CHARACTERISATION AND FUNCTIONAL ANALYSIS OF NOVEL GENES IN THE STRIATED MUSCULATURE

LRRC39 AND ITS ROLE FOR MUSCLE FUNCTION

INAUGURALDISSERTATION

zur Erlangung des Doktorgrades der Naturwissenschaften
- Doctor rerum naturalium -
(Dr. rer. nat.)

eingereicht am Fachbereich Biologie und Chemie der
Justus-Liebig-Universität Giessen

vorgelegt von
Katharina Jenniches

Bad Nauheim, 2012

Die vorliegende Arbeit wurde am Max-Planck-Institut für Herz- und Lungenforschung, W.G. Kerckhoff-Institut in Bad Nauheim angefertigt.

Erstgutachter:
Prof. Dr. Katja Becker
Nutritional Biochemistry, Interdisciplinary Research Center
Justus-Liebig-Universität Giessen
Heinrich-Buff-Ring 26-32, 35392 Giessen

Zweitgutachter:
Prof. Dr. Dr. Thomas Braun
Abteilung Entwicklung und Umbau des Herzens
Max-Planck-Institut für Herz- und Lungenforschung
Ludwigstraße 43, 61231 Bad Nauheim

Disputation am 29.08.2012

INDEX

LIST OF FIGURES AND TABLES

LIST OF ABBREVIATIONS

Å	Angstrom
aa	amino acids
ADP	adenosine diphosphate
ATP	adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
bp	base pairs
BSA	bovine serum albumine
cDNA	DNA complementary to mRNA
CTRL	Control
DAPI	4’, 6-diamidino-2-phenylindole
DNA	deoxyribonucleic acid
dpf	days post fertilisation
DSHB	Developmental Studies Hybridoma Bank
E	embryonic day
EAA	exon array analyser
ECC	excitation-contraction-coupling
EDTA	ethylenediaminetetraacetic acid
EST	Expressed sequence tag
EtOH	ethanol
FCS	fetal calf serum
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GFP	green fluorescence protein
GO	gene ontology
HEK	human embryonic kidney cell line
HMM	heavy meromyosin
hpf	hours post fertilisation
Ig	Immunglobulin
KO	knockout
kb	kilo base pairs
kDa	kilo dalton
LC-MS/MS	Liquid chromatography-mass spectrometry/mass spectrometry
LMM	light meromyosin
LRR	leucine-rich repeat domain
LRRC	leucine-rich repeat containing
LSM	Laser scanning microscope
MeSH	Medical subject headings
ml	milliliter

µl	microliter
mM	millimol
MO	morpholino oligonucleotide
mRNA	messenger RNA
MW	molecular weight
MyBP-C	Myosin binding protein C
MHC	myosin heavy chain
Myom	Myomesin
NCBI	National Center of Biotechnology Information
nm	nanometer
NMR	Nuclear magnetic resonance
PBS	Phosphate buffered saline
PFA	Paraformaldehyde
RNA	ribonucleic acid
RT-PCR	Reverse Transcriptase Polymerase Chain Reaction
SAGE	serial analysis of gene expression
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
TE	Tris-EDTA buffer
Tris	Tris-(hydroxymethyl)-aminomethan
TRITC	tetramethylrhodamine isothiocyanate

1 ZUSAMMENFASSUNG

Die quergestreifte Muskulatur besteht aus hoch organisierten Strukturen mit dem Sarkomer als kleinste kontraktile Einheit. Um effiziente Muskelfunktion zu ermöglichen, ist ein exakter Aufbau dieser Strukturen essentiell. Die zellulären Prozesse die notwendig sind um den präzise Sarkomerbildung zu ermöglichen, sowie deren regulatorische Abläufe sind bislang nur unvollständig verstanden. Um einen derart komplexen Vorgang im Detail verstehen zu können, müssten alle beitragenden Faktoren bekannt sein. Mit dem Ziel bisher unbekannte Faktoren zu erfassen, wurde ein komplexer Screen entwickelt, mit dem neuartige Gene identifiziert werden können, die in der quergestreifte Muskulatur exprimiert werden und im Laufe der Evolution konserviert wurden und somit höchst wahrscheinlich eine wichtige Rolle für den Muskel darstellen. Die ermittelten Gene besitzen derzeit keine beschriebene Funktion im Bereich der Muskel Funktion. Ein Kandidat dieser neuartigen Gruppe ist *leucine-rich repeat containing protein 39*, kurz *LRRC39*. Ziel dieser Arbeit ist es dieses Gen zu charakterisieren und deren Bedeutung für die Muskulatur zu untersuchen. Mittels Expressionsstudien konnte die Lokalisation von *LRRC39* im Herz-, sowie im Skelettmuskel bestätigt werden. Sie subzelluläre Lokalisation diese Proteins wurde im Bereich der sarcomerischen M-bande detektiert. Um eine mögliche funktionelle Rolle dieses Genes zu adressieren, wurden sogenannte *loss-of-function*-Experimente im Zebrabärblingsmodell durgeführt und resultierten in veränderten herzstrukturen und deformierter Skelettmuskulatur. Diese Beobachtungenweisen auf eine wichtige Bedeutung für die Muskelentwicklung lieferten. Detailierte Untersuchungen der Ultrastruktur des Skelettmuskels mittels Elektronenmikroskopie zeigten strukturelle Veränderungen in der Sarkomerorganisation. Um Einblick in den Entstehungsprozess dieses Phänotyps zu gewinnen, wurden Proteinexpressionsstudien mittels Massenspektrometrie durchgeführt, die drastische Veränderungen substanzieller Proteine der Sarkomerassemblierung zeigten. Um die primäre Rolle von Lrrc39 auf molekularer Ebende zu untersuchen, wurden Protein-Protein Interaktionsstudien durchgeführt, welche die α-helikale Region der Myosin schweren Ketten (MHC) als Bindungspartner identifizierten. Obgleich die Kopfdomäne der Myosine, mit ihrer kennzeichnenden Motorfunktion für die Muskelkontraktion intensiv beforscht wurde, sind die Eigenschaften der α-helikalen Region des Myosins nur wenig verstanden. Die Interaktion von Lrrc39 mit diesem Bereich des Myosins wurde in weiterführenden Experimenten untersucht und weist auf einen möglichen regulatorischen Einfluss von Lrrc39 auf MHC Proteine durch veränderte Phosphorylierung hin. Auch wenn der Mechanismus dieser Interaktion im zentralen Bereich des Sarkomers nicht final geklärt werden konnte, zeigen die Ergebnisse dieser Studie Lrrc39 als wichtigen Faktor in der Sarkomerassemblierung, möglicherweise durch Regulation der Myosin Funktion. Zusammenfassend konnte in dieser Arbeit ein neuartiges Protein des Sarkomers beschrieben werden, welches eine bedeutende Funktion in der Organisation der Sarkomere einnimmt.

2 ABSTRACT

The cross-striated musculature consists of highly organised structures with the sarcomere as the basic contractile unit. To allow efficient muscle function, exact formation of these structures is essential. The cellular processes, which are necessary to enable precise sarcomere organisation and their regulatory sequence of events, are not completely understood. In order to elucidate such complex processes in detail, all involved factors ought to be known. Aiming at the determination of previously unknown factors, a highly organised screening procedure has been developed to identify novel genes with still uncharacterised function in the striated muscle development. These genes were evolutionary conserved and foreseen to be expressed in the striated musculature, which suggested an important role for muscle function. One interesting candidate among them is the leucine-rich repeat containing protein 39 (LRRC39). This study is aimed to characterise this gene and analyses its impact on muscle functions. The data presented in this dissertation confirmed the expression patterns of LRRC39 in the heart and the skeletal muscle in model organisms. To address potential functions of this gene, loss-experiments in zebrafish were conducted. As result, deformations in the heart structure and skeletal muscle tissue were observed to indicate the importance of Lrrc39 for the muscle development of zebrafish. By electron microscopy, the disturbed sarcomere organisation of the zebrafish skeletal muscle was revealed to firmly confirm the functional role of Lrrc39. To elucidate the mechanisms behind the observed phenotypes upon the reduction of Lrrc39, transcriptomics and proteomics techniques were employed. A careful profiling by mass spectrometry was conducted, which showed drastic changes of essential proteins for the sarcomere assembly. Furthermore, subcellular localisation studies and protein-protein interaction analysis were conducted to identify the α-helical tail region of the well-known structure proteins myosin heavy chain (MHC) as interaction partners of LRRC39.

Over the years of researches in the field of muscle development and function revealed the importance of the head domain of MHC proteins in regards to their motor activity for muscle contraction. However, the properties of the α-helical tail region are less well-understood. To address this point, a set of experiments were performed to suggest the possible involvement of LRRC39 to the regulation of MHC proteins through altered phosphorylation patterns of MHC. This suggests a link between the sarcomere assembly and muscle contraction in the context of LRRC39 and MHC proteins. In sum, this study presented a novel protein of the sarcomere with its vital function for the sarcomere organisation.

„Dem Anwenden muss das Erkennen vorausgehen“
“Insight must precede application“
Max Planck

3 INTRODUCTION

The striated musculature is essential for any voluntary movements, as well as for a regular and powerful heartbeat. In order to assure this reliable and high-quality force production, not only the major structural proteins (e.g. actin, myosin and titin) but also proteins to regulate and control the appropriate contraction are needed. Additional proteins are necessary to set the orientation, to give stability by connecting to other proteins, to sense extracellular signals and to transport this information. Furthermore, it is a generally accepted view that muscle fibres do not increase in number (hyperplasia) after birth and therefore, a physiological replacement of the complete fibres seems unlikely. However, in order to maintain their functions, the frequently-used structural proteins have to be replaced constantly and in a timely manner. This requires a well organised turnover of structural proteins within the muscle cells. To maintain such functions, the force-producing muscle requires a complex network of events, which is not yet understood completely. Therefore, functional analysis of involved genes might help to understand the muscle function in detail.

Although many genomes are sequenced and protein coding sequences are known, their function still remains uncharacterised. Data from high-throughput screenings showing evidence for a large number of genes without any documented functions. In the human genome for example, approximately 37% of the genes are functionally not annotated (Wren, 2009). The constantly increasing amount of results from high-throughput technologies showing evidence for a relatively high number of uncharacterised genes, but even if these genes show tissue specificity or changes in their expression patterns in different organs and/or under a particular condition (e.g. hypertrophy), they are often removed from further analysis (Pawlowski, 2008). For this reason, a systematic way to characterise these genes is needed in order to identify novel key players of the tightly regulated processes of the musculature.

The first part of the introduction focuses on the structural and functional role of the striated muscle and illuminates some open research questions. The second part demonstrates one way to identify and annotate novel genes in order to create a better understanding of the complex muscle function.

3.1 Structure and function of the striated musculature

The musculature is an efficient and very precise tool with a complex cytoskeletal network, and from its contractile activity, it is regarded as the force-generating machinery of our body. The functional outcome of a muscle cell is always the same: Force.
In vertebrates, three main force productive cell types can be distinguished: smooth muscle, skeletal muscle and cardiac muscle (Brenner, 2010). Although the development of the heart and the skeletal muscle as well as their localisation and their conductive system are fundamentally different, the basic structures of cardiac and skeletal muscle are highly similar. These two cell types are composed of parallel arranged myofibrils, containing repetitive units of overlapping structural filaments and motor proteins, which leads to a cross-striated appearance at the microscopic level. Over the last decades, researches based on transcription factors and their regulation, have been very successful in elucidating components important for muscle differentiation and fuction. However, the initial process that forms the contractile units is less understood.

3.1.1 Architecture of the striated muscle

Cardiac and skeletal muscle are characterised by their regular cross striation, which can be visualised at the microscopic level. This feature is the result of their behaviour under polarised light. Isotropic regions, the so called I-bands, appear lighter and are alternating with anisotropic structures, called A-bands, which are comparably darker in their appearance (Hanson and Huxley, 1953). As shown in Figure 1, the I-band is divided by the dark appearing Z-line. The A-band contains a centralised lighter region, which is composed of the H-zone with the dark appearing M-band in the centre. The repeated alternations of these structures are aligned longitudinally to the fibre orientation, which lead to the observed cross striation. The muscle fibre consists of bundled myofibrils, which are composed of small repetitive subunits. The smallest functional subunit of the myofibril is the sarcomere, which is defined as the segment between two Z-lines (Ehler and Gautel, 2008). The sarcomere contains the three main protein filament systems: (1) The thin filament, which contains mainly actin, tropomyosin and troponin; (2) the thick filament, which holds predominantly myosins and (3) the giant protein titin that spans the half of sarcomere length from the Z-line to the M-band. Thin and thick filaments align parallel in a conserved hexagonal orientation at the distal regions of the A-band as shown in Figure 1B. This overlap is necessary for the formation of the cross bridges, which form the core for the contraction. The single filament systems are described in the following subsections.

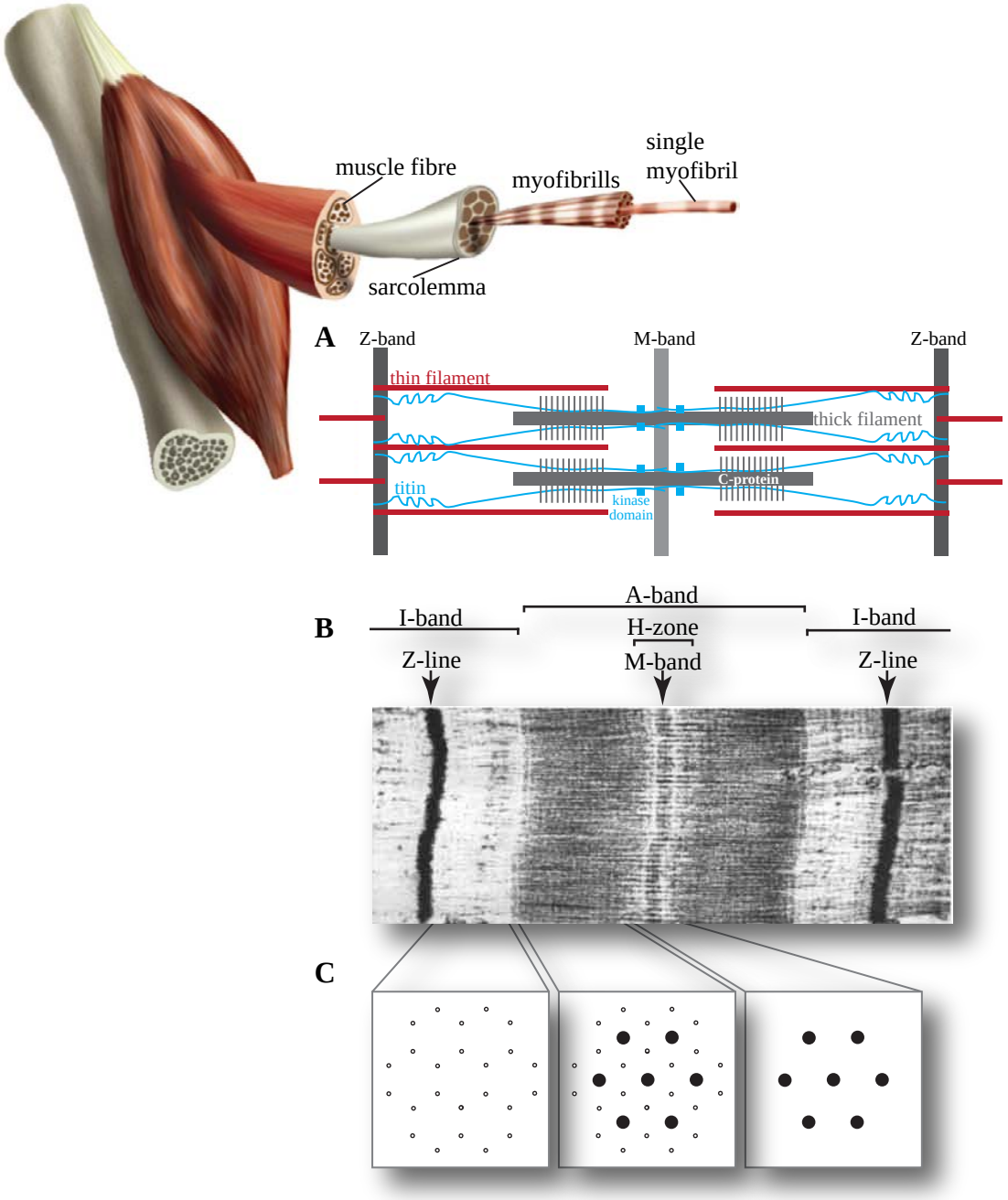


Figure 1: Composition of the striated muscle. Scheme of the skeletal muscle structure (modified from Bloom and Fawcett, 1975; Myhre and Pilgrim, 2012); A: Schematic view of the sarcomere (modified from Gregorio, 1999); C: Electron microscopy image of longitudinal section of a skeletal muscle myofibril showing one sarcomere bordered by two Z-lines (taken from H.E. Huxley 1958); D: Scheme of the sarcomere structure in transversal view showing hexagonal orientation. modified from (H.E. Huxley, 1958).

The thick filament

The thick filament consists predominantly of myosins. This large family consists of approximately 140 myosins and can be grouped by sequence analysis into seventeen distinct classes (Korn, 2000). The different classes generally vary in their properties at the amino-terminal head domain that is involved in diverse cellular processes, such as migration, adhesion, intracellular transport and signal transduction (Krendel and

Mooseker, 2005). The sarcomere contains class II myosins that take up about 40-50% of the total muscle proteins. The members of class II myosins vary mainly in their tissue specificity and ATPase activity (Clark, 2007). Myosin-1 (*MYH1*) as well as myosin-2 (*MYH2*) and myosin-4 (*MYH4*) are predominantly expressed in the adult skeletal muscle, whereas *MYH3* is found in embryonic muscle. Furthermore, *MYH8* is expressed in the perinatal skeletal muscle. The amount and distribution of these genes vary in different muscle cell types (e.g. slow versus fast fibre type) (Allen, 2000). In the adult heart, *MYH6* (also known as α-*MHC*) and *MYH7* (also known as β-*MHC*) are found. The expression patterns and the tissue specificities of these myosins are evolutionary conserved. Alteration in these specifications can lead to pathological changes. The ablation of *MYH6* in mice, for example, was shown to be lethal (Jones, 1996). The homozygous mice died of severe cardiac defects around embryonic day 12. The surviving heterozygous offspring developed dramatic cardiac abnormalities, suggesting that MYH7 cannot compensate the loss of its family member. Similar observations were shown in mice carrying null mutations in distinct adult skeletal muscle myosin isoforms. These mice have a decreased muscle mass and show substantial alterations in the skeletal muscle morphology (Allen, 2000). Consequently, this class of myosins appear to be functionally diverse, rather than functionally redundant.

Although species and fibre type specific differences of the musculature exist, the molecular structure of the class II myosin is highly conserved. Members of the class II myosin are characterised by four myosin light chains (each around 20 kDa), which are further subdivided into two essential- and two regulatory light chains, and by two myosin heavy chains (MHC) of around 200 kDa. The heavy chains form globular heads at the amino-terminal regions of myosin. This head region has a dimension of 165x 65 x 40 Å and acts through their actin binding and ATPase activity to provide the motor of muscle contraction. The two heavy chains form a α-helical coiled-coil structure at the carboxy-terminal region of 1600 Å in length, which is termed the rod domain. A repetitive sequence of seven amino acids within the myosin heavy chain builds a hydrophobic core, which is surrounded by alternating positive and negative charged chains at the surface of the structure. An additional repetitive region of 196 amino acids creates one turn of 14.3 nm in length. These arrangements cause a stable coiled-coil structure of the two myosin heavy chains with a diameter of 20 Å (Sjostrom and Squire, 1977a; McLachlan and Karn, 1982). Myosin can be further separated by enzymatic digestion in the heavy meromyosin fragment (HMM) and the α-helical rod part termed light meromyosin (LMM), as shown in Figure 2A. The HMM part contains the two heads (S1) and a part of the α-helical myosin heavy chain, which is called S2 domain. The LMM fragment, which contains the main part of the rod domain is essential for the stable construction of the myosin filaments. The three fragments are connected through so-called hinge regions at the S1/S2

and at the S2/LMM junction. These joints are important for the orientation of the myosin heads, which bind to the thin filament (H.E. Huxley, 1969).

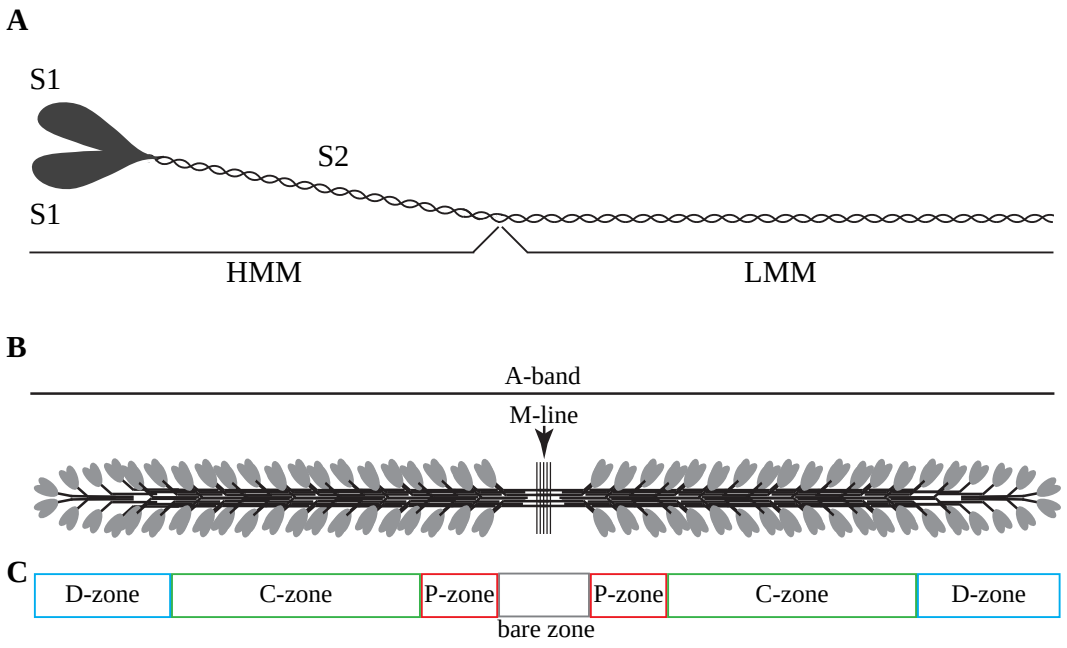


Figure 2: Scheme of the class II muscle myosin. A: Hexameric myosin molecule and their defined regions: HMM (heavy meromyosin) with S1 (the myosin heads) and S2 (first part of the a-helical heavy chain) and LMM (light meromyosin) **(**based on Zoghbi, 2008). B: Schematic view of the thick filament, representing the A-band and indicates the different defined areas of the myosin filament. C: From the M-line that marks the centre towards the distal ends, the regions are further divided into bare zone, P-zone, C-zone and D-zone (based on Sjostrom and Squire, 1977b; McLachlan and Karn, 1982).

Under normal physiological condition, myosin II molecules arrange spontaneously in filaments and form later the A-band of the sarcomere (Srikakulam and Winkelmann, 2004). During the myosin assembly, several proteins were identified to be involved in the organisation of this process (reviewed in Wang, 1998; Ehler and Gautel, 2008). Before the actual assembly initiates, myosin molecules need to be folded to their hexameric structure mentioned above. During the folding process, a complex of chaperones were identified to play a major role, including *HSP70* and *HSP90* (Srikakulam and Winkelmann, 2004) and a member of the UCS-domain containing family, *UNC-45*. These proteins interact with each other to regulate proper myosin arrangement. The signalling pathways of this complex are yet to be elucidated. However, loss of function studies resulted in abnormalities in myofibrillar organisation, including myosin aggregation, which support the important roles of these chaperones (Barral, 2002; Landsverk, 2007; Wohlgemuth, 2007). Once the myosin molecules are properly folded, they align to a bundled structure, with their C-termini associate with each other and build the final filament as illustrated in Figure 2B.

The anti-parallel organisation in the centre of the filament leads to an overlap of rod domain of the myosins, which forms the M-region. This region, also known as “bare zone”, lacks myosin heads due to the bipolarity of myosins, which are targeting towards opposite directions. The central part of the bare zone shows electron dense structures, which appear in electron microscopy as a dark line of approximately 100 nm width. The M-band serves as the primary anchor point for the thick filament. The thick filament contains 294 myosin molecules and has a precise length of 1.59 µm (McLachlan and Karn, 1982). The different defined areas within the thick filament are illustrated Figure 2C. The bare zone measures around 1600 Å in length (Al-Khayat, 2008). As mentioned above, the bare zone contains only the myosin rod part; thus, it is unable to form any myosin cross bridges. The first myosin heads appear with the start of the P-zone, which is neighbouring the bare zone. At a distance of 2040 Å, the C-zone starts, which is followed by a cylindrical ending area called the “D-zone” (Sjostrom and Squire, 1977b). These clearly defined zones are based on detailed structural analyses and not on their functionalities. Nevertheless, expression analysis for myosin binding protein-C (MyBP-C), which is located in the C-zone, verified this classification (Craig and Offer, 1976; Gilbert, 1999; Winegrad, 2003). Myosin binding protein C (*MyBP-C*) is localised in seven to nine stripes in the C-region of the A-band (Luther, 2008). After targeted gene knockout in mice, lack of *MyBP-C* resulted in dilated cardiomyopathy (McConnell, 1999) or hypertrophic cardiomyopathy, sudden death and reduced exercise capacity (Harris, 2002). The severity of the observed phenotypes proves the importance of these zones for the muscle. However, the mechanism of *MyBP-C* function is not understood in detail.

In the central part of the thick filament, additional proteins are identified. A detailed microscopic analysis of the M-band revealed five prominent M-lines with a distance of about 22 nm. The number of lines can vary dependent on the species and fibre type. It was suggested that the line density correlates with the physiological performance of different muscle types (Sjostrom and Squire, 1977a; Thornell, 1987). The appearance in cardiac sarcomeres, for example, correlates with the beating rate in different species. In extreme cases, like in embryonic heart and extraocular muscle, the sarcomere do not show an electron dense M-band (Pask, 1994; Smolich, 1995). Most common form of the M-band is the presence of five separate lines, defined as M1 in the centre, surrounded by M4, M4’, M6 and M6’. Although the lines are clearly detectable, corresponding proteins building these lines are not fully identified. Until now, the muscle-type creatine kinase (*MM-CK*), which is one of the key enzymes in the cellular energy supply, and myomesin (*MYOM)* have been found in this defined region (van der Ven, 1996; Hornemann, 2003). The *MM-CK* was detected by immunhistological stainings to be localised in the M4-lines (Wallimann, 1983). The myomesins were initially described by Eppenberger and colleagues in 1984. They identified a 165 kDa protein, which they named M-protein (*MYOM2*) and a 185 kDa protein called myomesin (*MYOM1*). Both proteins are localised

in the M-band (Grove, 1984). Furthermore, it was shown that myomesin dimerises and crosslinks myosin filamtents at the M-band (Lange, 2005b). This connection between myomesin and myosin forms a solid structure at the M-band and is thought to stabilise the thick filaments (Agarkova, 2003). A similar interaction was observed between skelemin (EH-myomesin) and M-protein, which are later identified to be homologs of myomesin (Masaki and Takaiti, 1974; Steiner, 1999).

Due to the lack of known annotated proteins in the M-band, the functional characterisation of this specific region still remains unclear. Based on microscopic observations, this structure seems essential for regular packing of the myosin filaments (Knappeis and Carlsen, 1968). This aspect is supported by a novel three-dimensional model by Lange and colleagues, which was developed based on interaction studies of myomesin, titin and myosin (Lange, 2005b). The M-band serves not only as an anchor for myosin filaments but participates as a structural component in the maintenance of the A-band lattice (Luther, 1981). During active muscle movements, it contributes to the distribution of mechanical stress across the lattice. In the recent years, an interest into uncovering the function of the M-band has increased. Additional constituents of the M-band were identified to interact with the thick filament and thereby regulating muscle function. These include MURFs, a group of muscle specific RING finger proteins. Through their ubiquitin ligase activity, they are implicated in diverse processes, including ubiqutination and transcriptional regulation. In the same context the protein Nbr1 (*neighbor of BRCA1 gene 1*) was identified (Perera, 2011). Other M-band related proteins are Obscurin (Fukuzawa, 2008), the Muscle LIM protein (*MLP*) (Ehler, 2001; Lange, 2005a) and Calpain 3 (Sarparanta, 2010). Calpain 3 (also known as p94) is a calcium-dependent protease (Roperto, 2010) that interacts with the M-band region of titin and is associated with muscular dystrophy (Y. Ono, 1999; Haravuori, 2001). The latest results suggest that this structure in the centre of the sarcomere is not only important for the stability but can actively regulate sarcomere function. These additionally identified proteins are currently investigated further to provide a better understanding of the organisation and function of this filament.

The thin filament

The main component of the thin filament is actin. This well-known structural protein has a molecular weight of 42 kDa and is ubiquitously expressed. It is involved in diverse cellular functions, such as motility and cytokinesis. In the striated muscle, actin filaments are comprised of globular actin molecules (G-Actin), which polymerise to filamentous actin (F-actin) and form polar, α-helical structures. Likewise myosin assembly, the chaperones Hsp70/ Hsp90 are important for proper folding and formation of actin. Additionally, the proteins GimC (Siegers, 1999) and TRiC (Grantham, 2002) are

involved. However, the precise mechanism of actin folding and its assembly is unclear. In isolated conditions, actin is able to self-assemble and forms filamentous structures, the F-actin. However under physiological conditions, many proteins are associated with this process and regulate thin filament assembly. These proteins have been described to add or exchange actin monomeres either at their C-terminal end, the *barbed end* or at their N-terminal, the *pointed end* (S. Ono, 2007). Tropomodulin (*TMOD*) binds the pointed ends and thereby protects the thin filament from polymerisation and accordingly depolymerisation (Broschat, 1990; Yamashiro, 2008). Other proteins at the same actin terminus are Leiomodin (*LMOD*), which leads after inhibition to severe changes in actin filament assembly (Chereau, 2008) and cofilin (*CFL*) (Van Troys, 2008). On the other site of the structure, *CAPZ* (capping protein muscle Z-line) binds to the barbed ends and inhibits actin polymerisation (Pappas, 2008). Expression analyses in cultured chicken cardiomyocytes identified an additional protein, N-RAP, to regulate the linkage between two barbed ends (Carroll, 2004). Through the binding of these proteins the length of actin filaments is controlled (reviewed in (Cooper and Schafer, 2000).

In its final form, actin filament is ~1 μm in length and has a diameter of approximately 7 nm. On its surface, myosin binding sites are positioned in order to connect thin filament to the thick filament. These sites are present in an interval of 2.7 nm. For this actin-myosin interaction, the actin-binding proteins tropomyosin (Tm) and troponins (Tn) are essential. Tm consists of two α-helical chains that integrate between the grooves of the α-helial actin filaments (Gordon, 2000). In this position, Tm stabilises the filament structure but also blocks the myosin-head binding sides on the surface of the thin filament. Through calcium-dependent activation of the nearby troponin complex, which consists of three subunits (TnI, TnC and TnT), Tm is shifted in its position and exposes the myosin binding sides (Herzberg and James, 1985). This important movement permits muscle contraction and will be described in the later section. Apart from the crucial role in contraction, the thin filament is essential for the sarcomere formation. The barbed ends of the polar structure are bound to α-actinin and crosslink at the Z-band. The rest of the filament extends with the pointed end in the direction of the M-band. Through intermediate filaments like desmin (*DES*) and Merosin, and adapter proteins like dystrophin (*DMD*) and vinculin (*VCL*), which are positioned at the Z-band, actin filaments are connected to the plasma membrane and neighbouring filament systems (Winder and Ayscough, 2005). The importance becomes evident by disruptions of these connections as seen in muscular dystrophies such as Duchenne- or Becker's muscular dystrophy, which cause muscle weakness and to an increasing degree degeneration of myofibres (Carmignac and Durbeej, 2012).

