hindlimb muscle tissue of mdx and control mice in a temporal study. Additionally, we have determined the gene expression profiles of differentiating human myoblasts in vitro, since regeneration processes recapitulate myogenesis. We also set out to compare gene expression levels of different mouse models for muscular dystrophy to find common and distinct molecular mechanisms that underlie different forms of muscular dystrophy. Accordingly, we first had to determine the effects of genetic background variation between inbred mouse strains, and to study the feasibility of alternative experimental designs.

The studies described above form the core of this thesis. The introduction gives an overview of myogenesis, muscle structure, muscular dystrophy, and gene expression profiling. The discussion describes the experimental outcome of the studies in a broader perspective.

Chapter 1

Introduction

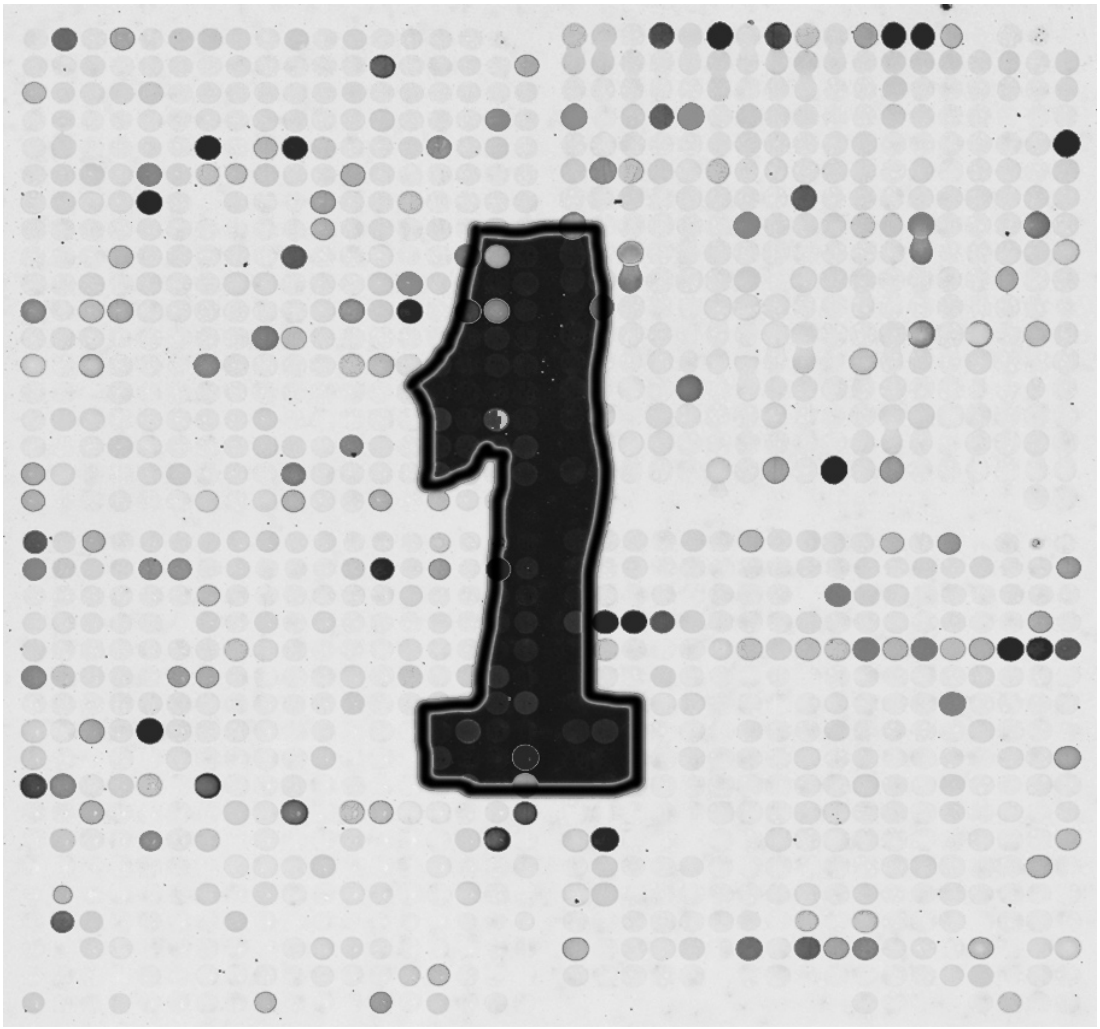


Table of Contents

1 Muscle

According to the Oxford dictionary, muscle is an elastic substance in the body that can be tightened or loosened to produce movement. More specifically, muscle tissue has the ability to contract and relax in order to produce motion. Muscle comes in three forms, namely skeletal, cardiac, and smooth muscle. The research in this thesis will mainly focus on skeletal muscle tissue. The function of skeletal muscle tissue concerns the ability to contract and extend in order to position bones or skin. This function is accomplished by the building blocks of muscle, the myofibers. Myofibers are long, cylindrical, multinucleated cells that are packed with contractable filaments. These filaments, or myofibrils, are primarily made of thin actin and thick myosin molecules, which are aligned precisely. Successively ordered contractile units, or sarcomeres, result in the banded, or striated, appearance of skeletal muscle tissue.

Each myofiber is wrapped in a sheet of connective tissue, also known as the endomysium. A bundle of myofibers forms a fascicle, which is surrounded by the perimysium. Finally, an individual muscle consists of multiple fascicles, is surrounded by the epimysium, and is connected to skeletal or skin-like structures via tendons (Figure 1). Thus, the anatomy of muscle tissue reflects an highly organized structure.

Muscle tissue does not consist of myofibers alone. Muscle action is controlled by the stimulation of a motor neuron via the neuromuscular junction. The axonal ending of the motor neuron transfers an action potential, which results in the contraction of the myofiber. Each myofiber is activated via a single neuromuscular junction. Furthermore, muscle tissue is perfused by the vascular system, which is responsible for both the delivery of oxygen and nutrients, and the disposal of metabolic waste. Within the muscle tissue, the internal innate defense is present, which helps to protect the muscle from pathogens, and helps tissue repair processes. The effector cells of the innate defense system are mainly represented by macrophages, dendritic cells, and B-cells[148].

Muscle tissue has a high regeneration potential. Subsequent to muscle injury by trauma or exercise, muscle precursor cells can proliferate and differentiate to form new myofibers, or to fuse with existing ones. This process demonstrates large similarities with embryonic myogenesis[42].

In this thesis the muscular dystrophies play a central role. Muscular dystrophies are characterized by progressive wasting of muscle tissue, which is replaced by adipose and fibrotic tissue. As a result, the affected muscle is not able to function properly due to weakness with often dramatic results. In Duchenne Muscular Dystrophy (DMD), the most severe muscular dystrophy, patients die in their twenties due to respiratory or cardiac failure.

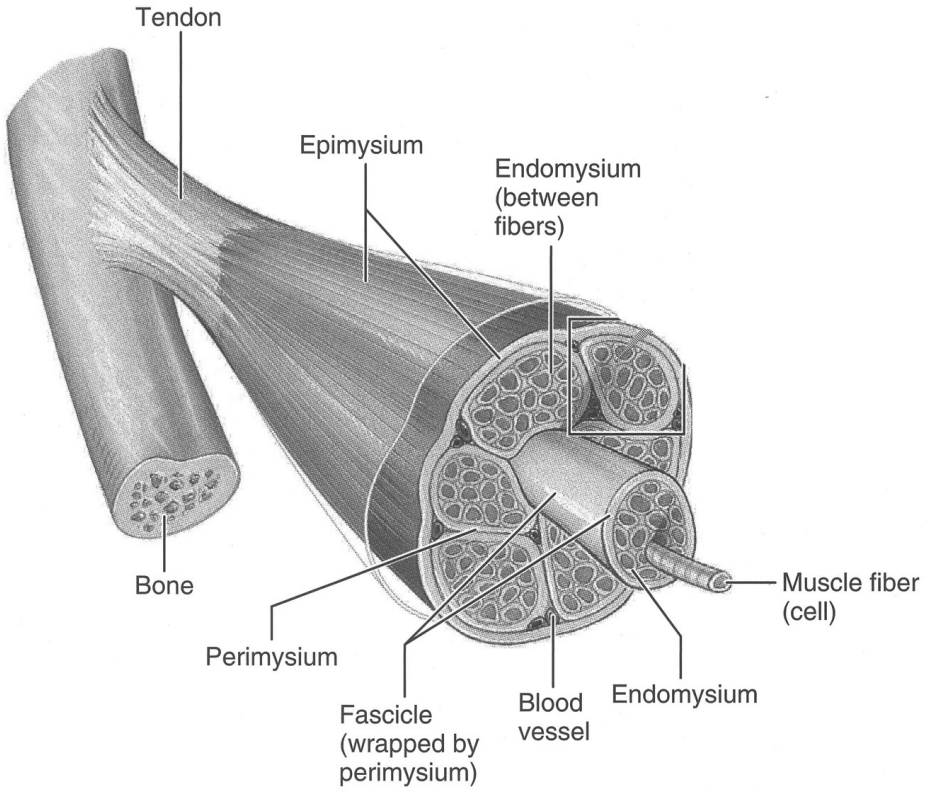


Figure 1 The anatomy of muscle tissue

The anatomy of muscle reflects a highly ordered sturcture, where separate muscle stuctures of different levels are packed in specific connective tissue layers.

(from Human Anatomy & Physiology, 6th Edition, Elaine N. Marieb, Pearson Education, Inc.)

1.1 Myogenesis

As with the formation of tissues in general, the development of muscle tissue is highly coordinated during the development of the embryo. In the adult, muscle is present throughout the entire body. Muscle derives from the mesoderm, which is a single embryonic layer. The mesoderm is the primary embryonic tissue for many tissues, such as the heart, blood, the vascular system, skin, and bone. During development, the paraxial mesoderm segments into somites, which are located on either side of the neural tube (Figure 2)[44]. The newly formed somites harbour specific cell populations, which function as progenitor cells for the development of specialized tissues. The specific cell populations are formed as the somites expand. A layer of pseudo-stratified columnar cells form the dermomyotome, which grows dorsomedially and ventrolaterally, thereby forming the epaxial and hypaxial myotome, respectively[217]. The epaxial myotome will, eventually, form the intrinsic musculature of the back, which surrounds and attaches to the vertebrae. The hypaxial myotome will provide the myogenic cells that will make up the musculature of the body wall, limbs, diaphragm, and neck[242]. The musculature of the head is formed by two separate systems. The tongue and laryngeal muscles are formed by the most anterior somites, whereas the rest of the head musculature is derived from unsegmented mesoderm present in the region that will develop into brain[69].

Muscle development from progenitor cell populations depends on specific temporal and spatial cell signalling, which is initiated and regulated by the expression of specific transcription factors. This cell signalling originates from surrounding embryonal tissues[69], and is called embryonal induction. Cell signalling is responsible for the delamination (or detachment), determination, migration, proliferation and differentiation of muscle progenitor cells (reviewed in Buckingham et al.[31] and Parker et al.[176]). One of the earliest transcription factors playing a role in myogenesis is Pax3, which is characterized by homeo- and paired-domain motifs[75]. Pax3 is already expressed in the presomitic mesoderm, but plays a critical role in the delamination of muscle precursor cells from the hypaxial myotome prior to migration. Interaction of c-met, a tyrosine kinase receptor, and its ligand hepatocyte growth factor (HGF) is necessary to determine the migratory route. HGF is produced by non-somitic mesodermal cells, which therefore direct the migration[63].

A specific group of transcription factors, which eventually determine the myogenic lineage, are the myogenic regulatory factors (MRFs). MRFs contain a basic-helix-loop-helix (bHLH) domain to which E-proteins can bind and form hetero-dimers. Subsequently, the heterodimers can bind to specific DNA sequences, called E-boxes, which are present in promoter and enhancer regions of skeletal muscle specific genes[29]. During development, the first MRF to be expressed is Myf5. The expression of Myf5 in the epaxial myotome is thought to be regulated by expression of Sonic hedgehog (Shh) from the notochord and the neural tube[25,87]. Determination towards the myogenic lineage requires the expression of MyoD, which belongs to the muscle regulatory factors. MyoD expression in the hypaxial myotome is stimulated by Wnt-signalling from the dorsal ectoderm, whereas MyoD expression is inhibited by transforming growth factor β like (TGF β like) signalling from the lateral-plate mesoderm[176]. Both Myf5 and MyoD, or 'primary' MRFs, are thought to function in the first step towards muscle formation, namely the determination which turns muscle precursor cells into myoblasts[188].

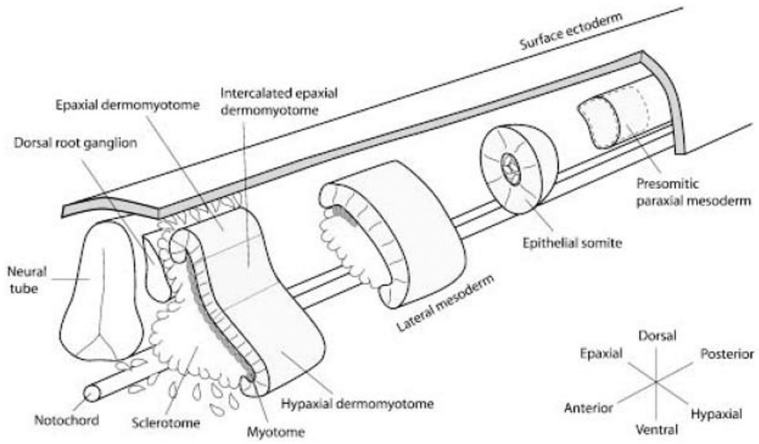


Figure 2 Embryonic development of muscle tissue

The paraxial mesoderm develops into somites, which initially consists of the sclerotome and the lateral mesoderm. The lateral mesoderm expands dorsomedially and ventrolaterally, thereby forming the epaxial and hypaxial dermomyotome, respectively. The epaxial dermomyotome curls inside to form the dorsomedial lip (DML), whereas the hypaxial dermomyotome curls inside to form the ventrolateral lip (VLL). Migrating hypaxial cells or myoblasts eventually develop into functional skeletal muscle tissue.

(from Buckingham et al., J Anat (2003), ***202****, pp59-68.)*

The second step concerns the differentiation of myoblasts, and the subsequent fusion to form myotubes and eventually myofibers. The molecular details of myoblast fusion are reviewed by Abmayr et al.[1]. Differentiation of myoblasts is coupled to withdrawal from the cell cycle, and thereby from the proliferative state. Differentiation is specified by the activation of contractile, scaffolding, and control protein genes of the myofibril[169]. Transcription of these proteins is carefully regulated by muscle specific transcription factors. Two of these transcription factors are the 'secondary' MRFs, Myogenin and Mrf4. Other important transcription factors for myogenesis are Mef2-isoforms and members of the Six-family of nuclear factors.

During limb formation, individual muscles are formed from single populations of myogenic cells, and are divided by connective tissue, which also ensures the attachment to the skeleton[20,28]. Muscle fiber formation is instigated by the fusion of myoblasts. The production of myofibers is biphasic; two waves of myoblast fusion can be determined[170,243]. The first wave produces the primary myofibers, which then play a critical role in the formation of a 'second wave' of secondary myofibers. Between 5 and 20 secondary myofibers are positioned around the primary myofiber, although the number can vary and can be higher in large animals[241]. At this stage, no additional muscle fibers are formed. Further growth (both lateral and longitudinal) of muscles is realized by the proliferation and subsequent fusion of myogenic cells to existing fibers[242].

1.2 Muscle Structure

Adult skeletal muscle tissue consists of a number of different cell types and structures, which contribute to the muscle structure and function, and are discussed below.

1.2.1 Myofiber type

Skeletal muscle is a versatile tissue, since it is able to perform multiple functions ranging from short bursts of activity to long lasting activity as seen in posture support or chewing. To make this broad range possible as well as efficient, muscles are built from specialized myofibers. Basically, there are two types of myofibers: slow contracting fibers with resistance to fatigue, and fast contracting fibers which are susceptible to fatigue. The localization and function of the muscle determines the presence of certain myofiber types, and therefore whether the muscle is slow, fast, or a mixture of both.

Both slow and fast myofibers are made from the same building blocks, but differ in the expression of specific protein isoforms. A well-characterized example is displayed by the Myosin heavy chain (MyHC) isoforms, which confer contractive force to the muscle.

The myofiber type of a muscle is determined by a number of factors. Fiber type specification might be initiated in the myoblast population from which a specific muscle is formed. Within a single muscle, different myofiber types may exist. In general, slow myofibers are localized in the interior of the muscle, whereas fast myofibers are located at the periphery. Neuronal activity by functional demands can also determine the fiber type. This effect is clearly shown by the application of exercise on muscle. Endurance training will transform muscle to a slow fiber type, whereas short bursts of activity (seen in sprinting, weight-lifting etc.) will produce a fast fiber type musculature. There is no evidence so far, that the innervation during development plays a role in the determination of the fiber types[242].

1.2.2 Muscle Precursor Cells

Muscle tissue has a high regenerative potential after injury, thereby preventing the loss of muscle mass. Muscle regeneration is achieved by the activation, proliferation, and differentiation of muscle precursor cell (satellite cells), and the subsequent fusion to existing or newly formed myofibers[200]. The population of satellite cells (SC) is maintained by the capability of self-renewal. Satellite cells were first described by Mauro et al.[142]. Satellite cells are a distinct population of mononuclear cells, different from embryonic and fetal myoblasts. The population as such is formed at the ‘last’ wave of myoblast fusion during the 10-14th week in human development, or from around E17.5 in murine embryos[50]. Approximately 30% of the sublaminar nuclei in postnatal muscle tissue belong to satellite cells. The cytoplasm of SC is sparse, but contains most of the organelles. The nucleus is oval, and consists of a large amount of heterochromatin. However, this does not indicate that the SC is inactive, since ribosomes are present, as well as rough endoplasmic reticulum and Golgi apparatus[18]. Satellite cells are generally localized near the neuromuscular junction[244].

Upon activation, SC migrate across the myofiber to the place of injury[198]. Furthermore, SC can also migrate between individual myofibers and fuse to them[121]. Satellite cells are able to escape from quiescence by re-entering the cell cycle due to a wide variety of conditions (injury, overwork, denervation, exercise, and stretch)[18]. Satellite cell activation is a dual process. First, the SC needs to abandon the G_0-phase and enter the G_1-phase by passing a certain restriction point. This process is mediated by so-called competence factors. Subsequently, progression factors are needed to keep the SC in a proliferative state[174]. Candidates for compentence and progression factors are fibroblast growth factor (FGF) and insulin-like growth factor (IGF), respectively[18].

Currently, a large number of SC activating factors from different origin are known (reviewed in Hawke et al.[95]). Examples of autocrine factors are IGF, FGF, HGF, and TGF-β like factors. These factors can also be released via the vascular system. Immune-related cells (macrophages, neutrophils, T-cells, etc.) can modulate activation of SC by releasing above factors, as well as platelet derived growth factor (PDGF), cytokines, or interleukines. Neurotrophic factors via motor neuron stimulation might activate SC. Furthermore, hormones (testosteron) or small metabolites can mediate the same effect.

1.2.3 Extra- and Intracellular structure

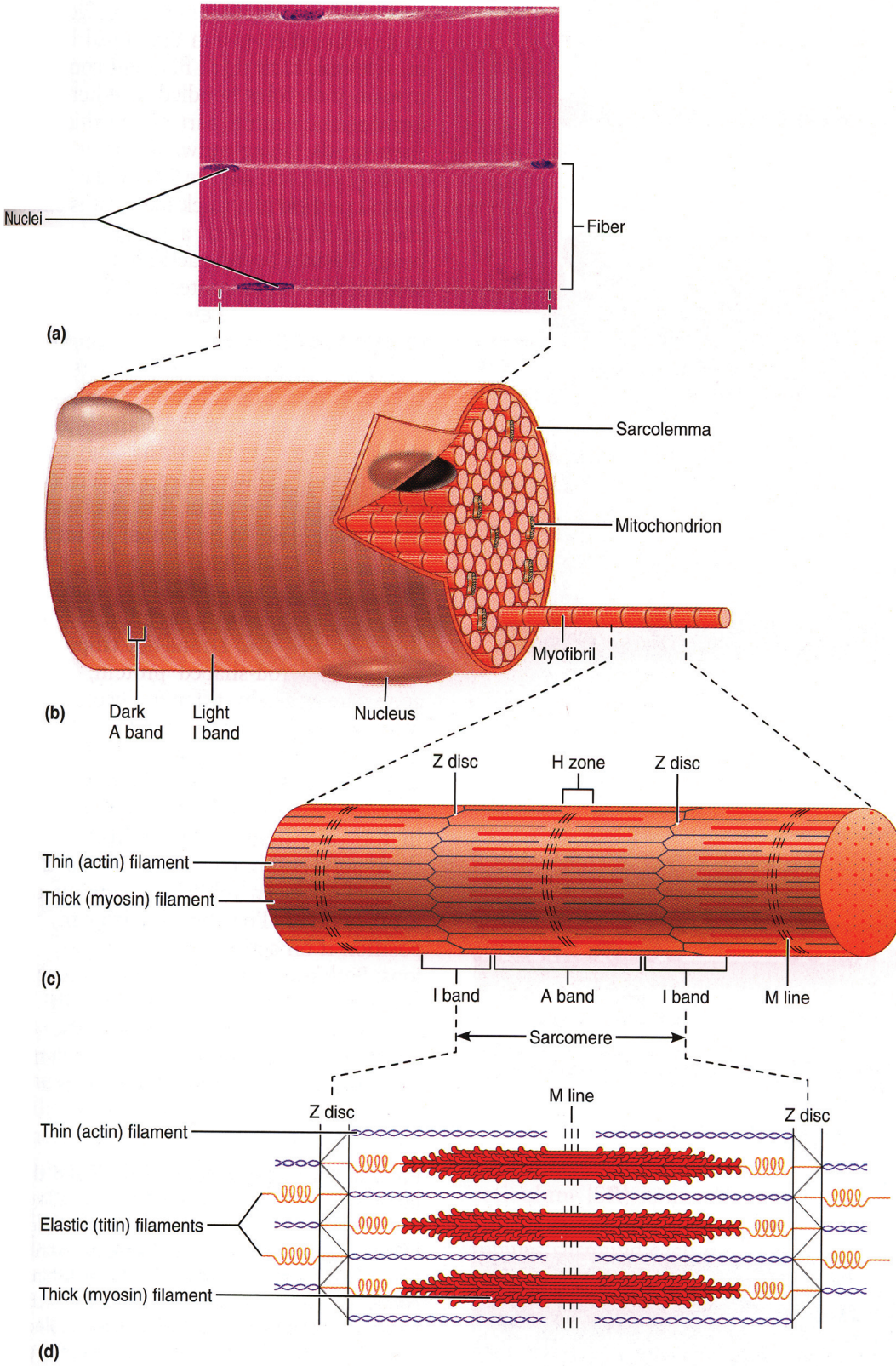
1.2.3.1 Sarcomere

The sarcomere is the basic contractile element in muscle. Successive sarcomeres, serially, form a myofibril, and multiple myofibrils, in parallel, constitute the contractile machinery of a myofiber (Figure 3). The sarcomere is a rodlike structure and contains specific structural elements. On the outsides of the sarcomere, a disc-like structure is present in a cross-sectional plane, which is called the Z-disk. The Z-disk contains a large number of structural molecules with a variety of functions. The major constituent of the Z-disk is the actin-binding α-actinin, which has a number of different functions. The periphery of the Z-disk contains intermediate filaments, which align the Z-disks of surrounding sarcomeres. Furthermore, these intermediate filaments bind the myofibrils indirectly to the sarcolemma. Actin filaments are positioned at a right angle to the Z-disk. Myosin filaments are positioned parallel to the actin filaments, and are set out from another plane situated between two successive Z-disks. This plane is called the M-line.

Figure 3 Myofiber composition

(a) Staining of a lateral section demonstrates the striated appearance of the myofiber. (b) The myofiber consists of a large number of myfibrils, which are the contractile elements. (c) Myofibrils are made of consecutively ordered sarcomeres. (d) Sarcomeres are flanked by Z-disks to which the thin (actin) filaments are bound. The thick (myosin) filaments are connected to the M-line. The striated appearance originates from the composition of the thin and thick filaments.

(Adapted from Human Anatomy & Physiology, 6th Edition, Elaine N. Marieb, Pearson Education, Inc.)


Nuclei
Fiber
(a)
Sarcolemma
Mitochondrion
Myofibril
(b)
Dark
A band
Light
I band
Nucleus
Z disc
H zone
Z disc
Thin (actin) filament
Thick (myosin) filament
(c)
I band
A band
I band
M line
Sarcomere
M line
Z disc
Z disc
Thin (actin) filament
Elastic (titin) filaments
Thick (myosin) filament
(d)

An action potential generated from the axon of a motor neuron results in releasing of the neurotransmitter acetylcholine (ACh) in the neuromuscular junction, which subsequently binds to the acetylcholine-receptor (AChR). Activation of the AChR leads to the opening of ion channels, allowing a flux of Na^+ and K^+. As a result, a change in membrane potential occurs, since more Na^+ flows in than K^+ flows out of the myofiber. The generated local current spreads throughout the sarcolemma, down the T tubules, and eventually results in the extracellular release of calcium from the sarcoplasmic reticulum via the terminal cisternae throughout the myofiber. A calcium-dependent interaction between actin- and myosin filaments results in a contraction between the two Z-disks of each sarcomere in the myofiber. The 'sliding filament' theory, regarding the interaction between actin and myosin, has been proposed by Huxley et al.[107,108], and will not be discussed in further detail. After contraction, the cytosolic Ca^{2+} is relocated in the sarcoplasmic reticulum by continuously active Ca^{2+} pumps, which leads to the relaxation of the sarcomere. Altogether, the generated force of the contraction is distributed from the Z-disks to the sarcolemma via a mesh of interacting proteins, called the costameres.

1.2.3.2 Costameres

Costameres are macromolecular protein structures localized between the sarcolemma and the sarcomeres, and function as positioned focal adhesion complexes to transmit the contractile force generated by the sarcomeres to the sarcolemma (Figure 4)[175]. Due to their position at both the M-line and Z-disk, and their indirect interaction with the extracellular matrix, the sarcomeres of individual myofibers are aligned[127]. This arrangement facilitates the uniform transmission of lateral forces between neighboring myofibers, which is necessary to maintain the consistency of the sarcolemmal lipid bilayer [76,61].
When approached from the sarcolemmal side, the costameres are attached to the cell membrane via two separate systems; the vinculin-talin-integrin system, and the dystrophin-glycoprotein-complex. A common property of the two systems is the binding of filamentous actin (F-actin) to a membrane-associated complex[5]. F-actin is part of the subsarcolemmal cytoskeleton. The intermediate filaments present in the costamere, such as desmin[98], synemin[98], and vimentin[82], are extensively interconnected to F-actin by the linker protein plectin[81,45]. Correspondingly, the intermediate filaments connect to the α-actinin network of the Z-disk, completing the link from the sarcolemma.

1.2.3.3 DGC

The dystrophin-glycoprotein complex (DGC) is positioned at the sarcolemma, where it functions as a bridging structure between the extracellular matrix and the intracellular cytoskeleton of the myofiber. The core of the DGC is formed by the structural trinity dystrophin-dystroglycan-laminin2 (Figure 5) (reviewed in Dalkilic et al.[58]).
Dystrophin is a 427 kD cytoskeletal protein, and belongs to the family of β-spectrin/α-actinin proteins[123]. Dystrophin has four distinct structural domains. The amino terminal domain has high homology to actin binding regions in proteins such as β-spectrin and α-actinin, and binds F-actin (reviewed in Rybakova et al.[187]). The rod domain consists of an array of 24 repeats, which form a nested helix, and has similarities to spectrin[56,116]. The cystein rich domain contains a number of binding sites to proteins belonging to the DGC. One of these domains is

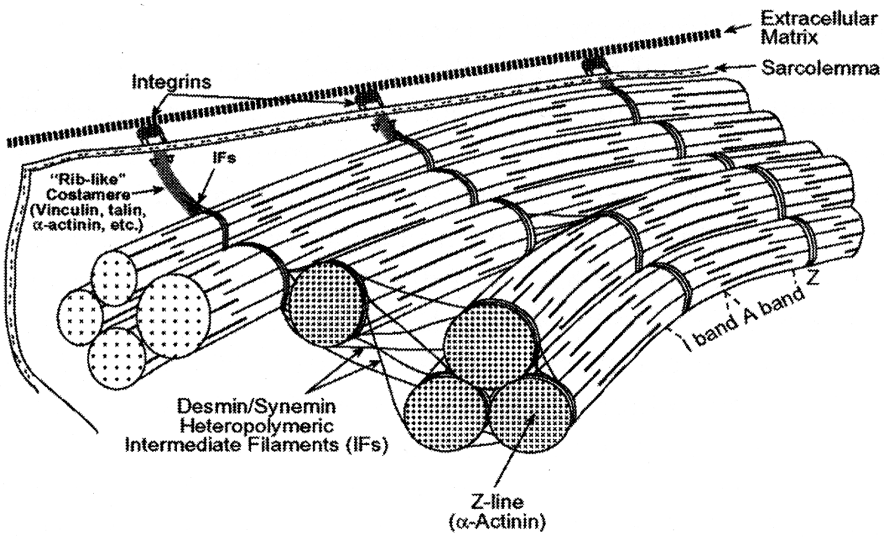


Figure 4 Costameres

The costameric protein network alligns the Z-disks of the individual myofibrils and connects them via the sarcolemma to the extracellular matrix. The DGC-related proteins are not shown.

(from Myology, 3rd edition, eds. Engel & Franzini-Armstrong, McGraw-Hill)

the WW domain, which recognizes the proline-containing motifs present in the PPXY motif of β-dystroglycan[105]. Furthermore, the binding to β-dystroglycan is mediated by a ZZ domain[114]. The carboxy terminal domain also contains binding sites for DGC-related proteins.
Dystroglycan consists of two subunits, which arise from a single gene-transcript[109,110]. This precursor protein is split into the two subunits in the endoplasmic reticulum[209,103,165]. α-Dystroglycan is localized at the extracellular side of the sarcolemma, is heavily glycosylated, and binds to laminin[243,130]. α- and β-dystroglycan are bound via multiple covalent bonds[201]. β-dystroglycan is a single-pass transmembrane protein with its carboxy terminal present at the cytosolic site of the sarcolemma. The COOH-terminal contains a PPXY motif to bind to dystrophin, and is capable to alternatively bind caveolin3 via its WW domain[213].

Laminin2 is a heterotrimeric extracellular matrix protein, which consists of three chains; α1, β1, and γ1. Laminin2 binds to collagen IV, which is the major constituent of the basement membrane. The interaction between laminin2 and collagenIV is stabilized by other proteins such that a scaffold is formed. Examples of these stabilizing proteins are perlecan, biglycan, and nidogen.

The trinary protein core of the DGC is stabilized by the transmembrane sarcoglycan-sarcospan complex (SGC), which contains four glycosylated subunits (α-, β-, γ-, and δ- sarcoglycan) and sarcospan[246,54]. The sarcoglycan subunits are all single-pass transmembrane proteins with a molecular weight of 50 kD, 43 kD, 35 kD, and 35 kD, respectively[92]. α-Sarcoglycan is a type I transmembrane protein with the aminoterminus on the extracellular side, whereas the other

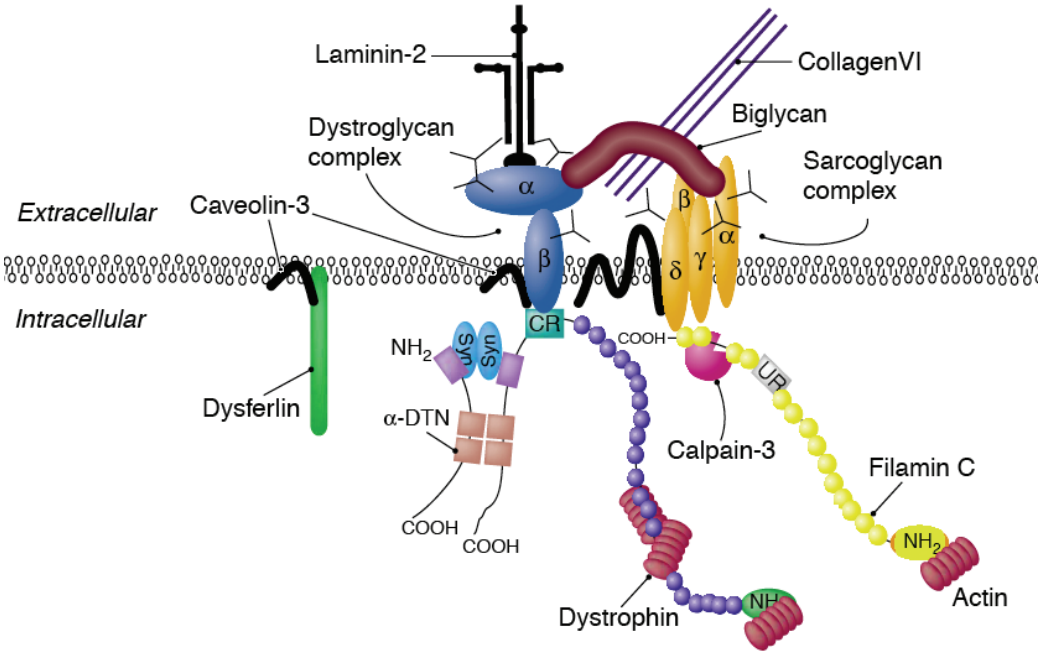


Figure 5 The dystrophin-glycoprotein complex

The core of the dystrophin-glycoprotein complex (DGC) is formed by dystrophin, dystroglycan, and laminin. The trio is stabilized by the sarcoglycan-sarcospan complex (SGC).

(from Dalkilic et al., Curr Opin Genet Dev (2003) ***13****(3) 231-238)*

sarcoglycans are type II transmembrane proteins with their aminoterminus at the cytosolic side of the sarcolemma[226]. The sarcoglycan subunits are tightly associated[90,40]. The sequence homology is highest between the γ-, and δ-sarcoglycan subunit[115].

1.2.3.4 The integrin-vinculin-talin system

The integrin-vinculin-talin system is a second protein complex involved in the binding of the costameres to the sarcolemma, and providing a linkage to the extracellular matrix[61]. Cytoskeletal F-actin binds to a cluster of focal adhesion proteins, in which talin plays a central role. Vinculin, focal adhesion kinase, actin, and integrins bind to talin (reviewed in Nayal et al.[162]). The complex is indirectly linked to the extracellular matrix via a transmembrane integrin-dimer[5].

1.2.3.5 Extracellular matrix

Each myofiber, fascicle, and muscle is surrounded by a layer of connective tissue, which consists of a protein- and carbohydrate-rich extracellular matrix, fibroblasts, macrophages, capillaries, and nerve branches. This layer contributes to the mechanical properties of muscle, and plays a role in myogenesis and regeneration. The surrounding connective tissue has a distinct

structure, since a number of layers with specific characteristics can be determined. The two most important layers are the basal lamina and the reticular lamina, which together form the basement membrane.
The basal lamina is closest to the sarcolemma, and consists of the electron-lucent lamina rara which is then covered by the electron-dense lamina densa[112]. The basal lamina contains predominantly collagen IV and laminin[191,46]. Collagen IV and laminin have the ability to self-assemble, and their networks are linked by the glycoprotein entacting/nidogen[223]. These components form a structure to which a large number of different proteins can be linked, such as other proteoglycans, components of the reticular lamina (collagen VI), and transmembrane components (integrins, dystroglycan)[191]. The reticular lamina is a dense fibrillar structure containing predominantly collagen and other fibrils.

1.2.3.6 Transverse fixation system

The contraction and stretching of skeletal muscle generates a force that is distributed across the muscle to the tendons, thereby providing the ability to move. The simultaneous contraction of the myofibers is closely orchestrated at several levels. Muscle contraction is initiated by a neuronal action potential leading to depolarization of the sarcolemma. The depolarization continues via the transverse tubules to the inner part of each fiber, which results in a uniform contraction throughout the myofiber. The myofibrils within a single myofiber are aligned along the Z-disks of the sarcomeres via an intricate system of intermediate filaments. The aligned myofibrils are linked to the sarcolemma via the costameric network, and are subsequently linked to the extracellular matrix. This transverse fixation systems facilitates the lateral transmission of force within a myofiber, and maintains the sarcolemmal integrity. Proteins functioning in the transverse fixation system often lead to muscular dystrophies or other myopathies when defective (Table 1).

1.3 Muscular Dystrophies

1.3.1 General Introduction

Muscular dystrophies are characterized by progressive irreversible degeneration processes, which results in weakness and wasting of muscle tissue. The muscular dystrophies share clinical symptoms, but differ largely in severity, age of onset, and distribution of affected muscle tissue[71]. To date, the mechanisms responsible for this divergence in pathology have not yet been identified in detail. The genetic defects causative for the majority of muscular dystrophies have been largely elucidated (Table 1).

1.3.2 Dystrophinopathies

Duchenne muscular dystrophy (DMD) and Becker muscular dystrophy (BMD) are caused by defects in the DMD gene, which encodes the subsarcolemmal protein dystrophin[101]. Since the DMD gene is located on the X-chromosome, primarily males are affected. Generally, DMD is caused by mutations that disturb the genetic reading frame, whereas BMD is caused by mutations that leave the reading frame intact[124,99,17,157]. DMD is the most common form of muscular dystrophy which occurs during childhood. The incidence is approximately 1 in 3000[70].

Table 1: Myopathies caused by costameric protein deficiencies

Myopathy	*Type*	*Mode of inheritance*	*Gene Product*
Duchenne and Becker MD	DMD / BMD	XR	Dystrophin
Limb-Girdle MD	LGMD1A	AD	Myotilin
	LGMD1B	AD	Lamin A/C
	LGMD1C	AD	Caveolin-3
	LGMD2A	AR	Calpain-3
	LGMD2B	AR	Dysferlin
	LGMD2C	AR	gamma-Sarcoglycan
	LGMD2D	AR	alpha-Sarcoglycan
	LGMD2E	AR	beta-Sarcoglycan
	LGMD2F	AR	delta-Sarcoglycan
	LGMD2G	AR	Telethonin
	LGMD2H	AR	TRIM32
	LGMD2I	AR	FKRP
	LGMD2J	AR	Titin
	LGMD2K	AR	POMT1
Congenital MD	MDC1A	AR	Laminin-2
	Itga7-deficiency	AR	Integrin Alfa7
	Ullrich Syndrome	AR	Collagen VI
Distal MD	Miyoshi Myopathy	AR	Dysferlin
Other	Bethlem Myopathy	AD	Collagen VI
	Epidermolysis Bullosa Simplex	AR	Plectin
	Myofibrillar Myopathy	AD/AR	Desmin
	Nemaline Myopathy	AR	Nebulin
	Nemaline Myopathy	AD/AR	alpha-Actin

MD = Muscular Dystrophy
XR = X-linked Recessive
AD = Autosomal Dominant
AR = Autosomal Recessive

Clinical phenotype

DMD was first described by Meryon (1852), and Duchenne (1868)[151,66]. The first symptoms of DMD are characterized by frequent falling, difficulty of getting up from a standing or lying posistion (demonstrating Gower's manoeuvre), and a waddling gait. Furthermore, the calf musculature is significantly enlarged by hypertrophy, fat infiltration, and accumulation of connective tissue. Approximately 20% of the patients has a mental impairment, which translates in an IQ of less than 70. Affected skeletal musculature is mainly proximal, and results in wheelchair dependence during the early teens. Patients often die of cardiac failure as an adolescent. Furthermore, infections, leading to pneumonia as a result of lack of ventilation, are often the cause of death, while respiratory care by assisted ventilation may increase the survival[208]. A major characteristic of DMD and other muscular dystrophies is displayed by muscle necrosis, which results in high serum levels of the isoenzyme creatine kinase (CK). CK 'leaks' from affected degenerating myofibers to the bloodstream in addition to other enzymes such as pyruvate kinase[211].

BMD was first described by Becker (1953) and Walton (1955)[16,237]. In general, BMD can be seen as a milder form of DMD with a later age of onset, loss of ambulation at an older age, and a later age of death (reviewed in Engel et al.[73]). However, the clinical phenotypes vary distinctly, and may in some cases be indistinguishable from DMD[163,33].

Pathological cause and effect

The widely accepted pathological model for DMD/BMD postulates that the absence (DMD) or instability (BMD) of the DGC results in membrane instability, making the sarcolemma vulnerable to rupture[156,196,37,239]. Plasma membrane defects are early and basic pathologic alterations, and are represented by lesions of various size[156]. The regions near small lesions contain dilated endocytotic vesicles, whereas large lesions harbour dilated SR vesicles, irregularly positioned sarcotubular components, degenerating mitochondria, and small clusters of glycogen[73]. As a result, extracellular calcium can enter the myofiber and cause an imbalance of the calcium homeostasis, which results in myofiber degeneration. Controversial evidence questions that an imbalance of calcium homeostasis leads to degeneration. It might be that degeneration preludes the extracellular calcium influx[85]. Hence, the degeneration leading to myofiber necrosis would be initiated via other mechanisms. The detection of apoptotic nuclei demonstrating DNA fragmentation in muscle biopsies from DMD patients led to another hypothesis[190,222,140]; absence of the DGC leads to sarcolemmal rupture and subsequent leaking of intracellular proteins, which are unknown to the immune system. This attracts and activates cytotoxic lymphocytes and helper T-cells, which initiate the apoptotic program in ruptured myofibers[214].

Myofiber necrosis

Dystrophin deficiency leads to irreversible myofiber cell death, or necrosis. The first necrotic myofibers can be demonstrated in the neonatal period, and are generally single[27]. With increasing age of the patient, necrotic myofibers appear in groups of 2-15 myofibers. Myofiber

necrosis is often segmental, which means that not the entire myofiber is affected by necrosis, but only part of it. The basal lamina surrounding the necrotic myofiber is not affected, remaining like an empty shell[26].

The early stage of necrosis is characterized by partially lysed and highly contracted myofibrils. Later stages demonstrate granular and filamentous debris and clumps of degenerating membranous organelles. Myofibrillar structures, as well as sarcotubular components can not be recognized[72]. Furthermore, the affected myofibers contain membrane attack complexes (MAC), which are the result of complement activation. The insertion of MAC in the sarcolemma creates holes, and causes cell-lysis.

1.3.3 Limb-Girdle Muscular Dystrophies

1.3.3.1 Sarcoglycanopathies

A distinct group of muscular dystrophies is classified as the Limb-Girdle Muscular Dystrophies (LGMD). Diagnosis of these muscular dystrophies is often difficult, because the different forms demonstrate a high heterogeneity within and between the diseases[34]. As a result, LGMD requires a more complex classification using clinical appearance, protein analysis, and genetic studies[35]. Both autosomal dominant (LGMD1) and autosomal recessive (LGMD2) forms exist (Table 1). Although the monogenetic causes for most of these disorders have been elucidated, more LGMD are found due to a higer level of specification in clinical determination, and subsequent finding of defective genes. LGMD differ from DMD/BMD in severity, age of onset, and distribution of affected muscle tissue.

A particular group within the autosomal recessive LGMDs is formed by the sarcoglycanopathies. Sarcoglycanopathies are characterized by deficiency of one of the sarcoglycan proteins (α-, β-, γ-, δ-sarcoglycan). These four isoforms form together with sarcospan the sarcoglycan-sarcospan protein complex (SGC). The transmembrane SGC stabilizes the dystrophin-dystroglycan-laminin bolt, which forms the core of the DGC. Deficiency of one of the sarcoglycans inhibits the proper formation of the SGC, and can lead to its marked reduction in the sarcolemma[102,165,206]. The loss of the entire complex is commonly seen in β- and δ-sarcoglycan deficiency. However, α- and γ-sarcoglycan deficiency show a more restricted loss of the SGC[227]. Rare cases of mutations in the γ-sarcoglycan gene have been reported that leave the SGC intact, but with probable dysfunction of the SGC, which therefore leads to muscular dystrophy[23,235]. In conclusion, defective expression (or absence) of the entire SGC might explain the large similarity in the pathological phenotypes of the sarcoglycanopathies[173].

Recently, two other sarcoglycans (ε- and ζ-sarcoglycan) have recently been reported[77,240]. ε-Sarcoglycan deficiency results in myoclonus-dystronia, which is a movement disorder with involuntary jerks and dystonic contractions[64]. Thus far, ζ-sarcoglycan has not been found to relate to a human disease. These two sarcoglycans will not be discussed in further detail.

Clinical phenotype

In sarcoglycanopathies, the proximal limb-girdle musculature is mainly affected. The age of onset is highly variable, but with a bias toward childhood around the age of 6 to 8 years[24]. A characteristic clinical feature is the presence of scapular winging due to affected periscapular musculature. The diseases are progressive by nature, but progression is again highly variable. Early age of onset does not necessarily indicates a rapid progression of the disease. The patients eventually present a loss of ambulation resulting in wheelchair dependence. As with DMD, sarcoglycan deficiency can lead to cardiomyopathy. This clinical feature is more common in β-, and δ-sarcoglycan deficiency[14].

Pathological cause and effect

On a histopathological level, a large similarity between the sarcoglycanopathies and dystrophinopathies is demonstrated, since muscle degeneration is demonstrated in both diseases. Loss of the SGC leads to a moderate reduction of dystrophin, and of the dystrophin-dystroglycan-laminin bolt subsequently[228]. Furthermore, physiological studies showed that sarcoglycan deficiency is not accompanied by aberrant force generation, thereby demonstrating that the remaining DGC is functional[90]. Thus, the mechanical role of the DGC, which is considered to provide stabilization of force transmission over the sarcolemma, does not have to be affected in sarcoglycan deficiency. Indeed, it has been argued that, since sarcoglycan deficiency does lead to muscular dystrophy, alternative or additional mechanisms are required to explain the mechanical defects in DMD[91]. Notably, dystrophin deficiency results in the absence of the entire DGC, including the SGC. The absence of the SGC can therefore contribute as secondary mechanism to the pathology of dystrophin deficiency.

The subunits of the sarcoglycan complex, and especially γ-sarcoglycan, have cystein-rich domains which demonstrate a homology to EGF-like cystein-rich domains. Since the EGF-like domain is located at the extracellular side of the sarcolemma, ligand-binding and signalling properties of the SGC are not excluded[144]. The possibility that the DGC has signalling function next to its mechanical function was further highlighted by demonstrating analogy to other signalling complexes[182].

1.3.3.2 Dysferlinopathies

Mutations leading to dysferlin deficiency cause the autosomal recessive Limb-Girdle muscular dystrophy (LGMD2B) or Miyoshi Myopathy (MM)[15,132]. Dysferlin is a transmembrane protein with a distibution similar to that of dystrophin[6]. Dysferlin deficiency does not lead to altered expression of the DGC[6].

Dysferlin shows a significant homology to FER-1, a Caenorhabditis elegans gene that plays a role in vesicle fusion to the plasma membrane[3]. In normal muscle, sarcolemmal injury is followed by the accumulation of vesicles at the ruptured site, which subsequently fuse with eachother and the ruptured plasma membrane to close the gap, leaving a ‘patch’ [146,147]. Dysferlin plays a critical role in the efficient docking and fusion of vesicles to the plasmamembrane in

a Ca^{2+}-dependent manner, and therefore in the repair process of membrane rupture[12]. Dysferlin deficiency leads to accumulation of subsarcolemmal vesicles at the site of rupture, which do not fuse with the membrane, leaving the ruptured site unrepaired[12]. As a consequence, the affected muscle fiber is prone to degeneration, characteristic of muscular dystrophy, but follows a different etiological pathway.

Clinical phenotype

The age of onset in dysferlinopathy is highly variable, but shows a slight tendency towards the age of 20 years[131,135,7]. A characteristic feature of the slowly progressive disease is the absence of symptoms prior to the onset. Muscular weakness in the LGMD variant usually starts in the pelvifemoral region with particular involvement of the quadriceps[10]. Scapular winging as seen in sarcoglycanopathy is not demonstrated. Miyoshi Myopathy shows a more distal presentation. The gastrocnemius and the soleus muscle are mainly affected during the onset of the disease. As a result, the patients are not able to walk on their toes[155,131]. Highly elevated CK levels mark the onset and the active phase of the disease. Cardiac involvement is not seen in dysferlinopathies[24]. This is consistent with the monocellular nature of heart muscle, in which the rupture/repair process is unlikely to play a role.

Pathological cause and effect

Rupture of the sarcolemma is common in normal muscle tissue. Small lesions are inflicted by exercise, growth, etc. Inability of membrane repair after rupture leads to a disturbance of the myofiber homeostasis. Dysferlin deficiency results in defects of the membrane repair system[13]. Ultrastructural analysis of dysferlin deficient muscle demonstrates a thickened basal lamina over –probably unrepairable- membrane lesions, as well as small vacuolar proliferations and degenerative papillary projections[202].

1.3.3.4 Autosomal dominant limb-girdle muscular dystrophies

LGMD1A is an autosomal dominant (AD) muscular dystrophy caused by genetic mutations in the myotilin gene[94]. Myotilin is localized to the Z-disk, where it presumably binds to filamin C[230]. Myotilin deficiency is characterized by an adult onset of proximal muscular weakness, which starts at the hip-region and progresses to the shoulder region[84].

Genetic mutations in caveolin 3 (Cav3) cause AD-LGMD1C[153]. Cav3 is the muscle specific isoform of the caveolin protein family[219]. Cav3 is localized at the sarcolemma, where it can interact with a large number of proteins, like DGC-components[145,213,55], dysferlin[139], neuronal nitric oxide synthase (nNOS)[232], and phosphofructokinase[212]. Currently, it is postulated that Cav3 functions as a molecular facilitator, bringing receptors and second messengers in close proximity to facilitate the assembly of signaling complexes[24]. Cav3 deficiency leads to moderate proximal muscle weakness with an average age of onset of 5 years. Furthermore, the patients showed calf hypertrophy, Gower's sign at adult life, and elevated serum CK-levels[153].

1.3.3.5 Autosomal recessive limb-girdle muscular dystrophies

Genetic mutations in the calcium-activated proteolytic enzyme Calpain 3 (Capn3) lead to autosomal recessive (AR) LGMD2A[184]. LGMD2A patients show an initial distribution of scapular-humeral-pelvic distribution of muscle weakness, which is sometimes followed by weakness of the distal muscles in the lower extremities[78]. Capn3 functions as an indirect regulator of anti-apoptotic processes, which might explain the apoptotic cell death of myofibers in patients with calpainopathy[11]. Furthermore, Capn3 might interfere with the interaction between filamin C and γ-, and δ-sarcoglycan, although this mechanism is not yet fully understood[88].

AR-LGMD2G is caused by genetic mutations in the telethonin gene, which translates into a sarcomeric protein localized at the Z-disk[160]. Telethonin is also known as titin-cap (TCAP), since one of its functions is to cap titin[161]. Furthermore, telethonin binds to myozenin, which interacts with α-actinin, calsarcin, and filamin C[218,79,80]. Likely, telethonin plays a role in the maintenance of the integrity of the sarcomere, as well as during its assembly[138]. The clinical presentation of telethonin deficiency is not precise, since only a limited number of families are described[160,159,247]. The muscular weakness in the studied families is predominantly proximal with an age of onset from 9-15 years. A number of patients displayed loss of ambulation[24].

1.3.3.6 Congenital muscular dystrophy

Laminin α2 deficiency is the cause of the autosomal recessive congenital muscular dystrophy 1A (MDC1A)[97]. Laminin α2 (Lama2) is an extracellular glycoprotein that forms a link between the basal lamina and the sarcolemma by binding to dystroglycan. Lama2 deficiency leads to progressive muscle degeneration (reviewed in Voit et al.[234]). Shortly after birth, patients show profound hypotonia, which are frequently accompanied by contractures. The weakness affects the facial, proximal, and distal musculature.
Ullrich Syndrome is classified as a congenital muscular dystrophy, and is a result of collagen VI deficiency[113,150]. The clinical phenotype is characterized by generalized muscle weakness, spine rigidity, and respiratory insufficiency. Collagen VI is present in microfibrillar structures in many tissues. Collagen VI might play a role in cell migration, differentiation, and embryonic development[74].

1.3.4 Animal models for muscular dystrophy

Animal models for human disease have greatly facilitated the study of human disease throughout the last decades. One of the largest achievements has been the ability to generate transgenic mice. This applies as well for the study of muscular dystrophies. It has been more than 20 years ago, that the first mouse model for muscular dystrophy (mdx) was discovered[32]. This mouse model occurred due to a spontaneous mutation; the technique to generate transgenic mice was not available yet. Since the mouse has been the major organism for studying muscular dystrophies, the different mouse models will be discussed below. Furthermore, a number of alternative animal models will be discusses as well.

1.3.4.1 Mouse models for dystrophinopathy

The mdx mouse model for DMD contains a point mutation in exon 23 of the murine homologue of the DMD gene resulting in a premature stop codon[207]. This mutation leads to the absence of dystrophin at the sarcolemma[22]. Similar to DMD, absence of dystrophin leads to a great reduction of the dystrophin-glycoprotein complex[168]. The histopathology of the mdx mouse is well described. The first clue of a mutant phenotype was obtained as a result of elevated levels of the muscle isoenzyme pyruvate kinase. Following histological characterization of muscle tissue, a temporal myopathy was revealed, which was characterized by excessive atrophy with loss of muscle fibers. Furthermore, a variation in fiber size, degeneration of fibers, and marked concentration of densely stained, proliferating, sarcolemmal nuclei with phagocytic cells in place of lost fibers was found[32].

Although the mdx mouse was presented at the moment of the discovery as a potential model for DMD, a remarkable feature doubted its authenticity[59]. After a period of extensive degeneration, the mdx mouse shows the ability to recover; a feature not seen in the lethal human DMD pathology. The elucidation of the DMD gene in both human and mice, however, confirmed the involvement of a homologous gene[30,100]. This led to a change in the view on dystrophin-deficiency. What mechanisms does the mdx mouse apply to circumvent the deficiency of dystrophin, and to generate functional muscle tissue? The first studies concentrated on histological examination of the pathological stages of the mdx mouse.

The pathology of the mdx mouse commences at approximately 2-4 weeks of age, when a widespread necrosis of myofibers occurs. This age corresponds with an increase of mechanical demands due to growth. The lack of dystrophin is likely to become critical at this stage[57]. As a result, the plasma membrane of myofibers becomes instable and has an elevated tendency to membrane rupture[186]. Due to a high extracellular concentration, an influx of of free Ca^{2+} ions disturbes the intracellular calcium homeostasis, which eventually leads to myofiber necrosis (for details see Chapter 1.3.6.1). Muscle fiber necrosis induces an inflammatory response, which is characterized by the influx of inflammatory cells at the periphery of necrotic regions. Necrotic myofibers are cleared via phagocytosis by macrophages.

As stated above, the mdx mouse has a successful regenerative response to the degeneration resulting from dystrophin deficiency. Muscle regeneration is hallmarked by the activation of normally quiescent satellite cells, their proliferation, fusion, and eventual replacement of affected myofibers. Regeneration is dependent on the successful completion of these separate processes. Regeneration is a consequence of degeneration, and commences as soon as degeneration occurs. In the mdx mouse, a temporal shift is displayed from degeneration towards regeneration. During this phase both processes occur, but in different proportions. From 2-4 weeks of age, degeneration prevails over regeneration. From approximately 8 weeks of age, regenerative processes overcome degeneration, which results in a stabilized condition at around 12 weeks of age. The eventual prevailing of regeneration might be due to a decrease in degenerative processes. However, a number of alternative hypotheses have been proposed. The replicative potential of murine satellite cells may exceed those of humans, although this difference as well as a change in the number of satellite cells during aging is still under de-

bate[167]. Furthermore, murine satellite cells may act more responsive to external signalling such as growth factors. It may also be possible, that the small caliber of murine muscle fibers makes them less susceptible to mechanical stress[117,118]. Fibertype alterations can also lead to a decrease in the probability of sarcolemmal rupture, and thereby the severity of the disease[179,136].

Alternative mouse models for dystrophinopathy

Since the original mdx mouse did not show progressive muscle weakness and wasting, alternative mouse models for muscular dystrophy were generated using chemical mutagenesis with N-ethylnitrosourea (ENU)[41]. Screening for variant alleles in the dystrophin locus combined with elevated levels of muscle enzymes in the blood, resulted in additional mouse models which were named mdx$^{2\text{-}5cv}$. The mutations of these mouse models were determined, and found to be point mutations in the Dmd gene leading to aberrant, non-functional transcripts[52,111]. These alternative mouse models displayed a dystrophic phenotype similar to the original mdx mouse[41,60,52]. Furthermore, two transgenic mice were generated by targeting exon 52 (mdx52)[9], and the first exon of the Dp71 isoform (Dp71-/-)[193]. The mdx52 showed a similar phenotype compared with the mdx mouse. On the other hand, the Dp71-/- mouse model did not show any dystrophic characteristics[193], which can be explained by the fact that the targeted isoform is not expressed in muscle[197]. In conclusion, alternative mouse models for dystrophinopathy, which affect several or all dystrophin isoforms, do not seem to differ in dystrophic phenotype (except for Dp71-/-). This feature was recently confirmed by another mouse model for dystrophinopathy. Kudoh et al. generated a mouse model for DMD by deleting the entire Dmd gene using Cre-lox-P recombination system[126]. The pathology of these mdx-null mice was virtually identical to the mdx mouse in both skeletal muscle and diaphragm.

Utrophin-deficient mice

A genetic sequence with a high homology to dystrophin was found to be ubiquitously expressed, particularly in several fetal tissues such as heart, placenta and intestine[133,134]. The 80 kDa protein product of the 4.8 kbp transcript was called utrophin[21]. Utrophin forms a protein complex with similar proteins as found in the DGC and is located at the neuromuscular junction (NMJ)[141]. Utrophin has a high similarity to dystrophin based on the primary structure[224].

In adult mdx mice, utrophin is localized across the sarcolemma, a phenomenon not seen in DMD patients. This might explain the milder phenotype of the mdx mouse; utrophin might replace the missing dystrophin. Utrophin-deficient mice do not show a severe phenotype[62]. However, mice lacking both utrophin and dystrophin have a very severe muscular dystrophy, indicating the complementary role of utrophin in the mdx mouse[62]. Furthermore, the likely complementary role of utrophin in the mouse might explain the severe phenotype of genetic defects in other components of the DGC compared to dystrophin-deficiency. These findings resulted in using utrophin as a target for therapeutic approaches (reviewed in[177,122]).

1.3.4.2 Mouse models for sarcoglycanopathy

Defective expression of any single sarcoglycan results in muscular dystrophy. To study this

group of diseases in more detail, mouse models have been generated, which are transgenic for one of the four sarcoglycan subunits.

α-*Sarcoglycan deficient mice*

Duclos et al. generated transgenic mice for α-sarcoglycan (Sgca) as a model for LGMD2D[67]. Both Sgca alleles were affected, resulting in a null-model. Sgca is only expressed in striated muscle tissue. The Sgca-null mice developed progressive muscular dystrophy which increased with age. Although the Sgca-null mice did not show any overt signs of muscular dystrophy, the first features of muscular dystrophy could be determined on histological level shortly after birth. Histological stainings demonstrated ongoing degeneration of skeletal muscle tissue (diapragm) with age. Typical features for human muscular dystrophy were detected, such as necrosis, inflammation, centrally located nuclei, fibrosis, atrophy, hypertrophy, fiber splitting, and dystrophic calcification[67].

Immunohistochemical analysis of both cardiac and skeletal muscle tissue of Sgca-null mice showed absence of Sgca, and a marked reduction of the other (β, γ, and δ) sarcoglycan subunits. Furthermore, immunohistochemical staining of dystrophin was patchy and reduced. Detection of α-dystroglycan by western blot demonstrated normal levels, although it was not tightly associated with the sarcolemma. These results corroborate the necessity of the SGC for the formation and stabilization of the DGC.

β-*Sarcoglycan deficient mice*

Two groups independently generated transgenic mice for β-sarcoglycan (Sgcb) as a model for LGMD2E[8,68]. Araishi et al. targeted exon 2 of Sgcb to disrupt the reading frame, whereas Durbeej et al. targeted exon 3-6. Both mouse models developed a severe, progressive muscular dystrophy.

Araishi et al. detected the first symptoms of muscular dystrophy in the β-sarcoglycan deficient mouse from 2 weeks of age. Histological analysis showed muscle degeneration and infiltration of mononuclear cells. The degenerative changes were most prominent between 4 and 8 weeks of age. At 8 weeks of age, an increase in the mass of connective tissue was seen. At 14 weeks of age the regenerative changes were predominant, and were associated with hypertrophy. At 20 weeks almost all muscle fibers (>95%) in the quadriceps femoris muscle showed centrally located nuclei, as well as variability in fiber size. The heart musculature was affected at 56 weeks of age, showing fibrotic patches on the heart wall. Cardiac pathology could not be detected at 6 weeks of age. Arashi et al. found no differences in SGC gene expression between Sgcb deficient mice and healthy wildtype mice. However, SGC protein expression could be scarcely detected in affected skeletal muscle tissue. Other components of the DGC than SGC components were expressed and could be detected in skeletal muscle tissue.

Durbeej et al. demonstrated that their Sgcb-null mouse presented similar dystrophic symptoms. Furthermore, additional dystrophic features like fiber splitting, extensive dystrophic calcification, endomysial fibrosis, and massive fatty infiltration were detected. The pathology

of the Sgcb-null mouse is more severe compared to mice lacking α-sarcoglycan, since larger areas of necrosis and fatty infiltration were detected. Durbeej et al. found small necrotic areas in the heart of Sgcb-null mice at 9 weeks of age. At 20 weeks of age, ischemic-like regions were detected. Further investigation showed that Sgcb-null mice lack expression of Sgcb in smooth muscle tissue. As a result, perturbations in smooth muscle tissue due to this lack lead to vascular irregularities. In the heart vasculature, these perturbations led to constrictions associated with pre- en poststenotic aneurisms, and preceeded the onset of ischemic-like lesions. Durbeej et al. concluded at that time that the effects of muscular dystrophy in striated muscle are exaggerated by vascular irregularities due to smooth muscle perturbations.

γ-*Sarcoglycan deficient mice*

Hack et al. generated transgenic mice deficient in γ-sarcoglycan (Sgcg), and found a mouse model for LGMD2C[90]. Exon 2 was targeted to disrupt the reading frame of Sgcg. The Sgcg-null mice developed a muscular dystrophy, which preferentially affected the proximal muscles. Sgcg-null mice displayed a stunted growth, an abnormal gait, and a relative slowness and inactivity. Furthermore, 50% of the Sgcg-null mice died before the age of 20 weeks. The Sgcg-null mice showed typical muscular dystrophy symptoms, such as necrosis, infiltration of inflammatory cells, centrally located nuclei, calcification, replacement of muscle by adipose and connective tissue, and regeneration. Sgcg-null mice developed cardiomyopathy, which was detected at 20 weeks of age. The affected heart showed ventricular fibrosis, and an increased right and left ventricular wall. Immunohistochemical staining showed the absence of SGC components in the sarcolemma of skeletal muscle. Sgcg-null mice demonstrated a muscular dystrophy, which is most progressive from 3-8 weeks of age.

Sasaoka et al. generated an alternative Sgcg deficient mouse model for LGMD2C, by targeting exon 3 of the Sgcg gene[194]. This model differed from the one Hack et al. had generated, by showing high survival rates (>1 year), and remarkable muscle hypertrophy and muscle weakness after 12 weeks of age. The observed hypertrophy is, according to Sasaoka et al., due to an increase in the number of regenerating myofibers and not due to fat infiltration or proliferation of connective tissue.

δ-*Sarcoglycan deficient mice*

A mouse model for LGMD2F was generated by Coral-Vazquez et al. by targeting exon 2 of δ-sarcoglycan (Sgcd), thereby disrupting the reading frame[49]. The first dystrophic features were detected at 2 weeks of age, and pathological alterations increased with age, showing a progressive course of the disease. Similar to the other mouse models for sarcoglycanopathy, Sgcd-null mice demonstrated necrosis, infiltration of inflammatory cells etc. Furthermore, the Sgcd-null mice developed cardiomyopathy and vascular irregularities similar to those of Sgcb-null mice. Coral-Vazquez et al. postulated the hypothesis that disruption of the SGC in vascular smooth muscle perturbs vascular function and induces ischemic-like lesions in the heart and exaggerates the dystrophic phenotype (see β-Sarcoglycan deficient mice).

1.3.4.3 Mouse models for dysferlinopathy

The first report of dysferlin deficiency in the mouse described a 171 bp in frame deletion in the dysferlin gene of the SJL wild type strain, resulting in a highly unstable molecule, which subsequently leads to a low abundance of the mutant protein[19]. The SJL mouse shows a progressive muscular dystrophy, which increases in severity at later stages of life. The first symptoms appear after approximately 3-4 weeks, and are marked by small myopathic lesions like small-calibre myofibers with centrally located nuclei. At 10 months of age skeletal muscle demonstrates typical dystrophic characteristics, like degeneration, inflammation and regeneration. At 15 months of age the pathology shows an advanced muscular dystrophy with fibrosis and fat infiltration in skeletal muscle tissue[19].

A transgenic mouse model for dysferlinopathy was generated by targeting the last three exons of dysferlin, which resulted in the complete loss of protein expression[12]. This dysferlin-null mouse also developed a slowly progressing muscular dystrophy with the first symptoms at 2 months of age. At 8 months of age the dysferlin-null mouse displayed the characteristics of muscular dystrophy, showing degeneration, inflammation, regeneration, fat infiltration, and fibrosis[12].

1.3.4.4 Other animal models for muscular dystrophy

Since mouse models for muscular dystrophy have been primarily used to study the pathological mechanisms as described in this thesis, these models have been described in detail. However, a number of other animal models have been used to study the muscular dystrophies, and will be discussed briefly. Canine and feline models for muscular dystrophy were recently reviewed in detail by Shelton and Engvall[205]. Compared to rodent animals for muscular dystrophy, large animal models provide a higher clinical relevance to study the research and testing of therapies due to a higher similarity towards humans[104].

Canine models for muscular dystrophy

Clinical observations of a myopathy in the Golden Retriever led to the finding of a canine model for muscular dystrophy[149,125,229]. The pathology of the canine model, also known as Golden Retriever Muscular Dystrophy (GRMD), is caused by a genetic mutation leading in the DMD gene, which leads to a lack in dystrophin expression[48]. The GRMD dog displays the characteristic clinical pathology of muscular dystrophy; progressive muscular weakness and wasting, muscle hypertrophy, elevated CK levels, and degeneration and regeneration cycles of affected muscle fibers. A number of other canine models for muscular dystrophy have been found recently. Next to canine models for dystrophinopathy, canine models for sarcoglycanopathy have also been found (reviewed in Shelton et al.[205]).

Feline models for muscular dystrophy

A number of muscular dystrophy-like disorders have been reported in cats[236,36,83]. The clinical phenotype is characterized by generalized skeletal muscle hypertrophy, excessive salivation,

reduced exercise tolerance, stiff gait and bunny-hopping when running, difficulty in jumping, vomiting/regurgitation, cardiomyopathy (reviewed in Vite[233]). The pathological manifestations are characterized by myofiber degeneration and regeneration, myofiber splitting, calcific deposists, and centrally located nuclei[205].

Rodent models for muscular dystrophy

The BIO14.6 hamster is an animal model for LGMD2F, since it has a genetic mutation in the δ-sarcoglycan gene[164]. Before the genetic cause was known, the hamster was widely used as a model for hypertrophic cardiomyopathy. The animals develop widespread muscle degeneration in both the myocardium and the skeletal musculature. The primary cause of death in the BIO14.6 hamster is heart failure.

Laminin2-deficiency leads to congenital muscular dystrophy type 1A (Chapter 1.3.5.3). A spontaneous genetic mutation in the murine laminin2 gene resulted in the dy mouse model for MDC1A. This mouse model was described as early as 1955, and was therefore the first mouse model for muscular dystrophy[152] (reviewed in Nonaka et al.[166]).

1.3.5 Secondary pathological processes

The pathological effects on cellular level show a high similarity between the muscular dystrophies. As a result of abberations in a variety of genes, which lead to muscular dystrophy, myofiber necrosis is the first secondary process to occur. Currently, the effects of apoptosis as a possibility for cell death are under investigation as well. Myofiber necrosis instigates an inflammatory reaction, and attracts inflammatory cells. Due to the activation of satellite cells, regenerative processes are started to repair or replace affected myofibers. These effects are reflected in the mouse models, and will be discussed in more detail in this chapter.

