INVESTIGATING TITIN SPLICING IN MOUSE MUSCLE

A THESIS PRESENTED TO THE DEPARTMENT OF BIOCHEMISTRY, CELL AND MOLECULAR BIOLOGY, UNIVERSITY OF GHANA

BY

JAMES OSEI OWUSU

(10284181)

THIS THESIS IS SUBMITTED TO THE UNIVERSITY OF GHANA, LEGON IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE AWARD OF THE MASTER OF PHILOSOPHY IN BIOCHEMISTRY DEGREE

JULY 2015

DECLARATION

I, James Osei Owusu (Department of Biochemistry, Cell and Molecular Biology, University of Ghana), hereby declare that this thesis is the outcome of my own research (except where stated) undertaken at the Gage's Laboratory (Northern Arizona University, USA) and the Caporaso Laboratory (Northern Arizona University, USA) under the supervision of Dr. Matthew Gage (University of Massachusetts, Lowell) and Prof. Naa Ayikailey Adamafio (University of Ghana, Legon) with assistance from Dr. Caporaso (Northern Arizona University, USA). To the best of my knowledge, this thesis contains neither materials which have been accepted for the award of any material previously published by another author, except where due reference is made in the text of the thesis.

……..

JAMES OSEI OWUSU

(CANDIDATE)

……..

DR. MATTHEW GAGE

(SUPERVISOR)

……..

PROF. NAA AYIKAILEY ADAMAFIO

(SUPERVISOR)

ABSTRACT

Titin is a giant protein that is known for its contribution to muscle elasticity and sarcomere assembly in vertebrate striated muscles. There is considerable interest in understanding the expression of the single gene (*TTN*) that encodes titin because mutations in the coding gene have been implicated in muscular dystrophy. Insights into the regulatory mechanisms involved in titin splicing would contribute to the development of effective treatments for titin associated muscular dystrophy. Therefore, there is a need to expand the current knowledge in titin isoform expression. For this reason, titin expression in skeletal muscles of *Mus musculus* was investigated. The *Mus musculus* is a good experimental model for studying the expression and functions of titin in the sarcomere due to the close similarities between the mouse titin orthologue and that of humans.

Titin splicing was investigated by sequencing the entire transcriptome of three skeletal muscles, namely *Elongus digitorum longus* (EDL), psoas and soleus. The experimental procedure involved total RNA extraction, mRNA purification and fragmentation, first and second strand cDNA synthesis, cDNA library preparation, cluster generation, sequencing and computational analysis.

Differential exon skipping was observed in the three skeletal tissues. The differential exon skipping led to the expression of different titin isoforms in the EDL, psoas and soleus. Out of 322 titin exons analyzed, 13 exons were skipped in the soleus titin, 14 exons were skipped in the EDL titin and 20 exons were skipped in the psoas titin. Therefore the soleus titin isoform was the largest isoform, followed by the EDL titin and then the psoas titin. The titin isoforms differed from each other in their PEVK domain

sequences. The differential expression of the PEVK domain in the three isoforms was speculated to be critical to the role of titin in the tissue in which it is expressed. The pattern of exon splicing in the titin isoforms suggested that there were splicing factors that were ubiquitous and targeted similar splice sites in the three isoforms. These speculated common splicing factors were implicated in the exclusion of the same exons in all three isoforms. On the other hand, the exclusion of specific exons in only one or two of the isoforms pointed to the existence of tissue-specific splicing factors which recognized different splice sites in the three tissues studied and might be responsible for the expression of the different titin isoforms.

DEDICATION

This work is dedicated to my family; Comfort Owusu Atinkah, Charles Owusu and Cynthia Owusu, for their love and support.

ACKNOWLEDGMENTS

To God Almighty, I express my deepest gratitude for His love and favors. Without Him, this would have not been possible.

Dr. Matthew Gage, thank you for taking me into your lab and exposing me to a new field of research. I am very grateful for your support, positivity and encouragement.

Professor Naa Ayikailey Adamafio, thank you very much. You have been very patient in your instructions and guidance to me. You are blessed with wells of knowledge that I can boldly say that I have learnt from.

Dr. Caporaso, thank you for your supervision in analyzing my sequence data and To John Chase, I am very grateful for your help in analyzing my sequence data.

I wish to greatly acknowledge my family for their support. It is safe to know that there is always a family to run to.