The giant protein titin

Titin, formerly known as connectin, is the largest known polypeptide. Although it is not a filamentous structure, it connects thin- to thick filaments, and is therefore regarded as the third filament system of the muscle. This giant protein contains around 30 000 amino acids and has a molecular weight of 3-3.7 MDa. Due to the presence of its isoforms, there exist a variety in weight and length of titin (Maruyama, 1977). Titin localises in parallel to the thin and thick filament and covers half of the sarcomere length from Z-line to M-band. It contains several interaction sites to sarcomeric proteins, such as actin, myosin and *MyBP*-C. The N-terminal end of titin is fixed to the thin filament at the Z-line through alpha-actinin, whereas the C-terminus of titin binds to the thick filament as well as to myomesin and thereby anchors at the M-band (Labeit, 1997). The A-band region of titin consists mainly of repetitive immunglobulinlike (Ig) and fibronectin-typIII (Fn3) domains. The I-band region, on the other hand, shows Ig-domains, which is termed "PEVK-segment" that consists of prolin (P), glutamate (E), valin (V) and lysine (K). In this region, variations are found, which influence the muscle elasticity. In addition to the PEVK segment, a second domain exists in the I-band part of titin, which modulates muscle elasticity called the "N2-region". In this region, there exist three alternative spliced forms: N2A, N2B and N2AB. The N2B variant is regarded as the stiffest form of all three isoforms and, together with N2AB, it is exclusively expressed in the heart (Voelkel and Linke, 2011). The combination of these domains lead to the function as a molecular spring that defines the sarcomeric stiffness, and thereby regulates passive force through flexible extension of the described domains (Linke, 1996; Linke, 2000).

The C-terminal region of titin, which is localised parallel to the A-band, consists mainly of fibronectin typ III (Fn3)-repeats and Ig-domains (Labeit, 1992), which interact with myosin and MyBP-C (Freiburg and Gautel, 1996). It was proposed that integration and localisation of these proteins are regulated through titin (Gregorio, 1999). Moreover, multiple protein binding sites are known at the C-terminus of titin, which localises in the M-band. In this region, oppositely directed titin molecules overlap to each other to form a dimer (Obermann 1996). Through the direct binding to the structural protein myomesin, this dimerised structure is stabilised. At the M-band region additional binding sites exist for Calpain 3 (Kinbara, 1997), FHL2 (Lange, 2002), Nbr1 (*Neighbor of Brca gene*) and p62 (*sequestom1*) and MURF (*muscle-upregulated RING finger protein)* (Gregorio, 1999). MURFs were shown to regulate transcriptional processes, and thereby act as important trigger for signal transduction (Arya, 2004; Lange, 2005). Thus, it is speculated that the cellular processes involving titin are regulated not only at the protein level but also at the gene level. Furthermore, Murf1 binds to titin at the highly evolutionary conserved kinase domain (Gautel, 1995). This kinase domain is positioned at the periphery of the M-band, and its structure is similar to myosin light chain kinase (MLCK). The classical MLCKs phosphorylate the regulatory myosin light chains at the head domain and thereby

myosin orientation (Kamm and Stull, 2001). However, the function at the C-terminal end by the titin kinase is currently not well understood. Detailed *in vitro* experiments revealed that titin kinase phosphorylates the protein telethonin (*TCAP*), which is located at the Z-band and crucial for sarcomere assembly (Zhang, 2009) The interaction of these proteins, which are positioned in two different regions of the sarcomere, indicated that titin kinase and *TCAP* are in close proximity, which may occur during reorganization of the cytoskeleton (Mayans, 1998).This suggests a function during sarcomere assembly.

3.1.2 Sarcomere assembly

The sarcomere assembly describes the process in which the thin- and the thick filament systems are combined to form sarcomeres, which than built up the myofibril. Actin- and myosin filament, titin and many of their associated proteins were intensively investigated. Several mouse models have underlined their importance for muscle function, and human diseases could be linked to specific mutations in sarcomeric proteins. Still, the process how these structures are formed initially is rather unknown. Evidence for an evolutional conservation is given by the highly consistent shape of the sarcomere of approximately 2µm length and 1µm in diameter. Regardless the size of an organism, the size of the sarcomere and its structures stays constant among species, ranging from Drosophila and zebrafish to human (Page and Huxley, 1963; Brenner, 2010).
Over the past years, several models have been proposed to explain the sequence of events that lead to a fully functional myofibre. The current understanding is mainly based on *in vitro* studies, which describe the formation of the initial structures in the cell periphery. Stress-fibre like structures (Schmidt and Hall, 1998) or small units of actin filaments with premature Z-bands (I-Z-I bodies) were observed as the first organised structures that are regarded as the premyofibrils (Rhee, 1994; Sanger and Sanger, 2008). These actin-containing structures accumulate and connect to the plasma membrane most likely through integrin adhesion sites (Pardo, 1983; Sparrow and Schock, 2009).

As illustrated in Figure 3, the cytoskeletal actins accumulate and integrate α-actinin at the overlapping N-termini, which form a premature Z-band. The fusion of α-actinin was observed in living cells by fluorescent labelled protein (McKenna, 1986) and show a first diffuse plain distribution which then organises to small regular lines. Observation in chicken myotubes confirmed this localisation and suggested dynamic regulations during the alignment of myofibrils (Dabiri, 1997). During the assembly capping proteins, like N-RAP and Tmod are bound to the terminal ends of premyofibrils to regulate the length and stability of filamentous actin (Gokhin and Fowler, 2011). Additionally, non-muscle myosins (NMM) integrate in the central region of the thin filament. How NMM incorporates in this structure is not clear, but *in vitro* study have shown that phosphorylation in their tail regions regulate the motility of this myosins, which effected the early filament assembly (Sabry, 1997). At this stage of the sarcomere assembly, an alternating pattern is visible, which is caused by the NMM and the premature Z-bands. titin associates with Z-band proteins and align in a stepwise fashion along the filament (Ehler and Gautel, 2008). This formation of titin is thought to regulated the spacing between actin- and myosin filaments (Whiting, 1989). The binding to the N-termini of actin in the early phase of the assembly, suggested that titin establishes the formation of the sarcomere. Various knockout experiments and transgenic models support an essential role of titin during sarcomere assembly (Person, 2000; van der Ven, 2000). In these models impairment of

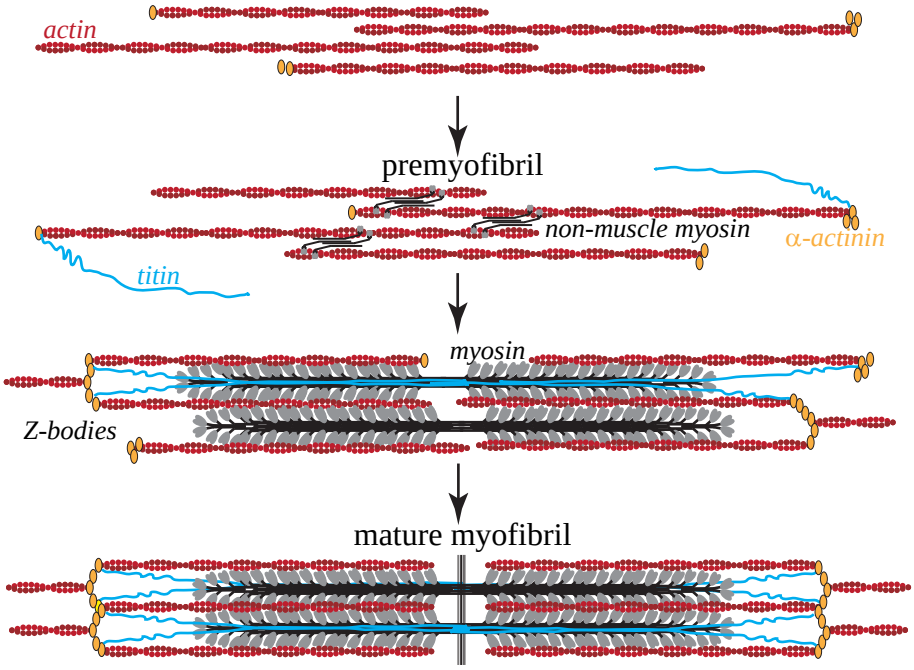


Figure 3: Modell of the sarcomere assembly. The assembly of sarcomere starts at the periphery of the muscle cell, where cytoskeletal actins accumulates and form the premyofibril. Actin-binding proteins and non-muscle myosin integrates in the forming filament. The premature Z-line (Z-bodies) is formed with titin binding at the N-terminus. Capping proteins regulate the length of the aligning actin filaments and muscle myosins exchange the non-muscle myosins. Additional proteins are incorporated and form the Z-line and the M-band, respectively. By contraction the structure linearize and forms the final sarcomere (modified from Sparrow and Schock, 2009)

the thick filament organisation was prominent, which supported the evidence that titin adjust the spacing between Z-bands and regulate the incorporation of myosin filaments and proposed that myosin filaments assemble timely after the formation of the initial cytoskeletal scaffold (Rhee, 1994). However, it remains unclear how titin aligns before its C-terminus can bind to the M-band. First when the NMM are exchanged by muscle class II myosins the final composition of the sarcomere is completed. To these myosins, M-band proteins i.e. myomesin integrate and build the M-band; the C-terminal part of titin associates. However, the role of the M-band for sarcomere assembly is not finally clarified. As the M-band serves as anchor for the C-terminal titin, which modulate muscle elasticity, it may counteract force imbalances and improve stability (Agarkova, 2003). This completes the scaffold and results in the mature sarcomere.

Apart from the sequential integration of sarcomeric proteins, contraction seems to be crucial for the uniform arrangement. It was an interesting finding that muscle cells assemble and develop better after electric pulse stimulation, which stimulated contractility (Fujita, 2007). Further, it was shown in cardiomyocyte cultures, inhibition of the contraction by different chemicals leads to disorganisation of the sarcomeres. As soon as the reagents were removed, the sarcomeres recover and show normal assembly (Simpson, 1993).

endurance performance. On the contrary, the fast twitch (type II) fibres have -in general- a higher glycolytic capacity. They utilize oxidative as well as anaerobic metabolisms for energy production (Dubowitz and Pearse, 1960). These fibres display a higher maximal shortening velocity and are faster fatigable (e.g. *Musculus extensor digitorum longus*).
To add even more complexity, different fibre types vary in their gene isoform composition. It is reported for MHC that each isoform show slightly different rates of cross-bridge cycling, with results for example in a different shortening velocity (Bottinelli, 2001). This may be explained by variations in size of the MHC lever arm, which allows larger shifts during the contraction process (Warshaw, 2000). Further, the troponin subunit C (TnC) exists in different isoforms, which vary in their Ca^{2+} sensitivity (Veigel, 1999). Due to the importance of the different cellular processes involved, it seems likely that additional genes exist to regulate or modify muscle function.

3.2 Screening for novel striated muscle genes

As explained in the previous section, the force production involves several diverse processes; and it forms a very complex network. Since the current knowledge about how such network forms is rather limited, further understanding through annotating of novel genes might shed a light on elucidating the detailed processes and linking the interactions among the different cellular processes. The existence of unknown factors becomes obvious by the calculated number of the “uncharacterised genes” shown in Table 1. Although the protein-coding sequences are known, the functional annotations of a large number of genes across species are still missing. In human, a comparatively small number of protein-coding genes have no publication (15.9 % of total genes). In comparison, the percentages of uncharacterized genes to the total number of genes are much higher in model organisms: 42.2 %, 67.4 %, 81.8 % and 74.5 % in mouse, rat, chicken and zebrafish, respectively. Such a high number of uncharacterised genes is surprising as years of extensive usages of high-throughput methods (e.g. microarrays, serial analysis of gene expression (SAGE), deep sequencing) have generated large lists of genes; however, there is a very little progress on giving functions to such identified genes.

Organism	Number of genes	Without publication
Human	20272	3233 (15.9 %)
Mouse	28766	12148 (42.2 %)
Rat	27506	18526 (67.4 %)
Chicken	17309	14156 (81.8 %)
Zebrafish	32032	23869 (74.5 %)

Table 1: Number of uncharacterised genes. The number of protein-coding genes for each organism (human, mouse, rat, chicken and zebrafish) is shown, followed by their publication record (modified from NCBI database). Publications that described more than 100 genes were excluded for this calculation (updated march 2012).

gene annotation was a motivation behind to combine available databases and screen for these unknown genes to identify new components of the striated musculature.

3.2.1 Databases and their yield of *in silico* data

To screen for uncharacterized genes, the following databases and data sets were utilized. UniGene (http://www.ncbi.nlm.nih.gov/unigene) (Wheeler, 2003) is a database of non-redundant sets of transcript sequences. Each entry is classified under the same transcription locus (gene or expressed pseudogene). By filtering these records for a specific tissue results in a set of genes, genes with a preferential expression for a certain tissue (e.g. heart) were identified. With the assumption that evolutionary conserved genes are likely to have an important function, the filtered genes of different organisms were combined using the HomoloGene database (http://www.ncbi.nlm.nih.gov/homologene) as a source of information about homologs. In order to filter for uncharacterised genes, literature

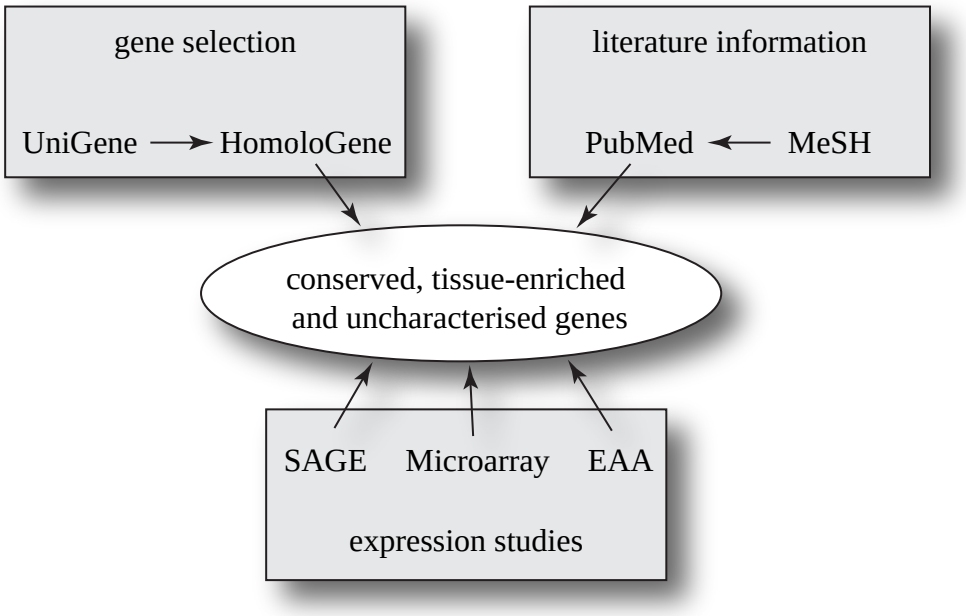


Figure 4: Scheme of database organisation. The combination of publically available databases for gene selection, literature information and expression studies allows to search for conserved, tissue-enriched as well as uncharacterised genes. MeSH (medical subject headings); SAGE (serial analysis of gene expression); EAA (exon array analysis).

information provided by the PubMed database (http://www.ncbi.nlm.nih.gov/pubmed/) was integrated. When a gene is characterised, results are usually published in scientific journals. Using this common practice as a basis for measurements, a gene with only one or no publication is regarded as “uncharacterised”. To make more educational filtering, an additional information was obtained through the Medical Subject Headings (MeSH) (http://www.ncbi.nlm.nih.gov/mesh) vocabulary system to index articles for a specific

topic, or here the target tissue. The systematic combination of these databases yields in screening for uncharacterised, evolutionary-conserved genes with tissue-enriched expression (Uchida, 2009). To add further information, data from high-throughput studies were integrated to obtain an overview of the expression patterns of uncharacterised genes. The following datasets were utilized: SAGE, microarrays and expression data from exon array analysis (EAA) that identified tissue-enriched alternative splicing isoforms (Gellert, 2009; Gellert, 2011).

Through this highly organized screening procedure (Gellert, 2010), a number of novel genes were identified. The next challenging step was the initial characterisation of these candidates. In order to test whether these candidates have an important functional impact, an appropriate model system is required.

3.2.2 Zebrafish as model system for muscle research

The zebrafish (*Danio rerio*) serves as a model organism in this study. This tropical freshwater fish belongs to the minnow family (*Cyprinidae*) and has become a widely used model organism in developmental biology and molecular genetics over the last years. The genome of the zebrafish is sequenced (www.sanger.ac.uk/Projects/D_rerio/) and shows evolutional conservation to humans, especially regarding the early vertebrate development. For example, the keysteps in the heart development, from the migration of precardiac cells to the looping process and the chamber differentiation, exist in equal measure in human, mouse and zebrafish (Stainier, 1993). Moreover, the developmental processes are very rapid compared to the mouse (*Mus musclulus)*. Further, physiological processes in zebrafish and human are remarkably similar and may be even better for comparison than that of mouse, due to the fact that the murine heart beats around 7 to 10 times faster than the human heart (Dahme, 2009), whereas those of zebrafish are more similar to those of human (Arnaout, 2007; Chan, 2009).
The zebrafish as a model organism offers a number of advantages, such as the transparency of the eggs that allows simple access for imaging, and the possibility of live monitoring of the developing organs. Furthermore, an efficient breeding results in hundreds of eggs per mating to superb number of specimens compared to any other model organisms. Furthermore, the external development of the embryo offers an easy access for genetic manipulation *in vivo*. One possibility for gene manipulation to analyse the function of a specific gene is to block its expression and investigate the thereby formed loss-of-function phenotype. To achieve this functional assay, gene specific oligonucleotides that bind to the complementary strand of the mRNA target gene can be used. When such oligonucleotides are present in the cell, ribosomes cannot bind for translation, which leads to specific inhibition of the target protein. For the usage of this antisense technology, Morpholino-phosphoroamidate antisense oligonucleaotides (MOs) are widely used in the zebrafish

research (Nasevicius and Ekker, 2000). These synthetic molecules copy the RNA structure and contain a morpholine ring instead of a ribose ring that binds the specific bases for the target gene. These morpholine rings are connected through phosphoroamidate, instead of phosphate as in the case of RNAs, which is not negatively charged and therefore, less prone to non-specific interactions with other cellular components (Stirchak, 1989; Heasman, 2002). This organisation allows for a standard base pairing. Furthermore, this construction is not recognised by the enzyme RNase H, an endonuclease that cleaves RNA, and thus MOs are very stable structures (J.E. Summerton, 2007). By the design of the specific target sequence, MOs can either interfere with translation or splicing of the target mRNAs; thereby gene expression is inhibited. For zebrafish manipulations, MOs are injected once at an early developmental stage, which this causes the dilution of the MO concentration during the growth of the embryo; this leads to transient knockdown of the gene.

As in any other transient *loss of function* assays, the specificity and the off-target effects through the injection of this artificial substance exist (Eisen and Smith, 2008). Thus, careful controls for these experiments are necessary. Taken all the above facts together, this technique offers an versatile way to block gene expression for loss-of-function studies. For this reason, the zebrafish animal model was used in this study to rapidly assess whether the predicted candidate genes play an important role in the muscle development or not.

3.3 Aim of the study

The striated musculature requires a tight cooperation of diverse processes to regulate its main function, which is the muscle contraction. The assembly of actin- and myosin filament form the basis of the sarcomere, which acts as smallest contractile unit of the striated muscle tissue. Their filament assembly, a controlled elongation and precise alignment within the sarcomeres as well as the connection to another and to the plasma membrane are essential for muscle function. The different processes depend on each other and need to be discussed in context rather than focusing on single aspects. However, the knowledge regarding these essential processes and their potential crosstalk is relatively limited.

One possibility to elucidate the lack of these connections is that further, still unknown factors are involved. In order to understand potential interactions between these involved processes requires the functional analysis of novel genes. Through a highly organized screening procedure a number of new genes were identified, which are foreseen to be evolutionary conserved and expressed in striated muscle tissue (Uchida, Jenniches 2009; Gellert, Jenniches 2010). These genes might be involved in muscle specific processes and their functional annotation might shed some light on the complex field of muscle function.

The primary aim of this study was the initial characterisation of one of these genes, namely *LRRC39*. Further aims of this work were:

- To evaluate its expression patterns and to define the subcellular localisations
- To investigate its functional importance by *loss of function* studies in zebrafish
- To determine its interaction partner and analyse its influence on muscle function.

4 MATERIAL AND METHODS

4.1 Molecular biology methods

4.1.1 Reverse transcriptase PCR (RT-PCR)

For the expression analysis in adult mouse tissue, RNA was isolated from fifteen different mouse organs: brain, eyes, lung, heart, liver, kidney, spleen stomach, small intestine, colon, white adipose tissue (WAT), brown adipose tissue (BAT), testis and uterus. The preparation of the tissues was done according to the Current Protocols in Molecular Biology: Tissue Collection for Systematic Phenotyping in the Mouse. The samples were collected, washed in ice-cold PBS (137 mM NaCl, 2.7 mM KCl, 4.3 mMNa_2HPO_4-$2H_2O$, 1.4 mM KH_2PO_4) and snap frozen in liquid nitrogen. The homogenisation of a tissue was conducted using Polytron in the presence of TRIzol reagent (Life Technologies, Darmstadt). After purification of RNA by following the protocol of TRIzol reagent, two µg of the purified RNA were reverse transcribed using SuperScriptII First-Strand Synthesis System (Life Technologies, Darmstadt) to synthesise the cDNA according to manufacturer's protocol. First-strand cDNA was diluted 1:5 in ddH2O. All reactions were carried out in 20 µl volume including:

2 µl	Taq polymerase buffer (10x)
1 µl	50 mM $MgCl_2$
1 µl	dNTPs (10 mM)
1 µl	primer mix (10 µM)
0.2 µl	Taq polymerase (Eppendorf)
1 µg	cDNA template

The following PCR reaction condition was used:

Step	Temperature	Time	
Initial denaturing	95 °C	2 min	
Denaturing	94 °C	30 sec	} x cycles
Annealing	58-62 °C	30 sec	
Extension	72 °C	1 min	
Final extension	72 °C	7 min	

The amplified PCR products were separated in 2% agarose gel, diluted in TE-buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0), containing ethidiumbromide and were visualised under UV light. The band size of the products was controlled by standard marker (GeneRuler™ DNA Ladder Mix, Fermentas). Specific primers for murine Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) and beta-actin were used to standardise RT-PCR experiments.

Table 2: RT-PCR primers. This table lists the primers used for RT-PCR analysis. Targeted genes are shown in the first column. The primer name, which is given according to the target exon (E) or exon junction, is shown together with the sequence for forward and reverse primers, respectively. The last column shows the expected PCR-product size for the corresponding primer set.

gene	forward primer	reverse primer	product size
*m*Lrrc39	E3/4 GGAAGACTCTCCCTTCTTCT	E6/7 CTTCGCATTCTGTCTAGTGA	454 bp
*m*Lrrc39	E3/4 GGAAGACTCTCCCTTCTTCT	E7/8 CCTCAGATTCGTCATTTCTT	599 bp
*m*GAPDH	E 5 ACCACAGTCCATGCCATCAC	E 5 CATGCCAGTGAGCTTCCCGT	168 bp
*m*β-actin	E 4 CAACGAGCGGTTCCGATG	E 4/5 GCCACAGGATTCCATACCCA	67 bp
*z*Lrrc39	E3/4 GTGGAAGACTCTTCCTCCAG	E7 CTCAAAACTAGCGTGTCCAG	546 bp
*z*GAPDH	E4/5 CTTCCTTCCTGGGTATGGAA	E5 CAATGCCAGGGTACATGGT	120 bp
*z*β-actin	E3/4 GAGCACCAGGTTGTGTCCAC	E4/5 CTCATTGTCATACCATGTGAC CAG	138 bp

For expression analysis in zebrafish (Tübinger strain), total RNA from different developmental stages were collected: 12 hours after fertilisation until day seven. The RNA was prepared using TRIzol Reagent and transcript to cDNA using SuperScriptII First-Strand Synthesis System as described above.

The DNA sequences for all used genes are accessible via Ensembl database (http://ensembl.org). For sequence analysis, the SeqBuilder software from DNASTAR lasergene 8 was used. For the designing of primers, Primer3 (http://frodo.wi.mit.edu/) was utilized. All primer sequences were tested for their specificities using BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) or BLAT for zebrafish sequences.

4.1.2 Generation of tagged fusion proteins

The fusion proteins are constructed for the expression analysis in cells. They allow observation of the subcellular localisation without immunostaining. Therefore the sequence for *Lrrc39* and *Myh1* (forward primer: GCAGAGGAGCGATGCGAC; reverse primer: GAGCTCAGATTTTTCCTGCTCAATTTG; product size: 1920 bp) were amplified by PCR and cloned into pGEM®-T Easy Vector System (Promega, Catalog #A1360). The verified sequence was subcloned in th following three vectors:

1. pEGFP-C1 (Clontech, Catalog # PT3028-5), which contains a N-terminal coupled eGFP
2. pEGFP-N3 (Clontech, Catalog #PT3054-5), which contains eGFP at the C-terminal end of the target protein
3. pcDNA™3.1/V5-His A (Life Technologies, Catalog # V810-20) containing C-terminal V5 epitope and polyhistidine tag

The insert was enzymatically cut by *BamH1* and *Sal1*, and ligated into the *Bgl2* and *Sal1*sites of pEGFP-C1 and –N3. The insert for pcDNA 3.1/V5-His was prepared using *BamHI* and *Sal1*, the vector with *BamH1* and *Xho1*, respectively. Insert and final vector were ligated with T4-DNA-Ligase (Promega) over night at 16 °C. For the transformation of the DNA in electro-competent cells (XLblue), GenePulserTM (Bio-Rad) was used. Cells were resuspended in 200 µl LB-medium (Lennox, ROTH) and plated on appropriate LB-antibiotic plates and incubated at 37 °C overnight. The bacterial cells carrying the transformed plasmid were antibiotic resistant and grew as single colonies, which were picked and inoculated in LB-medium at 37 °C and shaking at 210 rpm. Cells were harvested on the following day and the plasmid was extracted using NucleoSpin® Plasmid kit (Macherey-Nagel).
The construct was tested by transfection into HEK293T cells, followed by Western Bot analysis (described in section 2.3.3). To cite the pEGFP-N3 vector as an example see Figure 2, the transfection efficiency was tested in HEK293T cells and the corresponding protein lysate was analysed by immunoblotting with antibodies against the enclosed tag, here GFP. Additionally, the specificity of the commercially available antibody for Lrrc39 was tested with this approach.

4.1.3 Yeast-two hybrid

The full sequence of the murine *Lrrc39* was amplified using the primer Lrrc39BsmB1 (5'CGTATTTCGTCTCCCATGACAGAAAGTGCGGTTTGTACC3') and Lrrc39Sal1 (5' CGATACGTCGACTTATTACATTTTTCTGACCTCTGGTGG3'). The PCR were

carried out under following condition using Phusion Polymerase (Finnzymes) to reduce the mutation rate:

10 µl	Phusion buffer (5x)
1 µl	dNTP mix (10 mM)
2.5 µl	Primer Mix (10 µM)
1 µg	Template cDNA
0.5 µl	Phusion Polymerase (Finnzymes)
35 µl	ddH2O
Total volume 50 µl	

Following amplification conditions were chosen:

Step	Temperature	Time	
Denaturation	98 °C	30 sec	
Denaturation	98 °C	30 sec	40 cycles
Annealing	55 °C	50 sec	
Extension	72 °C	30 sec	
Extension	72 °C	7 min	
Storage	4 °C		

After amplification, the PCR reaction mix was separated in a 1 % agarose gel, containing 0.1 % ethidiumbromide for visualisation under UV light. The according band was isolated and purified using PCR clean-up Gel extraction Kit (Macherey-Nagel). To insert the PCR product into pGBKT7 vector (Clonetech), in the Y2H vector it was enzymatically digested with *BsmBI* (NEB) for 2 h at 55 °C followed by Sal1 digestion for 2 h at 37 °C. The produced insert with a size of 1014 bp was separated by gel electrophoresis and purified. The DNA was precipitated by Phenol/Chloroform/Isoamylalcohol. The DNA concentration was measured using spectrophotometer Nanodrop (PEQLAB Biotechnologie GmbH). Before ligation of the insert into the final vector at 16 °C overnight, the pGBKT7 vector was prepared using the restriction enzymes NcoI and SalI (both from NEB). The ligation reaction was transformed into XLblue by electroporation and plated on appropriate LB-antibiotic plates and incubated at 37°C overnight. The colonies were picked and inoculated as described before. The plasmid DNA was extracted using NucleoSpin® Plasmid kit (Macherey-Nagel). The final product was sequenced, verified and used as bait protein for the Y2H screen. The Matchmarker Two-Hybrid System 3(Clonetech) that was used contains the bait protein, which is fused to a GAL4 DNA-binding domain (DNA-BD) and the preys, which consist of a library of cDNA that is fused to the GAL4 activation domain (AD). When the bait protein, here Lrrc39 binds a specific library fusion protein, the two

fused domain are in close proximity and thereby activate the transcription of reporter genes. The interacting clones from the library were isolated and sequenced. Obvious false positive candidates, like non-coding parts of the cDNA were removed as well as non-specific preys. For Yeast two hybrid screening, the construct was sent for service to the DKFZ, Heidelberg.

4.1.4 Whole mount *in-situ* hybridisation

In-situ hybridisation is a method to visualise a specific RNA sequence in the entire tissue by labelled complementary RNA strands (probes). For this study it was used to analyse the expression pattern of *Lrrc39* in murine embryonic tissues. For sample preparation, pregnant wildtype mice were sacrificed at the corresponding day after vaginal plug. The whole mount mouse embryos at the developmental stage E10.5 and E11.5 were dissected of the uterus tissue and fixed in 4 % PFA at 4 °C overnight. After washing in ice-cold PBS for 30 min the samples were dehydrated through methanol series of 25 %, 50 %, 75 % and twice in 100 % at room temperature for 10 min each. The antisense RNA probes were generated and labelled with digoxigenin-UTP (Roche) by SP6 RNA polymerase (Promega). The detailed hybridisation procedure is described in (Wilkinson, 1992) and was done accordingly. Embryos were post-fixed in 4 % PFA and imaged using Leica M205 FA.

4.1.5 Gene expression profiling

The zebrafish embryos at 3 dpf were selected by phenotype and pooled in batches of fifteen individuals each. The following four different conditions were used: 1. without injection; 2. Standard control MO injection; 3. *zLrrc39* MO targeting the ATG site and 4. Control MO targeting the ATG site of a specific gene on a different chromosome.
All embryos were taken from the same zebrafish mating. The RNA was extracted from pooled, snap frozen samples and homogenised in 500 µl TRIZOL reagent (Life technologies, Darmstadt) and incubated for 5 min at room temperature. The gene expression profiling using GeneChip Zebrafish Genome Array (Affymetrix) was performed. The complete zebrafish cDNA was hybridized to the Affymetrix GeneChip following the manufacturer's protocol. The pre-processing of the data, including background subtraction was performed according to the robust multi-array average (RNA) procedure.