1.3.5.1 Necrosis / Apoptosis

The hallmark of muscular dystrophy is the propensity to cause myofiber cell death, according to Rando[182]. Muscular dystrophy seems to be a conditional disease, since not all muscles are affected, and the age of onset is highly variable. Although the genetic defect is generally present in all body cells, the pathology is only initiated by the degenerative necrosis of myofibers.
A common theme in muscular dystrophy is that the disease causing genes are involved in the structural maintenance of the myofiber, as well as the alignment with neighboring myofibers. As a consequence of the dysfunction of these proteins by genetic aberrations, the sarcolemma becomes unstable and prone to rupture. The myofiber can respond to the lesions in different ways (reviewed in McNeil et al.[147]). First, the lesions can be repaired by spontaneous resealing. This process, however, is incomplete, since it leaves a hole in the membrane[129]. Second, ruptures can be repaired by a process called exocytotic resealing. Membraneous vesicles are actively fused to the plasma-membrane nearby the rupture site, and thereby lower the tension of the membrane, making resealing possible[183]. Third, the rupture in the membrane can be filled with individual vesicles, which subsequently fuse, and form a 'patch'[146]. However, it might be possible that the myofiber is not capable of repairing the membrane lesions. As a

result, a continuous inward flow of extracellular cations disturbs the intracellular homeostasis (mainly Ca^{2+}), and initiates necrosis or apoptosis by activating Ca^{2+}-dependent proteolytic enzymes. Thus, myofiber degeneration is a result of an imbalanced homeostasis caused by membrane instability-induced sarcolemmal rupture.

An alternative hypothesis regarding the sequence of the above mentioned processes was postulated by Gillis[85]. This hypothesis questions the results of previous experiments concerning the increase of intracellular Ca^{2+} due to membrane rupture, and the effect it has on proteolytic activity. The experimental evidence concerning elevated intracellular Ca^{2+} is highly controversial (see[85]). Therefore, the degenerative effects might preceed the membrane rupture and the influx of Ca^{2+}. These degenerative effects may concern the activation of apoptotic processes, which might trigger the dystrophic processes (reviewed in Tews et al.[220]).

The study of apoptotic markers, such as DNA fragmentation and morphological criteria, in mdx mice revealed a peak at 2 weeks of age, prior to the massive necrosis seen during the degenerative phase[189,210]. Furthermore, apoptotic processes were detected in skeletal muscle from sarcoglycan-deficient and laminin2-deficient mice[90,154]. The molecular mechanisms behind apoptosis have become more clear recently (reviewed in Hay et al.[96]). Important apoptopic factors are represented by the bcl-2 family, with both pro-apoptotic members (bax) and anti-apoptotic members (bcl-2)[225], and the caspase protein family[53]. Interestingly, apoptotic events also play a role in the formation of multi-nuclear (syncytial) cells by fusion (reviewed in Huppertz et al.[106]). Therefore, it is not surprising that apoptosis has solely been found in regenerating myofibers in dystrophic muscle tissue[221].

Calcium plays a major role in the activation of apoptotic processes. Cellular overload of Ca^{2+}, or the perturbation of intracellular Ca^{2+} compertimentalization can cause cytotoxicity and trigger apoptotic cell death (reviewed in Orrenius et al.[172]). Autolysis of myofibers can be initiated by the activation of Ca^{2+}-dependent proteolytic enzymes (proteases), known as calpains. Calpains cleave myofibrillar and cytoskeletal proteins (reviewed in Goll et al.[86]). Therefore, the initiation of apoptotic processes, as a result of a disturbance of the intracellular calcium levels, is likely to occur in dystrophic myofibers.

In conclusion, the initiating mechanisms leading to necrosis or apoptosis during degeneration in muscular dystrophy remain to be further investigated.

1.3.5.2 Inflammation

The primary function of the immune system is to protect the organism from pathogens via an intricate system of highly specialized cells. Except for eliminating the pathogenic effects of infectious agens (bacterial, viral, fungal, and parasitic), the immune system is employed to repair injured tissue. By starting an inflammatory response, damaging agents can be confined to the the site of injury, affected tissue can be removed, and reparation can be initiated.

The degeneration of myofibers in muscular dystrophies is followed by a localized inflammatory response. The inflammatory response is initiated by the release of a chemical signal at

the site of injury. This signal originates from local immune cells, which are attracted by intracellular components leaking into the interstitium due to a ruptered sarcolemma. The first immune cells at the site of injury are polymorphonuclear leucocytes (i.e neutrophils), and a few macrophages. The number of macrophages increases significantly during the first 48 hours after injury[171,143]. The objective of macrophage presence is dual: macrophages remove necrotic myofibers by phagocytosis, and macrophages activate and attract satellite cells for regenerative purposes[185]. Depletion of the macrophage population by irradiation impairs the regeneration of muscle, whereas reconstitution of bone marrow restores the regenerative potential[128].

Necrosis of the myofiber leaves the endomysial surroundings intact, as well as the satellite cells at the site of degeneration. Although the necrotic myofiber is broken down by proteolytic activity, the remnants are cleared by phagocytosis of attracted macrophages. Since endogenous macrophages are relatively rare in normal adult muscle, preservation of the vascular supply is essential for the supply of blood-borne macrophages[18,2].

In conclusion, inflammation facilitates regenerative processes by clearance of degenerative myofibers, and subsequent activation of satellite cells.

1.3.5.3 Regeneration

Satellite cells (SC) are capable of restoring injured myofibers by activation, proliferation, differentiation and fusion to existing or new myofibers (reviewed in Charge et al.[42]). Activation of satellite cells from their normally quiescent state relies on the release of growth factors, also known as competence factors. The origin of these growth factors is thought to be plural. It ranges from autocrine release to vasculature, or immune related release (see Chapter 1.2.2).

Upon activation, satellite cells start to migrate to the site of injury. However, this migration depends on the vascularization of the necrotic area. Satellite cells at the site of injury first migrate to the periphery of the necrotic area[199], before remigrating towards the necrotic site during revascularization[180]. Just then, the satellite cells will start to proliferate. Therefore, it may be concluded that vascular presence is necessary, or at least facilitates, regeneration.

During proliferation, the number of satellite cells increases. A part of the proliferating satellite cell pool is used to repopulate the quiescent satellite cell population (self-renewal)[95]. The other part is used for regenerative purposes. The satellite cells are kept in a proliferative state due to the presence of proliferation-stimulating (or differentation-inhibiting) ligands, or progression factors. Well-studied progression factors are the insulin-like growth factors I and II (Igf-I and Igf-II), which are secreted by skeletal muscle. Overexpression of Igf-I results in enhanced satellite cell proliferation[39]. Igf-I participates in multiple signalling pathways, like the calcineurin/NFAT, MAP kinase and PI-3K pathways[38,47,203]. Different isoforms of fibroblast growth factor (FGF) are known to stimulate the proliferation of satellite cells[204]. Furthermore, an increase of FGF receptor expression enhances proliferation[195]. Members of the transforming growth factor β (Tgf-β) protein family are known to inhibit differentiation by inhibiting the expression of muscle regulatory factors as MyoD and myogenin[4,137,119]. However, the role of the above mentioned growth factors is not solely in proliferation. It has been shown that some

of these factors can initiate differentiation as well[4,39,248]. What causes the switch from proliferation to differentiation of satellite cells is not known in detail. Since growth factors play a role in both processes, the switch from proliferation to differentiation can probably not be seen as an on/off situation. The activity of growth factors may be changed in a more subtle manner.

Differentiation of satellite cells is marked by the expression of genes which code for contractile proteins of the sarcomere (i.e. MHC, desmin, actin-isoforms). This process is called terminal differentiation. The expression of contractile proteins is regulated by transcription factors of the MRF family. During proliferation, MyoD and Myf5 are expressed, which are markers for the early phase of myogenesis (reviewed in Sabourin et al.[188]). Terminal differentiation is marked by the expression of 'late phase' MRFs like myogenin and Mrf4. The contractile elements are eventually generated in a process called myofibrillogenesis (reviewed in Sanger et al.[192]). The myofibrils in the myofiber are gradually built by the maturation and elongation of sarcomere stretches to form a functional contractile myofiber. Before maturation of the myofiber, the differentiating satellite cells (or myoblasts) fuse to become a syncytial myofiber. Although the molecular mechanisms of myoblast fusion are not known in detail, a number of proteins essential for fusion are known. For instance, the muscle specific Ca^{2+}-dependent adhesion molecule m-cadherin is a satellite cell marker, and its expression is induced during myotube formation during myogenesis[65] and regeneration[158]. M-cadherin is present in a protein complex, and interacts with the cytoskeleton via microtubules, which function in the alignment of the fusing myoblasts[120]. Myoblast fusion is characterized by a number of processes; myoblasts have to (1) acquire fusion competence, (2) require the ability for recognition and adhesion, and (3) fuse by combining two lipid-bilayers into one[1]. Furthermore, when new myofibers are formed, innervation of these myofibers is required as well as proper vascularization[18]. Therefore, muscle regeneration is not limited to myofiber formation alone. In conclusion, successful regeneration requires optimal collaboration between different cell- and tissue types.

2 Gene expression profiling

The genome within each cell-type of an individual organism is initially identical, although dozens of point mutations occur as a result of cell duplication or growth of the organism. The combination of the manufactered proteins within each individual cell determines the function of the specific cell-type. The manufacture of proteins is regulated on a number of levels. First, the deoxyribonucleic acid (DNA) sequence which encodes the information for a certain protein needs to be physically accessible. Non-active DNA sequences are normally packaged tightly by histone structures. Furthermore, DNA methylation is an epigenetic mechanism to determine the activity of genes; methylated genes are not (or less) active. Secondly, the DNA sequence needs to be transcribed. Therefore, transcription factors bind to special binding sites, where a transcription complex is formed. An exact copy is produced in the form a ribonucleic acid (RNA). Subsequently the pre-messenger RNA molecule is processed, in a process called splicing, to remove non-coding sequence in order to remain only protein coding information. This messenger RNA (mRNA) is then stabilized by processing the ends of the molecules by capping the beginning of the molecule, and by addition of a polyA tail. The mRNA molecule is ready to be translated into the protein molecule by ribosomes and additional complexes.

The collection of mRNA molecules in a cell predominantly determines which proteins are expressed. Since the function of the cell is determined by the collection of proteins, this function can be determined indirectly via knowledge of available mRNA molecules. However, one must keep in mind that the correlation between mRNA and protein levels is not absolute, but can be influenced by factors such as protein translation rates, post-translational modifications, and half-lives of mRNA and proteins[89,181,238,51].

2.1 Introduction

In 1975, Ed Southern developed a technique in which double-stranded DNA fragments, separated by gel electrophoresis and denatured in situ, are transferred to a solid nitro-cellulose support by lateral diffusion. This technique was called the Southern blot. Soluble single-stranded DNA fragments, or probes, are able to hybridize to their complementary sequences on the blot. Hence, specific DNA sequences can be detected by labeling of the probes using radioactivity. The first descendent of the Southern blot was characterized by replacing DNA with cellular RNA as transferred material, and was called the Northern blot. As a result, the presence of specific mRNA molecules could be detected in a sample and, for the first time, gene expression could be measured.

The detection and quantification of specific RNA sequences is limited using the Northern blot. The technique is inefficient for determining the expression of a large number of genes, because the expression of genes needs to be determined one by one. Recently, this limitation is circumvented by two techniques, namely serial analysis of gene expression (SAGE) and the use of microarrays.

The SAGE method determines the abundancy of mRNA molecules by counting gene specific sequences[231]. In short, mRNA molecules are converted into double-stranded DNA molecules

by reverse transcription, and cleaved using a restriction enzyme. After the linking of a docking module to the cleaved site, a short piece of DNA is cut off using a restriction enzyme that leaves a specific overhang. This short piece, or tag, is approximately 10-14 basepairs long. Within the sample under study, thousands of tags are generated, which can be ligated to long concatemers using the specific overhang. Subsequently, the concatemers are cloned into plasmids, which are transfected to bacteria, and eventually sequenced. Using a computer program, the tags are linked to their gene of origin and the number of hits determines the abundance in the sample under study. This method can be used to compare the specific mRNA abundances in different samples on a large scale.

The microarray is basically a reverse Northern blot. The specific nucleotide sequences are bound to a solid support, while cellular RNA (or DNA complement thereof) is, generated, labeled and hybridized to detect gene expression. The microarray technique is the main technique used in this thesis and will, therefore, be explained in detail in the following chapter.

2.2 The technique

2.2.1 Target handling

Total RNA derived from a biological sample of interest functions as template material for the microarray technique. To acquire labeled target material, the mRNA molecules are reverse transcribed into complementary DNA strands, incorporating fluorescently labelled nucleotides. Alternatively, mRNA molecules can be amplified using an in vitro transcription (IVT) assay. mRNA is reverse transcribed into double-stranded cDNA molecules, which function as template for transcribing cRNA molecules, thereby amplifying the amount of starting material. During the transcription, fluorescently labelled nucleotides are incorporated. Fluorescent labels can also be linked indirectly to the cRNA via the formation of cis-platin complexes[93], or by chemical coupling[216]. The amplification technique can be used when the amount of starting material is limited.

Fluorescent labeled probes can be detected by laser excitation of the incorporated fluorochromes. Laser excitation yields an emission with a characteristic spectrum, which is measured using a scanning confocal laser microscope. Multiple differently labelled probes can be detected, because each fluorochrome emits its own spectrum. This allows the hybridization of differently labeled samples on a single microarray, thereby facilitating direct comparisons of gene expression levels between different samples. The technique is summarized in figure 6.

2.2.2 Platforms

Gene expression profiling can be performed on different experimental platforms. The platform is determined by the type of DNA sequence and the material on which these sequences are bound. In general, microarrays contain thousands of specific nucleotide sequences, which are localized in a known, ordered position. Hybridization of labeled probes is based on temperature and salt conditions. Proper adjustment of the hybridization condition provides a stringency, which allows specific complementary hybridization and concurrently prevents a-specific binding.

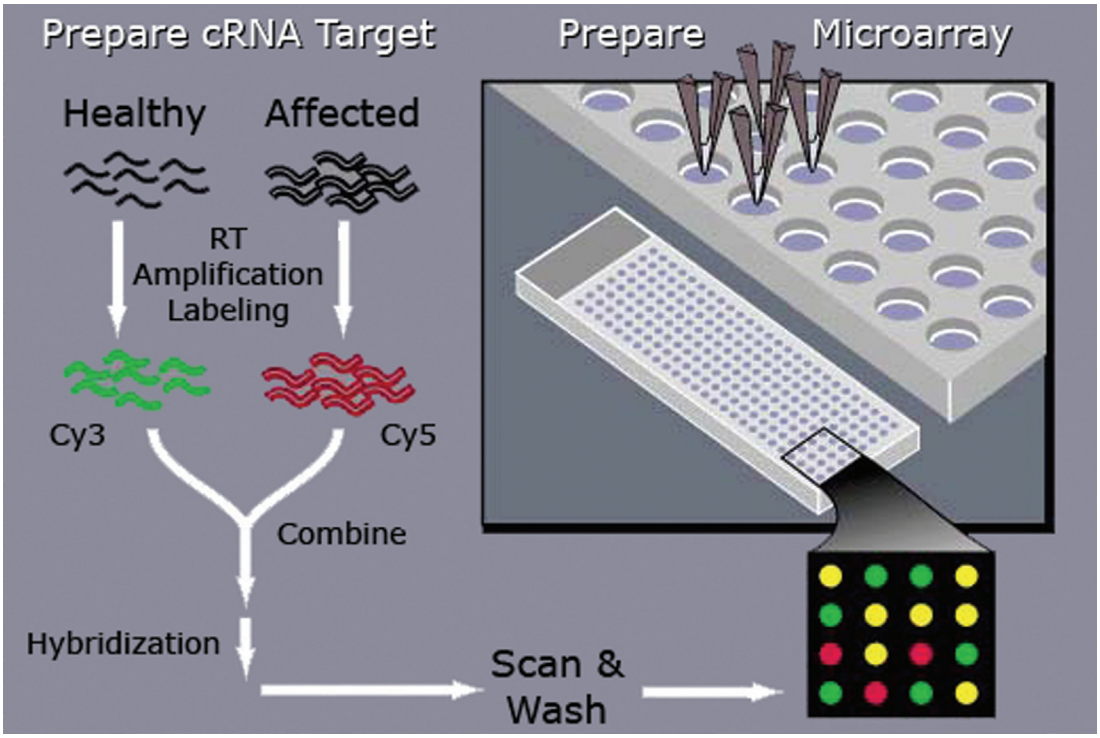


Figure 6 Overview of the microarray technique

Total RNA samples from healthy and affected origin are reverse transcribed, amplified, and labeled with a fluorescent dye into two separate probes. The probes are then combined and hybridized to a microarray. The microarray consists of thousands of different oligonucleotide sequences spotted in a ordered array on a coated microscope slide. The hybridization is followed by stringent washes and subsequent scanning using a confocal laser scanner.

(Adapted from Jakubowski (http://employees.csbsju.edu/hjakubowski/classes/ch331/bind/oldrugdevel.html))

2.2.2.1 cDNA microarrays

cDNA microarrays were the first type of microarrays available. The probes are based on cDNA sequences derived from bacterial clones harbouring a plasmid with a specific cDNA insert. The cDNA is amplified using the Polymerase Chain Reaction (PCR) using specific flanking sequences as priming sites. The length of the cDNA ranges between approximately 500 and 1500 basepairs. Following purification steps of the PCR product, the double stranded cDNA is spotted in a specific order onto coated microscope slides. The manufacture of cDNA microarrays is laborious, since PCR and subsequent purification on such a large scale is time-consuming. Furthermore, the required amount of cDNA to be spotted demands a highly efficient PCR. Although labor-intensive, the cDNA library provides a hypothetically inexhaustible source to generate amplicons.

The quality of the microarray is determined by the quality of the cDNA library. The quality control is arduous, because a subset of the PCR products need to be sequenced to check whether the cDNA sequence correlates with the given gene information. In addition, the PCR products need to checked using gel electrophoresis to determine contamination of the cDNA library by neighboring wells.

2.2.2.2 Spotted oligonucleotide microarrays

As an alternative for the use of cDNA, synthetic oligonucleotides can be used as probe. The specificity of the oligonucleotide is determined by both the sequence and the length of the oligomer. With the completion of genomic sequence information, and the engineering of software tools, gene-specific stretches of DNA sequence can be determined. Furthermore, this method is independent of the availabilty of full-length cDNA clones. As a result, region specific (i.e. 5', 3', exonic, intronic) oligonucleotides can be synthesized. Sets of oligonucleotide probes for manufacturing microarrays are commercially available in desired quantities. This circumvents the arduous amplification of cDNA probes, where the yield of the PCR and loss during purification restricts a high quality consistency. Furthermore, the quality control does not require as much effort as the cDNA microarrays. The commercially available oligonucleotide library is, in contrast to the cDNA library, a finite source for generating microarrays. As a result, the costs of generating oligonucleotide microarrays is relatively high.

2.2.2.3 In situ lithographic microarrays

The GeneChip microarrays developed by Affymetrix are manufactured by chemical synthesis of DNA fragments using in situ photolithography. The chemical synthesis starts with the coating of a quartz surface with a light-sensitive chemical compound that prevents the coupling of nucleotides to the wafer. Lithographic masks are then used to either block or transmit light onto specific locations on the quartz surface. As a result, active sites are generated to which a specific nucleotide (adenine, thymine, cytosine, or guanine) can bind. Each bound nucleotide contains a light sensitive protective group, so that only a single nucleotide can bind. After the binding of the first type of nucleotide, the process is repeated by using another lithographic mask and the next nucleotide. To generate the first nucleotide for each probe, a total of four masks need to be used. Subsequently, the other nucleotides are chemically coupled to generate

probes with a length of 65 nucleotides by using similar lithographic masks.

As a quality control, for each probe that represents a specific gene (perfect match), a so-called mismatch probe is generated. This mismatch probe contains a single mismatch at the middle of the sequence. Hence, the binding efficiency can be monitored and subsequently the fluorescence of the perfect-match can be corrected. The Affymetrix system has a one-color approach. Due to the large number of internal quality controls, the technical variability is monitored stringently. Furthermore, due to the one-color approach, the experimental design differs from two-color microarrays. In this thesis, Affymetrix GeneChips have not been used. Therefore, the technique will not be discussed in further detail.

2.2.3 Applications

Microarrays are based on deoxyribonucleic acid sequences, and the intrinsic property to hybridize with complementary sequences. As a result, microarrays can be employed on several levels. The most widely used application is gene expression profiling, where mRNA molecules (or cDNA derivates) are used as template. Furthermore, the technique can be used to detect changes in genomic DNA. To achieve this, genomic DNA is enzymatically fragemented and labelled with fluorochromes for detection.

Many applications exist for gene expression profiling using a two-color platform. The most basic application makes a direct comparison between two biological samples, for instance a diseased versus a healthy sample, or a treated versus an untreated sample. Determination of gene expression differences between more than two samples can easily be applied using a reference sample to which all samples are hybridized. Furthermore, it is possible to monitor gene expression levels over time by determining gene expression levels at different timepoints.

3 Aim and scope of the study

The genes, which cause muscular dystrophy when genetic aberrations are present, are currently known for most muscular dystrophy types (Table 1). In order to understand the pathological mechanisms, the first line of investigation was concentrated on the function of the protein product of the disease causing gene. A large number of muscular dystrophies are caused by defects in the costameric protein network, which facilitates the transverse fixation system[76]. In general, these proteins have a structural function by supporting the connection between the intracellular cytoskeleton and the extracellular matrix (such as members of the DGC). The lack of sarcolemmal stability and subsequent rupture when the DGC is not present, has become a paradigm for muscular dystrophy[37]. In other words, muscular dystrophy occurs, when the intracellular myofiber homeostasis is irreversibly altered by continuous or unrepaired ruptures of the sarcolemma. For instance, defects in genes involved in membrane repair (i.e. dysferlin) can also cause muscular dystrophy, because naturally occuring ruptures of the sarcolemma can not be repaired. However, this paradigm does not explain the occurrence of muscular dystrophies which do not demonstrate sarcolemmal rupture (i.e. laminin-α2 deficiency)[215].

Another argument against the paradigm comes from contraction induced injury during exercise. The dystrophin-deficient mdx mouse model shows an aggravation of the pathology during exercise, because the subsequent mechanical defect induces sarcolemmal rupture[178]. Hack et al. demonstrated that this feature could not be seen in γ-sarcoglycan deficient mice[91]. The authors suggested that there was no mechanical defect in γ-sarcoglycan deficient mice, since the core of the DGC was still functional. Therefore, they suggested that muscular dystrophy is induced by an alternative mechanism, likely to be involved in cell signalling[91]. This evolved from the notion that members of the DGC have potential signalling functions[245]. The hypothesis that muscular dystrophy is caused by alterations in cell signalling is supported by the similarities between the DGC and other sarcolemmal complexes involved in signalling (reviewed in Rando[182]).

Although the paradigm concerning the cause of muscular dystrophy has broadened, the effect of muscular dystrophy has undoubtedly remained. Muscular dystrophy results in the degeneration of muscle tissue. As pointed out by Rando[182], degeneration does not occur in all muscle tissues (or even myofibers) of affected individuals. Therefore, muscular dystrophy has to be seen as the 'propensity to cause muscle cell death […] rather than requisite cell death'[182]. As a result of degeneration, secondary processes are activated, which include inflammation, regeneration, fat infiltration, and fibrosis. Although degeneration initiates muscular dystrophy, these secondary processes eventually determine the severity of the pathology. Whereas normal muscle tissue has the ability to regenerate injured myofibers successfully, dystrophic muscle tissue demonstrates a replacement of affected myofibers by fat and fibrotic tissue, which indicates an unsuccessful regeneration. The underlying molecular mechanisms of these pathological processes in muscular dystrophy are not well understood.

Differences between muscular dystrophies are generally related to the genetic mutation underlying the disease. As a result, the muscular dystrophies differ in clinical presentation, age of onset, and severity. These differences are thought to be associated with the function of the

mutated gene. However, on the histological level affected muscle tissue demonstrates a high degree of similarity between the muscular dystrophies. As it is not fully understood why some muscles (or myofibers) within a single muscular dystrophy are affected and some or not, it is not fully understood why the muscular dystrophies differ as they do. The further study of muscular dystrophies is highly relevant as the alteration of the severity holds important therapeutic potential.

Aim of the study

This thesis describes the application of gene expression profiling to determine the molecular details of the processes leading to muscular dystrophy. The results of gene expression profiling of muscular dystrophies can be applied at two different levels. First, distinct gene expression profiles of each muscular dystrophy can be used as a diagnostic tool. This approach is based on pattern recognition by finding biomarkers for the individual muscular dystrophies. Second, studying the differentially expressed genes in muscular dystrophy can increase the understanding of the molecular details underlying the pathological processes in muscular dystrophy. The research which has been done to acomplish these goals will be outlined in the following chapters, and will be introduced here briefly.

To facilitate the biomedical research, a large number of wildtype mouse strains have been bred to obtain genetic homogeneity within each strain. These so-called inbred mouse strains have been used to generate animal models for human disease. We have implemented gene expression profiling to determine the effect of genetic background on gene expression levels. Therefore, we have compared the gene expression levels in skeletal muscle tissue of five wildtype mouse strains. Although the effects are limited compared to the number of differentially expressed genes between tissue affected by disease and healthy tissue, these results may facilitate the later, direct, comparison of mouse models with different genetic background by subtracting the genes, which expression deviates between mouse inbred strains. The results of this study are described in chapter 2.

To study the molecular details of the pathological processes in Duchenne muscular dystrophy, we have determined temporal gene expression changes in the mdx mouse compared to wildtype mice from the ages of 1 to 20 weeks, thereby comprising the most prominent stages of the pathology. Since the manifestation of dystrophin-deficiency in the mdx mouse, in contrast to the human pathology, does not result in lethality, we have focussed on the regenerative processes. The results of this study are described in chapter 3.

As discussed above, the various muscular dystrophies differ in genetic cause and clinical presentation. We have compared the gene expression profiles of several different mouse models for muscular dystrophies to determine the shared and distinct pathological processes. With this study, we have attempted to find biomarkers for the individual muscular dystrophies in order to facilitate the complicated diagnosis. The results of this study are described in chapter 4.

Since the regenerative processes in muscular dystrophy recapitulate myogenesis, we have determined gene expression profiles of primary myoblast cultures, which are forced into differentiation. The results of this study are described in chapter 5.

After the development and first implementations of the microarray technique, a large effort has been made to refine and optimize the application of microarrays. Due to an increased demand in the number of variables used in a typical microarray experiment, we set out to investigate a more flexible experimental design and subsequent method for data analysis. The experiments leading to this are discussed in chapter 6.

This thesis is concluded by a general discussion in which the obtained knowledge from the gene expression profiling studies is placed in a broader perspective. Furthermore, the advantages and limitations of the gene expression profiling technique are discussed. This discussion is the topic of chapter 7.

4 References

1. Abmayr SM, Balagopalan L, Galletta BJ, Hong SJ. Cell and molecular biology of myoblast fusion. Int Rev Cytol 2003;225:33-89.
2. Abood EA, Jones MM. Macrophages in developing mammalian skeletal muscle: evidence for muscle fibre death as a normal developmental event. Acta Anat (Basel) 1991;140(3):201-212.
3. Achanzar WE, Ward S. A nematode gene required for sperm vesicle fusion. J Cell Sci 1997;110 (Pt 9):1073-1081.
4. Allen RE, Boxhorn LK. Regulation of skeletal muscle satellite cell proliferation and differentiation by transforming growth factor-beta, insulin-like growth factor I, and fibroblast growth factor. J Cell Physiol 1989;138(2):311-315.
5. Anastasi G, Amato A, Tarone G, Vita G, Monici MC, Magaudda L, Brancaccio M, Sidoti A, Trimarchi F, Favaloro A, Cutroneo G. Distribution and localization of vinculin-talin-integrin system and dystro phin-glycoprotein complex in human skeletal muscle. Immunohistochemical study using confocal laser scanning microscopy. Cells Tissues Organs 2003;175(3):151-164.
6. Anderson LV, Davison K, Moss JA, Young C, Cullen MJ, Walsh J, Johnson MA, Bashir R, Britton S, Keers S, Argov Z, Mahjneh I, Fougerousse F, Beckmann JS, Bushby KM. Dysferlin is a plasma mem brane protein and is expressed early in human development. Hum Mol Genet 1999;8(5):855-861.
7. Aoki M, Liu J, Richard I, Bashir R, Britton S, Keers SM, Oeltjen J, Brown HE, Marchand S, Bourg N, Beley C, McKenna-Yasek D, Arahata K, Bohlega S, Cupler E, Illa I, Majneh I, Barohn RJ, Urtiz bereaJA, Fardeau M, Amato A, Angelini C, Bushby K, Beckmann JS, Brown RH, Jr. Genomic organization of the dysferlin gene and novel mutations in Miyoshi myopathy. Neurology 2001; 57(2):271-278.
8. Araishi K, Sasaoka T, Imamura M, Noguchi S, Hama H, Wakabayashi E, Yoshida M, Hori T, Ozawa E. Loss of the sarcoglycan complex and sarcospan leads to muscular dystrophy in beta-sarcoglycan-defi cient mice. Human Molecular Genetics 1999;8(9):1589.
9. Araki E, Nakamura K, Nakao K, Kameya S, Kobayashi O, Nonaka I, Kobayashi T, Katsuki M. Targeted disruption of exon 52 in the mouse dystrophin gene induced muscle degeneration similar to that observed in Duchenne muscular dystrophy. BiochemBiophysResComm 1997;238(2):492.
10. Argov Z, Sadeh M, Mazor K, Soffer D, Kahana E, Eisenberg I, Mitrani-Rosenbaum S, Richard I, Beck mann J, Keers S, Bashir R, Bushby K, Rosenmann H. Muscular dystrophy due to dysferlin deficiency in Libyan Jews. Clinical and genetic features. Brain 2000;123 (Pt 6):1229-1237.
11. Baghdiguian S, Martin M, Richard I, Pons F, Astier C, Bourg N, Hay RT, Chemaly R, Halaby G, Loiselet J, Anderson LV, Lopez de Munain A, Fardeau M, Mangeat P, Beckmann JS, Lefranc G. Calpain 3 deficiency is associated with myonuclear apoptosis and profound perturbation of the IkappaB alpha/NF-kappaB pathway in limb-girdle muscular dystrophy type 2A. Nat Med 1999;5(5):503-511.
12. Bansal D, Miyake K, Vogel SS, Groh S, Chen CC, Williamson R, McNeil PL, Campbell KP. Defective membrane repair in dysferlin-deficient muscular dystrophy. Nature 2003;423(6936):168.
13. Bansal D, Campbell KP. Dysferlin and the plasma membrane repair in muscular dystrophy. Trends Cell Biol 2004;14(4):206-213.
14. Barresi R, Di Blasi C, Negri T, Brugnoni R, Vitali A, Felisari G, Salandi A, Daniel S, Cornelio F, Morandi L, Mora M. Disruption of heart sarcoglycan complex and severe cardiomyopathy caused by beta sarcoglycan mutations. J Med Genet 2000;37(2):102-107.
15. Bashir R, Britton S, Strachan T, Keers S, Vafiadaki E, Lako M, Richard I, Marchand S, Bourg N, Argov Z, Sadeh M, Mahjneh I, Marconi G, Passos-Bueno MR, Moreira Ed, Zatz M, Beckmann JS, Bushby K. A gene related to Caenorhabditis elegans spermatogenesis factor fer-1 is mutated in limb-girdle muscular dystrophy type 2B. Nature Genetics 1998;20(1):37.
16. Becker PE. Dystrophia Musculorum Progressiva. Stuttgart: Thieme; 1953.
17. Beggs AH, Hoffman EP, Snyder JR, Arahata K, Specht L, Shapiro F, Angelini C, Sugita H, Kunkel LM. Exploring the molecular basis for variability among patients with Becker muscular dystrophy: dystro phin gene and protein studies. American Journal of Human Genetics 1991;49:54.
18. Bischoff R, Franzini-Armstrong C. Satellite and Stem Cells in Muscle Regeneration. In: Engel AG, Franzini-Armstrong C, editors. Myology. 3rd ed. Volume 1. New York: McGraw-Hill; 2004. p 66-86.

19. Bittner RE, Anderson LV, Burkhardt E, Bashir R, Vafiadaki E, Ivanova S, Raffelsberger T, Maerk I, Hoger H, Jung M, Karbasiyan M, Storch M, Lassmann H, Moss JA, Davison K, Harrison R, Bushby KM, Reis A. Dysferlin deletion in SJL mice (SJL-Dysf) defines a natural model for limb girdle muscular dystrophy 2B. Nature Genetics 1999;23(2):141.
20. Blagden CS, Hughes SM. Extrinsic influences on limb muscle organisation. Cell Tissue Res 1999;296(1):141-150.
21. Blake DR, Love DR, Tinsley J, Morris GE, Turley H, Gatter K, Dickson G, Edwards YH, Davies KE. Characterization of a 4.8 kb transcript from the Duchenne muscular dystrophy locus expressed in Schwannoma cells. Human Molecular Genetics 1992;1:103.
22. Bonilla E, Samitt CE, Miranda AF, Hays AP, Salviati G, Dimauro S, Kunkel LM, Hoffman EP, Rowland LP. Duchenne muscular dystrophy: Deficiency of dystrophin at the muscle cell surface. Cell 1988;54:447.
23. Bonnemann CG, Finkel RS. Sarcolemmal proteins and the spectrum of limb-girdle muscular dystro phies. Semin Pediatr Neurol 2002;9(2):81-99.
24. Bonnemann CG, Bushby K. The Limb-Girdle Muscular Dystrophies. In: Engel AG, Franzini-Armstrong C, editors. Myology. 3rd ed. Volume 2. New York: McGraw-Hill; 2004. p 1077-1121.
25. Borycki AG, Brunk B, Tajbakhsh S, Buckingham M, Chiang C, Emerson CP, Jr. Sonic hedgehog controls epaxial muscle determination through Myf5 activation. Development 1999;126(18):4053-4063.
26. Bowman W. On the minute structure and movements of voluntary muscle. PhilosTransRSocLondBiol 1840;130:457.
27. Bradley WG, Hudgson P, Larson PF, Papapetropoulos TA, Jenkison M. Structural changes in the early stages of Duchenne muscular dystrophy. J Neurol Neurosurg Psychiatry 1972;35(4):451-455.
28. Brand-Saberi B, Christ B. Genetic and epigenetic control of muscle development in vertebrates. Cell Tissue Res 1999;296(1):199-212.
29. Brennan TJ, Chakraborty T, Olson EN. Mutagenesis of the myogenin basic region identifies an ancient protein motif critical for activation of myogenesis. Proc Natl Acad Sci U S A 1991;88(13):5675-5679.
30. Brockdorff N, Cross GS, Cavanna JS, Fisher EM, Lyon MF, Davies KE, Brown SD. The mapping of a cDNA from the human X-linked Duchenne muscular dystrophy gene to the mouse X chromosome. Nature 1987;328(6126):166.
31. Buckingham M, Bajard L, Chang T, Daubas P, Hadchouel J, Meilhac S, Montarras D, Rocancourt D, Relaix F. The formation of skeletal muscle: from somite to limb. JAnat 2003;202(1):59.
32. Bulfield G, Siller WG, Wight PA, Moore KJ. X chromosome-linked muscular dystrophy (mdx) in the mouse. ProcNatlAcadSciUSA 1984;81:1189.
33. Bushby KM. Genetic and clinical correlations of Xp21 muscular dystrophy. J Inherit Metab Dis 1992;15(4):551-564.
34. Bushby KM. The limb-girdle muscular dystrophies-multiple genes, multiple mechanisms. Hum Mol Genet 1999;8(10):1875-1882.
35. Bushby KM. Making sense of the limb-girdle muscular dystrophies. Brain 1999;122 (Pt 8):1403.
36. Carpenter JL, Hoffman EP, Romanul FC, Kunkel LM, Rosales RK, Ma NS, Dasbach JJ, Rae JF, Moore FM, McAfee MB, et al. Feline muscular dystrophy with dystrophin deficiency. Am J Pathol 1989;135(5):909-919.
37. Carpenter S, Karpati G. Duchenne muscular dystrophy: plasma membrane loss initiates muscle cell necrosis unless it is repaired. Brain 1979;102(1):147.
38. Chakravarthy MV, Abraha TW, Schwartz RJ, Fiorotto ML, Booth FW. Insulin-like growth factor-I ex tends in vitro replicative life span of skeletal muscle satellite cells by enhancing G1/S cell cycle progres sion via the activation of phosphatidylinositol 3'-kinase/Akt signaling pathway. J Biol Chem 2000;275(46):35942-35952.
39. Chakravarthy MV, Davis BS, Booth FW. IGF-I restores satellite cell proliferative potential in immobi lized old skeletal muscle. J Appl Physiol 2000;89(4):1365-1379.
40. Chan YM, Bonnemann CG, Lidov HG, Kunkel LM. Molecular organization of sarcoglycan complex in mouse myotubes in culture. J Cell Biol 1998;143(7):2033-2044.
41. Chapman VM, Miller DR, Armstrong D, Caskey CT. Recovery of induced mutations for X chromosome-linked muscular dystrophy in mice. Proceedings of the national Academy of Sciences USA 1989;86:1292.

42. Charge SB, Rudnicki MA. Cellular and molecular regulation of muscle regeneration. Physiol Rev 2004;84(1):209.
43. Chiba A, Matsumura K, Yamada H, Inazu T, Shimizu T, Kusunoki S, Kanazawa I, Kobata A, Endo T. Structures of sialylated O-linked oligosaccharides of bovine peripheral nerve alpha-dystrogly can. The role of a novel O-mannosyl-type oligosaccharide in the binding of alpha-dystroglycan with laminin. J Biol Chem 1997;272(4):2156-2162.
44. Christ B, Ordahl CP. Early stages of chick somite development. Anat Embryol (Berl) 1995;191(5):381-396.
45. Clark KA, McElhinny AS, Beckerle MC, Gregorio CC. Striated muscle cytoarchitecture: an intricate web of form and function. Annu Rev Cell Dev Biol 2002;18:637-706.
46. Colognato H, Yurchenco PD. Form and function: the laminin family of heterotrimers. Dev Dyn 2000;218(2):213-234.
47. Coolican SA, Samuel DS, Ewton DZ, McWade FJ, Florini JR. The mitogenic and myogenic actions of insulin-like growth factors utilize distinct signaling pathways. J Biol Chem 1997;272(10):6653-6662.
48. Cooper BJ, Winand NJ, Stedman H, Valentine BA, Hoffman EP, Kunkel LM, Scott MO, Fischbeck KH, Kornegay JN, Avery RJ, et al. The homologue of the Duchenne locus is defective in X-linked muscular dystrophy of dogs. Nature 1988;334(6178):154-156.
49. Coral-Vazquez R, Cohn RD, Moore SA, Hill JA, Weiss RM, Davisson RL, Straub V, Barresi R, Bansal D, Hrstka RF, Williamson R, Campbell KP. Disruption of the sarcoglycan-sarcospan complex in vascular smooth muscle: a novel mechanism for cardiomyopathy and muscular dystrophy. Cell 1999;98(4):465.
50. Cossu G, Molinaro M. Cell heterogeneity in the myogenic lineage. Curr Top Dev Biol 1987;23:185-208.
51. Cox B, Kislinger T, Emili A. Integrating gene and protein expression data: pattern analysis and profile mining. Methods 2005;35(3):303-314.
52. Cox GA, Phelps SF, Chapman VM, Chamberlain JS. New mdx mutation disrupts expression of muscle and nonmuscle isoforms of dystrophin. Nature Genetics 1993;4:87.
53. Creagh EM, Martin SJ. Caspases: cellular demolition experts. Biochem Soc Trans 2001;29(Pt 6):696-702.
54. Crosbie RH, Heighway J, Venzke DP, Lee JC, Campbell KP. Sarcospan, the 25-kDa transmembrane component of the dystrophin- glycoprotein complex. Journal of Biological Chemistry 1997;272(50):31221.
55. Crosbie RH, Yamada H, Venzke DP, Lisanti MP, Campbell KP. Caveolin-3 is not an integral component of the dystrophin glycoprotein complex. FEBSLett 1998;427(2):279.
56. Cross RA, Stewart M, Kendrick-Jones J. Structural predictions for the central domain of dystrophin. FEBS Lett 1990;262(1):87-92.
57. Cullen MJ, Jaros E. Ultrastructure of the skeletal muscle in the X chromosome-linked dystrophic (mdx) mouse. Comparison with Duchenne muscular dystrophy. Acta Neuropathol (Berl) 1988;77(1):69-81.
58. Dalkilic I, Kunkel LM. Muscular dystrophies: genes to pathogenesis. CurrOpinGenetDev 2003;13(3):231.
59. Dangain J, Vrbova G. Muscle development in mdx mutant mice. Muscle Nerve 1984;7(9):700.
60. Danko I, Chapman V, Wolff JA. The frequency of revertants in mdx mouse genetic models for Duchenne muscular dystrophy. Pediatr Res 1992;32(1):128-131.
61. Danowski BA, Imanaka-Yoshida K, Sanger JM, Sanger JW. Costameres are sites of force transmission to the substratum in adult rat cardiomyocytes. J Cell Biol 1992;118(6):1411-1420.
62. Deconinck AE, Rafael JA, Skinner JA, Brown SC, Potter AC, Metzinger L, Watt DJ, Dickson JG, Tinsley JM, Davies KE. Utrophin-dystrophin-deficient mice as a model for Duchenne muscular dystrophy. Cell 1997;90(4):717-727.
63. Dietrich S, Abou-Rebyeh F, Brohmann H, Bladt F, Sonnenberg-Riethmacher E, Yamaai T, Lumsden A, Brand-Saberi B, Birchmeier C. The role of SF/HGF and c-Met in the development of skeletal muscle. Development 1999;126(8):1621-1629.
64. Doheny DO, Brin MF, Morrison CE, Smith CJ, Walker RH, Abbasi S, Muller B, Garrels J, Liu L, De Carvalho Aguiar P, Schilling K, Kramer P, De Leon D, Raymond D, Saunders-Pullman R, Klein C, Bressman SB, Schmand B, Tijssen MA, Ozelius LJ, Silverman JM. Phenotypic features of myoclonus-dystonia in three kindreds. Neurology 2002;59(8):1187-1196.

65. Donalies M, Cramer M, Ringwald M, Starzinski-Powitz A. Expression of M-cadherin, a member of the cadherin multigene family, correlates with differentiation of skeletal muscle cells. Proc Natl Acad Sci U S A 1991;88(18):8024-8028.
66. Duchenne GBA. Recherches sur la paralysie musculaire pseudohypertrophique ou paralysie myo-sclero sique. ArchGenMed 1868;11:5.
67. Duclos F, Straub V, Moore SA, Venzke DP, Hrstka RF, Crosbie RH, Durbeej M, Lebakken CS, Ettinger AJ, van der MJ, Holt KH, Lim LE, Sanes JR, Davidson BL, Faulkner JA, Williamson R, Campbell KP. Progressive muscular dystrophy in alpha-sarcoglycan-deficient mice. Journal of Cell Biology 1998;142(6):1461.
68. Durbeej M, Cohn RD, Hrstka RF, Moore SA, Allamand V, Davidson BL, Williamson RA, Campbell KP. Disruption of the beta-sarcoglycan gene reveals pathogenetic complexity of limb-girdle muscular dystrophy type 2E. MolCell 2000;5(1):141.
69. Emerson CP, Jr., Hauschka SD. Development, Differentiation, and Diversity. In: Engel AG, Franzini-Armstrong C, editors. Myology. 3rd ed. Volume 1. New York: McGraw-Hill; 2004. p 3-44.
70. Emery AE. Population frequencies of inherited neuromuscular diseases--a world survey. Neuromuscul Disord 1991;1(1):19-29.
71. Emery AE. The muscular dystrophies. Lancet 2002;359(9307):687.
72. Engel AG, Banker BQ. Ultrastructural Changes in Diseased Muscle. In: Engel AG, Franzini-Armstrong C, editors. Myology. 3rd ed. Volume 1. New York: McGraw-Hill; 2004. p 749-887.
73. Engel AG, Ozawa E. Dystrophinopathies. In: Engel AG, Franzini-Armstrong C, editors. Myology. 3rd ed. Volume 2. New York: McGraw-Hill; 2004. p 961-1025.
74. Engel J, Furthmayr H, Odermatt E, von der Mark H, Aumailley M, Fleischmajer R, Timpl R. Structure and macromolecular organization of type VI collagen. Ann N Y Acad Sci 1985;460:25-37.
75. Epstein JA, Shapiro DN, Cheng J, Lam PY, Maas RL. Pax3 modulates expression of the c-Met receptor during limb muscle development. Proc Natl Acad Sci U S A 1996;93(9):4213-4218.
76. Ervasti JM. Costameres: the Achilles' heel of Herculean muscle. J Biol Chem 2003;278(16):13591-13594.
77. Ettinger AJ, Feng G, Sanes JR. epsilon-Sarcoglycan, a broadly expressed homologue of the gene mutated in limb-girdle muscular dystrophy 2D. Journal of Biological Chemistry 1997;272(51):32534.
78. Fardeau M, Hillaire D, Mignard C, Feingold N, Feingold J, Mignard D, de Ubeda B, Collin H, Tome FM, Richard I, Beckmann J. Juvenile limb-girdle muscular dystrophy. Clinical, histopathological and genetic data from a small community living in the Reunion Island. Brain 1996;119 (Pt 1):295-308.
79. Faulkner G, Pallavicini A, Comelli A, Salamon M, Bortoletto G, Ievolella C, Trevisan S, Kojic S, Dalla Vecchia F, Laveder P, Valle G, Lanfranchi G. FATZ, a filamin-, actinin-, and telethonin-binding protein of the Z-disc of skeletal muscle. J Biol Chem 2000;275(52):41234-41242.
80. Frey N, Richardson JA, Olson EN. Calsarcins, a novel family of sarcomeric calcineurin-binding proteins. Proc Natl Acad Sci U S A 2000;97(26):14632-14637.
81. Fuchs E, Yang Y. Crossroads on cytoskeletal highways. Cell 1999;98(5):547-550.
82. Fulton AB. Spatial organization of the synthesis of cytoskeletal proteins. J Cell Biochem 1993; 52(2):148-152.
83. Gaschen FP, Hoffman EP, Gorospe JR, Uhl EW, Senior DF, Cardinet GH, 3rd, Pearce LK. Dystrophin deficiency causes lethal muscle hypertrophy in cats. J Neurol Sci 1992;110(1-2):149-159.
84. Gilchrist JM, Pericak-Vance M, Silverman L, Roses AD. Clinical and genetic investigation in autosomal dominant limb-girdle muscular dystrophy. Neurology 1988;38(1):5-9.
85. Gillis JM. Understanding dystrophinopathies: an inventory of the structural and functional consequences of the absence of dystrophin in muscles of the mdx mouse. J Muscle Res Cell Motil 1999;20(7):605-625.
86. Goll DE, Thompson VF, Li H, Wei W, Cong J. The calpain system. Physiol Rev 2003;83(3):731-801.
87. Gustafsson MK, Pan H, Pinney DF, Liu Y, Lewandowski A, Epstein DJ, Emerson CP, Jr. Myf5 is a direct target of long-range Shh signaling and Gli regulation for muscle specification. Genes Dev 2002;16(1):114-126.
88. Guyon JR, Kudryashova E, Potts A, Dalkilic I, Brosius MA, Thompson TG, Beckmann JS, Kunkel LM, Spencer MJ. Calpain 3 cleaves filamin C and regulates its ability to interact with gamma- and delta-sarcoglycans. Muscle Nerve 2003;28(4):472-483.
89. Gygi SP, Rochon Y, Franza BR, Aebersold R. Correlation between protein and mRNA abundance in yeast. Mol Cell Biol 1999;19(3):1720-1730.

90. Hack AA, Ly CT, Jiang F, Clendenin CJ, Sigrist KS, Wollmann RL, McNally EM. Gamma-sarcoglycan deficiency leads to muscle membrane defects and apoptosis independent of dystrophin. Journal of Cell Biology 1998;142(5):1279.
91. Hack AA, Cordier L, Shoturma DI, Lam MY, Sweeney HL, McNally EM. Muscle degeneration without mechanical injury in sarcoglycan deficiency. Proc Natl Acad Sci U S A 1999;96(19):10723-10728.
92. Hack AA, Lam MY, Cordier L, Shoturma DI, Ly CT, Hadhazy MA, Hadhazy MR, Sweeney HL, Mc Nally EM. Differential requirement for individual sarcoglycans and dystrophin in the assembly and function of the dystrophin-glycoprotein complex. Journal of Cell Science 2000;113 (Pt 14):2535.
93. Hagedoorn R, Joseph R, Kasanmoentalib S, Eilers P, Killian J, Raap AK. Chemical RNA labeling with out 3' end bias using fluorescent cis-platin compounds. Biotechniques 2003;34(5):974-976, 978, 980.
94. Hauser MA, Horrigan SK, Salmikangas P, Torian UM, Viles KD, Dancel R, Tim RW, Taivainen A, Bartoloni L, Gilchrist JM, Stajich JM, Gaskell PC, Gilbert JR, Vance JM, Pericak-Vance MA, Carpen O, Westbrook CA, Speer MC. Myotilin is mutated in limb girdle muscular dystrophy 1A. Hum Mol Genet 2000;9(14):2141-2147.
95. Hawke TJ, Garry DJ. Myogenic satellite cells: physiology to molecular biology. JApplPhysiol 2001;91(2):534.
96. Hay BA, Huh JR, Guo M. The genetics of cell death: approaches, insights and opportunities in Drosophila. Nat Rev Genet 2004;5(12):911-922.
97. Helbling-Leclerc A, Zhang X, Topaloglu H, Cruaud C, Tesson F, Weissenbach J, Tome FM, Schwartz K, Fardeau M, Tryggvason K, et al. Mutations in the laminin alpha 2-chain gene (LAMA2) cause mero sin-deficient congenital muscular dystrophy. Nat Genet 1995;11(2):216-218.
98. Herrmann H, Aebi U. Intermediate filaments and their associates: multi-talented structural elements specifying cytoarchitecture and cytodynamics. Curr Opin Cell Biol 2000;12(1):79-90.
99. Hodgson S, Hart K, Abbs S, Heckmatt J, Rodillo E, Bobrow M, Dubowitz V. Correlation of clinical and deletion data in Duchenne and Becker muscular dystrophy. J Med Genet 1989;26(11):682-693.
100. Hoffman EP, Monaco AP, Feener CA, Kunkel LM. Conservation of the Duchenne muscular dystrophy gene in mice and humans. Science 1987;238:347.
101. Hoffman EP, Fischbeck KH, Brown RH, Johnson M, Medori R, Loike JD, Harris JB, Waterston R, Brooke M, Specht L, Kupsky W, Chamberlain J, Caskey CT, Shapiro F, Kunkel LM. Characterization of dystrophin in muscle-biopsy specimens from patients with Duchenne's or Becker's muscular dystrophy. New EnglJMed 1988;318:1363.
102. Holt KH, Campbell KP. Assembly of the sarcoglycan complex. Insights for muscular dystrophy. J Biol Chem 1998;273(52):34667-34670.
103. Holt KH, Crosbie RH, Venzke DP, Campbell KP. Biosynthesis of dystroglycan: processing of a precursor propeptide. FEBS Lett 2000;468(1):79-83.
104. Howell JM, Fletcher S, Kakulas BA, O'Hara M, Lochmuller H, Karpati G. Use of the dog model for Duchenne muscular dystrophy in gene therapy trials. Neuromuscul Disord 1997;7(5):325-328.
105. Huang X, Poy F, Zhang R, Joachimiak A, Sudol M, Eck MJ. Structure of a WW domain containing fragment of dystrophin in complex with beta-dystroglycan. Nat Struct Biol 2000;7(8):634-638.
106. Huppertz B, Tews DS, Kaufmann P. Apoptosis and syncytial fusion in human placental trophoblast and skeletal muscle. IntRevCytol 2001;205:215.
107. Huxley AF, Niedergerke R. Structural changes in muscle during contraction; interference microscopy of living muscle fibres. Nature 1954;173(4412):971-973.
108. Huxley H, Hanson J. Changes in the cross-striations of muscle during contraction and stretch and their structural interpretation. Nature 1954;173(4412):973-976.
109. Ibraghimov-Beskrovnaya O, Ervasti JM, Leveille CJ, Slaughter CA, Sernett SW, Campbell KP. Pri mary structure of dystrophin-associated glycoproteins linking dystrophin to the extracellukar matrix. Nature 1992;355:696.
110. Ibraghimov-Beskrovnaya O, Milatovich A, Ozcelik T, Yang B, Koepnick K, Francke U, Campbell KP. Human dystroglycan: skeletal muscle cDNA, genomic structure, origin of tissue specific isoforms and chromosomal localization. Human Molecular Genetics 1993;2:1651.
111. Im WB, Phelps SF, Copen EH, Adams EG, Slightom JL, Chamberlain JS. Differential expression of dystrophin isoforms in in strains of mdx mice with different mutations. Human Molecular Genetics 1996;5(8):1149.

112. Inoue S. Ultrastructure of basement membranes. Int Rev Cytol 1989;117:57-98.
113. Ishikawa H, Sugie K, Murayama K, Ito M, Minami N, Nishino I, Nonaka I. Ullrich disease: collagen VI deficiency: EM suggests a new basis for muscular weakness. Neurology 2002;59(6):920-923.
114. Ishikawa-Sakurai M, Yoshida M, Imamura M, Davies KE, Ozawa E. ZZ domain is essentially required for the physiological binding of dystrophin and utrophin to beta-dystroglycan. Hum Mol Genet 2004;13(7):693-702.
115. Jung D, Duclos F, Apostol B, Straub V, Lee JC, Allamand V, Venzke DP, Sunada Y, Moomaw CR, Leveille CJ, Slaughter CA, Crawford TO, McPherson JD, Campbell KP. Characterization of delta-sarcoglycan, a novel component of the oligomeric sarcoglycan complex involved in limb-girdle muscu lar dystrophy. J Biol Chem 1996;271(50):32321-32329.
116. Kahana E, Gratzer WB. Minimum folding unit of dystrophin rod domain. Biochemistry 1995;34(25): 8110-8114.
117. Karpati G, Carpenter S. Small-caliber skeletal muscle fibers do not suffer deleterious consequences of dystrophic gene expression. Am J Med Genet 1986;25(4):653-658.
118. Karpati G, Carpenter S, Prescott S. Small-caliber skeletal muscle fibers do not suffer necrosis in mdx mouse dystrophy. Muscle Nerve 1988;11(8):795-803.
119. Katagiri T, Akiyama S, Namiki M, Komaki M, Yamaguchi A, Rosen V, Wozney JM, Fujisawa-Sehara A, Suda T. Bone morphogenetic protein-2 inhibits terminal differentiation of myogenic cells by suppressing the transcriptional activity of MyoD and myogenin. Exp Cell Res 1997;230(2):342-351.
120. Kaufmann U, Kirsch J, Irintchev A, Wernig A, Starzinski-Powitz A. The M-cadherin catenin complex interacts with microtubules in skeletal muscle cells: implications for the fusion of myoblasts. Journal of Cell Science 1999;112 (Pt 1):55.
121. Kennedy JM, Eisenberg BR, Reid SK, Sweeney LJ, Zak R. Nascent muscle fiber appearance in over loaded chicken slow-tonic muscle. Am J Anat 1988;181(2):203-215.
122. Khurana TS, Davies KE. Pharmacological strategies for muscular dystrophy. Nat Rev Drug Discov 2003;2(5):379-390.
123. Koenig M, Monaco AP, Kunkel LM. The complete sequence of dystrophin predicts a rod-shaped cyto skeletal protein. Cell 1988;53:219.
124. Koenig M, Beggs AH, Moyer M, Scherpf S, Heindrich K, Bettecken T, Meng G, Muller CR, Lindlof M, Kaariainen H, De la Chapelle A, Kiuru A, Savontaus ML, Gilgenkrantz H, Recan D, Chelly J, Kaplan JC, Covone AE, Archidiacono N, Romeo G, Liechti-Gallati S, Schneider V, Braga S, Moser H, Darras BT, Murphy P, Francke U, Chen JD, Morgan G, Denton MJ, Greenberg CR, Wrogemann K, Blonden LAJ, Van Paassen HMB, Van Ommen GJB, Kunkel LM. The molecular basis for Duchenne versus Becker muscular dystrophy: correlation of severity with type of deletion. American Journal of Human Genetics 1989;45:498.
125. Kornegay JN, Tuler SM, Miller DM, Levesque DC. Muscular dystrophy in a litter of golden retriever dogs. Muscle Nerve 1988;11(10):1056-1064.
126. Kudoh H, Ikeda H, Kakitani M, Ueda A, Hayasaka M, Tomizuka K, Hanaoka K. A new model mouse for Duchenne muscular dystrophy produced by 2.4Mb deletion of dystrophin gene using Cre-loxP re combination system. Biochem Biophys Res Commun 2005;328(2):507-516.
127. Lazarides E. Intermediate filaments as mechanical integrators of cellular space. Nature 1980;283(5744): 249-256.
128. Lescaudron L, Peltekian E, Fontaine-Perus J, Paulin D, Zampieri M, Garcia L, Parrish E. Blood borne macrophages are essential for the triggering of muscle regeneration following muscle transplant. Neuro muscul Disord 1999;9(2):72-80.
129. Lieber MR, Steck TL. A description of the holes in human erythrocyte membrane ghosts. J Biol Chem 1982;257(19):11651-11659.
130. Lim LE, Duclos F, Broux O, Bourg N, Sunada Y, Allamand V, Meyer J, Richard I, Moomaw C, Slaughter C, Tome C, Fardeau F, Jackson M, Beckmann CE, Campbell KP. Beta-sarcoglycan: character ization and role in limb-girdle muscular dystrophy linked to 4q12. Nature Genetics 1995;11(3):257.
131. Linssen WH, Notermans NC, Van der Graaf Y, Wokke JH, Van Doorn PA, Howeler CJ, Busch HF, De Jager AE, De Visser M. Miyoshi-type distal muscular dystrophy. Clinical spectrum in 24 Dutch pa tients. Brain 1997;120 (Pt 11):1989-1996.

132. Liu J, Aoki M, Illa I, Wu C, Fardeau M, Angelini C, Serrano C, Urtizberea JA, Hentati F, Hamida MB, Bohlega S, Culper EJ, Amato AA, Bossie K, Oeltjen J, Bejaoui K, McKenna-Yasek D, Hosler BA, Schurr E, Arahata K, De Jong PJ, Brown RH, Jr. Dysferlin, a novel skeletal muscle gene, is mutated in Miyoshi myopathy and limb girdle muscular dystrophy. Nature Genetics 1998;20(1):31.
133. Love DR, Hill DF, Dickson G, Spurr NK, Byth BC, Marsden RF, Walsh FS, Edwards YH, Davies KE. An autosomal transcript in skeletal muscle with homology to dystrophin. Nature 1989;339:55.
134. Love DR, Morris GE, Ellis JM, Fairbrother U, Marsden RF, Bloomfield JF, Edwards YH, Slater CP, Parry DJ, Davies KE. Tissue distribution of the dystrophin-related gene product and expression in the mdx and dy mouse. Proceedings of the national Academy of Sciences USA 1991;88:3243.
135. Mahjneh I, Marconi G, Bushby K, Anderson LV, Tolvanen-Mahjneh H, Somer H. Dysferlinopathy (LG MD2B): a 23-year follow-up study of 10 patients homozygous for the same frameshifting dysferlin mu tations. Neuromuscul Disord 2001;11(1):20-26.
136. Marshall PA, Williams PE, Goldspink G. Accumulation of collagen and altered fiber-type ratios as indicators of abnormal muscle gene expression in the mdx dystrophic mouse. Muscle Nerve 1989;12(7):528.
137. Martin JF, Li L, Olson EN. Repression of myogenin function by TGF-beta 1 is targeted at the basic helix-loop-helix motif and is independent of E2A products. J Biol Chem 1992;267(16):10956-10960.
138. Mason P, Bayol S, Loughna PT. The novel sarcomeric protein telethonin exhibits developmental and functional regulation. Biochem Biophys Res Commun 1999;257(3):699-703.
139. Matsuda C, Hayashi YK, Ogawa M, Aoki M, Murayama K, Nishino I, Nonaka I, Arahata K, Brown RH, Jr. The sarcolemmal proteins dysferlin and caveolin-3 interact in skeletal muscle. Hum Mol Genet 2001;10(17):1761-1766.
140. Matsuda R, Nishikawa A, Tanaka H. Visualization of dystrophic muscle fibers in mdx mouse by vital staining with Evans blue: evidence of apoptosis in dystrophin-deficient muscle. J Biochem (Tokyo) 1995;118(5):959-964.
141. Matsumura K, Ervasti JM, Ohlendieck K, Kahl SD, Campbell KP. Association of dystrophin-related protein with dystrophin-associated proteins in mdx mouse muscle. Nature 1992;360(6404):588-591.
142. Mauro A. Satellite cell of skeletal muscle fibers. J Biophys Biochem Cytol 1961;9:493-495.
143. McLennan IS. Degenerating and regenerating skeletal muscles contain several subpopulations of macro phages with distinct spatial and temporal distributions. J Anat 1996;188 (Pt 1):17-28.
144. McNally EM, Duggan D, Gorospe JR, Bonnemann CG, Fanin M, Pegoraro E, Lidov HG, Noguchi S, Ozawa E, Finkel RS, Cruse RP, Angelini C, Kunkel LM, Hoffman EP. Mutations that disrupt the car boxyl-terminus of gamma-sarcoglycan cause muscular dystrophy. Hum Mol Genet 1996;5(11):1841-1847.
145. McNally EM, de Sa M, Duggan DJ, Bonnemann CG, Lisanti MP, Lidov HGW, Vainzof M, Passos-Bueno MR, Hoffman EP, Zatz M, Kunkel LM. Caveolin-3 in muscular dystrophy. Human Molecular Genetics 1998;7(5):871.
146. McNeil PL, Vogel SS, Miyake K, Terasaki M. Patching plasma membrane disruptions with cytoplasmic membrane. J Cell Sci 2000;113 (Pt 11):1891-1902.
147. McNeil PL, Terasaki M. Coping with the inevitable: how cells repair a torn surface membrane. Nat Cell Biol 2001;3(5):E124-129.
148. Medzhitov R, Janeway C, Jr. Innate immunity. N Engl J Med 2000;343(5):338-344.
149. Meier H. Myopathies in the dog. Cornell Vet 1958;48(3):313-330.
150. Mercuri E, Yuva Y, Brown SC, Brockington M, Kinali M, Jungbluth H, Feng L, Sewry CA, Muntoni F. Collagen VI involvement in Ullrich syndrome: a clinical, genetic, and immunohistochemical study. Neurology 2002;58(9):1354-1359.
151. Meryon E. On granular and fatty degeneration of the voluntary muscles. Med Chir Trans 1852;35:72.
152. Michelson AM, Russell ES, Harman PJ. Dystrophia muscularis: a hereditary primary myopathy in the house mouse. Proc Natl Acad Sci U S A 1955;41:1079-1084.
153. Minetti C, Sotgia F, Bruno C, Scartezzini P, Broda P, Bado M, Masetti E, Mazzocco M, Egeo A, Donati MA, Volonte D, Galbiati F, Cordone G, Bricarelli FD, Lisanti MP, Zara F. Mutations in the caveolin-3 gene cause autosomal dominant limb- girdle muscular dystrophy. Nature Genetics 1998;18(4):365.
154. Miyagoe Y, Hanaoka K, Nonaka I, Hayasaka M, Nabeshima Y, Arahata K, Nabeshima Y, Takeda S. Laminin alpha2 chain-null mutant mice by targeted disruption of the Lama2 gene: a new model of merosin (laminin 2)-deficient congenital muscular dystrophy. FEBS Lett 1997;415(1):33-39.

155. Miyoshi K, Kawai H, Iwasa M, Kusaka K, Nishino H. Autosomal recessive distal muscular dystrophy as a new type of progressive muscular dystrophy. Seventeen cases in eight families including an autopsied case. Brain 1986;109 (Pt 1):31-54.
156. Mokri B, Engel AG. Duchenne dystrophy: electron microscopic findings pointing to a basic or early abnormality in the plasma membrane of the muscle fiber. Neurology 1975;25(12):1111-1120.
157. Monaco AP, Bertelson CJ, Liechti-Gallati S, Moser H, Kunkel LM. An explanation for the phenotypic differences between patients bearing partial deletions of the DMD locus. Genomics 1988;2:90.
158. Moore R, Walsh FS. The cell adhesion molecule M-cadherin is specifically expressed in developing and regenerating, but not denervated skeletal muscle. Development 1993;117(4):1409-1420.
159. Moreira ES, Vainzof M, Marie SK, Sertie AL, Zatz M, Passos-Bueno MR. The seventh form of auto somal recessive limb-girdle muscular dystrophy is mapped to 17q11-12. American Journal of Human Genetics 1997;61(1):151.
160. Moreira ES, Wiltshire TJ, Faulkner G, Nilforoushan A, Vainzof M, Suzuki OT, Valle G, Reeves R, Zatz M, Passos-Bueno MR, Jenne DE. Limb-girdle muscular dystrophy type 2G is caused by mutations in the gene encoding the sarcomeric protein telethonin. Nat Genet 2000;24(2):163-166.
161. Mues A, van der Ven PF, Young P, Furst DO, Gautel M. Two immunoglobulin-like domains of the Z-disc portion of titin interact in a conformation-dependent way with telethonin. FEBS Lett 1998; 428(1-2):111-114.
162. Nayal A, Webb DJ, Horwitz AF. Talin: an emerging focal point of adhesion dynamics. Curr Opin Cell Biol 2004;16(1):94-98.
163. Nicholson LVB, Johnson MA, Bushby KMD, Gardner-Medwin D, Curtis A, Ginjaar HB, den Dunnen JT, Welch JL, Butler TJ, Bakker E, Van Ommen GJB, Harris JB. Integrated study of 100 patients with Xp21 linked muscular dystrophy using clinical, genetic, immunochemical, and histopathological data. I. Trends across the clinical groups. JMedGenet 1993;30:728.
164. Nigro V, Okazaki Y, Belsito A, Piluso G, Matsuda Y, Politano L, Nigro G, Ventura C, Abbondanza C, Molinari AM, Acampora D, Nishimura M, Hayashizaki Y, Puca GA. Identification of the Syrian hamster cardiomyopathy gene. Hum Mol Genet 1997;6(4):601-607.
165. Noguchi S, Wakabayashi E, Imamura M, Yoshida M, Ozawa E. Formation of sarcoglycan complex with differentiation in cultured myocytes. EurJBiochem 2000;267(3):640.
166. Nonaka I. Animal models of muscular dystrophies. Lab Anim Sci 1998;48(1):8-17.
167. Oexle K, Kohlschutter A. Cause of progression in Duchenne muscular dystrophy: impaired differentiation more probable than replicative aging. Neuropediatrics 2001;32(3):123-129.
168. Ohlendieck K, Campbell KP. Dystrophin-associated proteins are greatly reduced in skeletal muscle from mdx mice. Journal of Cell Biology 1991;115(6):1685.
169. Okazaki K, Holtzer H. Myogenesis: fusion, myosin synthesis, and the mitotic cycle. Proc Natl Acad Sci U S A 1966;56(5):1484-1490.
170. Ontell M, Kozeka K. The organogenesis of murine striated muscle: a cytoarchitectural study. Am J Anat 1984;171(2):133-148.
171. Orimo S, Hiyamuta E, Arahata K, Sugita H. Analysis of inflammatory cells and complement C3 in bupivacaine-induced myonecrosis. Muscle Nerve 1991;14(6):515-520.
172. Orrenius S, Zhivotovsky B, Nicotera P. Regulation of cell death: the calcium-apoptosis link. Nat Rev Mol Cell Biol 2003;4(7):552-565.
173. Ozawa E, Noguchi S, Mizuno Y, Hagiwara Y, Yoshida M. From dystrophinopathy to sarcoglycanopathy: evolution of a concept of muscular dystrophy. Muscle Nerve 1998;21(4):421.
174. Pardee AB. G1 events and regulation of cell proliferation. Science 1989;246(4930):603-608.
175. Pardo JV, Siliciano JD, Craig SW. A vinculin-containing cortical lattice in skeletal muscle: transverse lattice elements ("costameres") mark sites of attachment between myofibrils and sarcolemma. Proc Natl Acad Sci U S A 1983;80(4):1008-1012.
176. Parker MH, Seale P, Rudnicki MA. Looking back to the embryo: defining transcriptional networks in adult myogenesis. Nat Rev Genet 2003;4(7):497-507.
177. Perkins KJ, Davies KE. The role of utrophin in the potential therapy of Duchenne muscular dystrophy. Neuromuscul Disord 2002;12 Suppl 1:S78-89.
178. Petrof BJ, Shrager JB, Stedman HH, Kelly AM, Sweeney HL. Dystrophin protects the sarcolemma from stresses developed during muscle contraction. ProcNatlAcadSciUSA 1993;90(8):3710.