I also appreciate all friends who have influenced my life. God bless you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

TTN	Titin gene
RNA	Ribonucleic Acid
RNA-Seq	RNA-Sequencing
DNA	Deoxyribonucleic Acid
cDNA	Complementary DNA
SBS	Sequencing by Synthesis
EDL	*Elongus Digitorum Longus*
PEVK	Proline, Glutamine, Valine and Lysine
ds	Double stranded
PCR	Polymerase Chain Reaction
TAE	Tris base, Acetic acid and EDTA
DEPC	Diethylpyrocarbonate
dNTPs	Deoxynucleotides

CHAPTER ONE

INTRODUCTION

Titin is a muscle protein that plays a critical role in muscle elasticity and sarcomere assembly (Linke *et al*., 1996; Linke and Granzier, 1998; Linke *et al*., 1999). Mutations in the gene encoding titin have been implicated in muscular dystrophy, a debilitating disease that weakens the musculoskeletal system, thereby hindering locomotion; a disease for which there is no known cure (Brown and Mendell, 2005). As many as 127 disease-causing mutations have been identified in the titin gene (Chauveau *et al*., 2014). To date, information on titin expression is sparse. Therefore, information on titin splicing patterns would contribute greatly to a better understanding of the mechanisms governing differential titin expression in muscles. Although studies on rabbits have shown that different isoforms of the protein occur in different tissues, little information is available on titin splicing patterns in skeletal muscles of *Mus musculus*. The titin orthologue in mouse presents a structure that is similar to that of humans and therefore serves as a good experimental system to study the functions of titin as well as the role of each titin domain in the sarcomere. For the reasons outlined, this project focused on titin expression in *Mus musculus*.

Titin is the largest protein known and it is found in vertebrate striated muscles. This giant protein is approximately 3.0 - 3.7 MDa in size and spans half of an entire sarcomere from the Z-disc to the M-line (Bang *et al*., 2001; Freiburg *et al*., 2000). Titin overlaps with adjacent titin molecules at both the Z-disc and M-line to form a continuous titin filament along the myofibril (Gregorio *et al*., 1999). In each half-sarcomere, titin is associated with

thick and thin filaments in the A-band and I-band regions respectively (Funatsu *et al.*, 1993). Studies have shown that the A-band region of titin is inextensible whereas the I-band region is extensible (Granzier *et al.*, 2014). The I-band is made up of three spring elements; the PEVK (Proline, Glutamine, Valine and Lysine) region, the N2A/ N2B element and the tandem-Ig segments. The spring elements have been implicated in the elasticity of the I-band. The extensibility of the I-band segment of titin allows it to function as a spring, contributing to muscle passive tension under physiological conditions (Furst *et al.*, 1988; Maruyama, 1994). Passive tension, commonly known as passive elasticity of a muscle, is the extent to which the muscle resists deformation in response to applied force. In addition to contributing to muscle passive tension, the efficient functioning of striated muscles also greatly depend on correct positioning and coordination of titin with other muscle filaments (Gregorio *et al.*, 2005). For instance, in muscle development, titin serves as a template for sarcomere assembly (Gregorio *et al.*, 1999). The significance of titin in the development and functioning of muscles makes it an indispensable muscle protein.

Titin is encoded by a single gene (*TTN*) with multiple isoforms observed in the human cardiac muscle, the mouse cardiac muscle and rabbit skeletal muscles (Freiburg *et al.*, 2000; Labeit *et al.*, 2000). Alternative splicing enhances protein diversity by generating multiple protein encoding transcripts from a single gene. Proteins expressed from such transcripts could have distinct or opposing functional differences that can significantly affect the phenotype of an organism (Graveley, 2001). Alternative splicing in *TTN* has revealed distinct I-band structures in muscle types with unique titin-based elastic properties in rabbits (Freiburg *et al.*, 2000). However, the mechanisms by which diverse titin isoforms with differential elasticity are expressed from the titin gene remain unclear. Therefore, the

purpose of the study was to generate information on the specific types of titin isoforms expressed in different skeletal muscles of *Mus musculus* by identifying exons of the mouse mRNA transcript that codes for titin in three different skeletal muscles; EDL, Psoas and Soleus. By sequencing the mRNA transcript, exons that are included and those that are excluded during alternative splicing can be identified and this would give insight into the exon distribution that gives rise to titin in the above mentioned muscle tissues.