4.1.6 *In-vitro* transcription

For zebrafish mRNA microinjection experiments, the full length messenger RNA of murine and zebrafish *Lrrc39* were cloned in pGEM-4Z vector (Promega) using following primers:

*m*lrrc39_forward:	5'-AAAAGATCTACCATGATGACAGAAAGTGCGGTT-3'
*m*lrrc39_reverse:	5'-AAAACTAGTTCAATTACTTTCTCCATCAGAATG-3'
*zlrrc*39_forward:	5'-AAAGATATCACCATGACGGGGGTCACGGTG-3'
*zlrrc*39_reverse:	5'- AAAGATATCCTATGCCGTTTCTGACACACC-3'

The plasmid DNA was linearized using *Sal1* digestion at 37 °C for 2 hours. The mRNA synthesis was performed using the Ambion mMESSAGE mMACHINE Kit (Ambion AM1340) according to the manufacture's protocol. The synthesised mRNA concentration was determined photometrically and diluted to 2µg/µl in RNase free water.

4.2 Cell biological methods

4.2.1 Morpholino oligonucleotide microinjection

The following Morpholino antisense oligonucleotide (MO) were designed as described in (J. Summerton and Weller, 1997) and purchased from GeneTools (www.gene-tools.com):

MO ^{lrrc39ATG1}:	5'-ATGATGATACGGAGCCGCAAATTTC-3'
MO ^{lrrc39ATG2}:	5'-CGCAAATTTCACTTCACACTCACAT-3'
MO ^{lrrc39Spl1}:	5'-TGCACGTACCTACCCACTGCTCTTC-3'
MO $^{standard\ CTRL}$:	5'-CCTCTTACCTCAGTTACAATTTATA -3'

The *lrrc39* RNA sequence was targeted at the start codon to block the entire translation with MO*lrrc39ATG1-2*, and within the sequence by MO*lrrc39 Spl* and 2. All MO stocks were diluted in ddH2O to 1mM. For all experiments, the optimal MO concentration was tested in three different concentration (high: 500 µM MO; medium: 200 µM MO; low: 100 µM MO). Zebrafish embryos were microinjected with a volume of 2nl of MO during one-cell-stage. Needles were pulled from glass capillary tubes using a flaming micropipette puller (Model P-97, Sutter Instruments Co.) and injections were performed using a microinjector. To control the injection procedure all MOs were diluted in phenol red (Sigma) for visualisation. A minimum of fifty eggs were used for each injection. In order to control the injection technique and the MO specificity, the following controls were used: Phenol Red without any MO; a standard control MO, directed against the human β-globin pre-mRNA (GeneTools) and MO targeting a specific non-related gene.
The capped mRNA of *lrrc39* was prepared as described in section 2.1.6 and was injected at the one to two cell stage. For rescue experiments co-injection of 200 mg/µl mRNA and 200 µM MO*lrrc39ATG1* were performed.

4.2.2 Electron microscopy

Transmission electron microscopy (TEM) uses a high voltage electron beam to examine a sample at a very fine scale. In order to analyse the fine structure of the sarcomere in zebrafish, the larvae were anaesthetised with Tricaine (Sigma; A5040) at 5 dpf, treated with 100 mM CdCl2 for equal contraction (Baker, 1944) and fixed in 3 % Glutaraldehyde in cacodylate buffer (21.4 g Na-cacodylate, 75 g sucrose/ liter) at 4 °C overnight. The sample preparation for transmission electron microscopy (TEM) was done essentially as described before (Rieger and Koster, 2007). The zebrafish larvae were washed twice for 10min each in sodium cacodylate buffer and incubated in 1% osmium tetroxide in veronal acetate buffer (14.7 g Na-Veronal; 9,7 g Na-acetate/500 ml) with 4% sucrose for 2 h at 4 °C. The samples were dehydrated in a gradient ethanol series (20, 30, 50, 70, 90 and 100 % for 15 min each). Ethanol was removed and exchanged by propyleneoxide (PO) and incubated for 30 min, followed by gradient exchange to Epon (4 g Epon, 3 g DDSA (Dodecenyl Succinic Anhydride), 1 g NMA (Methyl-5-Norbornene-2,3-Dicarboxylic Anhydride), 0.1 g DMP **(**2,2-dimethorypropane) per 10 ml). The Epon blocks polymerised at 60 °C for 24 hours before ultrathin sections of 70 nm were prepared using an ultramicrotome and mounted on grids. The sections were contrasted with uranyl acetate and analysed with a Philips CM10 TEM, operated at 100 kV. The tail skeletal muscle of two $MO^{control}$-injected- and two MO^{Lrrc39}-injected zebrafish larvae were examined.

4.2.3 Cell culture methods

C_2C_{12} and HEK293T cell lines were grown in DMEM (1 g/ml Glucose) media (GIBCO®), supplemented with 10 % FCS and 100U/ml penicillin, 100 µg/ml streptomycin and 300 µg/ml L-glutamine. Before transfection, cells were grown for 2-3 days at 37 °C and 5 % CO_2. One day prior transfection, cells were trypsinized (0.05 % trypsin, 0.02 % EDTA in 1x PBS) and plated at 20-30 % confluence. For transfection of each 10 cm culture dish, 0.5 ml 2xHBS (280 mM NaCl; 1.5 mM Na_2HPO_4-$7H_2O$; 50 mM HEPES; pH 7.0), were gently mixed with 10 µg DNA, 61 µl 2M Ca_2Cl_2 in 0.5 ml ddH_2O. This transfection mixture was immediately transferred to the cells by drop wise distribution. Cells were incubated at 37 °C and 5 % CO_2 for 24 hours. After media change the cells were maintained for further 24 hours. Additionally, cells were treated with 0.2 µM Jasplakinolide (ALX-350-275) for 1 h, or with 2 µM Lantrunculin A (ALX-350-130) for 30 min (both reagents from Enzo life sciences GmbH). Cells were washed in PBS and used for stainings or Western blot analysis.

For radiolabelling, which was used to document the incorporation of phosphate and thereby analyse changes in phosphorylation patterns in combination or absence of Lrrc39, HEK293T cells were re-suspended in 4 ml DMEM media, supplemented with 10 % FCS and 10 µl [γ-^{32}P] dATP (100 µCi) and incubated for 2 hours at 37 °C.

After immunoprecipitation of the transfected proteins and SDS gel electrophoresis the nitrocellulose membrane was documented by phosphoimager.

4.2.4 Isolation, cultivation and transfection of neonatal rat cardiomyocytes

Primary neonatal rat cardiomyocytes (NRC) were freshly isolated and plated on fibronectin (10 µl/ml) coated culture slides in Medium 199, containing 10 % horse serum, 1 % Penicillin-Streptamycin-Glutamine (Invitrogen), Ara C and AmphotericinB, and were incubated at 37 °C. On the second day after isolation, cells were transfected with full length *mLrrc39* in pEGFP-C1, pEGFP-N3 and His/V5 using transfection reagent ESCORT III (Sigma, L3037). The transfection procedure was done according to manufacturer's protocols. The transfected cardiomyocytes were cultured for another two days in Medium 199 (2 % FCS, 1 % Penicillin-Streptamycin-Glutamine (Invitrogen) before fixing in 4 % PFA followed by antibody staining etc. Lrrc39 expression was analysed by GFP signal on a fluorescent microscope (Zeiss Imager.Z1, Carl Zeiss MicroImaging, Inc.) using 40x or 100x objective. Micrographs were analysed using AxioVision Release 4.5 and ImageJ software.

4.2.5 Isolation, cultivation and transfection of primary chicken cardiomyocytes

Fertilized Chicken eggs were purchase from a local farm (Geflügelzucht Körner, Usingen). The fertilized eggs were incubated at 38 °C until sample preparation at day 11. After cleaning of the eggs with 70 % EtOH, the embryos were decapitated and the hearts were isolated and collected in a 10 cm preparation dish containing DMEM media. The primary explants were mechanically disaggregated and transferred to 50 ml falcons. The finely chopped tissues were additionally enzymatically disaggregated by 0.25 % trypsin for 15 min at 37 °C. The supernatant were discarded and the pellet was suspended in 0.15% collagenase, diluted in PBS and incubated for 30 min at 37 °C. The supernatant was collected in 10 cm culture dish. The collagenase digestion was repeated twice and the collected supernatant incubated for 30 min at 37 °C. The pre-plating step is necessary to minimise fibroblast contamination in the final culture. Due to their faster adherence, these cells were attached to the culture dish while the cardiomyocytes were still floating in the media. After final incubation, the supernatant was used for transfection.
The full-length mLrrc39-eGFP fusion protein as well as the empty vector (described in section 2.1.2) was used for electroporation with Amaxa® Rat Cardiomyocyte – Neonatal Nucleofector® Kit (Lonza, Cologne) according to the manufacturers' protocol. Transfected cells were cultured for 24 h in DMEM containing 10 % horse serum and 1 % Penicillin-Streptamycin-Glutamine (Invitrogen) before analysis.

4.3 Immunhistochemistry

4.3.1 Vibratome sections

The utilised transgenic zebrafish lines contain a fluorophore, which was analysed by confocal microscopy. Due to the thickness of the whole zebrafish, tissue sections were prepared using a vibrating microtome. Therefore, dechorinated zebrafish were anaesthetised with Tricaine (3-amino benzoic acidethylester, A-5040, Sigma) and fixed in 4 % PFA overnight at 4 °C. After washing in PBS, the samples were orientate in 4 % low melting agarose (Sigma) and left for harden in a wet chamber for at least six hours at 4 °C. The in agarose fixed samples were glued to the section base and cut in 50 µm thick slices with a in solution vibrating knife using Microtom VT 1000S (Leica). The transparent slices were transferred to objective slides and embedded in Mowiol. Images were taken using confocal microscopy (Leica TCS SP2 AOBS).

4.3.2 Immunostaining

The immunhistological experiments oblige the exact localisation of relevant molecules within the cells. Therefore the cultivated cells are fixed for 10 min at room temperature in 4% paraformaldehyde (PFA), which forms crossbridges of aminogroups and thereby stabilises the subcellular structures. After washing steps to remove the PFA, the samples were permeabilised with 0.05 % Triton-X100 for 5 min. This detergence increases permeability of the cell membranes and erleichtert antibodies to enter the cell. After permeabilisation and washing steps with PBS (137 mM NaCl, 2.7 mM KCl, 4.3 mMNa_2HPO_4-$2H_2O$, 1.4 mM KH_2PO_4), unspecific binding of antibodies were minimised by 0.1 % BSA (bovine serum albumin) treatment for 10 min. By this saturation of free binding sites background staining is reduced. Next step is the incubation of the primary antibody against the target protein. The antibodies were diluted in PBS as listed in Table 3. The incubation time depended on the quality of the antibody as well as the sample and was generally done at 4 °C overnight. After PBS washing steps, the incubation of the secondary antibody that was directly coupled to a fluorophore followed. In order to identify the cell shape additional fluorophore-labelled substances were used. To visualise structure proteins, the peptide Phalloidin (Phalloidin-TRITC, Sigma) was used, which binds F-actin (Cooper, 1987). For nuclear staining of the cells DAPI (4',6'-diamidino-2-phenylindol) was added. After staining and final washing steps the cells were embedded in Mowiol® (Calbiochem) and visualised using different microscopes. For each secondary antibody a negative control was used, with the same conditions but excluding the primary antibody in order to recognise unspecific binding.

The whole mount immunofluorescence staining for zebrafish followed the same principle. The dechorinated fish were collected into 24-well plates. The E3 medium was removed and the samples were washed once with PBS. Samples were fixed in 4 % PFA overnight at 4 °C. Zebrafish were washed in PBS and permeabilised for two hours at room temperature in PBS, containing 1 % Triton X-100 (Roche, Karlsruhe) and 0.1 % DMSO. Primary antibody were diluted in PBS as shown in Table 3 and incubated overnight at 4 °C following by washing steps and nuclear staining with DAPI (1:1000; Invitrogen). Samples were mounted on objective slides in mowiol and visualised using a Leica confocal microscope TCS SP2 AOBS with argon and helium-neon lasers and 63xlens. Images were processed with Leica imaging software and ImageJ.
For oil red O stainings, the fixed larvae were incubated in 60 % 2-propanol for 30 min, followed by 0.5 % oil red O (sigma) in 60 % 2-propanol.

Table 3: Antibodies for immunofluorescence staining. Utilised primary and secondary antibodies are listed by antibody name, dilution, company and order number.

Name	Dilution	Source, (order ID)
Lrrc39	1:50	Abnova (H00127495-B01P)
mMaC Myomesin B4	1:100	Developmental Studies Hybridoma Bank
Titin M8	1:50	Gift from Prof.M.Gautel
GFP	1:500	Abcam (ab6556)
α-actinin	1:200	Sigma (clone EA53)
F59	1:20	Developmental Studies Hybridoma Bank
MyBP-C	1:200	Gift from Prof.M.Gautel
Goat-anti-rabbit-Alexa 488	1:1000	Invitrogen (A11008)
Goat-anti-mouse-Alexa 594	1:1000	Invitrogen (A11005)
Donkey anti-mouse Cy5	1:200	Millipore (AP192S)

4.3.3 SDS gel electrophoresis and immunoblot analysis

To identify specific polypeptides within a complex protein mixture, SDS-PAGE (sodium dodecyl sulphate- polyacrylamide gel electrophoresis) and western blot were applied. This technique was used for the identification of proteins after immunprecipitation to verify potential interaction partners. The snap frozen tissues were homogenized in Extraction-buffer (100 mM Tris-HCl, 12.7 mM EDTA, 10 % SDS, pH 8.0) supplemented with Protease-Inhibitor-Cocktail-Tablets (1tablet/ 10 ml; Roche) using ultrasonic device (Sonoplus, Bandelin Electronic, Berlin).
To have equal amounts of samples per lane the protein concentration were measured by DC Protein Assay from BioRad based on the Lowry method. In alkali solution (reagent A) proteins form complexes with copper-II-ions, this is known as biuret-reaction that leads to the blue colouring of the sample. The so formed protein complex reacts with folin-ciocaltheau-phenol-reagent (reagent B) that enhances colour change and increases the sensitivity of the assay. The colour intensity is proportional to the amount of proteins

and can be measured by photometry at 750 nm (Fluorimeter FLUOstar Galaxy, Bio-Rad). The samples are diluted in Extraction buffer and Laemmli buffer (66.7 mM Tris-HCl, 0.2 % SDS, 27 % Glycerin, pH8.0) to a final concentration of 2 µg/µl.

The denaturated polyacrylamide gelelectophoresis (PAGE) after *Laemmli* (Laemmli, 1970) allows the separation of a complex protein mixture by the size of the polypeptides. Based on the application of sodiumdodecylsulfate (SDS), which binds amino acids and charges them negatively, the proteins can be separated in an electric field. Due to the correlation of molecule mass and electrophoretic mobility are separated by size. To reduce potential disulfide bonds of proteins strong reducing agents like dithiotreitol (DTT) were used. For protein separation, gradient Bis-Tris gels (NuPage 4-12 %, Life Technologies) were utilised. The electrophoresis was performed at 175 volts in semi-dry modules (Xcell SureLock™ Mini-Cell, Invitrogen) in MES-SDS running buffer. A prestained protein ladder (Novex®Sharp Pre-stained Standard, Invitrogen) was applied to estimate the band size. After electrophoresis, proteins were transferred to nitrocellulose-membrane (Amersham Bioscience) at 30 V for 2 h in Transfer-buffer (25 mM Bicine, 25 mM Bis-Tris, 0.1 mM EDTA, 20 % Methanol). In order to control the protein transfer, membranes were stained with RedAlert™ (Novagen) prior to Western Blot analysis. For immunodetection the membrane was saturated with 5 % milk pulver in TBST (137 mM NaCl, 20 mM Tris-HCl, 0.1 % Tween) for 1 h and incubated with specific primary antibodies (listed in Table 3) overnight at 4 °C. After washing steps in TBST, the secondary antibody was incubated for 1 h at room temperature; followed by washing steps. The membrane bound proteins were visualised by chemiluminescence via horseradishperoxidase (HRP)-coupled secondary antibodies and Supersignal West Pico Chemiluninescent Substrate (Pierce) and were detected by VersaDoc MP5000 (Bio-Rad).

Table 4: Antibodies for Western Blotting. Utilised primary and horse radish peroxidase (HRP)-conjugated secondary antibodies are listed by antibody name, dilution, company and order number.

Antibody	Dilution	Source, (order ID)
Lrrc39	1:1000	Abnova (H00127495-B01P)
Myomesin	1:1000	Developmental Studies Hybridoma Bank
Anti-GFP	1:10 000	Abcam (ab6556)
Anti-V5	1:5000	Novex®, Life technologies
MF20	1:1000	Developmental Studies Hybridoma Bank
Goat anti-mouse-HRP	1:1000	Pierce (1858413)
Goat ant-rabbit-HRP	1:1000	Pierce (1858415)

4.4 Biochemical methods

4.4.1 Immunoprecipitation

For in vitro analysis of Lrrc39 and its binding to myosin heavy chain immunoprecipitations were performed. Cultured chicken cells were transfected as mentioned in section 2.2.4. The utilised constructs are described in section 2.1.2. After harvest of the cells and preparation in lysis buffer (150 mM NaCl, 1% Triton® X-100 in 50 mM Tis-HCl, pH 8.0), the samples containing the epitope-tagged fusion protein were processed using µMACS GFP Isolation Kit (#130-091-125, Milteny Biotec) to isolate the specific protein of interest with their binding partners. Therefore, cells were incubated with anti-GFP microbeads and isolated by column. The samples were eluted in pre-heated buffer, containing 1 % SDS, 1 mM EDTA 50 mM DTT in 50 mM Tris-HCl (pH 6.8) and additionally for gel loading 10 % glycerol and bromphenol blue to visualise the sample. Afterwards, the samples were processed for Western Blot analysis as described in the previous section and the results were visualised in VersaDoc MP 5000 System (BioRad).

4.4.2 Mass spectrometric analysis of zebrafish tissue

Mass spectrometry identifies proteins by ionisation of chemical compounds to charge peptides and measurement of their mass-to charge ratio. This technique was utilised for proteome analyse of zebrafish samples. Therefore, the zebrafish tails were freshly isolated at 120 hpf to provide the aged matched condition to the electron microscopy experiments. For each sample to be read by mass spectrometry, twenty individual tails of zebrafish larvae were dissected using forceps and directly snap frozen in liquid nitrogen. The samples for MO^{Lrrc39} and a standard MO control zebrafish (MO^{Ctrl}) were pooled. Three replicates were prepared for each experiment.

To be able to precisely quantify the changes of protein expression, the stable isotope labelling (SILAC) approach was used for this study. This technique applies $13C^6$-Lysine (Lys-6) labelled bacteria, yeast or mouse cells to label whole zebrafishes, which represent the internal protein standard for protein quantification. Proteins of the heavy standard were mixed in one-to-one proportion with the control- and MO^{Lrrc39} group, respectively. A comparison of the ratios between samples and the heavy labelled fish allows precise quantification of protein expression. Protein samples were typsin digested and measured by LC-MS/MS. The pre-processing of the data, including background subtraction, elimination of contaminants and normalisation was performed using MaxQuant (maxquant.org) software (Cox, 2009). For final data analysis Microsoft Excel was utilised. Results were expressed as means. A two tailed t-test was performed to assess differences between the two groups. P-values < 0.05 were considered as significant.

4.4.3 Structural analysis by homology modelling

Visualisation of 3D structures requires NMR or crystallographic analysis, which was not available for LRRC39. To test the validity of the results and to visualise the binding of this protein, its three dimensional structure was predicted based on homology modelling. Assuming that, a similar primary structure is likely to form a similar conformation the protein internalin-A (PDB entry: 2omxA) that shows highly similar sequence patterns (Wollert, 2007) was chosen as template. For molecular visualisation the software PyMOL version 1.3r1 (pymol.org) was utilised. The structure information for α-helical myosin heavy chain (PDB entry: 3DTP) was taken from (Alamo, 2008).

4.5 Animal procedures

Adult Tübingen strain zebrafish *(Danio rerio)* and transgenic lines Tg(*myl7::eGFP-HsHRAS*)^{s883} (D'Amico, 2007), Tg(*aactin::eGFP*) (Higashijima, 1997) were maintained in system water at 28.5 °C according to standard methods (Westerfield, 1993). Embryos were taken from natural mated adult fish and raised in embryo media E3 (5 mM NaCl, 0,17 mM KCl, 0,33 mM $CaCl_2$, 0,33 mM $MgSO_4$). For analysis embryos were anesthetized with Tricaine (ethyl 3-aminobenzoate methanesulfonate , Sigma) before preparation.

5 RESULTS

The analysis of novel genes in the striated musculature started with a computer-based comparison of different databases in order to filter genes by their expressions and identify genes that are specific or enriched in heart- and/or skeletal muscle and are not published in the target field of research (Uchida, Jenniches, 2009; Gellert, Jenniches 2010). The identified genes were further selected by their known family members or their protein domains. This work focuses on one of such candidate genes termed *leucine rich repeat containing 39* (Lrrc39), which is highly conserved through different species.
In the first part of this section, the expression pattern of this gene was analysed in different tissues, in order to verify the *in silico* screen. Furthermore, the subcellular localisation of Lrrc39 was examined (Chapter 5.1).
The second part concentrates on the functional importance of Lrrc39 *in vivo*, including the phenotypic description after gene knock-down and detailed morphological analysis using immunhistological stainings, confocal microscopy as well as gene and protein analysis by microarray and mass spectrometry, respectively (Chapter 5.2).
The third part of this section focuses on the interaction to other proteins by analysing the binding partner of Lrrc39. Furthermore, the effect of this interaction was investigated *in vitro*, using freshly isolated neonatal myocytes.

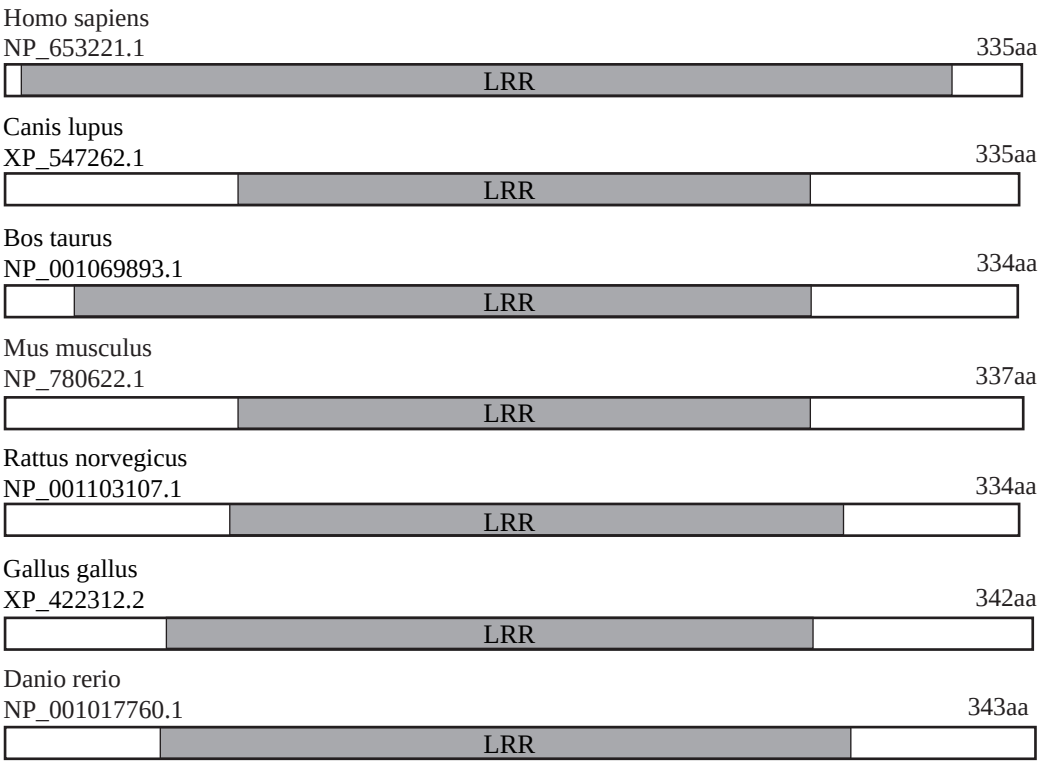
5.1 Expression pattern

5.1.1 Sequence and expression of *LRRC39* are highly conserved

LRRC39 is an evolutionary-conserved gene, which encodes the protein of 334 to 342 amino acids. Apart from a comparable short N-terminal and C-terminal region, the sequence of *LRRC39* shows a repetitive pattern of leucine-rich repeat (LRR) domains. This sequence is highly conserved across species ranging from human to zebrafish as shown in Figure 5A. The homologs of *LRRC39* were documented in seven different organisms from zebrafish to human, based on the results of HomoloGene database (www.ncbi.nlm.nih.gov/ homologene). To obtain the sequence similarity scores, ClustalW2 (www.ebi.ac.uk/ Tools/ msa/clustalw2) was utilised.

The human *LRRC39* (*Homo sapiens*: NP_653221.1) encodes 335 amino acids in length, which consists predominantly of LRR-domains. These conserved LRR-domains appear always in the central region of the sequence across species; however, it is slightly longer in human compared to all other species. The sequence length of *LRRC39* in *Canis lupus* is identical to the human form, but less LRR domains are detected. The homologs in *Bos Taurus* and *Rattus norvegicus* are one amino acid less than that of human, whereas those

A



B

Pairwise Alignment		Identity (%) Protein	DNA
M.musculus	vs. R.norvegicus	94,9	94,1
H.sapiens	vs. M.musculus	85,9	86,0
H.sapiens	vs. R.norvegicus	85,6	84,9
H.sapiens	vs. G.gallus	69,0	71,7
M.musculus	vs. G.gallus	67,9	67,7
R.norvegicus	vs. G.gallus	66,5	66,7
H.sapiens	vs. D.rerio	63,9	64,1
R.norvegicus	vs. D.rerio	62,1	62,7
M.musculus	vs. D.rerio	61,8	62,1
G.gallus	vs. D.rerio	60,3	61,4

Figure 5: Sequence alignment of *LRRC39*. A: Conservation of Lrrc39 protein sequence in different species NCBI reference sequence numbers are stated (modified from HomoloGene database entry). Centrally located leucine-rich repeat domains (LRR) exist in all shown sequences. B: Pairwise alignment scores among selected organisms indicate high sequence identity on protein as well as on DNA level.

of chicken (*Gallus gallus*), zebrafish (*Danio rerio*) and mouse (*Mus musculus*) are longer (342 aa, 343 aa and 337 aa, respectively). In mice, an isoform of 175 aa is present, which contains only three LRR-domains and is alternative spliced after the fifth exon.

The pairwise alignment of the human sequence and model organisms that were used in this study are shown in Figure 5B. Sequence identities, analysed by ClustaW ranging from 60 to 86 %. The highest similarity is between mouse and rat with 94.1 % identity at the level of DNA sequence and a slightly higher identity of the protein sequence with 94.9 %. The human sequence compared to those of mouse and rat yields in 85 to 86 %

identity. The weakest compliance shows the comparison between chicken and zebrafish with the identity of 60.3 %. However, the distribution of the LRR domains stays highly similar. Taken together, the alignments of the different *LRRC39* homologs indicate high conservation across species.

5.1.2 Expression of *LRRC39* in heart and skeletal muscle

To verify the analysed database entries for the *LRRC39* expression, RT-PCR experiment using fifteen different adult murine tissues was conducted. As shown in Figure 6A, a clear expression of the murine *Lrrc39* in heart and skeletal muscle tissues was recorded, which verify the predicted pattern based on our *in silico* results. Apart from the adult striated muscle tissues, the expression of *Lrrc39* was detected in the eyes and in the brown adipose tissue (BAT), whereas the white adipose tissue (WAT) shows no expression. In addition, a smaller band was observed in BAT to indicate the presence of an isoform as mentioned in the previous section.

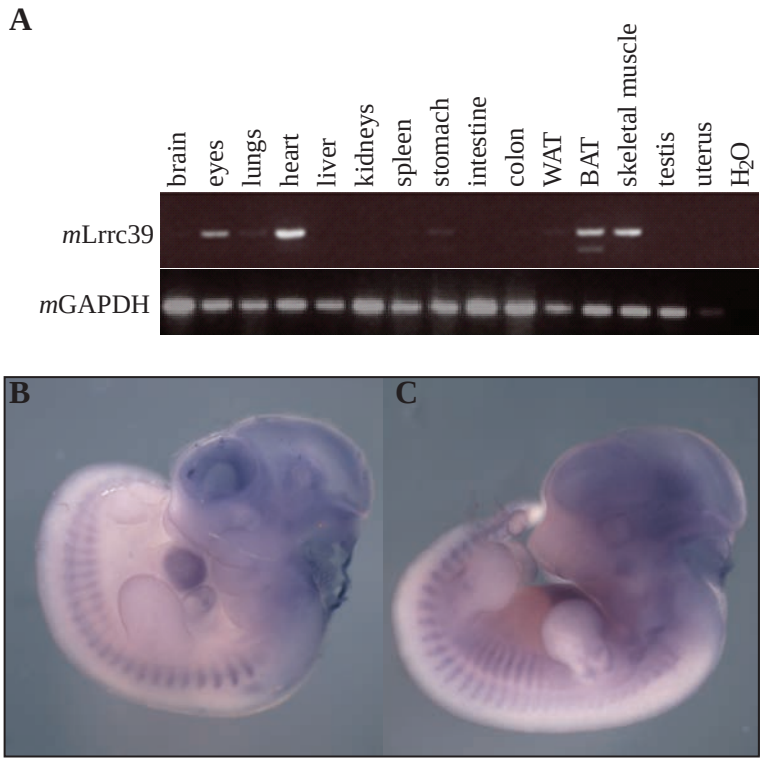


Figure 6: Expression pattern of *Lrrc39* in murine tissues. A: RT-PCR of adult mouse tissues. In adult mouse tissues, *Lrrc39* is expressed in eyes, heart, brown adipose tissue (BAT) and skeletal muscle (28 cycles of RT-PCR reaction). The murine *Gapdh* served as an internal control, and ddH2O was used as a negative control. Whole mount *in situ* hybridization of mouse embryos at developmental stage E10.5 (B) and E11.5 (C). A clear expression of Lrrc39 in heart and somites can be observed.

To further characterise the expression pattern of *Lrrc39*, whole mount *in situ* hybridisation experiment was performed using murine embryos. For this experiment, two different embryonic stages were analysed: 10.5 days after fertilisation (E10.5) and 11.5 days after fertilisation (E11.5), which are shown in Figure 6B and C, respectively. In both developmental stages, *Lrrc39* transcripts were detected in the developing heart as well as in the somites, which will form the skeletal muscle tissue in later time point of the development. In addition, a strong *Lrrc39* expression was detected in the ventricle of the heart. The expression in the somites seems to be restricted to the myotome at E10.5 and increased in size at E11.5. Additionally, *Lrrc39* transcripts are detectable in the fore limb. Some weak signals were detected in the eye and neck region. As it is clearly visible from the above figure, high expression of *Lrrc39* in the developing heart and somites to further confirm the results of our *in silico* screening.