179. Petrof BJ, Stedman HH, Shrager JB, Eby J, Sweeney HL, Kelly AM. Adaptations in myosin heavy chain expression and contractile function in dystrophic mouse diaphragm. Am J Physiol 1993;265 (3 Pt 1):C834-841.
180. Phillips GD, Lu DY, Mitashov VI, Carlson BM. Survival of myogenic cells in freely grafted rat rectus femoris and extensor digitorum longus muscles. Am J Anat 1987;180(4):365-372.
181. Pradet-Balade B, Boulme F, Beug H, Mullner EW, Garcia-Sanz JA. Translation control: bridging the gap between genomics and proteomics? Trends Biochem Sci 2001;26(4):225-229.
182. Rando TA. The dystrophin-glycoprotein complex, cellular signaling, and the regulation of cell survival in the muscular dystrophies. Muscle Nerve 2001;24(12):1575.
183. Raucher D, Sheetz MP. Characteristics of a membrane reservoir buffering membrane tension. Biophys J 1999;77(4):1992-2002.
184. Richard I, Broux O, Allamand V, Fougerousse F, Chiannilkulchai N, Bourg N, Brenguier L, Devaud C, Pasturaud P, Roudaut C, Hillaire D, Passos-Bueno MR, Zatz M, Tischfield JA, Fardeau M, Jackson CE, Cohen D, Beckmann JS. Mutations in the proteolytic enzyme calpain-3 cause Limb-girdle muscular dystrophy type 2A. Cell 1995;81:27.
185. Robertson TA, Maley MA, Grounds MD, Papadimitriou JM. The role of macrophages in skeletal muscle regeneration with particular reference to chemotaxis. Exp Cell Res 1993;207(2):321-331.
186. Rowland LP. Biochemistry of muscle membranes in Duchenne muscular dystrophy. Muscle Nerve 1980;3(1):3-20.
187. Rybakova IN, Amann KJ, Ervasti JM. A new model for the interaction of dystrophin with F-actin. JCell Biol 1996;135(3):661.
188. Sabourin LA, Rudnicki MA. The molecular regulation of myogenesis. Clin Genet 2000;57(1):16-25.
189. Sandri M, Carraro U, Podhorska-Okolov M, Rizzi C, Arslan P, Monti D, Franceschi C. Apoptosis, DNA damage and ubiquitin expression in normal and mdx muscle fibers after exercise. FEBS Lett 1995;373(3):291-295.
190. Sandri M, Minetti C, Pedemonte M, Carraro U. Apoptotic myonuclei in human Duchenne muscular dystrophy. Lab Invest 1998;78(8):1005-1016.
191. Sanes JR. The Extracellular Matrix. In: Engel AG, Franzini-Armstrong C, editors. Myology. 3rd ed. Volume 1. New York: McGraw-Hill; 2004. p 471-487.
192. Sanger JW, Chowrashi P, Shaner NC, Spalthoff S, Wang J, Freeman NL, Sanger JM. Myofibrillogenesis in skeletal muscle cells. Clin Orthop 2002(403 Suppl):S153-162.
193. Sarig R, Mezger-Lallemand V, Gitelman I, Davis C, Fuchs O, Yaffe D, Nudel U. Targeted inactivation of Dp71, the major non-muscle product of the DMD gene: differential activity of the Dp71 promoter during development. Hum Mol Genet 1999;8(1):1-10.
194. Sasaoka T, Imamura M, Araishi K, Noguchi S, Mizuno Y, Takagoshi N, Hama H, Wakabayashi-Takai E, Yoshimoto-Matsuda Y, Nonaka I, Kaneko K, Yoshida M, Ozawa E. Pathological analysis of muscle hypertrophy and degeneration in muscular dystrophy in gamma-sarcoglycan-deficient mice. Neuromuscul Disord 2003;13(3):193.
195. Scata KA, Bernard DW, Fox J, Swain JL. FGF receptor availability regulates skeletal myogenesis. Exp Cell Res 1999;250(1):10-21.
196. Schmalbruch H. Segmental fibre breakdown and defects of the plasmalemma in diseased human muscles. Acta Neuropathol (Berl) 1975;33(2):129-141.
197. Schofield JN, Blake DJ, Simmons C, Morris GE, Tinsley JM, Davies KE, Edwards YH. Apo-dystrophin 1 and apo-dystrophin 2, products of the Duchenne muscular dystrophy locus: expression during mouse embryogenesis and in cultured cell lines. Human Molecular Genetics 1994;3:1309.
198. Schultz E, Jaryszak DL, Valliere CR. Response of satellite cells to focal skeletal muscle injury. Muscle Nerve 1985;8(3):217-222.
199. Schultz E, Albright DJ, Jaryszak DL, David TL. Survival of satellite cells in whole muscle transplants. Anat Rec 1988;222(1):12-17.
200. Schultz E, McCormick KM. Skeletal muscle satellite cells. Rev Physiol Biochem Pharmacol 1994;123:213-257.
201. Sciandra F, Schneider M, Giardina B, Baumgartner S, Petrucci TC, Brancaccio A. Identification of the beta-dystroglycan binding epitope within the C-terminal region of alpha-dystroglycan. Eur J Biochem 2001;268(16):4590-4597.

202. Selcen D, Stilling G, Engel AG. The earliest pathologic alterations in dysferlinopathy. Neurology 2001;56(11):1472-1481.
203. Semsarian C, Wu MJ, Ju YK, Marciniec T, Yeoh T, Allen DG, Harvey RP, Graham RM. Skeletal muscle hypertrophy is mediated by a Ca2+-dependent calcineurin signalling pathway. Nature 1999;400(6744):576-581.
204. Sheehan SM, Tatsumi R, Temm-Grove CJ, Allen RE. HGF is an autocrine growth factor for skeletal muscle satellite cells in vitro. Muscle Nerve 2000;23(2):239-245.
205. Shelton GD, Engvall E. Canine and feline models of human inherited muscle diseases. Neuromuscul Disord 2005;15(2):127-138.
206. Shi W, Chen Z, Schottenfeld J, Stahl RC, Kunkel LM, Chan YM. Specific assembly pathway of sarco glycans is dependent on beta- and delta-sarcoglycan. Muscle Nerve 2004;29(3):409-419.
207. Sicinski P, Geng Y, Ryder-Cook AS, Barnard EA, Darlison MG, Barnard PJ. The molecular basis of muscular dystrophy in the mdx-mouse: a point mutation. Science 1989;244:1578.
208. Simonds AK, Muntoni F, Heather S, Fielding S. Impact of nasal ventilation on survival in hypercapnic Duchenne muscular dystrophy. Thorax 1998;53(11):949-952.
209. Smalheiser NR, Kim E. Purification of cranin, a laminin binding membrane protein. Identity with dystroglycan and reassessment of its carbohydrate moieties. J Biol Chem 1995;270(25):15425-15433.
210. Smith J. Muscle Growth Factors, Ubiquitin and Apoptosis in Dystrophic Muscle: Apoptosis Declines with Age in the mdx Mouse. Basic and Applied Myology 1996;6(4):279-284.
211. Somer H, Donner M, Murros J, Konttinen A. A serum isozyme study in muscular dystrophy. Particular reference to creatine kinase, aspartate aminotransferase, and lactic acid dehydrogenase isozymes. Arch Neurol 1973;29(5):343-345.
212. Song KS, Scherer PE, Tang Z, Okamoto T, Li S, Chafel M, Chu C, Kohtz DS, Lisanti MP. Expression of caveolin-3 in skeletal, cardiac, and smooth muscle cells. Caveolin-3 is a component of the sarcolemma and co- fractionates with dystrophin and dystrophin-associated glycoproteins. Journal of Biological Chemistry 1996;271(25):15160.
213. Sotgia F, Lee JK, Das K, Bedford M, Petrucci TC, Macioce P, Sargiacomo M, Bricarelli FD, Minetti C, Sudol M, Lisanti MP. Caveolin-3 directly interacts with the C-terminal tail of beta -dystroglycan. Identification of a central WW-like domain within caveolin family members. J Biol Chem 2000;275(48):38048-38058.
214. Spencer MJ, Walsh CM, Dorshkind KA, Rodriguez EM, Tidball JG. Myonuclear apoptosis in dystrophic mdx muscle occurs by perforin-mediated cytotoxicity. J Clin Invest 1997;99(11):2745-2751.
215. Straub V, Rafael JA, Chamberlain JS, Campbell KP. Animal models for muscular dystrophy show different patterns of sarcolemmal disruption. J Cell Biol 1997;139(2):375-385.
216. t Hoen PA, de Kort F, van Ommen GJ, den Dunnen JT. Fluorescent labelling of cRNA for microarray applications. Nucleic Acids Res 2003;31(5):e20.
217. Tajbakhsh S, Buckingham M. The birth of muscle progenitor cells in the mouse: spatiotemporal considerations. CP O, editor. London: Academic Press; 2000. pp. 225-268 p.
218. Takada F, Vander Woude DL, Tong HQ, Thompson TG, Watkins SC, Kunkel LM, Beggs AH. Myozenin: an alpha-actinin- and gamma-filamin-binding protein of skeletal muscle Z lines. Proc Natl Acad Sci U S A 2001;98(4):1595-1600.
219. Tang Z, Scherer PE, Okamoto T, Song K, Chu C, Kohtz DS, Nishimoto I, Lodish HF, Lisanti MP. Molecular cloning of caveolin-3, a novel member of the caveolin gene family expressed predominantly in muscle. J Biol Chem 1996;271(4):2255-2261.
220. Tews DS. Apoptosis and muscle fibre loss in neuromuscular disorders. Neuromuscul Disord 2002; 12(7-8):613-622.
221. Tews DS. Role of Apoptosis in Myopathies. Basic and Applied Myology 2003;13(4):181-190.
222. Tidball JG, Albrecht DE, Lokensgard BE, Spencer MJ. Apoptosis precedes necrosis of dystrophin-deficient muscle. J Cell Sci 1995;108 (Pt 6):2197-2204.
223. Timpl R, Brown JC. Supramolecular assembly of basement membranes. Bioessays 1996;18(2):123-132.
224. Tinsley JM, Blake DJ, Roche A, Fairbrother U, Riss J, Byth BC, Knight AE, Kendrick-Jones J, Suthers GK, Love DR, Edwards YH, Davies KE. Primary structure of dystrophin-related protein. Nature 1992;360:591.
225. Tsujimoto Y. Role of Bcl-2 family proteins in apoptosis: apoptosomes or mitochondria? Genes Cells 1998;3(11):697-707.

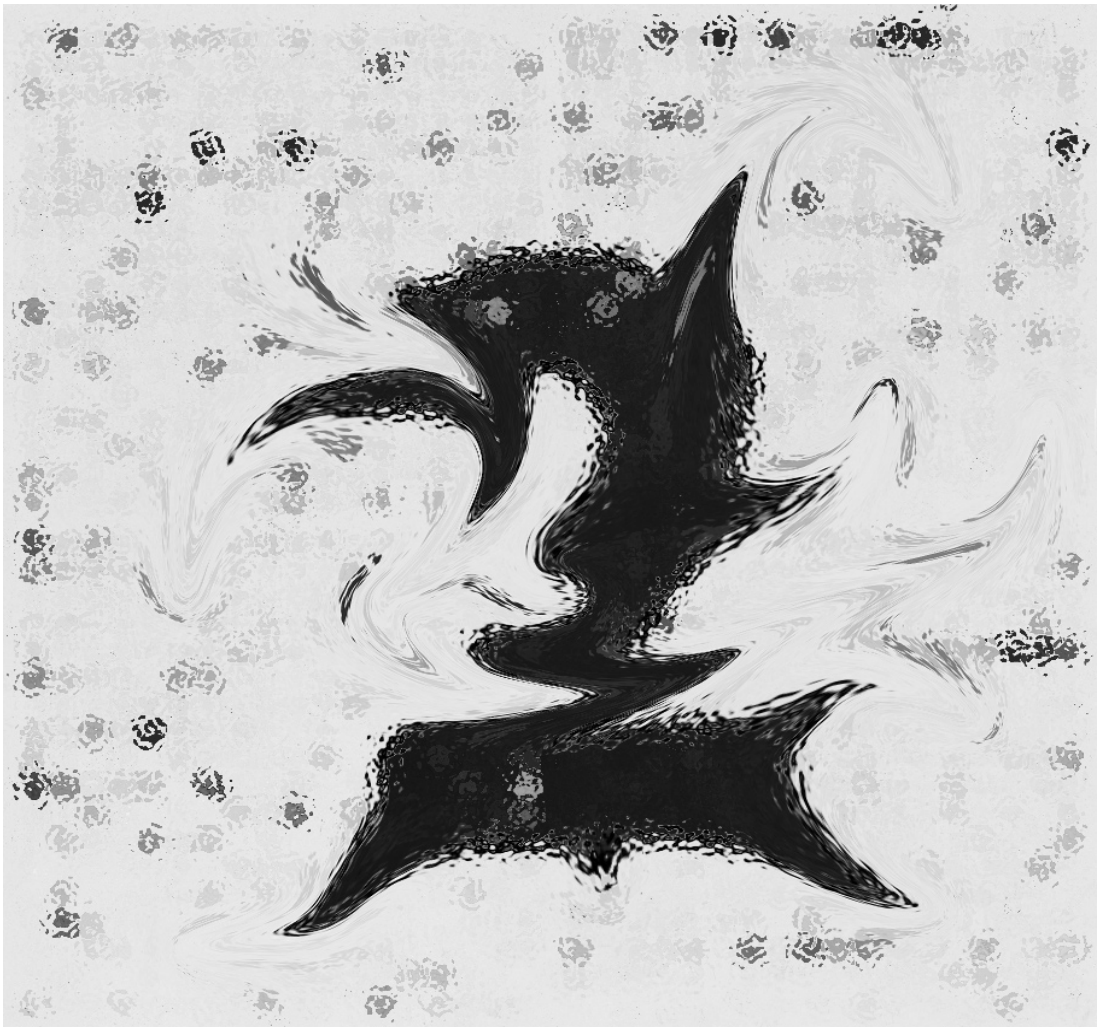
226. Ueda H, Ueda K, Baba T, Ohno S. delta- and gamma-Sarcoglycan localization in the sarcoplasmic reticulum of skeletal muscle. J Histochem Cytochem 2001;49(4):529-538.
227. Vainzof M, Passos-Bueno MR, Canovas M, Moreire\a ES, Pavanello RCM, Marie SK, Anderson LVB, Bonnemann CG, McNally EM, Nigro V, Kunkel LM, Zatz M. The sarcoglycan complex in the six autosomal recessive limb-girdle muscular dystrophies. Human Molecular Genetics 1996;5(12):1963.
228. Vainzof M, Moreira ES, Ferraz G, Passos-Bueno MR, Marie SK, Zatz M. Further evidence for the organisation of the four sarcoglycans proteins within the dystrophin-glycoprotein complex. Eur J Hum Genet 1999;7(2):251-254.
229. Valentine BA, Winand NJ, Pradhan D, Moise NS, de Lahunta A, Kornegay JN, Cooper BJ. Canine X-linked muscular dystrophy as an animal model of Duchenne muscular dystrophy: a review. Am J Med Genet 1992;42(3):352-356.
230. van der Ven PF, Wiesner S, Salmikangas P, Auerbach D, Himmel M, Kempa S, Hayess K, Pacholsky D, Taivainen A, Schroder R, Carpen O, Furst DO. Indications for a novel muscular dystrophy pathway. gamma-filamin, the muscle-specific filamin isoform, interacts with myotilin. J Cell Biol 2000;151(2):235-248.
231. Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. Serial analysis of gene expression. Science 1995;270:484.
232. Venema VJ, Ju H, Zou R, Venema RC. Interaction of neuronal nitric-oxide synthase with caveolin-3 in skeletal muscle. Identification of a novel caveolin scaffolding/inhibitory domain. J Biol Chem 1997;272(45):28187-28190.
233. Vite CH. Myopathic Disorders. Vite CH, editor. Ithaca, New York: International Veterinary Information Service; 2005.
234. Voit T, Tome FMS. The Congenital Muscular Dystrophies. In: Engel AG, Franzini-Armstrong C, editors. Myology. 3rd ed. Volume 2. New York: McGraw-Hill; 2004. p 1203-1238.
235. Vorgerd M, Gencik M, Mortier J, Epplen JT, Malin JP, Mortier W. Isolated loss of gamma-sarcoglycan: diagnostic implications in autosomal recessive limb-girdle muscular dystrophies. Muscle Nerve 2001;24(3):421-424.
236. Vos JH, van der Linde-Sipman JS, Goedegebuure SA. Dystrophy-like myopathy in the cat. J Comp Pathol 1986;96(3):335-341.
237. Walton JN, Race RR, Philip U. On the inheritance of muscular dystrophy; with a note on the blood groups, and a note on colour vision and linkage studies. Ann Hum Genet 1955;20(1):1-38.
238. Washburn MP, Koller A, Oshiro G, Ulaszek RR, Plouffe D, Deciu C, Winzeler E, Yates JR, 3rd. Protein pathway and complex clustering of correlated mRNA and protein expression analyses in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A 2003;100(6):3107-3112.
239. Weller B, Karpati G, Carpenter S. Dystrophin-deficient mdx muscle fibers are preferentially vulnerable to necrosis induced by experimental lengthening contractions. J Neurol Sci 1990;100(1-2):9-13.
240. Wheeler MT, Zarnegar S, McNally EM. Zeta-sarcoglycan, a novel component of the sarcoglycan complex, is reduced in muscular dystrophy. Hum Mol Genet 2002;11(18):2147-2154.
241. Wigmore PM, Dunglison GF. The generation of fiber diversity during myogenesis. Int J Dev Biol 1998;42(2):117-125.
242. Wigmore PM, Evans DJ. Molecular and cellular mechanisms involved in the generation of fiber diversity during myogenesis. Int Rev Cytol 2002;216:175-232.
243. Wilson SJ, McEwan JC, Sheard PW, Harris AJ. Early stages of myogenesis in a large mammal: formation of successive generations of myotubes in sheep tibialis cranialis muscle. J Muscle Res Cell Motil 1992;13(5):534-550.
244. Wokke JH, Van den Oord CJ, Leppink GJ, Jennekens FG. Perisynaptic satellite cells in human external intercostal muscle: a quantitative and qualitative study. Anat Rec 1989;223(2):174-180.
245. Yang B, Jung D, Motto D, Meyer J, Koretzky G, Campbell KP. SH3 domain-mediated interaction of dystroglycan and Grb2. J Biol Chem 1995;270(20):11711-11714.
246. Yoshida M, Suzuki A, Yamamoto H, Noguchi S, Mizuno Y, Ozawa E. Dissociation of the complex of dystrophin and its associated proteins into several unique groups by n-octyl beta-D-glucoside. Eur JBiochem 1994;222:1055.
247. Zatz M, Vainzof M, Passos-Bueno MR. Limb-girdle muscular dystrophy: one gene with different phenotypes, one phenotype with different genes. Curr Opin Neurol 2000;13(5):511-517.
248. Zentella A, Massague J. Transforming growth factor beta induces myoblast differentiation in the pres ence of mitogens. Proc Natl Acad Sci U S A 1992;89(11):5176-5180.

Chapter 2

Gene expression variation between mouse inbred strains

Turk R, ‘t Hoen PA, Sterrenburg E, de Menezes RX,
de Meijer EJ, Boer JM, van Ommen GJ, den Dunnen JT

BMC Genomics 2004 5(1):57



BMC Genomics

BioMed Central

Research article

Open Access

Gene expression variation between mouse inbred strains

Rolf Turk[1], Peter AC 't Hoen[1], Ellen Sterrenburg[1], Renée X de Menezes[2], Emile J de Meijer[1], Judith M Boer[1], Gert-Jan B van Ommen[1] and Johan T den Dunnen*[1]

Address: [1]Center for Human and Clinical Genetics, Leiden University Medical Center, Wassenaarseweg 72, 2333AL Leiden, Nederland and [2]Department of Medical Statistics, Leiden University Medical Center, Wassenaarseweg 72, 2333AL Leiden, Nederland

Email: Rolf Turk - r.turk@lumc.nl; Peter AC 't Hoen - p.a.c.hoen@lumc.nl; Ellen Sterrenburg - e.sterrenburg@lumc.nl; Renée X de Menezes - r.x.menezes@lumc.nl; Emile J de Meijer - e.j.de_meijer@lumc.nl; Judith M Boer - j.m.boer@lumc.nl; Gert-Jan B van Ommen - g.j.b.van_ommen@lumc.nl; Johan T den Dunnen* - j.t.den_dunnen@lumc.nl

* Corresponding author

Published: 18 August 2004

BMC Genomics 2004, **5**:57 doi:10.1186/1471-2164-5-57

Received: 13 May 2004

Accepted: 18 August 2004

This article is available from: http://www.biomedcentral.com/1471-2164/5/57

Abstract

Background: In this study, we investigated the effect of genetic background on expression profiles. We analysed the transcriptome of mouse hindlimb muscle of five frequently used mouse inbred strains using spotted oligonucleotide microarrays.

Results: Through ANOVA analysis with a false discovery rate of 10%, we show that 1.4% of the analysed genes is significantly differentially expressed between these mouse strains. Differential expression of several of these genes has been confirmed by quantitative RT-PCR. The number of genes affected by genetic background is approximately ten-fold lower than the number of differentially expressed genes caused by a dystrophic genetic defect.

Conclusions: We conclude that evaluation of the effect of background on gene expression profiles in the tissue under study is an effective and sensible approach when comparing expression patterns in animal models with heterogeneous genetic backgrounds. Genes affected by the genetic background can be excluded in subsequent analyses of the disease-related changes in expression profiles. This is often a more effective strategy than backcrossing and inbreeding to obtain isogenic backgrounds.

Background

Due to their isogenicity, inbred mouse strains demonstrate low biological variability within each strain[1,2]. Genetic variation between inbred strains is considerable and has recently been characterized in detail using single nucleotide polymorphisms[3]. Differences in genetic background between strains affect the gene expression levels of a subset of genes, which probably explains phenotypic differences. Indeed, several reports have been published in which gene expression profiles have been used as QTLs in genetic mapping studies to identify complex traits [4-6].

From literature [7-9], it appears that the subset of genes for which expression is significantly affected by genetic background is small. However, this has never been related to the extent of gene expression changes observed due to disease-causing mutations. We are studying differential gene expression between affected and healthy muscle in a range of murine models for neuromuscular disorders with

BMC Genomics 2004, **5**:57 http://www.biomedcentral.com/1471-2164/5/57

different genetic backgrounds (Turk *et al.*, manuscript in preparation). We, therefore, determined gene expression levels in hindlimb muscles from five frequently used wildtype mouse inbred strains, and compared these to the differential gene expression levels in affected muscle tissue from a mouse model (*mdx*) for Duchenne muscular dystrophy with healthy muscle tissue. Both the number of differentially expressed genes between strains as well as the fold-change levels are lower when compared to the differences found in affected versus healthy muscle tissue.

Results

Gene expression levels in hindlimb muscle tissue from five different inbred strains (CBA, BALB, BL6, DBA, and BL10) were determined. Total RNA from two individuals per strain was isolated, reversed transcribed, and subsequently labelled according to a recently developed protocol (adapted from Xiang *et al.*, 2002), which requires an input of only 1 μg total RNA. Labelled cDNA was hybridised to murine microarrays containing 7,776 65-mer oligonucleotides spotted in duplicate.

Significance levels (p-values) between the five mouse inbred strains were calculated using analysis of variance[10]. Significance levels among two individual mice within each strain were determined using a hierarchical *t*-test providing higher statistical power than conservative methods for low (2–4) replicate numbers[11]. The higher power is yielded by borrowing information across genes to produce a better expression variance estimator. The gain in power is reported via an increase in the degrees of freedom associated with the t-test. Differentially expressed genes for both computations were selected by controlling the false discovery rate (FDR), as suggested by Benjamini and Hochberg (1995), rather than using pre-defined cut-offs for p-values or corrections for multiple testing. The FDR represents an expectation of the proportion of false positives among the selected differentially expressed genes, which increases dramatically during multiple testing, inherent in microarray experiments[12].

Using an FDR of 10% we selected 88 out of 6144 (1.4%) expressed genes that are differentially expressed between strains (Fig. 1). A lower number of differentially expressed genes was found in the analysis of variation within strains with identical FDRs of 10% (Table 1). Results with other FDR levels are available online as additional file. Correlation between gene expression levels of the two samples from each strain was high (Pearson correlation coefficient ranging from 0.87 to 0.95), also indicating low internal variation (Table 1). A considerable amount of differentially expressed genes (718 genes) were selected when predefined cut-off values ($p < 0.05$) were used to determine the differential gene expression between strains. However, adjusted FDR levels indicated a proportion of false positives equal to 42%. On the other hand, adjusting for multiple testing using Bonferroni correction proved to be too stringent, leaving no or few differentially expressed genes. Controlling the FDR, therefore, appears to be an optimal method for both selecting differential gene expression and simultaneously determining the validity of the experimental outcome.

To put the influence of differential gene expression due to genetic background in perspective, we studied gene expression between affected and healthy tissue from hindlimb muscle derived from *mdx* mice, and from control mice with identical genetic backgrounds. Selection with an FDR of 10% resulted in 1298 differentially expressed genes. Differential gene expression between the two most divergent mouse inbred strains (BL6 and CBA, data not shown) was determined to allow a direct comparison with identical statistical methods. Selection with an FDR of 10% showed an approximately ten-fold decrease in the number of differentially expressed genes (126). Absolute fold changes were calculated and subsequently a comparison of the distribution was made (Fig. 2). Median gene expression levels are equal between affected/control and inbred/inbred. However, the number of large fold changes (>3) between affected/healthy (221) is much higher than between inbred/inbred (7), consistent with low contribution of differential expression due to genetic background.

Although overall expression levels are similar between strains, a relatively high number of differentially expressed genes was due to deviating gene expression levels in BL6. We performed quantitative real-time RT-PCR (qPCR) on five genes to verify our microarray data. Two genes myomesin 1 and tropomodulin 1, which were 2.2-fold and 1.8-fold lower expressed in BL6 compared to the other strains on our microarrays, were also found to be lower expressed (2.0-fold and 2.2-fold respectively) in our qPCR assay (Fig. 3). Three other genes (dysferlin, cystatin B, and thrombospondin 4) showed no differential expression between any strains.

Discussion

This study shows that variation in overall gene expression levels between mouse inbred strains is relatively low in hindlimb muscle tissue. This is particularly evident when the number of differentially expressed genes between two mouse inbred strains (C57 vs. Bl6, 126 genes with 7 genes having a fold-change > 3) is compared to that between diseased and healthy muscle tissue (*mdx* vs. wild-type, 1298 genes with 221 genes having a fold-change >3). Therefore, the use of mice with deviating genetic background may be justified in disease-related studies. Alternatively, strain-dependent gene expression differences

BMC Genomics 2004, **5**:57 http://www.biomedcentral.com/1471-2164/5/57

AJ290947	Stretch regulated skeletal muscle protein (Usmg5)	signal transducer	Mm.29722
X70032	Nebulin	muscle development	
U51744	TEF-1-related factor (FR-19)	biological_process unknown	
AF203898	Nebulin, partial cds	muscle development	
AJ290946	Stretch regulated skeletal muscle protein (Usmg4)	biological_process unknown	Mm.27881
M13967	4.5S RNA	biological_process unknown	
NM_009041	Radixin (Rdx)	cytoskeletal anchoring	Mm.7051
NM_007993	Fibrillin 1 (Fbn1)	skeletal development	Mm.735
NM_009429	Translationally regulated transcript (21 kDa) (Trt)	protein folding	Mm.254
NM_011671	Uncoupling protein 2, mitochondrial (Ucp2)	ATP/ADP exchange	Mm.144413
AF248643	Alpha-actinin 2 mRNA, complete cds	invasive growth	Mm.10711
U27315	Adenine nucleotide translocase-1 (Ant1)	ATP/ADP exchange	Mm.16228
NM_008118	Gastric intrinsic factor (Gif)	cobalt ion transport	Mm.456
NM_011694	Voltage-dependent anion channel 1 (Vdac1)	apoptotic program	Mm.3555
NM_010867	Myomesin 1 (Myom1)	muscle development	Mm.4103
U58108	Nebulin, partial cds	muscle development	
X73876	Phka1 unprocessed pre-Mrna	biological_process unknown	
AF059029	Calcium/calmodulin-dependent protein kinase II delta	learning and memory	Mm.34377
NM_010743	Lymphocyte antigen 84 (Ly84)	cell surface receptor linked signal transduction	Mm.35692
AF069744	Aip1 mRNA, complete cds	biological_process unknown	Mm.42086
NM_009299	Seminal vesicle antigen (Sva)	transmembrane receptor protein tyrosine phosphatase signaling pathway	Mm.4119
NM_010801	Myeloid leukemia factor 1 (Mlf1)	GTPase activator	Mm.10414
NM_008794	Proprotein convertase subtilisin/kexin type 7 (Pcsk7)	proteolysis and peptidolysis	Mm.3255
NM_016921	ATPase, H+ transporting, lysosomal I (Atp6i)	positive control of cell proliferation	Mm.19185
X83932	Ryanodine receptor type 1	calcium ion transport	Mm.4519
NM_008974	Protein tyrosine phosphatase 4a2 (Ptp4a2)	DNA replication	Mm.6355
NM_011236	RAD52 homolog, (S. cerevisiae) (Rad52)	DNA repair	Mm.149
X58196	H19	biological_process unknown	
AF150755	Microtubule-actin crosslinking factor (Macf)	cytoskeletal anchoring	Mm.3350
NM_009745	B-cell CLL/lymphoma 7B (Bcl7b)	biological_process unknown	Mm.31
NM_019628	Proline-rich protein PRP MP5 (PRPMP5)	cell growth and/or maintenance	Mm.4491
NM_007734	Procollagen, type IV, alpha 3 (Col4a3)	cell adhesion	Mm.8069
NM_009757	Bone morphogenetic protein 15 (Bmp15)	oogenesis	Mm.42160
NM_009694	Apolipoprotein B editing complex 2 (Apobec2)	mRNA editing	Mm.27822
NM_010592	Jun proto-oncogene related gene d1 (Jund1)	transcription regulation	Mm.1175
NM_009758	Bone morphogenetic protein receptor, type 1A (Bmpr1a)	TGFbeta receptor signaling pathway	
NM_007808	Cytochrome c, somatic (Cycs)	electron transport	Mm.35389
NM_010174	Fatty acid binding protein 3, muscle and heart (Fabp3)	negative control of cell proliferation	Mm.22220
AJ250691	sr513 protein	biological_process unknown	Mm.52047
AB034801	mRNA, complete cds, similar to L-amino acid oxidase	apoptosis	Mm.26698
NM_007959	ets related protein 71 (Etsrp71)	transcription regulation	Mm.4829
AF038896	Cyclic AMP specific phosphodiesterase	olfaction	
NM_020582	hypothetical protein, clone:2-31 (LOC57423)	hydrogen transport	Mm.1674
NM_011014	Opioid receptor, sigma 1 (Oprs1)	steroid biosynthesis	Mm.22745
NM_007385	Apolipoprotein C2 linked (Acl)	biological_process unknown	Mm.8239
NM_021883	Tropomodulin 1 (Tmod1)	actin filament organization	Mm.139690
NM_020045	HIRA-interacting protein 5 (HIRIP5)	iron homeostasis	Mm.23809
NM_010518	Insulin-like growth factor binding protein 5 (Igfbp5)	receptor	Mm.578
AL357197	EST from clone 640060, 5' end	biological_process unknown	
NM_010595	Potassium voltage gated channel, shaker related subfamily, member 1 (Kcna1)	potassium transport	Mm.2493
NM_008539	MAD homolog 1 (Drosophila) (Madh1)	common-partner SMAD protein phosphorylation	Mm.15185
NM_018780	Secreted frizzled-related sequence protein 5 (Sfrp5)	establishment and/or maintenance of cell polarity	Mm.103731
AF037256	ES2 protein (Es2)	leading strand elongation	Mm.10094
X78666	Ryr1 (BALB/c)	calcium ion transport	
NM_009760	BCL2/adenovirus E1B 19 kDa-interacting protein 1, NIP3 (Bnip3)	apoptosis	Mm.2159
NM_021041	ATP-binding cassette, sub-family C (CFTR/MRP), member 9 (Abcc9), transcript variant 2	potassium transport	Mm.35670
AB001456	Motor domain of KIF1C, partial cds	microtubule-based process	MM.90223
NM_009553	Zinc finger protein 29 (Zfp29)	transcription regulation	Mm.5010
NM_011732	Y box protein 1 (Ybx1)	response to pest/pathogen/parasite	Mm.21054
U25158	Dystonin isoform 2	developmental processes	
NM_011179	Prosaposin (Psap)	respiration	Mm.3363
X99751	HK11 gene, 3' UTR	biological_process unknown	
J02935	cAMP-dependent protein kinase type II regulatory subunit mRNA, 3' end	intracellular signaling cascade	Mm.3176
NM_016970	Killer cell lectin-like receptor subfamily G, member 1 (Klrg1)	cell surface receptor linked signal transduction	Mm.20434
U02971	2-oxoglutarate dehydrogenase E1 component (BALB/c)	glycolysis	
NM_009490	Vomeronasal organ family 2, receptor, 15 (V2r15)	ossification	Mm.34447
AF178752	Odorant receptor 83 (C57BL)	olfaction	
AB010374	mszf89	transcription regulation	
NM_010638	Kruppel-like factor 9 (Klf9)	transcription regulation	Mm.19788
L13171	Myocyte-specific enhancer factor 2 (MEF-2C) mRNA sequence	transcription regulation	Mm.24001
NM_009161	Sarcoglycan, alpha (50kDa dystrophin-associated glycoprotein) (Sgca)	muscle development	Mm.18709
NM_010412	Histone deacetylase 5 (Hdac5)	repression of transcription from Pol II promoter	Mm.12665
NM_009448	Tubulin alpha 6 (Tuba6)	nuclear congression	Mm.88212
NM_020604	Junctophilin 1 (Jp1-pending)	protein folding	Mm.143761
AB041351	Col4a6 mRNA for type IV collagen alpha 6 chain	cell adhesion	Mm.155586
NM_007969	Extracellular proteinase inhibitor (Expi)	developmental processes	Mm.1650
AB010370	mszf80-2	transcription regulation	
D87910	PA28 beta subunit	cell adhesion	Mm.626
AF084364	D6MM5E protein (D6Mm5e)	mitochondrial processing	Mm.100652
NM_007739	Procollagen, type VIII, alpha 1 (Col8a1)	cell adhesion	Mm.86813
NM_011258	Replication factor C, 140 kDa (Recc1)	DNA replication	Mm.148877
NM_016899	RAB25, member RAS oncogene family (Rab25)	intracellular protein traffic	Mm.26994
NM_021549	Polynucleotide kinase 3'- phosphatase (Pnkp)	histidine metabolism	Mm.29545
NM_011184	Proteasome (prosome, macropain) subunit, alpha type 3 (Psma3)	ubiquitin-dependent protein degradation	Mm.1007
NM_021311	MIWI protein (MIWI)	cell growth and/or maintenance	Mm.87282
NM_013555	Homeo box D9 (Hoxd9)	transcription regulation	Mm.57122
NM_011966	Proteasome (prosome, macropain) subunit, alpha type 4 (Psma4)	ubiquitin-dependent protein degradation	Mm.30270
NM_020506	Exportin 4 (Xpo4-pending)	protein modification	Mm.143753

Figure 1

Differentially expressed genes between mouse inbred strains Relative expression levels of differentially expressed genes between mouse inbred strains are depicted in colour as relative intensity levels. Shown for each gene are GenBank accession number, description, functional annotation according to Gene Ontology, and UniGene cluster IDs. Relative expression levels are calculated by subtracting the average intensity value per gene from the strain-dependent intensity values. Differential expression was determined by selecting p-values from analysis of variance based on a false discovery rate of 10%.

BMC Genomics 2004, **5**:57 http://www.biomedcentral.com/1471-2164/5/57

Table 1: Number of differentially expressed genes using several cut-off strategies

	Between strains MA-ANOVA	**Within strains** Hierarchical *t*-test				
		CBA	**BL10**	**BL6**	**DBA**	**BALB**
Correlation		0.95	0.95	0.87	0.87	0.92
Naive ($p < 0.05$)	718	737	610	963	1043	483
Bonferroni	0	2	4	1	0	3
FDR 10%	88	2	4	14	0	16

Correlation between two individuals per strain was calculated using Pearson's correlation coefficient. Significance levels (p-values) between strains were calculated with MA-ANOVA, and within strains using the hierarchical *t*-test. Differential gene expression was determined by selecting genes with p-values lower than a specified threshold. Thresholds were selected using three different strategies; naive, Bonferroni corrected, and False Discovery Rate (10%), and resulted in different numbers of significantly differentially expressed genes.

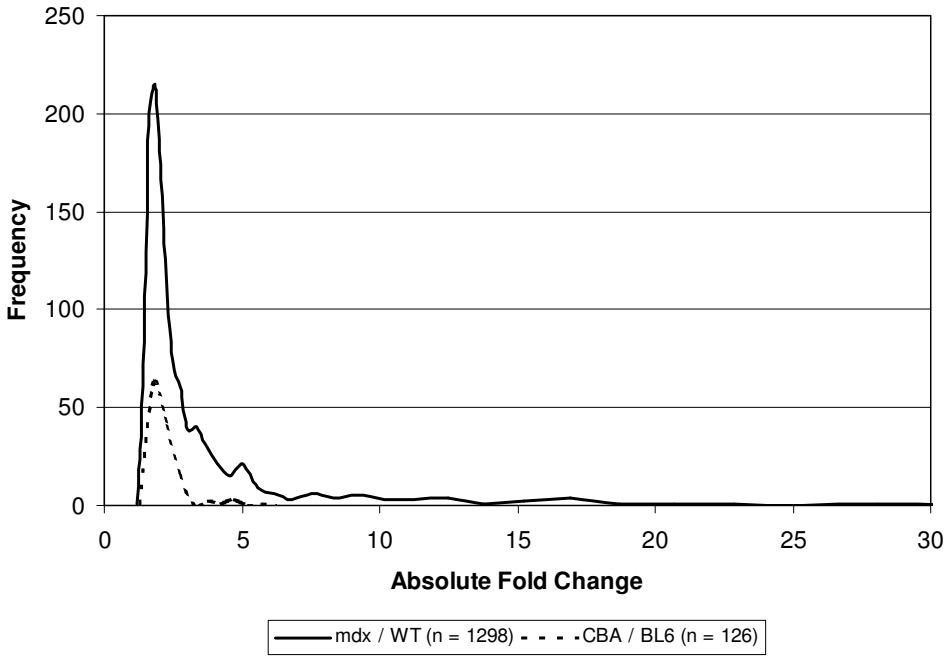


Figure 2
Effect of different genetic background on differential gene expression The distribution of absolute fold changes of differentially expressed genes (n = 1298) between affected (*mdx*) and healthy (WT) muscle were compared to the distribution of absolute fold changes of differentially expressed genes (n = 126) between two mouse inbred strains (CBA and BL6). Selections were based on a FDR of 10%.

BMC Genomics 2004, **5**:57 http://www.biomedcentral.com/1471-2164/5/57

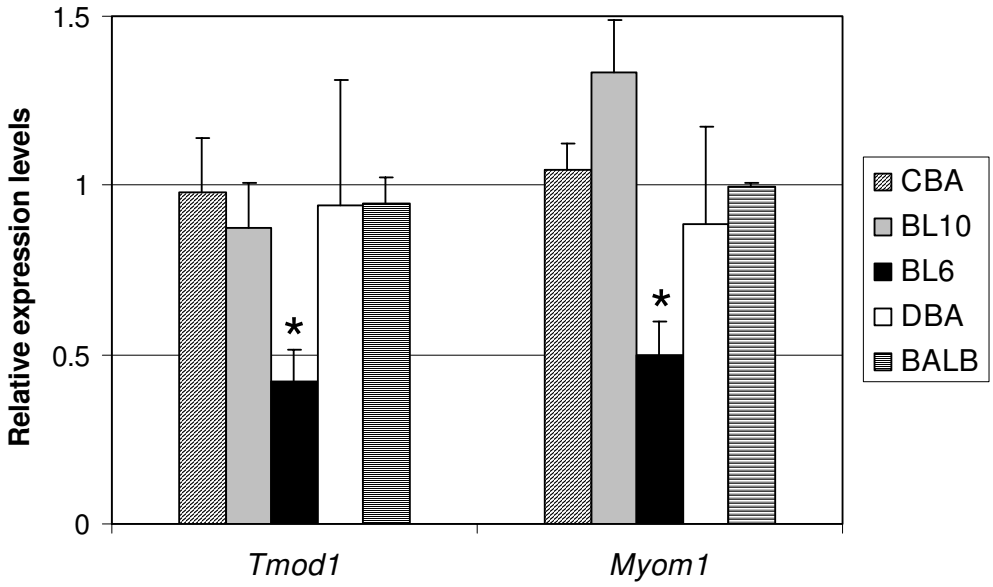


Figure 3
Validation of BL6-dependent gene expression with qPCR Relative gene expression levels between mouse inbred strains of tropomodulin 1 (*Tmod1*) and myomesin 1 (*Myom1*) as determined by quantitative RT-PCR. Significantly lower expression ($p < 0.01$, marked by *) for both genes was shown in BL6 compared to other strains.

may be evaluated in the initial study phase of gene targeting experiments, although the effect of hybrid backgrounds is difficult to assess.

Gene expression studies in the brain revealed that approximately 1% of expressed genes differ between two mouse strains[8]. Application of alternative statistical methods, similar to those used in our study, on this dataset resulted in an increase in the number of differentially expressed genes (approx. 3%) between the two mouse strains[7], demonstrating that the number of differentially expressed genes is highly dependent on the statistical criteria used. A similar number of differentially expressed genes was found in a comparison of hippocampal gene expression between 8 different mouse strains[9]. The results of our study in muscle tissue demonstrated that approximately 1.4% of the expressed genes show differential expression between mouse strains. Based on these results, strain differences in gene expression seem to have a similar magnitude across different tissues.

Genomic variability could be correlated with high levels of single nucleotide polymorphisms (SNPs) occurring in specific blocks between mouse inbred strains. The presence of cis-acting (single nucleotide) polymorphisms may be associated with regulatory variation affecting gene expression levels. It was estimated that probably a consistent amount (up to 6%) of the roughly estimated 35,000 mouse genes contain such functional regulatory variants[13]. We investigated if differentially expressed genes were localized in blocks with high genomic variability, but our number of differentially expressed genes was too low to obtain statistically significant answers (data not shown).

This study suggests an additional method for phenotyping mouse inbred strains and provides a list of genes with significant differential expression based upon false discovery rate selection. Although overall gene expression profiles are highly similar, most significant differences are determined by low gene expression levels of BL6 compared to the other strains. A large proportion of these BL6-specific genes function as structural muscle proteins (i.e. nebulin, alpha-actinin 2, myomesin 1 and radixin). To date, however, no major differences in muscle physiology in BL6-mice have been described which can be attributed to these reduced gene expression levels.

Perfectly isogenic backgrounds are sometimes difficult to obtain. This explorative study demonstrates that the effect of genetic background on muscle expression profiles is significant but rather limited compared to other effects, e.g. the dystrophic genetic defect (*mdx*) we study. As such, the genetic background will only marginally interfere with data analysis. Determination of gene expression profiles between mouse strains enables flagging a modest number of differentially expressed genes, and is an efficient and sensible approach to circumvent tedious backcrossings, necessary to obtain isogenic animals.

Methods

Mouse breeding, tissue preparation and total RNA isolation

We obtained CBA/CaOlaHsd (CBA), BALB/cOlaHsd (BALB), C57Bl/6JOlaHsd (BL6), DBA/2OlaHsd (DBA), and C57Bl/10ScSnOlaHsd (BL10) mice from Harland Laboratories, and C57Bl/10ScSn-Dmd$^{mdx/J}$ (*mdx*) mice from Jackson Laboratory at the age of 6 weeks. Mice were kept under standard conditions and were sacrificed by cervical dislocation when 8 weeks old. Hindlimb muscles (m. quadriceps femoris) were dissected and promptly snap-frozen in isopentane at -80°C. Total RNA was prepared by disrupting tissue using mortar and pestle and subsequent homogenisation by a rotor-stator homogenizor (Ultra-Turrax T25, Janke & Kunkel IKA-Labortechnik) in RNA-Bee (Campro Scientific) until uniformly homogenous (15–45 sec). Total RNA was isolated according to manufacturer's instructions followed by purification using RN-easy columns (Qiagen). Quality and yield was determined using Lab-on-a-chip (BioAnalyzer, Agilent).

Target preparation and hybridisation

Aminoallyl labelled cDNA (aa-cDNA) was prepared based on a previously described protocol[14]. Aliquots of 1 μg of total RNA in the presence of 2 μg amino-TN_6 primer (5'-NH_2-$(CH_2)_6$-TN_6, Eurogentec) were adjusted to a volume of 21 μl with DEPC-treated H_2O (diethyl pyrocarbonate, Sigma), heated for 10 minutes at 70°C and chilled on ice for 10 minutes. Reverse transcription mastermix (1.8 μl RevertAid RNaseH-M-MuLV reverse transcriptase (200 U/μl, MBI Fermentas), 6 μl 5x first-strand buffer (MBI Fermentas), and 1.2 μl 25x aa-dUTP / dNTP solution (2 μl 50 mM dATP, 2 μl 50 mM dCTP, 2 μl 50 mM dGTP, 1.2 μl 50 mM dTTP, 0.8 μl 50 mM aminoallyl-dUTP (Ambion)) was added per reaction and incubated at room temperature for 10 minutes followed by 2 hours at 42°C. RNA was hydrolysed by addition of 10 μl 0.5 M EDTA and 10 μl 1 M NaOH and incubation at 65°C for 30 minutes followed by neutralization by addition of 10 μl 1 M HCl. Aminoallyl labelled cDNA was then purified by combining 300 μl of PB-buffer (Qiagen) to 60 μl of the neutralized sample and centrifuged through a Qiaquick column (Qiagen) at 13000 rpm for 1 minute. Two washing steps were performed by spinning 500 μl of 75% EtOH at 13000 rpm for 1 minute while discarding the flow-through. To remove ethanol-traces the columns were centrifuged for an additional minute. cDNA was recovered by eluting three times using 30 μl basic H_2O (3.3 mM $NaHCO_3$ buffer, pH 9.0) and concentrated to a volume of 6.66 μl using a speedvac. Aliquots of Cy3 and Cy5 reactive dyes (PA23001, PA25001, Amersham) were prepared by dissolving each vial of monoreactive dye in 40 μl fresh anhydrous DMSO (Sigma) and dividing into aliquots of 2 μl followed by vaccuumdrying until dry and subsequent storage at 4°C in the presence of silica. Fluorescent dyes were coupled by adding 3.33 μl of bicarbonate buffer (1 M $NaHCO_3$ buffer, pH 9.0) to the aa-cDNA sample and dissolving the dried aliquot of reactive dye, followed by incubation at room temperature for 1 hour in the dark. To the samples 4.5 μl 4 M hydroxylamine (Sigma) was added and incubated at room temperature in the dark for 15 minutes, followed by addition of 186 μl TE^{-3}-buffer. Hybridisation mixtures were prepared by combining a Cy3-labeled cDNA sample with a Cy5-labeled cDNA sample and 10 μl Mouse-Hybloc (1 μg/μl, Applied Genetics Laboratories) followed by removing uncoupled dyes by spinning through a pre-wetted Microcon column (YM30, Amicon) for 8 minutes at 13000 rpm. Hybridisation mixture was washed by spinning 500 μl TE^{-3}-buffer through the column and discarding the flow-through. This step was repeated two times as 2 μl yeast-tRNA (10 μg/μl, Sigma) and 2 μl polyA-RNA (10 μg/μl, Sigma) were added during the last step. Mixture was collected by inverting the column and spinning for 1 minute at 13000 rpm. Hybridisation mixture was finalized by adding TE^{-3}-buffer to 84 μl together with 17 μl 20x SSC and 3 μl 10% SDS followed by denaturing at 100°C for 2 minutes, renaturing at room temperature for 15 minutes and spinning at 13000 rpm for 10 minutes. Labelled target was hybridised overnight on murine oligonucleotide microarrays (65-mer with 5'-hexylaminolinker, Sigma-Genosys mouse 7.5 K oligonucleotide library, spotted in duplicate). Hybridisation occurred in a automatic hybridisation station (GeneTac, Perkin Elmer) and was followed by washing with 5x

2xSSC + 0.1% SDS at 30°C, 5x 1xSSC at 30°C, 3x 0.2xSSC at 30°C, 1x 0.2xSSC at 65°C, 2x 0.2xSSC at 30°C, and subsequently scanned as described previously[15].

Experimental design, data extraction and analysis

Gene expression profiles from hindlimb muscle derived from 2 male animals of each strain were generated using dye-swap experiments. Subsequent duplicate spots on each array resulted in 8 replicate measurements per gene. Targets were assigned at random to the arrays, while avoiding co-hybridisation of samples from the same strain. GenePix Pro 3.0 (Axon) was used for feature extraction and quantification. Genes were considered as being expressed when the corresponding feature was not flagged by the algorithm provided by GenePix. Local background corrected spot intensities were normalized using Variance Stabilization and Normalization (VSN) in R [16]. Array data has been made available through the GEO data repository of the National Center for Biotechnology Information under series GSE662. Correlation between individuals was calculated using Pearson's correlation coefficient. Significantly differential expression levels were determined using MA-ANOVA (**MAANOVA2.0 The Jackson Laboratory** http://www.jax.org/staff/churchill/labsite/software/anova/), hierarchical *t*-test [11] and the False Discovery Rate [17] selection procedure.

Quantitative Reverse Transcription Polymerase Chain Reaction

qPCR was performed in duplicate for each individual resulting in four measurements per strain per gene. cDNA was prepared by reverse transcription using 1 μg total RNA as template. Random hexamers (40 ng) were used to prime the transcription after heating 10 minutes at 70°C followed by chilling on ice for 10 minutes. cDNA was synthesized by RevertAid RNaseH$^-$ MuLV reverse transcriptase and accompanying buffer (MBI-Fermentas) using 1 mM dNTPs. The mixture was incubated at room temperature for 10 minutes before a 2 hour incubation step at 42°C, followed by 10 minutes at 70°C. Quantitative PCR was performed using the Lightcycler (Roche). PCR mixture was prepared by combining cDNA dilution, 10 pmol forward and reverse primer, $MgCl_2$ (4 mM) with 4x homemade LC mastermix (0.9 mM dNTPs, BSA (1 μl/μl, Pharmacia Biotech), Taq polymerase (0.8 U/μl), 4x SYBR Green I (Molecular Probes), 4x AmpliTaq Reaction Buffer (Perkin Elmer)) to a total volume of 20 μl. Amplicons were generated during 45 cycles with annealing temperature set at 55°C. Optimal cDNA dilutions and relative concentrations were determined using a dilution series per gene. Replicate experiments (n = 4) were normalized to 1 and relative expression values were determined by calculating the ratio per gene over the average relative expression of genes, which show no differential expression on both microarray and qPCR (dysferlin, cystatin B, and thrombospondin 4). Significance levels were calculated with a one-sample *t*-test. PCR primer pairs were designed using the Primer3 search engine, available at: **Primer3 Software Distribution** http://frodo.wi.mit.edu/primer3/primer3_code.html. The screened genes and the oligonucleotide primer pairs used for each of the genes in this study correspond to the following nucleotides: myomesin1, 4761–4780 and 4865–4884 (NM_010867); tropomodulin1, 670–689 and 878–897 (NM_021883); dysferlin, 4218–4237 and 4353–4372 (AF188290); cystatinB, 3–22 and 151–170 (NM_007793); thrombospondin4, 2167–2186 and 2289–2308 (NM_011582).

Authors' contributions

RT carried out the tissue preparation, total RNA isolation, target preparation, hybridisations, experimental design, data extraction, data analysis, rt-PCR, and the drafting of the manuscript. PH participated in the experimental design, analysis, rt-PCR, and study coordination. ES participated in the experimental design and analysis. RM provided statistical support. EM was responsible for mouse breeding and tissue preparation. JM participated in experimental design. GO and JD coordinated the study. All authors read the final manuscript.

Additional material

Additional File 1

Differentially expressed genes between mouse inbred strains are selected with a false discovery rate of 10, 15, and 20%. Selected genes are indicated with 1, genes not selected by the specified criteria are indicated with 0. The mean of the relative gene expression levels of each of the five mouse strains is shown. For each gene the GenBank accession number is shown as well as the UniGene ID, gene description and the gene ontology description. The additional file is formatted as Comma Separated Values (CSV) file, and is named Turketal2004_Additional_File.csv.

Click here for file

[http://www.biomedcentral.com/content/supplementary/1471-2164-5-57-S1.csv]

Acknowledgements

We thank Claire Wade and Mark Daly (Whitehead/MIT Center for Genome Research) for providing SNP density levels, and Stefan White (LUMC) for critical comments on the manuscript. This work was supported by the Nederlandse Stichting voor Wetenschappelijk Onderzoek (NWO), the Center for Medical Systems Biology (CMSB) established by the Netherlands Genomics Initiative/Netherlands Organisation for Scientific Research (NGI/NWO), and the Muscular Dystrophy Campaign (UK). Technical support was provided by the Leiden Genome Technology Center.

References

1. Beck JA, Lloyd S, Hafezparast M, Lennon-Pierce M, Eppig JT, Festing MF, Fisher EM: **Genealogies of mouse inbred strains.** *Nat Genet* 2000, **24:**23-25.

2. Pritchard CC, Hsu L, Delrow J, Nelson PS: **Project normal: defining normal variance in mouse gene expression.** *Proc Natl Acad Sci U S A* 2001, **98:**13266-13271.
3. Wade CM, Kulbokas EJ,III, Kirby AW, Zody MC, Mullikin JC, Lander ES, Lindblad-Toh K, Daly MJ: **The mosaic structure of variation in the laboratory mouse genome.** *Nature* 2002, **420:**574-578.
4. Eaves IA, Wicker LS, Ghandour G, Lyons PA, Peterson LB, Todd JA, Glynne RJ: **Combining mouse congenic strains and microarray gene expression analyses to study a complex trait: the NOD model of type 1 diabetes.** *Genome Res* 2002, **12:**232-243.
5. Schadt EE, Monks SA, Drake TA, Lusis AJ, Che N, Colinayo V, Ruff TG, Milligan SB, Lamb JR, Cavet G, Linsley PS, Mao M, Stoughton RB, Friend SH: **Genetics of gene expression surveyed in maize, mouse and man.** *Nature* 2003, **422:**297-302.
6. Chesler EJ, Wang J, Lu L, Qu Y, Manly KF, Williams RW: **Genetic correlates of gene expression in recombinant inbred strains: a relational model system to explore neurobehavioral phenotypes.** *Neuroinformatics* 2003, **1:**343-357.
7. Pavlidis P, Noble WS: **Analysis of strain and regional variation in gene expression in mouse brain.** *Genome Biol* 2001, **2:**RESEARCH0042.
8. Sandberg R, Yasuda R, Pankratz DG, Carter TA, Del Rio JA, Wodicka L, Mayford M, Lockhart DJ, Barlow C: **Regional and strain-specific gene expression mapping in the adult mouse brain.** *Proc Natl Acad Sci U S A* 2000, **97:**11038-11043.
9. Fernandes C, Paya-Cano JL, Sluyter F, D'Souza U, Plomin R, Schalkwyk LC: **Hippocampal gene expression profiling across eight mouse inbred strains: towards understanding the molecular basis for behaviour.** *Eur J Neurosci* 2004, **19:**2576-2582.
10. Kerr MK, Martin M, Churchill GA: **Analysis of variance for gene expression microarray data.** *J Comput Biol* 2000, **7:**819-837.
11. de Menezes RX, Boer JM, Houwelingen JC: **Microarray data analysis: a hierarchical t test to handle heteroscedasticity.** *Applied Bioinformatics* 2004:in press.
12. Reiner A, Yekutieli D, Benjamini Y: **Identifying differentially expressed genes using false discovery rate controlling procedures.** *Bioinformatics* 2003, **19:**368-375.
13. Cowles CR, Joel NH, Altshuler D, Lander ES: **Detection of regulatory variation in mouse genes.** *Nat Genet* 2002, **32:**432-437.
14. Xiang CC, Kozhich OA, Chen M, Inman JM, Phan QN, Chen Y, Brownstein MJ: **Amine-modified random primers to label probes for DNA microarrays.** *Nat Biotechnol* 2002, **20:**738-742.
15. 't Hoen PA, de Kort F, van Ommen GJ, den Dunnen JT: **Fluorescent labelling of cRNA for microarray applications.** *Nucleic Acids Res* 2003, **31:**e20.
16. Huber W, Von Heydebreck A, Sultmann H, Poustka A, Vingron M: **Variance stabilization applied to microarray data calibration and to the quantification of differential expression.** *Bioinformatics* 2002, **18:**S96-S104.
17. Benjamini Y, Hochberg Y: **Controlling the false discovery rate: a practical and powerful appraoch to multiple testing.** *J Roy Stat Soc B* 1995, **57:**289-300.

Chapter 3

Muscle regeneration in dystrophin-deficient mdx mice studied by gene expression profiling

Turk R, Sterrenburg E, de Meijer EJ, van Ommen GJ,
den Dunnen JT, 't Hoen PAC

BMC Genomics 2005 6:98

BMC Genomics

BioMed Central

Research article

Open Access

Muscle regeneration in dystrophin-deficient *mdx* mice studied by gene expression profiling

R Turk[1,3], E Sterrenburg[1], EJ de Meijer[1], G-JB van Ommen[1], JT den Dunnen[1,2] and PAC 't Hoen*[1]

Address: [1]Center for Human and Clinical Genetics, Leiden University Medical Center, Wassenaarseweg 72, 2333 AL Leiden, Nederland, [2]Leiden Genome Technology Center, Leiden University Medical Center, Wassenaarseweg 72, 2333 AL Leiden, Nederland and [3]Department of Physiology and Biophysics, Howard Hughes Medical Institute, University of Iowa, 400 Eckstein Medical Research Building, Iowa City, IA52240-1101, U.S.A

Email: R Turk - rolf-turk@uiowa.edu; E Sterrenburg - e.sterrenburg@Lumc.nl; EJ de Meijer - edm@Lumc.nl; G-JB van Ommen - gjvo@Lumc.nl; JT den Dunnen - ddunnen@HumGen.nl; PAC 't Hoen* - p.a.c.hoen@Lumc.nl

* Corresponding author

Published: 13 July 2005

BMC Genomics 2005, **6**:98 doi:10.1186/1471-2164-6-98

Received: 08 February 2005

Accepted: 13 July 2005

This article is available from: http://www.biomedcentral.com/1471-2164/6/98

Abstract

Background: Duchenne muscular dystrophy (DMD), caused by mutations in the dystrophin gene, is lethal. In contrast, dystrophin-deficient *mdx* mice recover due to effective regeneration of affected muscle tissue. To characterize the molecular processes associated with regeneration, we compared gene expression levels in hindlimb muscle tissue of *mdx* and control mice at 9 timepoints, ranging from 1–20 weeks of age.

Results: Out of 7776 genes, 1735 were differentially expressed between *mdx* and control muscle at at least one timepoint ($p < 0.05$ after Bonferroni correction). We found that genes coding for components of the dystrophin-associated glycoprotein complex are generally downregulated in the *mdx* mouse. Based on functional characteristics such as membrane localization, signal transduction, and transcriptional activation, 166 differentially expressed genes with possible functions in regeneration were analyzed in more detail. The majority of these genes peak at the age of 8 weeks, where the regeneration activity is maximal. The following pathways are activated, as shown by upregulation of multiple members per signalling pathway: the Notch-Delta pathway that plays a role in the activation of satellite cells, and the Bmp15 and Neuregulin 3 signalling pathways that may regulate proliferation and differentiation of satellite cells. In DMD patients, only few of the identified regeneration-associated genes were found activated, indicating less efficient regeneration processes in humans.

Conclusion: Based on the observed expression profiles, we describe a model for muscle regeneration in *mdx* mice, which may provide new leads for development of DMD therapies based on the improvement of muscle regeneration efficacy.

Background

Duchenne muscular dystrophy (DMD) is caused by mutations in the gene encoding dystrophin, a subsarcolemmal protein functioning within the dystrophin-associated glycoprotein complex (DGC)[1,2]. This complex connects the intracellular cytoskeleton to the extracellular matrix. The DGC is concentrated at the Z-lines of the sarcomere and confers the transmission of force across the muscle

BMC Genomics 2005, **6**:98 http://www.biomedcentral.com/1471-2164/6/98

fibre[3]. Disruption of this link results in membrane instability, which eventually leads to sarcolemmal ruptures[4,5]. Influx of extracellular calcium alters molecular processes like muscle contraction and activates proteolytic activity. Affected muscle fibres become necrotic or apoptotic, and release mitogenic chemoattractants, which initiate inflammatory processes [6-8]. Cycles of degeneration and regeneration eventually lead to irreversible muscle wasting and replacement by fibrotic and adipose tissue.

Muscle has the potential to regenerate by activation of undifferentiated myogenic precursor cells (satellite cells), which are normally quiescent and situated between the basal membrane and the myofibers[9,10]. Upon activation, satellite cells proliferate and divide asymmetrically, with the daughter cells having divergent cell fates[11]. Only one of the daughter cells differentiates, progresses towards the myoblast-stadium, and subsequently fuses with other myoblasts or with damaged muscle fibres to induce muscle fibre repair. The other daughter cell remains in a proliferating state or returns to quiescence[12]. Genetic mutations responsible for DMD are also present in satellite cells. Hence, the ability to restore normal muscle function remains obstructed. A small number of muscle fibres are able to produce functional dystrophin, mostly due to secondary mutations in myogenic precursor cells which restore the reading frame[13]. However, these so-called revertant fibres are in a too small minority to alleviate the pathology of the dystrophin-deficiency. Exhaustion of the satellite cell pool due to degeneration and regeneration cycles is thought to critically contribute to the disease[14].

The *mdx* mouse model for DMD has a spontaneous mutation in exon 23 of the Dmd gene, introducing a premature stopcodon[15,16]. The pathology of the *mdx* mouse is characterized by histologically well-defined stages with similarity to the human pathology. Neonatal muscle tissue appears to be unaffected. Necrotic or apoptotic processes in combination with inflammation emerge at approximately 3 weeks of age[15]. Regeneration processes are initiated around the age of 6 weeks and continue while alternating with ongoing degeneration until 12 weeks of age [17-19]. Contrary to the lethal human pathology, the *mdx* mouse somehow recovers from the progressive muscle wasting, and does not show the accumulation of connective and adipose tissue[17,20]. However, *mdx* mice do show a decline in their regeneration capacity at advanced age (>65 weeks), while necrotic processes persist[21]. Since the degeneration processes are similar to those seen in human pathology, the regenerational differences may hold one of the clues of restoration of proper muscle function.

Although previous studies have studied gene expression levels in the *mdx* mouse [22-27], regeneration processes were not studied in full detail. We studied the regeneration process through genome-wide monitoring of gene expression levels[28] in healthy control and *mdx* mice at 9 time points from 1 to 20 weeks of age, while putting emphasis on time points where regenerative activity is maximal (6–12 weeks), a period which was not analysed in detail in a previous time course study[23]. According to the temporal gene expression profiles, we determined which pathways are active during regeneration with respect to normal muscle aging. The majority of identified genes presented in this study have not been described before and provide a substantial addition to the elucidation of the temporal phasing of degeneration and regeneration. By careful annotation based on existing literature, new light is shed on the pathology and subsequent recovery in the *mdx* mouse. Furthermore, we compared gene expression profiles to those of human DMD patients and found only modest overlap in regeneration-associated genes. This confirms that regeneration is no longer an active process at the age at which the patients were profiled (5–12 years old).

Results and discussion

Global comparison of mdx *and control mice*

Gene expression levels were determined in hindlimb muscle tissue from *mdx* and control mice at 9 time points, ranging from 1 to 20 weeks. Differential gene expression levels were calculated per time point by subtraction of the average normalized intensities of control samples from those of the *mdx* samples to correct for normal aging processes. The effects of the normal aging processes on gene expression are relatively minor, and are discussed below. Statistical significance was calculated per time point by performing a Student's *t*-test. Differential gene expression was considered significant when p-values were lower than 0.05 after applying a Bonferroni correction for multiple testing ($p \leq 6.43 \times 10^{-6}$). Out of 7,776 temporal gene expression profiles 1,735 were selected, which satisfied the significance criterion at one or more time points [see Additional file 1].

The number of differentially expressed genes per time point changes considerably during the time course, an effect also shown in a previous study by Porter *et al.*[27] (Figure 1). The number of differentially expressed genes peaks at the age of 8–12 weeks, coinciding with the period of maximal muscle regeneration. Interestingly, also at the first two time points (1 and 2.5 weeks of age), where the histology of the *mdx* muscle is not different from that of control mice, a large number of genes was differentially expressed, indicating differences in muscle development in dystrophin-deficient animals. The majority of these genes (553/677 and 355/407, respectively) also show

BMC Genomics 2005, **6**:98 http://www.biomedcentral.com/1471-2164/6/98

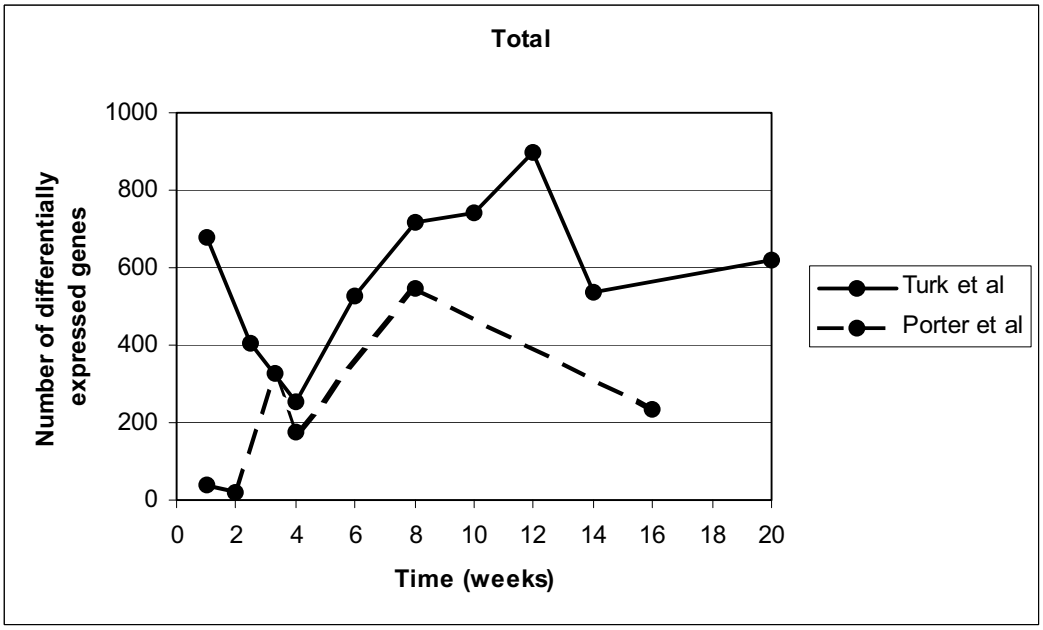


Figure 1
Amount of differentially expressed genes. The number of statistically significantly differentially expressed genes between *mdx* and control mice measured across 9 consecutive timepoints from 1 to 20 weeks in this study (continuous line) are compared to the number of differentially expressed genes found in the study of Porter *et al.*[27] (dashed line).

statistically significant differences in expression at later time points. This overlap can be explained by the assumption that the repertoire of gene products used for muscle growth and development also functions in muscle regeneration.