1.1 AIM

To investigate exon skipping patterns during titin gene expression in different skeletal muscle tissues of *Mus musculus.*

1.2 SPECIFIC OBJECTIVES

To:

- Sequence the transcript using Illumina sequencing technology and submit sequence to databank
- Analyze the transcriptome for Titin exons using computational tools to identify tissue specific isoforms

CHAPTER TWO

LITERATURE REVIEW

2.1 TITIN STRUCTURE

Titin, initially known as 'connectin', is the largest protein in mammals, ranging from 2970 to 3700 kDa, and the third most abundant protein after myosin and actin (LeWinter *et al.*, 2007; Maruyama *et al*., 1977). It is found in vertebrate striated muscles. It is estimated that an adult human of 62 kg body weight contains approximately 0.5 kg titin (Labeit *et al*., 1997). Titin is encoded by the *TTN* gene on chromosome 2 at position 2q31 as shown in Figure 2.1 (Bang *et al.*, 2001). This gene contains 363 exons in humans which is the largest number of exons found in any single gene (Bang *et al*., 2001; Udd *et al*., 2005). It spans half a complete sarcomere from the Z disc to the M-line or center of one complete sarcomere (Gregorio *et al*., 1999) (Figures 2.1b and 2.2).

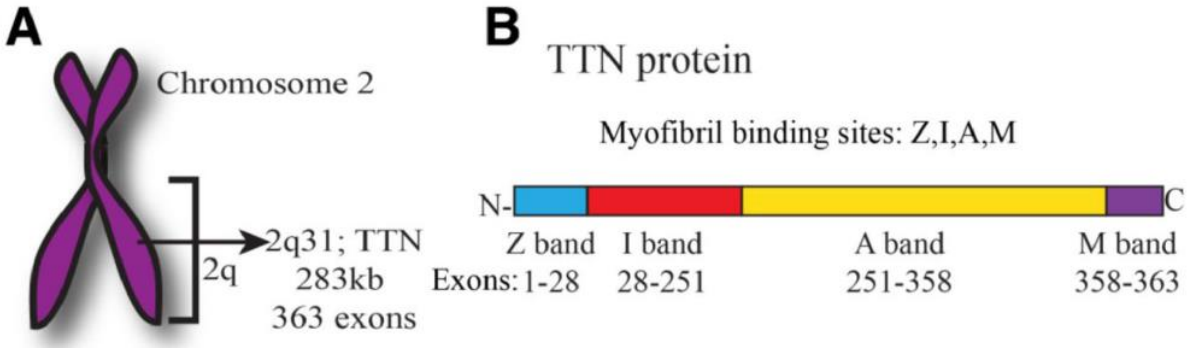


Figure 2.1: ***a) TTN at position 2q31 on the long arm of chromosome 2. b) N and C-terminal regions of titin protein, showing the various band segment of the protein, Z, I, A and M band segments and the number of exons involved in their expression (Punetha and Hoffman, 2013).***

The C and N-terminal ends of titin are located in the M-line and Z-disc respectively with the C-terminal end of one titin molecule overlapping with the adjacent C-terminal end of another titin molecule in the adjacent half of the sarcomere. Similarly, the N-terminus of one titin molecule overlaps with N-terminus of another titin molecule in the adjacent sarcomere (Obermann *et al*., 1996). This leads to the formation of a continuous titin filament along the myofibril. In each half-sarcomere, titin is associated with thick and thin filaments in the A-band and I-band respectively (Funatsu *et al*., 1993). The I-band region of titin consists of the proximal tandem Ig, N2A/ N2B elements, the PEVK region and distal tandem Ig segments as shown in Figure 2.2. The tandem Ig segments constitute serially linked immunoglobulin-like domains and the PEVK region is named so because of its high contents of Proline, Glutamine, Lysine and Valine. The A-band constitutes repeating fibronectin (FNIII) and Ig domains. The I-band segment has been shown to be extensible whereas the A-band is considered inextensible due to its tight association with the thick filament. The PEVK and the proximal tandem Ig have been shown to be responsible for the elasticity of the I-band (Gautel and Goulding, 1996; Linke *et al*., 1998).