As it was shown in Figure 5, *LRRC39* is conserved from lower species, including zebrafish. To verify its evolutional conservation, RT-PCR experiment was conducted using zebrafish tissues. For this purpose, total RNA of zebrafish larvae was collected at different developmental stages: from one hour post fertilisation (hpf) to seven days post fertilisation (dpf). Unlike mice, which can be inbred to establish genetic homogeneity, zebrafish are generally outbred (Monson and Sadler, 2010), and therefore, genetic variations are higher. To circumvent this problem, all samples mentioned above were taken from one Tübingen-strain mating. The reason behind choosing Tübingen over other strains is that the sequencing project of zebrafish was conducted with this strain (http://www.sanger.ac.uk/Projects /D_rerio/). By using this strain, in principle, it would be possible to base the understanding on the publicly-available sequences, which can be obtained from major public databases, such as ENSEMBL and Entrez Gene databases. As shown in Figure 7A, RT-PCR analysis in zebrafish larvae resulted in detection of *lrrc39* expressions starting relatively weak at the first day of development and increased to a plateau phase from two to four days post fertilisation, followed by a slight decrease from day five onwards. During these early developmental stages of the zebrafish, the organs are too small to gain enough material for a more detailed, tissue-specific analysis. Thus, although it is informative, RT-PCR experiment can only indicate the presence of *lrrc39* during the development of zebrafish larvae. As it has been done for mouse (Figure 6), whole-mount *in situ* hybridisation should provide spatial information of *lrrc39*. Conveniently, such datasets are available from high-throughput analysis by Thisse and Thisse (Thisse, 2004), which are accessible through the ZFIN organisation (www.zfin.org). These results are shown in Figure 7B and C. *Lrrc39* transcripts were detected in the somites and pectoral fin muscles, however no clear signal could be found in the heart at this early developmental stage. As shown in the above results, the expression patterns of *lrrc39* are similar to its mouse homolog.

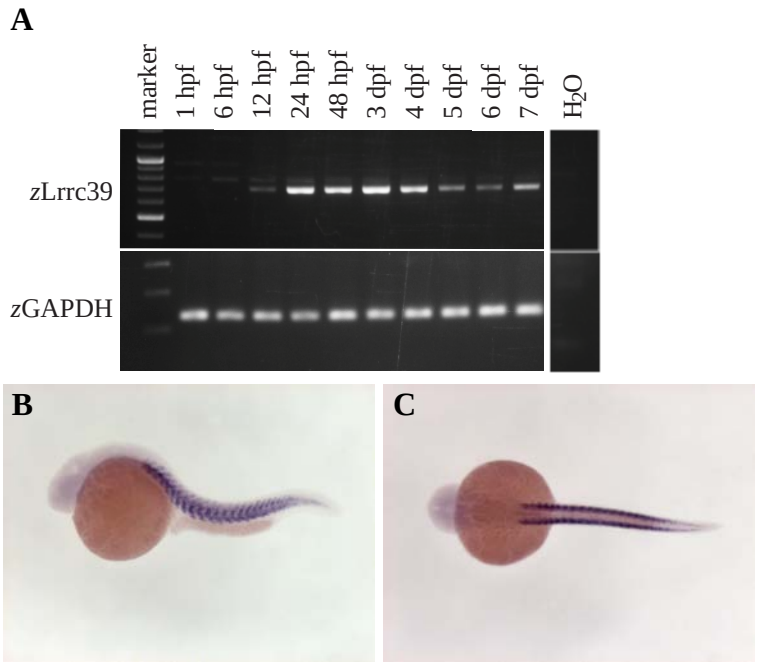


Figure 7: Expression pattern of *lrrc39* in zebrafish tissues. A: RT-PCR of zebrafish tissues showing *lrrc39* transcripts arising at 12 hours post fertilisation (hpf), increasing to a plateau phase from two to four dpf followed by a slight decrease from five to seven days post fertilisation (dpf) (40 cycles of RT-PCR reaction). The zebrafish GAPDH served as an internal control, and ddH2O was used as a negative control. Whole mount in situ hybridization of *lrrc39* from lateral view (B) and dorsal view (C) taken from (Thisse, et al., 2004).

5.1.3 Characteristic localisation of *LRRC39* in the sarcomere

The above results are informative about the expressions of *LRRC39* transcripts. Since proteins are the ones doing the job of muscle maintenance, the expressions of *LRRC39* were analysed at the protein level using fusion proteins and an antibody against *LRRC39*. First, in order to obtain subcellular localisation of *LRRC39*, GFP-fused Lrrc39 was over-expressed in various cell lines. As shown in Figure 8A, the fusion protein of Lrrc39 was localised in the cytoplasm of a cell as well as predominantly at the edges of the cell. The constructed fusion proteins are co-localised with structural protein F-actin, which is visualised by phalloidin staining. Next, in order to test the specificity of a commercially available *LRRC39* antibody (Abnova), co-localisation with different *Lrrc39*-fusion proteins over-expressed in various cell lines (HEK293T, NIH-3T3, P19CL6 and C2C12 cells) was performed, which confirmed its specificity. The specificity of anti-*LRRC39* antibody was further tested by Western Blot experiment using over-expressed fusion proteins in the same set of cells (data not shown).

Since Lrrc39 was localised in similar patterns like F-actin, its properties for actin polymerization and depolymerization were tested (Figure 8B). Cultured C_2C_{12} cells were treated with 0.2µM Jasplakinolide (Enzo life sciences GmbH), a reagent that promotes actin polymerisation (Holzinger, 2009), and on the contrary LatrunculinA (2µM;

Alexis®Biochemicals), which promotes actin depolymerisation (Chesarone, 2009). Lrrc39-eGFP fusion proteins were detected in cytoplasmic localisation of the all observed cells. DMSO treated cells served as control. In Jasplakinolide-treated cells, a clear increase in cell size was observed. Furthermore, Lrrc39 expression was increased compared to the control cells. The harsh conditions after the Latrunculin A treatment lead to reduced cell number and cell size, which is visualised by DIC images. Surprisingly, Lrrc39-tranfected cells were viable. In fact, all surviving cells were GFP-positive to indicate some functional roles of Lrrc39 for actin polymerization and depolymerization.

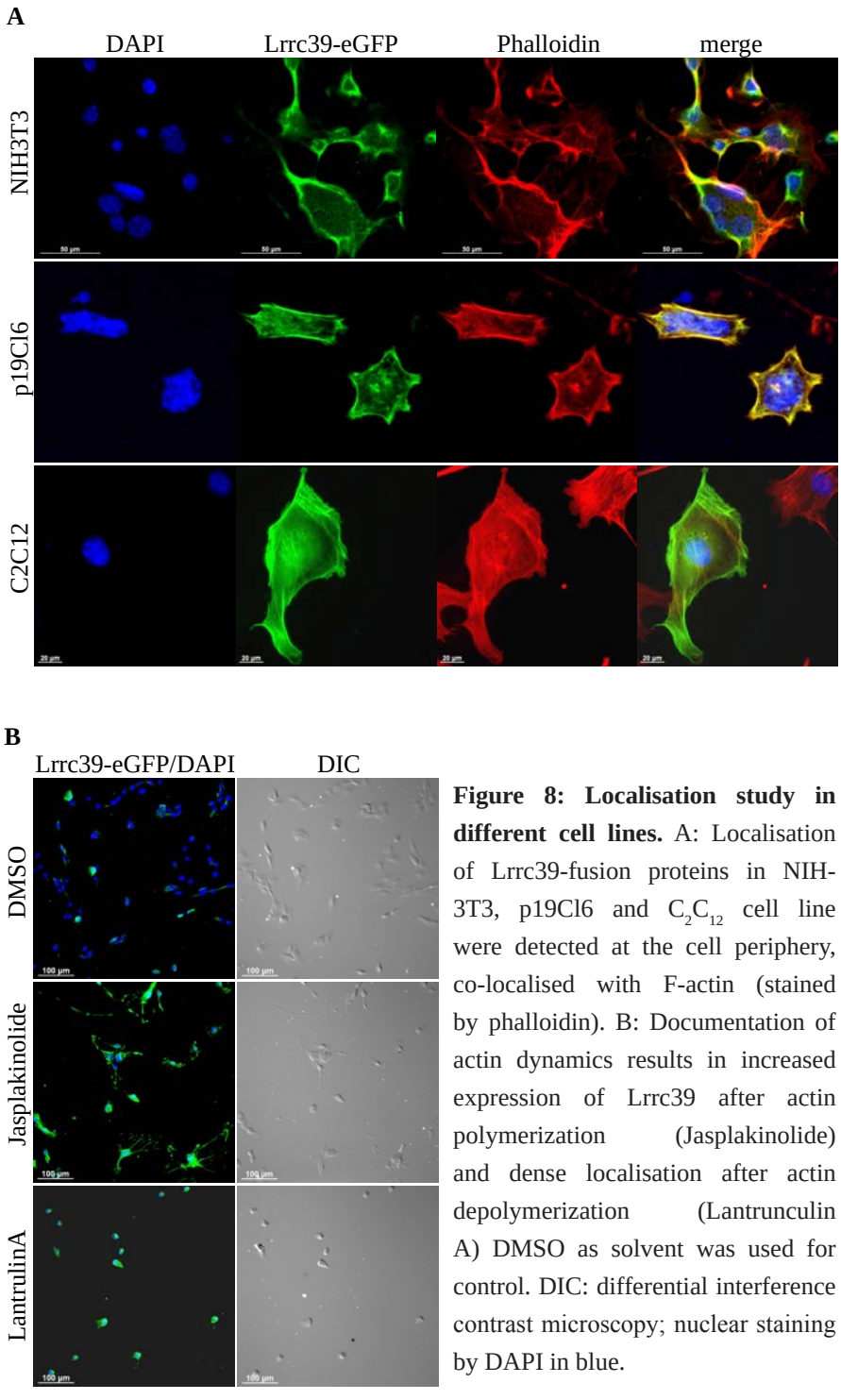


Figure 8: Localisation study in different cell lines. A: Localisation of Lrrc39-fusion proteins in NIH-3T3, p19Cl6 and C_2C_{12} cell line were detected at the cell periphery, co-localised with F-actin (stained by phalloidin). B: Documentation of actin dynamics results in increased expression of Lrrc39 after actin polymerization (Jasplakinolide) and dense localisation after actin depolymerization (Lantrunculin A) DMSO as solvent was used for control. DIC: differential interference contrast microscopy; nuclear staining by DAPI in blue.

Although co-staining with the cellular actin filaments was shown above, the utilised cell lines do not (or only weakly under certain culture conditions as in the case of C2C12 cells) contain the typical sarcomeric structure of the musculature. Therefore, primary cardiomyocytes were used for immunostainings. Unlike the above used cultivated cells, in freshly isolated murine adult cardiomyocytes, a regular, striated pattern of the expression of *Lrrc39* was observed (Figure 9). Identical expression patterns were obtained in the primary isolated neonatal rat cardiomyocytes. To define the region within this distinctive sarcomeric structure, antibodies of widely used markers were used for co-staining. An antibody against α-actinin that is exclusively expressed in a dense region along the Z-line of the sarcomere was used (Figure 9A and B), which resulted in no co-localisation. In contrast, the co-staining with an antibody against myomesin, which localises at the centre

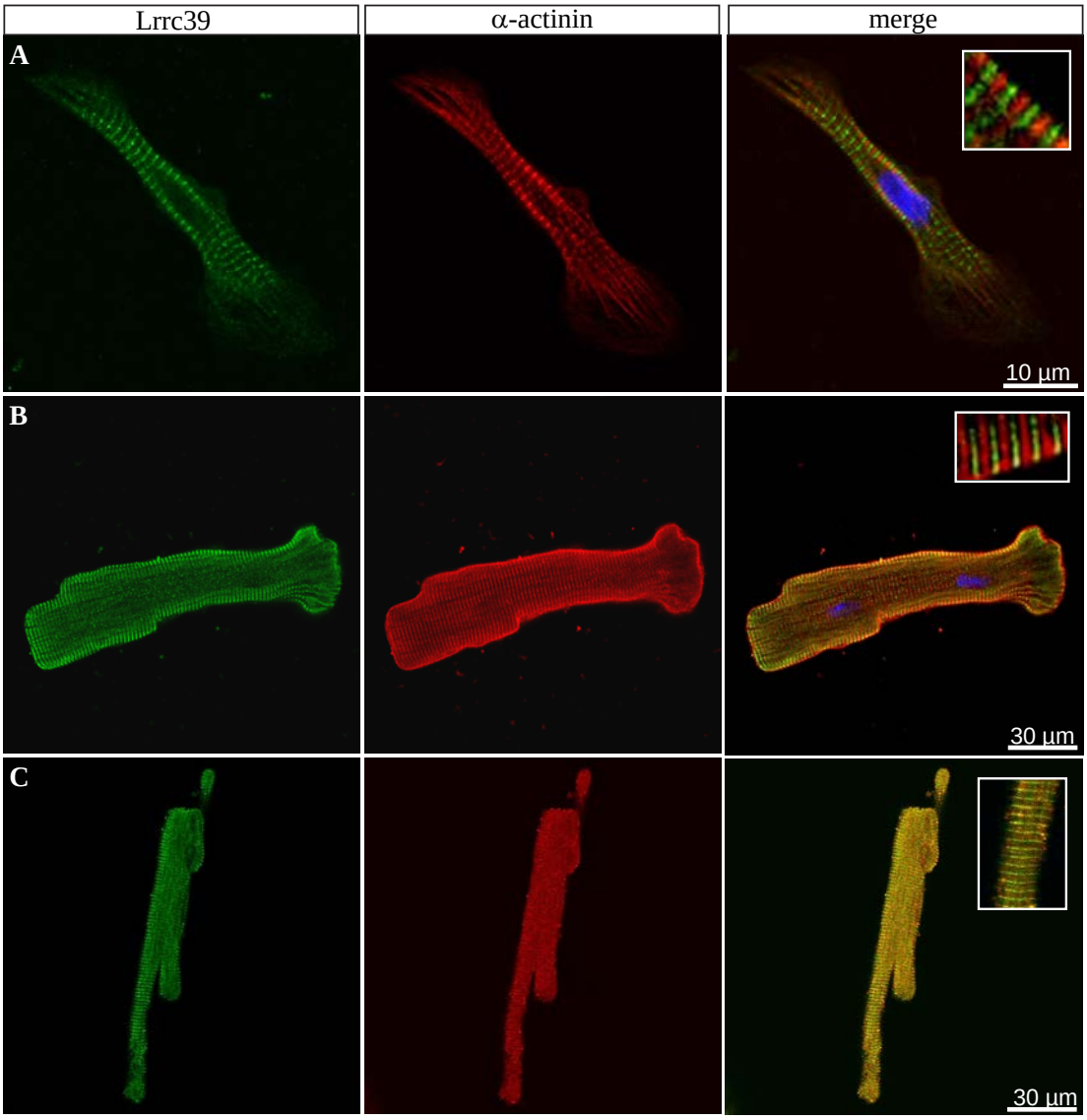


Figure 9: Subcellular localisation in cardiomyocytes. Endogenous Lrrc39 visualised by antibody staining in freshly isolated neonatal (A) and adult cardiomyocytes (B and C). Lrrc39 is expressed in a regular, striated pattern and is co-localised with the M-band marker myomesin (C). Lrrc39 (green), a-actinin and myomesin (red); DAPI (blue).

of the sarcomere, showed the overlapped staining pattern (Figure 9C). As described in Chapter 1.1.2, the composition and the function of this region are not fully understood. Therefore, the unique staining pattern of Lrrc39 indicates the possibility as a novel member of sarcomeric proteins.

5.2 Phenotypic description and morphological analysis

As shown in the above sections, the expression patterns of *Lrrc39* in the heart and skeletal muscle as well as a specific localisation within the sarcomeric structure were observed. In order to investigate a potential function of Lrrc39 in these tissues, the expression level was altered *in vivo*. For this purpose, zebrafish larvae were utilized as a model organism. In order to conduct a *loss-of-function* experiment, microinjections of morpholino oligonucleotides (MO) were performed to reduce the protein level of Lrrc39. On the contrary, an injection of capped mRNA of Lrrc39 to increase its protein level was performed to provide a *gain-of-function* experiment.

For the *loss-of-function* experiment, two independent MOs were used to target the start codon of *lrrc39*. Furthermore, third MO was utilized to provide mis-splicing of *lrrc39* to provide an additional mode of knockdown. As describe in section 2.2.1, three different concentrations of each MO were employed. At the highest concentration of 500 µM many embryos developed severe body edema as well as a high lethality rate was observed. At medium concentration (200 µM), the zebrafish developed cardiac edema and was viable during the sample period. Further, no stiking body edema or severe mutations as seen in the high concentration were detected. Low concentration of 100 µM developed relatively normal and no obvious phenotype was observed. To control the phenotype, age-matched larvae from the same mating without any injection and with injections of a standard MO at the same concentration were used.

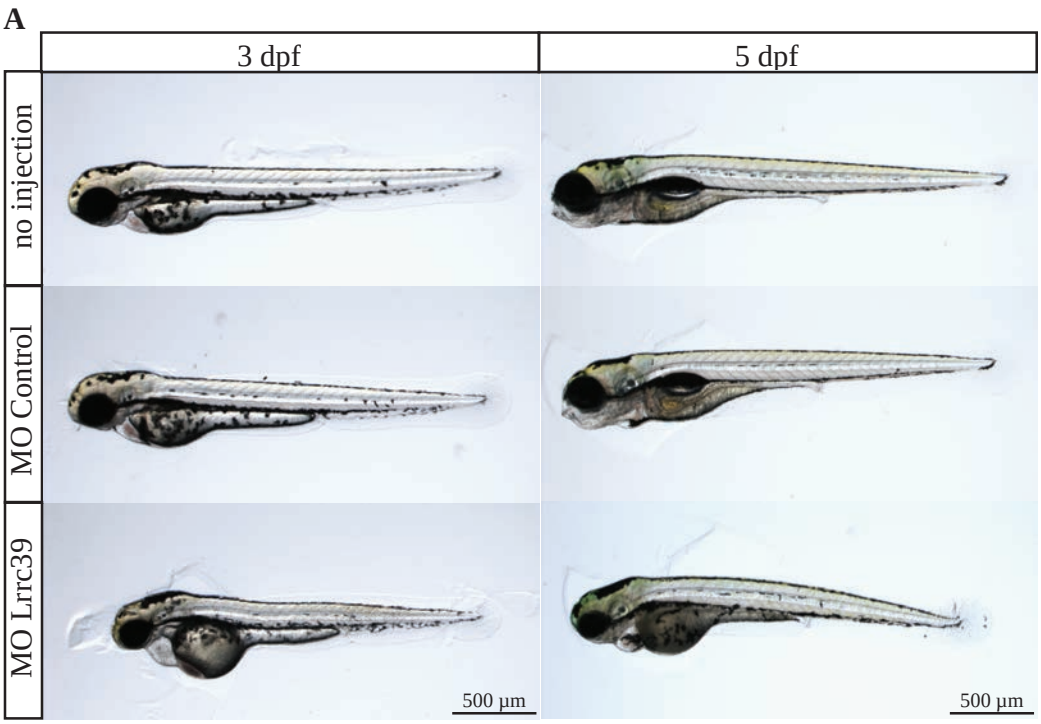
The effects of these different MOs result in the same reproducible phenotype. To provide detailed phenotypic analysis, different zebrafish strains and transgenic reporter lines were utilized as follows. As a wild type strain, Tübingen was used. The transgenic line *Tg(myl7::eGFP-HRAS)s*, which is driven by the myosin light chain 7 promoter, was employed to label cardiac cells. In this strain, the fusion of the GFP tag to the membrane-bound *ras* protein allows an easy observation of the shape and the orientation of cardiac cells. Furthermore, transgenic line, *Tg(acta::eGFP)* was utilized to label skeletal muscle cells through GFP expression under the control of the alpha actin promoter. In the following subsections, the detailed phenotypic analyses using the above techniques and strains are provided.

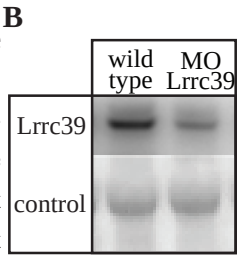
5.2.1 Loss of *LRRC39* leads to alterations of the ultrastructure

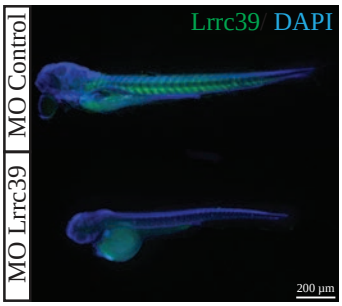
The loss-of-function of *lrrc39* in the zebrafish larvae showed phenotypic changes that are evident from 3 dpf on as shown in Figure 10. The controls develop normally and no differences in the eye, jaw, tail or cardiac development could be detected. Compared to the control injected embryos, the inhibition of *lrrc39* led to a slight reduction in the body size; the eyes appeared smaller; and the lower jaw was deformed. A reduced swimming activity was observed, and the larvae developed pericardiac edema, which is a sign of abnormal heart function. As embryos developed, the deformation in the tail was observed as well as the enlargement of pericardiac edema became evident. These results show that the functions of Lrrc39 are required for the proper development of zebrafish larvae.

In order to confirm the importance of Lrrc39 during the development of zebrafish, a gain-of-function experiment was performed. Although one would expect somewhat of the opposite effects than the loss-of-function experiment. However, as shown by the over-expression of GFP-fused *Lrrc39* in the previous sections, having more of this molecule in the cells than one necessary need might cause impairment in the proper development of zebrafish. To test this hypothesis, an overexpression of Lrrc39 in zebrafish larvae by mRNA injection was performed. Surprisingly, the gain-of-function resulted in the similar phenotype as in the case of the above mentioned loss-of-function experiment. Taken above two experiments together, the requirement of Lrrc39 during the development of zebrafish is inevitable.

To understand the phenotypes from the above two experiments further, a co-injection of *lrrc39* MO and mRNA was perform. This approach is used commonly as a rescue of the phenotype in the area of zebrafish research (Eisen and Smith, 2008). However, the compensation by two extremes (knockdown versus over-expression) did not work. This implies that there exists a delicate balance in which the expression level of *lrrc39* is tightly controlled. In other words, too less or too much of *lrrc39* expression leads to the deformation in the development of an embryo.







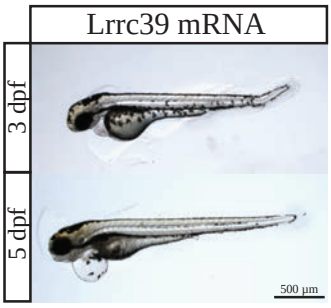


Figure 10: General phenotype after *lrrc39* alteration. A: Brightfield images of Tübingen-strain zebrafish document the development at 3 days post fertilisation (dpf) and 5 dpf. First row shows non-injected control zebrafish and the second row, control zebrafish injected with standard morpholino. These control embryos developed normally. The third row shows zebrafish after inhibition of Lrrc39 by morpholino injection. Reduced body size, smaller eyes, deformation in the lower jaw and pericardiac edema were observed. To control non-specific effects in loss-of-function experiments, age-matched larvea from the same mating without any injection and with injection of a standard MO at the same concentration were used. B: Knockdown efficiency of Lrrc39 was controlled in zebrafish embryos after 3 dpf by Western Blot using Lrrc39 antibody and the Red Alert (Novagen) stained proteins as loading control. C: Immunofluorescent analysis of seletced embryos after 3 dpf showed reduced Lrrc39 staining. D: Gain-of-function experiments show larvae after Lrrc39 overexpression using capped mRNA injection, which resulted in smaller eyes and pericardiac edema that increased during development. Same controls as seen in A have been used. Images were taken with Leica M205 FA.

Since cardiac edema in zebrafish was observed under various conditions and seems to be a rather general phenomenon, detailed phenotypic analysis of the heart was conducted. Microinjections in the transgenic zebrafish line *Tg(myl7::eGFP-HRAS)* resulted in the same phenotype as shown in Figure 10. In order to analyse the heart phenotype in detail, vibratome sections of the larvae were prepared, and the GFP signal was visualized using confocal microscopy. Figure 11 shows the morphological changes after *lrrc39* reduction in the heart of two days old larvae. At this developmental stage, the looping of the heart took place, the atrioventricular canal (AV-canal) is formed, the two chambers of the zebrafish heart are differentiated, and the heart is beating. These early developmental processes seem not to be influenced by the loss-of-function of Lrrc39. In the Lrrc39 MO-injected zebrafisch (MO^{Lrrc39}), the ventricle, which is located in the left part of the image (Figure 11), can be clearly distinguished from the thinner atrium, which is visual in the bottom right part. However, the surface of the ventricle seems in MO^{Lrrc39}. The single cardiac cells appear smaller in size and do not orientate in the direction of the AV-canal, which is shown in the control heart.

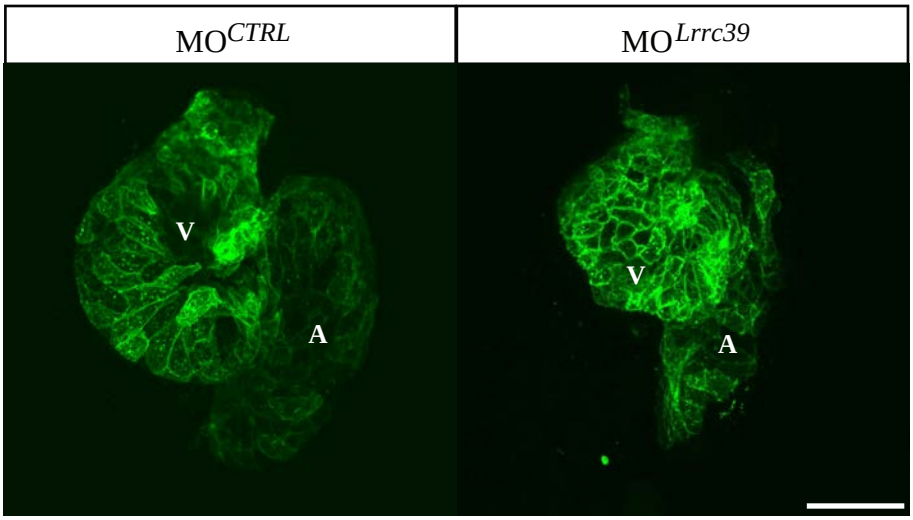


Figure 11: Cardiac phenotype after inhibition of *lrrc39*. The projection of *Tg(myl7::eGFP-HRAS)* zebrafish heart images at 48hpf are shown. The reduction of *lrrc39* (MO^{Lrrc39}) resulted in a deformed ventricle (V) whereas the atrium (A) appeared normal. The morphology of single cells of the heart were visualised by the expression of eGFP. MO^{Lrrc39} showed a reduced cell size and disrupted orientation of cardiac cells compared to the control-injected sibling (MO^{CTRL}). Leica TCS SP2 AOBS. Scale bar: 50 µm.

As Lrrc39 expression was detected in the skeletal muscle of the zebrafish, the tails of the larvae were investigated. As shown in Figure 12, in the absence of Lrrce39, deformations of the skeletal muscle fibres in the tail region were observed. Similar to those of the heart, skeletal muscle fibres appeared smaller and disorientated. Confocal imaging of the whole zebrafish tail allowed for a detailed morphological analysis. Immunostaining

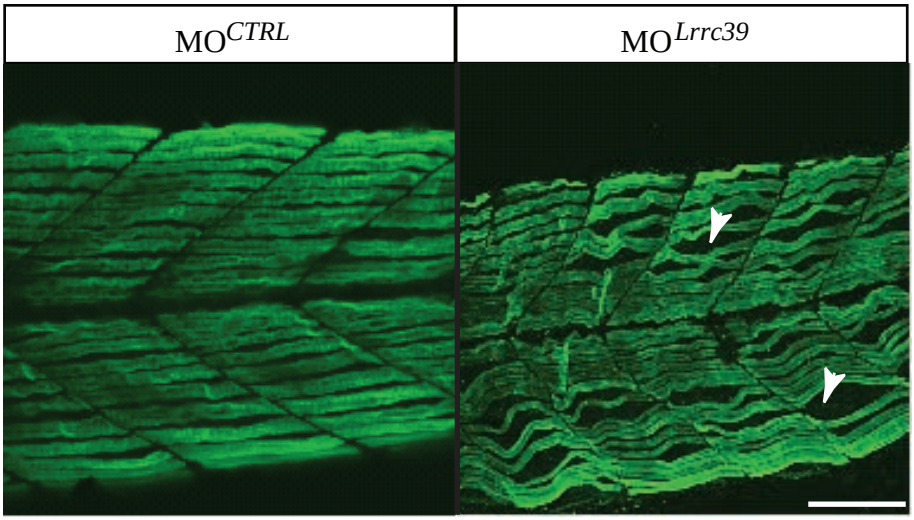


Figure 12: Disrupted skeletal muscle of zebrafish embryos after inhibition of *lrrc39*. The zebrafish tail regions were stained with F59 antibody to mark slow muscle myosins. MO^{Lrrc39} showed thinner muscle fibres with altered orientation. The myofibres appeared curved and showed gaps (arrow) compared to the tightly packed myofibres of the control-injected sibling (MO^{CTRL}). Images were taken with Leica TCS SP2 AOBS. Scale bar: 50 µm.

with F59 antibody were performed to visualise different muscle myosins (Myh1, 2, 4 and 6). Figure 12 shows an optical section of the longitudinal zebrafish tail at the lateral view with the dorsal site to the top. The control injected larvae showed a compact muscle layer. Each of the parallel orientated myofibrils spans one somite. At the same anatomical position, the MO^{Lrrc39} shows smaller somites but a comparable myofibre length as seen in the control. This phenotype is more evident at the macroscopic level as a wavy appearance of the muscle fibres was seen. Although the number of fibres per segment is constant, its thickness is decreased. As a result, unstained areas are visual between the fibres.

As disruptions in the heart- and the tail muscle were observed, additional analysis of the facial musculature were performed to test whether the absence of Lrrc39 leads to disturbance in all striated muscle tissues. The usage of the transgenic line *Tg(acta::eGFP)* allows a continuous observation of the muscle development. In Figure 13, the changes between the control and MO^{Lrrc39} injected zebrafish are shown. The images show the zebrafish larvae from the ventral side with the head of a larva at the top. The background staining in the lower part of the picture marks the yolk sac. At three days post fertilisation (3 dpf), the control larvae had developed the basic muscle fibres of the eyes, the pectoral fins and of the jaw area. In contrast to the control, MO^{Lrrc39} showed only weak GFP signals and the development of these muscles was not dectecable at this early stage . Two days later (5 dpf), a clear α-actin signal was visible at an equally high intensity. Although the muscle fibres of the eyes, jaw area and pectoral fins can be distinguished, the fibres appeared thinner and partially shorter. In addition, the different localization of muscle

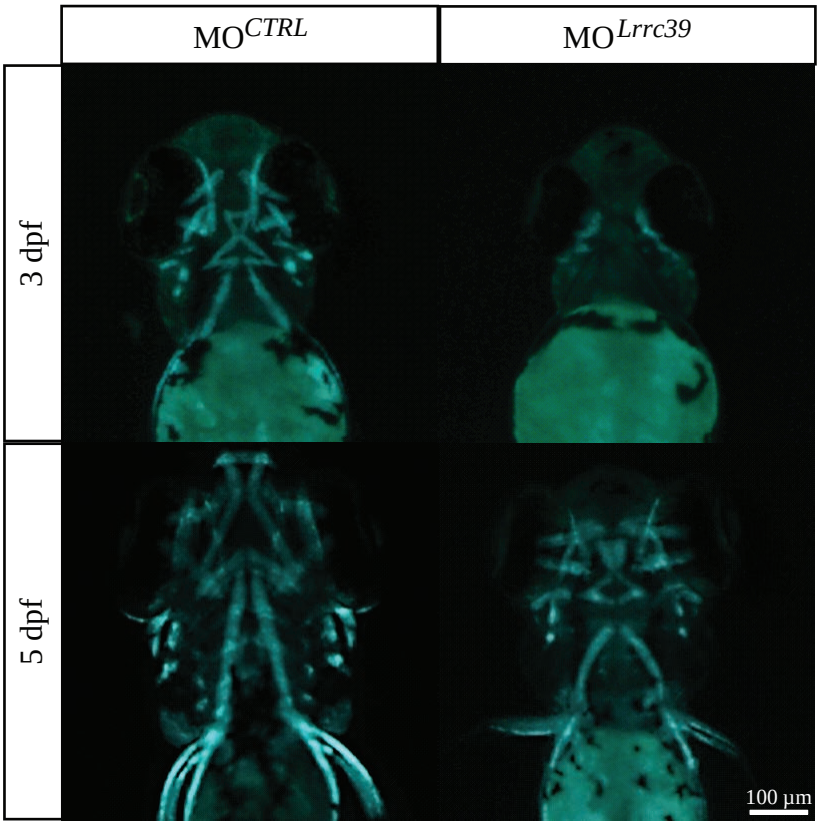


Figure 13: Comparison of facial musculature in control and MO^{Lrrc39} zebrafish. A: Whole mount images of eGFP expression visualised the facial muscles in *Tg(acta::eGFP)* zebrafish. The micrographs show the zebrafish head from ventral views and anterior to the top. The facial muscles of MO^{Lrrc39} were poorly developed after 3dpf in comparison to the control-injected sibling. At 5 dpf, the zebrafish head muscles MO^{Lrrc39} were detectable, but the muscle fibres appeared thinner and misaligned. Microscope: Leica M205 FA.

fibres was observed in MO^{Lrrc39} injected zebrafish compared to the control. This delayed expression of alpha-actin demonstrates an impaired development of facial muscles.