In this report we will describe the expression changes of two main categories in more detail: genes coding for proteins within the costamer and the dystrophin-associated glycoprotein complex (DGC), and genes involved in regeneration.

The Dystrophin-Glycoprotein Complex

The effect of dystrophin-deficiency on expression levels of dystrophin-glycoprotein complex (DGC)-related genes, or genes with associated functional relevance within the costamer has not been reported in previous gene expression profiling studies [22-26], with the exception of the study of Porter and co-workers[27]. In the Porter study, a downregulation of dystrophin was reported, but no changes in gene expression of other components of the DGC. A selection of 52 genes was made based on an overview by Ervasti *et al.* of members of the costameric protein network[29] (Additional file 4). According to the statistical selection criteria, 4 genes were upregulated (Figure 2A) and 12 were downregulated (Figure 2B) in the *mdx* mouse. Although a decrease in dystrophin expression was found in our study, the stringent statistical criteria were not met.

Upregulated DGC related genes

We find that the retina-specific isoform of dystrophin (Dp260) is expressed in skeletal muscle of *mdx* mice, whereas Dp260 cannot be detected in hindlimb muscle of control mice. Expression of Dp260 was detected by an oligonucleotide probe within the unique first exon of this transcript. The promoter of the Dp260 isoform resides in intron 29, downstream of the *mdx* mutation (exon23). Transgenic *mdx* mice, which overexpress Dp260 via an alpha-actin promoter, show a restoration of a stable association between costameric actin and the sarcolemma, a re-assembly of the DGC, and an overall alleviation of the

BMC Genomics 2005, **6**:98 http://www.biomedcentral.com/1471-2164/6/98

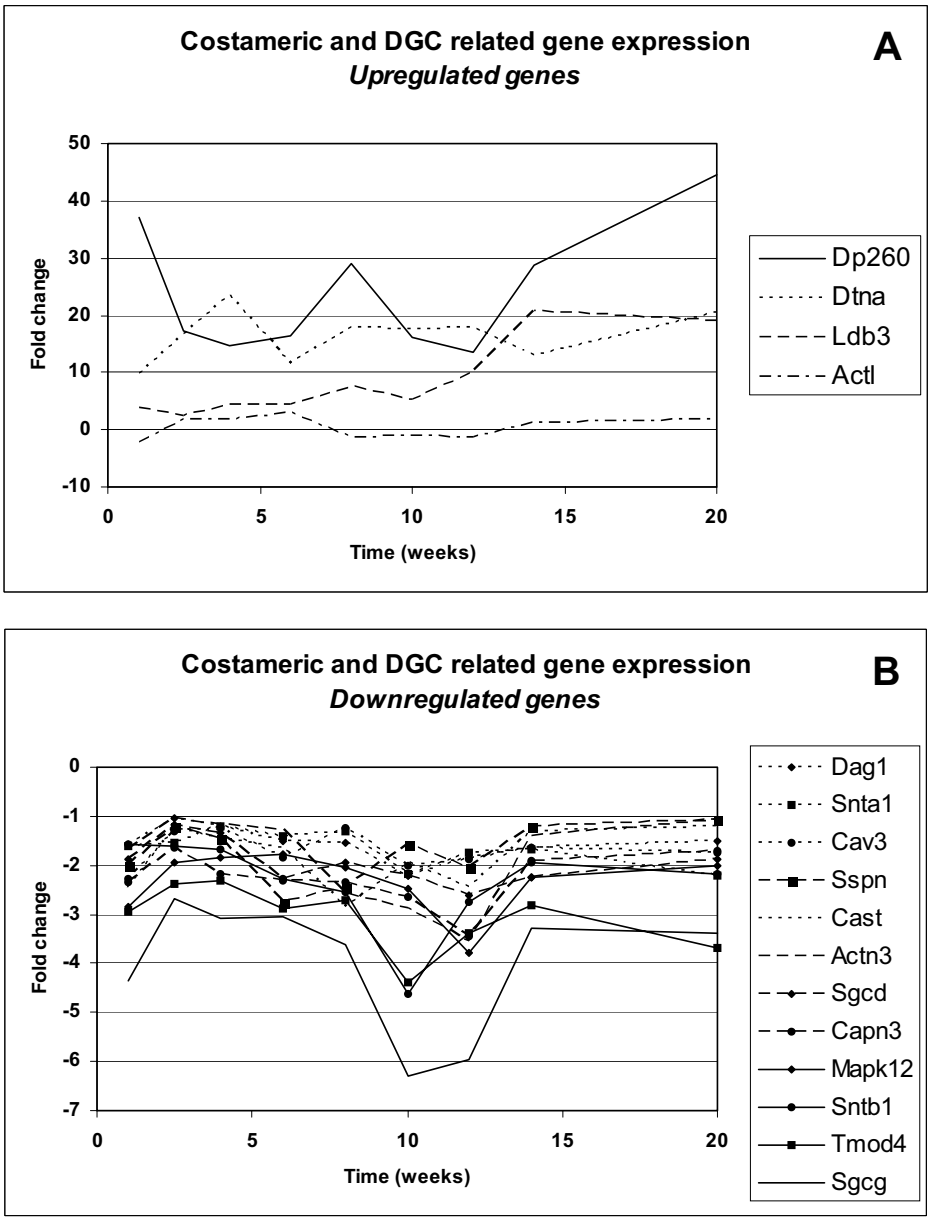


Figure 2
Costameric and DGC related gene expression. Fold change in gene expression levels between *mdx* and control muscle tissue measured across 9 consecutive timepoints from 1 to 20 weeks of costameric and DGC related genes. Figure 2A shows the fold changes of the statistically significantly upregulated genes over time; Figure 2B shows the downregulated genes.

BMC Genomics 2005, **6**:98 http://www.biomedcentral.com/1471-2164/6/98

pathology[30]. Increased transcription initiation of Dp260 might therefore be a natural adaptation for the lack of the muscle specific isoform of dystrophin. However, in contrast to the artificially raised expression by the alpha-actin promoter, the expression of Dp260 in the *mdx* mouse through the original promoter does not seem to be strong enough to compensate for loss of the full-length muscle specific isoform.

We found a continuous upregulation of alpha-dystrobrevin (Dtna) gene expression with maximum differential expression at 4 weeks in *mdx* mice. Dystrobrevin is a phosphotyrosine-containing protein localized at both the sarcolemma and the postsynaptic side of the neuromuscular junction (NMJ), where it binds to either dystrophin or utrophin [31-34]. Dtna has been described to function as a signalling mediator within the DGC[34]. Transcription of Dtna is activated when myoblasts differentiate into multinucleated myotubes[35]. Newey *et al.* reported that Dtna-protein levels are significantly reduced in the *mdx* mouse at the sarcolemma, whereas the protein level was unchanged at the NMJ. This would be consistent with a stabilizing action of Dtna upon binding to dystrophin or utrophin, since dystrophin is not present at the sarcolemma, whereas utrophin is expressed at the NMJ. They proposed a model, where localized translation of Dtna transcripts contributes to synapse formation[36]. Upregulation of Dtna in the *mdx* mouse might indicate an attempt to compensate for the increased turnover of the protein, in order to stabilize the post-synaptic side of neuromuscular junctions of affected muscle fibres, and retain neuronal connection.

Our results show a continuous upregulation of LIM domain protein 3 (Ldb3, Cypher/ZASP). Studies in Ldb3 knock-out mice demonstrated that ablation of Ldb3 eradicates the structural integrity of the Z-line in contracting striated muscle and causes a severe form of congenital myopathy[37]. Upregulation of Ldb3 indicates the necessity for stabilization of the Z-line in *mdx* mice, compensating the undermining effect of dystrophin-deficiency.

Downregulated DGC related genes

It can be seen that gene expression levels of several core-proteins of the DGC, e.g. the transmembrane proteins dystroglycan (Dag1), sarcospan (Sspn), and two members of the sarcoglycan-complex (Sgcd, Sgcg), are lower in *mdx* mice, over the whole time course. Lower expression levels were also detected in other members of the sarcoglycan complex (Sgca, Sgcb, Sgce) and in dystrophin (Dmd, oligonucleotide at the 3' end), but these were not statistically significant. The decrease in expression of DGC-related genes was most prominent during regeneration (8–12 weeks).

Interestingly, DGC related gene expression levels restore to pre-regeneration levels subsequent to the regeneration period, but remain lower than normal (control) level. Similarly, protein levels of core-proteins of the DGC have been shown to be severely reduced in dystrophin-deficient *mdx* mice[38]. It is suggested that the secondary displacement of DGC core-proteins is due to a decrease in protein synthesis and/or assembly, or due to an increase in protein degradation. Similarly, in sarcoglycan-deficiencies the absence of a single subunit causes the loss or strong reduction of the entire sarcoglycan protein complex [39-43]. Since our study reveals a downregulation of mRNA levels of the DGC core-proteins, we conclude that alterations in transcriptional activity also contribute to the decrease in protein levels. As transcription of members of the DGC is likely to be co-ordinately regulated[44], downregulation of these members as seen in *mdx* mice can occur via inhibition or downregulation of shared transcriptional activators.

Regeneration

In the *mdx* mouse, regeneration of affected muscle tissue is most prominent at the age of 6–12 weeks, after which a stabilized condition is reached. To identify pathways active in regeneration, we studied five categories of differentially expressed genes covering major functional characteristics of regenerative tissues (trophic factors, proteases, membrane associated proteins, signal transduction, and transcription, Additional file 5). This selection of 166 genes was typed for temporal effects during regeneration, and the pathways to which they belong. Since signal transduction pathways are still poorly annotated in current genomic databases, these pathways were constructed from the literature. In our study, a pathway is only considered activated or repressed, when multiple members show differential gene expression.

Temporal effects during regeneration

Differential gene expression profiles, based on the ratio between *mdx* and control mice, were scaled to the first time point. Differential expression profiles can therefore be compared independent of the ratio level, which enables the detection of temporal effects. Using k-means clustering (k = 6), differential gene expression profiles were classified according to their temporal similarity. The unscaled temporal effects of *mdx* and control gene expression profiles are shown per cluster for the up- and downregulated genes separately (Figure 3). Genes, which show an upregulation in gene expression during the regenerative phase, are present in clusters 1 (n = 27), cluster 2 (n = 36), and cluster 4 (n = 21). The temporal effect is determined by the gene expression profile of the *mdx* mouse, since gene expression is continuously low without temporal changes in the control mouse. Downregulation of gene expression in the *mdx* mouse during regeneration is

BMC Genomics 2005, **6**:98 http://www.biomedcentral.com/1471-2164/6/98

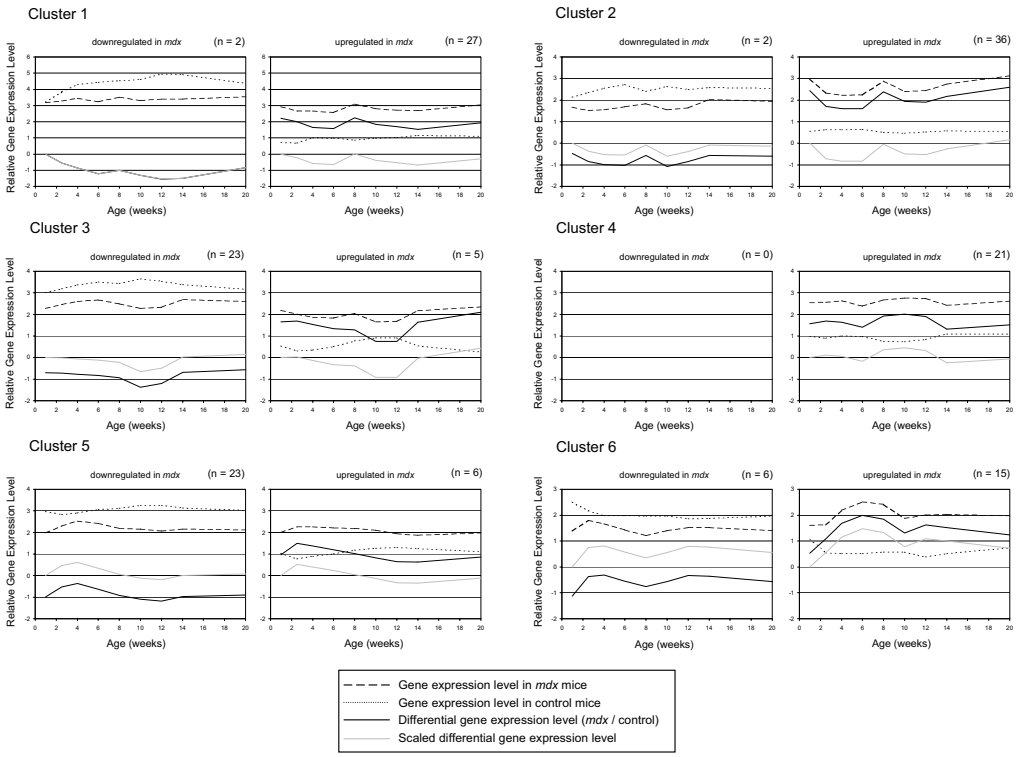


Figure 3

Temporal effects during regeneration. K-means clustering (k = 6) classifies gene expression profiles according to similarity in temporal patterns based on the scaled differential gene expression levels (grey line). For each cluster the up- and downregulated genes are shown separately. Unscaled differential gene expression levels are shown (black line), which are representative for the ratio between the *mdx* gene expression levels (dashed line) and control gene expression levels (small dashed line). Relative gene expression levels are obtained after normalization and coincide with the natural logarithm.

primarily seen in cluster 3 (n = 23). During normal aging, which can be seen in the control mouse, gene expression increases until the age of 10 weeks, followed by a slow decrease. During the regenerative phase in the *mdx* mouse, however, the expression of these genes is downregulated markedly.

Notch-Delta pathway

Gene expression levels of a number of genes functioning in the Notch-Delta pathway are upregulated (Notch1, Notch2, Hr), whereas others (Dxd26, Dvl, Dvl2) are downregulated in the *mdx* mouse at 8 weeks of age (Figure 4A). The gene expression of Dll3 and Numb are switched on in the *mdx* mouse, while no gene expression can be detected in the control mouse (Supplemental Table 1 [see Additional file 4]). The differential expression of the upregulated genes is mostly increased during the regeneration period (6 to 12 weeks) (Figure 4A). For several genes in the Notch-Delta pathway, quantitative RT-PCR experiments were performed to confirm the temporal expression profiles found on the microarray. In accordance with

BMC Genomics 2005, **6**:98 http://www.biomedcentral.com/1471-2164/6/98

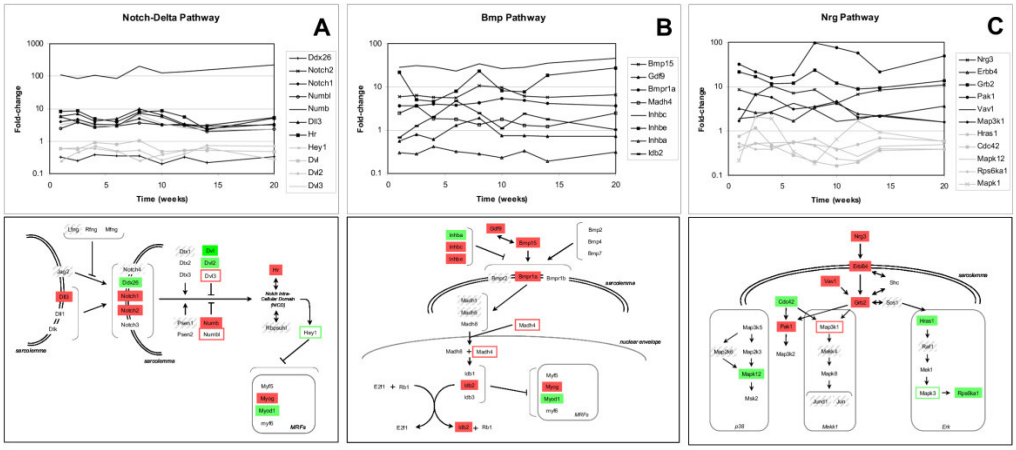


Figure 4

Reconstruction of active regeneration pathways. Regeneration-associated pathways were constructed based on differentially expressed genes, literature study and gene ontology. Expression levels of genes in the Notch-Delta pathway (Panel A), the Bmp15 pathway (Panel B), and the Neuregulin3 pathway (Panel C) are plotted as fold-changes between *mdx* and control mice, as a function of age. In the pathway diagrams, filled boxes refer to upregulation (red) and downregulation (green) at 8 weeks. Outlined boxes refer to upregulation (red) and downregulation (green) at other timepoints than 8 weeks. Shaded grey boxes represent genes which are not detected, or are not represented on the microarray. White boxes represent genes that show no differential expression between *mdx* and control mice.

the microarray results, quantitative RT-PCR experiments demonstrated higher expression of Notch2, Numb and myogenin in *mdx* than in control mice at all ages (Figure 5).

Previous work by Conboy *et al.*[11] indicates the role of the Notch-Delta signalling pathway in the regulation of proliferation versus differentiation of asymmetrically dividing satellite cells by Notch or Numb, respectively. According to Delfini *et al.*[45], Notch is expressed in immature myoblasts, while Delta (Dll) expressing cells are more advanced in myogenesis (post-mitotic myoblasts and muscle fibres). Notch activation is thought to inhibit transcription factors containing a basic helix loop helix domain (bHLH) [46-50], via the induction of Hairy and Enhancer of Split 1 (Hes1)[51], thereby inhibiting myogenic differentiation. Numb-expressing cells are able to undergo myogenic differentiation, because the Notch-Delta pathway is inhibited (Figure 4A). Based on our results, the Notch-Delta signalling pathway, notably the expression of Notch or Numb, is responsible for the determination towards proliferation or differentiation of activated satellite cells in the *mdx* mouse. Since gene expression profiling detects proliferation and differentiation processes simultaneously, satellite cell activation and commitment are ongoing, parallel processes.

Bmp pathway

Various members of a Bmp-associated pathway (Bmp15, Gdf9, Bmpr1a, Madh4, Inhbc, Inhbe, Inhba, and Idb2) are differentially expressed in the *mdx* mouse (Figure 4B). The gene expression of Bmp15, Bmpr1a, Inhbc, Inhbe, and Idb2 is switched on in the *mdx* mouse, while expression cannot be detected in the control mouse. Bone Morphogenetic Protein 15 (Bmp15) is a member of the transforming growth factor-β (TGF-β) family. Bmp15 induces transcription of Inhibitors of DNA Binding proteins (Idb1-3) via binding to Bone Morphogenetic Protein Receptor type I (Bmpr1a), and the downstream transportation of Smad-complex (Madh8-Madh4) to the nucleus[52,53]. Idb-proteins function as positive regulators of cell growth by binding to Retinoblastoma 1 (Rb1). This leads to the activation of the E2f transcription factor, which plays a role in cell-cycle regulation. Furthermore, Idb-proteins inhibit myogenic differentiation through binding to MRFs[54]. The activation of Idb2 in the pre-

BMC Genomics 2005, **6**:98 http://www.biomedcentral.com/1471-2164/6/98

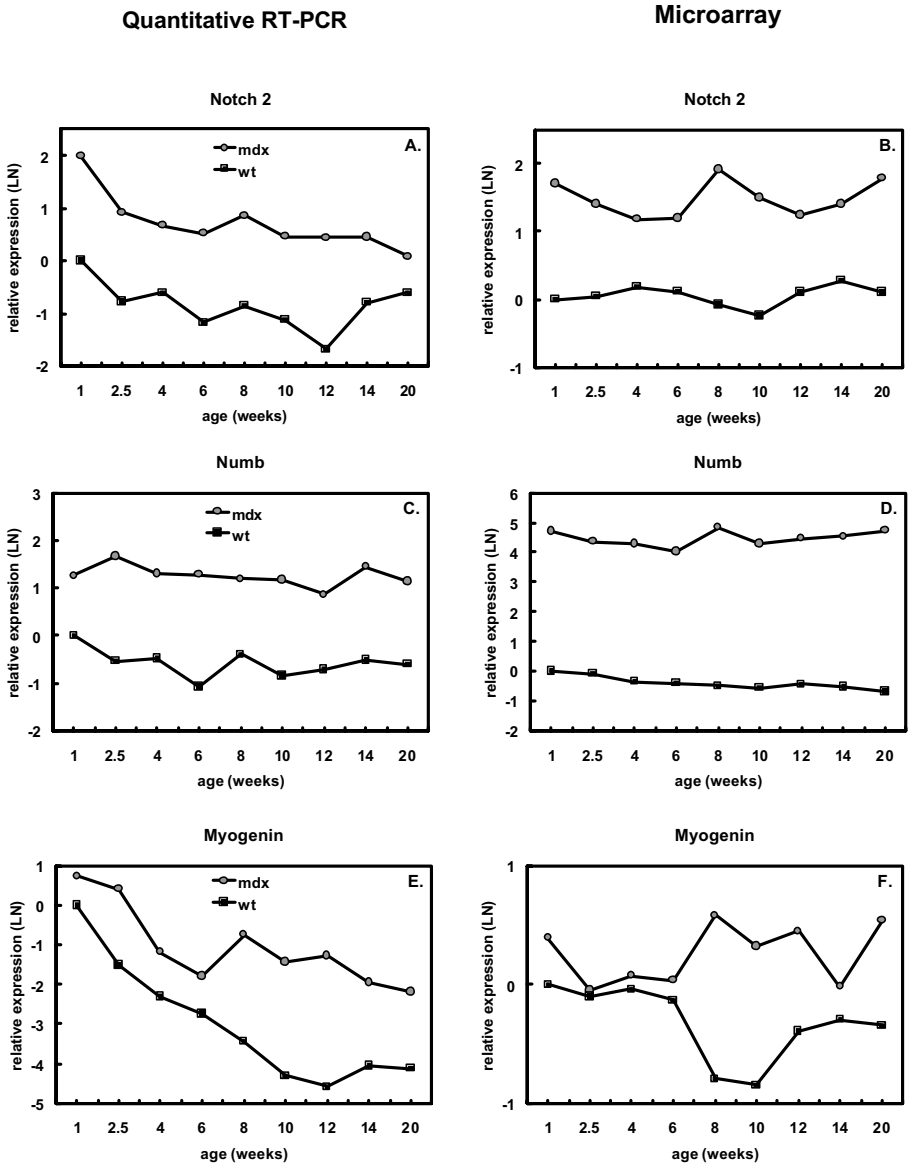


Figure 5
Confirmation of microarray data for genes in the Notch-Delta pathway by quantitative RT-PCR. Expression levels of Notch2 (Panel A and B), Numb (Panel C and D), and Myogenin (Panel E and F) in *mdx* (grey circles) and wild-type (black squares) mice at 1 to 20 weeks of age were measured by quantitative RT-PCR (Panel A, C, E) and expression microarrays (Panel B, D, F). Expression levels relative to those in 1 week-old wild-type mice are plotted on a logarithmic scale (natural logarithm).

BMC Genomics 2005, **6**:98 http://www.biomedcentral.com/1471-2164/6/98

regeneration period is indicative of an inhibition of the myogenic differentiation. This inhibition seems to be alleviated during the regeneration period by a decrease in differential expression of Idb2. Although most of the differentially expressed genes in the Bmp pathway are continuously upregulated, the expression of a number of genes peaks during regeneration (Inhbc, Inhbe, Bmp15, and Gdf9)(Figure 4B). The Inhibin proteins (Inhbc/e), likely to be antagonists of Bone Morphogenetic Proteins[55], are also upregulated. Altogether, this points to a positively and negatively controlled regulation of the Bmp15 pathway. Our data suggests that the Bmp15 pathway has an important function in the balancing of proliferation and differentiation of myoblasts, necessary for effective upscaling of muscle-mass.

Neuregulin pathway

In our study we found that several members of the Epidermal Growth Factor-like (EGF-like) Neuregulin pathway are differentially expressed (Figure 4C). The signalling cascade is activated by the binding of Neuregulin3 (Nrg3) to the extracellular domain of the upregulated protein tyrosine kinase v-erb-a erythroblastic leukemia viral oncogene homolog 4 (ErbB4)[56]. Both Nrg3 and ErbB4 are expressed in the *mdx* mouse and cannot be detected in the control mouse. The interaction between Nrg3 and ErbB4 activates epidermal growth factor-like signal transduction via binding of the adaptor protein Growth factor receptor bound protein 2 (Grb2), which peaks at the initiation of regeneration (Figure 4C). Grb2 can activate Mitogen activated kinase kinase 1 (Map3k1)[57], whose differential gene expression is increased at 2.5 weeks as well as during regeneration. The activation of the MAP kinase pathways eventually leads to transcriptional induction (reviewed in[58]) through members of Activating protein complex 1 (Ap1), like Jund1 and Jun.

Depending on the protein complexes formed, specific transcription activation will lead to different biological processes ranging from proliferation to differentiation. Furthermore, we find that other Grb-interacting proteins like Vav 1 oncogene (Vav1)[59], and p21-activated kinase 1 (Pak1)[60] are switched on and upregulated, respectively. The differential gene expression of Vav1 increases during the initiation of regeneration, and might play a role in the clustering of integrins for cell adhesion[61]. Pak1 differential gene expression is increased during regeneration. Downstream genes activated by Pak1 regulate cytoskeletal dynamics, proliferation and cell survival signalling[60]. Furthermore, our results show that the Erk pathway is downregulated in the *mdx* mouse as well as the p38 pathway.

Comparisons with other studies

In contrast to previously published studies of temporal gene expression profiling in the *mdx* mouse [25-27], we primarily focus on regeneration. The majority of differentially expressed genes in regeneration (148 out of 166) in our study have not been reported as differentially expressed in the *mdx* mouse in other studies. Apart from important differences in gene coverage (the Affymetrix U74v2 GeneChips used in the other studies lack probe sets for 22/166 genes), we explain the limited overlap by differences in cut-off levels: as we applied very stringent statistical tests, we could avoid setting a cut-off level for the fold change, thereby picking up genes with small but consistent fold changes, which can be biologically very relevant, especially in the case of transcription factors. This may also explain the large difference between the study of Porter *et al.* and our study in the number of genes found differentially expressed at the early timepoints (Figure 1), where mainly subtle expression changes are expected.

Goetsch *et al.* reported results from a gene expression profiling studies during muscle regeneration induced by cardiotoxin injection in wildtype mice[62]. The authors concluded that muscle regeneration is a complex process that requires the coordinated modulation of the inflammatory response, myogenic precursor cells, growth factors, and the extracellular matrix for complete regeneration of muscle architecture. A similar study of cardiotoxin-induced muscle regeneration, recently published by Zhao and Hoffman, reported that embryonic positional cues (Wnt, Shh, and Bmp) were not induced, whereas expression of factors involved in satellite cell proliferation and differentiation (MRFs, Pax, Notch1, and FGFR4) was recapitulated[63]. Our study, which also asserts satellite cell activation, proliferation, and differentiation, shows differences in muscle regeneration between *mdx* and wildtype mice. Bmp and EGF-like signalling pathways are activated during regeneration in the *mdx* mouse, as well as upregulation of members of the Notch-Delta pathway. In contrast to the upregulation of Pax7 in wildtype mice, we have found upregulation of Pax3.

Altogether, these findings suggest that dystrophin-deficiency might lead to enhanced regeneration processes in hindlimb muscles over and above those found in wildtype mice. It is likely that the regeneration pathways identified in our study are also active in the *mdx* diaphragm, given that the expression of their downstream targets, the muscle regulatory factors myf-5, myoD, and myogenin, are even more elevated in *mdx* diaphragm than in hindlimb[64]. As demonstrated in another recent study[65], the regeneration capacity per se is high and does not explain why the muscle wasting in the *mdx* diaphragm is more severe than in the hindlimb. Other factors

such as higher workload and different involvement of the immune system are likely to contribute.

To discern active processes between the lethal human and regenerative murine dystrophin-deficiency, the selected murine gene expression profiles at 8 weeks of age were compared to those of DMD patients[66]. Out of 166 regeneration-associated transcripts, 19 genes could be detected that are differentially expressed in both the human and murine muscular dystrophy (Additional file 5). Seven of these overlapping genes showed an opposite differential expression between human DMD and *mdx*, of which Platelet derived growth factor beta (Pdgfb) and Paired box 3 (Pax3) are discussed below.

Gene expression of Pdgfb is upregulated in the *mdx* mouse (18.6 fold), where it is downregulated (-1.7-fold) in DMD patients. In the *mdx* mouse, Pdgfb shows an increase in gene expression during regeneration (present in cluster 4, Figure 3). Pdgfb was immunolocalized in infiltrating macrophages, regenerating muscle fibres, and myofibre nuclei of affected dystrophic muscle tissue[67]. The mitogen Pdgfb stimulates myoblast proliferation, while inhibiting myoblast differentiation[68]. It seems to have a similar role during regeneration. Paired box 3 (Pax3) gene expression is activated in the *mdx* mouse relative to the control. Its gene expression increases during regeneration, peaking at 12 weeks of age, while hPax3 is downregulated in DMD patients (-1.6-fold). Pax3 is capable of activating the expression of the muscle regulatory factors Myod1, Myf5 or Myogenin, and thereby activating the myogenic program[69].

The limited amount of overlapping genes between *mdx* mice and human DMD patients, as well as a number of genes showing opposite expression (i.e. Pdgfb and Pax3), suggests that processes active in regenerating mouse muscle are not active in human patients at the time gene expression was profiled (Age: 5–12 years old). This corresponds with clinical findings that patients older than 5 years have surpassed active regeneration processes[70]. The discovery of genes showing opposite regulation may partly explain the differences in regeneration efficiency and lethal manifestation of the pathology between dystrophin-deficient human and murine muscles.

Conclusion

Mdx mice lack a functional DGC at the sarcolemma. As a consequence, gene expression of most DGC members is downregulated. *Mdx* mice suffer from massive muscle fibre necrosis starting at the age of 3 weeks. Regenerative processes, starting approximately at the age of 6 weeks, largely restore muscle tissue architecture, although muscle fibres remain centrally nucleated. Recovered muscles of *mdx* mice have slightly diminished strength and higher fatigability. By analysing temporal expression profiles in *mdx* and control mice, we have identified genes and pathways involved in regeneration. The expression of these genes peaks between the ages of 6–12 weeks. Based on the observation that several of these genes are not expressed in control muscle and based on gene ontology classification and further literature annotation, we suggest a role for these genes in activation, proliferation, or differentiation of satellite cells and myoblasts. We propose the following model (Figure 6). Muscular dystrophy leads to muscle fibre necrosis, which attracts inflammatory cells, and release of trophic factors. These factors activate quiescent satellite cells, which as a consequence start to proliferate and differentiate. This divergent cell-fate is controlled by the Notch-Delta pathway. Activated satellite cells differentiate to myoblasts, which proliferate and differentiate as well. The balance between these cell-fates may be regulated by the level of Numb and the activation of Bmp15 and Nrg3 signalling pathways. Differentiation of myoblasts eventually leads to fusion with affected muscle fibres or to the formation of new muscle fibres. The genes and pathways active in regeneration are reminiscent of embryonic myogenesis. We hypothesize that the newly formed or repaired muscle fibres are similar to those in the pre-necrotic phase, but are more able to adapt to dystrophin-deficiency through remodelling of muscle structure and fibre composition. Since many regeneration-related genes remain higher expressed in *mdx* than in control muscle, it seems that regeneration processes are active throughout the life span of the animal. Regenerative processes appear to be most effective when mice reach adulthood, and normal growth processes cease. Regeneration and muscle development are both dependent on the availability of satellite cells, and these processes will therefore compete for the satellite cell availability when activated simultaneously. In the human situation, regeneration processes seem to be exhausted before growth is finished. Together with the accumulation of fibrotic and adipose tissue, exhaustion is thought to be the reason of the lethal manifestation of the disease in human patients. Prolongation of the regenerative capacity by activating the described pathways and/or replenishment with a pool of 'regeneration-primed' cells may therefore provide an attractive strategy in the treatment of muscular dystrophy.

Methods

Target preparation and hybridisation

Hindlimb muscle tissue was isolated from *mdx* (C57Bl/10ScSn$^{mdx/J}$ (Jackson) × C57Bl/6NCrl (Charles River) × CBA/JCrl (Charles River)) and control (C57Bl/10ScSnOlaHsd, Harland) at the ages of 1, 2 1/2, 4, 6, 8, 10, 12, 14, and 20 weeks (2 individuals per time point per strain). Total RNA was isolated as described previously (Turk *et al.*, submitted). cRNA was prepared by linear

BMC Genomics 2005, **6**:98 http://www.biomedcentral.com/1471-2164/6/98

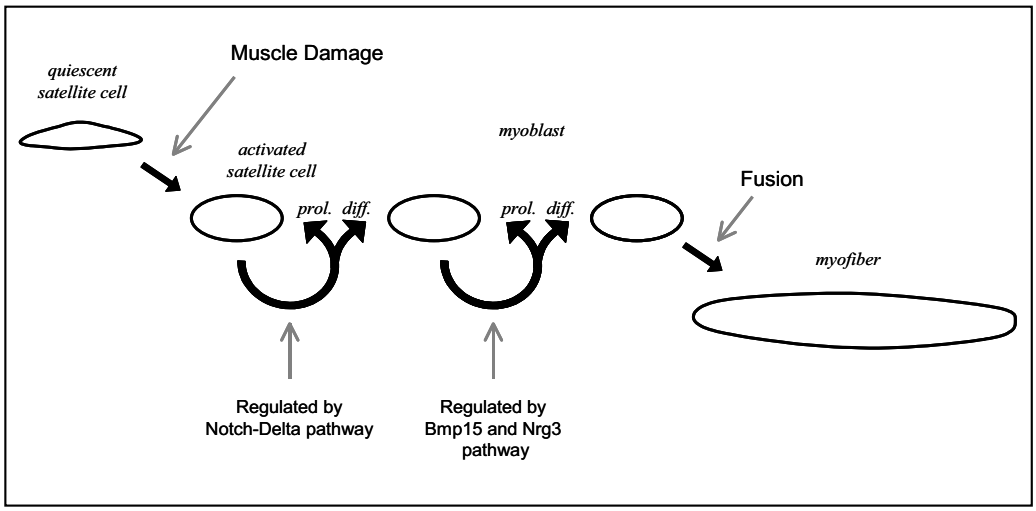


Figure 6
Schematic model of processes in regeneration.

amplification and concurrent incorporation of aminoallyl UTP, followed by chemical coupling to monoreactive Cy3 or Cy5 dyes[71]. Labelled targets (1.5 μg cRNA per target) were hybridised overnight on prehybridised murine oligonucleotide microarrays (65-mer with 5'-hexylaminolinker, Sigma-Genosys mouse 7.5K oligonucleotide library, printed in duplicate) using an automatic hybridisation station (GeneTac, Genomic Solutions). Posthybridisation washes were performed as described previously[71].

Data analysis

Hybridisations were performed in a dye-swap fashion using temporal loop-designs [72-74], enabling optimal detection of gene expression differences between adjacent time points for both *mdx* and control targets [see Additional file 2] Microarrays were analysed by GenePix Pro 3 feature extraction software (Axon). Local background-corrected, median spot intensities were normalized simultaneously for all microarray experiments using Variance Stabilization and Normalization (VSN) in R[75]. This transformation coincides with the natural logarithm for the high intensities. Array data has been made available through the GEO data repository of the National Centre for Biotechnology Information under series GSE1574. Averaged (arithmetic mean) normalized intensities were calculated per gene per time point for *mdx* and control samples based upon 8 data points (2 biological replicates with 4 technical replicates each). This method is more efficient than ratio-based calculations for each hybridisation[76]. Genes were considered expressed when the average normalized intensity was higher than the background level. The background level was determined by calculating the averaged normalized intensity of 157 empty spots in all experiments plus 3 standard deviations.

Fold-changes in gene expression were calculated on a linear scale by subtraction of averaged normalized intensities of control samples from *mdx* samples at each time point, followed by returning *e* raised to the power of the difference. Maximum fold-changes per gene were determined according to the highest fold-change within the time course. Statistically significant differential gene expression per gene per time point was calculated between *mdx* and control samples by performing a two-tailed Student's *t*-test assuming equal distributions. Significance levels were set at 0.05 after applying a Bonferroni correction for multiple testing. Gene expression profiles were taken in consideration when at least one time point showed statistically significant differential gene expression between *mdx* and control samples.

Comparisons with other gene expression studies were facilitated by the program GeneHopper[77] that links annotations for different platforms.

Clustering

Temporal differential gene expression profiles were scaled to time point t = 1 week. Selected profiles were grouped using k-means clustering into a predetermined number of clusters according to the correlation similarity measure (Spotfire DecisionSite 7.1.1, Functional Genomics package). Clustering was initiated using evenly spaced profiles as algorithm. This method generates profiles to be used as centroids that are evenly distributed between the minimum and maximum value for each variable in the selected profiles (from Spotfire DecisionSite User's Guide and Reference Manual).

Comparison with data from Duchenne patients

In a previous study, gene expression levels in 2 pools of muscle RNA from 5 Duchenne patients (aged 5–6 years and aged 10–12 years, respectively) and 2 pools of muscle RNA from non-dystrophic controls (aged 5–12 years and aged 4–13 years, respectively), were evaluated on Affymetrix U95A and U95Av2 GeneChips® [78]. We re-analysed the gene expression data to obtain expression levels for all genes, as well as the most recent annotation. To this end, publicly available cel-files http://pepr.cnmcresearch.org were loaded in Rosetta Resolver® Gene Expression Analysis System v4.0 (Rosetta Biosoftware Inc., Seattle, WA). Data were processed and normalized with the Rosetta error model for Affymetrix U95A Genechips. Genes that showed differential expression ($p < 0.001$ and absolute fold-change >1.5) between the pools of dystrophic patients and non-dystrophic controls were exported and linked to mouse UniGene clusters with GeneHopper, based on HomoloGene annotation.

Functional annotation

Gene Ontology annotation was developed by Compugen, using nomenclature obtained from the Gene Ontology consortium http://pepr.cnmcresearch.org. Additional information was retrieved via OMIM and LocusLink http://www.ncbi.nlm.nih.gov.

Quantitative RT-PCR

cDNA was synthesized by adding 40 ng of random hexamer primers to 1 μg of total RNA in a total volume of 11 μl. After denaturation for 10 min at 70°C and cooling for 10 min on ice, 4 μl of 5× first-strand buffer (MBI-Fermentas), 2 μl of 10 mM dNTPs and 2 μl (200 U/μl) RevertAid RNase H- (MBI-Fermentas) was added. The mixture was incubated for 10 min at room temperature 2 hours at 42°C. The cDNA synthesis was halted by heating at 70°>C for 10 min. Quantitative PCR assays, using 10 μl of 20× diluted cDNA, were run on a MyIQ real-time PCR detection system (BioRad), applying 36 cycles of 10 seconds denaturation (95°C), 20 seconds annealing (60°C), and 25 seconds extension (72°C). PCR mixtures contained 1× PCR Buffer (Roche), 3 mM $MgCl_2$, 225 μM of each dNTP, 250 μg/ml BSA, 1× SYBR-Green (diluted from 10,000× stock, Molecular Probes), 10 nM fluorescein (BioRad), 2.5 U homemade Taq polymerase, 0.25 U AmpliTaq (Roche), and 10 pmol of forward and reverse primers (for sequences see Additional file 3). The PCR efficiencies, determined by analysis of a dilution series of a mixture of all cDNA samples over 5 orders of magnitude, ranged from 94.5% to 98% for the different primer pairs. Melting curves were analyzed to confirm single product formation. Gene expression levels were calculated using the gene expression macro provided by BioRad and normalized to glyceraldehyde-3-phosphate dehydrogenase (GAPDH, stable expression in all samples) expression levels.

Abbreviations

DMD Duchenne muscular dystrophy

DGC Dystrophin-associated glycoprotein complex

NMJ Neuromuscular junction

VSN Variance stabilization and normalization

RT-PCR Reverse transcription followed by polymerase chain reaction

Authors' contributions

RT performed the microarray hybridisations, analysed the microarrays and wrote the draft of the paper. ES helped to set up the microarray technique and assisted with the analysis. EdM was involved in the breeding of the mice, the isolation of the tissues, and the quantitative PCR experiments. GJvO edited the manuscript. JdD conceived the study. P'tH was involved in the microarray analysis and participated in the writing of the paper.

Additional material

Additional File 1

Averaged expression and significance levels for 1735 genes differentially expressed at at least one time point.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-6-98-S1.xls]

Additional File 4

Costameric and DGC related gene expression.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-6-98-S4.xls]

Additional File 5

Regeneration associated genes.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-6-98-S5.xls]

Additional File 2

***Temporal loop design**. Hybridisations were done using a temporal loop design for the* mdx *and the control samples separately. The temporal loop design balances dyes and samples, and provides low variance between adjacent timepoints[73]. The order of age is maintained in the hybridisation scheme; each target is hybridised with the target of the following timepoint. Hybridisations are indicated by Hyb ID. Green and red boxes indicate labelling of target with Cy3 and Cy5 respectively. The number in the boxes indicates the age of the mouse. The boxes linked by an arrow are identical samples.*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-6-98-S2.pdf]

Additional File 3

Primer sequences used for quantitative RT-PCR assays.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-6-98-S3.xls]

Acknowledgements

This work was supported by the Nederlandse Stichting voor Wetenschappelijk Onderzoek (NWO) and the Center for Medical Systems Biology (CMSB) established by the Netherlands Genomics Initiative / Netherlands Organisation for Scientific Research (NGI/NWO).

References

1. Hoffman EP, Brown RHJ, Kunkel LM: **Dystrophin: the protein product of the Duchenne muscular dystrophy locus.** *Cell* 1987, **51:**919-928.
2. Ervasti JM, Campbell KP: **Membrane organization of the dystrophin-glycoprotein complex.** *Cell* 1991, **66:**1121-1131.
3. Straub V, Bittner RE, Leger JJ, Voit T: **Direct visualization of the dystrophin network on skeletal muscle fiber membrane.** *J Cell Biol* 1992, **119:**1183-1191.
4. Carpenter S, Karpati G: **Duchenne muscular dystrophy: plasma membrane loss initiates muscle cell necrosis unless it is repaired.** *Brain* 1979, **102:**147-161.
5. Petrof BJ, Shrager JB, Stedman HH, Kelly AM, Sweeney HL: **Dystrophin protects the sarcolemma from stresses developed during muscle contraction.** *Proc Natl Acad Sci U S A* 1993, **90:**3710-3714.
6. Bischoff R: **Control of satellite cell proliferation.** *Adv Exp Med Biol* 1990, **280:**147-157.
7. Chen G, Quinn LS: **Partial characterization of skeletal myoblast mitogens in mouse crushed muscle extract.** *J Cell Physiol* 1992, **153:**563-574.
8. Jimena I, Pena J, Luque E, Vaamonde R: **Muscle hypertrophy experimentally induced by administration of denervated muscle extract.** *J Neuropathol Exp Neurol* 1993, **52:**379-386.
9. Bischoff R: **The satellite cell and muscle regeneration.** In *Myology: basic and clinical Volume 3*. 2nd edition. Edited by: Engel AG and Franzini-Armstrong C. New York, McGraw-Hill; 1994:97-118.
10. Charge SB, Rudnicki MA: **Cellular and molecular regulation of muscle regeneration.** *Physiol Rev* 2004, **84:**209-238.
11. Conboy IM, Rando TA: **The regulation of Notch signaling controls satellite cell activation and cell fate determination in postnatal myogenesis.** *Dev Cell* 2002, **3:**397-409.
12. Conboy IM, Conboy MJ, Smythe GM, Rando TA: **Notch-mediated restoration of regenerative potential to aged muscle.** *Science* 2003, **302:**1575-1577.
13. Klein CJ, Coovert DD, Bulman DE, Ray PN, Mendell JR, Burghes AH: **Somatic reversion/suppression in Duchenne muscular dystrophy (DMD): evidence supporting a frame-restoring mechanism in rare dystrophin-positive fibers.** *Am J Hum Genet* 1992, **50:**950-959.
14. Jejurikar SS, Kuzon WMJ: **Satellite cell depletion in degenerative skeletal muscle.** *Apoptosis* 2003, **8:**573-578.
15. Bulfield G, Siller WG, Wight PA, Moore KJ: **X chromosome-linked muscular dystrophy (mdx) in the mouse.** *Proc Natl Acad Sci U S A* 1984, **81:**1189-1192.
16. Sicinski P, Geng Y, Ryder-Cook AS, Barnard EA, Darlison MG, Barnard PJ: **The molecular basis of muscular dystrophy in the mdx-mouse: a point mutation.** *Science* 1989, **244:**1578-1580.
17. Dangain J, Vrbova G: **Muscle development in mdx mutant mice.** *Muscle Nerve* 1984, **7:**700-704.
18. Anderson JE, Ovalle WK, Bressler BH: **Electron microscopic and autoradiographic characterization of hindlimb muscle regeneration in the mdx mouse.** *Anat Rec* 1987, **219:**243-257.
19. DiMario JX, Uzman A, Strohman RC: **Fiber regeneration is not persistent in dystrophic (MDX) mouse skeletal muscle.** *Dev Biol* 1991, **148:**314-321.
20. Torres LF, Duchen LW: **The mutant mdx: inherited myopathy in the mouse. Morphological studies of nerves, muscles and end-plates.** *Brain* 1987, **110 (Pt 2):**269-299.
21. Pastoret C, Sebille A: **mdx mice show progressive weakness and muscle deterioration with age.** *J Neurol Sci* 1995, **129:**97-105.
22. Tkatchenko AV, Le Cam G, Leger JJ, Dechesne CA: **Large-scale analysis of differential gene expression in the hindlimb muscles and diaphragm of mdx mouse.** *Biochim Biophys Acta* 2000, **1500:**17-30.
23. Porter JD, Khanna S, Kaminski HJ, Rao JS, Merriam AP, Richmonds CR, Leahy P, Li J, Guo W, Andrade FH: **A chronic inflammatory response dominates the skeletal muscle molecular signature in dystrophin-deficient mdx mice.** *Hum Mol Genet* 2002, **11:**263-272.
24. Rouger K, Le Cunff M, Steenman M, Potier MC, Gibelin N, Dechesne CA, Leger JJ: **Global/temporal gene expression in diaphragm and hindlimb muscles of dystrophin-deficient (mdx) mice.** *Am J Physiol Cell Physiol* 2002, **283:**C773-C784.
25. Boer JM, de Meijer EJ, Mank EM, van Ommen GB, den Dunnen JT: **Expression profiling in stably regenerating skeletal muscle of dystrophin-deficient mdx mice.** *Neuromuscul Disord* 2002, **12:**S118-S124.
26. Tseng BS, Zhao P, Pattison JS, Gordon SE, Granchelli JA, Madsen RW, Folk LC, Hoffman EP, Booth FW: **Regenerated mdx mouse skeletal muscle shows differential mRNA expression.** *J Appl Physiol* 2002, **93:**537-545.
27. Porter JD, Merriam AP, Leahy P, Gong B, Khanna S: **Dissection of temporal gene expression signatures of affected and spared muscle groups in dystrophin-deficient (mdx) mice.** *Hum Mol Genet* 2003, **12:**1813-1821.
28. Schena M, Shalon D, Davis RW, Brown PO: **Quantitative monitoring of gene expression patterns with a complementary DNA microarray.** *Science* 1995, **270:**467-470.
29. Ervasti JM: **Costameres: the Achilles' heel of Herculean muscle.** *J Biol Chem* 2003, **278:**13591-13594.
30. Warner LE, DelloRusso C, Crawford RW, Rybakova IN, Patel JR, Ervasti JM, Chamberlain JS: **Expression of Dp260 in muscle tethers the actin cytoskeleton to the dystrophin-glycoprotein complex and partially prevents dystrophy.** *Hum Mol Genet* 2002, **11:**1095-1105.
31. Wagner KR, Cohen JB, Huganir RL: **The 87K postsynaptic membrane protein from Torpedo is a protein-tyrosine kinase substrate homologous to dystrophin.** *Neuron* 1993, **10:**511-522.
32. Blake DJ, Nawrotzki R, Peters MF, Froehner SC, Davies KE: **Isoform diversity of dystrobrevin, the murine 87-kDa postsynaptic protein.** *J Biol Chem* 1996, **271:**7802-7810.
33. Sadoulet-Puccio HM, Rajala M, Kunkel LM: **Dystrobrevin and dystrophin: an interaction through coiled-coil motifs.** *Proc Natl Acad Sci U S A* 1997, **94:**12413-12418.
34. Peters MF, Sadoulet-Puccio HM, Grady MR, Kramarcy NR, Kunkel LM, Sanes JR, Sealock R, Froehner SC: **Differential membrane**

localization and intermolecular associations of alpha-dystrobrevin isoforms in skeletal muscle. *J Cell Biol* 1998, **142:**1269-1278.

35. Holzfeind PJ, Ambrose HJ, Newey SE, Nawrotzki RA, Blake DJ, Davies KE: **Tissue-selective expression of alpha-dystrobrevin is determined by multiple promoters.** *J Biol Chem* 1999, **274:**6250-6258.
36. Newey SE, Gramolini AO, Wu J, Holzfeind P, Jasmin BJ, Davies KE, Blake DJ: **A novel mechanism for modulating synaptic gene expression: differential localization of alpha-dystrobrevin transcripts in skeletal muscle.** *Mol Cell Neurosci* 2001, **17:**127-140.
37. Zhou Q, Chu PH, Huang C, Cheng CF, Martone ME, Knoll G, Shelton GD, Evans S, Chen J: **Ablation of Cypher, a PDZ-LIM domain Z-line protein, causes a severe form of congenital myopathy.** *J Cell Biol* 2001, **155:**605-612.
38. Ohlendieck K, Campbell KP: **Dystrophin-associated proteins are greatly reduced in skeletal muscle from mdx mice.** *J Cell Biol* 1991, **115:**1685-1694.
39. Roberds SL, Leturcq F, Allamand V, Piccolo F, Jeanpierre M, Anderson RD, Lim LE, Lee JC, Tome FMS, Romero NB, Fardeau M, Beckmann JS, Kaplan JC, Campbell KP: **Missense mutations in the adhalin gene linked to autosomal recessive muscular dystrophy.** *Cell* 1994, **78:**625-633.
40. Lim LE, Duclos F, Broux O, Bourg N, Sunada Y, Allamand V, Meyer J, Richard I, Moomaw C, Slaughter C, Tome C, Fardeau F, Jackson M, Beckmann CE, Campbell KP: **Beta-sarcoglycan: characterization and role in limb-girdle muscular dystrophy linked to 4q12.** *Nat Genet* 1995, **11:**257-265.
41. Noguchi S, McNally EM, Ben Othmane K, Hagiwara Y, Mizuno Y, Yoshida M, Yamamoto H, Bonnemann CG, Gussoni E, Denton PH, Kyriakides T, Middleton L, Hentati F, Ben Hamida M, Nonaka I, Vance JM, Kunkel LM, Ozawa E: **Mutations in the dystrophin-associated protein gamma-sarcoglycan in chromosome 13 muscular dystrophy.** *Science* 1995, **270:**819-822.
42. Bonnemann CG, Modi R, Noguchi S, Mizuno Y, Yoshida M, Gussoni E, McNally EM, Duggan DJ, Angelini C, Hoffman EP: **Beta-sarcoglycan (A3b) mutations cause autosomal recessive muscular dystrophy with loss of the sarcoglycan complex.** *Nat Genet* 1995, **11:**266-273.
43. Nigro V, de Sa ME, Piluso G, Vainzof M, Belsito A, Politano L, Puca AA, Passos-Bueno MR, Zatz M: **Autosomal recessive limb-girdle muscular dystrophy, LGMD2F, is caused by a mutation in the delta-sarcoglycan gene.** *Nat Genet* 1996, **14:**195-198.
44. Wakabayashi-Takai E, Noguchi S, Ozawa E: **Identification of myogenesis-dependent transcriptional enhancers in promoter region of mouse gamma-sarcoglycan gene.** *Eur J Biochem* 2001, **268:**948-957.
45. Delfini M, Hirsinger E, Pourquie O, Duprez D: **Delta 1-activated notch inhibits muscle differentiation without affecting Myf5 and Pax3 expression in chick limb myogenesis.** *Development* 2000, **127:**5213-5224.
46. Kopan R, Nye JS, Weintraub H: **The intracellular domain of mouse Notch: a constitutively activated repressor of myogenesis directed at the basic helix-loop-helix region of MyoD.** *Development* 1994, **120:**2385-2396.
47. Nye JS, Kopan R, Axel R: **An activated Notch suppresses neurogenesis and myogenesis but not gliogenesis in mammalian cells.** *Development* 1994, **120:**2421-2430.
48. Shawber C, Nofziger D, Hsieh JJ, Lindsell C, Bogler O, Hayward D, Weinmaster G: **Notch signaling inhibits muscle cell differentiation through a CBF1-independent pathway.** *Development* 1996, **122:**3765-3773.
49. Artavanis-Tsakonas S, Rand MD, Lake RJ: **Notch signaling: cell fate control and signal integration in development.** *Science* 1999, **284:**770-776.
50. Delgado I, Huang X, Jones S, Zhang L, Hatcher R, Gao B, Zhang P: **Dynamic gene expression during the onset of myoblast differentiation in vitro.** *Genomics* 2003, **82:**109-121.
51. Kuroda K, Tani S, Tamura K, Minoguchi S, Kurooka H, Honjo T: **Delta-induced Notch signaling mediated by RBP-J inhibits MyoD expression and myogenesis.** *J Biol Chem* 1999, **274:**7238-7244.
52. Moore RK, Otsuka F, Shimasaki S: **Molecular basis of bone morphogenetic protein-15 signaling in granulosa cells.** *J Biol Chem* 2003, **278:**304-310.
53. Miyazawa K, Shinozaki M, Hara T, Furuya T, Miyazono K: **Two major Smad pathways in TGF-beta superfamily signalling.** *Genes Cells* 2002, **7:**1191-1204.
54. Norton JD, Deed RW, Craggs G, Sablitzky F: **Id helix-loop-helix proteins in cell growth and differentiation.** *Trends Cell Biol* 1998, **8:**58-65.
55. Wiater E, Vale W: **Inhibin is an antagonist of bone morphogenetic protein signaling.** *J Biol Chem* 2003, **278:**7934-7941.
56. Zhang D, Sliwkowski MX, Mark M, Frantz G, Akita R, Sun Y, Hillan K, Crowley C, Brush J, Godowski PJ: **Neuregulin-3 (NRG3): a novel neural tissue-enriched protein that binds and activates ErbB4.** *Proc Natl Acad Sci U S A* 1997, **94:**9562-9567.
57. Pomerance M, Multon MC, Parker F, Venot C, Blondeau JP, Tocque B, Schweighoffer F: **Grb2 interaction with MEK-kinase 1 is involved in regulation of Jun-kinase activities in response to epidermal growth factor.** *J Biol Chem* 1998, **273:**24301-24304.
58. Hazzalin CA, Mahadevan LC: **MAPK-regulated transcription: a continuously variable gene switch?** *Nat Rev Mol Cell Biol* 2002, **3:**30-40.
59. Lee IS, Liu Y, Narazaki M, Hibi M, Kishimoto T, Taga T: **Vav is associated with signal transducing molecules gp130, Grb2 and Erk2, and is tyrosine phosphorylated in response to interleukin-6.** *FEBS Lett* 1997, **401:**133-137.
60. Puto LA, Pestonjamasp K, King CC, Bokoch GM: **p21-activated kinase 1 (PAK1) interacts with the Grb2 adapter protein to couple to growth factor signaling.** *J Biol Chem* 2003, **278:**9388-9393.
61. Krawczyk C, Oliveira-dos-Santos A, Sasaki T, Griffiths E, Ohashi PS, Snapper S, Alt F, Penninger JM: **Vav1 controls integrin clustering and MHC/peptide-specific cell adhesion to antigen-presenting cells.** *Immunity* 2002, **16:**331-343.
62. Goetsch SC, Hawke TJ, Gallardo TD, Richardson JA, Garry DJ: **Transcriptional profiling and regulation of the extracellular matrix during muscle regeneration.** *Physiol Genomics* 2003, **14:**261-271.
63. Zhao P, Hoffman EP: **Embryonic myogenesis pathways in muscle regeneration.** *Dev Dyn* 2004, **229:**380-392.
64. Porter JD, Merriam AP, Leahy P, Gong B, Feuerman J, Cheng G, Khanna S: **Temporal gene expression profiling of dystrophin-deficient (mdx) mouse diaphragm identifies conserved and muscle group-specific mechanisms in the pathogenesis of muscular dystrophy.** *Hum Mol Genet* 2004, **13:**257-269.
65. Matecki S, Guibinga GH, Petrof BJ: **Regenerative capacity of the dystrophic (mdx) diaphragm after induced injury.** *Am J Physiol Regul Integr Comp Physiol* 2004, **287:**R961-R968.
66. Chen YW, Zhao P, Borup R, Hoffman EP: **Expression profiling in the muscular dystrophies: identification of novel aspects of molecular pathophysiology.** *J Cell Biol* 2000, **151:**1321-1336.
67. Zhao Y, Haginoya K, Sun G, Dai H, Onuma A, Iinuma K: **Platelet-derived growth factor and its receptors are related to the progression of human muscular dystrophy: an immunohistochemical study.** *J Pathol* 2003, **201:**149-159.
68. Jin P, Sejersen T, Ringertz NR: **Recombinant platelet-derived growth factor-BB stimulates growth and inhibits differentiation of rat L6 myoblasts.** *J Biol Chem* 1991, **266:**1245-1249.
69. Maroto M, Reshef R, Munsterberg AE, Koester S, Goulding M, Lassar AB: **Ectopic Pax-3 activates MyoD and Myf-5 expression in embryonic mesoderm and neural tissue.** *Cell* 1997, **89:**139-148.
70. Hoffman EP: **Genotype/phenotype correlations in Duchenne/becker dystrophy.** In *Molecular and Cell Biology of Muscular Dystrophy Volume 2*. 1st edition. Edited by: Partridge TA. London, Chapman & Hall; 1993:12-36.
71. 't Hoen PA, de Kort F, van Ommen GJ, den Dunnen JT: **Fluorescent labelling of cRNA for microarray applications.** *Nucleic Acids Res* 2003, **31:**e20.
72. Churchill GA: **Fundamentals of experimental design for cDNA microarrays.** *Nat Genet* 2002, **32:**490-495.
73. Yang YH, Speed T: **Design issues for cDNA microarray experiments.** *Nat Rev Genet* 2002, **3:**579-588.
74. Townsend JP: **Multifactorial experimental design and the transitivity of ratios with spotted DNA microarrays.** *BMC Genomics* 2003, **4:**41.
75. Huber W, Von Heydebreck A, Sultmann H, Poustka A, Vingron M: **Variance stabilization applied to microarray data calibration and to the quantification of differential expression.** *Bioinformatics* 2002, **18:**S96-S104.

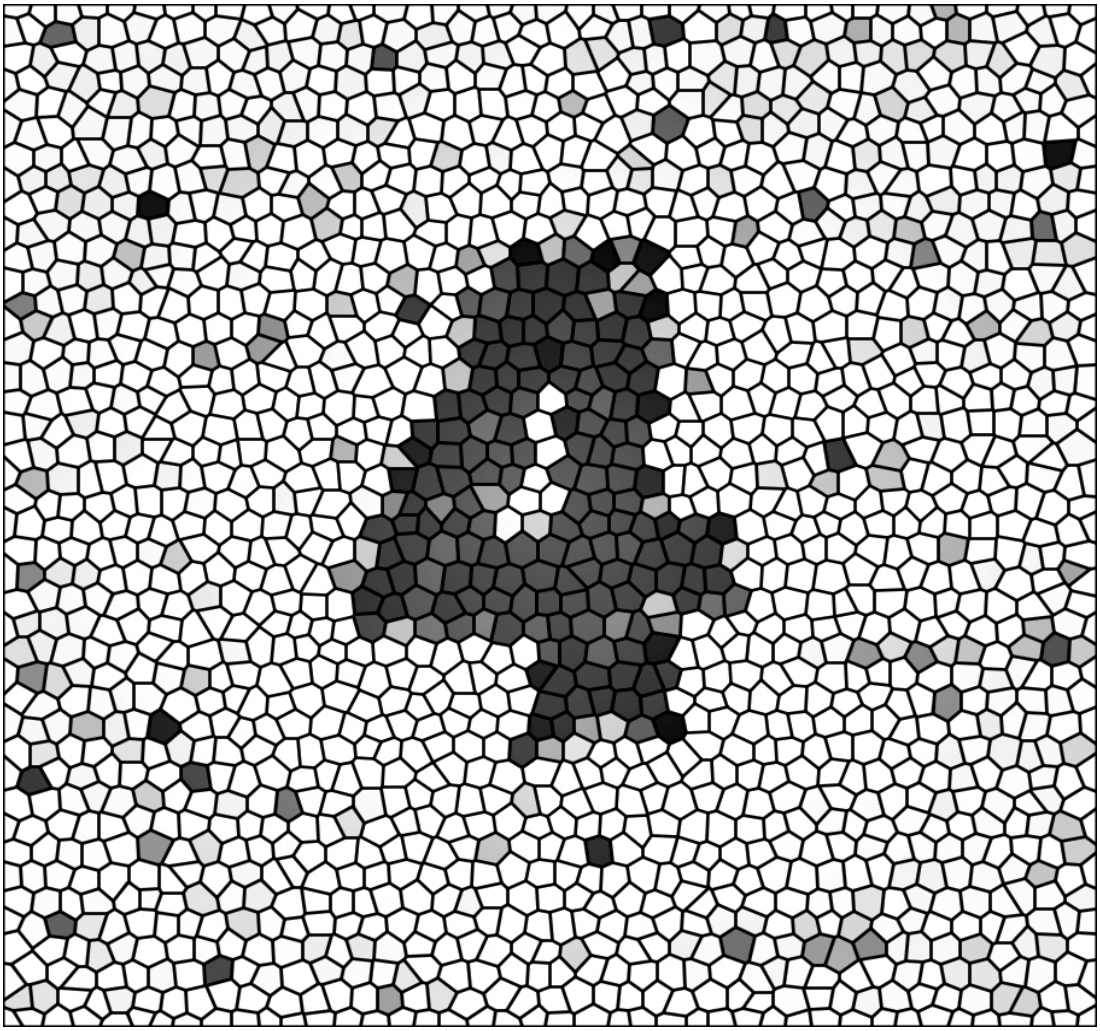
76. 't Hoen PA, Turk R, Boer JM, Sterrenburg E, de Menezes RX, van Ommen GJ, den Dunnen JT: **Intensity-based analysis of two-colour microarrays enables efficient and flexible hybridization designs.** *Nucleic Acids Res* 2004, **32:**e41.
77. Svensson BA, Kreeft AJ, van Ommen GJ, den Dunnen JT, Boer JM: **GeneHopper: a web-based search engine to link gene-expression platforms through GenBank accession numbers.** *Genome Biol* 2003, **4:**R35.
78. Bakay M, Zhao P, Chen J, Hoffman EP: **A web-accessible complete transcriptome of normal human and DMD muscle.** *Neuromuscul Disord* 2002, **12:**S125-S141.

Chapter 4

Common pathological mechanisms in mouse models for muscular dystrophies

R. Turk, E. Sterrenburg, C.G.C van der Wees, E.J. de Meijer,
R.X. de Menezes, S. Groh, K.P. Campbell, S. Noguchi,
G.J.B. van Ommen, J.T. den Dunnen, and P.A.C. ‘t Hoen

FASEB Journal 2006 20(1):127-129



The FASEB Journal • FJ Express

Common pathological mechanisms in mouse models for muscular dystrophies

R. Turk,*,† E. Sterrenburg,* C. G. C. van der Wees,* E. J. de Meijer,* R. X. de Menezes,* S. Groh,† K. P. Campbell,† S. Noguchi,‡ G. J. B. van Ommen,* J. T. den Dunnen,* and P. A. C. 't Hoen*,[1]

*Leiden University Medical Center, Center for Human and Clinical Genetics, Leiden, The Netherlands; †Howard Hughes Medical Institute, Department of Physiology and Biophysics, Iowa City, Iowa, USA; and ‡National Institute of Neuroscience, Department of Neuromuscular Research, Tokyo, Japan

To read the full text of this article, go to http://www.fasebj.org/cgi/doi/10.1096/fj.05-4678fje; doi: 10.1096/fj.05-4678fje

SPECIFIC AIMS

The aims of this study are to *1)* find common and distinct molecular mechanisms that underlie different forms of Duchenne/Becker (DMD/BMD) and limb-girdle muscular dystrophies (LGMD) and *2)* explore the feasibility of the use of microarrays for differential diagnosis of muscular dystrophies by establishing expression signatures in a large number of mouse models for muscular dystrophy.

PRINCIPAL FINDINGS

1. Expression profiling distinguishes severely affected dystrophin- and sarcoglyan-deficient mice from mildly affected dysferlin-deficient mice and control mice

We have carried out a comparative gene expression profiling of hind limb muscles of the following mouse models: dystrophin-deficient (*mdx, mdx3cv,* models for DMD), sarcoglycan-deficient (Sgca null, Sgcb null, Sgcg null, Sgcd null; models for LGMD-2C-F), dysferlin-deficient (Dysf null, SJL^{Dysf}; models for LGMD-2B), and sarcospan-deficient (Sspn null; no human disease known) mice. Two wild-type strains (C57Bl/6, C57Bl/10) were included as controls. All mice were 8 wk of age. A total of 2171 genes demonstrated differential expression between the 11 models ($P<0.05$, Benjamini and Hochberg correction for multiple testing). Hierarchical clustering of the expression profiles clearly separates two groups of mice. The first group contains the models for dystrophinopathy and sarcoglycanopathy, which have a severe dystrophic phenotype at the age of 8 wk and whose expression profiles are remarkably similar. The second group consists of models with a mild or unaffected phenotype.

2. Muscular dystrophies share many molecular and cellular responses

Analysis of the lists of 535 up-regulated and 493 down-regulated genes most discriminative for severely affected mice identified several major biological processes that are commonly altered in these muscular dystrophies (**Fig. 1**): increased cell adhesion (up-regulation of Icam1, P-selectin, and several integrins, laminins, and thrombospondins), inflammation (increased levels of many chemokines, cytokines, cytokine receptors, lymphocyte antigens), and regulation of muscle contraction (Atpa2, Calsequestrin 1 and 2, and Troponin C, I, T1 and T2). Many of the up-regulated genes localized either in the extracellular matrix (including 12 different collagens) or the lysosome. The most striking feature of the list of down-regulated genes is the participation of 53% of the genes (259 genes) in highly diverse metabolic processes, indicative of an overall decline in metabolic activity in dystrophic tissue.