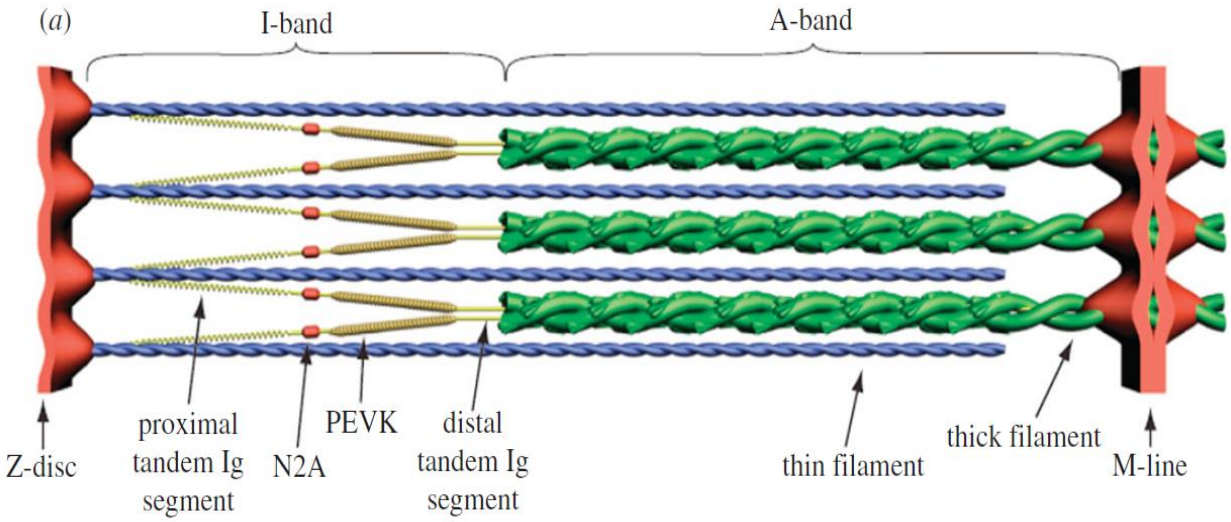


Figure 2.2: ***Schematic of a skeletal muscle half-sarcomere illustrating the layout of titin (yellow with a red N2A segment). Each titin molecule is bound to the thin filaments (blue) in the I-band, and to the thick filaments (green) in the A-band. Note that, for simplicity, thick filaments are illustrated as double-stranded, whereas in vertebrate skeletal muscle, they are actually triple-stranded. The N2A segment is located between the proximal tandem Ig segment and the PEVK segment*** **(Nishikawa *et al.*, 2012).**

2.1.1 Immunoglobulin (Ig) and Fibronectin Type III (FNIII) Domains

An immunoglobulin domain is a protein domain made up of a 2-layer sandwich, usually between 7 and 9 anti-parallel beta-strands in a pair of beta sheets, linked by a disulphide bond and hydrophobic interactions (Bork *et al.*, 1994; Brümmendorf and Rathjen, 1995). The immunoglobulin domain is found in numerous proteins and there are over 750 different genes in the human genome that contain sequences encoding immunoglobulin domains (Berg *et al.*, 2002). Examples of proteins that have the immunoglobulin domain include antibodies, T-cell receptors, antigen presenting molecules, receptor tyrosine kinases and titin molecules. Immunoglobulin domains are mainly known

to function as recognition and adhesion molecules and hence are highly significant in protein-protein interactions.

The Ig domain forms a super repeat pattern in both the I-band and the A-band of titin. In the I-band there is a proximal and distal tandem segment as shown in Figure 2.2. The tandem Ig chain has been reported to lengthen during sarcomere stretch (Gautel and Goulding, 1996). However, the extensibility of the tandem Ig has been reported to contribute less significantly to the generation of passive tension relative to the contribution of the PEVK domain during sarcomere stretch (Bennett *et al*., 1997; Linke *et al*., 1996).

Fibronectin is a dimeric glycoprotein composed of disulfide-linked subunits and it is made up of homologous repeats of three types: type I, II and III (Petersen *et al*., 1983; Baron and Norman, 1991). Fibronectin type III domain forms repeats similar to the Ig domain, but the repeats are only restricted to the A-band of titin. Fibronectin type III is one of the most common domains in proteins (Berry *et al.*, 2003). This domain is present in as much as 2% of proteins found in animals (Bork and Doolittle, 1992). Proteins that contain fibronectin type III domains are beta 4 subunit of integrin, titin, cytokine receptors, chaperonins, cell surface proteins and many more. The fibronectin type III structure resembles that of the Ig domain in such that it also has a beta-sandwich. This domain also binds to other proteins and molecules and hence this makes it really significant in protein-protein or cell-cell interactions.