To summarise, the reduction of Lrrc39 in zebrafish larvae leads to wide-ranging malformations of the musculature as shown in Figure 9. In general, the muscle fibres appear smaller and disorganised, which became evident through careful observational studies in the heart (Figure 11), tail muscle (Figure 12) and facial muscle (Figure 13). As a conclusion, *LRRC39* is a necessary gene for the proper development of the striated musculature (both heart and skeletal muscles).

To investigate whether the absence of *lrrc39* is accompanied by changes in sarcomere architecture, transmission electron microscopy (TEM) was performed. For this experiment, the skeletal muscle of the zebrafish tail was chosen for the higher density of sarcomeres within the tissue, which should allow for any structural changes to be detected. As shown in Figure14, MO^{Lrrc39} showed massive changes in the sarcomere organisation at 5 dpf.

The skeletal muscle of the control larvae (MO^{Ctrl}) showed the regular appearance of the sarcomere units. The Z-band can be identified by the localisation of T-tubules as well as the surrounding, light I-band. In the centre of the darker A-band, the thin, dark M-band is visual, surrounded by a narrow light area, the bare zone or H-zone. While the I-band is clearly visible in the control, it seems absent in the MO^{Lrrc39}. Apart from this, the diameter of the myofibres seems irregular and tends to be smaller. Light appearing gaps, surrounding the T-tubules, can be observed. Between the single myofibres, an accumulation of fine structured material can be detected. Figure 14B shows a single sarcomere of the image in detail. The sarcomere length is constant as well as the position of the Z- and M-band are persistent. However, the Z-band region was not clearly structured in MO^{lrrc39}, the M-band appeared darker and the light area of the bare zone around the M-band was not distinguishable. Through the careful analyses at the level of electron microscopy, a striking malformation of the sarcomere organisation in the absence of Lrrc39 was observed.

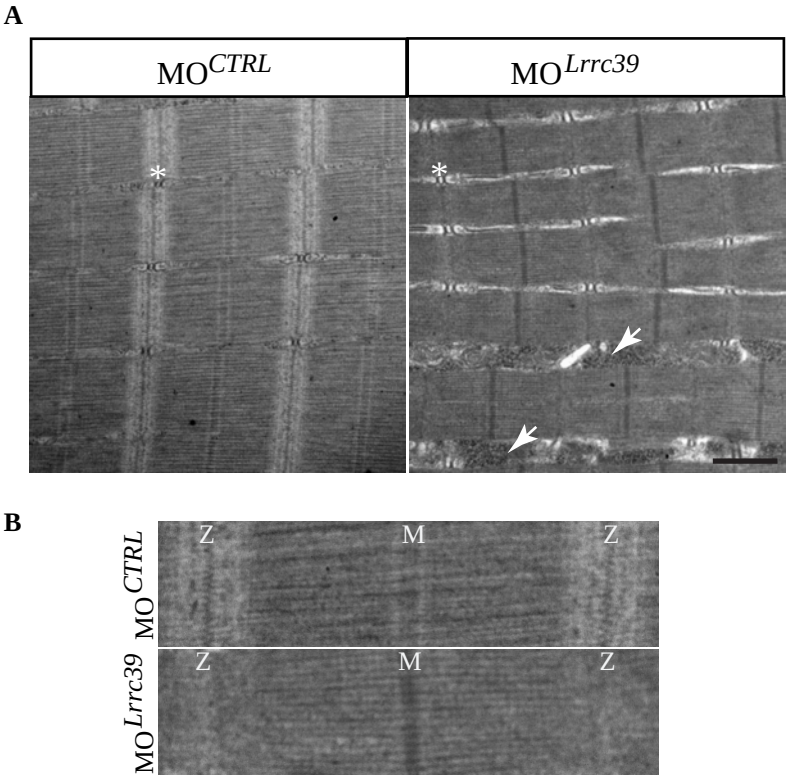


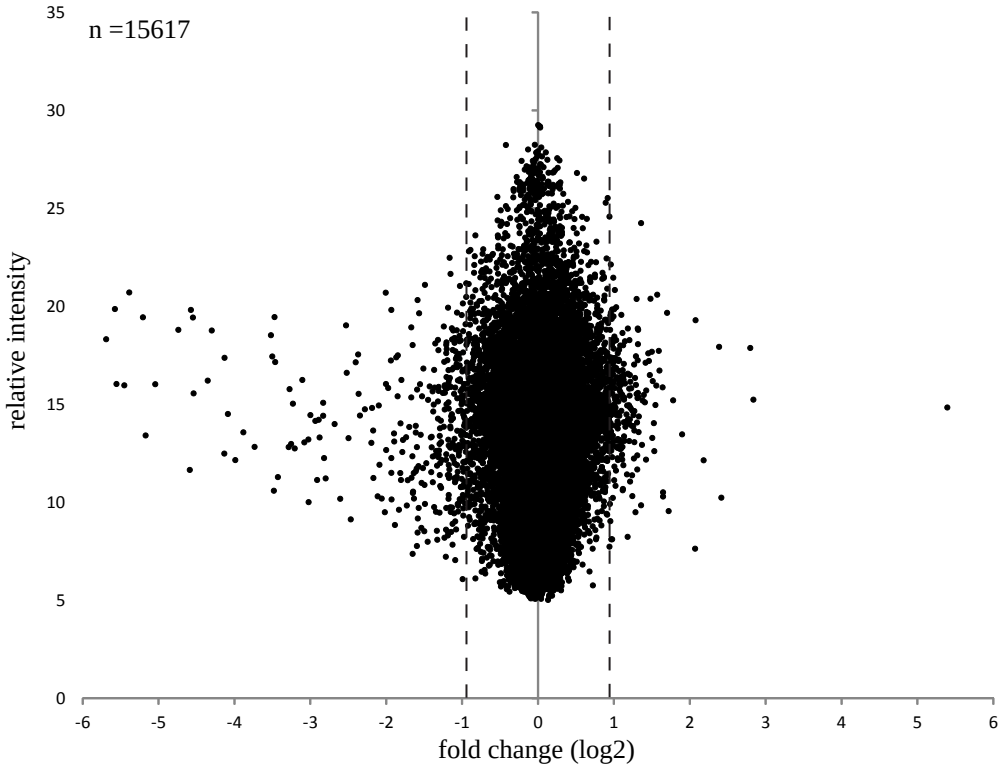
Figure 14: Ultrastructural analysis of the sarcomere. Electron microscopic analysis of zebrafish skeletal muscle of control-injected (MO^{CTLR}) and Lrrc39 morpholino-injected larvaes (MO^{Lrrc39}) at 5dpf revealed in disrupted sarcomeric structure. A: Skeletal muscle of MO^{CTLR} showed regularly organised sarcomeres with persistent myofibre diameter. In contrast, the fibre diameter in MO^{Lrrc39} appeared irregular and gaps around the T-tubules (asterisk) were partially filled with fine structured granules (arrow). B: Detailed images show one single sarcomere defined by two Z-bands (Z) with the centrally located M-band (M). The two sarcomeres shown were consisted in length. The I-band around the Z-band was clearly detectable in MO^{CTLR}, however in MO^{Lrrc39} this region were diffuse and the M-band appeared denser. Electron microscopy Philips CM10 TEM; scale bare: 1 µm

5.2.2 Transcriptomics approach based on microarray technology

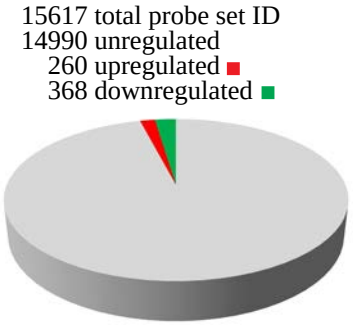
In order to elucidate the reason behind the disturbed muscle development in MO^{Lrrc39}, transcriptomics techniques were employed. Microarray experiments allow a global analysis of the expression intensity of genes. With this techonology, differential gene expression between the two zebrafish samples can be analysed. For the expression analysis, the Affimetrix GeneChip Zebrafish Genome Array was used. This chip contains 15,617 probe sets, which represent to 14,900 zebrafish specific gene transcripts. For this experiment, three zebrafish groups were prepared: (1) Standard MO injected larvae ($MO^{Standard}$); (2) MO^{Lrrc39}; and (3) Specific MO injected larvae targeting a different sarcomeric gene as an internal control (MO^{Ctrl}). Since the earliest visible phenotypes were observed at 3 dpf, this time point was chosen to profile the above mentioned experimental groups. In order to minimise the variation among samples, 15 zebrafish larvae were pooled for each condition. Thus, the replication of microarray experiments were omitted.

Figure 15 shows the distribution of the expression intensities from 15,617 probe set IDs. For the further analysis, expression changes higher than two fold of the standard deviation were considered. This resulted in the selection of 628 regulated probe sets as documented in Figure 15B. Using the complete larvae for the experiment caused a high background mainly from eye related genes, which fitted to the observed phenotype (described in chapter 3.2.1). By sorting the genes for their Gene Ontology (GO) terms, it was possible to filter the results for heart and muscle related candidates. This filter excludes all genes without annotated GO terms for further analysis. As shown in Figure 15C, 196 genes, which corresponds to 1.26% of the total probe set, were selected. Among them, nine genes were differentially expressed with a higher change than two-fold standard deviation. These genes are listed in Figure 15D. Among them, two different ATPases were detected, which were down regulated. These are a calcium-transporting ATPase (*atp2a2*) and a sodium-potassium ATPase (*atpX*). Both ATPases are essential for muscle contraction. As for the up-regulated genes, five genes were identified. One of them is *desmin a,* which is an intermediate filament of the muscle that connects myofibrils to the sarcolemma and investigations in zebrafish reported a crucial role in sarcomere alignment in heart and skeletal muscle (Vogel, 2009). Two genes related to myosin folding, *Unc-45* and *Hsp90,* were also up-regulated. *Unc-45* targets unassembled myosin for proteasome degradation (Kim, 2008) and binds to the chaperone *Hsp90,* which is important for myosin folding. Expression of the transcription factor *foxn4* was increased. In the zebrafish heart, *foxn4* was shown to regulate the AV-canal formation through the expression of *tbx2b* (Chi, 2008). However, the role of *foxn4* in skeletal muscle is currently unknown. The second transcription factor, which was up-regulated was *no tail a,* which is a homolog of *T, brachyury (mouse).* This gene contains a conserved T-box DNA binding domain, which is known for transcriptional regulation, and has been shown to be important for somite patterning in zebrafish (Halpern, 1993).

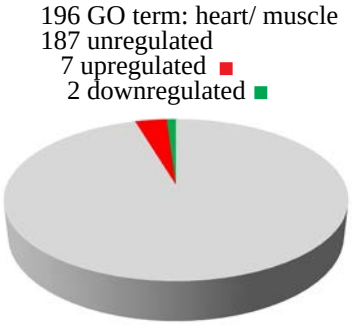
A



B



C



D

gene symbol	gene title	fold change
atp1b2b	ATPase, Na+/K+ transporting, beta 2b polypeptide	0,44
atp2a2a	ATPase, Ca++ transporting, cardiac muscle, slow twitch 2a	0,50
birc5a	baculoviral IAP repeat-containing 5a	1,92
desma	desmin a	2,00
foxn4	forkhead box N4	2,88
hsp90a.1	heat shock protein 90-alpha 1	2,49
noc3l	nucleolar complex associated 3 homolog (S. cerevisiae)	1,98
ntla	no tail a	3,13
unc45b	unc-45 homolog B (C. elegans)	2,60

Figure 15: Gene expression analysis by microarray. A: Comparison of gene expression changes after Lrrc39 reduction. Microarray analyses were conducted on whole larvae at three days post fertilisation. Genes with changes of >2-fold are considered for analysis. B: Distribution of regulated genes in the total probe set and C after filtering for Gene Ontology terms *heart* and/or *muscle*. D: Gene final changes in relation to heart and muscle.

Although the intensities of interesting genes might be masked by more abundant genes (e.g. eye-related genes), it was possible to describe the phenotypes after loss of *lrrc39* to a certain extent when appropriate filters were applied to dissect the transcriptomics data carefully. Given that only nine genes out of 14,900 zebrafish genes on the array were found to be dysregulated might indicate that potential changes in the musculature were obscured by using the whole organism for transcriptional profiling. Furthermore, the miner regulations might argue for a function of *lrrc39* at the level of proteins, instead of RNA, through such events as post-transcriptional modifications. In the end, it is not possible to give clear indications for major pathways involved and makes it difficult to explain the observed phenotype to a full extent. To overcome this problem and with regard to post-transcriptional regulations, zebrafish muscle of control- and MO^{lrrc39}-injected larvae were isolated and prepared for the proteome analysis by mass spectrometry.

5.2.3 Loss of Lrrc39 leads to altered expression of proteins involved in muscle sarcomere assembly

To screen for changes at the protein level, mass spectrometry technology was utilized. Since the above mentioned transcriptomics approach based on microarray technology indicated the usage of whole organisms results in unsatisfying results, the skeletal muscle was chosen to allow for a direct comparison to the electron microscopy results (Figure 14). The zebrafish tails were freshly isolated at 120 hpf to provide the age matched condition to the electron microscopy experiments. For each sample to be read by mass spectrometry, twenty individual tails were pooled for MO^{Lrrc39} and a standard MO control zebrafish. Three replicates were prepared for each experiment. To be able to precisely quantify the changes of protein expressions, the stable isotope labelling (SILAC) (Westman-Brinkmalm, 2011) approach was employed. This technique employs $13C^6$-Lysine (Lys-6) labelled zebrafishes as internal protein standard for protein quantification. The aged-matched, “heavy protein samples” from the SILAC-labelled zebrafish were mixed in a one-to-one dilution with each of the “light samples” from the experiment (MO^{Ctrl} and MO^{Lrrc39}). A comparison of the ratios between samples and the heavy-labelled fish allows precise quantification of protein expressions. The samples were measured using LC-MS/MS. The pre-processing of the data, including background subtraction, elimination of contaminants and normalisation was performed using MaxQuant (maxquant.org) software (Cox, 2009).

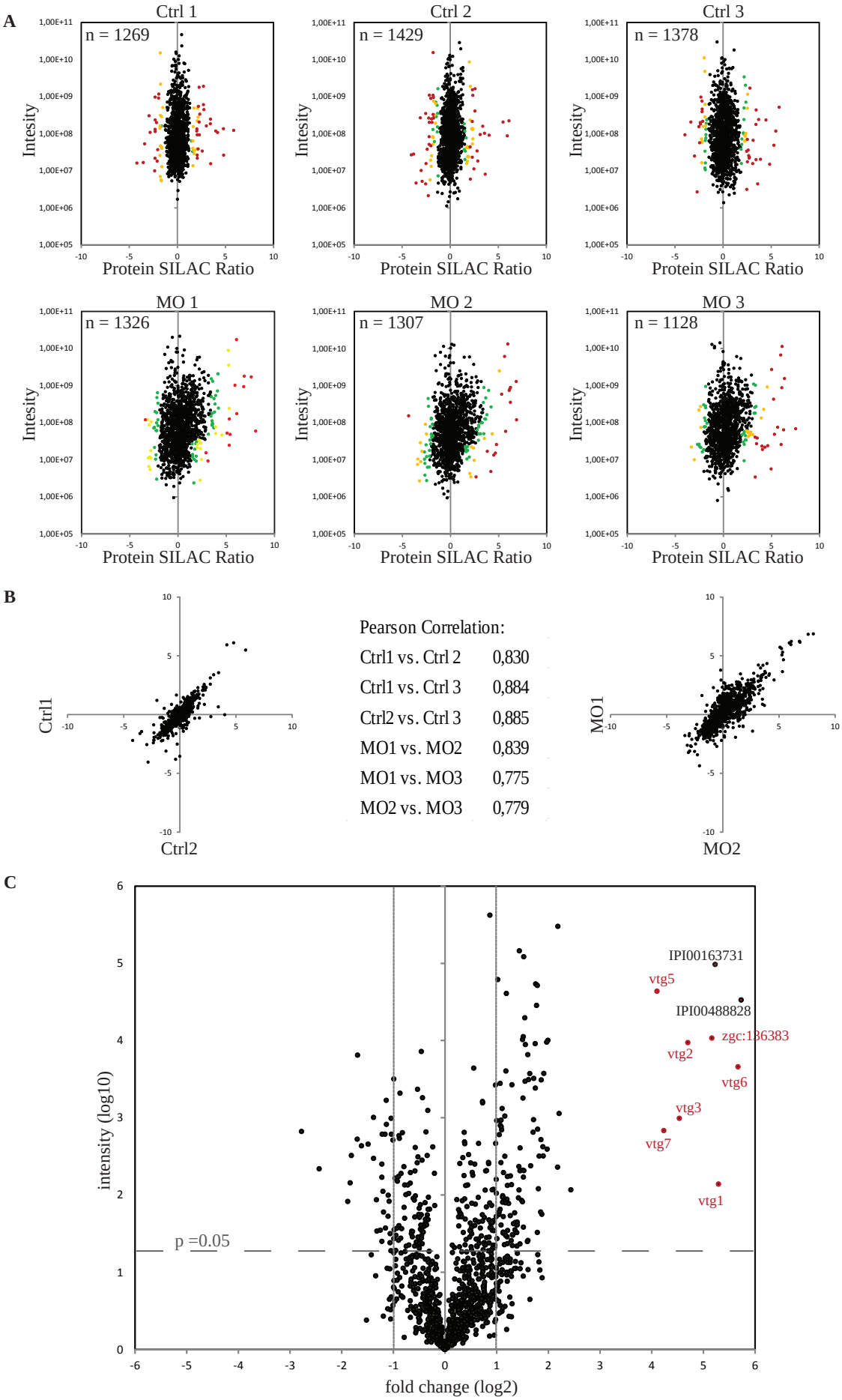

A
Ctrl 1
n = 1269
Ctrl 2
n = 1429
Ctrl 3
n = 1378
MO 1
n = 1326
MO 2
n = 1307
MO 3
n = 1128
Intesity
Protein SILAC Ratio
B
Ctrl1
Ctrl2
Pearson Correlation:
Ctrl1 vs. Ctrl 2 0,830
Ctrl1 vs. Ctrl 3 0,884
Ctrl2 vs. Ctrl 3 0,885
MO1 vs. MO2 0,839
MO1 vs. MO3 0,775
MO2 vs. MO3 0,779
MO1
MO2
C
intensity (log10)
fold change (log2)
p =0.05
IPI00163731
vtg5
IPI00488828
zgc:136383
vtg2
vtg6
vtg3
vtg7
vtg1

Figure 16: Proteome analysis. A: Correlation of experimental zebrafish sample with the SILAC labelled control. First row shows three control replicates (MO^{Ctrl}), second row the MO^{Lrrc39} replicates. Significant changes were marked (green: p<0.05-0.01; yellow: p:0.01 to 0.001; red: >0.001). B: The correlation of the replicates to each other showed the reproducibility of the results. C: Volcano-plot illustrates the expression intensity of analysed proteins. Results were expressed as means; proteins with p< 0.05 and more than two fold change were regarded as significantly changed.

First, the validity of this approach was assessed by inspecting the correlations among samples and internal standard as shown in Figure 16A and B. The ratios of heavy to light sample materials were plotted against the relative intensity. To visually represent the results, significant changes within this groups are marked according to their significance level (green: p<0.05-0.01; yellow: p=0.01-0.001; red: <0.001), and unchanged proteins are marked in black. The altered proteins represent mainly specific compounds of other tissues (e.g. the eyes) and were excluded for the final analysis. The control groups (in the upper panel of Figure 16A) show higher correlations than those of MO^{Lrrc39} groups (bottom panel in the same figure). Such a trend is expected as morpholino injections result in variability in phenotypes, whereas those of heavy-labelled zebrafish are of wild type without injections. To further assess the sample quality, the relationship between each biological replicate was checked. As shown in Figure 16B, correlations of 0.75 to 0.89 confirmed the reproducibility of the results. To give a statistical basis of such correlations, a two tailed t-test was performed between the two groups (P-values < 0.05).

After quality control and removal of ambiguous protein groups, a total of 3,217 proteins was detected Figure 16 C. Among them, 344 proteins were identified to be statistically significant in their expressions (p<0.05) between the two groups (MO^{Ctrl} versus MO^{Lrrc39}). Most outstanding was a group of proteins with increased intensity as high as 17.16 to 50.98 fold (Figure 16C, marked in red). All of these proteins belong to the family of vitellogenins (from lat. *vitellum*, yolk) and were initially described as egg/yolk proteins (Serrano-Pinto, 2004). As the strong increase of these yolk sac proteins differed extremely from the rest of the data set distribution, further investigations were needed to control this result. Vitellogenins are produced as precursors for lipid- and phosphoproteins, threrefore the drastic increase would correlate with the amount of lipoproteins. To examine whether such an increase could be observed, the zebrafish were stained for lipoproteins by Oil-Red-O, see Figure 17.

The comparison between MO^{ctrl} and MO^{lrrc39} showed clear difference between the staining, which correlated with the size of the yolk sac. Since the prepared tail region contained parts of the yolk, changes in lipid- and lipoprotein distribution were likely. This control experiment suggests the increase of vitellogenins as consequence of the sample preparation and may be due to enlarged yolk sac material. Therefore, vitellogenins were excluded from further analysis.

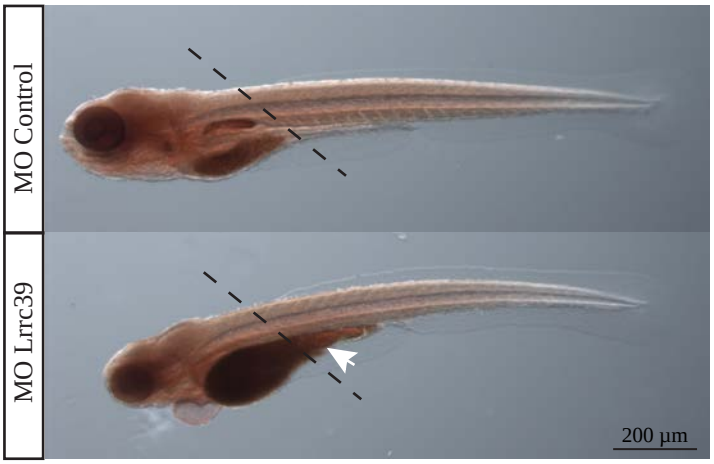
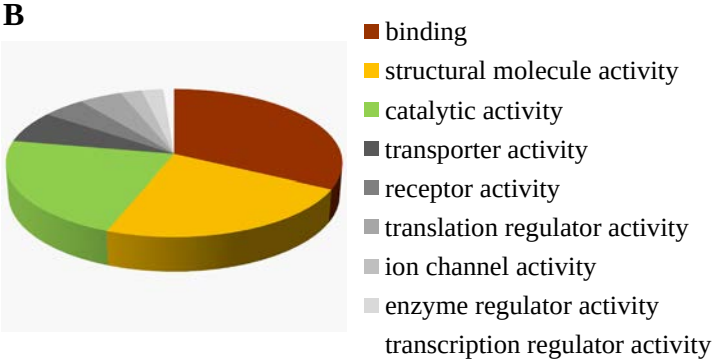


Figure 17: Distribution of lipoproteins in the zebrafish. Oil-Red-O in control larvae (upper panel) shows staining in the yolk. The MOLrrc39 group have comparable bigger yolk sacs with correspondingly increased distribution of the dye. The part of the zebrafish tails that were utilised for proteome analysis is marked by dashed lines and included parts of the yolk sac (arrowhead).

For the further analysis, proteins with more than two fold differences between two groups were selected (157 proteins). To grasp an overview about these proteins, the Panther database (www.pantherdb.org) was utilized to search for Gene Ontology terms. Since not all the proteins have been categorized by the Gene Ontology Consortium (http://www.geneontology.org/GO.consortiumlist.shtml), 71 out of 157 significant changed proteins were identified by the database and clustered by the corresponding GO terms for their molecular function as shown in Figure 18.

A

B



C

gene symbol	gene title	fold change
(CAPN3)	Novel protein similar to vertebrate calpain 3	0.28
cacna1s	Skeletal muscle L-type calcium channel alpha1S subunit	0.28
(myom2)	similar to myomesin 2, partial	0.31
dmd	Dystrophin	0.37
Si:dkey-16k6.1	Novel protein similar to human and mouse capping protein (Actin filament) muscle Z-line, alpha (CAPZA)	0.43
nrap	NRAP protein	0.44
tmod4	Tropomodulin 4 (Muscle)	0.45
(CACNA2D1)	Novel protein similar to vertebrate calcium channel, alpha 2/delta1	0.49
hsp90a.1	Heat shock protein HSP 90-alpha 1	2.23
hsp90ab1	Heat shock protein 90kDa alpha, class B member 1	2.31
hsc70	Heat shock 70kDa protein 8	2.62
hsp10	Heat shock protein 10 (Chaperonin)	3.12

Figure 18: Proteome changes clustered by Gene Ontology terms for their molecular function. A: Distribution of regulated genes in the total probe set; B: Gene Ontology descriptions for the all significant changed proteins ($p<0.05$). C: Selection of identified proteins with significant changes; sorted by increased fold change. A gene symbol in parenthesis represents the gene symbol of a homologous gene.

As shown in Figure 18A, the 71 analysed proteins with significant changes were grouped in nine GO terms. Of them, 56 proteins were grouped in the GO term “binding” (GO:0005488), 38 proteins in “structural molecule activity” (GO:0005198) and 38 proteins in “catabolic activity” (GO:0003824). The sum of these numbers indicates that proteins were clustered in more than one group, which occurred due to multifunctional proteins. Several candidates with structural molecule activity are binding to other proteins,

therefore an overlap of these groups was observed. However, these data point to that MO^{Lrrc39} regulated mainly binding protein, proteins with structural activity and proteins with catabolic activity (Figure 18B). Among these regulated candidates a group of actin-binding proteins were identified, including the following (see Figure 18C):

- Calpain3 (capn3) with decreased expression of 0.28 fold. This protein is located in the sarcomeric Z-and M-band and with its protease activity calpain3 cleaves cytoskeletal protein (Taveau, 2003).
- Si:dkey-16k6.1, the zebrafish homolog for capz (capping protein muscle Z-line) with decreased expression of 0.43 fold. This protein binds actin filaments at barbed ends and protects from polymerisation of depolymerisation of actin (Schafer, 1995).
- N-RAP with decreased expression of 0.44 fold. This scaffolding protein is expressed in the early phase of myofibril assembly and was reported to bind a-actinin and to establish actin polarity (Manisastry09)
- Tropomodulin4 (Tmod4) with decreased expression of 0.45 fold, which binds to pointed ends of the actin filament and regulates thin filament elongation.

The collective importance of these proteins for initial processes during myofibril assembly and their generally decreased expression in MO^{Lrrc39} samples indicated an important regulatory function for Lrrc39.

However, by this way of clustering 86 regulated protein entries were excluded, due to their missing GO annotation. Manually sorting discovered, apart from the mentioned group of vitellogenins, the chaperone hsp90 which are known for role of myosin folding. The hsp90a1 and hsp90-ab1 were up regulated by 2.31 and 2.23 fold, respectively. As it was described in the introduction, the chaperone hsp90 forms a complex hsp70 and unc-45. According to this, the expression of hsp70 shows in the absence of Lrrc39 an increase of 2.62 fold. For the second protein of this complex no change in protein expression was detected. Though, another member of the heat shock family, hsp10 was increased.
Apart from this, the expression of two calcium channels were reduced, which accorded with the microarray results. In the context of the phenotypic analysis of the sarcomere, which showed changes in the M-band appearance, alterations in corresponding proteins were investigated, but no increases were found. Contrary, a protein similar to myomesin 2 was detected with decreased expression. Furthermore, the protein dystrophin, which is known to cause severe muscle damages, was reduced by 0.37 fold change. Taking together, proteome analysis of the zebrafish skeletal muscle in the absence on Lrrc39 resulted in a decrease of actin binding proteins, as well as calcium channels, and in an increased expression of heat shock proteins. These results suggest that Lrrc39 interact with other sarcomeric proteins.

5.3 Molecular mechanism

The proteome analysis of zebrafish muscle after reduction of Lrrc39 uncovered changes is a group of proteins, which were categorised by the GO term “binding”. However, a mere analysis of protein concentration levels does not allow to distinguish between direct and secondary or side effects. In order to understand the mechanism that leads to the distortion of the musculature, the next chapter will focus on the direct interactions of Lrrc39 to other proteins.

5.3.1 Identification of binding partner by yeast-two-hybrid assay

The characteristic leucine-rich repeat (LRR) domain of Lrrc39 is foreseen for protein-protein interactions (Bella, 2008) and together with the information from the mass spectrometric analysis it was likely that this protein is binding. In order to utilize such information, a screening for binding partners of Lrrc39 was performed using yeast-two hybrid method. The full length mouse *Lrrc39* was cloned into the pGBKT7 vector (Clontech) and used as bait protein. The yeast-two-hybrid screening was performed at the DKFZ, Heidelberg.

This screen for the murine Lrrc39 resulted in eight interacting proteins, shown in Figure 19A. Out of these eight candidates, five belong to the muscle-myosin class II, which was introduced in Chapter 1.1. Among them: the human MYH1 as well as the murine homolog Myh1, Myh3, MYH7 and Myh8. The binding of Lrrc39 occurred in all five proteins in identical positions between 846 and 1530 aa, shown in Figure 19B. These regions correspond to the coiled-coil myosin heavy chain tail domain (also known as rod domain), which provides the structural backbone of the thick filament in the muscle sarcomere (Zoghbi, et al., 2008). Moreover, the binding region is located in the S2 fragment of the tail domain, which connects the globular myosin heads and the rod region. Apart from the myosin heavy chains, three additional binding partners were detected: H19, Rbm25 and a riken clone 2900073G15Rik. The identified *2900073G15Rik* gene has two homologs in human as well in zebrafish namely myosin light chain 12a and 9, regulatory (MLC12 and MLC9). These two genes show sequence identity of 90%, using ClustaW for sequence alignment. Although, the screen resulted in additional potential binding proteins, the striking group of MHC was further investigated.