3. Inflammation and remodeling of the extracellular matrix, sarcomere, and cytoskeleton are also evident in dysferlin-deficient mice, though present at lower levels, in agreement with the later age of onset and the earlier stage of the disease analyzed

The majority of up-regulated (101/154) and down-regulated (88/167) genes in dysferlin-deficient compared with wild-type strains were also differentially expressed between severely affected and mildly or nonaffected mouse models. Since the latter group includes the dysferlin-deficient mice, these are genes with subtle changes in dysferlin-deficient mice and higher fold-changes in the more severe models. The

[1] Correspondence: Center for Human and Clinical Genetics, Leiden University Medical Center, Wassenaarseweg 72, Leiden 2333 AL, Nederland. E-mail: p.a.c.hoen@lumc.nl

Upregulated genes

Level 1	Level 2	Level 3	Level 4	p-value	number of genes
Biological process	Cellular process				
		Cell Communication			
			Cell adhesion	1.58E-04	36
	Physiological Process				
		Response to stimulus		1.68E-05	82
			Response to stress	1.96E-04	44
				Response to wounding	7.20E-04	21
			Response to biotic stimulus	1.49E-07	55
				Defense respons	4.94E-08	51
	Organismal physiological process			1.73E-04	73
		Muscle Contraction			
			Regulation of muscle contraction	4.25E-04	6
Cellular Component	***Extracellular region***			1.49E-05	145
		Extracellular matrix		1.09E-04	32
			Basement membrane	5.60E-04	9
			Collagen	9.07E-08	12
		Extracellular space		1.52E-05	135
	Intracellular				
		Intracellular organelle			
			Intracellular membrane-bound organelle		
				Lysosome	4.46E-04	13
		Cytoplasm			
			Contractile fiber		
				Troponin complex	7.37E-04	4
Molecular Function	Binding				
		Protein binding		2.39E-04	138
	Signal transducer activity				
		Receptor activity			
			Transmembrane receptor activity		
				Hematopoietin/interferon class (D200-domain) cytokine receptor activity	8.85E-04	10
	Structural molecule activity			4.19E-04	37
		Extracellular matrix structural constituent		1.71E-08	16
			Extracellular matrix structural constituent conferring tensile strength	1.91E-07	11

Downregulated genes

Level 1	Level 2	Level 3	Level 4	p-value	number of genes
Biological process	Physiological process				
	Cellular physiological process			6.97E-06	328
		Cellular metabolism		1.56E-09	259
			Cofactor metabolism	1.28E-06	19
			Organic acid metabolism	1.04E-04	28
			Lipid metabolism	2.71E-05	35
			Alcohol metabolism	2.88E-07	10
			Carbohydrate metabolism	2.48E-12	40
Cellular component	Intracellular			5.26E-11	264
	Intracellular organelle			3.07E-08	228
		Mitochondrion		5.48E-24	72
			Mitochondrial membrane	3.29E-11	23
	Cytoplasm				
		Cytosol			
			Proteasome core complex	5.13E-04	6
Molecular function	***Catalytic activity***			3.01E-15	223
	Hydrolase activity				
		Endopeptidase activity			
			Threonine endopeptidase activity	8.31E-04	6
	Isomerase activity			1.35E-04	14
			Heme-copper terminal oxidase activity	8.46E-05	6
	Oxidoreductase activity			1.89E-10	52
		Oxidoreductase acting on CH-OH group of donors		1.59E-04	12
			Acting on CH-OH group of donors, NAD or NADP as acceptor	3.78E-04	11
		Oxidoreductase acting on heme group of donors		8.46E-06	6
			Acting on heme group of donors, oxygen as acceptor	8.46E-06	6
		Oxidoreductase acting on the aldehyde or oxo group of donors		1.49E-04	7
	Transferase activity			7.40E-06	80

Figure 1. Functional classification of significantly up- and down-regulated genes in severely affected mouse models. Genes up-regulated in dystrophic muscle from severely affected mouse models were grouped according to biological process, cellular component, and molecular function based on Gene Ontology classifications. Only branches of the GO-tree containing categories that were significantly overrepresented (displayed in italics; $P<0.001$) in the list of up- and down-regulated genes are shown. Listed P values are from a hypergeometric test that compares, for each category, the number of genes in the set of up- or down-regulated genes with the total number of genes present on the array in that category. Number of genes refers to the number of genes in the list of up- or down-regulated genes in a specific category

major pathogenic mechanisms in which these genes participate are depicted in the schematic diagram (**Fig. 3**). It appears that the molecular and cellular details of the pathological processes secondary to the different primary defects are highly similar for the muscular dystrophies analyzed, and are likely to contribute to the progression of these diseases.

4. Based on the differences in severity of the animal models analyzed, we identified biomarker genes for disease progression

Since the pathology is more progressive in sarcoglycan- and dystrophin-deficient animals than in dysferlin-deficient animals, we statistically identified a set of 33 biomarker genes whose expression levels correlate with disease progression and severity (**Fig. 2**).

5. For the first time, we show a molecular phenotype for sarcospan-deficient animals

Two genes up-regulated in both mildly and severely affected animal models, *S100a9*, coding for a phagocytic protein known as calgranulin B, and *Spp1*, coding for a cytokine known as osteopontin, were also significantly up-regulated in sarcospan-deficient mice. This observation suggests for the first time a subtle molecular phenotype for these mice. With quantitative RT-PCR, we confirmed the much higher expression of *Spp1* and *S100a9* in sarcospan-deficient mice (29- and 14-fold

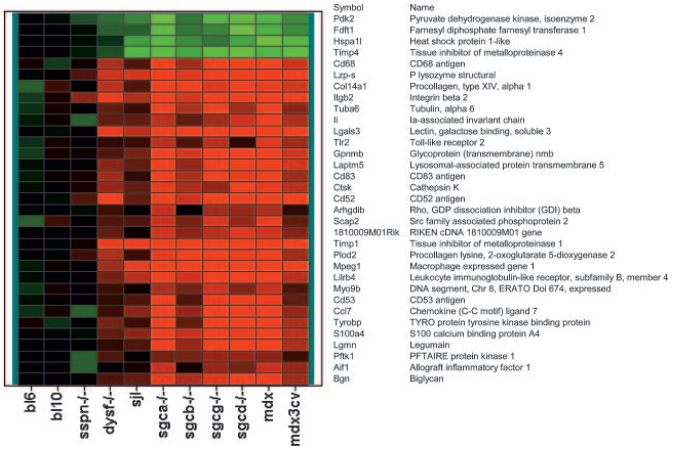


Figure 2. Biomarker genes for disease progression in muscular dystrophy. Heat map of averaged expression levels (relative to levels in wild-type mice) of genes that correlate with disease progression. The top 4 genes demonstrate significantly lower expression (displayed in green) in dysferlin-deficient mice compared with wild-type and sarcospan-deficient mice, and even lower expression levels in the more severely affected mouse models, whereas the bottom 29 genes demonstrate significantly higher (displayed in red) expression in dysferlin-deficient mice compared with wild-type and sarcospan-deficient mice, and even higher expression levels in the more severely affected mouse models.

increase over controls, respectively), as well as in the other muscular dystrophy models.

6. Based on identified differences in expression signatures between the muscular dystrophies and on the correlation with studies in human muscular dystrophy patients, we think that it would be possible to build a diagnostic classifier

A comparison was made between sarcoglycan-deficient and dystrophin-deficient mice, revealing 46 differentially expressed genes. This indicates there are subtle differences between the different muscular dystrophies. However, given the observed significant inter-individual variability in affected mice and the expected even greater variability in humans, construction of a diagnostic classifier would necessitate the analysis a large number of samples per disease. The analyzed expression patterns in mouse models may speed up the selection of genes relevant for classification.

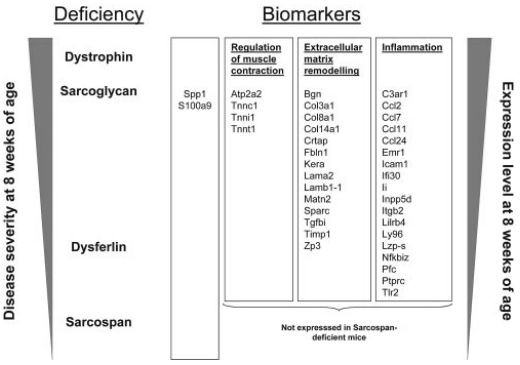


Figure 3. Genes associated with major pathogenic processes that are already apparent in dysferlin-deficient mice but more pronounced in dystrophin- and sarcoglycan-deficient mice at the age of 8 wk. *Spp1* and *S100a9* expression levels are also elevated in sarcospan-deficient mice and may be the first visible markers for muscular dystrophy.

CONCLUSIONS AND SIGNIFICANCE

We have found remarkable similarity in the expression patterns of dystrophin-, sarcoglycan-, and dysferlin-deficient mice. Genes functioning in the inflammatory response and structural organization are significantly up-regulated compared with wild-type mice, whereas metabolism genes are down-regulated. We conclude that common pathogenic mechanisms underlie the onset and progression of different forms of muscular dystrophy in mice (**Fig. 3**). Given the similarity with published human studies on specific forms of muscular dystrophy, these pathogenic mechanisms may also contribute to the different forms of muscular dystrophy in humans. Moreover, we have identified sets of biomarker genes that can be used to monitor disease progression in muscular dystrophies. Finally, by recognition of disease-specific expression signatures, we took the first step toward the development of expression profiling-based classification as a powerful diagnostic approach for muscular dystrophies. FJ

The FASEB Journal express article 10.1096/fj.05-4678fje. Published online November 23, 2005.

Common pathological mechanisms in mouse models for muscular dystrophies

R. Turk,*,† E. Sterrenburg,* C. G. C. van der Wees,* E. J. de Meijer,* R. X. de Menezes,* S. Groh,† K. P. Campbell,† S. Noguchi,‡ G. J. B. van Ommen,* J. T. den Dunnen,* and P. A. C. 't Hoen*

*Leiden University Medical Center, Center for Human and Clinical Genetics, Leiden, Nederland; †Howard Hughes Medical Institute, Department of Physiology and Biophysics, Iowa City, Iowa; and ‡National Institute of Neuroscience, Department of Neuromuscular Research, Tokyo, Japan

Corresponding author: Dr. P.A.C. 't Hoen, Center for Human and Clinical Genetics, Leiden University Medical Center, Wassenaarseweg 72, 2333 AL Leiden, Nederland. E-mail: p.a.c.hoen@lumc.nl

ABSTRACT

Duchenne/Becker and limb-girdle muscular dystrophies share clinical symptoms like muscle weakness and wasting but differ in clinical presentation and severity. To get a closer view on the differentiating molecular events responsible for the muscular dystrophies, we have carried out a comparative gene expression profiling of hindlimb muscles of the following mouse models: dystrophin-deficient (*mdx*, *mdx*^{3cv}), sarcoglycan-deficient (Sgca null, Sgcb null, Sgcg null, Sgcd null), dysferlin-deficient (Dysf null, SJL^{Dysf}), sarcospan-deficient (Sspn null), and wild-type (C57Bl/6, C57Bl/10) mice. The expression profiles clearly discriminated between severely affected (dystrophinopathies and sarcoglycanopathies) and mildly or nonaffected models (dysferlinopathies, sarcospan-deficiency, wild-type). Dystrophin-deficient and sarcoglycan-deficient profiles were remarkably similar, sharing inflammatory and structural remodeling processes. These processes were also ongoing in dysferlin-deficient animals, albeit at lower levels, in agreement with the later age of onset of this muscular dystrophy. The inflammatory proteins Spp1 and S100a9 were up-regulated in all models, including sarcospan-deficient mice, which points, for the first time, at a subtle phenotype for Sspn null mice. In conclusion, we identified biomarker genes for which expression correlates with the severity of the disease, which can be used for monitoring disease progression. This comparative study is an integrating step toward the development of an expression profiling-based diagnostic approach for muscular dystrophies in humans.

Key words: microarray • dystrophin-glycoprotein complex • inflammation • extracellular matrix • biomarker

Muscular dystrophies are a heterogeneous group of inherited neuromuscular disorders characterized by progressive muscle wasting and weakness. The genetic defects underlying many muscular dystrophies have been elucidated (1, 2). A particular subset

of muscular dystrophies is caused by mutations in genes coding for constituents of the dystrophin-associated glycoprotein complex (DGC). The DGC is a multimeric protein complex composed of integral, peripheral, and cytoplasmic proteins expressed at the sarcolemma of muscle fibers. One likely role of the DGC is to maintain the sarcolemma stability by providing a physical link between the extracellular matrix and the actin cytoskeleton. This link occurs through the trio dystrophin-dystroglycan-laminin-2. Dystrophin binds to cytoskeletal F-actin and the transmembrane β-subunit of dystroglycan. Dystroglycan binds to laminin-2 in the extracellular matrix via its α-subunit (3). The trio dystrophin-dystroglycan-laminin-2 is stabilized by other DGC components such as the sarcoglycan-sarcospan complex (SGC), which resides in the sarcolemma. In skeletal muscle, the SGC contains four glycosylated subunits (α-, β-, γ-, and δ-sarcoglycan) and sarcospan (4).

Although genetic mutations in one of the DGC components frequently result in destabilization or mislocalization of the entire complex, the clinical presentation of the gene defects is variable: dystrophin-deficiency (dystrophinopathy) results in either the lethal Duchenne or the milder Becker muscular dystrophy (DMD/BMD) (5), dystroglycan-deficiency is likely to be embryonically lethal (6), and laminin 2-deficiency results in congenital muscular dystrophy (MDC1A) (7). Mutations in sarcoglycan-proteins (sarcoglycanopathies) are responsible for several recessive forms of limb-girdle muscular dystrophy (LGMD-2C-F, reviewed in refs 8, 9), whereas sarcospan has not been associated with human disease (10).

Mutations in dysferlin, a muscle membrane protein that plays a role in membrane repair (11), cause the non-DGC related muscular dystrophies LGMD-2B and Myoshi myopathy (12). In these relatively late-onset diseases, the defective membrane repair system is ultimately unable to cope with contraction-induced injuries to the sarcolemma, explaining the observed muscle degeneration (11, 13).

The pathological manifestation of the different muscular dystrophies at the histological level is similar. Dystrophic muscle tissue is hallmarked by myofiber degeneration, infiltration of inflammatory cells, and subsequent formation of foci of fibrotic and adipose tissue. These manifestations are secondary to the primary defect. The decisive secondary factors responsible for the variability in the clinical phenotypes are still unknown.

Muscular dystrophies can be clinically grouped based on several indicators, namely age of onset, severity, presentation, affected musculature, and genetic inheritance (14). The clinical variability within each muscular dystrophy complicates the diagnosis of muscular dystrophies, particularly in young children. Consequently, a large array of molecular diagnostic methods, frequently combining analyses at the protein, DNA, and mRNA level, has to be applied to come to a specific diagnosis (15).

Gene expression profiling studies have generated more detailed insight in the molecular processes underlying DMD (16–19). However, few microarray datasets have been published on other muscular dystrophies (13, 20–23). The goal of the present study is to assess whether the muscular gene expression patterns in the individual muscular dystrophies are sufficiently different to allow for microarray-based classification, which could greatly facilitate diagnosis. In addition, comparison of the different gene expression patterns in muscular dystrophies will help to understand the molecular mechanisms underlying the specificity of the clinical symptoms.

Human muscle biopsies display large inter-individual variation in gene expression due to differences in genetic make-up, age, and exposure to environmental factors, necessitating the evaluation of large patient groups. To facilitate the delineation of the molecular mechanisms underlying the phenotypic characteristics, we decided to perform large-scale gene expression profiling of mouse models known to recapitulate different human muscular dystrophies (Fig. 1). Mice with early onset/severe phenotypes, represented by dystrophinopathies (*mdx* and *mdx*^{3cv}) and sarcoglycanopathies (Sgca null, Sgcb null, Sgcg null and Sgcd null), and late onset/mild phenotypes, represented by dysferlinopathies (SJL^{Dysf} and Dysf null), were included in the study (Table 1). Sarcospan-deficient (Sspn null) mice, with no apparent muscular phenotype, were also included. Two wild-type strains (C57Bl/6 and C57Bl/10) served as controls. Although these mouse models have been extensively studied on phenotypic characteristics, like severity, age of onset, loss of the DGC, lifespan, and the temporal manifestation of the pathology (reviewed in ref 24), a large-scale comparison of gene expression patterns has not yet been performed. As a starting point, and to avoid age-dependent differences in gene expression, we evaluated young adult male mice (8 wk of age). At this age, *mdx* mice show most prominent changes in gene expression compared with wild-type (25) and sarcoglycan-deficient animals are severely dystrophic. On the contrary, the dysferlin-deficient mice are only mildly affected at this age (11, 26). Our data demonstrate that gene expression patterns classify the animals according to the severity of the disease and point at common secondary disease mechanisms.

MATERIALS AND METHODS

Target preparation and hybridization

Hindlimb muscle tissue (m. quadriceps femoris) was isolated from the following mice at the age of 8 wk (2 individuals per strain): *mdx* (*mdx*) (C57Bl/10ScSn-mdx/J, The Jackson Laboratory x C57Bl/6NCrl, Charles River x CBA/JCrl, Charles River), *mdx*^{3cv} (*mdx*^{3cv}) (27), Sgca null (Sgca–/–) (28), Sgcb null (Sgcb–/–) (29), Sgcb null-2 (Sgcb–/–2) (30), Sgcg null (Sgcg–/–) (31), Sgcd null (Sgcd–/–) (32), Dysf null (Dysf–/–) (11), SJL (SJL^{Dysf}) (SJL/J, The Jackson Laboratory), Sspn null (Sspn–/–) (33), C57Bl/10 (Bl10) (C57Bl/10ScSnOlaHsd, Harland), C57Bl/6 (Bl6) (C57Bl/6JOlaHsd, Harland), and hDMD mice, which contain the full-length 2.3 Mb human dystrophin gene (hDMD, 't Hoen et al., unpublished observations). Expression profiles of the hDMD mice will be discussed in a separate manuscript. The expression profiles of the Sgcb null-2 mice were not included in the analysis, for reasons indicated in the Discussion. Total RNA was isolated as described previously (34). cRNA was prepared by linear amplification and concurrent incorporation of amino-allyl UTP, followed by chemical coupling to monoreactive Cy3 or Cy5 dyes (35). Labeled targets (1.5 µg cRNA per target) were hybridized overnight on prehybridized murine oligonucleotide microarrays (65-mer with 5′-hexylaminolinker, Sigma-Genosys mouse 7.5 K oligonucleotide library, printed in duplicate) using an automatic hybridization station (GeneTac, Genomic Solutions, Ann Arbor, MI). Posthybridization washes were performed as described previously (35).

Data analysis

Hybridizations were performed in a dye-swap fashion using a randomized design, while avoiding co-hybridization of samples from the same model (Supplemental Table 1). This experimental design generates eight data points per gene per mouse model (2 biological replicates with 4

technical replicates each). The randomized experimental design facilitates an intensity-based analysis procedure, since ratio-based analysis procedures become complicated and inefficient in multi-class comparisons (36). Spot intensities were evaluated with feature extraction software (GenePix Pro 5, Axon, Union City, CA). Local background corrected, median spot intensities were normalized simultaneously for all microarray experiments using variance stabilization and normalization (VSN) as described previously (37). This transformation coincides with the natural logarithm for the high intensities. Two filtering criteria were applied. First, genes that were flagged in at least four out of eight observations for a particular genotype were excluded to minimize the influence of unreliable spots (specks, high background). Since samples were randomly hybridized with each other, and spots were flagged only when signals in both channels were below background, it is very unlikely that genes that are expressed only in one or a few models were excluded from the analysis. Second, 88 genes previously shown to demonstrate variation in expression in the muscle due to differences in the genetic background (34) were excluded from the analysis. In the end, 2751 genes were included in the analysis.

Microarray data have been made available through the GEO data repository of the National Center for Biotechnology Information under series GSE2112.

Evaluation of differential expression between mouse models

Significance levels of differential gene expression differences between different classes of mouse models were calculated in R (38) using linear regression models that take the dye effect into account. All fixed-effects model were of this form:

$$(1)\ Y = \alpha + \beta \cdot \text{class} + \gamma \cdot \text{dye}$$

In all comparisons, the expression levels of Cy3-labeled targets were captured in the intercept (dye=0), whereas dye equaled 1 for Cy5-labeled targets.

In the comparison of the different strains, the expression levels of C57Bl/10 were captured in the intercept (class=0), and different β coefficients were estimated for the other strains.

In the comparison of severely affected (*mdx*, *mdx*^{3cv}, Sgca null, Sgcb null, Sgcd null, Sgcg null) and non- or mildly affected (C57Bl/10, C57Bl/6, Sspn null, Dysf null, SJL^{Dysf}) models, the expression levels of non- or mildly affected were captured in the intercept (class=0), and class equaled 1 for severely affected mouse models.

Similary, in the comparison of dystrophin-deficient (*mdx*, *mdx*^{3cv}) and sarcoglycan-deficient (Sgca null, Sgcb null, Sgcd null, Sgcg null) mouse models, expression levels of dystrophin-deficient mice were captured in the intercept (class=0), and class equaled 1 for sarcoglycan-deficient mice.

Similarly, in the comparison of dysferlin-deficient (Dysf null, SJL^{Dysf}) vs. control (C57Bl/10, C57Bl/6) mouse models, expression levels of control mice were captured in the intercept (class=0), and class equaled 1 for dysferlin-deficient mice.

P values for class were corrected for multiple testing using the method proposed by Benjamini and Hochberg (BH-correction) (39).

To find biomarker genes for disease progression, we fitted the following models:

$$(2)\ Y = \alpha_2 + \beta_2 \cdot \text{class2} + \gamma_2 \cdot \text{dye}$$

$$(3)\ Y = \alpha_3 + \beta_3 \cdot \text{class3} + \gamma_3 \cdot \text{dye}$$

In model 2, we compared nonaffected (intercept; C57Bl/10, C57Bl/6, Sspn null) with mildly affected (class2=1; Dysf null, SJL^{Dysf}), and in model 3, we compared mildly affected (intercept; Dysf null, SJL^{Dysf}) with severely affected (class3=1; *mdx*, mdx^{3cv}, Sgca null, Sgcb null, Sgcd null, Sgcg null). We selected genes with significant *P* values (BH-corrected *P* value<0.05) in both comparisons and then identified up-regulated biomarker genes as genes with both β_2 and β_3 greater than zero, whereas for down-regulated biomarker genes both β_2 and β_3 were smaller than zero.

Clustering

Unsupervised hierarchical clustering was performed using the Functional Genomics package within Spotfire DecisionSite 7.3. Expression levels were scaled such that the average expression per gene over all conditions was zero. Average linkage was the clustering method applied, and Euclidean distance was taken as similarity measure.

Functional annotation

Recent functional annotation of genes was obtained from the SOURCE database (http://source.stanford.edu) (40). Genes were classified based on gene ontology using the web-based tool Gene Ontology Tree Machine (GOTM; http://genereg.ornl.gov/gotm) (41). Using this tool, we identified the functional processes that were overrepresented in lists of up- and down-regulated genes compared with the list of all the genes on the array using a hypergeometric test. Only categories containing at least three differentially expressed genes and with a *P* value < 0.001 were considered.

Quantitative RT-PCR

Quantitative RT-PCR was performed as described previously (25). Gene expression levels were calculated using the gene expression macro provided by Bio-Rad (Bio-Rad, Hercules, CA) and normalized to glyceraldehyde-3-phosphate dehydrogenase (GAPDH, stable expression in all samples) expression levels. Primer sequences are available on request.

RESULTS

We applied large-scale gene expression profiling to make a global inventory of differences in gene expression levels between mouse models for different muscular dystrophies. Nine mouse models with different genetic defects (Table 1) and two wild-type mouse strains were compared. Total RNA, isolated from hindlimb muscle of two individual mice, was amplified and hybridized to murine microarrays containing 7445 gene-specific oligonucleotides, spotted in duplicate. A dye-balanced, randomized hybridization design was chosen to avoid potential dye, sample processing, and hybridization biases (Supplemental Table 1). Combined with an intensity-based

analysis procedure, this experimental design is most efficient when comparing large numbers of groups (36).

After being filtered for unreliable spots and for genes the expression of which was known to be affected by differences in genetic backgrounds (34), 2171 genes demonstrated differential expression in the 11 models ($P<0.05$, BH-correction). These genes, as well as their expression levels in the different models, are listed in Supplemental Table 2. We applied unsupervised hierarchical clustering on the profiles of the 300 most significantly differentially expressed genes to determine the similarity between the mouse models (Fig. 2*A*). According to the hierarchical clustering, two groups can be clearly discerned. The first group (I) contains the models for dystrophinopathy and sarcoglycanopathy, which have a severe dystrophic phenotype at the age of 8 wk. The second group (II) consists of models with a mild or unaffected phenotype.

The identity of the genes that contribute most to the distinction between group I and group II was determined by calculating per gene the statistical significance between the two groups with a linear regression model. After correction for multiple testing, 1028 genes appear to function as general markers for dystrophic muscles (Supplemental Table 3). In Table 2, we present the genes that are at least twofold up-regulated or down-regulated (128 and 46 genes, respectively) in the mouse models for dystrophinopathy and sarcoglycanopathy, compared with mildly affected or unaffected muscle tissue.

To determine the biological variation in gene expression between the two individual mice from each strain, we performed a hierarchical clustering on the 300 most significant genes using the average gene expression levels of all separate individuals (Fig. 2*B*). Within the group of mildly and nonaffected mouse models (group II), both individuals of each strain clustered together, demonstrating a high level of similarity in their gene expression profiles. In contrast, within the group of severely affected mouse models (group I), the gene expression profiles of the biological replicates were located in different clusters, which is indicative of a higher variability in gene expression profiles between individuals than between models.

To assess whether the absence of one DGC component affects the expression of other DGC components, we evaluated the expression levels of seven genes involved in the DGC (*Dmd*, *Sgca*, *Sgcb*, *Sgcg*, *Sgcd*, *Sspn*, and *Dysf*; Fig. 3). The knockout mouse models demonstrated clearly reduced mRNA levels of the targeted gene, with an exception for *Sgcd* in Sgcd null mice and *Dysf* in dysferlin-deficient mice, in which both only a relatively modest decrease in signal intensity is shown. The effect of a point mutation (*mdx*, *mdx*^{3cv}, and SJL^{Dysf}) seems to have a milder influence on the mRNA levels, probably due to less efficient degradation by nonsense-mediated mRNA decay (NMD), which depends on the location of the mutation (42, 43). In the mouse models for sarcoglycanopathy, a decrease in expression levels from other DGC components was observed, with largest effects in Sgca null mice. The gene expression of dysferlin, on the other hand, was up-regulated in severely dystrophic mouse models, with the exception of Sgcb null mice.

We implemented a recently published bioinformatics tool based on gene ontology (GOTM; ref 41) to investigate the functional processes active in severely affected dystrophic muscle tissue (group I). This tool associates sets of genes (e.g., up-regulated genes) with functional processes and determines whether these processes are represented at a higher frequency than can be

expected by chance, i.e., as represented in the reference gene list, consisting of all genes on the microarray. The results of this analysis are presented in Fig. 4. The most strikingly overrepresented biological processes in the list of up-regulated genes are cell adhesion (*Icam1*, *P-selectin* and several integrins, laminins, and thrombospondins), inflammation (many chemokines, cytokines, cytokine receptors, lymphocyte antigens), and regulation of muscle contraction (*Atpa2, Calsequestrin 1* and *2,* and *Troponin C, I, T1,* and *T2*). Many of the up-regulated genes localized either in the extracellular matrix (including 12 different collagens) or in the lysosome. The most striking feature of the list of down-regulated genes is the participation of 53% of these genes (259 genes) in highly diverse metabolic processes, indicative of an overall decline in metabolic activity in dystrophic tissue. A large number of the gene products of these genes are localized in the mitochondrion (72 genes; *P* value for overrepresentation: 5.48×10^{-24}). The individual genes in the listed functional categories are given in Supplemental Table 4.

Although gene expression profiles of mouse models for dystrophinopathy show high similarity with those of mouse models for sarcoglycanopathy, some genes displayed statistically significant differential gene expression between these groups. With the linear regression model described in Materials and Methods, we found 46 differentially expressed genes ($P<0.05$ after BH-correction for multiple testing, Supplemental Table 5). A selection of four genes with greater than twofold lower expression and five genes with greater than twofold higher expression in mouse models with sarcoglycanopathy compared with mouse models for dystrophinopathy is presented in Table 3.

The clustering diagrams in Fig. 2 suggest that gene expression patterns of mildly affected dysferlin-deficient animals can be discerned from those of healthy mice. With a linear regression model, 321 genes were identified that displayed statistically significant ($P<0.05$ after BH-correction) differences in expression levels between dysferlin-deficient (SJL^{Dysf} and Dysf null mice) and the two wild-type strains (Supplemental Table 6). Interestingly, the majority of up-regulated (101/154) and down-regulated (88/167) genes were also differentially expressed between group I (severely affected) and group II (mildly or nonaffected). Since the latter group includes the dysferlin-deficient mice, these are genes with subtle changes in dysferlin-deficient mice and higher fold changes in the more severe models. We statistically evaluated which genes show a significant trend in expression level from nonaffected to mildly and to severely affected animal models. These genes can be seen as biomarker genes for disease severity. The heat map in Fig. 5 illustrates the expression levels of the 33 identified biomarker genes.

Two genes up-regulated in both mildly and severely affected animal models, *S100a9,* coding for a phagocytic protein known as calgranulin B, and *Spp1*, coding for a cytokine known as osteopontin, were also up-regulated in sarcospan-deficient mice ($P<10^{-7}$ after correction for multiple testing). This observation suggests for the first time a subtle molecular phenotype for these mice. With quantitative RT-PCR, we confirmed the much higher expression of *Spp1* and *S100a9* in sarcospan-deficient mice (29- and 14-fold increase over wild-types, respectively), as well as in the other muscular dystrophy models (Supplemental Fig. 1).

DISCUSSION

In this report, we have classified muscular dystrophies based on gene expression profiles in skeletal muscle of 8-wk-old mice. Cluster analysis readily distinguished severely affected mouse

models (dystrophinopathies and sarcoglycanopathies) from dysferlin- and sarcospan-deficient and wild-type mice. Earlier studies demonstrated a high similarity in the histology of muscles from mouse models for dystrophinopathy and sarcoglycanopathies at the age of 8 wk (27–29, 32, 44, 45). We now show that common denominators in dystrophinopathies and sarcoglycanopathies are also apparent at the molecular level. The robustness of this classification is further illustrated by the fact that one of the two mouse models for β-sarcoglycan deficiency (29, 30) that were originally included in the study, clustered with the non- or mildly affected mice. Since this was highly unexpected, a detailed analysis was performed using histological techniques and RT-PCR, showing that one of the two mice analyzed was incorrectly genotyped as a Sgcb null mouse instead of wild-type. As a consequence, we had to exclude the second Sgcb null model from further analysis.

Gene expression levels of DGC components decrease in DGC-related muscular dystrophies

Although the entire DGC is absent from the sarcolemma in dystrophinopathies, whereas only the SGC is lost in sarcoglycanopathies, we observed a general decrease in mRNA levels of the DGC components in both dystrophin- and sarcoglycan-deficient mice. The instability of the DGC apparently triggers a negative feedback to transcriptional levels of the DGC components. Notable is the up-regulation of dysferlin, which is indicative for an increased need for membrane repair systems to prevent membrane leaking or rupture, a common feature for muscular dystrophies (46, 47). Other important processes secondary to the genetic defect and the loss of a functional DGC are induction of an immune response, increased expression of cellular adhesion and extracellular matrix proteins, and changes in cytoskeletal organization. These processes were also shared between human patients with DMD and LGMD-2D (α-sarcoglycan deficiency) (16).

The inflammatory response in muscular dystrophy can be divided into multiple components

By looking at markers for different types of immune cells (48), we found evidence for the infiltration and accumulation of macrophages [*Cd68*, *Lgals3* (*Mac-2*), *Mpeg1*, *Clecsf12*], B-cells (*Cd83, Blnk*), T-cells (*CD8b1, Ly6e*), and NK-cells (*Ypel1*). The recruitment of inflammatory cells is probably mediated by CC-class chemokines that are up-regulated (*Ccl2, Ccl6, Ccl7, Ccl8*, and *Ccl9*), as was reported for *mdx* mice (25, 49, 50). Furthermore, an activation of components of the complement system (*C1qa, C1qg, C1qr1, C1qTNF3*, and *C3ar1*) was observed, which contributes to the inflammatory response in dystrophic muscle by further damaging cell membranes and releasing complement split products that attract macrophages for phagocytosis.

Genes involved in sarcomeric organization and extracellular matrix formation are up-regulated in both sarcoglycan- and dystrophin-deficient muscles

Among the up-regulated extracellular matrix proteins are 12 different collagens, laminin α4, B1 and gamma1, and collagen binding proteins such as biglycan. The up-regulated cytoskeletal proteins include the intermediate filament vimentin, the microtubular components tubulin α-1, -2, -6, and β-5, the actin-interacting protein transgelin2 and the sarcomeric troponins I (skeletal, slow), T1 (skeletal, slow), T2 (cardiac), C (cardiac / slow skeletal), most of which having been reported before as up-regulated in *mdx* mice (25, 50–52). Such up-regulation of extracellular and

intracellular structural proteins, which is also found in human DMD patients (17–19, 53) might compensate for the loss of force-generating capacity due to the instability of the DGC-mediated link between the cytoskeleton and the extracellular matrix.

Dystrophic muscle tissue experiences a metabolic crisis

The large number of down-regulated genes functioning in diverse metabolic processes reflects probably the metabolic crisis also seen in human DMD, LGMD-2D, FSHD, and nemaline myopathy patients (16, 22, 23). From our and data of others (51, 52), we conclude that the metabolic crisis in *mdx* mice is less severe than in sarcoglycan-deficient mice. However, a reduction in the respiratory rate of mitochondria in skeletal muscle of both *mdx* mice and DMD patients was also found in previous studies using other techniques (54, 55).

Differential gene expression between dystrophinopathies and sarcoglycanopathies

Despite the high similarities between dystrophinopathies and sarcoglycanopathies at 8 wk of age, 46 genes were found significantly differentially expressed between these groups. This may indicate that there are some subtle differences between dystrophinopathies and sarcoglycanopathies. These differences may become more prominent at higher age, since the *mdx* mice, in contrast to the sarcoglycan-deficient animals, make an almost complete recovery. A more extensive temporal gene expression profiling study together with functional analysis of the identified proteins is necessary to characterize the differences (work in progress). Cysteine and glycine-rich protein 3 (*Csrp3*), a positive regulator myogenesis (56), was the most prominent, higher expressed gene in sarcoglycan-deficient compared with dystrophin-deficient animals. Interestingly, it has been found before that *Csrp3* is also up-regulated in human FSHD but not DMD patients (22). This exemplifies that muscular dystrophies share some aberrations in gene expression and differ in others, and that it would be feasible to find a specific disease signature based on the measurement of the expression of combinations of genes.

Vascular irregularities are not likely to contribute to muscular dystrophy

Mutations in smooth muscle SGC-components (Sgcb, Sgcg, Sgcd) lead to vascular irregularities, which are thought to exacerbate the cardiac pathology seen in mice deficient for the individual components (29, 57, 58). Alpha-sarcoglycan is not expressed in smooth muscle and therefore the Sgca null mouse does not have altered SGC expression in smooth muscle. We show that the gene expression profiles of Sgca null mice are highly similar to the gene expression profiles of the other sarcoglycan-deficient mouse models, corroborating the suggestion of Durbeej et al. that vascular irregularities are not likely to contribute to the skeletal muscle pathology (24).

Differential gene expression precedes histological changes in dysferlinopathies

The phenotype of dysferlin-deficient mice is less progressive than that of the dystrophin- and sarcoglycan-deficient mice (11, 26). HE staining shows only few necrotic fibers at the age of 8 wk, corresponding with the higher age of onset of dysferlinopathies, both in mice and humans. Still, we observe significant differences in gene expression between wild-type mice and dysferlin-deficient mice at the age of 8 wk. Among the up-regulated genes that have also been found in an earlier study in SJL mice (13) are: Troponins T1, C, and I, *Ccl2*, *Cd53*, *Cd68*, *Cd83*, P lysozyme structural, tissue inhibitor of metalloproteinase 1 and 4. Many of these genes (Fig. 5)

are more severely changed in the severely affected mouse models. Thus, in dysferlin-deficient mice, similar to dystrophin and sarcoglycan-deficient mice, inflammatory and muscle remodeling processes play an important role (confirmed by gene ontology analysis with GOTM, not shown). Although differences in the expression of these genes are still limited at the age of 8 wk, they may become more prominent at higher ages. In the *mdx* mouse, however, the expression of these genes is already increased at 4 wk and peaks at 6 wk (25), in agreement with the earlier age of onset in dystrophinopathies. Consequently, these genes can be regarded as markers for disease progression.

A molecular phenotype for sarcospan-deficient mice

Genes, coding for secreted phosphoprotein 1 (Spp1 or Osteopontin), functioning in chemotaxis, and the phagocytic S100 calcium binding protein A9 (S100a9 or Calgranulin B), are also up-regulated in sarcospan-deficient mice. Differential expression of these genes indicates that sarcospan-deficiency may induce a subtle immune response in muscle. *Spp1* seems to be an extremely sensitive marker for muscle pathology, since it was picked up in other gene expression profiling studies in dystrophin-deficient humans and mice (18, 51).

Inter-individual variation in gene expression levels

Our study shows that the inter-individual differences in expression levels are considerable, even in genetically identical mice. Consequently, the subtle differences that discriminate between the muscular dystrophies will only become significant when large sample groups are analyzed. We suppose that the inter-individual differences in the severely affected mice are related to the physical behavior or the amount of exercise. An increase in muscle activity leads to an increase in sarcolemmal rupture, which initiates the secondary processes responsible for the severity of the disease (59, 60). Recently observed inter-individual variability in satellite cell number (61) and possibly regenerative potential in dystrophic mice, may also contribute to the observed variability. Due to the interindividual variability, at least 4 animals per group were required for meaningful comparisons. This is why classes of genotypes were combined in the performed statistical tests (e.g., the class of dysferlin-deficient mice, consisting of the 2 SJL^{Dysf} and 2 dysferlin null individuals). In humans, many more genetic and environmental factors will contribute to the inter-individual variation (62), and building of a robust classifier necessary for muscular gene expression profiling-based diagnosis of muscular dystrophies would necessitate the analysis a large number of samples per group. On the other hand, analysis of disease progression using the biomarker set described here, would be more readily achievable, although more definitive correlations between expression levels of the identified biomarkers and other pathological parameters need to be established. This type of analysis would be useful to monitor response to treatment.

CONCLUSION

In summary, we have found remarkable similarity in the expression patterns of dystrophin-, sarcoglycan-, and dysferlin-deficient mice. Genes functioning in the inflammatory response and structural organization are significantly up-regulated compared with wild-type mice, whereas metabolism genes are down-regulated. We conclude that common pathogenic mechanisms underlay the onset and progression of different forms of muscular dystrophy in mice. Given the

similarity with published human studies on specific forms of muscular dystrophy, these pathogenic mechanisms may also contribute to the different forms of muscular dystrophy in humans. Furthermore, we have identified sets of biomarker genes that can be used to monitor disease progression in muscular dystrophies, thereby taking the first step toward the development of a diagnostic approach for muscular dystrophies by expression profiling-based classification.

ACKNOWLEDGMENTS

This work was supported by the Nederlandse Stichting voor Wetenschappelijk Onderzoek (NWO), the Center for Biomedical Genetics (CBG), and the Centre for Medical Systems Biology (CMSB) established by the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research (NGI/NWO).

REFERENCES

1. Cohn, R. D., and Campbell, K. P. (2000) Molecular basis of muscular dystrophies. *Muscle Nerve* **23,** 1456–1471

2. Dalkilic, I., and Kunkel, L. M. (2003) Muscular dystrophies: genes to pathogenesis. *Curr. Opin. Genet. Dev.* **13,** 231–238

3. Ozawa, E., Yoshida, M., Suzuki, A., Mizuno, Y., Hagiwara, Y., and Noguchi, S. (1995) Dystrophin-associated proteins in muscular dystrophy. *Hum.Mol.Genet.* 4 Spec No, 1711-1716

4. Straub, V., Duclos, F., Venzke, D. P., Lee, J. C., Cutshall, S., Leveille, C. J., and Campbell, K. P. (1998) Molecular pathogenesis of muscle degeneration in the delta-sarcoglycan-deficient hamster. *Am. J. Pathol.* **153,** 1623–1630

5. Hoffman, E. P., Brown, R. H., Jr., and Kunkel, L. M. (1987) Dystrophin: the protein product of the Duchenne muscular dystrophy locus. *Cell* **51,** 919–928

6. Williamson, R. A., Henry, M. D., Daniels, K. J., Hrstka, R. F., Lee, J. C., Sunada, Y., Ibraghimov-Beskrovnaya, O., and Campbell, K. P. (1997) Dystroglycan is essential for early embryonic development: disruption of Reichert's membrane in Dag1-null mice. *Hum. Mol. Genet.* **6,** 831–841

7. Tome, F. M. S., Evangelista, T., Leclerc, A., Sunada, Y., Manole, E., Estournet, B., Barois, A., Campbell, K. P., and Fardeau, M. (1994) Congenital muscular dystrophy with merosin deficiency. *Comp.Rend.Acad.Sci.(Paris)* **317,** 351–357

8. Ozawa, E., Noguchi, S., Mizuno, Y., Hagiwara, Y., and Yoshida, M. (1998) From dystrophinopathy to sarcoglycanopathy: evolution of a concept of muscular dystrophy. *Muscle Nerve* **21,** 421–438

9. Lim, L. E., and Campbell, K. P. (1998) The sarcoglycan complex in limb-girdle muscular dystrophy. *Curr. Opin. Neurol.* **11,** 443–452

10. Crosbie, R. H., Lim, L. E., Moore, S. A., Hirano, M., Hays, A. P., Maybaum, S. W., Collin, H., Dovico, S. A., Stolle, C. A., Fardeau, M., et al. (2000) Molecular and genetic characterization of sarcospan: insights into sarcoglycan-sarcospan interactions. *Hum. Mol. Genet.* **9,** 2019–2027

11. Bansal, D., Miyake, K., Vogel, S. S., Groh, S., Chen, C. C., Williamson, R., McNeil, P. L., and Campbell, K. P. (2003) Defective membrane repair in dysferlin-deficient muscular dystrophy. *Nature* **423,** 168–172

12. Bashir, R., Britton, S., Strachan, T., Keers, S., Vafiadaki, E., Lako, M., Richard, I., Marchand, S., Bourg, N., Argov, Z., et al. (1998) A gene related to Caenorhabditis elegans spermatogenesis factor fer-1 is mutated in limb-girdle muscular dystrophy type 2B. *Nat. Genet.* **20,** 37–42

13. Lennon, N. J., Kho, A., Bacskai, B. J., Perlmutter, S. L., Hyman, B. T., and Brown, R. H., Jr. (2003) Dysferlin interacts with annexins A1 and A2 and mediates sarcolemmal wound-healing. *J. Biol. Chem.* **278,** 50,466–50,473

14. Emery, A. E. (2002) The muscular dystrophies. *Lancet* **359,** 687–695

15. Bushby, K. M. (1999) Making sense of the limb-girdle muscular dystrophies. *Brain* **122,** 1403–1420

16. Chen, Y. W., Zhao, P., Borup, R., and Hoffman, E. P. (2000) Expression profiling in the muscular dystrophies: identification of novel aspects of molecular pathophysiology. *J. Cell Biol.* **151,** 1321–1336

17. Bakay, M., Zhao, P., Chen, J., and Hoffman, E. P. (2002) A web-accessible complete transcriptome of normal human and DMD muscle. *Neuromuscul. Disord.* **12,** S125–S141

18. Haslett, J. N., Sanoudou, D., Kho, A. T., Bennett, R. R., Greenberg, S. A., Kohane, I. S., Beggs, A. H., and Kunkel, L. M. (2002) Gene expression comparison of biopsies from Duchenne muscular dystrophy (DMD) and normal skeletal muscle. *Proc. Natl. Acad. Sci. USA* **99,** 15,000–15,005

19. Noguchi, S., Tsukahara, T., Fujita, M., Kurokawa, R., Tachikawa, M., Toda, T., Tsujimoto, A., Arahata, K., and Nishino, I. (2003) cDNA microarray analysis of individual Duchenne muscular dystrophy patients. *Hum. Mol. Genet.* **12,** 595–600

20. Campanaro, S., Romualdi, C., Fanin, M., Celegato, B., Pacchioni, B., Trevisan, S., Laveder, P., De Pitta, C., Pegoraro, E., Hayashi, Y. K., et al. (2002) Gene expression profiling in dysferlinopathies using a dedicated muscle microarray. *Hum. Mol. Genet.* **11,** 3283–3298

21. Tsukahara, T., Tsujino, S., and Arahata, K. (2002) CDNA microarray analysis of gene expression in fibroblasts of patients with X-linked Emery-Dreifuss muscular dystrophy. *Muscle Nerve* **25,** 898–901

22. Winokur, S. T., Chen, Y. W., Masny, P. S., Martin, J. H., Ehmsen, J. T., Tapscott, S. J., Van Der Maarel, S. M., Hayashi, Y., and Flanigan, K. M. (2003) Expression profiling of FSHD muscle supports a defect in specific stages of myogenic differentiation. *Hum. Mol. Genet.* **12,** 2895–2907

23. Sanoudou, D., Frieden, L. A., Haslett, J. N., Kho, A. T., Greenberg, S. A., Kohane, I. S., Kunkel, L. M., and Beggs, A. H. (2004) Molecular classification of nemaline myopathies: "nontyping" specimens exhibit unique patterns of gene expression. *Neurobiol. Dis.* **15,** 590–600

24. Durbeej, M., and Campbell, K. P. (2002) Muscular dystrophies involving the dystrophin-glycoprotein complex: an overview of current mouse models. *Curr. Opin. Genet. Dev.* **12,** 349–361

25. Turk, R., Sterrenburg, E., De Meijer, E., Van Ommen, G.-J. B., Den Dunnen, J. T., and 't Hoen, P. A. (2005) Muscle regeneration in dystrophin-deficient mdx mice studied by gene expression profiling. *BMC Genomics* **6,** 98

26. Bittner, R. E., Anderson, L. V., Burkhardt, E., Bashir, R., Vafiadaki, E., Ivanova, S., Raffelsberger, T., Maerk, I., Hoger, H., Jung, M., et al. (1999) Dysferlin deletion in SJL mice (SJL-Dysf) defines a natural model for limb girdle muscular dystrophy 2B. *Nat. Genet.* **23,** 141–142

27. Cox, G. A., Phelps, S. F., Chapman, V. M., and Chamberlain, J. S. (1993) New mdx mutation disrupts expression of muscle and nonmuscle isoforms of dystrophin. *Nat. Genet.* **4,** 87–94

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

29. Araishi, K., Sasaoka, T., Imamura, M., Noguchi, S., Hama, H., Wakabayashi, E., Yoshida, M., Hori, T., and Ozawa, E. (1999) Loss of the sarcoglycan complex and sarcospan leads to muscular dystrophy in beta-sarcoglycan-deficient mice. *Hum. Mol. Genet.* **8,** 1589–1598

30. Durbeej, M., Cohn, R. D., Hrstka, R. F., Moore, S. A., Allamand, V., Davidson, B. L., Williamson, R. A., and Campbell, K. P. (2000) Disruption of the beta-sarcoglycan gene reveals pathogenetic complexity of limb-girdle muscular dystrophy type 2E. *Mol. Cell* **5,** 141–151

31. Sasaoka, T., Imamura, M., Araishi, K., Noguchi, S., Mizuno, Y., Takagoshi, N., Hama, H., Wakabayashi-Takai, E., Yoshimoto-Matsuda, Y., Nonaka, I., et al. (2003) Pathological analysis of muscle hypertrophy and degeneration in muscular dystrophy in gamma-sarcoglycan-deficient mice. *Neuromuscul. Disord.* **13,** 193–206

32. Coral-Vazquez, R., Cohn, R. D., Moore, S. A., Hill, J. A., Weiss, R. M., Davisson, R. L., Straub, V., Barresi, R., Bansal, D., Hrstka, R. F., et al. (1999) Disruption of the sarcoglycan-

sarcospan complex in vascular smooth muscle: a novel mechanism for cardiomyopathy and muscular dystrophy. *Cell* **98,** 465–474

33. Lebakken, C. S., Venzke, D. P., Hrstka, R. F., Consolino, C. M., Faulkner, J. A., Williamson, R. A., and Campbell, K. P. (2000) Sarcospan-deficient mice maintain normal muscle function. *Mol. Cell. Biol.* **20,** 1669–1677

34. Turk, R., 't Hoen, P. A., Sterrenburg, E., de Menezes, R. X., de Meijer, E. J., Boer, J. M., van Ommen, G. J., and den Dunnen, J. T. (2004) Gene expression variation between mouse inbred strains. *BMC Genomics* **5,** 57

35. 't Hoen, P. A., de Kort, F., van Ommen, G. J., and den Dunnen, J. T. (2003) Fluorescent labelling of cRNA for microarray applications. *Nucleic Acids Res.* **31,** e20

36. 't Hoen, P. A., Turk, R., Boer, J. M., Sterrenburg, E., de Menezes, R. X., van Ommen, G. J., and den Dunnen, J. T. (2004) Intensity-based analysis of two-colour microarrays enables efficient and flexible hybridization designs. *Nucleic Acids Res.* **32,** e41

37. Huber, W., Von Heydebreck, A., Sultmann, H., Poustka, A., and Vingron, M. (2002) Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* **18,** S96–S104

38. R: Development core team (2004) *R: A language and environment for statistical computing.* Vienna, Austria

39. Benjamini, Y., and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful appraoch to multiple testing. *J. Roy. Stat. Soc. B.* **57,** 289–300

40. Diehn, M., Sherlock, G., Binkley, G., Jin, H., Matese, J. C., Hernandez-Boussard, T., Rees, C. A., Cherry, J. M., Botstein, D., Brown, P. O., et al. (2003) SOURCE: a unified genomic resource of functional annotations, ontologies, and gene expression data. *Nucleic Acids Res.* **31,** 219–223

41. Zhang, B., Schmoyer, D., Kirov, S., and Snoddy, J. (2004) GOTree Machine (GOTM): a web-based platform for interpreting sets of interesting genes using Gene Ontology hierarchies. *BMC Bioinformatics* **5,** 16

42. Nagy, E., and Maquat, L. E. (1998) A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. *Trends Biochem. Sci.* **23,** 198–199

43. Inoue, K., Khajavi, M., Ohyama, T., Hirabayashi, S., Wilson, J., Reggin, J. D., Mancias, P., Butler, I. J., Wilkinson, M. F., Wegner, M., et al. (2004) Molecular mechanism for distinct neurological phenotypes conveyed by allelic truncating mutations. *Nat. Genet.* **36,** 361–369

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

45. Noguchi, S., Wakabayashi-Takai, E., Sasaoka, T., and Ozawa, E. (2001) Analysis of the spatial, temporal and tissue-specific transcription of gamma-sarcoglycan gene using a transgenic mouse. *FEBS Lett.* **495,** 77–81

46. Carpenter, S., and Karpati, G. (1979) Duchenne muscular dystrophy: plasma membrane loss initiates muscle cell necrosis unless it is repaired. *Brain* **102,** 147–161

47. Bansal, D., and Campbell, K. P. (2004) Dysferlin and the plasma membrane repair in muscular dystrophy. *Trends Cell Biol.* **14,** 206–213

48. Abbas, A. R., Baldwin, D., Ma, Y., Ouyang, W., Gurney, A., Martin, F., Fong, S., van Lookeren, C. M., Godowski, P., Williams, P. M., et al. (2005) Immune response in silico (IRIS): immune-specific genes identified from a compendium of microarray expression data. *Genes Immun.* **6,** 319–331

49. Porter, J. D., Guo, W., Merriam, A. P., Khanna, S., Cheng, G., Zhou, X., Andrade, F. H., Richmonds, C., and Kaminski, H. J. (2003) Persistent over-expression of specific CC class chemokines correlates with macrophage and T-cell recruitment in mdx skeletal muscle. *Neuromuscul. Disord.* **13,** 223–235

50. Porter, J. D., Merriam, A. P., Leahy, P., Gong, B., and Khanna, S. (2003) Dissection of temporal gene expression signatures of affected and spared muscle groups in dystrophin-deficient (mdx) mice. *Hum. Mol. Genet.* **12,** 1813–1821

51. Porter, J. D., Khanna, S., Kaminski, H. J., Rao, J. S., Merriam, A. P., Richmonds, C. R., Leahy, P., Li, J., Guo, W., and Andrade, F. H. (2002) A chronic inflammatory response dominates the skeletal muscle molecular signature in dystrophin-deficient mdx mice. *Hum. Mol. Genet.* **11,** 263–272

52. Rouger, K., Le Cunff, M., Steenman, M., Potier, M. C., Gibelin, N., Dechesne, C. A., and Leger, J. J. (2002) Global/temporal gene expression in diaphragm and hindlimb muscles of dystrophin-deficient (mdx) mice. *Am. J. Physiol. Cell Physiol.* **283,** C773–C784

53. Haslett, J. N., and Kunkel, L. M. (2002) Microarray analysis of normal and dystrophic skeletal muscle. *Int. J. Dev. Neurosci.* **20,** 359–365

54. Kuznetsov, A. V., Winkler, K., Wiedemann, F. R., von Bossanyi, P., Dietzmann, K., and Kunz, W. S. (1998) Impaired mitochondrial oxidative phosphorylation in skeletal muscle of the dystrophin-deficient mdx mouse. *Mol. Cell. Biochem.* **183,** 87–96

55. Sharma, U., Atri, S., Sharma, M. C., Sarkar, C., and Jagannathan, N. R. (2003) Skeletal muscle metabolism in Duchenne muscular dystrophy (DMD): an in-vitro proton NMR spectroscopy study. *Magn. Reson. Imaging* **21,** 145–153

56. Arber, S., Halder, G., and Caroni, P. (1994) Muscle LIM protein, a novel essential regulator of myogenesis, promotes myogenic differentiation. *Cell* **79,** 221–231

57. Coral-Vazquez, R., Cohn, R. D., Moore, S. A., Hill, J. A., Weiss, R. M., Davisson, R. L., Straub, V., Barresi, R., Bansal, D., Hrstka, R. F., et al. (1999) Disruption of the sarcoglycan-sarcospan complex in vascular smooth muscle: a novel mechanism for cardiomyopathy and muscular dystrophy. *Cell* **98,** 465–474

58. Cohn, R. D., Durbeej, M., Moore, S. A., Coral-Vazquez, R., Prouty, S., and Campbell, K. P. (2001) Prevention of cardiomyopathy in mouse models lacking the smooth muscle sarcoglycan-sarcospan complex. *J. Clin. Invest.* **107,** R1–R7

59. McArdle, A., Edwards, R. H., and Jackson, M. J. (1995) How does dystrophin deficiency lead to muscle degeneration? – evidence from the mdx mouse. *Neuromuscul. Disord.* **5,** 445–456

60. Brussee, V., Tardif, F., and Tremblay, J. P. (1997) Muscle fibers of mdx mice are more vulnerable to exercise than those of normal mice. *Neuromuscul. Disord.* **7,** 487–492

61. Schafer, R., Zweyer, M., Knauf, U., Mundegar, R. R., and Wernig, A. (2005) The ontogeny of soleus muscles in mdx and wild-type mice. *Neuromuscul. Disord.* **15,** 57–64

62. Emery, A., and Muntoni, F. (2003) Clinical features. In: *Duchenne muscular dystrophy* pp. 26–45, Oxford University Press, Oxford.

Received July 15, 2005; accepted September 22, 2005.

Table 1

Phenotype of studied mouse models

				Histopathological parameters at 8 wk					
Disease	**Model**	**Affected gene**	**Age of onset**	**Skeletal dystrophy**	**Inflammation**	**Central nuclei**	**DGC loss**	**SGC loss**	**Reference**
DMD	mdx	Dystrophin	2-3 wk	Severe	+	+	y	y	(44)
DMD	mdx^{3cv}	Dystrophin	2-3 wk	Severe	+	+	y	y	(27)
LGMD-2D	Sgca null	α-Sarcoglycan	1 wk	Severe	+	+	n	highly reduced	(28)
LGMD-2E	Sgcb null	β-Sarcoglycan	at least 4 wk	Severe	+	+	n	y	(29,30)
LGMD-2C	Sgcg null	gamma-Sarcoglycan	2 wk	Severe	+	+	n	highly reduced	(31)
LGMG-2F	Sgcd null	delta-Sarcoglycan	2 wk	Severe	+	+	n	highly reduced	(32)
LGMD-2B	Dysf null	Dysferlin	8 wk	Mild/moderate	-	+/–	n	n	(11)
LGMD-2B	Sjl^{Dysf}	Dysferlin	3 wk	Mild/moderate	n/a	+/–	n/a	n/a	(26)
not known	Sspn null	Sarcospan	None	None	-	-	n	n	(33)

Table 2

Differentially expressed genes in severely vs. mildly and nonaffected models

Up-regulated genes

Accession	UGCluster	Description	Symbol	*P* Value*	Fold-change I vs. II
X16834	Mm.248615	Lectin, galactose binding, soluble 3	Lgals3	4.03E-14	10.29
NM_009263	Mm.288474	Secreted phosphoprotein 1	Spp1	2.17E-07	9.28
AJ251685	Mm.302602	Glycoprotein (transmembrane) nmb	Gpnmb	6.08E-12	7.63
L20315	Mm.3999	Macrophage expressed gene 1	Mpeg1	3.26E-14	6.69
NM_008590	Mm.1089	Mesoderm specific transcript	Mest	4.53E-20	6.46
NM_020008	Mm.239516	C-type lectin domain family 7, member a	Clecsf12	4.18E-12	6.40
NM_010745	Mm.2639	Lymphocyte antigen 86	Ly86	6.89E-14	6.33
NM_021443	Mm.42029	Chemokine (C-C motif) ligand 8	Ccl8	2.49E-12	5.88
NM_011593	Mm.8245	Tissue inhibitor of metalloproteinase 1	Timp1	6.97E-11	5.61
NM_011333	Mm.290320	Chemokine (C-C motif) ligand 2	Ccl2	1.06E-08	5.28
U66888	Mm.2254	EGF-like module containing, mucin-like, hormone receptor-like sequence 1	Emr1	2.80E-10	5.18
M55561	Mm.24130	CD52 antigen	Cd52	4.12E-10	4.82
X58196	Mm.14802	H19 fetal liver mRNA	H19	5.79E-16	4.68
NM_009853	Mm.15819	CD68 antigen	Cd68	5.06E-08	4.60
NM_009779	Mm.2408	Complement component 3a receptor 1	C3ar1	2.80E-10	4.58
NM_013590	Mm.177539	P lysozyme structural	Lzp-s	4.81E-09	4.22
NM_007739	Mm.371554	Procollagen, type VIII, alpha 1	Col8a1	3.76E-09	4.21
NM_008871	Mm.250422	Serine (or cysteine) proteinase inhibitor, clade E, member 1	Serpine1	6.30E-11	4.18
NM_008404	Mm.1137	Integrin beta 2	Itgb2	1.82E-08	4.18
NM_007572	Mm.370	Complement component 1, q subcomponent, alpha polypeptide	C1qa	2.55E-15	4.15
NM_013532	Mm.34408	Leukocyte immunoglobulin-like receptor, subfamily B, member 4	Lilrb4	1.48E-11	4.05
NM_011662	Mm.46301	TYRO protein tyrosine kinase binding protein	Tyrobp	2.39E-08	4.02
AF263458	Mm.34609	Placenta-specific 8	Plac8	2.54E-08	3.99
AF246265	Mm.280158	C1q and tumor necrosis factor related protein 3	C1qtnf3	9.81E-04	3.84
M33863	Mm.14301	2′-5′ oligoadenylate synthetase 1G	Oas1a	5.31E-14	3.81
NM_011619	Mm.247470	Troponin T2, cardiac	Tnnt2	4.44E-15	3.73
L04694	Mm.341574	Chemokine (C-C motif) ligand 7	Ccl7	4.53E-08	3.71
NM_011175	Mm.17185	Legumain	Lgmn	2.62E-12	3.54
NM_010686	Mm.271868	Lysosomal-associated protein transmembrane 5	Laptm5	5.30E-12	3.51
NM_009814	Mm.15343	Calsequestrin 2	Casq2	8.21E-10	3.51

AF000427	Data not found	B144 mRNA, m17r splice variant		2.48E-11	3.51
NM_021467	Mm.44379	Troponin l, skeletal, slow 1	Tnni1	2.30E-06	3.38
NM_011210	Mm.130953	Protein tyrosine phosphatase, receptor type, C	Ptprc	8.19E-08	3.36
NM_008761	Mm.1870	FXYD domain-containing ion transport regulator 5	Fxyd5	3.31E-05	3.28
NM_011208	Data not found	Protein tyrosine phosphatase, nonreceptor type substrate 1	Ptpns1	6.28E-10	3.27
NM_015783	Mm.358664	Interferon, alpha-inducible protein	G1p2	2.18E-07	3.13
NM_011999	Mm.47384	C-type lectin domain family 4, member a2	Clecsf6	6.69E-14	3.10
D00208	Mm.3925	S100 calcium binding protein A4	S100a4	3.83E-11	3.08
NM_010496	Mm.34871	Inhibitor of DNA binding 2	Idb2	1.30E-09	3.05
D90156	Mm.16528	Myogenin	Myog	3.75E-10	3.03
NM_010730	Mm.248360	RIKEN cDNA C430014K04 gene	Anxa1	7.92E-07	3.03
AJ131395	Mm.297859	Procollagen, type XIV, alpha 1	Col14a1	1.86E-11	2.99
NM_009856	Mm.57175	CD83 antigen	Cd83	1.81E-07	2.97
NM_007599	Mm.18626	Capping protein (actin filament), gelsolin-like	Capg	5.57E-10	2.95
NM_009696	Mm.305152	Apolipoprotein E	Apoe	2.01E-09	2.91
NM_010253	Mm.4655	Galanin	Gal	7.07E-05	2.90
NM_008328	Mm.261270	Interferon activated gene 203	Ifi203	1.34E-09	2.89
NM_007542	Mm.2608	Biglycan	Bgn	1.74E-12	2.86
NM_009970	Data not found	colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage)	Csf2ra	3.67E-07	2.85
NM_007574	Mm.3453	Complement component 1, q subcomponent, gamma polypeptide	C1qg	6.48E-08	2.85
NM_011607	Mm.980	Tenascin C	Tnc	4.28E-05	2.81
NM_009403	Mm.4664	Tumor necrosis factor (ligand) superfamily, member 8	Tnfsf8	9.31E-08	2.80
NM_009448	Mm.88212	Tubulin, alpha 6	Tuba6	9.17E-09	2.78
NM_011701	Mm.268000	Vimentin	Vim	4.81E-09	2.75
NM_011961	Mm.79983	Procollagen lysine, 2-oxoglutarate 5-dioxygenase 2	Plod2	6.89E-14	2.73
NM_009139	Mm.137	Chemokine (C-C motif) ligand 6	Ccl6	3.14E-07	2.69
NM_007793	Mm.6095	Cystatin B	Cstb	4.43E-07	2.61
NM_011671	Mm.171378	Uncoupling protein 2 (mitochondrial, proton carrier)	Ucp2	7.33E-08	2.61
AF149291	Mm.271711	Transgelin 2	Tagln2	3.02E-08	2.58
NM_007651	Mm.316861	CD53 antigen	Cd53	1.44E-07	2.58
M22432	Mm.138471	Similar to elongation factor 1-alpha 1 (EF-1-alpha-1)	Eef1a1	7.41E-07	2.57
NM_011035	Mm.260227	P21 (CDKN1A)-activated kinase 1	Pak1	4.62E-02	2.57
M27347	Mm.2131	Elastase 1, pancreatic	Ela1	3.38E-07	2.54

NM_009468	Mm.8180	Dihydropyrimidinase-like 3	Dpysl3	8.62E-12	2.51
X00496	Mm.276499	Ia-associated invariant chain	Ii	3.69E-10	2.44
NM_009192	Mm.7601	Src-like adaptor	Sla	2.58E-05	2.44
NM_011116	Mm.6483	Phospholipase D3	Pld3	2.75E-06	2.42
M18933	Mm.249555	Procollagen, type III, alpha 1	Col3a1	1.42E-04	2.41
NM_007802	Mm.272085	Cathepsin K	Ctsk	2.16E-07	2.36
NM_011338	Mm.2271	Chemokine (C-C motif) ligand 9	Ccl9	4.56E-04	2.36
NM_011727	Mm.347026	X-linked lymphocyte-regulated 3b	Xlr3b	6.57E-04	2.36
NM_018820	Mm.153684	SERTA domain containing 1	Sertad1	3.31E-05	2.36
NM_011206	Mm.361	Protein tyrosine phosphatase, non-receptor type 18	Ptpn18	2.33E-06	2.35
NM_020575	Mm.260635	Membrane-associated ring finger (C3HC4) 7	Axot	3.69E-04	2.33
NM_009151	Mm.332590	Selectin, platelet (p-selectin) ligand	Selpl	1.12E-05	2.33
X04231	Data not found	4.5S small RNA associated with poly-(a)-containing RNAs		2.69E-05	2.32
NM_016904	Mm.3049	CDC28 protein kinase 1b	Cks1b	2.96E-04	2.31
AB031386	Mm.28385	RIKEN cDNA 1810009M01 gene	1810009M01Rik	2.25E-10	2.29
NM_019467	Mm.10747	Allograft inflammatory factor 1	Aif1	9.04E-08	2.28
NM_011653	Mm.371591	Tubulin, alpha 1	Tuba1	4.89E-07	2.28
AF203898	Data not found	Nebulin		4.15E-05	2.27
AF188290	Mm.220982	Dysferlin	Dysf	7.93E-07	2.26
NM_010634	Mm.741	Fatty acid binding protein 5, epidermal	Fabp5	1.89E-05	2.26
NM_019389	Data not found			6.47E-07	2.25
NM_007737	Mm.10299	Procollagen, type V, alpha 2	Col5a2	1.31E-12	2.25
NM_008597	Mm.243085	Matrix gamma-carboxyglutamate (gla) protein	Mglap	3.10E-11	2.25
NM_009515	Mm.4735	Wiskott-Aldrich syndrome homolog (human)	Was	1.06E-08	2.24
NM_009037	Mm.4876	Reticulocalbin 1	Rcn1	3.66E-09	2.24
NM_010740	Mm.681	Complement component 1, q subcomponent, receptor 1	C1qr1	6.82E-09	2.24
NM_009364	Mm.25612	Tissue factor pathway inhibitor 2	Tfpi2	2.09E-11	2.22
NM_008409	Mm.193	Integral membrane protein 2A	Itm2a	3.29E-10	2.22
AJ006837	Mm.84664	RNA, U17d small nucleolar	Rnu17d	2.99E-06	2.21
NM_008538	Mm.30059	Myristoylated alanine rich protein kinase C substrate	Marcks	3.22E-07	2.18
NM_008035	Mm.2724	Folate receptor 2 (fetal)	Folr2	1.48E-05	2.17
NM_008047	Mm.182434	Follistatin-like 1	Fstl1	1.47E-11	2.17
NM_010566	Mm.15105	Inositol polyphosphate-5-phosphatase D	Inpp5d	1.44E-07	2.15

NM_011618	Mm.358643	Troponin T1, skeletal, slow	Tnnt1	6.73E-03	2.13
NM_008516	Mm.158868	Leucine rich repeat protein 1, neuronal	Lrrn1	5.68E-05	2.13
L10905	Data not found			1.32E-04	2.13
AF068182	Mm.9749	B-cell linker	Blnk	7.74E-04	2.12
AF064749	Mm.7562	Procollagen, type VI, alpha 3	Col6a3	1.67E-08	2.12
NM_008660	Data not found	Myosin IF	Myo1f	1.38E-08	2.11
NM_019755	Mm.18565	Proteolipid protein 2	Plp2	1.01E-07	2.11
NM_013762	Mm.290899	Ribosomal protein L3	Rpl3	2.22E-05	2.11
NM_021394	Mm.116687	Z-DNA binding protein 1	Zbp1	7.99E-04	2.10
X17501	Mm.1738	CD48 antigen	Cd48	9.67E-06	2.09
NM_011313	Mm.100144	S100 calcium binding protein A6 (calcyclin)	S100a6	7.79E-07	2.08
NM_021372	Mm.339676	SERTA domain containing 2	Sertad2	7.39E-10	2.07
NM_010276	Mm.247486	GTP binding protein (gene overexpressed in skeletal muscle)	Gem	4.64E-08	2.06
NM_011065	Mm.7373	Period homolog 1 (Drosophila)	Per1	7.58E-05	2.06
NM_018773	Mm.221479	Src family associated phosphoprotein 2	Scap2	5.48E-08	2.05
NM_013672	Mm.4618	Trans-acting transcription factor 1	Sp1	4.90E-05	2.05
NM_021474	Mm.276367	Epidermal growth factor-containing fibulin-like extracellular matrix protein 2	Efemp2	3.34E-08	2.04
NM_016778	Mm.3295	Bcl-2-related ovarian killer protein	Bok	5.94E-06	2.03
NM_015766	Mm.256798	Epstein-Barr virus induced gene 3	Ebi3	7.77E-08	2.03
NM_008638	Mm.443	Methylenetetrahydrofolate dehydrogenase (NAD+ dependent)	Mthfd2	9.42E-06	2.03
NM_015807	Mm.307174	5′,3′-nucleotidase, cytosolic	Nt5c	2.97E-05	2.03
NM_008972	Mm.19187	Prothymosin alpha	Ptma	3.66E-04	2.02
NM_011655	Mm.273538	Tubulin, beta 5	Tubb5	2.10E-03	2.02
NM_010422	Mm.27816	G elongation factor, mitochondrial 2	Hexb	1.55E-07	2.02
NM_009750	Mm.90787	Nerve growth factor receptor (TNFRSF16) associated protein 1	Ngfrap1	3.51E-05	2.02
X62154	Mm.4502	Minichromosome maintenance deficient 3 (S. cerevisiae)	Mcm3	1.79E-04	2.01
AF188504	Mm.35600	Chemokine-like factor super family 7	Cklfsf7	5.66E-05	2.01
U79144	Mm.250492	Lysyl oxidase-like 1	Loxl1	2.20E-05	2.01
NM_007570	Mm.239605	B-cell translocation gene 2, anti-proliferative	Btg2	2.28E-06	2.01
NM_017380	Mm.38450	Septin 9	Sep9	1.07E-09	2.00
NM_019727	Mm.271891	Sorting nexin 1	Snx1	6.53E-03	2.00
AF309649	Mm.30241	Interferon gamma inducible protein 30	Ifi30	1.99E-05	2.00

Down-regulated genes

Accession	UGCluster	Description	Symbol	*P* Value*	Fold-change I vs. II
NM_009463	Mm.4177	Uncoupling protein 1 (mitochondrial, proton carrier)	Ucp1	1.49E-11	−9.53
NM_007702	Mm.449	Cell death-inducing DNA fragmentation factor, alpha subunit-like effector A	Cidea	1.09E-14	−9.28
AF155353	Mm.27159	Ankyrin repeat and SOCS box-containing protein 2	Asb2	1.09E-14	−3.38
NM_011281	Mm.4372	RAR-related orphan receptor gamma	Rorc	1.93E-06	−3.24
NM_020589	Mm.358722	Zinc finger protein 467	Zfp467	5.31E-14	−3.03
NM_008615	Mm.148155	Malic enzyme, supernatant	Mod1	7.91E-09	−2.84
AF060220	Mm.218857	Erythroid associated factor	Eraf	9.06E-03	−2.77
NM_021391	Mm.143788	Protein phosphatase 1, regulatory (inhibitor) subunit 1A	Ppp1r1a	3.59E-11	−2.72
NM_015763	Mm.153625	Lipin 1	Lpin1	3.85E-10	−2.68
NM_008975	Mm.153891	Protein tyrosine phosphatase 4a3	Ptp4a3	2.00E-05	−2.68
AF032131	Mm.23296	E2F transcription factor 6	E2f6	3.80E-05	−2.59
NM_008768	Mm.4777	Orosomucoid 1	Orm1	4.50E-04	−2.54
AJ278735	Mm.38330	Chaperone, ABC1 activity of bc1 complex like (S. pombe)	Cabc1	3.99E-05	−2.54
NM_013849	Mm.317764	Dual specificity phosphatase 13	Dusp13	4.42E-08	−2.52
NM_010585	Mm.227912	Inositol 1,4,5-triphosphate receptor 1	Itpr1	1.45E-09	−2.48
NM_009204	Mm.10661	Solute carrier family 2 (facilitated glucose transporter), member 4	Slc2a4	5.30E-12	−2.47
NM_009161	Mm.18709	Sarcoglycan, alpha (dystrophin-associated glycoprotein)	Sgca	4.59E-03	−2.45
NM_009349	Mm.299	Indolethylamine *N*-methyltransferase	Inmt	2.41E-05	−2.44
NM_018819	Mm.288510	RIKEN cDNA A630084N20 gene	Brp44l	9.93E-11	−2.35
NM_007489	Mm.12177	Aryl hydrocarbon receptor nuclear translocator-like	Arntl	1.44E-07	−2.35
NM_009624	Mm.310036	Adenylate cyclase 9	Adcy9	1.18E-05	−2.34
NM_008748	Mm.39725	Dual specificity phosphatase 8	5530400B01Rik	2.80E-10	−2.33
NM_008288	Mm.28328	Hydroxysteroid 11-beta dehydrogenase 1	Hsd11b1	1.96E-08	−2.30
NM_021430	Mm.41180	RIKEN cDNA 2900002H16 gene	2900002H16Rik	1.33E-09	−2.27
AF282730	Mm.255607	Tissue inhibitor of metalloproteinase 4	Timp4	1.98E-11	−2.26
NM_009953	Mm.236081	Corticotropin releasing hormone receptor 2	Crhr2	2.98E-08	−2.21
D00926	Data not found	Transcription factor S-II-related protein		7.92E-03	−2.18
NM_013558	Mm.14287	Heat shock protein 1-like	Hspa1l	5.86E-09	−2.18
NM_019435	Mm.30084	Nuclear protein 15.6	Np15	2.90E-03	−2.18
NM_007469	Mm.182440	Apolipoprotein C-I	Apoc1	8.17E-06	−2.17
NM_008377	Mm.245210	Leucine-rich repeats and immunoglobulin-like domains 1	Lrig1	5.14E-06	−2.17
NM_016917	Mm.28756	Solute carrier family 40 (iron-regulated transporter), member 1	Slc40a1	1.92E-07	−2.16
AF288783	Mm.256926	Liver glycogen phosphorylase	Pygl	1.56E-05	−2.16
NM_007377	Mm.6826	Apoptosis-associated tyrosine kinase	Aatk	5.86E-09	−2.15
NM_013459	Mm.4407	Adipsin	Adn	1.20E-02	−2.13
AF267660	Mm.29768	Pyruvate dehydrogenase kinase, isoenzyme 2	Pdk2	8.45E-08	−2.11
NM_009861	Mm.1022	Cell division cycle 42 homolog (S. cerevisiae)	Cdc42	1.62E-06	−2.07

NM_017479	Mm.248967	MYST histone acetyltransferase monocytic leukemia 4	Myst4	2.38E-07	−2.07
NM_011697	Mm.15607	Vascular endothelial growth factor B	Vegfb	1.18E-05	−2.05
X74504	Mm.265990	DNA segment, Chr 16, human D22S680E, expressed	D16H22S680E	1.58E-09	−2.04
X98848	Mm.249131	6-Phosphofructo-2-kinase/fructose-2,6-biphosphatase 1	Pfkfb1	1.20E-07	−2.04
NM_011428	Mm.45953	Synaptosomal-associated protein 25	Snap25	1.59E-07	−2.03
NM_010191	Mm.371560	Farnesyl diphosphate farnesyl transferase 1	Fdft1	5.75E-10	−2.03
NM_016772	Mm.291776	Enoyl coenzyme A hydratase 1, peroxisomal	Ech1	1.45E-10	−2.02
X93035	Mm.38274	Chitinase 3-like 1	Chi3l1	1.46E-06	−2.01
NM_008428	Mm.1482	Potassium inwardly rectifying channel, subfamily J, member 8	Kcnj8	1.71E-09	−2.00

Table 3

Differentially expressed genes between Mdx and Sarcoglycan-deficient mice

Lower expressed in sarcoglycan-deficient models

Accession	UGCluster	Name	Symbol	*P* Value*	Fold Change SG-mdx
NM_009441	Mm.213408	Tetratricopeptide repeat domain 3	Ttc3	2.44E-02	-3.12
NM_019567	Mm.297078	Apoptotic chromatin condensation inducer 1	Acin1	8.33E-03	-3.02
NM_009484	Mm.20477	Ubiquitously transcribed tetratricopeptide repeat gene, Y chromosome	Uty	8.38E-03	-2.73
NM_007788	Mm.298893	Casein kinase II, α 1 polypeptide	Csnk2a1	1.63E-03	-2.05

Higher expressed in sarcoglycan-deficient models

Accession	UGCluster	Name	Symbol	*P* Value*	Fold Change SG-mdx
NM_010217	Mm.1810	Connective tissue growth factor	Ctgf	1.45E-03	2.14
NM_021503	Mm.141157	Myozenin 2	Myoz2	1.85E-03	2.15
NM_020033	Mm.143737	Ankyrin repeat domain 2 (stretch responsive muscle)	Ankrd2	1.84E-05	2.22
NM_013558	Mm.14287	Heat shock protein 1-like	Hspa1l	1.71E-02	2.23
NM_013808	Mm.17235	Cysteine and glycine-rich protein 3	Csrp3	4.53E-04	3.73

Fig. 1

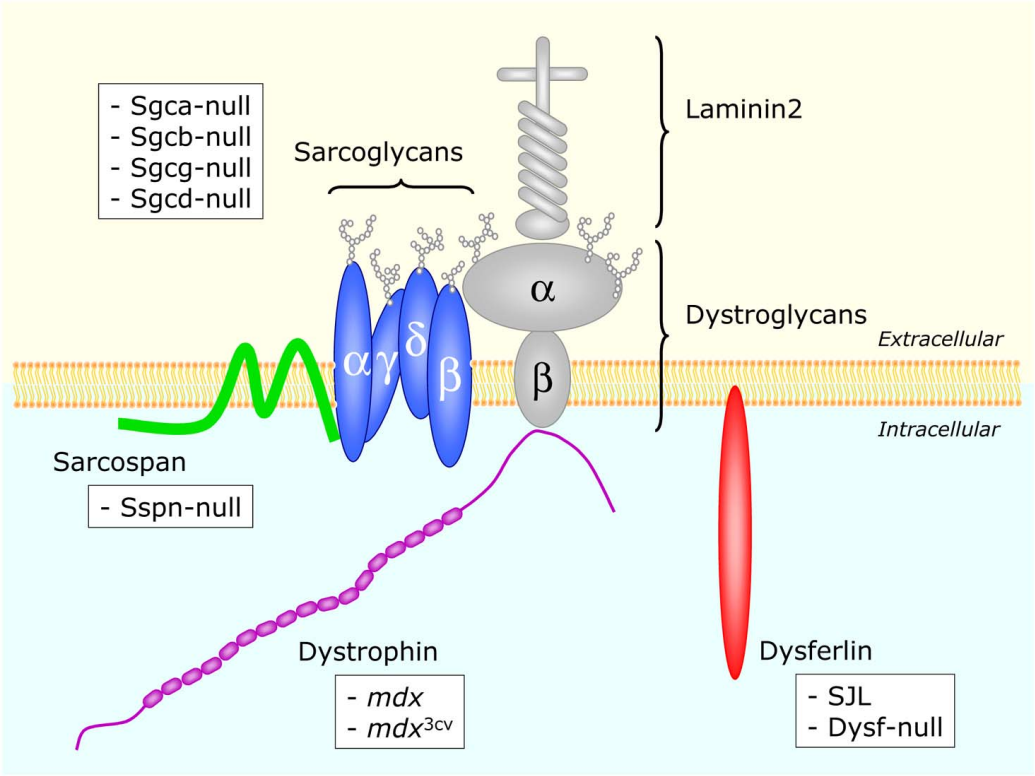


Figure 1. Mouse models for muscular dystrophy related to DGC.