2.1.2 PEVK Domain

As the name suggests, it constitutes high contents of Proline, Glutamine, Valine and Lysine. This region is part of the I-band and is known to contribute significantly to titin elasticity and generation of passive tension in the sarcomere. There are 112 exons of the titin gene that have been shown to be differentially spliced in the expression of the PEVK domain (Westermann *et al*., 2007). Differential splicing of the PEVK domain is influenced by the tissue in which it is expressed. For instance the PEVK domain of titin in the soleus is longer than the PEVK of titin in the psoas of rabbits (Freiburg *et al*., 2000). In the mammalian heart, two titin variants are co-expressed; the N2B and the N2BA isoforms. The N2B is known to contain a shorter PEVK domain than the N2BA isoform. For instance, the dog titin has an N2B titin with a shorter PEVK while the N2BA isoform has a longer PEVK domain with a length of 703 to 900 amino acids (Greaser *et al*., 2002). Skeletal muscles express the N2A isoform and contain a much longer PEVK domain of a few thousand residues (Bang *et al*., 2001). The report by Greaser (2001) demonstrated that there are repeating motifs of 26-28 amino acids occurring in the PEVK domain. The secondary/tertiary structures of the motifs have not been fully elucidated. One of the motifs is the P-P-A-K, so named because of the main amino acids that constitute the beginning of the motif. This motif exists in groups of 2-12 which are separated by the polyE segment. The polyE segment is made up of a high number of glutamate residues (approximately 45%). The PPAK motif has been found in the short N2B-PEVK domain of the cardiac muscle but no polyE motif has been identified in N2B-PEVK. On the other hand, both PPAK and polyE motifs have been identified in the skeletal muscle PEVK domain (Greaser, 2001). The arrangement of the residues in the PEVK has been proposed to be of

significance to the domain's structural and functional properties. Research by Gutierrez-Cruz *et al.* (2001) also suggested the existence of polyproline-type II helix-like configuration in skeletal fetal PEVK. In 2001, Li *et al.* also suggested that cardiac PEVK may be a random coil. This and many other structural descriptions of the PEVK domain have been studied through the use of PEVK-specific antibodies, circular dichroism and electron microscopy. More work is required in this field to fully elucidate the importance of the structure of the PEVK domain and its significance to the function of titin in the sarcomere. Nevertheless, the PEVK domain has been shown to be the main determinant of titin elasticity and passive tension generation in the sarcomere (Gautel and Goulding, 1996; Linke *et al*., 1996). The PEVK domain has been reported to function as a largely unfolded polypeptide that stretches at low force levels and thereby acts as a spring at physiological sarcomere conditions (Kulke *et al*., 2001; Yamasaki *et al*., 2001; Labeit *et al*., 2003). Similar data by Linke *et al.* (2002) supported the view that the PEVK domain acts as an entropic spring with properties of a random coil that exhibits mechanical conformations of different flexibility. Loss of the PEVK domain in cardiac muscle titin has been reported to lead to diastolic dysfunction development which indicates the significance of the PEVK domain to the function of titin in the muscle (Granzier *et al*., 2009). Elastic and force-extension properties of the PEVK domain has been investigated using methods such as single molecule atomic force spectroscopy, immunofluorescence and immunoelectron microscopy with isolated myofibril mechanics (Labeit and Kolmerer, 1995; Li *et al.*, 2001). Novel roles of the PEVK domain in hypertrophy signaling have also been identified by others where the PEVK domain was shown to induce the hypertrophy and fetal gene response, including the up-regulation of four and half LIM domain (FHL) proteins

(Granzier *et al*., 2009). PEVK domain has also been shown to influence and also be influenced by actin sliding and Ca^{2+}. Findings by Linke *et al*. in 2002 showed that cardiac PEVK has an effect on actin sliding but this effect was partially suppressed by physiological Ca^{2+} whereas skeletal PEVK was not influenced by physiological Ca^{2+} concentration and also had a lesser effect on actin sliding. The mechanism by which PEVK functions and influences other physiological processes is still not clear. More work is required in this field to investigate the roles of the PEVK domain in the efficient functioning of titin. This would enhance understanding of the PEVK domain and titin in the sarcomere.

2.1.3 N2A Region

The N2A region is found in the I-band segment of titin and it is ubiquitously expressed in all titin isoforms except the N2B titin found in cardiac tissue. It consists of four Ig domains (I-80, I-81, I-82 and I-83) in mouse. Between the I-80 and I-81 is a 117 amino acid unique insertion sequence known as 'IS' (Hayashi *et al*., 2008). The structure of N2A has not been fully elucidated but a model for the 'IS' structure has been created. The 'IS' region is in an extended conformation and constitutes predominantly alpha helices. The N and C-terminal ends are disordered and these ends can be coerced to form more stabilized helices. This model was generated through Circular Dichroism (CD) and fluorescence spectroscopy studies (Tiffany, 2014).