A

Prey gene symbol	Description
Interacting Proteins:	
MYH1	myosin, heavy chain 1, skeletal muscle, adult
Myh8	myosin, heavy polypeptide 8, skeletal muscle, perinatal
MYH7	myosin, heavy chain 7, cardiac muscle, beta
Myh1	myosin, heavy polypeptide 1, skeletal muscle, adult
Myh3	myosin, heavy polypeptide 3, skeletal muscle, embryonic
H19	H19 fetal liver mRNA
Rbm25	RNA binding motif protein 25
2900073G15Rik	RIKEN cDNA 2900073G15 gene
Uncertain interactors: Preys that that were isolated only once with this bait	
COX7A1	cytochrome c oxidase subunit VIIa polypeptide 1 (muscle)
Myh6	myosin, heavy polypeptide 6, cardiac muscle, alpha
Likely false Positives: promiscuous preys	
Atp5a1	ATP synthase, H+ transporting, mitochondrial F1 complex, alpha 1
PRDX6	peroxiredoxin 6
RPS20	ribosomal protein S20
COX2	cytochrome c oxidase subunit II
PGM1	phosphoglucomutase 1
COX1	cytochrome c oxidase subunit I
False Positives: fragments from the 3' UTR	
SLC31A1	solute carrier family 31 (copper transporters), member 1
Dab2	disabled homolog 2 (Drosophila)
Wdr68	WD repeat domain 68

B

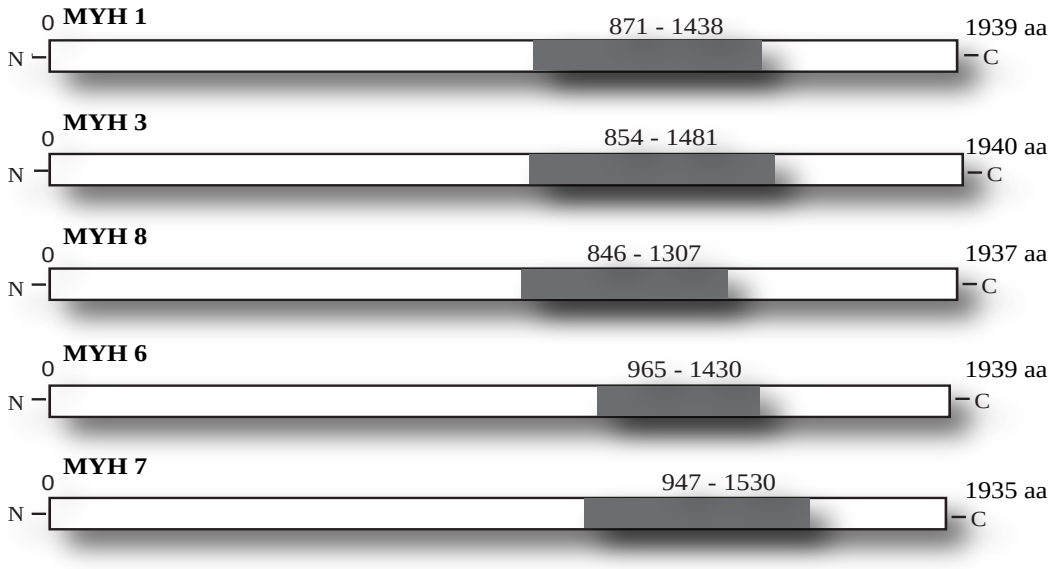


Figure 19: Binding partners of Lrrc39 by Yeast-two-Hybrid screening A: List of binding partners from Y2H screen (provided by DKFZ service). The results were subdivided in definite- and uncertain interacting proteins and likely false positive and definite false positive interacting proteins. Among the first two groups six myosin heavy chain proteins were recognized. B: The Lrrc39 binding region (grey box) of the different myosin heavy chains (*MYH1, 3, 8, 6* and *7*) resulted in a similar pattern at the myosin tail region.

5.3.2 Co-immunoprecipitation experiments confirm an interaction of Lrrc39 with type II myosins

In order to verify the binding partners from the yeast-two-hybrid screen, immunoprecipitations of Lrrc39 were performed. Since the Lrrc39 antibody was not suitable for classical immuno-precipitations, the labelled fusion constructs were utilised. To exclude any disturbance of the interaction by the tag, different fusion constructs were designed and tested as described in Chapter 2.1.2. The full length *Lrrc39* as well as the potential binding region (939-1579 bp) of *Myh1*were prepared with N-terminal GFP fusion (*Lrrc39-eGFP; Myh1-eGFP*) and with N-terminal fusion of V5 and 6xHistidine (*Lrrc39-V5/his; Myh1-V5/his*), respectively. To confirm the interaction, freshly isolated chicken cardiomyocytes were co-transfected with *Lrrc39-eGFP* and *Myh1-V5/his* and analysed after GFP pulldown assay. To control the experiment, empty vectors were included as negative control.

As shown in Figure 20, after precipitating Lrrc39-eGFP, GFP antibody staining detects a

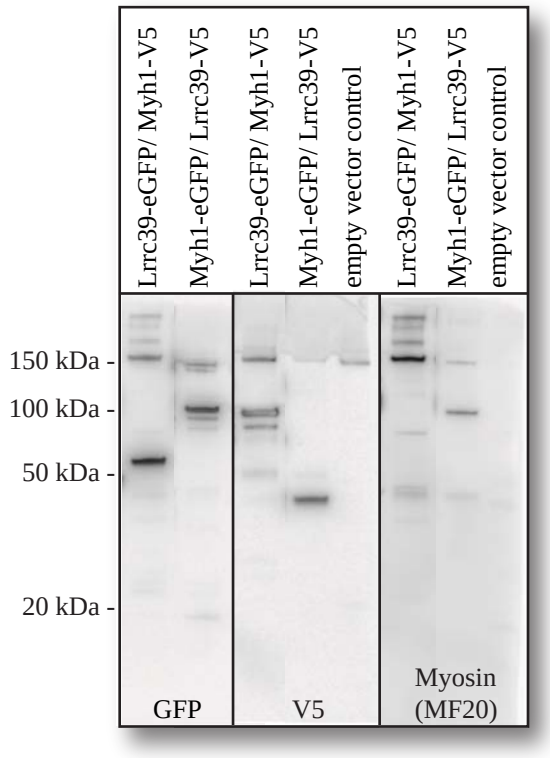


Figure 20: Lrrc39-Myosin interaction by immunoprecipitation in primary chicken cardiomyocytes. Western Blot analysis after GFP-precipitation of co-transfected cells show: 1. GFP antibody staining (Abcam) to control the precipitation procedure; 2. The V5-labeled binding partner and 3. Endogenous myosin stained with MF20 (Hybridoma). After precipitation of Lrrc39-GFP, Myh1-V5 was detected (1st lanes). Vice versa, after Myh1-GFP precipitation, Lrrc39 is detected by V5 antibody (Novex) staining (2nd lanes). Empty vector was used as negative control. Utilised tagged-fusion constructs are shown above.

strong band of ~66 kDa, which is the expected size of the fusion protein (39 kDa + GFP epitope of ~27 kDa). This proves that the experimental setup is working. V5 antibody staining detects a strong band with a size of around 75 kDa, a weaker second band below, and a third band at ~150 kDa. The prepared Myh1binding region contains 1920 bp, which results in a calculated band size of ~71 kDa, plus the fused V5 epitope and the histidine tag of ~2.2 kDa. The strongest band at ~75 kDa was slightly above the expected size of ~73.2 kDa for Myh1. To prove the binding to myosin, additional antibody staining for MF20 (Hybridoma), which recognise muscle myosins was performed. This antibody did not recognise the Myh1 fragment at ~75 kDa, but detected a strong band at ~160 kDa, followed by additional bands between ~180 and ~250 kDa in size. These bands with high molecular weight were not detected by V5 antibody. This result confirms the interaction between Lrrc39 and Myosin. To control the assay furthermore, the experiment was repeated with exchanged fusion tags. Accordingly, Myh1-eGFP was precipitated and Lrrc39 was read out by V5 antibody staining (2nd lanes). The GFP staining to control the precipitation detected a band at ~100 kDa, which matched the calculated size of Myh1-eGFP (99 kDa). The V5 antibody staining detected a ~40 kDa band, which showed Lrrc39-eGFP. MF20 staining revealed in two weak bands at ~100 kDa and 150 kDa.
With this approach, the interaction between Lrrc39 and the predicted region on Myh1 was confirmed. After precipitation of Lrrc39 and analysis of the eluent, Myh1 was identified and vice versa. Furthermore, by utilizing the ability of MF20 to recognize the heavy chain of sarcomeric myosins (Gonzalez-Sanchez and Bader, 1984), the binding of Lrrc39 to various myosin heavy chain proteins, which were identified by Y2H, was confirmed. Taken together, the view that Lrrc39 is a novel component of sarcomeres was validated. Further experiments in HEK293Tcells entail the same results.

5.3.3 Confirmation of binding by structural modelling

To understand how Lrrc39 interacts with myosins, computational predictions based on homology modelling was conducted. This *in silico* prediction allows for further understanding of Lrrc39 as integral part of the sarcomere.
Visualisation of 3D protein structures requires NMR or crystallographic analysis, which was not available for LRRC39. To test the validity of the results and to visualise the binding of this protein, its three dimensional structure was predicted based on homology modelling. Assuming that, a similar primary structure is likely to form a similar conformation the protein internalin-A (PDB entry: 2omxA) that shows highly similar sequence patterns (Wollert, 2007) was chosen as template. For molecular visualisation the software PyMOL version 1.3r1 (pymol.org) was utilised. The predicted 3D structure of Lrrc39 forms 12 parallel orientated β-sheets with an arc-like shape. This arrangement leads to a concave appearance of the protein and creates a potential binding pocket with a diameter of approximately 30 Å. Within this region conserved amino acids are straightened towards the core of the binding pocket. These preferentially aromatic amino

acids like tryptophan and phenylalanine, which are marked in Figure 21A, are likely to play an active role in the binding-process.

In order to receive an impression about the binding between Lrrc39 and MHC, the two protein structure information were modelled in concert. The structure information for α-helical myosin heavy chain (PDB entry: 3DTP) was taken from (Alamo, 2008). As shown in Figure 21B and C, the model of Lrrc39 shaped exactly around the α-helical coiled-coil domain of MHC. This view supported this protein-protein interaction and was

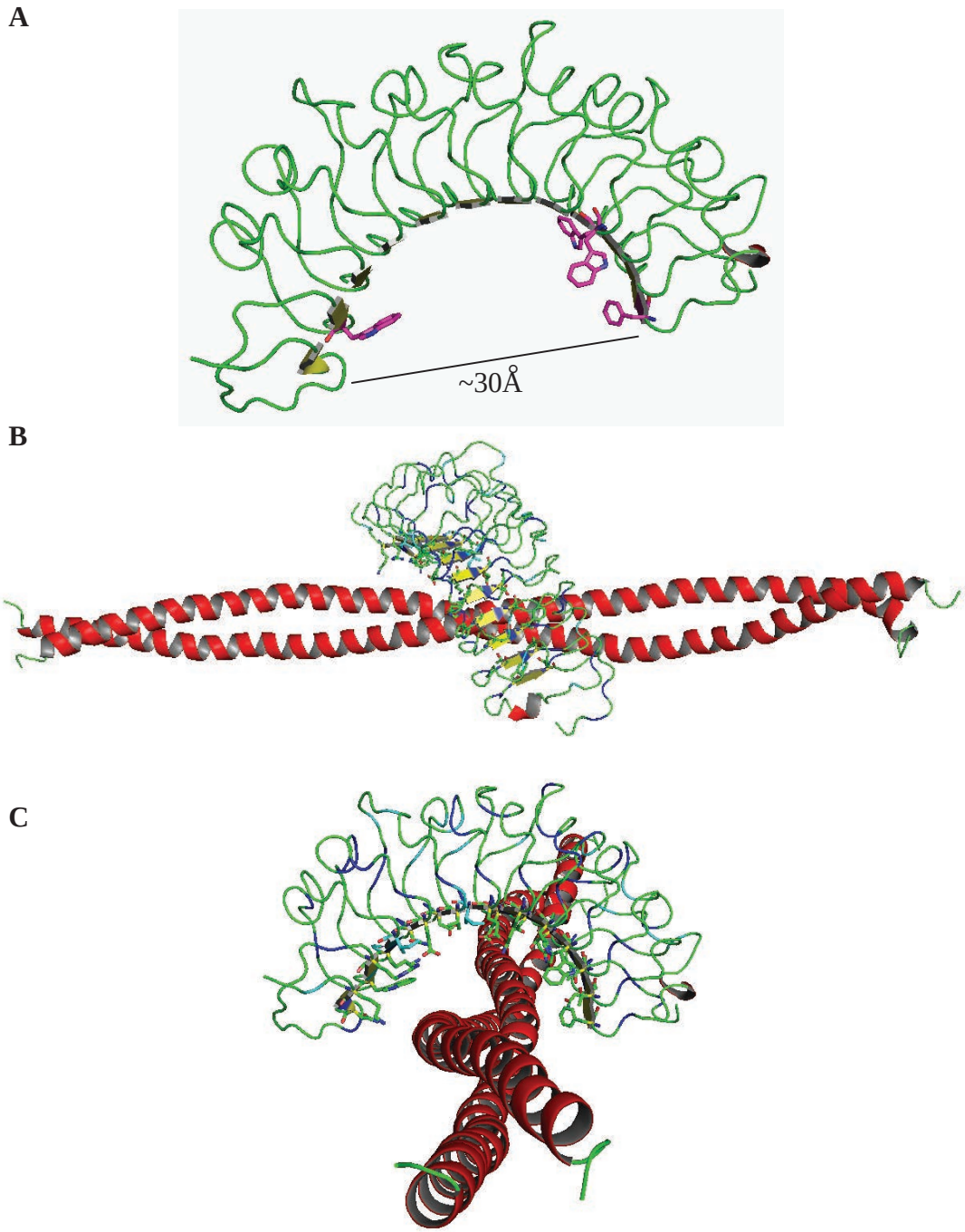


Figure 21: Predicted protein structure of Lrrc39. A: Ribbon diagram of calculated Lrrc39 protein structure viewed from the frontal side, i.e. from the n-terminus on the left to the c-terminus at the right side of the image. B: Structural modelling of Lrrc39 and its binding partner MHC (red) from lateral view and C from frontal view.

further suggestive of a regulatory action through the aromatic amino acids in the core of the binding region.

The results of the yeast-two-hybrid screen identified a conspicuously conserved binding region on diverse muscle myosins that are illustrated in Figure 19B. This conservation suggested a functional role of this region. To examine whether any known characteristics

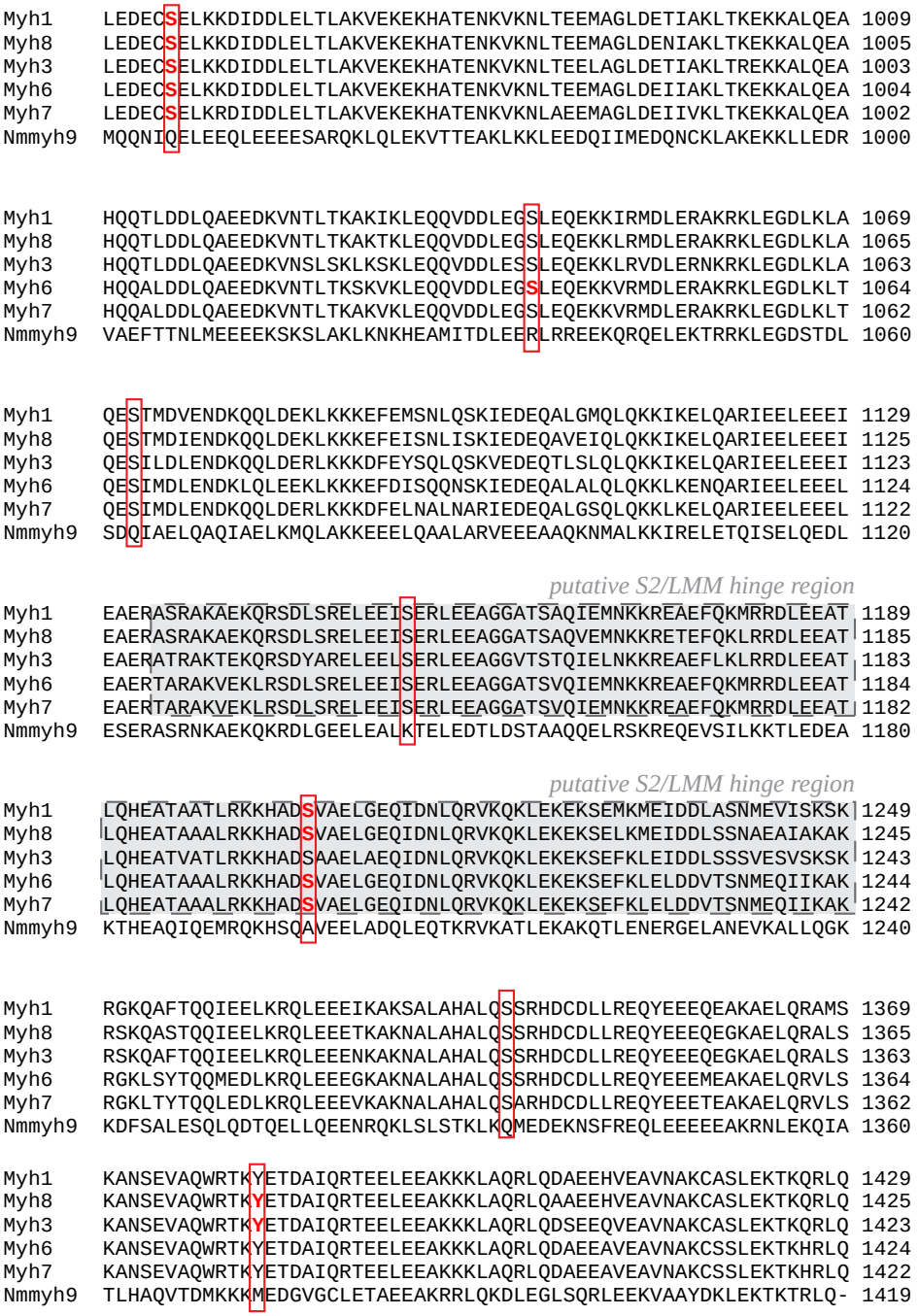


Figure 22: Conserved phosphorylation sites of myosin heavy chains. The sequence of Myosin binding region between ~950 and ~1420 bp comprises conserved localisation of phosphorylation sites (marked in boxes). The known phosphorylations of serine and threonine are labelled in red. The hinge region between the S2 fragment and the light meromyosin (LMM) region is marked in grey (Lu and Wong, 1985). The sequence of non-muscle myosin heavy chain 9 (Nmmyh9) served as control.

are located on MHC that could be influenced by Lrrc39 binding, the protein sequence was further analysed by bioinformatical means. Within the Lrrc39 binding areas on myosin heavy chains, a highly conserved pattern of mapped phosphorylation sites (Phosphosite.org) were detected. As shown in Figure 22, several predicted phosphorylation sites were identified within the binding region.

To investigate whether Lrrc39 interaction could influence these identified phosphorylation sites, potential changes in the phosphorylation pattern were investigated. Hek293T cells were co-transfected with tagged version of Lrrc39 and the binding region of *myh1* fused to GFP. Gamma-ATP supplementation *in vitro* allows the detection of general phosphorylation. After immunoprecipitation of GFP, the phosphorylation status of the MHC binding region changed significantly depending on the presence of Lrrc39 (Figure 23).

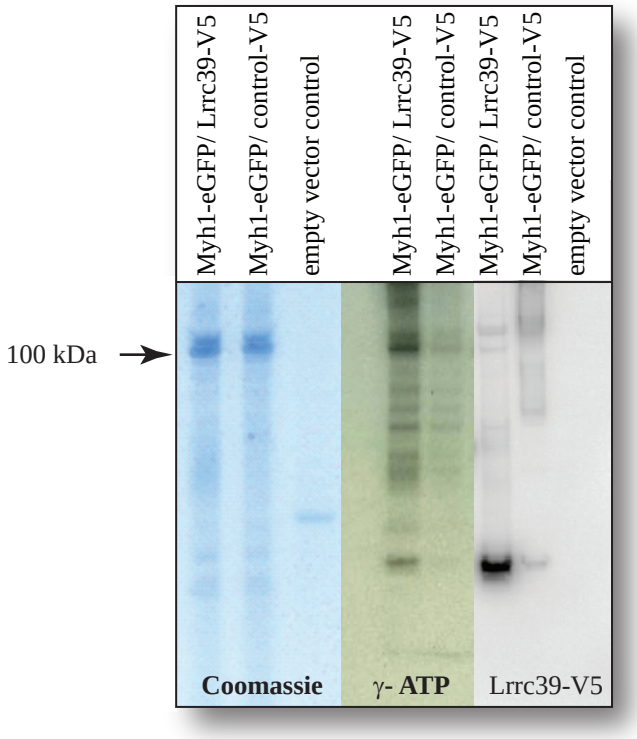


Figure 23: Alteration of phosphorylation status through Lrrc39 binding. Hek293Tcells expressing Myh1-GFP fragment with Lrrc39-V5 (first lanes) or Control vector (second lanes). A: Coomassie staining showed the Myh1-fragment at 100 kDa band size in equal amounts of loaded proteins. B: Darker bands represent increased 32P incorporation to show phosphorylation level C: Western blot against V5 detects Lrrc39 fusion protein.

Altogether, the binding partner analysis of Lrrc39 results in the protein myosin heavy chain, which was identified by Yeast-two-Hybrid screen, confirmed through immunoprecipitations and the predicted structural view completes the picture. Given the recent increasing interests in the field of proteomics and modifications thereof of various proteins, this phosphorylation of Myh1 by the overexpression of Lrrc39 shed a new light on the possible functions of yet uncharacterised modifications of Myh1.

5.3.4 Differential localisation of the titin M-band region

To investigate the role of the Lrrc39 interaction, the protein was over-expressed in neonatal rat cardiomyocytes, using a GFP- fusion protein. In the same experimental setup, different subcellular localisations were observed, which was unexpected due to the consistent results from the endogenous Lrrc39 localisation study (shown in Chapter 3.1.3) To minimise any potential interference of the GFP-tag regarding protein-protein interaction, three different fusion constructs were used. The N-terminal and the C-terminal fusion of GFP as well as the N-terminal fusion of V5 and His led to the same result. To summarise the results after several repetitions of the transfection with the different constructs, the subcellular localisation of Lrrc39 appeared dynamic. The fluorescent signal of the C-terminal fused GFP of Lrrc39 was analysed in neonatal rat cardiomyocytes. As shown in Figure 24, the protein was localised in the periphery and appeared in a dense and intensive signal at the edges of the cell, which was observed likewise in pre-experiments in C2C12 cell line.
In the same experiment the overexpressed Lrrc39 fusion protein was found in a 'patchy' pattern in the cytoplasm, with a distribution preferentially in the cell periphery (Figure 24B). The localisation in Figure 24C was equally distributed in the cytoplasm, in relatively thick aligning patches. In Figure 24D the signal appeared more structured, in fine repetitive lines, which were observed likewise using antibody staining for endogenous Lrrc39.

Counter staining were performed in the same cells using antibody against Titin M8, a C-terminal domain of Titin, which is located at the M-region. The expected pattern of M8 was observed in untransfected cells, shown in Figure 24E in the upper cell. However, in Lrrc39 transfected cells this pattern was disrupted. Instead small aggregates in the cytoplasm were detected. To assess whether this unexpected finding holds true for myosin-binding proteins, counter staining against MyBP-C were performed, shown in Figure 24F. The MyBP-C staining pattern showed no difference between Lrrc39 transfected and non-transfected cell. Apart from this, the normal staining pattern of MyBP-C in the C-zone of the thick filament alternated, with the fine structured signal of the Lrrc39-eGFP fusion protein, indicating that overexpression of Lrrc39 prevented a specific interaction of MHC with the M8 domain of titin, but did not disrupt all protein-protein interactions of MHC in the M-region.

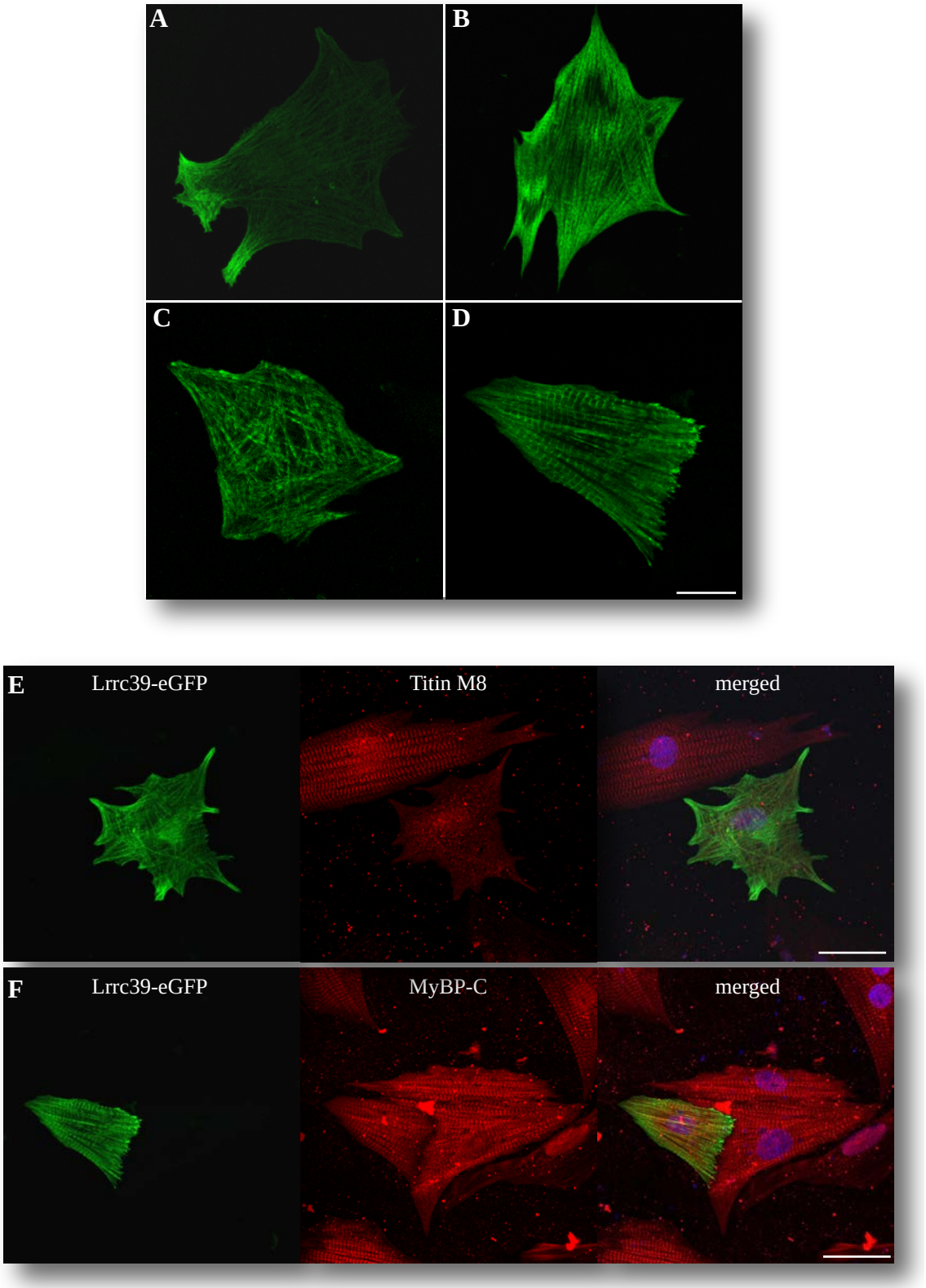


Figure 24: Effect of Lrrc39 overexpression in neonatal rat cardiomyocytes. A- D: Diverse expression patterns of Lrrc39-GFP fusion protein suggest dynamic localisation (scale bar: 10 µm). E: Overexpression of Lrrc39 correlated with disruptions of the c-terminal end of titin, stained with titin M8 antibody; F: Staining pattern of MyBP-C was not influenced by Lrrc39 overexpression (scale bar: 30 µm).

The overexpression of Lrrc39 in neonatal rat cardiomyocytes results in two main observations: First, the subcellular localisation of Lrrc39 fusion-protein appeared in different patterns. And second, increased expression of the protein correlated with disruption of Titin M8 expression pattern. The latter observation suggests that Lrrc39 might have occupied the required space of titin M8, which in turn, disrupts its normal expression.

To sum the results, Lrrc39 is an evolutionary conserved gene, which is highly expressed in the muscle structures (e.g. heart and skeletal muscle). The Lrrc39 protein is necessary for proper development of zebrafish larvae and has to be maintained within a narrow concentration range as shown by loss- and gain-of-function experiments. As demonstrated Lrrc39 resides in the M-band and interact with MHC. Such interaction was visualized by computational simulations. One possible outcome of this interaction is the control of phosphorylation of myosins, which may affect the proper function of the muscles.

„Die Wissenschaft fängt eigentlich erst da an interessant zu werden, wo sie aufhört"
"Science starts becoming interesting just where it ends"
Justus Freiherr von Liebig (1803-1873) aus den 'Chemischen Briefen'

6 DISCUSSION

This study characterises a novel protein of the sarcomere termed Lrrc39 (leucine-rich repeat containing protein39). The presented results evaluate its predicted expression patterns in heart and muscle tissues by different techniques as follows: RT-PCR and whole mount *in situ* hybridisation experiments revealed the evolutionary-conserved expression patterns in the striated musculature. To investigate the functional impact of this gene, *loss of function* experiments in zebrafish were conducted, which revealed disruptions of cardiac and skeletal muscle tissues. More specifically, disorientated muscle cells, which were reduced in size, were observed. Detailed analysis of structural components by electron microscopy exposed disturbed sarcomere organisation. To elucidate the mechanisms behind the causes of such phenotypes upon the loss of Lrrc39, transcriptomics and proteomics analyses were performed, which indicated expression changes of key proteins for the myofibril assembly. Although OMICS approaches employed here yielded hints about the function of Lrrc39, it remains unclear how this protein initiates this changes and if the regulations are due to direct interactions or caused by secondary effect. To address this point, subcellular localisation studies and protein-protein interaction analysis were conducted. Remarkably, the binding partner screening resulted in the identification of the α-helical tail region of the well-known structure proteins myosin heavy chain (MHC). Although the head domain of MHC proteins has been investigated extensively in regards to their motor activity for muscle contraction, the properties of the α-helical tail region is less understood. To address this point, the binding site of Lrrc39 to MHC proteins was simulated by computation and further tested biochemically in regards to possible phosphorylation sites of MHC proteins. Taken the above results together, it can be concluded that Lrrc39 is a novel member of sarcomere whose functions are important for the development and maintenance of muscle structures. To further elucidate the function of Lrrc39 and its impact on the sarcomere, the following aspects will be reconsidered: 1) Where exactly is Lrrc39 localised, 2) how does it impact sarcomere assembly, 3) what are the effects of the binding between Lrrc39 and MHC and 4) what does this interaction means for future investigations of the muscle function?