Fig. 2

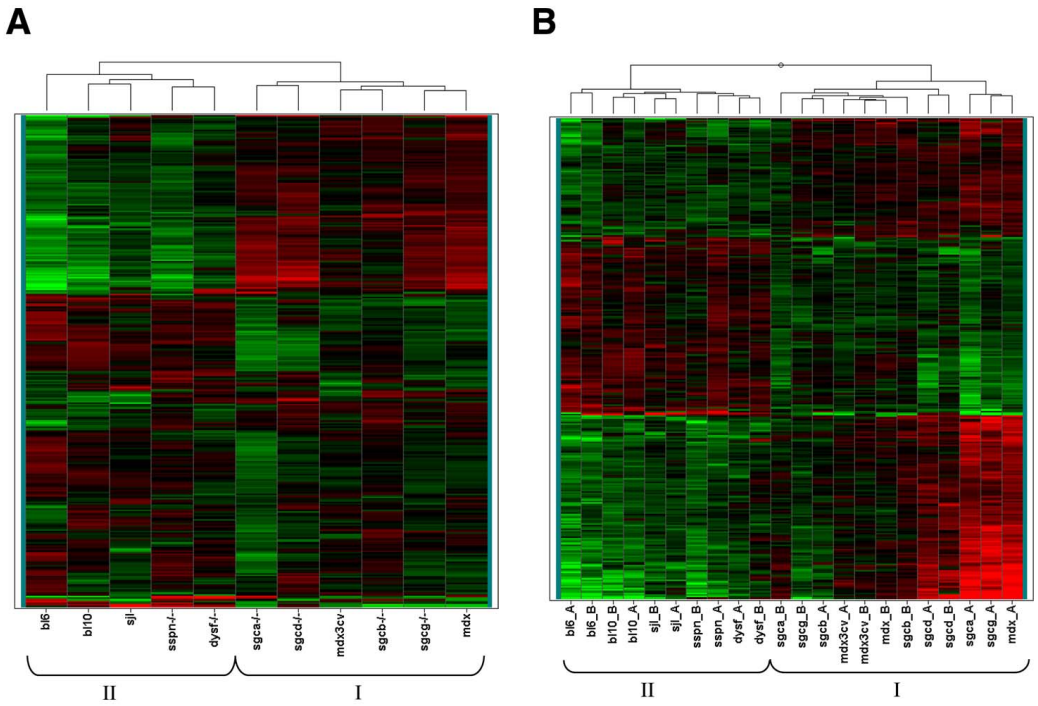


Figure 2. Unsupervised hierarchical clustering of gene expression patterns. ***A)*** 2-dimensional, unsupervised hierarchical clustering was performed on averaged expression levels of each strain for 300 genes that display most significant (BH-corrected P value < 4.5×10^{-8}) differences in expression between strains. For better visualization of up- and down-regulation, gene expression levels were scaled to an average value of 0. Euclidean distance was used as a distance measure. Based on clustering, 2 major groups can be discerned (I and II). ***B)*** Unsupervised hierarchical clustering was performed on averaged normalized gene expression levels per individual of 300 genes that display most significant (BH-corrected P value < 4.1×10^{-5}) differences in expression between severely affected and non- or mildly affected animal models. The clustering and visualization method applied was identical to that in ***A***.

Fig. 3

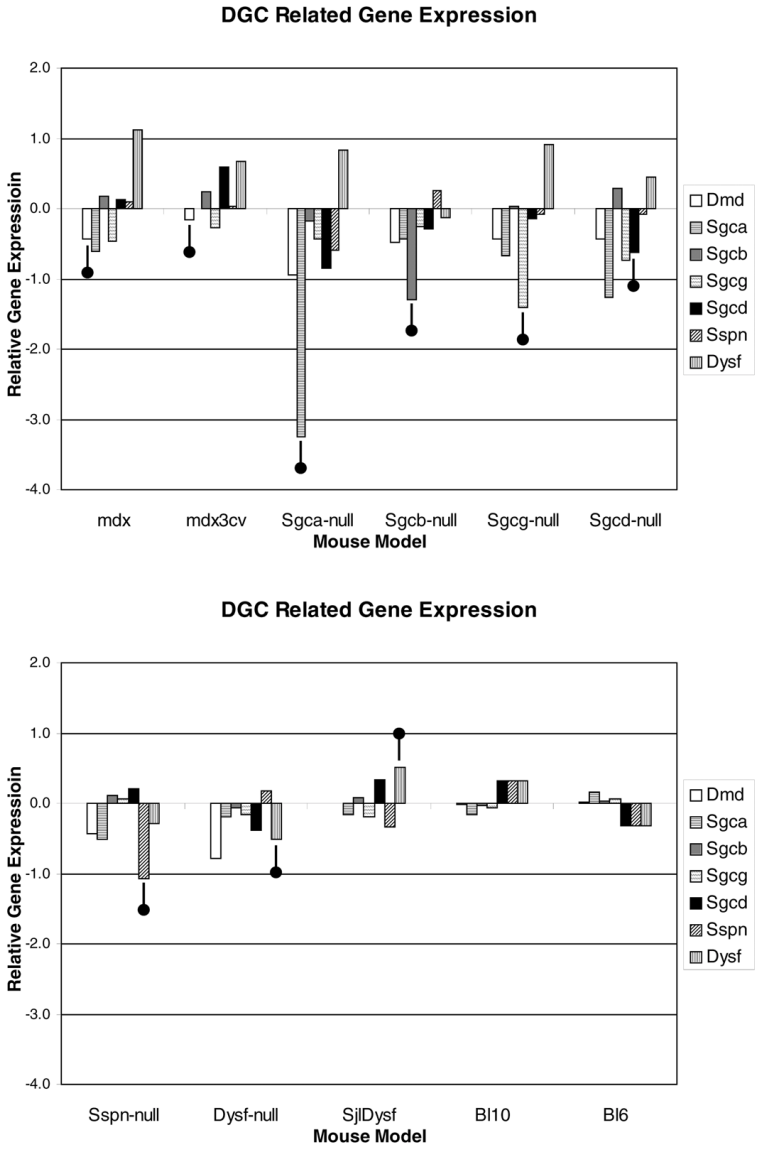


Figure 3. DGC-related genes are down-regulated in severely affected mouse models for muscular dystrophy. Differential gene expression levels between mouse models and wild-type mice were calculated by subtraction of average gene expression level of the 2 wild-type mice (Bl6 and Bl10) from gene expression levels of all models (including wild types). Differential expression levels of 7 DGC-related genes are shown (Dmd: dystrophin; Sgca: α-sarcoglycan; Sgcb: β-sarcoglycan; Sgcg: gamma-sarcoglycan; Sgcd: delta-sarcoglycan; Sspn: sarcospan; Dysf: dysferlin). Gene mutated in accompanied model is indicated with ●.

Fig. 4

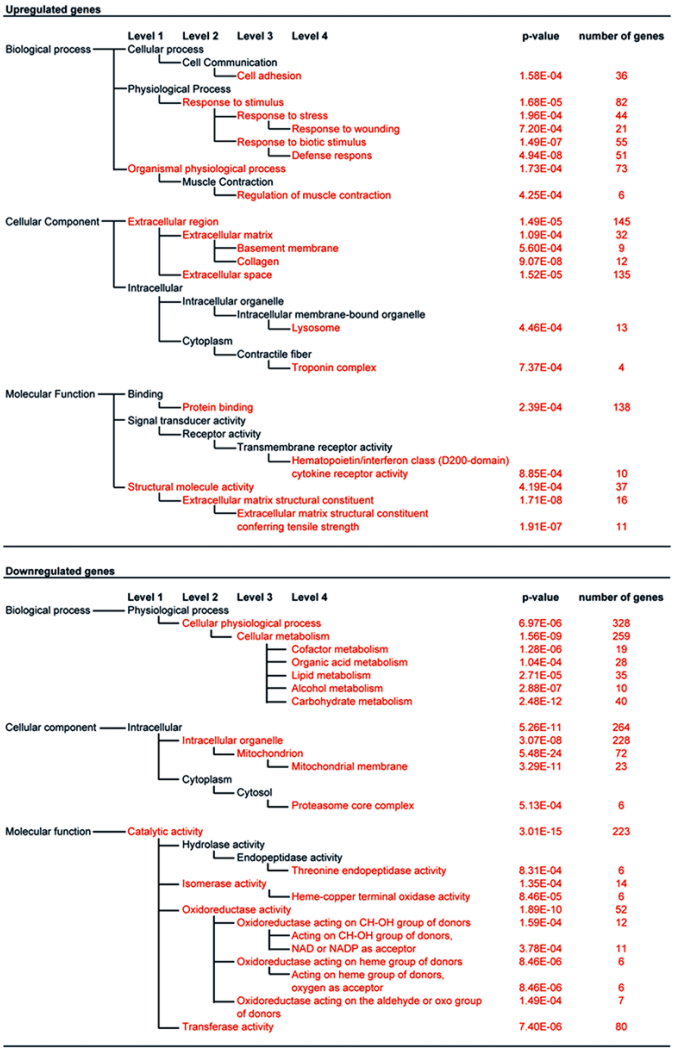
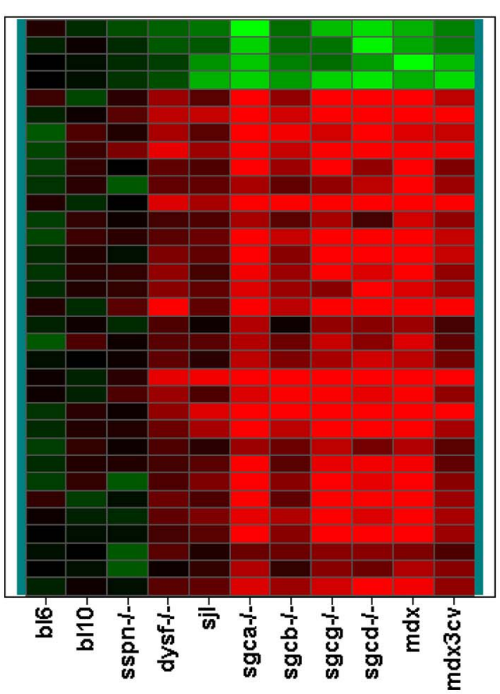


Figure 4. Functional classification of significantly up- and down-regulated genes in severely affected mouse models. Genes up-regulated in dystrophic muscle from severely affected mouse models were grouped according to biological process, cellular component, and molecular function, based on Gene Ontology classifications. Only branches of the GO-tree containing categories that were significantly overrepresented (displayed in red; $P<0.001$) in the list of up (***upper panel***)- and down-regulated (***lower panel***) genes are shown. Listed *P* values are from a hypergeometric test that compares, for each category, number of genes in the set of up- or down-regulated genes with total number of genes present on the array in that category. Number of genes refers to number of genes in the list of up- or down-regulated genes in a specific category. A specification of genes present in overrepresented categories can be found in Supplemental Table 4.

Fig. 5



Symbol	Name
Pdk2	Pyruvate dehydrogenase kinase, isoenzyme 2
Fdft1	Farnesyl diphosphate farnesyl transferase 1
Hspa1l	Heat shock protein 1-like
Timp4	Tissue inhibitor of metalloproteinase 4
Cd68	CD68 antigen
Lzp-s	P lysozyme structural
Col14a1	Procollagen, type XIV, alpha 1
Itgb2	Integrin beta 2
Tuba6	Tubulin, alpha 6
Ii	Ia-associated invariant chain
Lgals3	Lectin, galactose binding, soluble 3
Tlr2	Toll-like receptor 2
Gpnmb	Glycoprotein (transmembrane) nmb
Laptm5	Lysosomal-associated protein transmembrane 5
Cd83	CD83 antigen
Ctsk	Cathepsin K
Cd52	CD52 antigen
Arhgdib	Rho, GDP dissociation inhibitor (GDI) beta
Scap2	Src family associated phosphoprotein 2
1810009M01Rik	RIKEN cDNA 1810009M01 gene
Timp1	Tissue inhibitor of metalloproteinase 1
Plod2	Procollagen lysine, 2-oxoglutarate 5-dioxygenase 2
Mpeg1	Macrophage expressed gene 1
Lilrb4	Leukocyte immunoglobulin-like receptor, subfamily B, member 4
Myo9b	DNA segment, Chr 8, ERATO Doi 674, expressed
Cd53	CD53 antigen
Ccl7	Chemokine (C-C motif) ligand 7
Tyrobp	TYRO protein tyrosine kinase binding protein
S100a4	S100 calcium binding protein A4
Lgmn	Legumain
Pftk1	PFTAIRE protein kinase 1
Aif1	Allograft inflammatory factor 1
Bgn	Biglycan

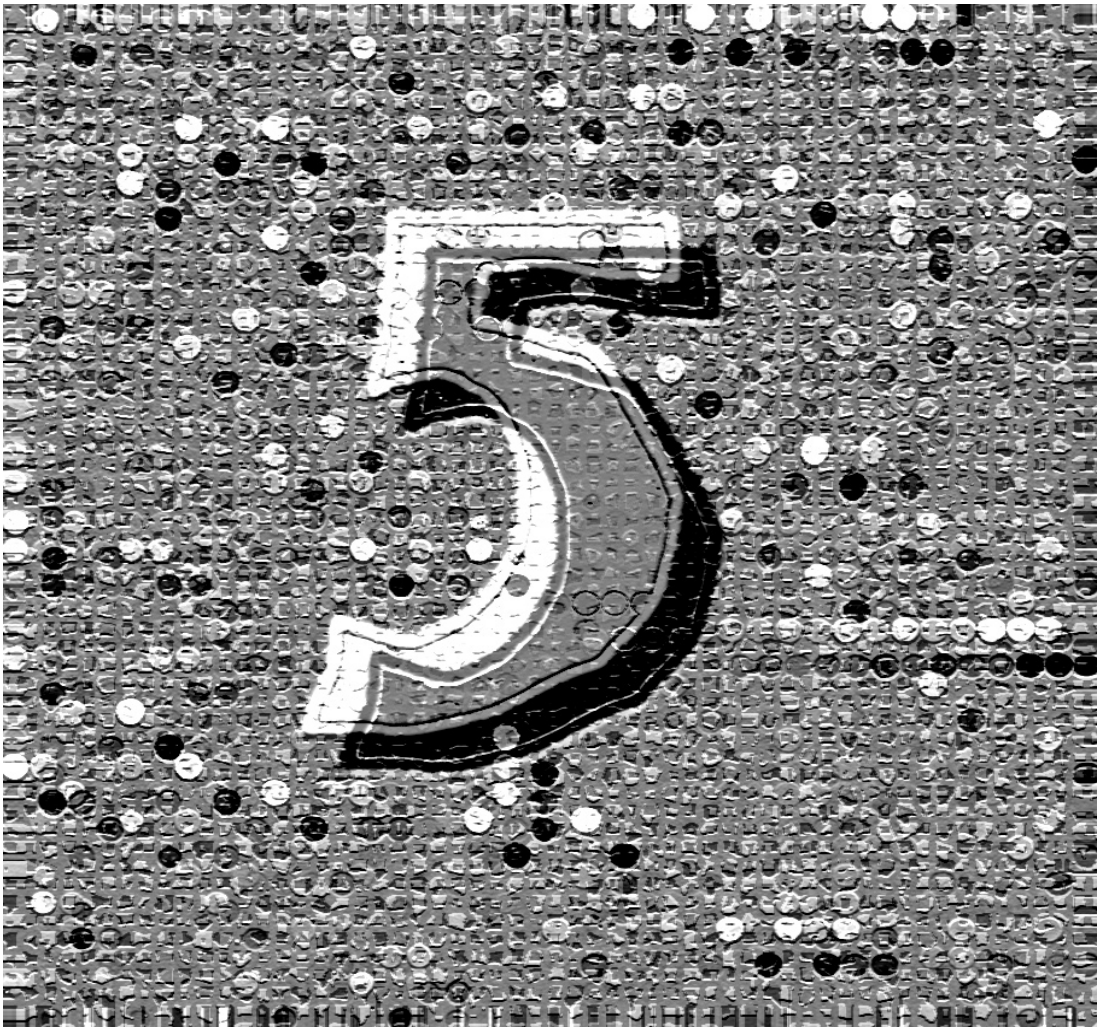
Figure 5. Biomarker genes for disease progression in muscular dystrophy heat map of averaged expression levels (relative to levels in wild-type mice) of genes that correlate with disease progression. The top 4 genes demonstrate significantly lower expression (displayed in green) in dysferlin-deficient mice compared with wild-type and sarcospan-deficient mice and even lower expression levels in the more severely affected mouse models, whereas the bottom 29 genes demonstrate significantly higher (displayed in red) expression in dysferlin-deficient mice compared with wild-type and sarcospan-deficient mice and even higher expression levels in the more severely affected mouse models.

Chapter 5

Large-scale gene expression analysis of human skeletal myoblast differentiation

Sterrenburg E, Turk R, ‘t Hoen PA, van Deutekom JC,
Boer JM, van Ommen GJ, den Dunnen JT

Neuromuscular Disorders 2004 14(8-9):507-18



PERGAMON

Neuromuscular Disorders 14 (2004) 507–518

www.elsevier.com/locate/nmd

Large-scale gene expression analysis of human skeletal myoblast differentiation

Ellen Sterrenburg, Rolf Turk, Peter A.C. 't Hoen, Judith C.T. van Deutekom, Judith M. Boer, Gert-Jan B. van Ommen, Johan T. den Dunnen*

Center for Human and Clinical Genetics, Leiden University Medical Center, Wassenaarseweg 72, 2333 AL Leiden, The Netherlands

Received 9 January 2004; received in revised form 16 March 2004; accepted 24 March 2004

Abstract

To study pathways involved in human skeletal myogenesis, we profiled gene expression in human primary myoblast cells derived from three individuals using both oligonucleotide and cDNA microarrays. Following stringent statistical testing (false-positive rate 0.4%), we identified 146 genes differentially expressed over time. Interestingly, 86 of these genes have not been reported to be involved in myogenesis in mouse cell lines. This demonstrates the additional value of human primary cell cultures in the study of muscle differentiation. Many of the identified genes play a role in muscle regeneration, indicating the close relationship of this process with muscle development. In addition, we found overlap with expression profiling studies in muscle from Duchenne muscular dystrophy patients, confirming ongoing muscle regeneration in Duchenne muscular dystrophy. Further study of these genes can bring new insights into the process of muscle differentiation, and they are candidate genes for neuromuscular disorders with an as yet unidentified cause.

Keywords: Myogenesis; Muscle differentiation; Gene expression; Microarray

1. Introduction

Skeletal muscle cell differentiation is a multistage process that has been studied extensively over the years. Early myogenesis involves the commitment of mesodermal cells to myogenic progenitors and ultimately myoblasts. In late myogenesis, myoblasts withdraw from the cell cycle, become longer and fuse to form multinucleated myotubes [1]. During muscle regeneration, the second stage of myogenesis is mimicked when satellite cells, present between the basal lamina and the connective tissue, become activated myoblasts and either fuse with existing myotubes or form new myotubes [1]. The myogenic regulatory factors (MRFs), including MyoD, Myf5, Myogenin and Mrf4, are known to be key regulators in the initiation and progression of myogenesis [2–6]. They contain a helix–loop–helix motif for heterodimerization with E-proteins. When heterodimerized, they can bind at sites known as E-boxes (CANNTG) in the promoter and enhancer regions of most skeletal muscle-specific genes [7]. Other important genes and gene families involved in myogenesis are MEF2 transcription factors, the Pax family, Sonic hedgehog and the Wnt genes [8–13]. TGF-β superfamily members and the Id family of helix–loop–helix proteins negatively regulate myogenesis [14,15].

The MRFs and other myogenesis factors have been identified by conventional techniques, often on a one-by-one basis. Current high throughput genomics approaches, such as microarray analysis, increasingly bring a more integrated overview of the control of muscle differentiation. As we are interested in neuromuscular disorders and the initial disturbances during myogenesis in these patients, we first set out to learn how myogenesis is normally regulated. Previously, gene expression profiling studies have been performed by inducing myogenic differentiation in mouse fibroblasts or myoblast cell lines [16–18]. The present study is the first to use primary human myoblast cultures in a time course experiment. This provides a robust and informative model for studying differential gene expression during late myogenesis in humans. We used a general 20K human oligonucleotide microarray as well as a 5K muscle-related cDNA microarray to obtain a complete representation of the gene expression changes during muscle cell maturation.

* Corresponding author. Tel.: +31-71-527-6105; fax: +31-71-527-6075.
E-mail address: ddunnen@lumc.nl (J.T. den Dunnen).

0960-8966/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.nmd.2004.03.008

2. Materials and methods

2.1. Cell culture

Healthy primary human myoblasts were isolated from skeletal muscle biopsies [19] of three healthy individuals (KM109, KM108 and HPP4 [20,21]). Cell cultures were grown at 37 °C and 5% CO_2 in proliferation medium (PM) consisting of Nut.Mix F-10 (HAM) with GlutaMax-1 (Gibco BRL) supplemented with 100 U/ml penicillin, 100 μg/ml streptomycin and 20% heat-inactivated fetal bovine serum (Gibco BRL) on collagen-coated culture flasks/dishes [20,21]. When cells were 80% confluent, they were shifted to differentiation medium (DM) consisting of DMEM without phenol red, supplemented with 1% glucose, 2% GlutaMax-1, 100 U/ml penicillin, 100 μg/ml streptomycin and 2% heat-inactivated fetal bovine serum [20]. All cell cultures used for the experiments had relatively low passage numbers between 6 and 10.

2.2. Immunohistochemical analysis

Myoblast/myotube cultures were fixed (six replicates) in −20 °C methanol at day 0, 1, 2, 4, 6, 10, 14, 19 and 22 after serum deprivation. Immunohistochemical staining of desmin and myosin proteins was performed as described previously [21]. Myogenicity of the culture was determined by calculating the percentage of desmin-positive cells at $t = 0$. KM109 contained 83%, KM108 32% and HPP4 67% desmin-positive cells.

2.3. Array fabrication and pre-hybridization

Cleaning of the glass slides (cut edges, 3 × 1 in.; Menzel) and coating with poly-L-lysine was performed as described previously (http://cmgm.stanford.edu/pbrown/) [22]. In this study two different types of arrays were used. The first is the Sigma-Genosys human oligonucleotide library (18.8K, 60mer with 5′-hexylaminolinker), which was printed over two slides at the Leiden Genome Technology Center. Preparation and printing of the oligonucleotides was performed as described [23]. The second are human cDNA microarrays containing 4417 muscle-related genes and ESTs from a human sequence-verified 40K I.M.A.G.E. cDNA library (Research Genetics) [24]. Clones were selected by combining information from other subsets (Telethon, Italy [25] and Genethon, France) and our own list (GenBank accession numbers available at http://145.88.211.102/humanegenetica/). cDNAs were PCR amplified and purified as described [26] and were spotted in triplicate with an Omnigrid 100® microarrayer (Genemachines). Pre-hybridization was as described for both arrays, except with cross-linking energy of 250 mJ/cm^2 for the oligonucleotide arrays [26].

2.4. Target preparation and hybridization

Total RNA was isolated from myoblasts/myotubes as described [26] just before serum deprivation (day 0) and at day 1, 2, 4, 6, 10, 14, 19 and 22 after serum deprivation.

2.4.1. Oligonucleotide arrays

Total RNA of the KM109 culture (0.5 μg/sample, $t = 0$, 1, 4, 6 and 14) was amplified with the Message Amp kit (Ambion) and the cRNA was labeled through incorporation of aa-UTP (ratio aa-UTP:UTP = 2:3) and coupling with Amershams monoreactive Cy3 and Cy5 dyes before hybridization to the oligonucleotide array [23]. Each sample (750 ng) was labeled with Cy5 and hybridized against a sample of day 0 (labeled with Cy3) and dye-swap experiments were performed.

2.4.2. cDNA arrays

Total RNA (0.5 μg/sample) of all three cell cultures and all timepoints was amplified with the Message Amp kit (Ambion), according to the manufacturer's protocol. For the cDNA arrays separate RNA preparations were used. The three individual cultures were processed separately. After amplification, the samples were labeled by incorporation of Cy3 or Cy5-dUTPs (NEN) during first strand synthesis and hybridized (1.5 μg) to the cDNA array. The samples were hybridized against a common cDNA reference sample with the reference sample always labeled with Cy3 and the target sample with Cy5 [26].

The quality and quantity of the total RNA and cRNA was checked with the Bioanalyzer Lab-on-a-Chip RNA nano assay (Agilent Technologies).

2.5. Data analysis

All slides were scanned with an Agilent scanner (Model 2565BA) and spot intensities were quantified with the GenePix Pro 3.0 program (Axon).

2.5.1. Oligonucleotide arrays

For the hybridizations to the oligo arrays, raw intensity files were imported into Rosetta Resolver® v3.2 (RosettaBio, US) and normalized with the Axon/Genepix error model. Genes that followed two criteria were analyzed: (1) the normalized signal intensity had to be higher than the mean + 2 standard deviations (SDs) of the negative array controls (antisense oligos) and (2) less than a 2-fold change in two self–self hybridizations (day 0). The first criterion should be consistent between the dye-swap hybridizations and on at least one timepoint of the time series. One-way ANOVA was performed with time as variable and genes were considered differentially expressed when the P-value $< 5 \times 10^{-6}$ (Bonferroni corrected) and at least a 2-fold change between one of the timepoints and $t = 0$ was observed. K-means clustering with $K = 8$ (cluster initialization: datacentroid based

search, similarity measure: euclidean distance) was performed with the Functional Genomics application of Spotfire Decision Site 7.1.1 software using normalized ratios (log 2) of genes that were differentially expressed. Functional annotation was determined according to LocusLink and OMIM databases. GenePix data and the normalized log 10 ratio of the hybridizations were submitted to the GEO database, accession numbers GSE908 and GSE909 [27].

2.5.2. cDNA arrays

As the hybridization design of the cDNA arrays did not allow us to analyze gene expression patterns in Rosetta Resolver®, analysis was performed as follows. Local background corrected target intensities of the cDNA hybridizations were normalized using variance stabilization and normalization (VSN) [28]. This transformation h corrects for array-specific spatial deviations and coincides with the natural logarithmic transformation for the large intensities. To identify genes differentially expressed in time, Z-test calculations were performed [17]. In short, three replicate spots present on the arrays were used and timepoint 0 was compared with all other timepoints. This resulted in comparing two groups of three measurements each (per gene, per timepoint). An average expression level from the six experimental measurements (E_{avg}) was calculated for each gene. The deviation (D) of each individual value (E_{ind}) from the average across all experiments (E_{avg}) was also calculated ($D = E_{ind} - E_{avg}$). A Z-score was calculated (Z_D) by standardizing the individual Ds relative to all of the other D values on the array. Features with a $|Z_D| > 1.96$ were considered significantly above or below the average expression level of that gene and were designated as Sig + or Sig − , respectively. The Z-score calculation was repeated, removing the genes already identified as Sig + or Sig − from the distribution, and the $|Z_D| > 1.96$ threshold was applied to the newly calculated Z-scores. This process was repeated through 20 iterations. Having calculated 'Sig' calls for each measurement, all six were combined and a gene was considered significantly up or downregulated when there was a Sig designation in at least five out of six measurements, with three calls in the same direction for one timepoint and at least two calls in the opposite direction for the other timepoint. By randomizing the datasets the number of false-positives was estimated to be 1 in every 230 genes (0.4%). The Z-test was performed for each cell culture separately. The results of the differentially expressed genes in each individual cell culture were combined to obtain a list of genes that showed a distinct expression pattern in time. To further minimize the number of false-positives, a gene had to be significantly up- or downregulated in all three cell cultures in at least one timepoint to be considered differentially expressed.

Clustering was performed with the Functional Genomics application of Spotfire Decision Site 7.1.1 software using the normalized signal intensities of the genes that were differentially expressed. Before performing hierarchical clustering (clustering method: complete linkage, similarity measure: euclidean distance), genes were scaled to make the individual cultures comparable. Before performing K-means clustering with $K = 8$ (cluster initialization: datacentroid based search, similarity measure: euclidean distance) each intensity value was scaled to the average for timepoint 0, obtaining ratio values (Δh). Functional annotation was determined according to LocusLink and OMIM databases. GenePix data and the normalized VSN values of the hybridizations were submitted to the GEO database, accession number GSE906 [27].

2.6. Cross-platform comparisons

The results from the cDNA arrays were compared with the oligonucleotide arrays by linking the UniGene clusters in the GeneHopper program available at www.lgtc.nl/GeneHopper [29]. Comparisons between our data and previously published data were also performed with this program.

2.7. Quantitative reverse transcription polymerase chain reaction

cDNA was prepared (using the total RNA of the KM109 time course) by reverse transcription using 0.5 μg total RNA as template. Random hexamers (40 ng) were used to prime the transcription for 10 min at 70 °C followed by chilling on ice for 10 min. cDNA was synthesized by RevertAid RNaseH$^-$ MuLV reverse transcriptase and accompanying buffer (MBI-Fermentas) using 1 mM dNTPs. The mixture was incubated at room temperature for 10 min before a 2 h incubation step at 42 °C, followed by 10 min at 70 °C. Quantitative PCR was done using the Lightcycler (Roche). PCR was performed on diluted cDNA, with 10 pmol forward and reverse primer, 4 mM $MgCl_2$, 0.225 mM dNTPs, BSA (0.25 μg/μl, Pharmacia Biotech), Taq polymerase (0.2 U/μl), 1 × SYBR Green I (Molecular Probes), 1 × AmpliTaq Reaction Buffer (Perkin Elmer)) in a total volume of 20 μl. A 35-cycle reaction was performed with annealing temperature set at 55 °C. Optimal cDNA dilutions and relative concentrations were determined using a dilution series per gene. All PCR experiments were repeated three times (mean and SD were calculated). Each gene was normalized to the abundance of glyceraldehyde-3-phosphate dehydrogenase mRNA (shows constant expression over time on the arrays). PCR primer pairs were designed using the Primer3 search engine at http://www-genome.wi.mit.edu/cgi-bin/primer/primer3.cgi/. The screened genes and the oligonucleotide primer pairs used for each of the genes in this study corresponded to the following nucleotides: glyceraldehyde-3-phosphate dehydrogenase, 510–529 and 625–644 (NM_002046); myogenic factor 5, 728–747 and 845–864 (NM_005593); transgelin 2, 866–885 and 967–986 (NM_003564); laminin, alpha 4, 3400–3419 and 3511–3530 (NM_002290).

3. Results

Primary human myoblast cultures were forced into differentiation by serum deprivation. Due to the primary nature of the cultures, the cell population was a mixture of mainly myoblasts and fibroblasts. Only myogenic cells are able to differentiate into myotubes, and immunohistochemical stainings were carried out to follow the differentiation process morphologically (Fig. 1). In time it could be seen that myogenic, desmin-positive cells became longer at day 1 and 2 after serum deprivation, and from day 4 cell fusion started and multinucleated, myosin-positive cells became visible. We isolated RNA at day 0, 1, 2, 4, 6, 10, 14, 19 and 22 after serum deprivation for global gene expression analysis.

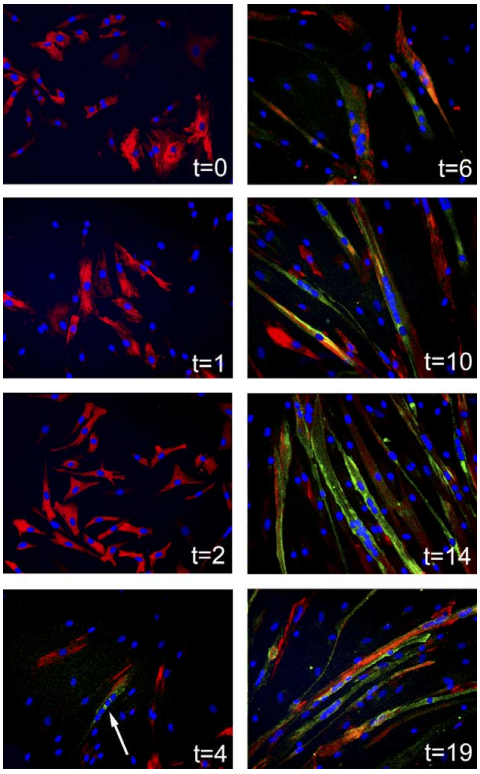


Fig. 1. Immunohistochemical staining of myoblasts/myotubes. Cells of the KM109 cell culture in different stages of myogenesis (time in days after serum deprivation) were stained with DAPI (blue) and antibodies to desmin (red) and myosin (green). In this culture, 85% of the cells were desmin-positive, indicating myogenic potential. Multinucleated, myosin-positive cells can be seen from day 4 (arrow).

3.1. Analysis of oligonucleotide arrays

As a pilot study, five timepoints (day 0, 1, 4, 6 and 14) of one of the cell cultures (KM109) were hybridized to a general human oligonucleotide array (20K). As only one of the three cell lines was hybridized to the oligonucleotide arrays, a stringent data selection was performed ($P < 5 \times 10^{-6}$, at least 2-fold change in one of the timepoints) to obtain a list of differentially expressed genes. This selection revealed that 139 genes were differentially expressed, of which, over time, 67 were up- and 72 were downregulated (Supplemental Table 2, available at http://145.88.211.102/humanegenetica/). K-means clustering with eight clusters (a–h) divided genes into groups with distinct expression patterns (Fig. 2). The figure shows one cluster of genes, which were upregulated shortly after induction of differentiation (Fig. 2, cluster b), whereas three groups showed a delayed reaction with upregulation from day 4 (Fig. 2, clusters a, c and d). All genes that were downregulated show an immediate effect from day 1, except for the genes in cluster g (Fig. 2, cluster e–h).

3.2. Analysis of muscle-related cDNA arrays

We hybridized time course samples of all three different myoblast cell cultures to a muscle-related cDNA array that contained a selection of 5000 clones with presumed expression in muscle. To minimize the number of false-positives, the selection of differentially expressed genes was such that only genes that were consistently up- or downregulated in all three cultures were included. Analysis of the hybridizations on the 5K cDNA array (see Section 2) shows that, during the time course, 78 genes were upregulated and 68 genes were downregulated in all three

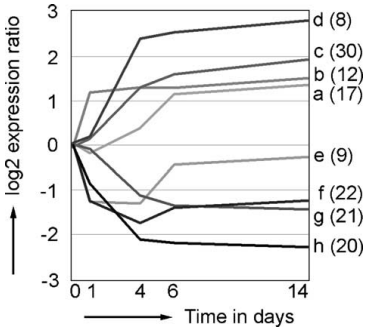


Fig. 2. K-means clustering of genes that are differentially expressed in time according to oligonucleotide arrays hybridized with one primary human cell culture (KM109). Log 2 ratios are shown. Clusters a–d show upregulation in time, whereas clusters e–h show downregulation in time (number of genes for each cluster between brackets).

cell cultures (Table 1). Highest fold changes were found in the cell culture that was the most myogenic. This suggests that differential expression is mainly occurring in myoblasts and not fibroblasts. To determine the synchronicity in expression changes during the differentiation of the three individual cell cultures, we performed hierarchical clustering on the differentially expressed genes (Fig. 3). For timepoints 0, 4 and 14, the cell cultures form separate clusters. However, timepoints 1 and 2 are not separated by the clustering. This is indicative of differences in differentiation rate between the individual cultures in the first stages of differentiation. The observation that the clusters of timepoints 4 and 14 are close together (can be seen in the dendrogram of Fig. 3), indicates that only minor changes in gene expression are found in the later stages of differentiation. The normalized expression ratios of the differentially expressed genes were grouped in eight clusters (a–h) using K-means clustering (Fig. 4). In general, the cDNA clusters have the same shape as the oligonucleotide clusters, except for two oligo clusters that do not have counterparts with similar expression profiles on the cDNA arrays (Fig. 2, cluster a and e). Cluster a shows a delayed upregulation starting at days 4–6 and cluster e shows a temporary downregulation of a group of genes.

Differentially expressed genes were functionally annotated using Locuslink and OMIM databases (Table 1) and the majority of the downregulated genes was assigned to be involved in cell growth and maintenance, metabolism and cell cycle progression. It also shows that genes immediately upregulated in the time course belong to the categories adhesion/matrix, cell cycle/DNA replication and structural/cytoskeletal (Fig. 4, clusters a, b and e). The clusters of genes, which are upregulated in a later stage of myogenesis, belong mainly to the structural/cytoskeletal category and nuclear regulatory factors (Fig. 4, clusters c and d). Of the 146 differentially expressed genes, about 30% had no functional assignment.

3.3. Confirmation by quantitative RT-PCR

To confirm the data obtained by microarray analysis, quantitative RT-PCR was performed. We picked three genes that were differentially expressed in time (two with known function and one with unknown function) and tested their temporal expression. The results show that the general expression patterns in time of laminin alpha 4, myogenic factor 5 and transgelin 2 are comparable with the array results (Fig. 5).

4. Discussion

In this study, we have identified genes involved in human myogenesis by determining gene expression profiles from three human primary myoblast cell cultures, differentiated in vitro. To our knowledge, our study is the first to analyze human skeletal muscle cell differentiation in primary human myoblasts on a genome-wide scale. Although culturing primary human myoblast cells and triggering differentiation is not the easiest way to study the process of myogenesis, it is probably the best in vitro model system to do so. Passaging of immortalized cell lines for long periods leads to cell selection and gives rise to spontaneously transformed cells that do not represent the original cell population present in vivo [30,31]. Furthermore, expression of genes in primary cell cultures may be lost in immortalized cell lines, giving an incomplete picture of the expression profile [32]. In our study, temporal differential expression was identified for 146 genes in all three cell cultures, of which 86, mostly with unknown function, have not been previously reported to be involved in myogenesis. These genes may be unique for human myogenesis but it is also possible that they were missed by previous studies because immortalized cell lines were used.

4.1. General expression pattern in time

Changes in expression levels take place immediately following induction of differentiation through serum withdrawal, demonstrated by the separation of timepoints 0 and 1 and 2 in hierarchical clustering (Fig. 3). Variable myogenicity of the cell cultures and culture-specific fusion rates probably lead to dissimilar expression patterns in the individual cultures at the early time points (Fig. 3, timepoints 1 and 2). Consequently, timepoint 1 of one culture can be more similar to timepoint 2 of another culture. The observation that the three cell cultures were not fully synchronous provides a possible explanation why temporarily up- or downregulated genes (at a single timepoint) were not found. As shown by K-means and hierarchical clustering, most of the changes in gene expression are complete after day 4, resulting in an almost constant state of gene expression in all three cell cultures. Strikingly, however, at this timepoint still only few myotubes are visible (Fig. 1) thus fusion is only just starting. This indicates that alterations in gene expression are especially critical in early myogenesis, while post-transcriptional remodeling of cells is more important during the later phase of myoblast differentiation.

4.2. Genes differentially expressed according to the oligonucleotide arrays

In initial analysis on the oligonucleotide arrays 139 genes were found to be differentially expressed. About 30% of these genes are not represented on the cDNA arrays but these may be equally relevant for human myogenesis. For instance, the strongly downregulated *ID1*, *ID3* and *TGFβI* which previously have been indicated to be downregulated during myogenesis [15,33]. *Triadin*, which is highly expressed in skeletal muscle and is proposed to play a structural role by anchoring Calsequestrin to ryanodine-sensitive calcium

Table 1
Genes differentially expressed during human myogenesis

GenBank[a]	Symbol[b]	Name and function[c]	Clusters in K-means clustering		Implied in:		
			cDNA arrays[d]	Oligo arrays[e]	Myo[f]	Reg[g]	DMD[h]
Structural/cytoskeletal							
AA669126	PPP1R12A	Protein phosphatase 1, regulatory (inhibitor) subunit 12A	a ↑	c ↑			
AA447737	CALD1	Caldesmon 1	a ↑				
N70734	TNNT2	Troponin T2, cardiac	a ↑	c ↑	[18]	[46]	
AA418414	SARCOSIN	Sarcomeric muscle protein	a ↑				
AA972352	ALP	Alpha-actinin-2-associated LIM protein	b ↑	c ↑	[18]		
AA828221	MYH8	Myosin, heavy polypeptide 8, skeletal muscle, perinatal	c ↑				[47,48]
N78927	MYL2	Myosin, light polypeptide 2, regulatory, cardiac, slow	c ↑				
M12126	TPM2	Tropomyosin 2 (beta)	c ↑			[46]	
AA677258	MYL4	Myosin, light polypeptide 4, alkali; atrial, embryonic	d ↑		[18]		[49]
AI005197	HUMMLC2B	Myosin light chain 2	d ↑	d ↑	[18]		
U34976	SGCG	Sarcoglycan, gamma	d ↑				[49]
AA705225	MYL4	Myosin, light polypeptide 4, alkali; atrial, embryonic	d ↑	d ↑			[49]
AA449932	TNNT3	Troponin T3, skeletal, fast	d ↑	c ↑	[17,18]		
AA872006	TTN	Titin	d ↑				
K00558	K-ALPHA-1	Tubulin, alpha, ubiquitous	f ↓				[47]
NM_0060	TUBB2	Tubulin, beta, 2	f ↓				
AA699926	SNTA1	Syntrophin, alpha 1 (dystrophin-associated protein A1, 59 kDa, acidic component)	f ↓				[47]
R22977	MSN	Moesin	h ↓				
Nuclear regulatory factors							
AA187933	TAZ	Transcriptional co-activator with PDZ-binding motif (TAZ)	c ↑				[48]
AA453175	BIN1	Bridging integrator 1	c ↑		[17]	[46]	
AA600217	ATF4	Activating transcription factor 4 (tax-responsive enhancer element B67)	c ↑				
W96114	HNRPH1	Heterogeneous nuclear ribonucleoprotein H1 (H)	f ↓		[18]		
AA464856	ID4	Inhibitor of DNA binding 4, dominant negative helix–loop–helix protein	f ↓				
AA504656	LTBP1	Latent transforming growth factor beta binding protein 1	f ↓				
H27564	DDX5	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 5 (RNA helicase, 68 kDa)	f ↓		[18]		
X14894	MYF5	Myogenic factor 5	h ↓*	h ↓			
Adhesion/matrix							
AA453712	LUM	Lumican	b ↑				[47,48]
N73836	FN1	Fibronectin 1	b ↑				[47]
R43734	LAMA4	Laminin, alpha 4	b ↑*				
H22914	BPAG1	Bullous pemphigoid antigen 1, 230/240 kDa	a ↑				
Cell cycle/DNA replication							
AA292054	GAS1	Growth arrest-specific 1	a ↑			[46]	
AA083032	CCNG1	Cyclin G1	b ↑				
AA676387	CPR2	Cell cycle progression 2 protein	b ↑				
N70463	BTG1	B-cell translocation gene 1, anti-proliferative	b ↑				
H59203	CDC6	CDC6 cell division cycle 6 homolog	g ↓				
R25788	CCNB1	Cyclin B1	g ↓				
AA663995	MCM6	MCM6 minichromosome maintenance deficient 6	g ↓	f ↓			
AA774665	CCNB2	Cyclin B2	g ↓				
AA458994	CCNA2	Cyclin A2	h ↓				
AA397813	CKS2	CDC28 protein kinase regulatory subunit 2	h ↓	h ↓		[46]	

(*continued on next page*)

Table 1 (*continued*)

GenBank[a]	Symbol[b]	Name and function[c]	Clusters in K-means clustering		Implied in:		
			cDNA arrays[d]	Oligo arrays[e]	Myo[f]	Reg[g]	DMD[h]
Cell growth and or maintenance							
AA292226	SLC6A8	Solute carrier family 6 (neurotransmitter transporter, creatine), member 8	a ↑				
AA999990	EIF4A2	Eukaryotic translation initiation factor 4A, isoform 2	a ↑				
AA486626	PABPC1	Poly(A) binding protein, cytoplasmic 1	a ↑				
AA481438	SERPING1	Serine (or cysteine) proteinase inhibitor, clade G (C1 inhibitor), member 1	d ↑				
AA917861	CLIC4	Chloride intracellular channel 4	f ↓				
AA437212	AP1S2	Adaptor-related protein complex 1, sigma 2 subunit	f ↓				
AI369144	EIF4EBP1	Eukaryotic translation initiation factor 4E binding protein 1	f ↓				
T59055	XPO1	Exportin 1 (CRM1 homolog, yeast)	f ↓				
R32756	EWSR1	Ewing sarcoma breakpoint region 1	f ↓			[46]	
AA054287	RBM3	RNA binding motif protein 3	f ↓				
AA044059	VDAC1	Voltage-dependent anion channel 1	g ↓				
AA865872	YARS	Tyrosyl-tRNA synthetase	g ↓				
AI017703	EIF3S3	Eukaryotic translation initiation factor 3, subunit 3 gamma, 40 kDa	g ↓				
AA598400	SFRS3	Splicing factor, arginine/serine-rich 3	g ↓		[18]	[46]	
AA663986	FBL	Fibrillarin	h ↓			[46]	
AA480866	RBM3	RNA binding motif protein 3	h ↓			[46]	
AI668800	H2AFZ	H2A histone family, member Z	h ↓				
Metabolism							
NM_004265	FADS2	Fatty acid desaturase 2	a ↑				
AA446822	LPIN1	Lipin 1	a ↑				
AA668425	AGL	Amylo-1, 6-glucosidase, 4-alpha-glucanotransferase	a ↑				
R93124	AKR1C1	Aldo-keto reductase family 1, member C1	b ↑	a ↑			
AI361530	FACL2	Fatty-acid-coenzyme A ligase, long-chain 2	b ↑				
H38650	SLC2A5	Solute carrier family 2 (facilitated glucose/fructose transporter), member 5	c ↑				
AA156571	AARS	Alanyl-tRNA synthetase	f ↓				
AA460115	ODC1	Ornithine decarboxylase 1	f ↓			[46]	
AA664101	ALDH1A1	Aldehyde dehydrogenase 1 family, member A1	f ↓		[18]		
R91438	PPP1CA	Protein phosphatase 1, catalytic subunit, alpha isoform	f ↓				
AA599092	PPP2CA	Protein phosphatase 2 (formerly 2A), catalytic subunit, alpha isoform	f ↓		[18]		
T77281	ALDOC	Aldolase C, fructose-bisphosphate	f ↓				
AA894927	ASNS	Asparagine synthetase	f ↓		[18]		
AA029851	GALNT1	UDP-*N*-acetyl-alpha-D-galactosamine:polypeptide *N*-acetylgalactosaminyltransferase 1	f ↓		[18]		
AI291445	PPP1R1A	Protein phosphatase 1, regulatory (inhibitor) subunit 1A	f ↓			[46]	
AI362803	PRPS1	Phosphoribosyl pyrophosphate synthetase 1	g ↓		[18]	[46]	
N33274	PAICS	Phosphoribosylaminoimidazole carboxylase	h ↓				
Receptors/signalling							
AI263201	CXCL12	Chemokine (C-X-C motif) ligand 12 (stromal cell-derived factor 1)	a ↑				
AI261580	ACVR1	Activin A receptor, type I	a ↑				
AA447115	CXCL12	Chemokine (C-X-C motif) ligand 12 (stromal cell-derived factor 1)	a ↑				
AA629897	LAMR1	Laminin receptor 1 (ribosomal protein SA, 67 kDa)	g ↓			[46]	
R19628	BIRC2	Baculoviral IAP repeat-containing 2	g ↓				
H22826	LMO7	LIM domain only 7	g ↓				
AA873060	STMN1	Stathmin 1/oncoprotein 18	g ↓				

(*continued on next page*)

Table 1 (*continued*)

GenBank[a]	Symbol[b]	Name and function[c]	Clusters in K-means clustering		Implied in:		
			cDNA arrays[d]	Oligo arrays[e]	Myo[f]	Reg[g]	DMD[h]
Proteolysis/apoptosis/chaperone							
AA449361	RNF13	Ring finger protein 13	a ↑				
AI572217	HSPB3	Heat shock 27 kDa protein 3	c ↑	c ↑			
AA865265	CYCS	Cytochrome *c*, somatic	f ↓				[47]
H99681	DP1	Likely ortholog of mouse deleted in polyposis 1	f ↓				
AF070561	OK/SW-cl.56	Beta 5-tubulin	g ↓				
H37989	OK/SW-cl.56	Beta 5-tubulin	g ↓				
Others and unknown							
T59873	HSPC134	HSPC134 protein	a ↑				
N54901	FRCP2	FRCP2 likely ortholog of mouse fibronectin type III repeat containing protein 2	a ↑				
R51218	KIAA0092	Translokin	a ↑				
AA156251	SPUF	Secreted protein of unknown function	a ↑				
AI017846	PHGDHL1	Phosphoglycerate dehydrogenase like 1	a ↑				
AA488084	SOD2	Superoxide dismutase 2, mitochondrial	a ↑				
AA130584	CEACAM5	Carcinoembryonic antigen-related cell adhesion molecule 5	a ↑				
AA702561	ZNF288	Zink finger protein 288	a ↑				
AI015986	RASSF2	ras association (RalGDS/AF-6) domain family 2	a ↑				
AA454668	PTGS1	Prostaglandin-endoperoxide synthase 1 (prostaglandin G/H synthase and cyclooxygenase)	a ↑				
N58145	LHFP	Lipoma HMGIC fusion partner	a ↑				
AA464246	HLA-C	Major histocompatibility complex, class I, C	a ↑				
AA991810	CDK11	Cyclin-dependent kinase (CDC2-like) 11	b ↑				
AA046700	FBXO32	F-box only protein 32	b ↑				
AA074535	PBXIP1	Pre-B-cell leukemia transcription factor interacting protein 1	b ↑				
R70518	OPTN	Optineurin	b ↑				
W69743		Sapiens mRNA of muscle-specific gene M1	c ↑	d ↑			
AA464691	DKFZp564I1922	Adlican	c ↑				[48]
AA401441	BF	B-factor, properdin	c ↑				
AA933056	RASSF4	ras association (RalGDS/AF-6) domain family 4	d ↑				
N52254	SH3BGR	sh3 domain binding glutamic acid-rich protein	d ↑		[18]		
T62048	C1S	Complement component 1, s subcomponent	e ↑				[47]
AA448599	F13A1	Coagulation factor XIII, A1 polypeptide	e ↑	d ↑			[47]
W30988	ANGPTL4	Angiopoietin-like 4	f ↓				
N32919	GNPNAT1	Glucosamine-phosphate *N*-acetyltransferase 1	f ↓				
AA447782	SCYL1	SCY1-like 1	f ↓				
H93393	FUS	Fusion, derived from t(12;16) malignant liposarcoma	f ↓		[18]		
AA490059	ENAH	Enabled homolog (*Drosophila*)	f ↓				
H20652	ARL6IP	ADP-ribosylation factor-like 6 interacting protein	f ↓			[46]	
H71092	ZD47C12		f ↓				
H08424	SMYD2	SET and MYND domain containing 2	f ↓				
AA192419	BLVRA	Biliverdin reductase A	f ↓		[18]		
T62060	SERPINC1	Serine (or cysteine) proteinase inhibitor, clade C (antithrombin), member 1	f ↓				
H08564	TAGLN2	Transgelin 2	g ↓*				
AA489609	KIAA0864	rho interacting protein 3	g ↓				
AI362799	D21S2056E	DNA segment of chromosome 21 (unique) expressed sequence	g ↓				
H59915	CD24	CD24 antigen (small lung carcinoma cluster 4 antigen)	g ↓				
AA456868	LMNB2	Lamin B2	h ↓				
N64508	PODXL	Podocalyxin-like	h ↓	h ↓		[46]	

(*continued on next page*)

Table 1 (*continued*)

GenBank[a]	Symbol[b]	Name and function[c]	Clusters in K-means clustering		Implied in:		
			cDNA arrays[d]	Oligo arrays[e]	Myo[f]	Reg[g]	DMD[h]
ESTs							
R98407	EST		a ↑				
H29604	EST		a ↑				
AA286819	FLJ12436	Hypothetical protein FLJ12436	a ↑				
AA405488	MGC16063	Hypothetical protein MGC16063	a ↑				
AA417956	FLJ38973	Hypothetical protein FLJ38973	a ↑				
AA733003	EST		a ↑				
AA460708	FLJ14834	Hypothetical protein FLJ14834	a ↑				
AA630373	FLJ90798	Hypothetical protein FLJ90798	a ↑				
AA058578	EST		a ↑				
R92352	DKFZp762H185	Hypothetical protein DKFZp762H185	a ↑				
H72027	No cluster		a ↑				
AA486085	No cluster		a ↑				
AA479883	FLJ21127	Hypothetical protein FLJ21127	b ↑				
AA927761	LOC128977	Hypothetical protein LOC128977	b ↑				
AI335831	EST		b ↑				
AA115749	EST		b ↑	b ↑			
AA910255	LOC56757	Hypothetical protein LOC56757	b ↑				
AA115749	EST		b ↑				
H14810	EST		b ↑				
AA630373	FLJ90798	Hypothetical protein FLJ90798	b ↑				
AA489055		Sapiens, clone IMAGE: 3891285, mRNA	f ↓				
AI291262	EST		f ↓				
R21530	EST		f ↓				
H56918	EST		f ↓				
AA027168	EST		f ↓				
AA456646	EST		g ↓				
AA136125	No cluster		g ↓				

[a] Genbank accession number.
[b] Gene symbol.
[c] Gene name and function.
[d] Cluster number (see Fig. 4) from K-means clustering cDNA arrays (*confirmed by quantitative RT-PCR).
[e] Cluster number (see Fig. 2) from K-means clustering oligonucleotide arrays.
[f] Genes differentially regulated in mouse myogenesis according to: Bergstrom et al. [17] and Delgado et al. [18].
[g] Genes differentially regulated during mouse regeneration according to: Yan et al. [46].
[h] Genes differentially regulated in DMD patients according to: Chen et al. [47], Haslett et al. [48], and Noguchi et al. [49].

channels, was found to be upregulated on the oligonucleotide arrays [34]. *Triadin* belongs to a group of genes (Fig. 2, cluster a), which show upregulation at a later stage in the myogenesis (from $t = 4–6$). In concordance with these results, *Calsequestrin* also shows an upregulation in expression (Fig. 2, cluster a) and is known to play a role in the storage of Ca^{2+} [35]. Striking is that this cluster (Fig. 2, cluster a) is one of the two that do not correlate to any cluster of the hybridization results of the cDNA array. When looking at the biological functions of the genes in this group it is obvious that mainly structural and membrane related genes are in this cluster together with some ESTs. For the other cluster (Fig. 2, cluster e), a large functional diversity was found.

4.3. Expression on the cDNA arrays and the overlap with oligonucleotide arrays

After a general analysis of myogenesis on the oligonucleotide arrays, a more extensive and focused analysis was performed using a muscle-related cDNA array. Of the 146 genes differentially regulated in time, as revealed by the cDNA arrays, 59 could be analyzed on the oligonucleotide arrays. The overlap in differentially up- and downregulated genes between the two platforms was 15 genes (Table 1). Although no inconsistencies were found, the low number of differentially expressed genes identified on both types of arrays is probably due to the different methods of data analysis of the two platforms, which may contribute to variability in the results. The inclusion of two extra cell cultures with lower myogenicity in the cDNA analysis will also result in the detection of only highly differentially expressed genes during myogenesis.

4.4. Functional annotation of differentially expressed genes

Functional annotation of the differentially regulated genes indicates that the shift from proliferation to differentiation takes place immediately following serum deprivation. Genes that primarily function in the stimulation

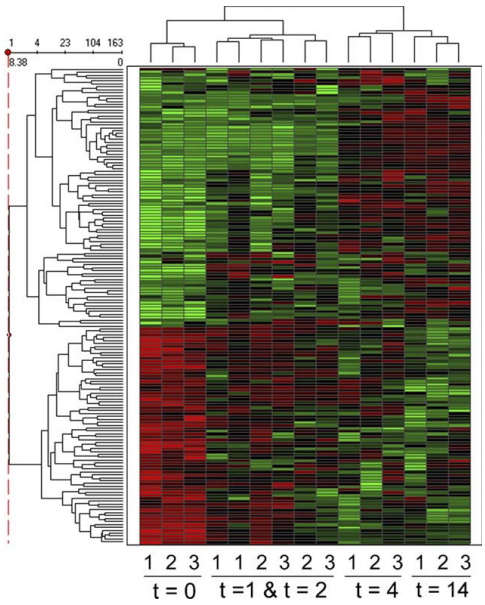


Fig. 3. Hierarchical clustering of genes up- (red) and down-(green) regulated in time ($t = 0$, $t = 1$, $t = 2$, $t = 4$ and $t = 14$) according to cDNA arrays hybridized with three primary human muscle cell cultures (1 = KM109, 2 = KM108, 3 = HPP4). Rows are differentially expressed genes. Columns are individual cell cultures.

of proliferation are downregulated (e.g. *CDC6* [36]). Accordingly, genes that are immediately upregulated after serum deprivation are involved in inhibition of proliferation and stimulation of differentiation, which is also indicative of the switch towards more specialized cells (e.g. *BTG1*

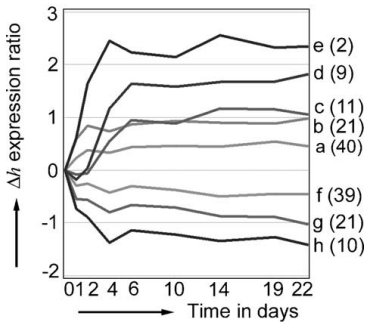


Fig. 4. K-means clustering of genes that are differentially expressed in time according to cDNA arrays hybridized with three primary human cell lines. Results are shown from the highest myogenic cell culture (KM109). Δh ratios are shown which coincide with the natural logarithm. Clusters a–e show upregulation in time whereas clusters f–h show downregulation in time (number of genes for each cluster between brackets).