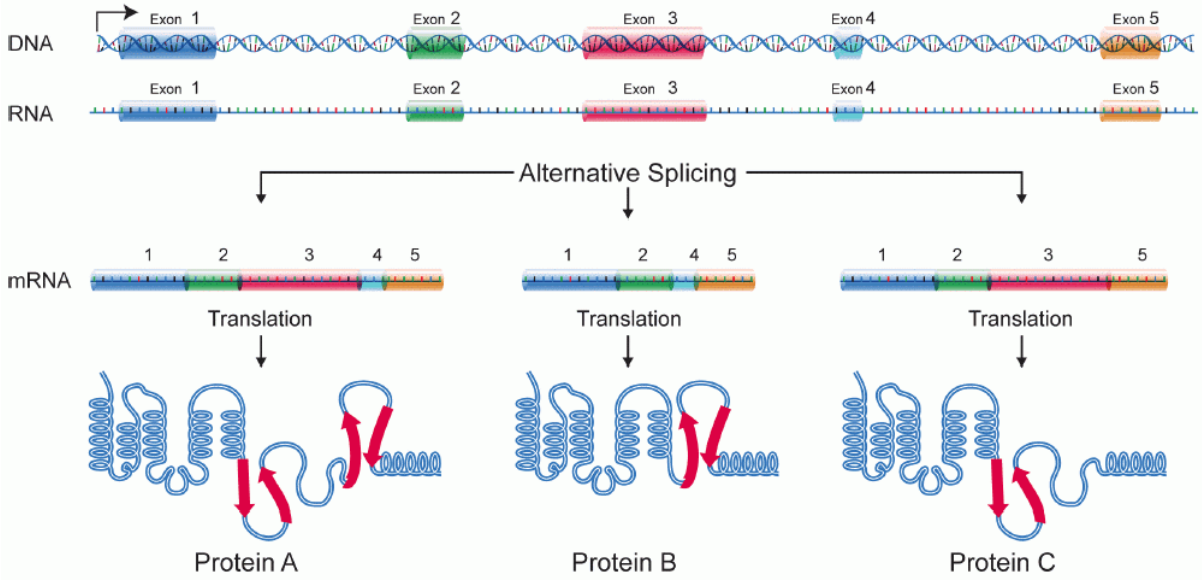
The N2A region is located in the extensible region (I-band segment) of titin and it has been reported to act as a spring element of titin. In addition, the N2A region seems to be involved in signal transduction in striated muscles via interactions with a number of

proteins. Muscle Ankyrin Repeat Proteins (MARPs), which are involved in cellular stress response, have been found to bind to the N2A region. The cardiac ankyrin repeat protein (CARP) and the two closely related proteins, ankrd2 and diabetes ankyrin repeat protein (DARP) are the ankyrin proteins that have been reported to bind to the N2A domain. In the N2A domain, the ankyrin proteins interact with a tyrosine-rich binding motif between the Ig80 and Ig81. This interaction is very important in the sarcomere since it has been reported to be involved in stretch signaling (Miller *et al*., 2003). Skeletal muscle-specific calpain homologue p94 also binds to multiple sites of titin, including the N2A region. Interaction between N2A and p94 has been proposed to suppress p94 autolysis and this protects titin from proteolysis (Hayashi *et al*., 2008).

2.2 ALTERNATIVE SPLICING

Alternative splicing is a key process by which cells regulate protein expression. Alternative splicing allows more than one protein to be expressed from a single gene (Figure 2.3), which promotes diversity and contradicts the initial hypothesis that stated that a single gene expresses only one protein, also known as one-gene-one-enzyme hypothesis (Beadle and Tatum, 1941). Genes are made up of introns and exons. Post transcriptional events lead to the splicing out of introns and the retention of exons. In the processing of the mRNA, some of the exons are excluded from the transcript and the skipping of particular exons determines the particular isoform being expressed. Through the regulation of this key process by the cell, multiple proteins are expressed from a single gene and this has even led some to propose that alternative splicing enhances economical storage of data by

eukaryote cells. Even though exon skipping is the most common alternative splicing mechanism observed, four other mechanisms also exist. They include mutually exclusive exon, alternative donor or acceptor site and intron retention. The mutually exclusive exon involves the retention of one out of two exons in the mRNA after splicing, but not both exons (Olson *et al*., 2007). Alternative donor or acceptor site mechanisms involve the use of alternative 3' and 5' splice junctions respectively. This mechanism changes the 5' boundary of the downstream exon in the case of the alternative donor site and changes the 3' boundary of the upstream exon in the case of alternative acceptor site (Breitbart *et al.*, 1987). Intron retention is the rarest mechanism of alternative splicing observed in mammals. As the name suggests, an intron is retained in the primary transcript in the expression of a protein (Black, 2003). Protein isoforms produced through alternative splicing differ in their amino acid sequence and consequently in their biological functions.