6.1 Differential localisation of Lrrc39 in the sarcomere

Lrrc39 was originally identified by *in silico* screening as heart and muscle enriched gene of unknown function (Uchida, Jenniches, 2009). To evaluate this screen, detailed sequence alignments of *LRRC39* in seven different organisms was performed (Figure 5), which proved the evolutional conservation of this gene. To validate the expression pattern, RT-PCR experiments using 15 adult murine tissues were conducted, which resulted in clear detection of Lrrc39 in cardiac and skeletal muscle (Figure 6). Thereby, it confirmed the results of the *in silico* screen. Apart from heart and muscle tissues, the expression of Lrrc39 was detected also in eyes and brown adipose tissue (BAT) of adult mice. Since this technique is very sensible, the expression in eyes may be caused by traces of surrounded ocular muscles, which could not be excluded. Apart from this, a role of Lrrc39 in non-muscle cells might be possible. Unlike in any other investigated tissue, the RT-PCR experiments detected a shorter isoform of Lrrc39 in the BAT. This finding is interesting in the context of skeletal muscle as a recent report (Seale, 2008) suggests the common developmental origin and the presence of shared progenitor cells in the BAT and skeletal muscle. A further investigation is necessary to understand the function of this shorter isoform in the BAT.
Apart from this, the expression of Lrrc39 was observed during development of embryos using whole mount *in situ* hybridisation experiments in mouse embryos (Figure 6) and RT-PCR experiments during the early development of zebrafish embryos (Figure 7), which were further confirmed by the published results of Thisse *et al.* (2004) (Figure 7).
As with the case of all of the above employed techniques, the exact localisation of Lrrc39 within the cell is unknown. To address this point, the subcellular localisation in primary muscle cells was investigated using an antibody against Lrrc39. Indirect immunofluorescence in cardiomyocytes revealed a clear, alternating pattern of Lrrc39 (Figure 9). Co-stainings with an antibody against α-actinin, which is localised in the Z-band, revealed that Lrrc39 does not reside in the Z-band. To understand the localisation of Lrrc39, an antibody against myomesin, which is expressed in the M-band, was utilized. As shown in Figure 9, the signal of Lrrc39 overlapped that of myomesin to demonstrate the localisation of Lrrc39 in the M-band. While this study was underway, another group reported the same expression patterns in cardiac cells (Will, 2010), which confirms the findings of this study.

Although the identification of the subcellular localisation of Lrrc39 based on co-staining with marker proteins, i.e. myomesin is promising, upon closer inspection, it contradicts to the very consistent results of the Yeast-two-Hybrid screen (section 3.3.1), which indicated Lrrc39 binding at the S2 region of MHC. This S2 region connects the myosin head domain

with the tail domain, which forms the rod of the filament. The opposing orientation of the myosin filaments leads to the formation of the *bare zone*, where only the myosin rod part is localised including the centred M-band. Therefore, the S2 region is expected to localise everywhere in the thick filament but not in the M-band. As it was shown by Will *et al.*, the Lrrc39 antibody does not co-localise with MyBP-C, which is located in the C-zone of the thick filament. Regarding this, a closer inspection of the thick filament is necessary:
The M-band is a structure of approximately 100 nm, and the subcellular localisation was analysed in this study using antibody staining and confocal microscopy. Given that the resolution of this technique is limited by the wavelength of the light and the numeric aperture of the objective, a diffraction limit of such assay is 200 nm, which is twice as large as ~100 nm for the M-band; thus, smaller structures cannot be distinguished with this technique. Therefore, it can be hypothesised that the observed fluorescent signal of the Lrrc39 antibody might be misleading as the actual subcellular localisation of Lrrc39 is in the proximal region next to the M-band, where the first S2 units appear. One possibility to circumvent this technical problem, immuno-electron microscopy could be utilized to provide a higher resolution and to allow for a detailed inspection of the subcellular localisation of Lrrc39 within the sarcomere.
Taken together, the above interpretation of results opens up the following interesting question to be pursued further: Why Lrrc39 appears in the thin band next to the centre of the myosin filament and not distributed in the complete rest of the thick filament? Two possible explanations can be conceivable: a) Lrrc39 is localised exclusively in the P-zone of myosin or b) Lrrc39 is localised in all S2 regions distributed among the thick filament, but due to the compact structure in the C- and D- zones, the antigen is not accessible for antibodies.
To address the second point, the subcellular localisation studies in neonatal rat cardiomyocytes were conducted (Figure 24). Through the usage of Lrrc39-GFP fusion proteins, a potential interference with the antigen was bypassed. With this approach, Lrrc39 was detected in different localisation patterns: in dense or patchy pattern at the edges of the cell to a fine-lined structure in the cytoplasm (described in Chapter 3.3.4). The appearance of this defined thin localisation of Lrrc39-GFP fusion protein argues against the point that antigens could not have been recognised. This finding supports the following notion: The observed variations in the localisation pattern occur during the sarcomere development. Lrrc39 expression starts in the premature myosin filament and specifies to the localisation in the P-zone as the sarcomere matures. The antibody staining for Lrrc39 in adult cardiomyocytes revealed in the condensed expression at the central part of the myosin filament (as shown in Figure 9) and did not give any evidence for the broader localisation. To verify this assumption, additional experiments are necessary. For example, live-imaging of Lrrc39-GFP fusion protein in transfected cells during their differentiation process, may provide an answer to the question of if the localisation

of Lrrc39 dynamically changes from the broad pattern at the cell periphery to its final position in the thin pattern at the P-zone. If indeed, Lrrc39 localises specifically to the MHC P-zone of the mature sarcomere, this protein might be better renamed in Myosin binding protein-P (MyBP-P).

Furthermore, the localisation patterns of Lrrc39 in neonatal muscle cells (Figure 24A-D) resemble those of α-actinin. The protein α-actinin was reported in several studies to locate in a plainer pattern at the cell periphery where it scaffolds sarcomere assembly and aligns stepwise in fine thin lines (McKenna, 1986; Dabiri, 1997; Gokhin and Fowler, 2011). Therefore, it is tempting to speculate that Lrrc39, with comparable subcellular localisation, might be involved in the important process of the sarcomere organisation.

6.2 The impact of Lrrc39 on myofibril assembly

To analyse a potential function of Lrrc39 for sarcomere organisation, loss-of-function studies in zebrafish were conducted. The phenotypic observations (Chapter 3.2.1) resulted in malformations of the heart shape and disorientated cardiac cells (Figure 11), disruptions in skeletal muscle tissue in the tail of the zebrafish (Figure 12) and delayed formation of the facial musculature, followed by disorientated muscle fibres (Figure 13). These phenotypes in different tissues, which contain cross-striated musculatures, argue for a general role of Lrrc39 during the development of muscles. The analysis of the fine structure on the zebrafish skeletal muscle using electron microscopy provided evidence that sarcomere organisation is altered by the absence of Lrrc39 (Figure 14).

The sarcomere is structured by the actin filaments and the myosin filament. During muscle contraction the actin filament slide along the intervening myosin filament, which lead to shortening of the I-band as well as of the H-zone. The length of the A-band that reflects the myosin filament is not changed. Due to the slightly different sizes in length (actin filament ~1 µm each and myosin filament ~1.57 µm) the I-band should be visible even in complete contracted sarcomeres. Although the observations in the skeletal muscle (Figure 14) showed that the length of the sarcomeres was not significantly changed between MO^{Ctrl} and MO^{Lrrc39}, the Z-band of MO^{Lrrc39} appeared diffuse and the I-band surrounding the Z-band could not be defined. This observation suggests a larger A-band that may explain the reduced size of the I-band.
To elucidate these changes the proteomics approach (Chapter 3.2.3) was utilised. The aged-matched zebrafish skeletal muscle analysis on protein level revealed reduced expression of several sarcomeric proteins. Among these proteins, was the skeletal muscle-specific protease calpain3 (CAPN3). Calpain3 is a muscle specific protease that cleaves cytoskeletal proteins at the Z- and M-Band (Taveau, 2003). This property has

been shown to be crucial during muscle atrophy. It was reported that reduction of calpain3 correlates with muscle degradation (Bartoli and Richard, 2005). Also, targeted mutations of calpain3 in mice resulted in misalignment of the thick filament (Kramerova, 2004). Further, reduced activity of this protein causes a specific form of muscular dystrophy, limb-girdle muscular dystrophy 2A (Richard, 1995).

Apart from this, Si:dkey-16k6.1, a homolog of CapZ (capping protein muscle Z-line) was found reduced in the MO^{Lrrc39} embryos. This protein caps actin filaments at barbed ends and protects from polymerisation of depolymerisation of actin (Schafer, 1995). Reductions of CapZ are associated with a loss of uniform alignment of the actin filament (Pappas, 2008). A further protein is N-RAP that scaffolds actin filaments into the early premyofibril (Manisastry, 2009). It was reported that actin-binding protein N-RAP, which is positioned at the Z-band, modifies the distribution of muscle myosins and non-myosin IIb by post-transcriptional modification (Dhume, 2006). The documented reduced expression of N-RAP may be responsible for the disruption of sarcomere organisation in the MO^{Lrrc39} zebrafish. Furthermore, tropomodulin (Tmod) that was downregulated on protein level. Tmod caps the pointed ends of the actin filaments and thereby protects the thin filament form polymerisation or depolymerisation. Inhibition of Tmod results in elongated thin filaments (Gregorio, 1995).

Interestingly, *in vitro* stimulation of actin dynamics using reagents to stimulate either polymerisation or depolymerisation correlated with Lrrc39 expression (Figure 8). This phenomenon argues for a role of Lrrc39 for actin filament formation and might explain the disrupted sarcomere structure of MO^{Lrrc39} zebrafish detected by the electron microscopy analysis. These interesting observations can be put together to support the fundamental role of Lrrc39 in the sarcomere organisation.

Alterations of the mentioned actin binding proteins were all associated with changes in the Z-band formation. Such changes could interrupt the connection to the sarcolemma of cell-cell interactions due to their important role as anchor point for intermediate filaments that provide the structural link between sarcomeres. Disrupted Z-bands may cause reduced intermediate filament contacts and thereby the contact to the sarcolemma might be reduced. The observed alteration of the protein dystrophin supports this course of events. Dystrophin connects the sarcomeres to the sarcolemma and showed after loss of Lrrc39 in the zebrafish skeletal muscle significant downregulation. Decreased or mutated dystrophin causes the muscle diseases Duchenne- or Becker's muscular dystrophy. This may explain the observed 'dystrophic' muscle phenotype of the MO^{Lrrc39} zebrafish.

These disarrangements of the sarcomere structure and a reduced contact to the sarcolemma may lead to secondary changes in the muscle tissue. Since these changes were observed in vivo, secondary and compensatory mechanisms cannot be excluded. Therefore, it remains unclear to what extent the regulation of calcium channels are direct effects of Lrrc39 reduction. The essential role of calcium transport for the muscle cell is beyond question,

and it is not surprising that even small alterations significantly influence muscle function. (Berchtold, 2000) The observed reduction of calcium channels may contribute to the phenotypic changes in the zebrafish model. Moreover, reductions of all these important factors are likely to cause the disruption of the sarcomere organisation. As a cellular response to the disturbed assembly, the expression of *heat shock proteins* (HSP) (Figure 18) may reflect a general stress reaction (Liu and Steinacker, 2001).
The increased expression of Hsp90a was observed on transcript as well as protein level. Although, HSP are ubiquitous expressed, hsp90alpha1 was demonstrated to play a crucial role during the assembly of sarcomeric myosin filaments. In zebrafish experiments, it was shown by Shao Jun Du and colleagues that a reduction of this chaperone causes significant myosin degradation and increases unc45b expression. Zebrafish developed poorly organised myofibrils (Du, 2008). In *loss of function* experiments of the co-chaperone, unc45b was reported with similar effects on myosin assembly (Bernick, 2010). Even though, a comparable phenotype was observed after reduction of Lrrc39 in zebrafish, investigations documented an increase of hsp90alpha1 and unc-45expression on protein level. It seems contradictive to the observations of Du and Bernick. Unfortunately, *gain of function* studies in zebrafish regarding hsp90 were not yet approached. Therefore, it remains unclear if the observed imbalance of heat shock proteins caused the alterations in thick filament assembly.

Together these data showed a reduction of important proteins for actin filament assembly upon Lrrc39 inhibition. These proteins were shown to be essential for the formation of the premyofibrils. Since disruptions in this initial structure may alter the incorporation of the myosin filament, and the increased expression of heat shock proteins may support misalignments of thick filaments in the developing sarcomere, I would like to conclude that Lrrc39 is required for sarcomere assembly. However, electron microscopy and proteomics do not reveal whether these changes were caused directly by Lrrc39 interaction or were due to secondary changes.

6.3 Interaction between Lrrc39 and myosin heavy chain

To address a direct role of Lrrc39 in the process of sarcomere assembly, its binding partners were screened. As shown in Chapter 3.3.2, the binding of Lrrc39 by Yeast-two-Hybrid revealed different sarcomeric muscle myosin heavy chains (MHC) (skeletal muscle as well as heart isoforms). This suggested a role in different tissues and is in agreement with the observed phenotype in zebrafish. The binding to MHC was validated through different approaches: First, the immunoprecipitations in chicken cardiomyocytes (shown in Figure 20) resulted in the corresponding binding partner, which further proves this interaction. After precipitation of Lrrc39, the co-transfected Myh1-fragment as well as endogenous

myosins (detected by MF20 antibody) were detected by Western blot analysis. Second, the predicted protein structure of Lrrc39 that based on homology modelling revealed in a conformation that perfectly matches to the size of the a-helical coiled-coil domain of the MHC rod. Third, indirect immunofluorescence of primary isolated muscle cells show co-localisation of Lrrc39 with the myosin filament in the sarcomere. These results confirm the Lrrc39 binding to MHC. More important seems the question, what is the function of Lrrc39 at this position?

Lrrc39 binds to a protruding position on the a-helical coiled-coil tail domain of MHC, termed the S2 fragment. This intermediate region connects the myosin head (S1) with the rod part of MHC (LMM) and was proposed to act as lever arm of MHC in order to connect to the actin filament for contraction. Thereby two joints form one at the S2/S1 conjunction and a second at the S2/LMM bond. The connections to the S2 fragment were proposed to act as hinge region as early as 1969 by H.E. Huxley. These flexible parts enable the myosin heads to disconnect from the myosin filament in order to reach the actin filament for binding (H.E. Huxley, 1969). S2 regions were described to be loosely attached to the MYH rod (McLachlan and Karn, 1982). Studies by Sanford Bernstein and colleagues focused on the region between S2 and LMM in Drosophila melanogaster. By exchanging the S2 fragment of a fast muscle to one from a slow muscle myosin of the flies, they observed profound changes in the myofibrilar organisation and alterations in myosin kinetics, such as increased actin affinity or myosin attachment time (Suggs, 2007).
By the interaction of Lrrc39 to this important region of the MHC muscle contraction might be influenced. Regarding this, the additionally identified interaction partner *2900073G15Rik* (Figure 19) may be of further interest. The *2900073G15Rik* gene has two homologs in human as well in zebrafish namely myosin light chain 12a and 9, regulatory (MLC12 and MLC9). Since the regulatory light chains modulate the movement of the myosin heads, an interaction by Lrrc39 seems likely to influence the process and thereby a regulation of the myosin-actin interaction might be possible. Indirect evidence form Margossian *et al.* supports this notion: In their experiments they inhibited the hinge region S1/S2 by specific antibodies and observed decreased muscle contraction (Margossian, 1991).

In sum, the binding to MHC at the S2 region is prone to influence the mobility of the lever arm and may cause the myosin head region to stick out in order to reach the actin filament for contraction. To understand how Lrrc39 regulates this movement further studies are necessary. However, closer inspections of the MHC binding region of Lrrc39 revealed conserved phosphorylation sites. This phosphorylation pattern within the binding region leads to the question if there could be an active role of Lrrc39?

6.4 Modification of myosin heavy chain phosphorylation

The documented phosphorylation sites in the Lrrc39 binding region were conserved among all MHC binding partners (Figure 22). To investigate whether Lrrc39 interaction alters phosphorylation, radiolabelling by γ-ATP supplementation in co-transfected cells was utilised. Transfection of the MHC binding region (myh1-GFP) with or without Lrrc39, followed by myh1-GFP precipitation and analysis of the ^{32}P incorporation was approached. The co-expression of Lrrc39 resulted in higher signal at the expected band size of the constructed myh1 binding region (~100 kDa) and thereby confirmed that binding of Lrrc39 leads to higher phosphorylation. Also additional bands were detected with increased ^{32}P-incorporation after Lrrc39 binding. Clearly, this experiment needs further confirmation and so far, the additional proteins could not be identified. However this pilot test showed clearly that Lrrc39 binding increases phosphorylation patterns. Due to the unknown kinase that phosphorylates MHC in this region and the unclear function of MHC phosphorylation, the precise mechanism of Lrrc39 remains open. One potential candidate that may influence this process is the titin kinase domain, which is localised in close proximity to Lrrc39 in the M-band region. It was shown that this titin region is involved in early muscle formation and severe phenotypes could be observed after titin knockout or deficiency in mice, including destroyed sarcomere formation (Pernigo, 2010). Since the *in vitro* data of Lrrc39 overexpression (Figure 24E) resulted in a loss of the C-terminal region of titin, it might be worthwhile to extent this experiment in order to clarify whether the strong increase of Lrrc39 proteins, which may interfere with titin binding at this position, or an increase of MHC phosphorylation was the cause of this disruption in titin expression pattern. Further, it remains to be determined whether titin kinase is altered by Lrrc39 expression. This could be addressed by analysing the titin kinase activity in the established *in vitro* system. Although the interaction partner of Lrrc39 was established and differences in phosphorylation patterns were observed, the mechanism how Lrrc39 can influences sarcomere assembly and thereby muscle function remains unclear.

To summarise the presented results of this study and combining the subcellular localisation of Lrrc39 with the observed changes in sarcomere assembly and the binding to MHC, I propose the following model to elucidate a potential role of Lrrc39 for sarcomere organisation. Figure 25 illustrates the localisation of Lrrc39 (green) in the proximal M-band region of the thick filament (grey), bound to the S2 region of MHC. The S2region was described as lever arm to connect the myosin head to the actin filament (red) in order to promote muscle contraction. From the proposed localisation of Lrrc39 in the MHC P-zone it can be postulated that the movement of the last myosins influence the N-terminus of the actin filament. The striking effect after Lrrc39 inhibition on actin-

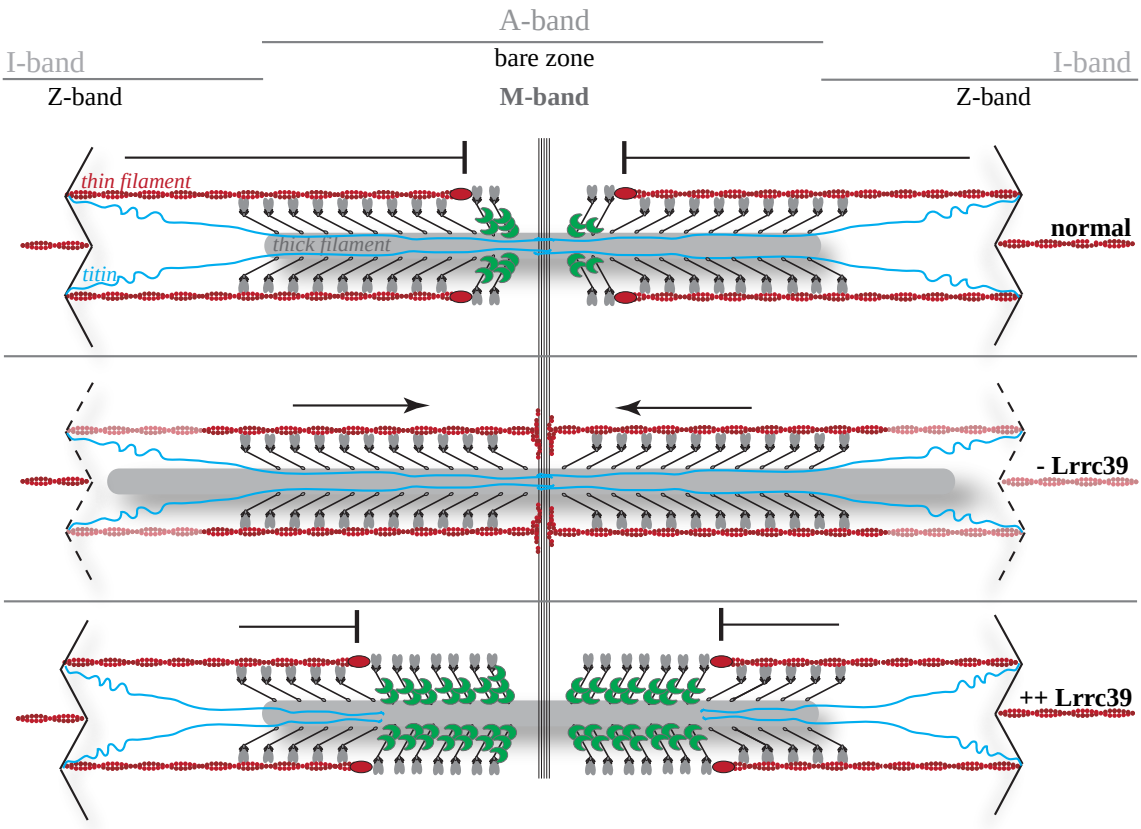


Figure 25: Model of hypothetical function of Lrrc39. The schematic drawing shows the localisation of Lrrc39 ( ) in the P-zone of the sarcomere positioned on the S2 region of MHC. Through interference in this region the lever arm is moved towards the actin filament (red). The thereby outwards sticking myosin signals the actin filament that the end is reached. The absence of Lrrc39 causes reduced ability of the lever arm to stick out from the myosin filament (grey). Myosin in the P-zone fails in signalling the actin filament to stop elongating and the both filament systems slide too far into each other. On the contrary, overexpression of Lrrc39 may result in enhanced mobility of the lever arms in a broader region, which influences filament elongation and may lead to reduced contractility. Further, the overexpression may impede the binding of the c-terminal end of titin (blue). Both changes might disrupt sarcomere organisation and cause secondary changes.

regulate actin elongation, supports this notion. Thereby, myosins in the P-zone may block further actin elongation and regulate the alignment of the filaments during sarcomere assembly. The loss of Lrrc39 would lead to reduced mobility of the central located myosins, followed by uncontrolled elongation of the actin filament and consequently to an imbalance of the filament systems. Thereby the extension in which myosin filaments slide into actin filaments during contraction is altered. On the contrary, overexpression of Lrrc39 may cause increased movement of the myosin lever arm, which may inhibit the actin filament elongation and consequently influence the sarcomere organisation. Additionally, the increase of Lrrc39 proteins may impede c-terminal titin binding to the M-band.

Since the large protein titin is regarded as a molecular spring that controls the sarcomere extension and prevents actin- and myosin filaments from sliding too far away from each other. Consequently an opposite mechanism must exist that controls the filaments to not slide too far into each other. The proposed model suggests the MHC at the P-zone of the myosin filament for this unique function. From the proposed role on filament assembly, it can be postulated that a balanced expression of Lrrc39 is required for normal sarcomere assembly.

6.5 Final conclusion and outlook

The data presented here expand our understanding of the sarcomere architecture. The analysed expression patterns and the investigated subcellular localisation of Lrrc39 result in a conserved position at the proximal M-region of the striated musculature. Investigations of phenotypic changes after loss-of-function experiments in zebrafish provided evidence for an impact on muscle development. Further, regulation of actin binding proteins and heat shock proteins provided evidence for a distinct role in sarcomere assembly.

One of the most important results of this study is the identification of the interaction domain within MHC, which interacts with Lrrc39. The identified S2 region seems crucial for the binding between myosin- and actin filament to allow spatially controlled muscle contraction. My results localise Lrrc39 to the P-zone of the thick filament. Together with the described effects on sarcomere assembly, these data argue for a crosstalk between structural assembly and contraction. Previous reports have shown that the integration of sarcomeric proteins during assembly correlate with the contraction of the cell (Simpson, 1993; Fujita, 2007). Therefore it was postulated that contraction is required for a uniform and proper arrangement of the sarcomere (De Deyne, 2000). Together with these data, it can be postulated that Lrrc39 influence sarcomere assembly by modulation of muscle contraction. Further work to understand the mechanisms of Lrrc39-MHC interaction is of paramount importance, in particular in terms of MHC phosphorylation in the S2 region. Further work will focus on the effect of Lrrc39 inhibition on force production. The studies of Dou and Arner on zebrafish musculature offer a promising tool for further investigations (Dou, 2008) in this respect.

It is important to note, that sarcomere assembly is not only restricted to the early muscle development but essential for muscle maintenance, physiological muscle growth and under pathological conditions like hypertrophy or remodelling processes after atrophy. Given the importance of sarcomere assembly for muscle function, I showed that Lrrc39 as integral part of the sarcomere with its regulatory properties is worth studying in detail. Since the expression analysis of Lrrc39 is not only expressed in early embryonic stage, but also in adult cardiac and skeletal muscle tissues, investigations of later developmental stages seem promising. To understand the basic function of Lrrc39 and to analyse the role in the adult tissue a mouse knock-out project has been initiated. With this model it would be possible to investigate the impact on sarcomere assembly during muscle maintenance and clarify *in vivo* whether Lrrc39 regulates muscle function in the adult tissue.

Apart from this, not all proteins involved in the process of sarcomere assembly have yet been characterised. It seems likely that the initial *in silico* screen, which yielded Lrrc39, still holds similarly important proteins. Additional functional studies, similar to the

approach presented in this thesis are necessary to explore this possibility, which might be important for muscle function. Work in this direction has already been started focusing on the *LMOD3* and *USP28* genes. I trust that the annotation of new genes is an important step in order to understand the mechanisms underlying force generation in the musculature.

7 REFERENCES

Agarkova, I, Ehler, E, Lange, S, Schoenauer, R and Perriard, JC. M-band: a safeguard for sarcomere stability? *J Muscle Res Cell Motil* **24**, 191-203, (2003).

Al-Khayat, HA, Morris, EP, Kensler, RW and Squire, JM. Myosin filament 3D structure in mammalian cardiac muscle. *J Struct Biol* **163**, 117-126, (2008).

Alamo, L, Wriggers, W, Pinto, A, Bartoli, F, Salazar, L, Zhao, FQ, . . . Padron, R. Three-dimensional reconstruction of tarantula myosin filaments suggests how phosphorylation may regulate myosin activity. *J Mol Biol* **384**, 780-797, (2008).

Allen, DL, Harrison, BC and Leinwand, LA. Inactivation of myosin heavy chain genes in the mouse: diverse and unexpected phenotypes. *Microsc Res Tech* **50**, 492-499, (2000).

Arnaout, R, Ferrer, T, Huisken, J, Spitzer, K, Stainier, DY, Tristani-Firouzi, M and Chi, NC. Zebrafish model for human long QT syndrome. Proc Natl Acad Sci U S A 104, 11316-11321, (2007).

Arya, R, Kedar, V, Hwang, JR, McDonough, H, Li, HH, Taylor, J and Patterson, C. Muscle ring finger protein-1 inhibits PKC{epsilon} activation and prevents cardiomyocyte hypertrophy. *J Cell Biol* **167**, 1147-1159, (2004).

Baker, JR. The Structure and Chemical Composition of the Golgi Element. *J. Micr. Sci* **85**, 1-71, (1944).

Barral, JM, Hutagalung, AH, Brinker, A, Hartl, FU and Epstein, HF. Role of the myosin assembly protein UNC-45 as a molecular chaperone for myosin. *Science* **295**, 669-671, (2002).

Bartoli, M and Richard, I. Calpains in muscle wasting. *Int J Biochem Cell Biol* **37**, 2115-2133, (2005).

Beis, I and Newsholme, EA. The contents of adenine nucleotides, phosphagens and some glycolytic intermediates in resting muscles from vertebrates and invertebrates. *Biochem J* **152**, 23-32, (1975).

Bella, J, Hindle, KL, McEwan, PA and Lovell, SC. The leucine-rich repeat structure. *Cell Mol Life Sci* **65**, 2307-2333, (2008).

Berchtold, MW, Brinkmeier, H and Muntener, M. Calcium ion in skeletal muscle: its crucial role for muscle function, plasticity, and disease. *Physiol Rev* **80**, 1215-1265, (2000).

Bernick, EP, Zhang, PJ and Du, S. Knockdown and overexpression of Unc-45b result in defective myofibril organization in skeletal muscles of zebrafish embryos. *BMC Cell Biol* **11**, 70, (2010).

Boateng, SY and Goldspink, PH. Assembly and maintenance of the sarcomere night and day. *Cardiovasc Res* **77**, 667-675, (2008).

Bottinelli, R. Functional heterogeneity of mammalian single muscle fibres: do myosin isoforms tell the whole story? *Pflugers Arch* **443**, 6-17, (2001).

Brenner, B. in *Physiologie* Vol. 6., vollst. überarb. Aufl. (ed Klinke) Ch. 4, 98-127 (Stuttgart; New York, NY:Thieme, 2010).

Broschat, KO. Tropomyosin prevents depolymerization of actin filaments from the pointed end. *J Biol Chem* **265**, 21323-21329, (1990).

Carmignac, V and Durbeej, M. Cell-matrix interactions in muscle disease. *J Pathol* **226**, 200-218, (2012).

Carroll, S, Lu, S, Herrera, AH and Horowits, R. N-RAP scaffolds I-Z-I assembly during myofibrillogenesis in cultured chick cardiomyocytes. *J Cell Sci* **117**, 105-114, (2004).

Chan, PK, Lin, CC and Cheng, SH. Noninvasive technique for measurement of heartbeat regularity in zebrafish (Danio rerio) embryos. BMC Biotechnol 9, 11, (2009).

Chereau, D, Boczkowska, M, Skwarek-Maruszewska, A, Fujiwara, I, Hayes, DB, Rebowski, G, . . . Dominguez, R. Leiomodin is an actin filament nucleator in muscle cells. *Science* **320**, 239-243, (2008).

Chesarone, M, Gould, CJ, Moseley, JB and Goode, BL. Displacement of formins from growing barbed ends by bud14 is critical for actin cable architecture and function. *Dev Cell* **16**, 292-302, (2009).

Chi, NC, Shaw, RM, De Val, S, Kang, G, Jan, LY, Black, BL and Stainier, DY. Foxn4 directly regulates tbx2b expression and atrioventricular canal formation. *Genes Dev* **22**, 734-739, (2008).

Clark, K, Langeslag, M, Figdor, CG and van Leeuwen, FN. Myosin II and mechanotransduction: a balancing act. *Trends Cell Biol* **17**, 178-186, (2007).

Coffey, VG and Hawley, JA. The molecular bases of training adaptation. *Sports Med* **37**, 737-763, (2007).

Cooper, JA. Effects of cytochalasin and phalloidin on actin. *J Cell Biol* **105**, 1473-1478, (1987).

Cooper, JA and Schafer, DA. Control of actin assembly and disassembly at filament ends. *Curr Opin Cell Biol* **12**, 97-103, (2000).

Cox, J, Matic, I, Hilger, M, Nagaraj, N, Selbach, M, Olsen, JV and Mann, M. A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat Protoc* **4**, 698-705, (2009).

Craig, R and Offer, G. The location of C-protein in rabbit skeletal muscle. *Proc R Soc Lond B Biol Sci* **192**, 451-461, (1976).

D'Amico, L, Scott, IC, Jungblut, B and Stainier, DY. A mutation in zebrafish hmgcr1b reveals a role for isoprenoids in vertebrate heart-tube formation. *Curr Biol* **17**, 252-259, (2007).

Dabiri, GA, Turnacioglu, KK, Sanger, JM and Sanger, JW. Myofibrillogenesis visualized in living embryonic cardiomyocytes. *Proc Natl Acad Sci U S A* **94**, 9493-9498, (1997).