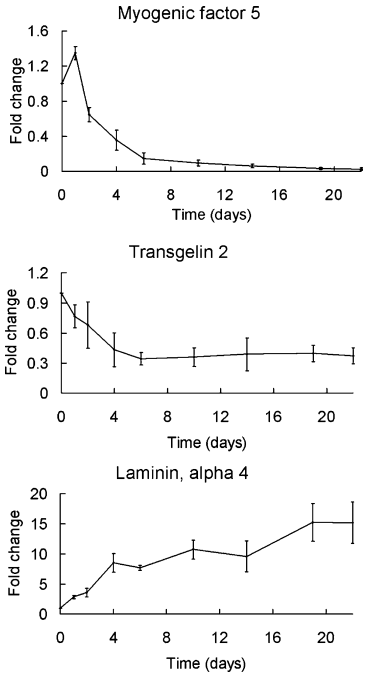


Fig. 5. Quantitative RT-PCR of three genes (*Myogenic factor 5*, *Transgelin 2*, and *Laminin alpha 4*) in the KM109 cell culture shows comparable gene expression results with the microarray expression data. *Myogenic factor 5* is in cluster h, *Transgelin 2* in cluster g and *Laminin alpha 4* in cluster b of Fig. 4.

and *BIN1* resp. [37,38]). Another immediate change following serum deprivation is the upregulation of specific adhesion genes and genes involved in the extracellular matrix (Table 1). This confirms the transformation of the myoblasts towards differentiated myotubes, as it is known that skeletal muscle cells synthesize basal lamina-type macromolecules and incorporate them into an insoluble, extracellular matrix during their maturation [39]. An interesting finding is that the structural/cytoskeletal genes are upregulated in two phases; one group shows an early, immediate upregulation (day 1) and another group a late upregulation (day 4). The early phase is mainly characterized by smooth or cardiac muscle-specific genes, which have also been implicated in the initiation of skeletal myogenesis (e.g. *TNNT2* and *CALD1* [40,41]). The genes upregulated from day 4 are primarily genes that determine the final structure of skeletal muscle (e.g. *TNNT3* and *SGCG* [42,43]). Of the 146 genes differentially expressed in time, 43 could not yet be functionally assigned to a specific group. However, clustering of temporal expression patterns assists in the allocation of a potential function to these genes and ESTs. For instance, *transgelin 2*, with unknown function, is a structural homolog of *transgelin* (*TAGLN/SM22α*),

which is highly expressed in both embryonal and adult smooth muscle cells, and only transiently detected in skeletal and heart myogenic lineages during muscle cell development [44]. In this time course, *transgelin 2* is also highly expressed in the myoblasts and downregulated upon myoblast differentiation (confirmed by quantitative RT-PCR), suggesting that it is a structural as well as a functional homolog of *transgelin*.

4.5. Cross-species comparison of genes involved in myogenesis

Our results are in agreement with data on gene expression during murine myogenesis (Table 1) [17,18]. For instance, the structural/cytoskeletal category is also predominantly upregulated at later timepoints in those studies, and consists largely of structural muscle genes. Overlap is also seen in the upregulation of adhesion/matrix genes and downregulation of metabolic genes [17,18].

4.6. Myogenesis and muscle regeneration

Several studies have been performed in which muscle damage was induced in mice in order to perform time course analysis of muscle regeneration [45,46]. Comparing our data to the results of Yan et al., regarding effects of satellite cells only, shows that the late phase of muscle differentiation studied clearly resembles the process of muscle regeneration. Consistent with our data, upregulation of muscle-specific genes and genes that are known to be induced during skeletal muscle differentiation is observed. This supports the relevance of myoblast differentiation in vitro as a model system for studying in vivo regeneration by satellite cells.

4.7. Muscle regeneration in DMD patients

In addition, our results partly overlap with results from expression profiling studies in skeletal muscle from Duchenne muscular dystrophy (DMD) patients (Table 1). DMD muscle overexpresses structural muscle genes, which were also found to be upregulated in our study [47–49]. There is only a small group of structural genes being downregulated during our time course, and strikingly Chen et al. found most of these genes to be also downregulated in DMD and α-sarcoglycan deficient patients (*α-1 Syntrophin* and *α-Tubulin* [47]). This overlap confirms that, in patients with certain neuromuscular dystrophies, muscle regeneration is also likely to be an ongoing process.

4.8. Future prospects

Finally, some genes in our study with unknown function are also differentially expressed in mouse myogenesis and muscle regeneration. Since they are likely to play a (key) role in muscle cell differentiation, functional studies of these genes in myogenesis will probably be highly rewarding. Notably, these are potential candidate genes for several neuromuscular disorders with an as yet unidentified genetic basis.

4.9. Note added in proof

During the submission of this manuscript, a paper was published by Tomczak et al. [50]. In this paper, they describe an interesting gene expression profiling study of mouse C2C12 cells during differentiation. We compared our results to the results presented in this paper and 30% of the genes that we present here, is also present in their list of differentially expressed genes. Furthermore, the shapes of the clusters of the differentially expressed genes that we present, closely resembles theirs, again showing large similarities between human and mouse myogenesis [50].

Acknowledgements

We would like to thank Renée X. de Menezes (Medical Statistics, LUMC) and Stefan White (Human Genetics, LUMC) for critical reading of the manuscript and valuable suggestions. We would like to thank the Leiden Genome Technology Center for providing the oligonucleotide arrays. This work was supported by grants from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO, NL) and the Muscular Dystrophy Campaign (MDC, UK).

References

[1] Muntoni F, Brown S, Sewry C, Patel K. Muscle development genes: their relevance in neuromuscular disorders. Neuromuscul Disord 2002;12:438–46.
[2] Davis RL, Weintraub H, Lassar AB. Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell 1987;51:987–1000.
[3] Braun T, Buschhausen-Denker G, Bober E, Tannich E, Arnold HH. A novel human muscle factor related to but distinct from MyoD1 induces myogenic conversion in 10T1/2 fibroblasts. Eur Mol Biol Org J 1989;8:701–9.
[4] Edmondson DG, Olson EN. A gene with homology to the myc similarity region of MyoD1 is expressed during myogenesis and is sufficient to activate the muscle differentiation program. Genes Dev 1989;3:628–40.
[5] Braun T, Bober E, Winter B, Rosenthal N, Arnold HH. Myf-6, a new member of the human gene family of myogenic determination factors: evidence for a gene cluster on chromosome 12. Eur Mol Biol Org J 1990;9:821–31.
[6] Miner JH, Wold B. Herculin, a fourth member of the MyoD family of myogenic regulatory genes. Proc Natl Acad Sci USA 1990;87:1089–93.
[7] Sabourin LA, Rudnicki MA. The molecular regulation of myogenesis. Clin Genet 2000;57:16–25.
[8] Molkentin JD, Black BL, Martin JF, Olson EN. Cooperative activation of muscle gene expression by MEF2 and myogenic bHLH proteins. Cell 1995;83:1125–36.

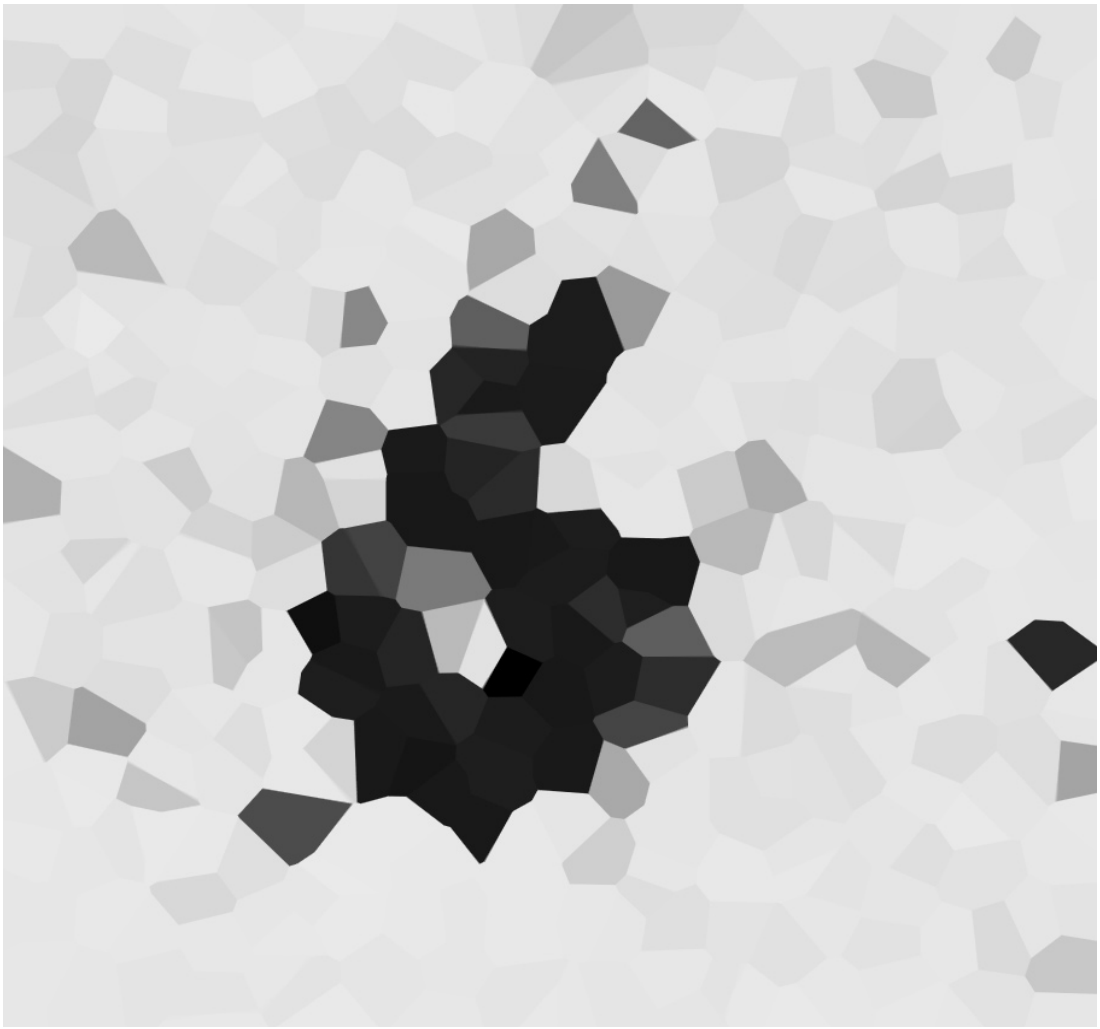
[9] Goulding M, Lumsden A, Paquette AJ. Regulation of Pax-3 expression in the dermomyotome and its role in muscle development. Development 1994;120:957–71.
[10] Williams BA, Ordahl CP. Pax-3 expression in segmental mesoderm marks early stages in myogenic cell specification. Development 1994; 120:785–96.
[11] Bober E, Franz T, Arnold HH, Gruss P, Tremblay P. Pax-3 is required for the development of limb muscles: a possible role for the migration of dermomyotomal muscle progenitor cells. Development 1994;120:603–12.
[12] Amthor H, Christ B, Weil M, Patel K. The importance of timing differentiation during limb muscle development. Curr Biol 1998;8: 642–52.
[13] Munsterberg AE, Kitajewski J, Bumcrot DA, McMahon AP, Lassar AB. Combinatorial signaling by Sonic hedgehog and Wnt family members induces myogenic bHLH gene expression in the somite. Genes Dev 1995;9:2911–22.
[14] Martin JF, Li L, Olson EN. Repression of myogenin function by TGF-beta 1 is targeted at the basic helix–loop–helix motif and is independent of E2A products. J Biol Chem 1992;267:10956–60.
[15] Benezra R, Davis RL, Lockshon D, Turner DL, Weintraub H. The protein Id: a negative regulator of helix–loop–helix DNA binding proteins. Cell 1990;61:49–59.
[16] Jost JP, Oakeley EJ, Zhu B, Benjamin D, Thiry S, Siegmann M, Jost Y-C. 5-Methylcytosine DNA glycosylase participates in the genome-wide loss of DNA methylation occurring during mouse myoblast differentiation. Nucleic Acids Res 2001;29:4452–61.
[17] Bergstrom DA, Penn BH, Strand A, Perry RL, Rudnicki MA, Tapscott SJ. Promoter-specific regulation of MyoD binding and signal transduction cooperate to pattern gene expression. Mol Cell 2002;9: 587–600.
[18] Delgado I, Huang X, Jones S, Zhang L, Hatcher R, Gao B, Zhang P. Dynamic gene expression during the onset of myoblast differentiation in vitro. Genomics 2003;82:109–21.
[19] Rando TA, Blau HM. Primary mouse myoblast purification, characterization, and transplantation for cell-mediated gene therapy. J Cell Biol 1994;125:1275–87.
[20] Aartsma-Rus A, Bremmer-Bout M, Janson AA, Den Dunnen JT, van Ommen GJ, van Deutekom JC. Targeted exon skipping as a potential gene correction therapy for Duchenne muscular dystrophy. Neuromuscul Disord 2002;12(Suppl. 1):S71–7.
[21] Aartsma-Rus A, Janson AA, Kaman WE, et al. Therapeutic antisense-induced exon skipping in cultured muscle cells from six different DMD patients. Hum Mol Genet 2003;12:907–14.
[22] DeRisi JL, Iyer VR, Brown PO. Exploring the metabolic and genetic control of gene expression on a genomic scale. Science 1997;278:680–6.
[23] 't Hoen PA, de Kort F, van Ommen GJ, Den Dunnen JT. Fluorescent labelling of cRNA for microarray applications. Nucleic Acids Res 2003;31:e20.
[24] Lennon G, Auffray C, Polymeropoulos M, Soares MB. The I.M.A.G.E. consortium: an integrated molecular analysis of genomes and their expression. Genomics 1996;33:151–2.
[25] Lanfranchi G, Muraro T, Caldara F, et al. Identification of 4370 expressed sequence tags from a 3′-end-specific cDNA library of human skeletal muscle by DNA sequencing and filter hybridization. Genome Res 1996;6:35–42.
[26] Sterrenburg E, Turk R, Boer JM, van Ommen GB, Den Dunnen JT. A common reference for cDNA microarray hybridizations. Nucleic Acids Res 2002;30:e116.
[27] Edgar R, Domrachev M, Lash AE. Gene expression omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 2002;30:207–10.
[28] Huber W, Von Heydebreck A, Sultmann H, Poustka A, Vingron M. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. Bioinformatics 2002; 18(Suppl. 1):S96–104.
[29] Svensson BA, Kreeft AJ, van Ommen GJ, Den Dunnen JT, Boer JM. GeneHopper: a web-based search engine to link gene-expression platforms through GenBank accession numbers. Genome Biol 2003;4: R35.
[30] Todaro GJ, Green H. Quantitative studies of the growth of mouse embryo cells in culture and their development into established lines. J Cell Biol 1963;17:299–313.
[31] Aaronson SA, Todaro GJ. Basis for the acquisition of malignant potential by mouse cells cultivated in vitro. Science 1968;162:1024–6.
[32] Lawson D, Harrison M, Shapland C. Fibroblast transgelin and smooth muscle SM22alpha are the same protein, the expression of which is down-regulated in many cell lines. Cell Motil Cytoskeleton 1997;38: 250–7.
[33] Massague J, Cheifetz S, Endo T, Nadal-Ginard B. Type beta transforming growth factor is an inhibitor of myogenic differentiation. Proc Natl Acad Sci USA 1986;83:8206–10.
[34] Guo W, Campbell KP. Association of triadin with the ryanodine receptor and calsequestrin in the lumen of the sarcoplasmic reticulum. J Biol Chem 1995;270:9027–30.
[35] Raichman M, Panzeri MC, Clementi E, et al. Differential localization and functional role of calsequestrin in growing and differentiated myoblasts. J Cell Biol 1995;128:341–54.
[36] Yan Z, DeGregori J, Shohet R, Leone G, Stillman B, Nevins JR, Williams RS. Cdc6 is regulated by E2F and is essential for DNA replication in mammalian cells. Proc Natl Acad Sci USA 1998;95: 3603–8.
[37] Rouault JP, Rimokh R, Tessa C, et al. BTG1, a member of a new family of antiproliferative genes. Eur Mol Biol Org J 1992;11: 1663–70.
[38] Wechsler-Reya RJ, Elliott KJ, Prendergast GC. A role for the putative tumor suppressor Bin1 in muscle cell differentiation. Mol Cell Biol 1998;18:566–75.
[39] Beach RL, Burton WV, Hendricks WJ, Festoff BW. Extracellular matrix synthesis by skeletal muscle in culture. Proteins and effect of enzyme degradation. J Biol Chem 1982;257:11437–42.
[40] Townsend PJ, Farza H, MacGeoch C, et al. Human cardiac troponin T: identification of fetal isoforms and assignment of the TNNT2 locus to chromosome 1q. Genomics 1994;21:311–6.
[41] Sobue K, Sellers JR. Caldesmon, a novel regulatory protein in smooth muscle and nonmuscle actomyosin systems. J Biol Chem 1991;266: 12115–8.
[42] Wu QL, Jha PK, Raychowdhury MK, Du Y, Leavis PC, Sarkar S. Isolation and characterization of human fast skeletal beta troponin T cDNA: comparative sequence analysis of isoforms and insight into the evolution of members of a multigene family. DNA Cell Biol 1994;13:217–33.
[43] Ervasti JM, Campbell KP. Membrane organization of the dystrophin–glycoprotein complex. Cell 1991;66:1121–31.
[44] Li L, Miano JM, Cserjesi P, Olson EN. SM22 alpha, a marker of adult smooth muscle, is expressed in multiple myogenic lineages during embryogenesis. Circ Res 1996;78:188–95.
[45] Zhao P, Iezzi S, Carver E, Dressman D, Gridley T, Sartorelli V, Hoffman EP. Slug is a novel downstream target of MyoD. Temporal profiling in muscle regeneration. J Biol Chem 2002;277: 30091–101.
[46] Yan Z, Choi S, Liu X, et al. Highly coordinated gene regulation in mouse skeletal muscle regeneration. J Biol Chem 2003;278:8826–36.
[47] Chen YW, Zhao P, Borup R, Hoffman EP. Expression profiling in the muscular dystrophies: identification of novel aspects of molecular pathophysiology. J Cell Biol 2000;151:1321–36.
[48] Haslett JN, Sanoudou D, Kho AT, et al. Gene expression comparison of biopsies from Duchenne muscular dystrophy (DMD) and normal skeletal muscle. Proc Natl Acad Sci USA 2002;99:15000–5.
[49] Noguchi S, Tsukahara T, Fujita M, et al. cDNA microarray analysis of individual Duchenne muscular dystrophy patients. Hum Mol Genet 2003;12:595–600.
[50] Tomczak KK, Marinescu VD, Ramoni MF, et al. Expression profiling and identification of novel genes involved in myogenic differentiation. Fed Am Soc Exp Biol J 2004;18:403–5.

Chapter 6

Intensity-based analysis of two-colour microarrays enables efficient and flexible hybridization designs

Peter A.C. 't Hoen, Rolf Turk, Judith M. Boer, Ellen Sterrenburg,
Renee X. de Menezes, Gert-Jan B. van Ommen and
Johan T. den Dunnen

Nucleic Acids Research 2004 32(4) e41



Published online February 24, 2004

Nucleic Acids Research, 2004, Vol. 32, No. 4 ***e41***
DOI: 10.1093/nar/gnh038

Intensity-based analysis of two-colour microarrays enables efficient and flexible hybridization designs

Peter A. C. 't Hoen[1,*], Rolf Turk[1], Judith M. Boer[1], Ellen Sterrenburg[1], Renée X. de Menezes[3], Gert-Jan B. van Ommen[1] and Johan T. den Dunnen[1,2]

[1]Center for Human and Clinical Genetics, [2]Leiden Genome Technology Center and [3]Department of Medical Statistics, Leiden University Medical Center, Wassenaarseweg 72, 2333 AL Leiden, The Netherlands

Received December 17, 2003; Revised January 30, 2004; Accepted January 31, 2004

ABSTRACT

In two-colour microarrays, the ratio of signal intensities of two co-hybridized samples is used as a relative measure of gene expression. Ratio-based analysis becomes complicated and inefficient in multi-class comparisons. We therefore investigated the validity of an intensity-based analysis procedure. To this end, two different cRNA targets were hybridized together, separately, with a common reference and in a self–self fashion on spotted 65mer oligonucleotide microarrays. We found that the signal intensity of the cRNA targets was not influenced by the presence of a target labelled in the opposite colour. This indicates that targets do not compete for binding sites on the array, which is essential for intensity-based analysis. It is demonstrated that, for good-quality arrays, the correlation of signal intensity measurements between the different hybridization designs is high (R > 0.9). Furthermore, ratio calculations from ratio- and intensity-based analyses correlated well (R > 0.8). Based on these results, we advocate the use of separate intensities rather than ratios in the analysis of two-colour long-oligonucleotide microarrays. Intensity-based analysis makes microarray experiments more efficient and more flexible: It allows for direct comparisons between all hybridized samples, while circumventing the need for a reference sample that occupies half of the hybridization capacity.

INTRODUCTION

DNA microarrays are widely used to measure genome-wide changes in mRNA expression levels across conditions such as developmental stages, disease states, drug treatment and gene disruption (1–5). Affymetrix GeneChips, prepared by photolithography, and spotted cDNA and 50–70mer oligonucleotide microarrays are currently the most frequently used platforms. The GeneChip is a one-colour system based on the immunofluorescent detection of biotinylated nucleic acids. The difference in perfect and mismatch probe intensities is used for gene expression measurements (6). Spotted microarrays are commonly hybridized with two samples labelled with two different fluorophores. For these arrays, the ratio of the signal intensities in the two channels is a relative measure of gene expression.

Normalization is essential to remove systematic biases in microarray data. For two-colour arrays, normalization algorithms can be applied to (log-transformed) ratios (7) (e.g. using a LOWESS algorithm). Alternatively, ANOVA models that account for array, dye and spot effects can be applied to the individual signal intensities on all the arrays (8). In both cases, after normalization, the ratio of the co-hybridized samples is usually calculated to minimize the influence of spatial variation in spot morphology and hybridization efficiency on the experimental outcome. Furthermore, some suggest that ratio-based analysis is important because of possible competitive hybridization of the two targets due to saturation of binding sites on the array (9).

Ratio-based analysis can be applied to experiments with a reference or loop design (10,11). A disadvantage of the reference design is that half of the acquired data represent only one sample that is often not biologically relevant, thereby doubling the number of arrays required (10,11). A loop design has other disadvantages (11). The calculated ratios have variable levels of precision since some samples are more directly related than others, and the set of hybridizations cannot be extended. This has important implications for studies in which not all samples become available at the same time; new samples could only be included in the experiment via forming new subloops, and only if biological material from the earlier samples is still available.

An intensity-based analysis in which the signal intensities in the two channels are kept separately, also after normalization, would allow for hybridization designs that are more efficient than the reference design and more flexible than the loop design. We designed a set of experiments to determine whether an intensity-based analysis would be justified for our spotted long-oligonucleotide microarrays. Our aims are 2-fold: first, to investigate whether hybridization patterns are sufficiently uniform across arrays; secondly, to verify if there is evidence for competition between targets for binding sites

*To whom correspondence should be addressed. Tel: +31 71 527 6611; Fax: +31 71 527 6075; Email: p.a.c.hoen@lumc.nl

e41 Nucleic Acids Research, 2004, Vol. 32, No. 4

on the array. We run two parallel statistical analyses, one ratio-based and the other one intensity-based, and compare their results.

MATERIALS AND METHODS

Microarray and target preparation

Murine oligonucleotide microarrays were produced in the Leiden Genome Technology Center by spotting the Sigma-Genosys mouse 7.5K oligonucleotide library (v. 1.0) (65mer, 20 µl in 50% DMSO) in duplicate on poly-L-lysine-coated slides (12). RNA was isolated from hind limb muscles of 20-week-old (sample A) and 8-week-old (sample B) C57Bl/10ScSn-DMD$^{mdx/J}$ (*mdx*) mice (R. Turk, E. Sterrenberg, E. de Meijer, G.J.B. van Oumen, J.T. den Dunnen and P.A.C. 't Hoen, in preparation). A reference RNA was created by pooling RNA from the following mouse tissues: muscle, testis, kidney, liver, brain, heart, spleen, ovary, uterus and whole embryos. The RNA was amplified by T7-polymerase-driven linear amplification and labelled through incorporation of aminoallyl-UTP and subsequent coupling to Cy3 or Cy5 monoreactive dyes, as described previously (12). Microarrays were hybridized with 1.5 µg of the indicated (Table 1) cRNA targets. Hybridization and washing was done in a GeneTAC hybstation (Genomic Solutions) (12).

Feature extraction and data analysis

Feature extraction was performed with GenePix 3.0 software (Axon Instruments Inc.). Spots with intensities lower than background or aberrant spot shape were flagged by the software and checked manually. Only spots that were not flagged on any of the analysed arrays were taken into account in further analyses, leaving 2224 data points per array. Local background-subtracted median signal intensities were used as intensity measures. Scaled gene expression ratios in samples A and B were calculated after transformation (natural logarithm) of the background-corrected intensities and subtraction of the average of the LN-transformed intensities (linear scaling).

Statistical analysis

SPSS (version 10.0.7) was used for standard statistical tests. To determine the significance level in comparisons of groups of Pearson correlation coefficients (r), the r values were first transformed using the formula

$$r(N-1)^{1/2}/(1-r^2)^{1/2}.$$

This was done to normalize the distribution of the correlation coefficients (13). The significance level used in all tests was 0.05.

Data analysis with MAANOVA (v2.0) was performed using Matlab (6.5R13, The MathWorks Inc). Using the MAANOVA package (www.jax.org/staff/churchill/labsite/software/anova/), we have fitted linear fixed-effects models, taking dye and average gene effects into account. The hyb-design was used as the test variable. The array effect could not be fitted because it is confounded with the hyb-design. Two models were compared: the null model where no effect due to hyb-design is assumed, and the alternative model in which an effect of hyb-design is expected. Hyb-design-specific p values were obtained via F-statistics for the individual features of each individual target (A and B) separately. F-statistics and corresponding p values are based upon the F2 statistic available in the MAANOVA package, which is a shrunk version of the classic F-statistic. To avoid distributional assumptions, the package offers the possibility of computing p values for hypothesis tests via permutation methods. We have chosen to perform the F-test with restricted residual shuffling and 1000 iterations. For more technical details about the model-fitting procedure, we refer to the package documentation.

This approach yielded a list of gene-specific p values relating to the hyb-design effect. In order to test if there was an overall hyb-design effect across all genes, we compared the proportion of computed p values of <0.05 with the expected one under no overall effect, 5%, via a conventional normal hypothesis test.

Table 1. Overview of used Hyb-designs and ratio calculations

Hyb Design	Array	Cy3	Cy5	Ratio
CoHyb	1	A	B	R1
	2	B	A	R2
ComRef	3	A	REF	R3
	4	REF	A	R4
	5	B	REF	R5
	6	REF	B	R6
OneColour	7	A	–	R7
	8	–	B	
	9	B	–	R8
	10	–	A	
SelfSelf	11	A	A	
	12	B	B	

R1–R8 are calculated from scaled LN-transformed background-subtracted intensities.
Average ratios are then calculated according to:
LN(RatioCoHyb): 0.5*[LN(R1) – LN(R2)]
LN(RatioComRef): 0.5*[LN(R5) – LN(R6)] – 0.5*[LN(R3) – LN(R4)]
LN(RatioOneColour): 0.5*[LN(R8) – LN(R7)].

RESULTS

We performed a set of experiments to assess the influence of co-hybridization on microarray gene expression measurements. Two different mouse RNA samples and a RNA reference were amplified separately in the presence of aminoallyl-UTP (12). Each cRNA target was independently labelled with Cy3 and Cy5. Aliquots were hybridized to murine 65mer oligonucleotide microarrays, according to the scheme in Table 1. This yielded the following hyb-designs: co-hybridization of sample A and B (CoHyb), hybridization of A and B against the common reference (ComRef), hybridization of A and B separately (OneColour) and self–self hybridization (SelfSelf). Dye swaps were performed for each hyb-design.

Analysis of raw background-corrected signal intensities for sample A and B suggests that the distribution of intensity measurements is not influenced by the hyb-design (Fig. 1). To confirm this, we fitted a linear mixed-effects model to the data using MAANOVA (8) and calculated p values corresponding to the hyb-design effect per feature. Since no more p values of <0.05 were found than would be expected by chance (2.2%

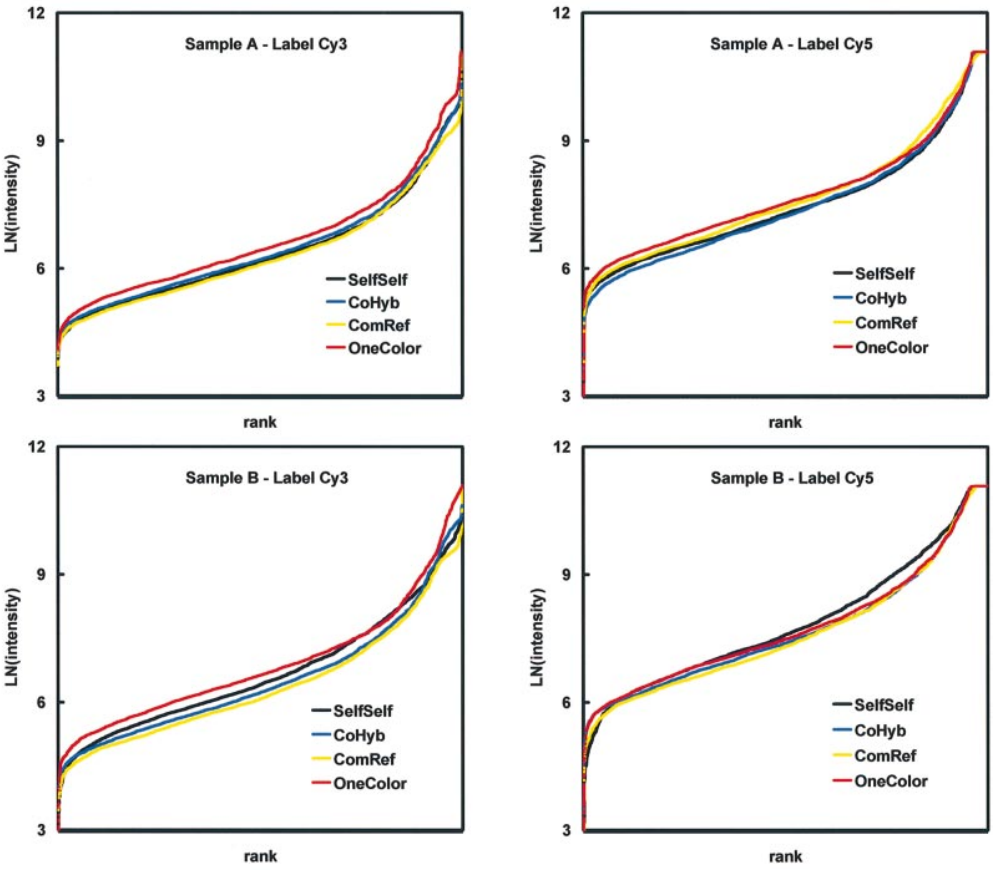


Figure 1. Effect of hybridization design on signal intensity distributions. The LN-transformed background-corrected Cy3 and Cy5 signal intensities from samples A and B, observed in the co-hybridizations (CoHyb, arrays 1 and 2, pink), hybridizations with the common reference (ComRef, arrays 3–6, yellow), one-colour hybridizations (OneColour, arrays 7–10, red) and self–self hybridizations (SelfSelf, arrays 11 and 12, blue), were ranked in rising order and plotted. The Cy5 signal levels off at high intensities due to scanner saturation.

and 5.3% for samples A and B, respectively), we concluded that the hyb-design did not significantly affect the measured signal intensity. We searched for indications of competition effects that would specifically affect measurements of highly abundant RNA species. To this end, we ranked the features according to average LN-transformed signal intensity on all arrays and made groups of 25 subsequent features. Then we performed a pairwise comparison of the signal intensities of the individual features in the different hyb-designs with Student's *t*-test. We found no statistically significant differences in intensity values between the one- and two-colour hybridizations in either of the signal intensity ranges. This indicates that, on spotted oligonucleotide microarrays, there is, even for high abundant RNA targets, no competition for binding sites. The presence of a co-hybridized target labelled in a different colour will, therefore, not affect signal intensities.

The intensity measurements on the arrays with different hyb-designs were highly consistent. To show this, the Pearson correlation coefficients of the separated Cy3 and Cy5 background-corrected intensities were calculated for each hybridization (Tables 2 and 3). Correlation coefficients for intensity measurements of a sample labelled in the same colour and hybridized to different arrays ranged between 0.88 and 0.98. We found that the correlations of signal intensities on one array (self–self hybridizations) were significantly stronger than the correlations of signal intensities from the same sample hybridized to different arrays ($p = 0.002$, Student's *t*-test). This confirms earlier observations that gene expression comparisons on the same array are more accurate than comparisons between different arrays (11). In addition, the signal intensities from targets labelled in the same colour are correlated significantly more strongly than the signal intensities from targets labelled in opposite colours ($p = 0.023$ and p

Table 2. Pearson correlation coefficients of raw Cy3 or Cy5 signal intensities from sample A, hybridized using different designs

CoHyb-3								
CoHyb-5	0.951							
ComRef-3	0.909	0.864						
ComRef-5	0.928	0.919	0.882					
OneColour-3	0.933	0.893	0.879	0.856				
OneColour-5	0.958	0.977	0.873	0.923	0.893			
SelfSelf-3	0.964	0.939	0.932	0.907	0.954	0.946		
SelfSelf-5	0.951	0.976	0.872	0.920	0.905	0.980	0.967	
Sample A	CoHyb-3	CoHyb-5	ComRef-3	ComRef-5	OneColour-3	OneColour-5	SelfSelf-3	SelfSelf-5

Pearson correlation coefficients are coloured according to strength: dark grey, $R > 0.95$; light grey, $0.90 < R < 0.95$; white, $R < 0.90$.

Table 3. Pearson correlation coefficients of raw Cy3 or Cy5 signal intensities from sample B, hybridized using different designs

CoHyb-3								
CoHyb-5	0.951							
ComRef-3	0.929	0.904						
ComRef-5	0.915	0.946	0.901					
OneColour-3	0.950	0.919	0.879	0.891				
OneColour-5	0.949	0.981	0.909	0.946	0.919			
SelfSelf-3	0.750	0.817	0.771	0.802	0.730	0.825		
SelfSelf-5	0.747	0.822	0.758	0.798	0.709	0.827	0.966	
Sample B	CoHyb-3	CoHyb-5	ComRef-3	ComRef-5	OneColour-3	OneColour-5	SelfSelf-3	SelfSelf-5

Pearson correlation coefficients are coloured according to strength: dark grey, $R > 0.95$; light grey, $0.90 < R < 0.95$; white, $R < 0.90$.

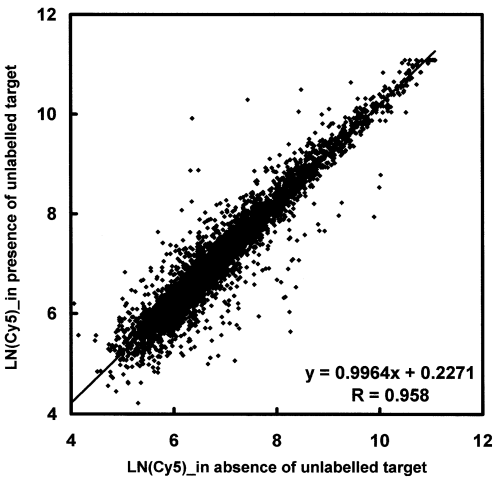


Figure 2. Effect of addition of unlabelled template on signal intensity measurements. Cy5-labelled cRNA was hybridized in the absence [*x*-axis, array 17 (table S1)] and presence [*y*-axis, array 18 (table S1)] of an equal amount of unlabelled cRNA. LN-transformed background-corrected signal intensities are plotted, together with the calculated regression line and the Pearson correlation coefficient.

= 0.037 for samples A and B, respectively, paired Wilcoxon test). This is indicative of a small dye effect and stresses the importance of dye swaps and balanced designs.

The experiment was repeated in a confined fashion with similar results (Supplementary tables S1 and S2 and fig. S1). In this second experiment, we also investigated the effect of co-hybridization with an identical unlabelled target. Again, we found a strong correlation of the intensity readings in hybridizations with and without the unlabelled target and only a very slight effect on absolute intensity values (Fig. 2).

To detect influences of the hyb-design on ratio calculations, we determined the individual gene expression ratios between sample A and B in three different ways: (i) averaging of the scaled ratios in the two co-hybridization experiments; (ii) calculating the ratios of the individual samples over the common reference and subsequently eliminating the common reference to obtain the ratio of A to B; (iii) calculating the ratios based on the intensity readings in the OneColour hybridizations (Table 1). We observed a high degree of correlation in the three types of ratio determinations ($R \geqslant 0.8$) (Fig. 3).

DISCUSSION

The main conclusion from our study is that co-hybridized targets do not influence each other's hybridization to spotted 65mer oligonucleotide microarrays. Even for highly abundant RNAs, we found no evidence for competition of the two targets for binding sites on the array. From an approximate calculation, we can see that indeed the number of binding sites generally exceeds the number of target molecules that can reach their complementary binding sites. According to the manufacturer, our pins deposit 0.7 nl of the 20 µM oligonucleotide solution onto the glass surface which amounts to 8×10^9 probe molecules, of which 20–50% is suggested to be available for binding (9). When we apply 2×1.5 µg of RNA and take into account that the cRNA is on average 1 kb in length, and that an abundant RNA species may comprise up to 1% of the RNA applied, there are ~5×10^{10} copies of this specific RNA molecule in solution. Owing to limiting diffusion distances of the molecules over the glass surface, only a fraction of these will reach the spot on the array. For PCR products spotted at a concentration of 0.5 µg/µl, the amount of binding sites is estimated to be only 10–20% of that

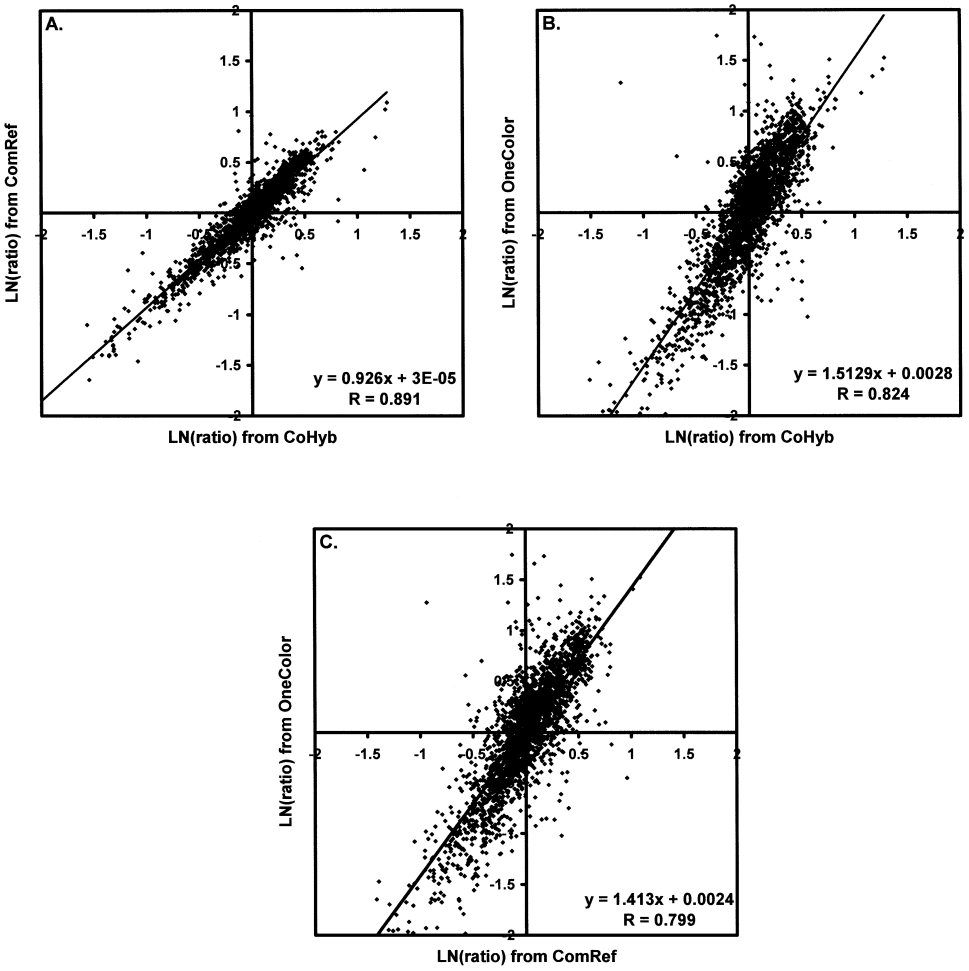


Figure 3. Effect of hybridization design on ratio calculation. After LN-transformation, background-corrected intensity values were linearly scaled by subtracting the mean LN-signal intensity in the separate channels of the arrays. Subsequently, gene expression ratios for samples A and B were calculated from the co-hybridized targets (CoHyb, arrays 1 and 2), targets hybridized with the common reference (ComRef, arrays 3–6) and the one-colour hybridizations (arrays 7–10), as described in the footnote to Table 1.

on the oligonucleotide arrays, indicating that competition may play a role, but only for higher abundant RNAs on cDNA arrays (14). Furthermore, the variation in concentration of spotted PCR products is larger than for oligonucleotide arrays, and therefore the influence of competition on the signal may be spot dependent, partly explaining the probe-concentration-dependent accuracy of cDNA arrays (15). The hybridization kinetics/thermodynamics on spotted microarrays are different from Affymetrix GeneChips, where saturation of binding sites is clearly observed and hybridization is adequately described by the Langmuir adsorption formula (16–18). This is probably due to the higher amount of target molecules applied (~15 µg cRNA), the substantially lower amount of binding sites [~10^7 per feature (19)] and the extensive mixing.

The independence of the measurements in the two channels of the microarray indicates that signal intensities from the two channels can be taken into subsequent analyses separately. After simultaneous normalization of the separated intensities on all arrays in the experiment, expression ratios can be calculated across all experimental groups. This is supported by our observation that the resulting gene expression ratios of sample A to sample B are independent of the hyb-design and the associated method of calculation. Obviously, this procedure critically depends on the uniformity of the arrays; spatial

hybridization effects should be absent and the spotting should be highly reproducible between different arrays. Judged from the high correlation of signal intensity measurements across different arrays, this uniform hybridization seems achievable.

In multi-class experiments, it is difficult to decide which samples should be co-hybridized. Since this and other studies show that differential expression measurements from co-hybridized samples are more accurate than from samples on different arrays, it has been suggested that the comparisons in which the experimenter is most interested should be hybridized to the same array (11,20). Co-hybridization effects may be included in mixed ANOVA models (21). Since this will be at the expense of a large reduction in the degrees of freedom in the experiment, it is best to avoid a co-hybridization bias. This can be achieved by co-hybridizing each group as many times to one group as to all the other groups (blocking of the effect), but this might not be possible when there are many different experimental groups. We suggest, therefore, that samples are randomly assigned from the different experimental groups to the arrays. This attributes an extra level of randomization in microarray designs above those suggested by Churchill (20), i.e. randomization of treatment and sampling, dye assignments, slides in a batch and spots on the array. To eliminate the dye effect that we and others (12,22,23) have demonstrated, it is essential to keep the design balanced with respect to dye, i.e. to label samples from a certain group as many times in one colour as in the other.

In summary, we show that the intensity measurements on spotted oligonucleotide arrays are not affected by co-hybridized targets and demonstrate the validity of separate analysis of the signal intensity readings from the different channels of the array. An intensity-based type of analysis has considerable advantages over ratio-based analysis in experiments that include many different groups. First, it enables comparisons between samples which were not hybridized to the same array without the need to relate the samples to a common reference or other samples in the same series. Secondly, the set of hybridizations is extendable, as long as temporal effects in array and target preparation and hybridization remain negligible. In future, multi-colour hybridizations (15), applying more than two targets on the same array, may become more general. For these multi-colour experiments, an intensity-based analysis would be even more favourable than for two-colour arrays, since a ratio-type of analysis will be complicated by the possible calculation of many different ratios from one array.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

REFERENCES

1. Van 't Veer,L.J., Dai,H., van de Vijver,M.J., He,Y.D., Hart,A.A., Mao,M., Peterse,H.L., van der,K.K., Marton,M.J., Witteveen,A.T. *et al.* (2002) Gene expression profiling predicts clinical outcome of breast cancer. *Nature*, **415**, 530–536.
2. Arbeitman,M.N., Furlong,E.E., Imam,F., Johnson,E., Null,B.H., Baker,B.S., Krasnow,M.A., Scott,M.P., Davis,R.W. and White,K.P. (2002) Gene expression during the life cycle of *Drosophila melanogaster*. *Science*, **297**, 2270–2275.
3. Boer,J.M., de Meijer,E.J., Mank,E.M., van Ommen,G.B. and den Dunnen,J.T. (2002) Expression profiling in stably regenerating skeletal muscle of dystrophin-deficient mdx mice. *Neuromusc. Disord.*, **12**, S118–S124.
4. Giaever,G., Chu,A.M., Ni,L., Connelly,C., Riles,L., Veronneau,S., Dow,S., Lucau-Danila,A., Anderson,K., Andre,B. *et al.* (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature*, **418**, 387–391.
5. Ross,D.T., Scherf,U., Eisen,M.B., Perou,C.M., Rees,C., Spellman,P., Iyer,V., Jeffrey,S.S., Van de,R.M., Waltham,M. *et al.* (2000) Systematic variation in gene expression patterns in human cancer cell lines. *Nature Genet.*, **24**, 227–235.
6. Lockhart,D.J., Dong,H., Byrne,M.C., Follettie,M.T., Gallo,M.V., Chee,M.S., Mittmann,M., Wang,C., Kobayashi,M., Horton,H. *et al.* (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol.*, **14**, 1675–1680.
7. Quackenbush,J. (2002) Microarray data normalization and transformation. *Nature Genet.*, **32** (Suppl.), 496–501.
8. Kerr,M.K., Martin,M. and Churchill,G.A. (2000) Analysis of variance for gene expression microarray data. *J. Comput. Biol.*, **7**, 819–837.
9. Wang,Y., Wang,X., Guo,S.W. and Ghosh,S. (2002) Conditions to ensure competitive hybridization in two-color microarray: a theoretical and experimental analysis. *Biotechniques*, **32**, 1342–1346.
10. Dobbin,K., Shih,J.H. and Simon,R. (2003) Statistical design of reverse dye microarrays. *Bioinformatics*, **19**, 803–810.
11. Yang,Y.H. and Speed,T. (2002) Design issues for cDNA microarray experiments. *Nature Rev. Genet.*, **3**, 579–588.
12. 't Hoen,P.A., de Kort,F., Van Ommen,G.J. and den Dunnen,J.T. (2003) Fluorescent labelling of cRNA for microarray applications. *Nucleic Acids Res.*, **31**, e20.
13. Cox,D.R. and Hinckley,D.V. (1974) *Theoretical Statistics.* Chapman & Hall, London, UK.
14. Stillman,B.A. and Tonkinson,J.L. (2001) Expression microarray hybridization kinetics depend on length of the immobilized DNA but are independent of immobilization substrate. *Anal. Biochem.*, **295**, 149–157.
15. Hessner,M.J., Wang,X., Khan,S., Meyer,L., Schlicht,M., Tackes,J., Datta,M.W., Jacob,H.J. and Ghosh,S. (2003) Use of a three-color cDNA microarray platform to measure and control support-bound probe for improved data quality and reproducibility. *Nucleic Acids Res.*, **31**, e60.
16. Hekstra,D., Taussig,A.R., Magnasco,M. and Naef,F. (2003) Absolute mRNA concentrations from sequence-specific calibration of oligonucleotide arrays. *Nucleic Acids Res.*, **31**, 1962–1968.
17. Held,G.A., Grinstein,G. and Tu,Y. (2003) Modeling of DNA microarray data by using physical properties of hybridization. *Proc. Natl Acad. Sci. USA*, **100**, 7575–7580.
18. Chudin,E., Walker,R., Kosaka,A., Wu,S.X., Rabert,D., Chang,T.K. and Kreder,D.E. (2002) Assessment of the relationship between signal intensities and transcript concentration for Affymetrix GeneChip arrays. *Genome Biol.*, **3**, RESEARCH0005.
19. Lockhart,D.J. and Winzeler,E.A. (2000) Genomics, gene expression and DNA arrays. *Nature*, **405**, 827–836.
20. Churchill,G.A. (2002) Fundamentals of experimental design for cDNA microarrays. *Nature Genet.*, **32** (Suppl.), 490–495.
21. Cui,X. and Churchill,G.A. (2003) Statistical tests for differential expression in cDNA microarray experiments. *Genome Biol.*, **4**, 210.
22. Tseng,G.C., Oh,M.K., Rohlin,L., Liao,J.C. and Wong,W.H. (2001) Issues in cDNA microarray analysis: quality filtering, channel normalization, models of variations and assessment of gene effects. *Nucleic Acids Res.*, **29**, 2549–2557.
23. Yang,Y.H., Dudoit,S., Luu,P., Lin,D.M., Peng,V., Ngai,J. and Speed,T.P. (2002) Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.*, **30**, e15.

Chapter 7

Discussion

Table of Contents

Discussion

This thesis describes the application of gene expression profiling to study the molecular mechanisms of the pathology of muscular dystrophies. In this section, the experimental outcome of the research will be discussed in a broader perspective.

Gene expression profiling has allowed the identification of genes and subsequent processes or pathways in which these genes play a role. The primary goals of the research leading to this thesis were the determination of molecular mechanisms in the pathology of the mouse model for Duchenne Muscular Dystrophy (Chapter 3), and the comparison of gene expression profiles of different mouse models for muscular dystrophy (Chapter 4). In the process of pursuing these goals, a number of other experiments had to be performed to allow the successful achievement of the primary goals. These additional experiments included the determination of differential gene expression as a result of variations in genetic background between different mouse inbred strains (Chapter 2), and the application of an alternative experimental design and data analysis method in gene expression profiling experiments with a large number of variables (Chapter 5). Furthermore, we were interested in the molecular details of myogenesis (Chapter 6) and its relation to regeneration.

Although gene expression profiling certainly has proved its potential, a number of limitations were apparent during its application. This section will start with discussing the advantages and limitations of gene expression profiling in the study of muscular dystrophy, and will be followed by discussing the experimental outcome of gene expression profiling studies of muscular dystrophies.

1 Gene Expression profiling

1.1 Challenges of gene expression profiling

Before the introduction of high throughput techniques in molecular biology, gene expression levels were mainly determined using Northern blots. The statistical analysis of Northern blot experiments could suffice with basic statistical methods. Changes in gene expression were compared with the expression of a single or sometimes two 'housekeeping' genes, which expression was thought to be stable between the compared samples. The major disadvantage of the Northern blot was the incompetence of determining the gene expression levels of a large number of genes without performing a large number of experiments.

The application of high throughput techniques such as serial analysis of gene expression (SAGE, see Introduction Chapter 2.1) and microarrays put an end to the limitation of analyzing the gene expression levels for only a few genes in a single experiment[67,57,10]. As a result, the high-throughput approach allows an experimental design based on the discovery of expression levels of genes not yet associated with the biological sample under investigation, instead of the hypothesis-driven 'candidate-gene' approach, where gene expression levels are determined using Northern blots. Furthermore, differential gene expression levels could be determined based on the expression of a large number of genes, instead of only one or two 'housekeeping' genes. As a result, the possibility to determine differential gene expression was increased. The determination of differential gene expression, however, demanded a far more complex data analysis with accompanying statistical applications. This chapter will discuss the challenges of determining differential gene expression using high throughput gene expression profiling with microarrays.

1.1.1 Image acquisition

The main microarray platform used in our studies consist of a microscope slide containing spotted oligonucleotide sequences with a length of 65 nucleotides. A population of mRNA molecules is reverse transcribed, amplified using in vitro transcription, and labeled with fluorescent dye, before being hybridized to the microarray. The hybridized, fluorescently labeled molecules are detected using a laser-scanner, which scans the entire microarray. When the laser beam excites a fluorescent dye, a photon is released, which is detected using a photo multiplier tube (PTM) and a sensitive CCD camera. As a result, scanning of the microarray results in the localization and determination of the intensity of the fluorescent dyes on the microarray. By combining the localization and identity of the spotted oligonucleotide sequences and the intensity of the fluorescent signal, a relative determination can be made of the amount of fluorescent molecules present. Currently, laser-scanners are able to detect up to four fluorescent labels by using different wavelengths. This allows the hybridization of four different samples on each microarray. The simultaneous application of four dyes enables a high throughput and minimizes the costs[61]. In a dual color experiment, a direct comparison can be made regarding the amount of fluorescent molecules in two biological samples by determining the two different fluorescent labels.

Scanning of a microarray renders a digital image displaying hybridized fluorescently labeled molecules to localized oligonucleotide sequences. Two-color microarray experiments render two digital images, one for each fluorescent dye. The digital image consists of pixels with certain intensity. The mainstream scanners have a 16-bit CCD camera, which has an intensity range of 0 to 65535 (equals 2^{16} possibilities). Software has been developed to determine the exact outline of the spotted oligonucleotides based on foreground/background differences within the spot. Furthermore, the software is capable of distinguishing signals derived from dust, scratches, and other anomalies, based on recognition of deviating signal patterning. A typical spot on our platform, also known as feature, harbors around 200-300 pixels. The software determines the mean or median intensity of the foreground pixels (inside of the spot) and the background pixels (outside of the spot). Furthermore, the software links the location of the spot to the information of the spotted oligonucleotide sequence. Hence, in a typical dual color microarray experiment the feature extraction software determines a number of parameters. Thus, the sequence information is linked with the background-corrected intensity of each color for each spot.

A spotted microarray can contain several thousand features. The feature extraction software is designed to facilitate the automatic detection of the spot on the microarray. However, the spot finding algorithm is not flawless. Therefore, the outline of the spots needs to be checked manually. This remains laborious, although the bulk of the features are correctly found. The finding algorithm is based on the assumption that the intensity of the pixels (signal) within a spot is higher compared to the intensity of the pixels outside a spot. A spot is flagged when this assumption cannot be fulfilled. Flagging occurs when there is no signal within the spot, because no fluorescently labeled mRNA has been hybridized. Flagging also occurs when the background intensity is as high as the foreground intensity, or when dust, speckles, or scratches alter the shape of the spot. Since the flagging algorithm does not diverge between the origins of the flagging, it is difficult to discriminate spots with no signal and spots with high background. We decided to leave out all spots that were flagged. As a consequence, the number of false positives was kept low at the cost of increasing the number of false negatives.

The quality of the microarray is also determined by the overall ratio between the signal intensity and its variation within each spot and the signal intensity and its variation of the associated background. This ratio is also known as the signal to noise ratio. The signal to noise ratio of the oligonucleotide platforms was higher than the signal to noise ratio of the cDNA microarrays as were used in Chapter 6. We found that the signal to noise ratio within a single microarray needs to be evenly dispersed to obtain reliable results (data not shown). In other words, the signal to noise ratio between different parts of a single microarray needs to be low. As a result, microarray experiments needed to be replicated, when this was not the case.

1.1.2 Normalization

Normalization is employed to correct for biases in the analyzed dataset, which are unrelated to the intrinsic biological differences between the individual samples (reviewed in[53]). These biases are mainly caused by the use of different fluorescent dyes, and differences in signal intensity distributions within a single microarray and between different microarrays.

In a two-color microarray experiment, two different fluorescent dyes are used to compare the gene expression patterns of two biological samples. The standard dyes currently used, are Cy3 and Cy5. These two dyes differ chemically, which results in a different behavior during direct labeling due to steric hindrance. This problem can largely be resolved by applying indirect labeling methods such as aminoallyl-labeling. Furthermore, the excitation efficiency of the individual dyes is shaped according to a normal distribution. Cy3 and Cy5 have overlapping excitation distributions, which causes excitation interference. The excitation levels, which depend on the strength of the laser, of the two fluorescent dyes in our study also differ in their quantum yield. In other words, the signal intensity of one fluorescent dye is significantly higher compared to the other. Furthermore, the average background levels between the two dyes differ, because the autofluorescence level of Cy3, which has a lower excitation wavelength, is higher than the autofluorescence level of Cy5. The biases caused by the different dyes result in a skewed intensity output. The application of dye-swap experiments, where the dyes are switched between the two samples on the microarray, results in a balanced design and reduces the extent of normalization required (reviewed in[74]).

Normalization is based on the transformation of the numerical data output after feature extraction. The most straightforward method for normalization is log transformation. However, this method increases the variance in the lower regions of the signal intensity. To compensate this increase in variance, a different algorithm should be used. An optimal strategy of normalization would normalize the lower intensities using a linear factor, whereas the higher intensities are normalized using a logarithmic factor. This algorithm is best approached by application of an arcsin(h)-based algorithm. Using this normalization method, the variance is stabilized over the entire range of signal intensities. This method, developed by Wolfgang Huber, is called Variance Stabilization and Normalization (VSN)[30]. In other words, the advantage of using this method is that the normalized dataset is less affected by increasing levels of variance on both sides of the intensity range, which occurs with either logarithmic or linear transformation. As a consequence, the normalized dataset will be more reliable, and will facilitate the finding of differentially expressed genes in both the lower and higher signal intensities.

1.1.3 Experimental design

Gene expression profiling can be applied to determine differences in gene expression levels between two biological samples using a single microarray. However, this strategy is not advisable, because it is not possible to determine biological variation of individual samples, or variation based on technical grounds.

The experimental design should be based on the biological questions and the amount of biological and technical variation[74]. In other words, obtaining the most reliable answer requires a proper adaptation of the experimental design. Furthermore, the limited availability of template material of biological samples, as well as the cost of microarray experiments, necessitates an efficient design of microarray experiments. A proper experimental design should minimize the effect of technical variation on gene expression levels. One example is the elimination of fluorescent dye effects by balancing these in a dye-swap experiment. Technical variation due to the use of multiple microarrays can be determined more accurately by increasing the number

of technical replicate experiments. The increase of replicate experiments can also be applied to better determine the biological variation. The difference, however, lies in the elimination of the variation. Replicate experiments restrict the influence of technical variation, so that the biological variation can be estimated more accurately. Increasing the number of biological replicates decreases the amount of false positives and false negatives.

The complexity of experimental design makes the analysis of ratio-based data more tedious and complex as the number of microarrays increase. This is explained more easily with a number of examples. In microarray experiment X, the gene expression levels of sample A (A') are compared with those of sample B (B'), resulting in the ratio A'/B'. To create a balanced experimental design, the fluorescent dyes are swapped (resulting in the gene expression levels A'' and B''), and again a ratio of A''/B'' can be calculated. If the gene expression levels of three samples are determined (K', L', M'), The following six (dye-swap) experiments need to be performed to determine the differential gene expression levels between the different samples: K'/L', K''/L'', K'/M', K''/M'', L'/M', and L''/M''. In the case of four variables, 12 dye-swap experiments need to be performed to determine all the possible ratios. With five variables, 20 microarray experiments are necessary, etc. The number of microarray experiments (y) can be calculated according to the following formula for pronic numbers, where n is the number of variables.

$$y = n(n-1)$$

In short, the number of microarray experiments will increase substantially with an increasing number of variables. Fortunately, other ratio-based experimental designs have been proposed. One of the alternative experimental designs is based on the use of a common reference sample, where each biological sample is hybridized to a common sample. As common reference, a number of possibilities exist; one can use the first sample in a timecourse experiment, or one can use a mixture of different samples. An alternative approach for using a common reference was developed in our lab[62]. As common reference for the cDNA microarrays, a pool of the spotted cDNA sequences was fluorescently labeled and hybridized with each biological sample. This method guarantees the calculation of a proper ratio, since a signal is always present in the common reference sample. This solution addresses an important problem in ratio-based calculations. Ratios are determined by two numbers. If one of the numbers is low or missing due to a very low or absent signal, unreliable ratios will arise. Our method eliminates this problem. Nonetheless, a major drawback of using a common reference is always the loss of half of the capacity of the microarray.

Since conventional experimental design requires a large number of microarray experiments and the use of a common reference results in the loss of half of the microarray capacity, we wondered if we could treat the signals from the two fluorescent dyes of the microarray independently. In other words, we wondered if we could use intensities instead of ratios, and thereby circumvent tedious calculations and using of a large number of microarrays (Chapter 5). The major question of this unconventional approach concerns the interaction of the two samples on the binding to specific sequences on the microarray. Hybridization is a reversible reaction, which results in an equilibrium. Since the target sequences (mRNA) of both samples can bind

to the same probe (oligonucleotide), a competition of the target sequences may take place for the available binding sites of the probe. The acquired signals, and subsequent ratio, from both samples therefore depends on the mutual target presence and the binding capacity of the probe. However, our study successfully demonstrated that the individual signal intensities of both samples do not influence each other. As a result, the signal intensities can be treated separately. This greatly facilitates the use of complex experimental designs. Indeed, by this method, it is not longer required to perform all possible combinations of the different variables directly. The only requirement is the use of a balanced design to eliminate dye-effects. In other words, for each biological sample the same number of experiments needs to be performed for both fluorescent dyes. Furthermore, the samples can be randomly distributed in the experimental design. Another advantage is that additional experiments can be added to the experimental design, which makes this method much more flexible than more conventional experimental designs.

In the temporal experiment described in this thesis (Chapter 3), we have applied a so-called loop design. This design has a number of advantages, compared to more conventional experimental designs. The number of microarrays is optimally used; the number is as low as possible without loss of experimental data by excluding the use of a reference sample (36 versus 72 microarrays, respectively). The experimental design remains balanced as well. The use of loop-designs in temporal studies provides the possibility to co-hybridize successive timepoints. This experimental design results in a higher precision between two consecutive timepoints than other strategies, because the differences in gene expression levels are usually smaller between consecutive timepoints than between non-consecutive timepoints[74].

1.1.4 Data analysis

Gene expression profiling using microarrays generates a large number of datapoints. For instance, the experiment described in Chapter 3 resulted in more than two million datapoints. This number of data points requires the use of relevant statistical methods to acquire a degree of reliability concerning the significance of the outcome. Statistical methods had already been developed for the handling of large amounts of data, and were basically applicable to microarray analysis. However, the existing algorithms had to be adapted to the variables used in microarray experiments.

To avoid high numbers of false positives in the comparison of gene expression levels in two or more biological samples, rigorous statistics are required. Statistical significance values (p-values) can be calculated using a number of statistical tests. Statistical significance between two biological variables can be calculated using the Student's t-test. Statistical significance between more than two biological variables can be calculated using an analysis of variance (ANOVA) approach. Both statistical analyses are based on the normal distribution of datasets. When the data is not distributed normally, other statistical tests, such as Wilcoxon or permutation tests can be applied to determine significance levels.

The p-values in a microarray experiment need to be corrected for multiple testing. Statistical methods are applied to determine whether a difference in the gene expression levels of any given gene can be attributed by chance alone or not. By increasing the number of tests

(genes), the chance that a significant outcome in one of the tests will appear, increases as well. Therefore, a correction for multiple testing is used. The most straightforward method is the application of a family-wise error rate (FWER) control method such as the Bonferroni correction for multiple testing, which divides the minimal significance level by the number of measurements. This correction method is extremely stringent, and may result in a high amount of false negatives. Another method is based on adjusting the tolerance of false positives, namely the false discovery rate (FDR)[3]. FDR is based on the distribution of significance levels for each individual experiment. The two different methods were compared in Chapter 2. Since both methods can be used in any microarray experiment, it depends on the biological question, which method is used. If one wants to control the number of false positives with the risk of increasing the number of false negatives, the FDR method is advised. In general, the degree in the tolerance of false positives depends on the effort it will cost to validate differentially expressed genes. In other words, if the experiment allows the validation of a large number of genes, the tolerance of false positives can be large. When the experiment allows only a small number of genes to be validated, the tolerance of false positives needs to be kept at a minimum.

Although the existing methods for calculating statistical significance suffices, more sophisticated methods are needed for complex experimental designs. Significance levels in the temporal gene expression profiling study in this thesis (Chapter 3) have only been calculated for each individual timepoint. Integration of both the experimental (loop) design and the time-course experiment into a single statistical test is ongoing, and might reveal more genes with significant differentially expressed genes[68].

1.1.5 Gene ontology

The aim of gene expression profiling is to determine the transcriptome of the biological samples under investigation. Comparing the transcriptome of two (or more) biological samples will provide genes that are differentially expressed. To understand why certain genes are differentially expressed, functional information about the protein products of those genes has to be obtained by performing literature studies, or querying biological databases. Information databases, as provided by the National Center for Biotechnology Information (NCBI), part of the National Institute of Health (NIH), are available to combine the biological information at different levels. The functional biological information is based on protein and genetic sequences and subsequent motifs, experimental data (for instance, where and when certain genes or proteins are expressed), and comparisons between different organisms.

A collaborative effort has been established by the Gene Ontology Consortium (www.geneontology.org) to provide a consistent description of gene products between different databases and between different species. The function of each gene is described in three structured, hierarchical vocabularies (ontologies). The description depends on the current understanding of the functional, biological and molecular role of the protein product of each gene. As with the genetic information, the gene ontology description for each protein is far from complete. However, conclusions are drawn based on this incomplete information. Therefore, one must always keep in mind that the functional annotation is prone to change. The analysis of microarray experiments based on gene ontology might thus change in time. Furthermore, the under-

standing of differentially expressed genes is dependent on the interpretation of the scientist.

1.1.5.1 Pathway analysis

Tools and information databases to functionally annotate gene expression data have recently become more accessible. They were developed as a result of an increasing demand due to the development of high-throughput techniques. Protein-protein interactions, which are based on experimental data, can be described by placing them in a functional pathway. Pathways are mostly constructed based on the biological process the proteins are active in, for instance metabolic pathways, signal transduction pathways, translational pathways etc. The construction of pathways highly facilitates the understanding of biological processes. With a continuously increasing knowledge of protein-protein interactions, the number of pathways is increasing. One of the earliest pathways were constructed and provided by the Kyoto Encyclopaedia of Genes and Genomes (KEGG) (www.genome.ad.jp/kegg). Still, only a minority of the annotated genes is placed into functional pathways, which is caused by a lack of experimental data or because the pathway has not been constructed yet.

A typical gene expression experiment can result in hundreds of differentially expressed genes. Pathway analysis of the experimental data is laborious, since the majority of the genes cannot be placed in functional pathways. As a result, the pathways have to be constructed subsequently. The construction of pathways is mainly based on published literature. Although literature studies are essential to derive pathway information, the amount of information increases enormously. Recently, tools have been developed that can perform literature studies by determining the relationships of search objects computationally[40,46,66]. However, this technique is still under development[59], and will not be discussed in detail.

1.1.5.2 Pitfalls

Gene ontology and the functional annotation of genes are closely connected to the DNA sequence of each gene. Proper management of sequence data is therefore detrimental to the functional interpretation of microarray experiments.

Nucleotide sequences of different organisms are stored in various databases, of which a number are publicly available. One of those public databases is maintained by the NCBI. The nucleotide sequences at the NCBI are identified by a unique GenBank Accession Number. Upon alignment of multiple GenBank accession numbers, ultimately the reference sequence of the mRNA molecule is determined. By alignment of the GenBank accession numbers, a biologically non-redundant molecule is sought to obtain. Before this completion, the accession numbers are grouped in specific UniGene clusters, where each cluster represents a single gene. At this stage, the individual accession numbers can cross-link different UniGene clusters, since the alignment is not comprehensive. Insufficient overlap of sequences within a single cluster can be one of the causes leading to the moving of GenBank accession numbers. Furthermore, intrinsic properties of individual genes, like alternative splicing, can make proper alignment difficult.

The UniGene database was founded approximately 10 years ago to cluster the highly redundant Expressed Sequence Tags (ESTs). Specific study of the DNA sequence can give indications about the function of the potential gene in each UniGene cluster based on the found sequences and motifs within the compiled sequence. With the completion of the sequencing of organismal genomes, the reference sequence of a growing number of genes is determined. Comparison of the genomic and cDNA sequences between organisms can allow one to identify homologous genes. Ultimately, the mRNA sequences of all genes, including their alternatively spliced products will become complete. Until this point, the databases provide, although incomplete, essential information about genes and their function.

A growing number of the probes used for the generation of microarrays are based on the final reference sequence of each mRNA molecule. However, a (large) number of the probes are based on GenBank accession numbers. One must keep in mind that the information given for a gene is based on the UniGene cluster. Therefore, the gene information of the probes should be checked regularly, because the GenBank accession numbers can switch from cluster to cluster.

1.2 Utilization of gene expression profiling

The possibility to answer biological questions using gene expression profiling depends on the availability of biological samples. There is a gradient of availability of fresh tissue, animal models, and cell cultures from an array of diseases, such as cancer, autoimmune diseases, brain diseases, etc. For the study of muscular dystrophies, human muscle biopsies, animal models, and muscle cell cultures are available, which can lead to insights into the pathogenesis of the disease.