Source: [National Human Genome Research Institute](http://www.genome.gov/Images/EdKit/bio2j_large.gif) - http://www.genome.gov/Images/EdKit/bio2j_large.gif

Figure 2.3: ***Expression of multiple proteins from a single gene through alternative splicing. Different exons are involved in the expression of each isoform.***

Alternative splicing also plays a significant role in the complexity of eukaryotes. The discovery of alternative splicing partially explains the presence of enormous amounts of proteins that far exceeds the number of genes that make up the genome of eukaryotes (Nilson and Graveley, 2010). This complexity is explained by the ability of a single gene to variably express proteins and this variability in protein expression through alternative splicing has been shown to greatly contribute to biodiversity of proteins in eukaryotes (Black, 2003).

The variability that occurs in protein expression through alternative splicing can also be an avenue through which numerous diseases can arise. Mutation in a single gene that expresses multiple proteins would eventually lead to the expression of multiple mutant

proteins from that single mutant gene. Such proteins might be non-functional and hence can be the cause of numerous genetic diseases (Fackenthal and Godley, 2008). Also, since alternative splicing and its regulatory mechanisms are inherited, any changes in the regulatory processes would also be inherited and is currently known as one of the major causes of hereditary diseases (Lim *et al*., 2011). A typical example is cancer; abnormal splicing events have been associated with most cancerous cells (Skotheim and Nees, 2007; He *et al*., 2009).

Detailed mechanism of alternative splicing remains unclear. Understanding and regulating alternative splicing would have tremendous impact on public health. More work is therefore required in this field to further enhance the understanding of alternative splicing that has been accumulated since its first observation in 1977 (Berget *et al*., 1977).

2.2.1 Alternative Splicing In Titin

Even though titin is expressed from a single gene, it occurs in different isoforms as a result of alternative splicing. Alternative splicing as described earlier is the expression of multiple proteins from a single gene (Figure 2.3). The isoforms are produced in different striated muscle tissues and exhibit diverse elasticity. The high exon number of 363 is suggestive of the expression of numerous isoforms from the *TTN* gene (Figure 2.4) (Bang *et al.*, 2001). Freiburg *et al*. (2000) showed that differentially expressed tandem-Ig, PEVK and N2B spring elements of human titin are coded by 158 exons and that the exons are contained within a 106 kb genomic segment. Through their research, it was also found out that 155 exons were excluded from the expression leading to a 2.97 MDa cardiac titin isoform in the ventricular heart muscle. In mammalian atria heart muscle, 90 to 100 exons

were excluded in the expression, leading to a 3.3 MDa cardiac titin isoform. Through the same research, skeletal muscles of rabbits, the psoas and soleus were also shown to exhibit differential exon-skipping mechanisms that led to the expression of a 3.7 and 3.35 MDa titin isoforms respectively. The I-band of cardiac titin is known to constitute both the N2B and N2A elements whereas the I-band of skeletal titin is known to contain only the N2A element (Figure 2.5). The myocardium co-expresses a titin isoform that contains only the N2B element and another that contains both the N2A and N2B elements, known as N2BA (Labeit *et al.*, 2000). The N2B isoform is smaller with a molecular weight of 2970 kDa whereas the N2BA isoform is larger, ranging from approximately 3200 to 3400 kDa. Apart from the differences in the N2A and N2B elements, differences in titin isoforms are also seen in varying lengths of Ig domains and the PEVK regions. Even though such isoforms have been identified by their differential mobility in gel electrophoresis, the mechanisms by which the isoforms arise still remain unclear (Labeit and Kolmerer, 1995; Centner *et al*., 2000).