Dahme, T, Katus, HA and Rottbauer, W. Fishing for the genetic basis of cardiovascular disease. *Dis Model Mech* **2**, 18-22, (2009).

De Deyne, PG. Formation of sarcomeres in developing myotubes: role of mechanical stretch and contractile activation. *Am J Physiol Cell Physiol* **279**, C1801-1811, (2000).

Dhume, A, Lu, S and Horowits, R. Targeted disruption of N-RAP gene function by RNA interference: a role for N-RAP in myofibril organization. *Cell Motil Cytoskeleton* **63**, 493-511, (2006).

Dou, Y, Andersson-Lendahl, M and Arner, A. Structure and function of skeletal muscle in zebrafish early larvae. *J Gen Physiol* **131**, 445-453, (2008).

Du, SJ, Li, H, Bian, Y and Zhong, Y. Heat-shock protein 90alpha1 is required for organized myofibril assembly in skeletal muscles of zebrafish embryos. *Proc Natl Acad Sci U S A* **105**, 554-559, (2008).

Dubowitz, V and Pearse, AG. A comparative histochemical study of oxidative enzyme and phosphorylase activity in skeletal muscle. *Z Zellforch Microsk Anat Histochem* **2**, 105-117, (1960).

Dulyaninova, NG, House, RP, Betapudi, V and Bresnick, AR. Myosin-IIA heavy-chain phosphorylation regulates the motility of MDA-MB-231 carcinoma cells. *Mol Biol Cell* **18**, 3144-3155, (2007).

Ehler, E and Gautel, M. The sarcomere and sarcomerogenesis. *Adv Exp Med Biol* **642**, 1-14, (2008).

Ehler, E, Horowits, R, Zuppinger, C, Price, RL, Perriard, E, Leu, M, . . . Perriard, JC. Alterations at the intercalated disk associated with the absence of muscle LIM protein. *J Cell Biol* **153**, 763-772, (2001).

Eisen, JS and Smith, JC. Controlling morpholino experiments: don't stop making antisense. *Development* **135**, 1735-1743, (2008).

Freiburg, A and Gautel, M. A molecular map of the interactions between titin and myosin-binding protein C. Implications for sarcomeric assembly in familial hypertrophic cardiomyopathy. *Eur J Biochem* **235**, 317-323, (1996).

Fujita, H, Nedachi, T and Kanzaki, M. Accelerated de novo sarcomere assembly by electric pulse stimulation in C2C12 myotubes. *Exp Cell Res* **313**, 1853-1865, (2007).

Fukuzawa, A, Lange, S, Holt, M, Vihola, A, Carmignac, V, Ferreiro, A, . . . Gautel, M. Interactions with titin and myomesin target obscurin and obscurin-like 1 to the M-band: implications for hereditary myopathies. *J Cell Sci* **121**, 1841-1851, (2008).

Gautel, M, Castiglione Morelli, MA, Pfuhl, M, Motta, A and Pastore, A. A calmodulin-binding sequence in the C-terminus of human cardiac titin kinase. *Eur J Biochem* **230**, 752-759, (1995).

Gellert, P, Jenniches, K, Braun, T and Uchida, S. C-It: a knowledge database for tissue-enriched genes. *Bioinformatics* **26**, 2328-2333, (2010).

Gellert, P, Teranishi, M, Jenniches, K, De Gaspari, P, John, D, Grosse Kreymborg, K, . . . Uchida, S. Gene Array Analyzer: alternative usage of gene arrays to study alternative splicing events. *Nucleic Acids Res*, (2011).

Gellert, P, Uchida, S and Braun, T. Exon Array Analyzer: a web interface for Affymetrix exon array analysis. *Bioinformatics* **25**, 3323-3324, (2009).

Gilbert, R, Cohen, JA, Pardo, S, Basu, A and Fischman, DA. Identification of the A-band localization domain of myosin binding proteins C and H (MyBP-C, MyBP-H) in skeletal muscle. *J Cell Sci* **112 (Pt 1)**, 69-79, (1999).

Gokhin, DS and Fowler, VM. Tropomodulin capping of actin filaments in striated muscle development and physiology. *J Biomed Biotechnol* **2011**, 103069, (2011).

Gonzalez-Sanchez, A and Bader, D. Immunochemical analysis of myosin heavy chains in the developing chicken heart. *Dev Biol* **103**, 151-158, (1984).

Gordon, AM, Homsher, E and Regnier, M. Regulation of contraction in striated muscle. *Physiol Rev* **80**, 853-924, (2000).

Grantham, J, Ruddock, LW, Roobol, A and Carden, MJ. Eukaryotic chaperonin containing T-complex polypeptide 1 interacts with filamentous actin and reduces the initial rate of actin polymerization in vitro. *Cell Stress Chaperones* **7**, 235-242, (2002).

Gregorio, CC, Granzier, H, Sorimachi, H and Labeit, S. Muscle assembly: a titanic achievement? *Curr Opin Cell Biol* **11**, 18-25, (1999).

Gregorio, CC, Weber, A, Bondad, M, Pennise, CR and Fowler, VM. Requirement of pointed-end capping by tropomodulin to maintain actin filament length in embryonic chick cardiac myocytes. *Nature* **377**, 83-86, (1995).

Grove, BK, Kurer, V, Lehner, C, Doetschman, TC, Perriard, JC and Eppenberger, HM. A new 185,000-dalton skeletal muscle protein detected by monoclonal antibodies. *J Cell Biol* **98**, 518-524, (1984).

Halpern, ME, Ho, RK, Walker, C and Kimmel, CB. Induction of muscle pioneers and floor plate is distinguished by the zebrafish no tail mutation. *Cell* **75**, 99-111, (1993).

Hanson, J and Huxley, HE. Structural basis of the cross-striations in muscle. *Nature* **172**, 530-532, (1953).

Haravuori, H, Vihola, A, Straub, V, Auranen, M, Richard, I, Marchand, S, . . . Udd, B. Secondary calpain3 deficiency in 2q-linked muscular dystrophy: titin is the candidate gene. *Neurology* **56**, 869-877, (2001).

Harris, SP, Bartley, CR, Hacker, TA, McDonald, KS, Douglas, PS, Greaser, ML, . . . Moss, RL. Hypertrophic cardiomyopathy in cardiac myosin binding protein-C knockout mice. *Circ Res* **90**, 594-601, (2002).

He, ZH, Bottinelli, R, Pellegrino, MA, Ferenczi, MA and Reggiani, C. ATP consumption and efficiency of human single muscle fibers with different myosin isoform composition. *Biophys J* **79**, 945-961, (2000).

Heasman, J. Morpholino oligos: making sense of antisense? *Dev Biol* **243**, 209-214, (2002).

Herzberg, O and James, MN. Structure of the calcium regulatory muscle protein troponin-C at 2.8 A resolution. *Nature* **313**, 653-659, (1985).

Higashijima, S, Okamoto, H, Ueno, N, Hotta, Y and Eguchi, G. High-frequency generation of transgenic zebrafish which reliably express GFP in whole muscles or the whole body by using promoters of zebrafish origin. *Dev Biol* **192**, 289-299, (1997).

Hocking, AM, Shinomura, T and McQuillan, DJ. Leucine-rich repeat glycoproteins of the extracellular matrix. *Matrix Biol* **17**, 1-19, (1998).

Holzinger, A. Jasplakinolide: an actin-specific reagent that promotes actin polymerization. *Methods Mol Biol* **586**, 71-87, (2009).

Hornemann, T, Kempa, S, Himmel, M, Hayess, K, Furst, DO and Wallimann, T. Muscle-type creatine kinase interacts with central domains of the M-band proteins myomesin and M-protein. *J Mol Biol* **332**, 877-887, (2003).

Huxley, AF and Niedergerke, R. Structural changes in muscle during contraction; interference microscopy of living muscle fibres. *Nature* **173**, 971-973, (1954).

Huxley, H and Hanson, J. Changes in the cross-striations of muscle during contraction and stretch and their structural interpretation. *Nature* **173**, 973-976, (1954).

Huxley, HE. The contraction of muscle. *Sci Am* **199**, 67-72 passim, (1958).

Huxley, HE. The mechanism of muscular contraction. *Science* **164**, 1356-1365, (1969).

Jones, WK, Grupp, IL, Doetschman, T, Grupp, G, Osinska, H, Hewett, TE, . . . Robbins, J. Ablation of the murine alpha myosin heavy chain gene leads to dosage effects and functional deficits in the heart. *J Clin Invest* **98**, 1906-1917, (1996).

Kamm, KE and Stull, JT. Dedicated myosin light chain kinases with diverse cellular functions. *J Biol Chem* **276**, 4527-4530, (2001).

Kim, J, Lowe, T and Hoppe, T. Protein quality control gets muscle into shape. *Trends Cell Biol* **18**, 264-272, (2008).

Kinbara, K, Sorimachi, H, Ishiura, S and Suzuki, K. Muscle-specific calpain, p94, interacts with the extreme C-terminal region of connectin, a unique region flanked by two immunoglobulin C2 motifs. *Arch Biochem Biophys* **342**, 99-107, (1997).

Knappeis, GG and Carlsen, F. The ultrastructure of the M line in skeletal muscle. *J Cell Biol* **38**, 202-211, (1968).

Korn, ED. Coevolution of head, neck, and tail domains of myosin heavy chains. *Proc Natl Acad Sci U S A* **97**, 12559-12564, (2000).

Kramerova, I, Kudryashova, E, Tidball, JG and Spencer, MJ. Null mutation of calpain 3 (p94) in mice causes abnormal sarcomere formation in vivo and in vitro. *Hum Mol Genet* **13**, 1373-1388, (2004).

Krendel, M and Mooseker, MS. Myosins: tails (and heads) of functional diversity. *Physiology (Bethesda)* **20**, 239-251, (2005).

Labeit, S, Gautel, M, Lakey, A and Trinick, J. Towards a molecular understanding of titin. *EMBO J* **11**, 1711-1716, (1992).

Labeit, S, Kolmerer, B and Linke, WA. The giant protein titin. Emerging roles in physiology and pathophysiology. *Circ Res* **80**, 290-294, (1997).

Laemmli, UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685, (1970).

Landsverk, ML, Li, S, Hutagalung, AH, Najafov, A, Hoppe, T, Barral, JM and Epstein, HF. The UNC-45 chaperone mediates sarcomere assembly through myosin degradation in Caenorhabditis elegans. *J Cell Biol* **177**, 205-210, (2007).

Lange, S, Auerbach, D, McLoughlin, P, Perriard, E, Schafer, BW, Perriard, JC and Ehler, E. Subcellular targeting of metabolic enzymes to titin in heart muscle may be mediated by DRAL/FHL-2. *J Cell Sci* **115**, 4925-4936, (2002).

Lange, S, Agarkova, I, Perriard, JC and Ehler, E. The sarcomeric M-band during development and in disease. *J Muscle Res Cell Motil* **26**, 375-379, (2005a).

Lange, S, Himmel, M, Auerbach, D, Agarkova, I, Hayess, K, Furst, DO, . . . Ehler, E. Dimerisation of myomesin: implications for the structure of the sarcomeric M-band. *J Mol Biol* **345**, 289-298, (2005b).

Lange, S, Xiang, F, Yakovenko, A, Vihola, A, Hackman, P, Rostkova, E, . . . Gautel, M. The kinase domain of titin controls muscle gene expression and protein turnover. *Science* **308**, 1599-1603, (2005c).

Linke, WA. Stretching molecular springs: elasticity of titin filaments in vertebrate striated muscle. *Histol Histopathol* **15**, 799-811, (2000).

Linke, WA, Ivemeyer, M, Olivieri, N, Kolmerer, B, Ruegg, JC and Labeit, S. Towards a molecular understanding of the elasticity of titin. *J Mol Biol* **261**, 62-71, (1996).

Liu, Y and Steinacker, JM. Changes in skeletal muscle heat shock proteins: pathological significance. *Front Biosci* **6**, D12-25, (2001).

Luther, PK, Bennett, PM, Knupp, C, Craig, R, Padron, R, Harris, SP, . . . Moss, RL. Understanding the organisation and role of myosin binding protein C in normal striated muscle by comparison with MyBP-C knockout cardiac muscle. *J Mol Biol* **384**, 60-72, (2008).

Lu, RC and Wong, A. The amino acid sequence and stability predictions of the hinge region in myosin subfragment 2. *J Biol Chem* **260**, 3456-3461, (1985)

Luther, PK, Munro, PM and Squire, JM. Three-dimensional structure of the vertebrate muscle A-band. III. M-region structure and myosin filament symmetry. *J Mol Biol* **151**, 703-730, (1981).

Lymn, RW and Taylor, EW. Mechanism of adenosine triphosphate hydrolysis by actomyosin. *Biochemistry* **10**, 4617-4624, (1971).

Manisastry, SM, Zaal, KJ and Horowits, R. Myofibril assembly visualized by imaging N-RAP, alpha-actinin, and actin in living cardiomyocytes. *Exp Cell Res* **315**, 2126-2139, (2009).

Margossian, SS, Krueger, JW, Sellers, JR, Cuda, G, Caulfield, JB, Norton, P and Slayter, HS. Influence of the cardiac myosin hinge region on contractile activity. *Proc Natl Acad Sci U S A* **88**, 4941-4945, (1991).

Maruyama, K, Matsubara, S, Natori, R, Nonomura, Y and Kimura, S. Connectin, an elastic protein of muscle. Characterization and Function. *J Biochem* **82**, 317-337, (1977).

Masaki, T and Takaiti, O. M-protein. *J Biochem* **75**, 367-380, (1974).

Mayans, O, van der Ven, PF, Wilm, M, Mues, A, Young, P, Furst, DO, . . . Gautel, M. Structural basis for activation of the titin kinase domain during myofibrillogenesis. *Nature* **395**, 863-869, (1998).

McConnell, BK, Jones, KA, Fatkin, D, Arroyo, LH, Lee, RT, Aristizabal, O, . . . Seidman, JG. Dilated cardiomyopathy in homozygous myosin-binding protein-C mutant mice. *J Clin Invest* **104**, 1235-1244, (1999).

McKenna, NM, Johnson, CS and Wang, YL. Formation and alignment of Z lines in living chick myotubes microinjected with rhodamine-labeled alpha-actinin. *J Cell Biol* **103**, 2163-2171, (1986).

McLachlan, AD and Karn, J. Periodic charge distributions in the myosin rod amino acid sequence match cross-bridge spacings in muscle. *Nature* **299**, 226-231, (1982).

Monson, CA and Sadler, KC. Inbreeding depression and outbreeding depression are evident in wild-type zebrafish lines. *Zebrafish* **7**, 189-197, (2010).

Myhre, JL and Pilgrim, DB. At the Start of the Sarcomere: A Previously Unrecognized Role for Myosin Chaperones and Associated Proteins during Early Myofibrillogenesis. *Biochem Res Int* **2012**, 712315, (2012).

Nasevicius, A and Ekker, SC. Effective targeted gene 'knockdown' in zebrafish. *Nat Genet* **26**, 216-220, (2000).

Ng, AC, Eisenberg, JM, Heath, RJ, Huett, A, Robinson, CM, Nau, GJ and Xavier, RJ. Human leucine-rich repeat proteins: a genome-wide bioinformatic categorization and functional analysis in innate immunity. *Proc Natl Acad Sci U S A* **108 Suppl 1**, 4631-4638, (2011).

Ono, S. Mechanism of depolymerization and severing of actin filaments and its significance in cytoskeletal dynamics. *Int Rev Cytol* **258**, 1-82, (2007).

Ono, Y, Sorimachi, H and Suzuki, K. New aspect of the research on limb-girdle muscular dystrophy 2A: a molecular biologic and biochemical approach to pathology. *Trends Cardiovasc Med* **9**, 114-118, (1999).

Page, SG and Huxley, HE. Filament Lengths in Striated Muscle. *J Cell Biol* **19**, 369-390, (1963).

Pappas, CT, Bhattacharya, N, Cooper, JA and Gregorio, CC. Nebulin interacts with CapZ and regulates thin filament architecture within the Z-disc. *Mol Biol Cell* **19**, 1837-1847, (2008).

Pardo, JV, Siliciano, JD and 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* **80**, 1008-1012, (1983).

Pask, HT, Jones, KL, Luther, PK and Squire, JM. M-band structure, M-bridge interactions and contraction speed in vertebrate cardiac muscles. *J Muscle Res Cell Motil* **15**, 633-645, (1994).

Pawlowski, K. Uncharacterized/hypothetical proteins in biomedical 'omics' experiments: is novelty being swept under the carpet? *Brief Funct Genomic Proteomic* **7**, 283-290, (2008).

Perera, S, Holt, MR, Mankoo, BS and Gautel, M. Developmental regulation of MURF ubiquitin ligases and autophagy proteins nbr1, p62/SQSTM1 and LC3 during cardiac myofibril assembly and turnover. *Dev Biol* **351**, 46-61, (2011).

Pernigo, S, Fukuzawa, A, Bertz, M, Holt, M, Rief, M, Steiner, RA and Gautel, M. Structural insight into M-band assembly and mechanics from the titin-obscurin-like-1 complex. *Proc Natl Acad Sci U S A* **107**, 2908-2913, (2010).

Person, V, Kostin, S, Suzuki, K, Labeit, S and Schaper, J. Antisense oligonucleotide experiments elucidate the essential role of titin in sarcomerogenesis in adult rat cardiomyocytes in long-term culture. *J Cell Sci* **113 Pt 21**, 3851-3859, (2000).

Rhee, D, Sanger, JM and Sanger, JW. The premyofibril: evidence for its role in myofibrillogenesis. *Cell Motil Cytoskeleton* **28**, 1-24, (1994).

Richard, I, Broux, O, Allamand, V, Fougerousse, F, Chiannilkulchai, N, Bourg, N, . . . et al. Mutations in the proteolytic enzyme calpain 3 cause limb-girdle muscular dystrophy type 2A. *Cell* **81**, 27-40, (1995).

Rieger, S and Koster, RW. Preparation of zebrafish embryos for transmission electron microscopy. *CSH Protoc* **2007**, pdb prot4772, (2007).

Roperto, S, De Tullio, R, Raso, C, Stifanese, R, Russo, V, Gaspari, M, . . . Roperto, F. Calpain3 is expressed in a proteolitically active form in papillomavirus-associated urothelial tumors of the urinary bladder in cattle. *PLoS One* **5**, e10299, (2010).

Sabry, JH, Moores, SL, Ryan, S, Zang, JH and Spudich, JA. Myosin heavy chain phosphorylation sites regulate myosin localization during cytokinesis in live cells. *Mol Biol Cell* **8**, 2605-2615, (1997).

Sanger, JM and Sanger, JW. The dynamic Z bands of striated muscle cells. *Sci Signal* **1**, pe37, (2008).

Sarparanta, J, Blandin, G, Charton, K, Vihola, A, Marchand, S, Milic, A, . . . Udd, B. Interactions with M-band titin and calpain 3 link myospryn (CMYA5) to tibial and limb-girdle muscular dystrophies. *J Biol Chem* **285**, 30304-30315, (2010).

Schafer, DA, Hug, C and Cooper, JA. Inhibition of CapZ during myofibrillogenesis alters assembly of actin filaments. *J Cell Biol* **128**, 61-70, (1995).

Schmidt, A and Hall, MN. Signaling to the actin cytoskeleton. *Annu Rev Cell Dev Biol* **14**, 305-338, (1998).

Seale, P, Bjork, B, Yang, W, Kajimura, S, Chin, S, Kuang, S, . . . Spiegelman, BM. PRDM16 controls a brown fat/skeletal muscle switch. *Nature* **454**, 961-967, (2008).

Serrano-Pinto, V, Landais, I, Ogliastro, MH, Gutierrez-Ayala, M, Mejia-Ruiz, H, Villarreal-Colmenares, H, . . . Vazquez-Boucard, C. Vitellogenin mRNA expression in Cherax quadricarinatus during secondary vitellogenic at first maturation females. *Mol Reprod Dev* **69**, 17-21, (2004).

Shih, WM, Gryczynski, Z, Lakowicz, JR and Spudich, JA. A FRET-based sensor reveals large ATP hydrolysis-induced conformational changes and three distinct states of the molecular motor myosin. *Cell* **102**, 683-694, (2000).

Siegers, K, Waldmann, T, Leroux, MR, Grein, K, Shevchenko, A, Schiebel, E and Hartl, FU. Compartmentation of protein folding in vivo: sequestration of non-native polypeptide by the chaperonin-GimC system. *EMBO J* **18**, 75-84, (1999).

Simpson, DG, Decker, ML, Clark, WA and Decker, RS. Contractile activity and cell-cell contact regulate myofibrillar organization in cultured cardiac myocytes. *J Cell Biol* **123**, 323-336, (1993).

Sjostrom, M and Squire, JM. Cryo-ultramicrotomy and myofibrillar fine structure: a review. *J Microsc* **111**, 239-278, (1977a).

Sjostrom, M and Squire, JM. Fine structure of the A-band in cryo-sections. The structure of the A-band of human skeletal muscle fibres from ultra-thin cryo-sections negatively stained. *J Mol Biol* **109**, 49-68, (1977b).

Smolich, JJ. Ultrastructural and functional features of the developing mammalian heart: a brief overview. *Reprod Fertil Dev* **7**, 451-461, (1995).

Sparrow, JC and Schock, F. The initial steps of myofibril assembly: integrins pave the way. *Nat Rev Mol Cell Biol* **10**, 293-298, (2009).

Srikakulam, R and Winkelmann, DA. Chaperone-mediated folding and assembly of myosin in striated muscle. *J Cell Sci* **117**, 641-652, (2004).

Stainier, DY, Lee, RK and Fishman, MC. Cardiovascular development in the zebrafish. I. Myocardial fate map and heart tube formation. *Development* **119**, 31-40, (1993).

Steiner, F, Weber, K and Furst, DO. M band proteins myomesin and skelemin are encoded by the same gene: analysis of its organization and expression. *Genomics* **56**, 78-89, (1999).

Stirchak, EP, Summerton, JE and Weller, DD. Uncharged stereoregular nucleic acid analogs: 2. Morpholino nucleoside oligomers with carbamate internucleoside linkages. *Nucleic Acids Res* **17**, 6129-6141, (1989).

Suggs, JA, Cammarato, A, Kronert, WA, Nikkhoy, M, Dambacher, CM, Megighian, A and Bernstein, SI. Alternative S2 hinge regions of the myosin rod differentially affect muscle function, myofibril dimensions and myosin tail length. *J Mol Biol* **367**, 1312-1329, (2007).

Summerton, J and Weller, D. Morpholino antisense oligomers: design, preparation, and properties. *Antisense Nucleic Acid Drug Dev* **7**, 187-195, (1997).

Summerton, JE. Morpholino, siRNA, and S-DNA compared: impact of structure and mechanism of action on off-target effects and sequence specificity. *Curr Top Med Chem* **7**, 651-660, (2007).

Syska, H, Wilkinson, JM, Grand, RJ and Perry, SV. The relationship between biological activity and primary structure of troponin I from white skeletal muscle of the rabbit. *Biochem J* **153**, 375-387, (1976).

Taveau, M, Bourg, N, Sillon, G, Roudaut, C, Bartoli, M and Richard, I. Calpain 3 is activated through autolysis within the active site and lyses sarcomeric and sarcolemmal components. *Mol Cell Biol* **23**, 9127-9135, (2003).

Thisse, B, Heyer, V, Lux, A, Alunni, V, Degrave, A, Seiliez, I, . . . Thisse, C. Spatial and temporal expression of the zebrafish genome by large-scale in situ hybridization screening. *Methods Cell Biol* **77**, 505-519, (2004).

Thornell, LE, Carlsson, E, Kugelberg, E and Grove, BK. Myofibrillar M-band structure and composition of physiologically defined rat motor units. *Am J Physiol* **253**, C456-468, (1987).

Thornell, LE and Eriksson, A. Filament systems in the Purkinje fibers of the heart. *Am J Physiol* **241**, H291-305, (1981).

Thorrez, L, Van Deun, K, Tranchevent, LC, Van Lommel, L, Engelen, K, Marchal, K, . . . Schuit, F. Using ribosomal protein genes as reference: a tale of caution. *PLoS One* **3**, e1854, (2008).

Uchida, S, Schneider, A, Wiesnet, M, Jungblut, B, Zarjitskaya, P, Jenniches, K, . . . Braun, T. An integrated approach for the systematic identification and characterization of heart-enriched genes with unknown functions. *BMC Genomics* **10**, 100, (2009).

van der Ven, PF, Bartsch, JW, Gautel, M, Jockusch, H and Furst, DO. A functional knock-out of titin results in defective myofibril assembly. *J Cell Sci* **113 (Pt 8)**, 1405-1414, (2000).

Van Der Ven, PF, Obermann, WM, Weber, K and Furst, DO. Myomesin, M-protein and the structure of the sarcomeric M-band. *Adv Biophys* **33**, 91-99, (1996).

Van Troys, M, Huyck, L, Leyman, S, Dhaese, S, Vandekerkhove, J and Ampe, C. Ins and outs of ADF/cofilin activity and regulation. *Eur J Cell Biol* **87**, 649-667, (2008).

Veigel, C, Coluccio, LM, Jontes, JD, Sparrow, JC, Milligan, RA and Molloy, JE. The motor protein myosin-I produces its working stroke in two steps. *Nature* **398**, 530-533, (1999).

Voelkel, T and Linke, WA. Conformation-regulated mechanosensory control via titin domains in cardiac muscle. *Pflugers Arch* **462**, 143-154, (2011).

Wallimann, T, Doetschman, TC and Eppenberger, HM. Novel staining pattern of skeletal muscle M-lines upon incubation with antibodies against MM-creatine kinase. *J Cell Biol* **96**, 1772-1779, (1983).

Wang, SM, Lo, MC, Shang, C, Kao, SC and Tseng, YZ. Role of M-line proteins in sarcomeric titin assembly during cardiac myofibrillogenesis. *J Cell Biochem* **71**, 82-95, (1998).

Warshaw, DM, Guilford, WH, Freyzon, Y, Krementsova, E, Palmiter, KA, Tyska, MJ, . . . Trybus, KM. The light chain binding domain of expressed smooth muscle heavy meromyosin acts as a mechanical lever. *J Biol Chem* **275**, 37167-37172, (2000).

Westerfield, M. *The zebrafish book : a guide for the laboratory use of zebrafish (Brachydanio rerio)*. (M. Westerfield, 1993).

Westerlund, M, Belin, AC, Anvret, A, Bickford, P, Olson, L and Galter, D. Developmental regulation of leucine-rich repeat kinase 1 and 2 expression in the brain and other rodent and human organs: Implications for Parkinson's disease. *Neuroscience* **152**, 429-436, (2008).

Westman-Brinkmalm, A, Abramsson, A, Pannee, J, Gang, C, Gustavsson, MK, von Otter, M, . . . Zetterberg, H. SILAC zebrafish for quantitative analysis of protein turnover and tissue regeneration. *J Proteomics* **75**, 425-434, (2011).

Wheeler, DL, Church, DM, Federhen, S, Lash, AE, Madden, TL, Pontius, JU, . . . Wagner, L. Database resources of the National Center for Biotechnology. *Nucleic Acids Res* **31**, 28-33, (2003).

Whiting, A, Wardale, J and Trinick, J. Does titin regulate the length of muscle thick filaments? *J Mol Biol* **205**, 263-268, (1989).

Wilkinson, D. in *In situ hybridization: A practical approach.* Vol. 1992 (ed DG) 75-84 (Oxford University Press, 1992).

Will, RD, Eden, M, Just, S, Hansen, A, Eder, A, Frank, D, . . . Frey, N. Myomasp/LRRC39, a heart- and muscle-specific protein, is a novel component of the sarcomeric M-band and is involved in stretch sensing. *Circ Res* **107**, 1253-1264, (2010).

Winder, SJ and Ayscough, KR. Actin-binding proteins. *J Cell Sci* **118**, 651-654, (2005).

Winegrad, S. Myosin-binding protein C (MyBP-C) in cardiac muscle and contractility. *Adv Exp Med Biol* **538**, 31-40; discussion 40-31, (2003).

Wohlgemuth, SL, Crawford, BD and Pilgrim, DB. The myosin co-chaperone UNC-45 is required for skeletal and cardiac muscle function in zebrafish. *Dev Biol* **303**, 483-492, (2007).

Wollert, T, Heinz, DW and Schubert, WD. Thermodynamically reengineering the listerial invasion complex InlA/E-cadherin. *Proc Natl Acad Sci U S A* **104**, 13960-13965, (2007).

Wren, JD. A global meta-analysis of microarray expression data to predict unknown gene functions and estimate the literature-data divide. *Bioinformatics* **25**, 1694-1701, (2009).

Yamashiro, S, Cox, EA, Baillie, DL, Hardin, JD and Ono, S. Sarcomeric actin organization is synergistically promoted by tropomodulin, ADF/cofilin, AIP1 and profilin in C. elegans. *J Cell Sci* **121**, 3867-3877, (2008).

Zhang, R, Yang, J, Zhu, J and Xu, X. Depletion of zebrafish Tcap leads to muscular dystrophy via disrupting sarcomere-membrane interaction, not sarcomere assembly. *Hum Mol Genet* **18**, 4130-4140, (2009).

Zoghbi, ME, Woodhead, JL, Moss, RL and Craig, R. Three-dimensional structure of vertebrate cardiac muscle myosin filaments. *Proc Natl Acad Sci U S A* **105**, 2386-2390, (2008).

8 APPENDIX

ACKNOWLEDGEMENTS

First and foremost I am grateful to my supervisor Thomas Braun who supported me by any means, and to Shizuka Uchida who was a great mentor during the last years. I am grateful to Katja Becker for her supervision and for reviewing this thesis.

I also would like to thank the people supported the presented work:
All my colleagues at the Max Planck Institute for Heart and Lung Research for the good working atmosphere and making me feel welcome in my 'second home' and for sharing my daily life. Special thanks to Piera De Gaspari, who shared the bench with me for the last years. Thank you for your support, your criticism, your entertainment and your chocolate. Many thanks to Mizue Teranishi and Pascal Gellert, for the help analysing zebrafish gene expression using microarray.

Mathias Gautel from the King's College London for his collaboration, very helpful discussions, his scientific expertise and the antibodies.
Benno Jungblut for introducing me to the zebrafish research, for his helpful discussions and good advice during the course of this work.
Anne Konzer for sharing the SILAC zebrafish for the MS analysis and for discussions.
Renate Möhren for the ultrathin zebrafish sections and for spending hours with me in front of the electron microscope.
Marion Wiesnet for the perfect isolation of adult mouse cardiomyocytes.
Jan Heering from the Institute of Biophysical Chemistry at the Goethe University Frankfurt for his input and help with structure modelling.
Steffi Goffart for help with the radiolabelling to analyse the phosphorylation pattern.
Rita Retzloff for the always reliable and everlasting care of the zebrafishes and for her great support, and thanks to the zebrafishes for their willingness to produce countless eggs.
Jan Medenbach, without whom I would not have started this thesis.
I also would like to thank Tina E. Trenczek and Reinhard Dammann for their willingness to examine my dissertation.

And last but not least my family and friends for believing in me and supporting me whenever things went wrong, for pushing me whenever it was needed and for the enormous patience during the last months.

Finally, I am grateful for the International Max Planck Research School for Heart and Lung Research (IMPRS) as well as the International Giessen Graduate Centre for Life Sciences (GGL) for supporting this project.

EIDESSTATTLICHE ERKLÄRUNG

„Ich erkläre: Ich habe die vorgelegte Dissertation selbständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind, eingehalten.“

Bad Nauheim, den 24.04.2012