Although muscular dystrophy does not only affect muscle tissue, the primary symptoms leading to the disease are progressive muscle weakness and wasting. To address the primary symptoms, we decided to focus primarily on skeletal muscle tissue. The clinical presentation of muscular dystrophies demonstrates significant variation between the various muscular dystrophies. This variation mainly concerns spatial (muscle groups) and temporal variation, and variation in severity. Furthermore, the variation between patients having a similar type of muscular dystrophy is significant.

We have employed two different model systems to circumvent the high variation displayed in human patients and to facilitate the application of gene expression profiling to determine the molecular mechanisms in muscular dystrophy. The first model system is based on the use of animal models for muscular dystrophy; the second is based on the use of in vitro primary cell cultures of human myoblasts. Both systems will be discussed below.

1.2.1 Animal models for muscular dystrophy

One of the major drawbacks of gene expression profiling in human muscular dystrophy patients, is the necessity of muscle biopsies. With a single biopsy, temporal processes, which are critical for the course of the disease, cannot be analyzed. Furthermore, obtaining a controlled

group of biopsies (i.e. similar age, or stage of the disease) is difficult, because a large group of patients is often not available.

Ethical issues based on the burden of the patient will rise with taking multiple biopsies in time. Furthermore, the majority of patients are diagnosed after the manifestation of the disease, making the initiation of the disease difficult to study. Other difficulties in obtaining biopsies are caused by the lack of presymptomatic tissue, the low number of myofibers in late stage muscle biopsies, and the fact that some muscular dystrophies have a very low incidence.

Determination of the pathological processes by gene expression profiling demands a controlled population of muscle biopsies with similar states of pathology, which need to be confirmed based on clinical examination and muscle histology. Furthermore, a significant amount of muscle biopsies is necessary to exclude inter-individual variability. As a result, the number of microarray experiments will increase significantly.

To obtain insight in the in vivo processes leading to muscular dystrophy, we have chosen to use animal models for muscular dystrophy. The use of animal models circumvents the issues concerning the obtaining of human muscle biopsies, and enables the collection of specimens at specific stages of the disease. Although a number of animal models for muscular dystrophy are available (see Introduction Chapter 1.3.4), we have chosen the mouse as animal model. The mouse has proven its value as a mammalian model for human disease in general over the last decades. Since mice have been used to generate a large number of models for muscular dystrophy (reviewed by Durbeej et al.[19]), this also facilitates the comparison of the different muscular dystrophies. For our in vivo studies, we have used hindlimb muscle tissue (quadriceps femoris), because this muscle group is affected by the disease, is relatively easy to dissect and renders a sufficient amount of tissue, especially in young mice.

1.2.2 Primary muscle cell culture

Muscular dystrophies can be studied by the use of in vitro model systems, like primary muscle cell cultures. This model system has a number of advantages. Primary muscle cell cultures are derived from undifferentiated satellite cells present in muscle biopsies. These cells can be held in a proliferative state by growing them in a medium rich in Fetal Calf Serum. Withdrawal of the serum (lowering the percentage from 20 to 2%) results in the differentiation of these satellite cells to myotubes, and then to myofibers. As a result, the molecular and cellular processes active in these cells can be controlled and studied at will.

1.2.3 Muscle biopsies versus cell cultures

Muscle is composed of a large variety of cell types, like myofibers, immune-related cells, nerve-related cells, and vascular structure-related cells. Gene expression profiling of muscle biopsies reflects the expression of genes of all cell types present. Therefore, it is not possible to directly determine the origin of differentially expressed genes. Functional annotation may provide information concerning the origin of expressed genes, when those genes are cell-type specific. However, the functional annotation is not yet complete (see Discussion Chapter

1.1.5.2). Furthermore, a large number of genes are ubiquitously expressed due to the nature of their function (i.e. cell maintenance, replication, transcription etc.).

Gene expression profiling of a mixture of cell types can be circumvented by the application of other techniques such as laser-capture microdissection (LCM). However, in our studies, which were done in a very early phase of gene expression profiling, we have not been able to apply this technique. LCM has a large potential, because nowadays it is possible to generate enough template material for gene expression profiling by performing multiple rounds of mRNA amplification[49]. The application of this technique makes it possible to isolate myofibers of specific stages in vivo. In contrast to using biopsies, this technique represents a cellular approach similar to cell cultures, but with the advantage of the in vivo environment. As a result, the outcome of gene expression profiling experiments will not produce 'gene expression noise' from other cell types, like neutrophils, macrophages, or vascular or neuronal cell types.

2 Processes in muscular dystrophy

The pathology of the different muscular dystrophies shares a high level of similarity and can be described by a cascade of processes. The disease is initiated by myofiber degeneration, which is followed by inflammation, and regeneration. The molecular details of these processes are not fully understood. Therefore, we have applied gene expression profiling to determine the genes which play a significant role in these processes. We have determined temporal gene expression levels in the mdx mouse model for Duchenne Muscular Dystrophy (Chapter 3). Furthermore, we have compared gene expression levels between different mouse models for muscular dystrophy (Chapter 4).

Although the mdx mouse model seems to recover from the pathology, in contrast to the lethal human disease, a number of the characteristic dystrophic processes take place. The mdx mouse suffers from considerable muscle degeneration in an early stage of life, which is followed by inflammatory and regenerative processes. As the mdx mouse reaches adulthood, regeneration supersedes degeneration which leads to a stable condition. In order to study these processes in more detail, we have determined gene expression levels of hindlimb muscles from both mdx and wildtype mice ranging from 1 to 20 weeks of age (see Chapter 3).

The various mouse models for muscular dystrophy differ in the age of onset, the severity of the disease, and which muscle groups are affected. However, the affected muscles undergo the same pathological cellular processes. To date, the reason behind the typical clinical pathology of the different muscular dystrophies is unknown. We set out to study the different muscular dystrophies on gene expression level to find common and distinct pathological processes in order to delineate the reasons of the divergence. Although the severity varies per muscle group and varies over time, we started with a single muscle group at a single timepoint. We determined the gene expression profiles of hindlimb muscle tissue of 8-week-old mouse models for muscular dystrophy (Chapter 4). However, by using a single timepoint a bias might be generated concerning the severity of the mouse models, i.e. the mouse models for dysferlinopathy develop a severe phenotype from 4-6 months of age. Therefore, a temporal study will be more efficient to circumvent these variables. Currently, these timecourse experiments are performed in our lab following the initial single-timepoint study.

Based on the two completed studies described above, we studied the individual processes characteristic for muscular dystrophy. The outcome is discussed in more detail in this chapter.

2.1 Degeneration

The current hypothesis of myofiber degeneration in muscular dystrophy is based on the assumption that defects in, or absence of, protein-products of genes with muscular dystrophy causing mutations result in membrane instability[45,58,13,72]. This instability eventually leads to unrepairable ruptures of the sarcolemma. The subsequent influx of calcium-ions from the interstitium, caused by a difference in extracellular and intracellular Ca^{2+} concentrations, leads to numerous events such as hypercontraction of the muscle fiber, the activation of calcium

dependent proteases and deregulated metabolic processes. These processes eventually lead to muscle cell death. However, not all the genetic defects leading to muscular dystrophy can be directly linked to muscle degeneration as a consequence of membrane instability. Genes encoding extracellular proteins, cytoplasmic proteins, glycosylating enzymes, Golgi-related proteins, or nuclear membrane proteins can cause muscular dystrophy as well. The precise degeneration-initiation mechanisms of these muscular dystrophies are not yet understood. In our comparative study of mouse models for muscular dystrophy, we were only able to include models related to membrane instability or membrane repair. Additional gene expression profiling of mouse models related to cytoplasmic proteins, glycosylating enzymes etc. might be beneficial to study alternative causes for muscular dystrophy.

2.1.1 Initiation of degeneration

Since the histological aspect of mdx muscle after birth appears to be normal, dystrophin-deficiency does not seem to lead to muscle deformation during development[16]. However, during neonatal growth the first symptoms of the disease are presented. Degeneration in the mdx mouse starts at approximately 2-4 weeks of age. This coincides with increased muscle activity due to increased ability of ambulation, and becomes evident at histological level. The initiation is characterized by myofiber necrosis, and a subsequent inflammatory response.

Our temporal study in the mdx mouse demonstrates for the first time a large difference in gene expression levels before the first histologically visible degenerative symptoms occur (Chapter 3). The majority of the differentially expressed genes are downregulated and generally remain downregulated during the studied timecourse. This observation leads to a number of questions. Are these genes related, and in what processes do these genes function? Why does the mdx mouse display this downregulation? Functional annotation and subsequent pathway analysis did not identify specific processes in which these genes function. This might be explained by subtle difference of gene expression levels not picked up by the microarray technique. As a result, functional analyses did not come up with enough candidates to determine the (in)activation of specific pathways. Increasing the number of biological replicates might address this limitation. Furthermore, the functional information of the differentially expressed genes is not yet complete (see Discussion Chapter 1.1.5.2), which can lead to the inability to find pathways.

Based on individual genes instead of pathways, differential gene expression at the first week after birth indicates a deficit in muscle development, since genes involved in cell growth are downregulated (data not shown). Hence, this might indicate that there is a difference between pre- and postnatal muscle development, which might be caused by changes in hormone regulation or the developing immune system. However, this should be investigated by additional gene expression profiling experiments on embryonic/fetal muscle tissue of wildtype and mdx mice.

Since the absence of the DGC likely leads to a stagnation of normal muscle growth/postnatal development, a negative feedback mechanism might be initiated for terminal differentiation into normal adult muscle tissue. As a model for myogenesis, myoblast cultures have been used

to study differences in the myogenic capacity of DMD patients versus healthy control individuals[6,5,18]. The DMD derived myoblast cultures displayed changes in morphology, reduced adhesiveness, lower fusion rates, lower numbers of nuclei, and a reduction in the number of population doublings. This supports the hypothesis that a stagnation of muscle growth/postnatal development occurs in the mdx mouse. Gene expression profiling of additional, earlier timepoints of the various mouse models for muscular dystrophy might further substantiate this hypothesis to a more general process in muscular dystrophy.

2.1.2 Inflammatory response

Myofiber necrosis initiates an inflammatory reaction, which removes the necrotic debris in order to facilitate the regeneration of muscle tissue. The first inflammatory cells to arrive at the necrotic location are neutrophils, which are followed by macrophages[43]. Neutrophils are likely to enhance the degeneration of affected myofibers by inducing a strong inflammatory response. The macrophages remove the debris by phagocytosis, and play a role in the activation of satellite cells (see Discussion Chapter 2.1.3).

The first paper to employ gene expression profiling to study muscular dystrophy compared gene expression levels of muscle biopsies from DMD and LGMD patients to those of healthy individuals[15]. This benchmark paper by Chen et al. demonstrated novel associated proteins involved in the early stages of myofiber necrosis, as well as proteins in local inflammatory and bystander responses (complement factor genes, cytokine-inducible nuclear protein). Similar results were found by Haslett et al. in the comparison of gene expression profiles between DMD patients and healthy individuals[23,24].

Recent studies by Porter et al. in the mdx mouse demonstrated a large number of genes involved in the inflammatory process[50,51]. Using different methods such as gene expression profiling, quantitative RT-PCR, and immunoblotting, the authors focused on chemokine ligands and receptors involved in the pathophysiology of muscular dystrophy. They found a number of chemo- and cytokine ligands and receptors (Ccl2, Ccl5, Ccl6, Ccl7, Ccl8, and Ccl9) that are regulated in a temporal and spatial fashion[51]. Multiple markers for macrophages and T-cells remained upregulated in the mdx mouse after the major inflammatory response between 4-6 weeks of age[52]. The prolonged presence of inflammatory cells suggests that inflammation is an ongoing process in muscular dystrophy, which was shown before by histological observations. Other gene expression profiling studies using the mdx mouse as a model for DMD resulted in the identification of more genes involved in inflammation[8,56,64]. These studies corroborated the presence of inflammatory cells such as neutrophils and macrophages.

In our studies, the inflammatory response in the mdx mouse is represented by the detection of differentially expressed genes, which are specifically expressed by inflammatory cells. We also found a large number chemokine and cytokine genes that were differentially expressed (Ccl2, Ccl6, Ccl9, and Ccr8). These markers peak at approximately 6 weeks of age by showing the highest differential gene expression. In our comparative study of the various mouse models, we found that similar chemokines were upregulated as well (Ccl2, Ccl6, Ccl7, Ccl8, Ccl9). These chemotactic molecules can be released from different tissues. For instance, chemokine

expression has been detected in differentiating myoblasts (Chapter 5). Other inflammatory genes, which are upregulated in the mdx mouse and the other models for muscular dystrophy, are interleukins and their receptors (Il2ra, Il4ra, Il6ra, Il17ra, and Il18), and complement factors (C1qa, C1qg, C1qr1, C1qtnf3, C3ra). The complement factors are part of the innate immune system, which is activated by releasing of intracellular components into the interstitium. The complement system consists of a cascade of sequential reactions, which eventually leads to chemotactic attraction of neutrophils and macrophages. Since the inflammatory processes are spatially and temporally regulated, the consecutive nature of the different invading inflammatory cells is difficult to determine using whole tissue as starting material. As a result, markers of different inflammatory cells are detected simultaneously.

2.1.3 Inflammation as activator for regeneration

The role of the inflammatory response is dual; it manages the removal of necrotic debris and it initiates, or at least stimulates, regeneration. Previous studies revealed the necessity of inflammatory cells for regeneration[38]. Macrophages present at the necrotic site activate the quiescent satellite cells and bring them in a proliferative state[44]. The activation of proliferation of satellite cells can occur by several means. Inflammatory cells can release activating factors such as growth factors (Fgf, Hgf, Igf, etc.) and chemo- and cytokines (see Discussion Chapter 2.1.2) that activate satellite cell activation (reviewed in[25]). A number of these molecules was also found to be differentially expressed in the mdx mouse. For instance, fibroblast growth factor 4 (FGF4) was upregulated at 6 weeks of age. This growth factor has been described as a potent activator of myoblast differentiation[21]. Furthermore, we found that the Fgf4 receptor was significantly upregulated in the mouse models for muscular dystrophy compared to wildtype mice.

Satellite cells can also be activated indirectly by the release of proliferation and differentiation factors, which are stored in the extracellular matrix between the myofibers[26,14]. Proteolytic enzymes are released by inflammatory cells and degenerating myofibers. Those enzymes will break down proteins that reside in the extracellular matrix, such as proteoglycan molecules. In healthy muscle tissue, the proteoglycan molecules bind growth factors[32]. A function of the proteolytic enzymes present at the site of necrosis is to partly break up the extracellular matrix to facilitate the migration of satellite cells. This function is mediated by specific proteases of which two important families are the matrix metalloproteinases (MMP) and the ‘a disintegrin and metalloproteinase domain’ (ADAM) containing proteins[36,7]. Members of these specific proteinases are upregulated in the mdx mouse (Mmp8, Mmp14, and Adam11). Mmp14 was found to be upregulated in all mouse models for muscular dystrophy compared to wildtype mice. Thus, the inflammatory reaction activates satellite cells and leads to degradation of the extracellular matrix, which results in the possibility of satellite cell migration and to the release of the proteoglycan-bound satellite cell-activating molecules. However, this last group is not likely to be detected by gene expression profiling, because the proteins are already stored in the extracellular matrix.

2.2 Regeneration

A remarkable characteristic of the mdx mouse is the ability to overcome the lethality of dystrophin-deficiency as seen in humans. Where in humans the disease slowly progresses to increased severity, the mdx mouse recovers after a period of extensive degeneration by successful regeneration[16]. The regenerative period in the mdx mouse is accompanied by a large increase in differential gene expression (Chapter 3). Genes were found that demonstrated specific differential expression during the regenerative period using clustering algorithms. We also applied another approach to study regeneration by studying genes that are involved in myogenesis. Regeneration recapitulates myogenesis in that similar processes are activated[11].

2.2.1 Regeneration in dystrophic muscle tissue

The mechanisms that lead to successful recovery in the mdx mouse are not well understood. An area, in which an extensive amount of research has been done, concerns the satellite cell (SC) population, since these cells are primarily responsible for muscle regeneration[42]. An explanation for the progressive nature of DMD was thought to concern a difference in the number of satellite cells in affected versus normal muscle tissue. Pioneering work from Blau et al. demonstrated a decrease in the number of satellite cells in dystrophic muscle tissue[6]. The progressiveness of DMD was explained as a regenerative deficit by a diminishing SC population due to replicative aging. This aging was thought to be caused by the high turnover of satellite cells. However, telomere-length shortening, which is a hallmark for replicative aging, could not be determined[48]. Furthermore, other studies demonstrated contradictory results by showing elevated SC numbers in dystrophic muscle tissue[33,71,39]. As an alternative cause of the disease, Oexle et al. postulated impairment in differentiation of dystrophin-deficient SC, which highly depends on the myogenic environment. The progressive effect is determined by the myogenic environment, which deteriorates with age.

Altogether, the molecular mechanisms concerning the activation, proliferation, and differentiation of satellite cells in general are still not completely elucidated. As a result, Sterrenburg et al. started to investigate these molecular mechanisms by applying gene expression profiling in human primary myoblast cultures that are forced into differentiation[63]. The results of this study will be discussed in detail in the thesis of Sterrenburg (to be published).

Myofiber degeneration is symptomatic for muscular dystrophy, and activates regenerative processes (see Chapter 3). The regenerative response is, however, not sufficient to counter the degenerative processes on the long term. This insufficiency is proven by the progressive nature of the muscular dystrophies, especially in the severe types like Duchenne Muscular Dystrophy. The major cause of this insufficiency lies in the fact that the regenerated myofibers are still deficient in the protein responsible for the disease. Therefore, the regenerated myofibers are still susceptible for degeneration.

The regenerative response does not only replace affected, degenerative myofibers. As a result of the regenerative processes, the affected muscle tissue changes at several points. Histological analysis demonstrates that the composition of dystrophic myofibers changes in affected tissue. The myofibers show variation in size, there is an increase in endomysial fibrosis and

adipose tissue, and in some muscles the myofiber type changes from fast-twitch (type II) to slow-twitch (type I) (see Discussion Chapter 2.2.4). Furthermore, the interaction with the extracellular matrix seems to change. These changes are probably a result of the muscle tissue to compensate for the degenerative processes other than by regeneration alone. The above-mentioned processes are described in more detail below.

2.2.2 Histological changes of dystrophic myofibers

Regenerated muscle fibers can be defined according to a number of histological parameters. The most evident feature is the location of the myofiber nuclei. In normal muscle, the nuclei are located at the periphery of the myofiber. In contrast, regenerated myofibers have centrally localized nuclei. During maturation of healthy myofibers, the nuclei migrate to reside at the sarcolemma. The central localization might indicate that the certain processes of terminal differentiation of myofibers are not activated or are stalled[11]. Alternatively, centrally located nuclei most likely contribute to the stability of the entire myofiber. In general, the stability of the myofiber is provided by the highly organized intracellular cytoskeleton, and the relation to the extracellular matrix. Mutations in structural proteins of the nuclear envelope can cause muscular dystrophy, indicating that the structural proteins located at other sites than the sarcolemma or the costameres and sarcomeres play a role in the structure of the myofiber. Hence, the central location of the nuclei might contribute to the overall stability of the myofiber per se by generating an alternative scaffold for cytoskeletal stability via the structural proteins of the nuclear envelope.

Dystrophic muscle tissue displays a large variety in myofiber size. Although the precise mechanisms behind this variation are not known, it is likely a consequence of degenerative and regenerative processes. During terminal differentiation of myogenesis, myofibers grow in size by increasing the number of myofibrils. Due to the constant generation of new myofibers by regeneration, the smaller sized myofibers are newly formed myofibers, whereas the large myofibers are full-grown. Alternatively, the ongoing degenerative processes may limit the growth of newly formed fibers in regenerating muscle by altering the microenvironment surrounding the regenerating myofibers.

2.2.3 Altered extracellular matrix

Endomysial fibrosis is a pathological feature of muscular dystrophies, which is most likely a result of the ongoing degenerative processes[4]. The ongoing degeneration leads to a persistent presence of inflammatory cells, which release a number of growth factors amongst which are TGF-β related proteins (Chapter 3). The proliferative effect of the releasing growth factors does not only stimulate satellite cell proliferation, but also stimulates the proliferation of fibroblasts[9]. The proliferation of fibroblasts leads to an increased deposition of extracellular matrix proteins, which cause the extensive fibrosis seen in dystrophic muscle tissue.

2.2.4 Fibertype switch to slow oxidative

Degeneration of myofibers in muscular dystrophy is a result of membrane instability and subsequent rupture of the sarcolemma. In DMD patients and in the mdx mouse, the fast glycolytic

myofibers are more affected than slow oxidative myofibers, since the number of fast glycolytic myofibers declines with age[70,12]. This implies that the fast glycolytic myofibers are more prone to sarcolemmal rupture due to their nature. Alternatively, regenerating myofibers can be predominantly slow oxidative. This hypothesis is supported by the fact that slow oxidative myofibers contain 3-4 times more satellite cells compared to fast muscle fibers.

The distribution of the Dystrophin-glycoprotein complex and the Integrin-Vinculin-Talin complex is different between fast and slow myofiber types[1]. In slow myofibers, the two complexes are localized above the I-band (Z-line). In fast myofibers, the complex is localized above the A-band (M-line). Differences in localization of these cell-signaling complexes might therefore result in altered cell signaling. Consequently, when presence and localization of these complexes is incorrect (as in muscular dystrophy), a difference in cell signaling might contribute to the disease.

Another possibility might be that activation of the Calcineurin cell signaling pathway regulates the slow myofiber phenotype[29,60]. Calcineurin is indirectly activated by intracellular Ca^{2+}-levels, which are increased due to rupture of the sarcolemma. Intracellular Ca^{2+}-levels determine the activity of Calmodulin to activate Calcineurin. Calcineurin subsequently activates NFAT, which then localizes to the nucleus. NFAT, together with other co-factors like MEF2, HDAC, GATA, and MRFs can induce transcription of slow fibertype specific genes.

Alternatively, specific expression of fibertype regulating genes might influence the fibertype switch seen in dystrophic muscle tissue. Expression of specific myogenic regulatory factors, under influence of innervation and hormone levels, can determine the fibertype[31]. More downstream of the MRFs, the level of expression of dystrophin or utrophin might be able to influence fibertype specificity[54]. Furthermore, intracellular calcium storage related genes, such as Ryr1 and Ip3r1, can influence the fibertype as well[34,35].

2.2.5 Human regeneration

The inability to regenerate dystrophic muscle successfully in DMD patients was assumed to be due to replicative exhaustion or aging of satellite cells. When the satellite pool is depleted or exhausted, dystrophic muscle tissue is not able to grow anymore. As an eventually lethal alternative, muscle tissue is then replaced by fibrotic and adipose tissue.

The latest results indicate that the size of the satellite cell pool is not decreased, and that satellite cells are still able to replicate and proliferate at increasing age. Ineffective regeneration seems to be caused by impaired differentiation of satellite cells[48,63]. This impaired differentiation can be explained by a number of causes. Although these causes can function separately, the possibility of a multifactorial pathology is not to be ruled out.

Aging can lead to a change in the microenvironment in the dystrophic muscle. The decisive processes leading to activation of satellite cells are not yet fully known. However, a number of growth factors have the ability to activate these processes (PDGF, IGF1, etc.). These growth factors can be released by the affected myofiber, inflammatory cells, or released from the extracellular matrix. Cycles of degeneration and regeneration can influence the processes

or availability of various cell types, thereby changing the microenvironment. The activating agents probably do not concentrate on satellite cells alone. Dystrophic muscle tissue has a high number of proliferating fibroblasts, which might be activated by similar mechanisms[5].

Impaired differentiation of satellite cells can be caused by a stall in proliferation. This can be caused by transforming growth factor β-related proteins, which keep the satellite cells in a proliferative state by blocking differentiation pathways. Dystrophin-deficiency can also directly lead to impaired differentiation if dystrophin plays a role in the differentiation process as a mechanical linker in the sarcolemma to the ECM, where it might function in a signaling complex. Furthermore, aggregates of acetylcholine receptors of junctional folds can be decreased due to a lack of dystrophin, which leads to altered signaling between nerve and muscle.

Primary in vitro cell cultures of human myoblasts consist of satellite cells and fibroblasts. Comparison of myoblast cultures derived from DMD patients with healthy controls showed that the ability to fuse was reduced[17]. Sterrenburg et al. used gene expression profiling to study the differences between primary in vitro cell cultures of human myoblast of both DMD patients and healthy individuals (manuscript accepted for publication)

Concluding, the progressive nature of Duchenne Muscular Dystrophy does not seem to be caused by a decrease in the satellite cell population, but due to a stall in terminal differentiation of activated myoblasts. The reason that skeletal muscle develops rather normally in DMD patients can be caused by relative inactiveness of skeletal muscle during the first years in life. The progressive nature starting from the onset of the disease commences with physical activity in that stage of life. This is in correlation with the current hypothesis that DMD is caused by rupture of the sarcolemma due to physical stress.

2.2.6 Regeneration during muscle growth

Although the severity of muscular dystrophy mainly depends on the affected gene or protein, the age of onset is likely to influence the severity of the disease, since it will affect normal development. The muscular dystrophies with an early age of onset are the dystrophinopathies and the congenital muscular dystrophies (CMD). The muscular dystrophies with a later age of onset are Limb-Girdle-, Facioscapulohumeral-, and Oculopharyngeal Muscular Dystrophy. This latter group has a milder severity compared to the first.

Muscle growth during postnatal development requires the activation, proliferation and differentiation of satellite cells to increase the number of myofibers. Regeneration of muscle tissue requires the same pool of satellite cells. When both processes (i.e. growth and regeneration) occur simultaneously, the demand on the satellite cell pool might become too large. As a result, both processes will compete, and not function adequately. The demand of satellite cells for regenerative purposes can be decreased by countering the inflammatory response, since inflammatory cells are able to activate satellite cells (see Discussion Chapter 2.1.3). Clinical trials with DMD patients using glucocorticoid corticosteroids (prednisone and deflazacort) showed a significant and sustained beneficial effect in slowing the decline of muscle strength as well as increasing the muscle mass (reviewed in[73,41]). This indicates that muscle growth via satellite cell activation can resume due to a decrease in regenerative demand. However, other

immunosuppressive drugs (i.e. azathioprine) did not have similar beneficial effect on muscle growth[37], indicating additional effects of prednisone and deflazacort. What those effects precisely are, remains unclear.

In human DMD patients regenerative processes can be demonstrated up to the age of 5 years, since the regenerated myofibers up to then express the characteristic fetal isoform of MyHC[27]. Growth is, however, active into young adulthood. The precise reasons for the termination of the regenerative processes are not yet fully known. As referred to in Discussion Chapter 2.2.5, the terminal differentiation of activated satellite cells may be affected. Furthermore, the myogenic environment might be altered irreversibly, resulting in impaired differentiation. This also might explain the ongoing wasting of myofibers after loss of ambulation.

Effective regeneration in mdx mice coincides with reaching adulthood, when the mice are full-grown. The remaining SC population can then be solely available for regenerative purposes. This might explain the why the mouse model for DMD shows a more successful regeneration compared to the human pathology of DMD.

2.3 Comparison between mice and men

With the introduction of high-throughput gene expression profiling, the study of muscular dystrophies has been given another direction, towards elucidation of the secondary mechanisms leading to the disease. Comparing gene expression profiles between human individuals affected by a muscular dystrophy and the corresponding animal model has potential to contribute substantially to the pathophysiological processes responsible for the disease, and to circumvent the high variation seen in humans. Unfortunately, a direct comparison between human samples and samples derived from animal models has not yet been performed. As a consequence, determination of common and shared gene expression profiles is more difficult due to the presence of a higher number of variables, such as the used platform, stages of the disease, and used muscle types. Furthermore, functional annotation between species might differ, and complicate pathway analysis.

Gene expression profiling studies have identified the molecular processes of a large number of the secondary effects of the disease (see previous chapters). Large similarities have been found between humans and animal models affected by the disease on pathophysiological levels. With an increasing amount of data becoming publically available from an increasing number of studies, the finding of similarities, and -equally informative- differences, in the molecular details of these processes will become increasingly useful.

This thesis does not discuss a direct comparison between the molecular details of shared and common processes in human and animal muscular dystrophy. The main reason is that to accomplish such a direct comparison, the existing datasets must be combined, and a direct comparison must be made based on computational analysis, which has not been done. This comparison requires a highly specialized bioinformatics approach, for which fundamental difficulties (such as gene annotation and the use of different platforms) must be resolved first. However, the results presented in this thesis are aimed at contributing to an understanding in more depth of the pathophysiological processes in muscular dystrophy.

3 Future leads

The development of the gene expression profiling technique in general has resulted in an augmentation of understanding molecular processes due to the high-throughput nature of the technique. The technique can be applied to both generate and confirm hypotheses. Although this technique is solely based on gene expression levels, it assists in a deeper insight into the regulation of biological processes and functional organization both on cellular and tissue levels from healthy and diseased organisms. Application of gene expression profiling (the most popular method in transcriptomics) is therefore a first step towards a comprehensive knowledge of organismal function, which is the ultimate goal of systems biology. The integration of other rapidly developing 'omics', such as proteomics, metabolomics, cytomics, and nutrigenomics will ultimately realize in the accomplishment of this goal.

Gene expression profiling as a technique has matured over the last decade. The initial difficulties concerned the miniaturization at methodological level (robotic printing of nanodrop quantities, fluorescent based laser scanning) in order to make an upscaling (high-throughput character) possible. When determination of relative gene expression levels was realized, statistical procedures had to be designed to provide reliable conclusions. As a result, the experimental design of microarray-based experiments was optimized for balancing technical and biological effects. At this moment, the modeling of large amounts of gene expression data into molecular and cellular processes based on functional annotation and meta-analysis (also known as inference) has become a major topic towards understanding gene expression profiles.

With the maturation of gene expression profiling and the involvement of other fields outside biology, such as methodology and statistics, a more critical view towards the application of the technique and the subsequent interpretation of the results is rising. Particularly, gene expression profiling used for classification of diseases such as cancer has received major critisism[47,55,20]. By no means will this criticism lead to the downfall of gene expression profiling, it will lead to further optimization of the technique and the reliability of its results.

3.1 Linking Gene Expression to Pathology

Gene expression profiling of muscle tissue affected by muscular dystrophy has resulted in a greater understanding of the pathological processes leading to the disease. The most apparent contribution lies in the elucidation of genes involved in different secondary processes of the disease, like degeneration, inflammation, fibrosis, and regeneration. These secondary processes are the major symptoms of muscular dystrophy, and are described in detail from a histological point of view. The major contribution of the technique in elucidating these secondary processes is the generation of potential leads for future therapy.

Differentially expressed genes in muscular dystrophy can be classified according to the known secondary processes. This is one strategy of resolving molecular details. Another strategy is to find unknown pathways related to muscular dystrophy based on gene expression profiling. However, finding those pathways is a laborious task since it depends on available information of genes. The functional annotation of genes according to gene ontology in combination with published data will highlight pathways not earlier described in muscular dystrophy. Under-

standably, this strategy depends on the availability of gene information. Biological databases for functional annotation, gene and protein information, and literature information like NCBI, Kegg, and Expasy are therefore indispensable in the interpretation of gene expression results. As the amount of information in public databases increases, so will the interpretation of gene expression data. Although the reason behind the differential expression of all genes cannot be explained at this moment, future analysis of gene expression profiles, even the ones generated earlier, will therefore result in a greater understanding of the processes playing a role in muscular dystrophy.

From the clinical perspective, a better understanding of muscular dystrophy will be aided by increasing the number of analyzed patients, and subsequent classification according to clinical phenotype. Thus, stage and age dependent gene expression profiles can be determined as well as mutation-related profiles. To find disease specific gene expression profiles, it is also necessary to study a large number of patients with different muscular dystrophy types.

3.2 And vice versa...

Clinical differentiation between the various muscular dystrophies is a laborious task, where the outcome cannot always be confirmed. In general, gene expression profiling is starting to be widely used to identify biomarkers, which are specifically expressed in a given type of disease or at the stage of a disease[22,65,69]. However, one must keep in mind that the application of the technique is still emerging, and the success of finding biomarkers lies in the elucidation of all possible variables. Finding biomarkers for the muscular dystrophy types will eventually facilitate and increase the speed of diagnosis of muscular dystrophy and its accuracy. The first studies have been performed on finding biomarker genes for muscular dystrophy[2,28]. However, a large-scale study as is performed on cancer patients has not been the case for the human dystrophies.

This thesis described the first study towards a fast and straightforward differential diagnosis of muscular dystrophy based on gene expression profiles (Chapter 4). The study was initiated to differentiate various mouse models for muscular dystrophy according to gene expression profiles. Although the study is not complete, it provides the first set of biomarkers for muscular dystrophy. To achieve a clinical diagnosis based on gene expression profiling, a number of additional experiments are required. First, the number of biological replicates needs to be increased to gain accuracy regarding the potential biomarkers. Second, the number of timepoints needs to be increased to gain a better understanding about the stages in the pathology within and between muscular dystrophies. Third, differential gene expression levels need to be determined between different mouse models of a single muscular dystrophy to determine the effect of different mutations on clinical phenotype and subsequent gene expression. Fourth, these results need to be extrapolated to, and confirmed in, human patients. Fifth, gene expression data of both human patients and animal models need to be stored in a specialized muscular dystrophy database with optimized functional annotation.

With the growing interest in gene expression profiling of muscular dystrophy, these goals can be realized. Nonetheless, the realization of these goals acquires the collaboration of many laboratories and the possibility to access the data publicly. Furthermore, the completion of this

project will increase the likelihood of understanding the details of the disease, and the subsequent development of stage- or disease-specific medication.

4 References

1. Anastasi G, Amato A, Tarone G, Vita G, Monici MC, Magaudda L, Brancaccio M, Sidoti A, Trimarchi F, Favaloro A, Cutroneo G. Distribution and localization of vinculin-talin-integrin system and dystro phin-glycoprotein complex in human skeletal muscle. Immunohistochemical study using confocal laser scanning microscopy. Cells Tissues Organs 2003;175(3):151-164.
2. Arning L, Jagiello P, Schara U, Vorgerd M, Dahmen N, Gencikova A, Mortier W, Epplen JT, Gencik M. Transcriptional profiles from patients with dystrophinopathies and limb girdle muscular dystrophies as determined by qRT-PCR. Journal of Neurology 2004;251(1):72.
3. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful appraoch to multiple testing. JRoyStatSocB 1995;57:289.
4. Bernasconi P, Torchiana E, Confalonieri P, Brugnoni R, Barresi R, Mora M, Cornelio F, Morandi L, Mantegazza R. Expression of transforming growth factor-beta 1 in dystrophic patient muscles corre lates with fibrosis. Pathogenetic role of a fibrogenic cytokine. J Clin Invest 1995;96(2):1137-1144.
5. Blau HM, Webster C, Chiu CP, Guttman S, Chandler F. Differentiation properties of pure populations of human dystrophic muscle cells. Exp Cell Res 1983;144(2):495-503.
6. Blau HM, Webster C, Pavlath GK. Defective myoblasts identified in Duchenne muscular dystrophy. Proc Natl Acad Sci U S A 1983;80(15):4856-4860.
7. Blobel CP. Functional and biochemical characterization of ADAMs and their predicted role in protein ectodomain shedding. Inflamm Res 2002;51(2):83-84.
8. Boer JM, de Meijer EJ, Mank EM, van Ommen GB, den Dunnen JT. Expression profiling in stably regenerating skeletal muscle of dystrophin-deficient mdx mice. NeuromusculDisord 2002;12:S118.
9. Border WA, Noble NA. Transforming growth factor beta in tissue fibrosis. N Engl J Med 1994;331(19):1286-1292.
10. Brown PO, Botstein D. Exploring the new world of the genome with DNA microarrays. Nature Genetics 1999;21(1 Suppl):33.
11. Carlson BM. The regeneration of skeletal muscle. A review. Am J Anat 1973;137(2):119-149.
12. Carnwath JW, Shotton DM. Muscular dystrophy in the mdx mouse: histopathology of the soleus and extensor digitorum longus muscles. J Neurol Sci 1987;80(1):39-54.
13. Carpenter S, Karpati G. Duchenne muscular dystrophy: plasma membrane loss initiates muscle cell necrosis unless it is repaired. Brain 1979;102(1):147.
14. Casar JC, Cabello-Verrugio C, Olguin H, Aldunate R, Inestrosa NC, Brandan E. Heparan sulfate proteoglycans are increased during skeletal muscle regeneration: requirement of syndecan-3 for successful fiber formation. J Cell Sci 2004;117(Pt 1):73-84.
15. Chen YW, Zhao P, Borup R, Hoffman EP. Expression profiling in the muscular dystrophies: identifica tion of novel aspects of molecular pathophysiology. Journal of Cell Biology 2000;151(6):1321.
16. Dangain J, Vrbova G. Muscle development in mdx mutant mice. Muscle Nerve 1984;7(9):700.
17. Delaporte C, Dehaupas M, Fardeau M. Comparison between the growth pattern of cell cultures from normal and Duchenne dystrophy muscle. J Neurol Sci 1984;64(2):149-160.
18. Delaporte C, Dautreaux B, Rouche A, Fardeau M. Changes in surface morphology and basal lamina of cultured muscle cells from Duchenne muscular dystrophy patients. J Neurol Sci 1990;95(1):77-88.
19. Durbeej M, Campbell KP. Muscular dystrophies involving the dystrophin-glycoprotein complex: an overview of current mouse models. CurrOpinGenetDev 2002;12(3):349.
20. Ein-Dor L, Kela I, Getz G, Givol D, Domany E. Outcome signature genes in breast cancer: is there a unique set? Bioinformatics 2005;21(2):171-178.
21. Fraidenraich D, Iwahori A, Rudnicki M, Basilico C. Activation of fgf4 gene expression in the myo tomes is regulated by myogenic bHLH factors and by sonic hedgehog. Dev Biol 2000;225(2):392-406.
22. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES. Molecular classification of cancer: class dis covery and class prediction by gene expression monitoring. Science 1999;286(5439):531-537.
23. Haslett JN, Sanoudou D, Kho AT, Bennett RR, Greenberg SA, Kohane IS, Beggs AH, Kunkel LM. Gene expression comparison of biopsies from Duchenne muscular dystrophy (DMD) and normal skeletal muscle. ProcNatlAcadSciUSA 2002;99(23):15000.
24. Haslett JN, Sanoudou D, Kho AT, Han M, Bennett RR, Kohane IS, Beggs AH, Kunkel LM. Gene expression profiling of Duchenne muscular dystrophy skeletal muscle. Neurogenetics 2003.
25. Hawke TJ, Garry DJ. Myogenic satellite cells: physiology to molecular biology. JApplPhysiol

2001;91(2):534.
26. Haynes LW. Fibroblast (heparin-binding) growing factors in neuronal development and repair. Mol Neurobiol 1988;2(4):263-289.
27. Hoffman EP, Partridge TA. Genotype/phenotype correlations in Duchenne/becker dystrophy. In: Wright DJM, Archard LC, editors. Molecular and Cell Biology of Muscular Dystrophy. Volume 1st, Molecular and Cell Biology of Human Diseases. London: Chapman & Hall; 1993. p 12.
28. Hoffman EP, Rao D, Pachman LM. Clarifying the boundaries between the inflammatory and dystrophic myopathies: insights from molecular diagnostics and microarrays. Rheum Dis Clin North Am 2002;28(4):743-757.
29. Horsley V, Jansen KM, Mills ST, Pavlath GK. IL-4 acts as a myoblast recruitment factor during mam malian muscle growth. Cell 2003;113(4):483.
30. Huber W, Von Heydebreck A, Sultmann H, Poustka A, Vingron M. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. Bioinformatics 2002;18: S96.
31. Hughes SM, Taylor JM, Tapscott SJ, Gurley CM, Carter WJ, Peterson CA. Selective accumulation of MyoD and myogenin mRNAs in fast and slow adult skeletal muscle is controlled by innervation and hormones. Development 1993;118(4):1137-1147.
32. Iozzo RV. Basement membrane proteoglycans: from cellar to ceiling. Nat Rev Mol Cell Biol 2005;6(8):646-656.
33. Ishimoto S, Goto I, Ohta M, Kuroiwa Y. A quantitative study of the muscle satellite cells in various neuromuscular disorders. J Neurol Sci 1983;62(1-3):303-314.
34. Jordan T, Jiang H, Li H, DiMario JX. Inhibition of ryanodine receptor 1 in fast skeletal muscle fibers induces a fast-to-slow muscle fiber type transition. J Cell Sci 2004;117(Pt 25):6175-6183.
35. Jordan T, Jiang H, Li H, DiMario JX. Regulation of skeletal muscle fiber type and slow myosin heavy chain 2 gene expression by inositol trisphosphate receptor 1. J Cell Sci 2005;118(Pt 10):2295-2302.
36. Kieseier BC, Schneider C, Clements JM, Gearing AJ, Gold R, Toyka KV, Hartung HP. Expression of specific matrix metalloproteinases in inflammatory myopathies. Brain 2001;124(Pt 2):341-351.
37. Kissel JT, Lynn DJ, Rammohan KW, Klein JP, Griggs RC, Moxley RT, 3rd, Cwik VA, Brooke MH, Mendell JR. Mononuclear cell analysis of muscle biopsies in prednisone- and azathioprine-treated Duchenne muscular dystrophy. Neurology 1993;43(3 Pt 1):532-536.
38. Lescaudron L, Peltekian E, Fontaine-Perus J, Paulin D, Zampieri M, Garcia L, Parrish E. Blood borne macrophages are essential for the triggering of muscle regeneration following muscle transplant. Neuromuscul Disord 1999;9(2):72-80.
39. Maier F, Bornemann A. Comparison of the muscle fiber diameter and satellite cell frequency in human muscle biopsies. Muscle Nerve 1999;22(5):578-583.
40. Maier H, Dohr S, Grote K, O'Keeffe S, Werner T, Hrabe de Angelis M, Schneider R. LitMiner and WikiGene: identifying problem-related key players of gene regulation using publication abstracts. Nucleic Acids Res 2005;33(Web Server issue):W779-782.
41. Manzur AY, Kuntzer T, Pike M, Swan A. Glucocorticoid corticosteroids for Duchenne muscular dystro phy. Cochrane Database Syst Rev 2004(2):CD003725.
42. Mauro A. Satellite cell of skeletal muscle fibers. J Biophys Biochem Cytol 1961;9:493-495.
43. McLennan IS. Degenerating and regenerating skeletal muscles contain several subpopulations of mac rophages with distinct spatial and temporal distributions. J Anat 1996;188 (Pt 1):17-28.
44. Merly F, Lescaudron L, Rouaud T, Crossin F, Gardahaut MF. Macrophages enhance muscle satellite cell proliferation and delay their differentiation. Muscle Nerve 1999;22(6):724.
45. Mokri B, Engel AG. Duchenne dystrophy: electron microscopic findings pointing to a basic or early abnormality in the plasma membrane of the muscle fiber. Neurology 1975;25(12):1111-1120.
46. Muller HM, Kenny EE, Sternberg PW. Textpresso: an ontology-based information retrieval and extrac tion system for biological literature. PLoS Biol 2004;2(11):e309.
47. Ntzani EE, Ioannidis JP. Predictive ability of DNA microarrays for cancer outcomes and correlates: an empirical assessment. Lancet 2003;362(9394):1439-1444.
48. Oexle K, Kohlschutter A. Cause of progression in Duchenne muscular dystrophy: impaired differentia tion more probable than replicative aging. Neuropediatrics 2001;32(3):123-129.
49. Ohyama H, Zhang X, Kohno Y, Alevizos I, Posner M, Wong DT, Todd R. Laser capture microdissec tion-generated target sample for high-density oligonucleotide array hybridization. Biotechniques 2000;29(3):530-536.

50. Porter JD, Khanna S, Kaminski HJ, Rao JS, Merriam AP, Richmonds CR, Leahy P, Li J, Guo W, Andrade FH. A chronic inflammatory response dominates the skeletal muscle molecular signature in dystrophin-deficient mdx mice. Human Molecular Genetics 2002;11(3):263.
51. Porter JD, Guo W, Merriam AP, Khanna S, Cheng G, Zhou X, Andrade FH, Richmonds C, Kaminski HJ. Persistent over-expression of specific CC class chemokines correlates with macrophage and T-cell recruitment in mdx skeletal muscle. NeuromusculDisord 2003;13(3):223.
52. Porter JD, Merriam AP, Leahy P, Gong B, Khanna S. Dissection of temporal gene expression sig natures of affected and spared muscle groups in dystrophin-deficient (mdx) mice. Human Molecular Genetics 2003;12(15):1813.
53. Quackenbush J. Microarray data normalization and transformation. Nat Genet 2002;32 Suppl:496-501.
54. Rafael JA, Townsend ER, Squire SE, Potter AC, Chamberlain JS, Davies KE. Dystrophin and utrophin influence fiber type composition and post-synaptic membrane structure. Hum Mol Genet 2000;9(9):1357-1367.
55. Ransohoff DF. Rules of evidence for cancer molecular-marker discovery and validation. Nat Rev Can cer 2004;4(4):309-314.
56. Rouger K, Le Cunff M, Steenman M, Potier MC, Gibelin N, Dechesne CA, Leger JJ. Global/tempo ral gene expression in diaphragm and hindlimb muscles of dystrophin-deficient (mdx) mice. AmJPhysiol Cell Physiol 2002;283(3):C773.
57. Schena M, Shalon D, Heller R, Chai A, Brown PO, Davis RW. Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. Proceedings of the national Academy of Sciences USA 1996;93(20):10614.
58. Schmalbruch H. Segmental fibre breakdown and defects of the plasmalemma in diseased human muscles. Acta Neuropathol (Berl) 1975;33(2):129-141.
59. Schuemie MJ, Weeber M, Schijvenaars BJ, van Mulligen EM, van der Eijk CC, Jelier R, Mons B, Kors JA. Distribution of information in biomedical abstracts and full-text publications. Bioinformatics 2004;20(16):2597-2604.
60. Schulz RA, Yutzey KE. Calcineurin signaling and NFAT activation in cardiovascular and skeletal muscle development. Dev Biol 2004;266(1):1-16.
61. Staal YC, van Herwijnen MH, van Schooten FJ, van Delft JH. Application of four dyes in gene expres sion analyses by microarrays. BMC Genomics 2005;6:101.
62. Sterrenburg E, Turk R, Boer JM, van Ommen GB, den Dunnen JT. A common reference for cDNA microarray hybridizations. Nucleic Acids Res 2002;30(21):e116.
63. Sterrenburg E, Turk R, t Hoen PA, van Deutekom JC, Boer JM, van Ommen GJ, den Dunnen JT. Large-scale gene expression analysis of human skeletal myoblast differentiation. NeuromusculDisord 2004;14(8-9):507.
64. Tseng BS, Zhao P, Pattison JS, Gordon SE, Granchelli JA, Madsen RW, Folk LC, Hoffman EP, Booth FW. Regenerated mdx mouse skeletal muscle shows differential mRNA expression. JApplPhysiol 2002;93(2):537.
65. van ‘t Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002;415(6871):530-536.
66. Van Mulligen EM, Van Der Eijk C, Kors JA, Schijvenaars BJ, Mons B. Research for research: tools for knowledge discovery and visualization. Proc AMIA Symp 2002:835-839.
67. Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. Serial analysis of gene expression. Science 1995;270:484.
68. Vinciotti V, Liu X, Turk R, de Meijer EJ, t Hoen PA. Exploiting the full power of temporal gene expression profiling through a new statistical test: Application to the analysis of muscular dystrophy data. BMC Bioinformatics 2006;7(1):183.
69. Wadlow R, Ramaswamy S. DNA microarrays in clinical cancer research. Curr Mol Med 2005;5(1): 111-120.
70. Watkins SC, Cullen MJ. Histochemical fibre typing and ultrastructure of the small fibres in Duchenne muscular dystrophy. Neuropathol Appl Neurobiol 1985;11(6):447-460.
71. Watkins SC, Cullen MJ. A quantitative study of myonuclear and satellite cell nuclear size in Duch enne’s muscular dystrophy, polymyositis and normal human skeletal muscle. Anat Rec 1988;222(1):6-11.

72. Weller B, Karpati G, Carpenter S. Dystrophin-deficient mdx muscle fibers are preferentially vulnerable to necrosis induced by experimental lengthening contractions. J Neurol Sci 1990;100(1-2):9-13.
73. Wong BL, Christopher C. Corticosteroids in Duchenne muscular dystrophy: a reappraisal. J Child Neurol 2002;17(3):183-190.
74. Yang YH, Speed T. Design issues for cDNA microarray experiments. NatRevGenet 2002;3(8):579.

Chapter 8

Summary

Samenvatting

Summary

Muscular dystrophies are characterized by progressive irreversible degeneration processes, which results in weakness and wasting of muscle tissue. The muscular dystrophies share clinical symptoms, but differ largely in severity, age of onset, and distribution of affected muscle tissue. The genetic defects causative for the majority of muscular dystrophies have been largely elucidated. However, the mechanisms responsible for the pathology and its divergence have not yet been identified in detail. In this thesis, molecular mechanisms in muscular dystrophies are studied in detail.

Each cell in an organism contains the same hereditary information encoded by the DNA, which is known as the genome. DNA uses a four-letter code to describe this information. This code contains four different molecules, the nucleotides adenine, thymine, guanine, and cytosine, which form pairs (basepairs). A sequence of basepairs is called the DNA sequence. The human genome contains twenty-two different autosomal chromosomes and two sexchromosomes. A chromosome is made of one strand of DNA, of which the largest contains approximately 250 million basepairs.
The human genome harbours approximately 22.000 genes, which are the blueprints of proteins. In general, each gene contains the information to build a specific protein. This information is not directly read from the DNA, but via an intermediate form called the messenger RNA (mRNA). Not all possible proteins are made in each individual cell of a multicellular organism. By expressing a specific subset of proteins, cells are able to show diversity and to specialize to perform a specific function in the organism. For instance, different proteins are expressed in a neuron compared to a muscle cell.
The expression of proteins is regulated by controlling which genes are activated and subsequently transcribed into the intermediate mRNA molecules. Therefore, by studying the population of mRNA molecules it is possible to know which proteins are expressed in a cell.

Changes in the DNA sequence, also known as mutations, can lead to changes in proteins or even absence of proteins when these changes occur in the coding sequence of genes. These changes can be beneficial to an organism and this is the keystone for evolution. However, these changes can also have devastating effects such as in cancer. When disease-causing mutations are passed on from parent to offspring, the disease is known as a hereditary disease. To compensate for the loss of a functional protein, the organism can activate or deactivate the production of other proteins. Furthermore, in a multicellular organism, other cells can be influenced by the loss of a certain protein. The results of mutations can be studied at different levels. Based on the clinical diagnosis, candidate genes for the human disease might be present. DNA analysis can be used for confirmation. To study the effects of the mutation, for instance in the interaction of the affected protein with other proteins, studies can be done on the protein level. Until recently, this was based on research that investigated the biological processes step by step. Only one or a small number of genes/proteins were studied at a time. Recently, techniques have been developed to study biological processes on a much larger scale. For instance, one can think about the sequencing of the genome of different organisms. This also accounts for studying the intermediate mRNA molecules.

Gene expression profiling is a technique to determine gene expression levels on a large scale. The technique is based on the ability of nucleotide sequences to bind to their complemetary strand in a process called hybridization. Hybridization occurs when a nucleotide sequence finds another nucleotide sequence with which it can form basepairs across the entire length. Gene expression profiling can be done by using microarrays. Microarrays are devices onto which tens-of-thousands of specific nucleotide sequences are spotted in a known fashion. Each spot contains the information of a single gene. When mRNA is labelled with a specific dye and subsequently hybridyzed to a microarray, it is possible to determine which spots are labelled and therefore which genes are expressed. When the mRNA population of a biological sample is determined, one has determined the gene expression profile of that sample. With this technique it is possible to compare the gene expression profiles of different biological samples. For instance, it is possible to contrast the mRNA molecules present in a biological samples from healthy and diseased subjects. This is known as differential gene expression. The platforms used in this thesis are home-made two-colour microarrays based on long or short DNA sequences. The long sequences are copies of total mRNA molecules (cDNA), whereas the short sequences are based on mRNA specific sequences with a length of 65 nucleotides (oligonucleotides). Two differently labelled mRNA samples are hybridized simultaneously using these type of microarrays, hence the two-colour.

The analysis of gene expression data resulting from microarray experiments is a challenge. The large amount of data requires a thorough statistical approach, whereas the biological outcome requires a reductionistic approach to understand the underlying cellular and molecular processes. To tackle to statistical analsis, we have studied the effect of experimental design on measured gene expression levels (Chapter 6). We concluded that two-color microarray data can be treated separately, which makes the design of microarrays and the statistical analysis more flexible .

Muscle tissue has the ability to regenerate after damage. The responsible cells are quiescent satellite cells, which are undifferentiated. By multiplication, a process called proliferation, and by maturation into muscle cells, a process called differentiation, damaged muscle tissue can be replaced. These processes are also active in muscle tissue affected by muscular dystrophy. However, the genetic defect causative of muscular dystophy is also present in the regenerated tissue. This contributes to the progressive nature of the disease. Regeneration capitulates processes that are active during the development of muscle tissue, also known as myogenesis. To study these processes in more detail we have determined the genes that are active during myogenesis. These experiments were done in satellite cells that are cultered in vitro (Chapter 5).

To study hereditary diseases, animal models can be generated, which have directed changes in their DNA sequence. As a result, the human disease is mimicked. The development of animal models have greatly facilitated the study of human disease throughout the last decades. One of the largest achievements has been the ability to generate transgenic mice. Although Duchenne Muscular Dystrophy is the most severe muscular dystrophy, leading to early death in human patients, mice with a similar genetic defect (mdx mice) do not die prematurely. The underlying mechanisms for the apparent survival are not yet known. Although the mouse does not represent the human disease in this respect, it might give a clue for therapy in human patients, because other dystrophic processes are active. We have focussed on the regeneration

phase in the mdx mouse to determine the processes involved in regeneration (Chapter 3). As expected we found that a large number of genes as well as pathways essential in mogenesis are expressed during regeneration.

By using gene expression profiling, we have set out to study similarities and differences between the muscular dystrophies (Chapter 4). Therefore, we have used mouse models for different muscular dystrophies. Based on our results we found that a large number of processes are shared between the different muscular dystrophies. These processes are mainly seondary effects of the primary disease causing mutation. The current hypothesis describing the cause of muscular dystrophy, is that specific mutations can lead to instability of the muscle fibers. As a result, the muscle fibers are very fragile and are more readily subject to cell death than healthy muscle fibers. Next to finding the shared gene expression profiles, we set out to find specific differences between the muscular dystrophies. This might be valuable for diagnostic purposes. The amount of differences we found were relatively small. This might be attributed to the fact that we only determined the gene expression profiles at a single timepoint. Further studies are necessary to find more differences, and this is a project we are working on currently.

Concluding, using gene expression profiling we have started to unravel the multiple processes active in muscular dystrophy. These results contribute to the understanding of the disease, and hopefully to the development of therapies. Furthermore, by studying mouse models for different muscular dystrophies, clues regarding the differences in age of onset, the pattern of affected muscle groups, and the severity of the disease might be understood eventually by further research.

Samenvatting

Spierdystrofieën worden gekarakteriseerd door progressieve irreversibele degeneratie van spierweefsel, die resulteert in zwakte and verwoesting. De spierdystrofieën hebben een groot aantal klinische symptomen gemeen, maar verschillen met name in ernst, aanvang van de eerste symptomen, en de distributie van aangedaan spierweefsel. De genetische defecten die de meeste spierdystrofieën veroorzaken zijn grotendeels bekend. De mechanismen die ten grondslag liggen aan het ziektebeeld en de verschillen daarin zijn daarentegen nog niet in detail bekend. Deze samenvatting bespreekt de experimenten die beschreven zijn in dit proefschrift om de moleculaire mechanismen in spierdystrofie te bestuderen.

Elke cel in een organisme bevat dezelfde erfelijke informatie in de vorm van DNA wat bekend staat als het genoom. DNA gebruikt een 4-letter code om deze informatie te beschrijven. Deze code bestaat uit 4 verschillende moleculen, namelijk adenine, thymine, guanine, en cytosine (nucleotiden), die in paren voorkomen (baseparen). De volgorde van baseparen wordt ook wel de DNA sequentie genoemd. Het humane genoom is opgedeeld in 22 verschillende autosomale chromosomen en 2 geslachtschromosomen. Een chromosoom bestaat uit één lange DNA streng, waarvan de grootste ongeveer 250 miljoen baseparen lang is.
Het menselijke genoom bevat ongeveer 22.000 genen, die de blauwdrukken zijn van eiwitten. In het algemeen bevat elk gen de informatie om een specifiek eiwit te produceren. Deze informatie wordt niet direct van het DNA afgelezen, maar via een intermediaire vorm, genaamd het boodschapper (messenger) RNA, ofwel mRNA. Niet in elke cel van een organisme wordt elk mogelijk eiwit geproduceerd. Door het kiezen van een bepaalde selectie van eiwitten is het mogelijk om diversiteit tussen cellen te creëren, alsmede om te specialiseren tot het uitvoeren van een bepaalde functie. Er zijn bijvoorbeeld andere eiwitten in een spiercel aanwezig in vergelijking tot een zenuwcel.
De productie van eiwitten wordt bepaald door te controleren welke genen worden geactiveerd en vertaald in mRNA. Door de populatie van mRNA moleculen te bestuderen, is het mogelijk om te weten welke eiwitten er worden geproduceerd.

Veranderingen in de DNA sequentie, ook bekend als mutaties, kunnen leiden tot veranderingen in of zelfs het ontbreken van eiwitten. Deze veranderingen kunnen het organisme ten goede komen; dit is het principe van evolutie. Deze veranderingen kunnen daarentegen ook verwoestende effecten hebben, zoals in kanker vaak het geval is. Wanneer ziekteveroorzakende mutaties worden doorgegeven van ouder op kind, dan is de ziekte een erfelijke ziekte.
Het organisme kan de productie van bepaalde eiwitten activeren of stoppen om te compenseren voor het verlies van een eiwit. In een multicellulair organisme is het zelfs mogelijk dat andere cellen kunnen worden beïnvloed door het verlies van een bepaald eiwit. Het gevolg van mutaties kan op verschillende niveaus worden bestudeerd. Op basis van de klinische diagnose kunnen genen worden geselecteerd waarvan bekend is dat ze bepaalde symptomen veroorzaken. Vervolgens kan DNA analyse worden toegepast om mutaties te valideren. Door middel van eiwit studies kan worden bepaald of het veranderde of verloren eiwit interacties heeft met andere eiwitten. Tot op voor kort was onderzoek gelimiteerd tot het bestuderen van enkele eiwitten tegelijk. Hedentendage is het mogelijk om biologische processen op een veel grotere schaal te bestuderen. Men kan hier bijvoorbeeld denken aan het bepalen van de totale humane DNA sequentie. Het op grotere schaal bestuderen is ook mogelijk voor de intermediaire

mRNA moleculen.

Gen expressie profilering is een techniek om gen expressie niveaus te bepalen op een grote schaal. De techniek is gebaseerd op het vermogen van DNA om te binden met complementaire sequenties in een proces dat hybridisatie wordt genoemd. Hybridisatie vindt plaats wanneer een streng nucleotiden een andere streng nucleotiden vindt waarmee over de gehele lengte baseparen kunnen worden gevormd. Gen expressie profilering kan worden gedaan met behulp van micro-arrays. Micro-arrays bevatten tienduizenden specifieke DNA sequenties op aparte, bekende locaties (spots). Elke spot bevat een DNA sequentie die een bepaald gen representeert. Wanneer een populatie mRNA moleculen wordt gemarkeerd met een specifieke kleur en vervolgens wordt gehybridiseerd op een micro-array, dan is het mogelijk om te bepalen welke spots zijn gekleurd, en dus welke genen er aanwezig zijn in de mRNA populatie. Wanneer een populatie van mRNA moleculen in een biologisch monster is bepaald, dan spreekt met over het gen expressie profiel. Met deze techniek is het mogelijk om gen expressie profielen van verschillende biologische monsters te vergelijken. Men kan bijvoorbeeld kijken naar een monster van een gezond weefsel en deze vergelijken met een ziek weefsel. Het verschil in gen expressie wordt ook wel differentiële gen expressie genoemd. De micro-arrays die in dit proefschrift voornamelijk worden gebruikt zijn zelfgemaakte, twee-kleuren micro-arrays gebaseerd op mRNA specifieke DNA sequenties van 65 nucleotiden lang (oligonucleotiden). Twee verschillend gekleurde mRNA monsters kunnen tegelijkertijd worden gehybridiseerd op dit type micro-arrays; vandaar de twee kleuren.

De data analyse van gen expressie profielen is een uitdaging. De grote hoeveelheid data vereist een stringente statistische aanpak, terwijl de biologische uitkomst een reductionistische aanpak vereist om de onderliggende cellulaire en moleculaire processen te begrijpen. Om de statistische aanpak te vergemakkelijken, hebben we het experimentele ontwerp bestudeerd (Hoofdstuk 6) . We hebben geconcludeerd dat de data afkomstig van een twee-kleuren experiment afzonderlijk behandeld kan worden. Dit heeft als gevolg dat het experimentele ontwerp meer flexibel is.

Spierweefsel heeft de mogelijkheid om te regenereren na het oplopen van schade. De verantwoordelijke cellen hiervoor zijn zogenaamde satelliet cellen, die normaal in een ongedifferentieerde rustfase verkeren. Door vermeerdering, ook wel proliferatie genaamd, en het veranderen in volwassen spiercellen, ofwel differentiatie, kunnen deze satelliet cellen uiteindelijk spierweefsel herstellen. Deze processen zijn ook actief in spierweefsel dat aangedaan is door spierdystrofie. Desondanks blijft het genetisch defect aanwezig in de nieuw aangelegde spiercellen. Dit draagt bij aan de progressieve natuur van de ziekte. Regeneratie vertoont grote vergelijkingen met processen die actief zijn tijdens de spierontwikkeling, ook wel myogenese genaamd. Door deze processen in meer detail te kunnen bestuderen, hebben we ons gericht op genen die actief zijn tijdens myogenese. Deze experimenten zijn gedaan in gekweekte satelliet cellen (Hoofdstuk 5).

Diermodellen voor genetische ziekten zijn gemaakt door gerichte veranderingen in de DNA sequentie om erfelijke ziekten beter te kunnen bestuderen. Deze veranderingen hebben als gevolg dat ze humane ziekten nabootsen. In de afgelopen decennia heeft de ontwikkeling van deze diermodellen enorm bijgedragen aan het bestuderen van humane ziekten. Een van de

grootste mijlpalen betreft de mogelijkheid om transgene muizen te genereren. Duchenne spierdystrofie is de meest ernstige vorm en leidt tot een vroege dood in patiënten. Het muismodel voor deze ziekte (mdx muis) vertoont, daarentegen, deze vroege dood niet. De onderliggende mechanismen verantwoordelijk voor het overleven zijn nog niet bekend. Ondanks dat het muismodel op dit vlak de humane ziekte niet representeert, kan het model aanwijzingen geven voor de humane ziekte daar andere vergelijkbare symptomen aanwezig zijn. Wij hebben ons gericht op het bestuderen van de regeneratie fase in mdx muizen om te bepalen welke processen betrokken zijn bij regeneratie (Hoofdstuk 3). Zoals verwacht zijn veel genen betrokken bij regeneratie ook actief tijdens myogenese.

Met behulp van gen expressie profilering hebben we getracht om gelijkenissen en verschillen tussen spierdystrofiën te ontdekken (Hoofdstuk 4). Hiervoor hebben we verschillende muismodellen voor spierdystrofie getest. We hebben een groot aantal processen gevonden die overeenkomstig zijn tussen de verschillende spierdystrofiën. Deze processen zijn met name secundaire effecten van de primaire ziekteveroorzakende genen. De huidige hypothese met betrekking tot het ontstaan van spierdystrofie is dat specifieke gen mutaties uiteindelijk leiden tot instabiliteit van spiercellen. Dit heeft als gevolg dat dystrofische spiercellen fragiel zijn en leiden tot een vergrote kans op celdood in vergelijking met gezonde spiercellen. Naast het bestuderen van processen die aanwezig zijn in alle spierdystrofiën, hebben we getracht processen te vinden die specifiek zijn voor bepaalde spierdystrofiën. Deze gegevens kunnen van belang zijn voor diagnostische doeleinden. De hoeveelheid genen die actief zijn in een enkele spierdystrofie is echter erg laag. Dit kan te wijten zijn aan het feit dat we de gen expressie profielen op één tijdpunt hebben bepaald. Verder onderzoek is nodig om meer verschillen te kunnen vinden, en daar wordt op dit moment aan gewerkt.

Samenvattend, met behulp van gen expressie profilering zijn we begonnen met het ontdekken van de vele processen die in spierdystrofie actief zijn. De resultaten beschreven in dit proefschrift dragen bij aan het begrijpen van deze groep aandoeningen, en leiden hopelijk tot de ontwikkeling van therapieën. Daarbij kunnen toekomstige studies van muismodellen aanwijzingen geven omtrent de verschillen in de ernst van de ziekte, het patroon van aangedane spieren en het tijdstip waarop de eerste symptomen ontstaan.

Bibliography

Ellen Sterrenburg, Caroline G van der Wees, Stefan J White, **Rolf Turk**, Renée X de Menezes, Gert-Jan B van Ommen, Johan T den Dunnen, Peter AC ‘t Hoen.
Gene expression profiling highlights defective myogenesis in DMD patients and a possible role for bone morphogenetic protein 4. Neurobiology of Disease. 23(1):228-236 (2006)

‘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)

Vinciotti V, Lui X, **Turk R**, de Meijer EJ, ‘t Hoen PA.
Exploiting the full power of temporal gene expression profiling through a new statistical test: Application to the analysis of muscular dystrophy data. BMC Bioinformatics. 7:183 (2006)

Turk R, Sterrenburg E, van der Wees CG, de Meijer EJ, de Menezes RX, Groh S, Campbell KP, Noguchi S, van Ommen GJ, den Dunnen JT, ‘t Hoen PA.
Common pathological mechanisms in mouse models for muscular dystrophies. FASEB J. 20(1):127-9. Epub 2005 Nov 23. (2006)

Turk R, Sterrenburg E, de Meijer EJ, van Ommen GJ, den Dunnen JT, ‘t Hoen PA.
Muscle regeneration in dystrophin-deficient mdx mice studied by gene expression profiling. BMC Genomics. 6:98 (2005)

R Turk, PAC ‘t Hoen, E Sterrenburg, RX de Menezes, GJB van Ommen, JT den Dunnen.
Gene expression variation between mouse inbred strains. BMC Genomics. 5:57 (2004)

E Sterrenburg, **R Turk**, PAC ‘t Hoen, JCT van Deutekom, JM Boer, GB van Ommen, JT den Dunnen.
Large-scale gene expression analysis of human skeletal myoblast differentiation. Neuromuscular Disorders. 14 507-518 (2004)

PA ‘t Hoen, **R Turk**, JM Boer, E Sterrenburg, RX de Menezes, GJ van Ommen, JT den Dunnen.
Intensity-based analysis of two-colour microarrays enables efficient and flexible hybridization designs. Nucleic Acids Res. 32(4): e41 (2004)

Sterrenburg E, **Turk R**, Boer JM, van Ommen GB, den Dunnen JT.
A common reference for cDNA microarray hybridizations. Nucleic Acids Res. 30 (21): e116 (2002).

Curriculum Vitae

Rolf Turk werd op 23 januari 1975 geboren te Leiden. Hij behaalde zijn diploma voor het Voortgezet Wetenschappelijk Onderwijs in 1993 aan het Rijnlands Lyceum Sassenheim. Van 1993 tot 1999 volgde hij de studie Medische Biologie aan de Universiteit van Amsterdam, alwaar hij zijn doctoraal diploma behaalde.
Aansluitend was hij werkzaam als promovendus bij het Centrum voor Humane en Klinische Genetica, onderdeel van het Leids Universitair Medisch Centrum, waar hij werd begeleid door Dr. Peter-Bram ‘t Hoen, Dr. Johan den Dunnen en Prof. Dr. Gert-Jan van Ommen. Vervolgens werkte hij als bioinformaticus bij het bedrijf Service XS.
Sinds 1 april 2005 is hij werkzaam als postdoctoraal onderzoeker in het lab van Dr. Kevin Campbell, aangesloten bij de Universiteit van Iowa in de Verenigde Staten.