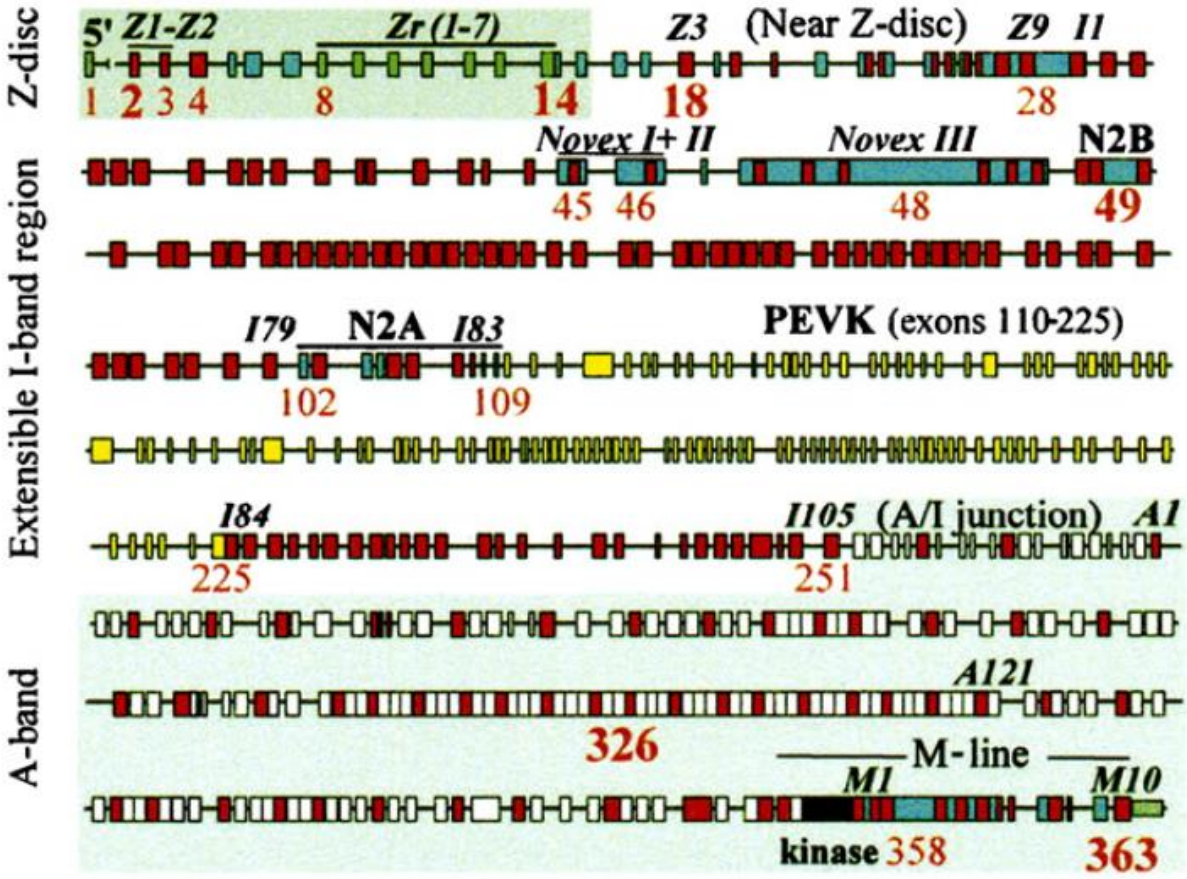


Figure 2.4: ***Outlined titin gene exons. This scheme illustrates the exons that are involved in the expression of both the cardiac titin and skeletal muscle titin. (LeWinter et al., 2007).***

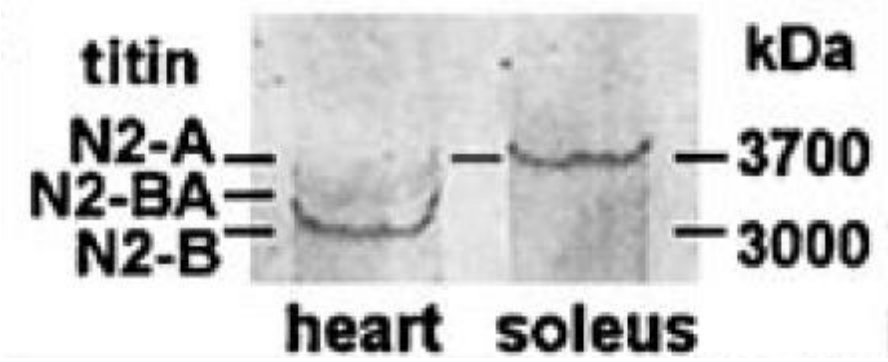


Figure 2.5: ***Isoforms of titin in the heart and the skeletal muscles. Rabbit left ventricle expresses mainly the N2B-titin isoform, as detected on 2% SDS-polyacrylamide gels (Linke et al., 1997). Rabbit soleus expresses N2A-titin*** **(Linke *et al.*, 2002).**

2.3 ELASTICITY AND FUNCTION OF TITIN

Titin functions as a molecular spring that develops passive tension when sarcomeres are stretched (Linke *et al*., 1996; Linke and Granzier, 1998; Linke *et al*., 1999). Passive tension is the force that is created when a sarcomere is stretched but not activated and it is this force that resists deformation when external force is applied to the sarcomere (Granzier and Labeit, 2004). The tension produced by titin filaments have been attributed to the elasticity of the I-band segment. Previous studies have also shown that the I-band of titin is expressed in varying lengths. The variability in the lengths of the I-band has also been shown to result in diverse elasticity of titin, leading to generation of varying passive force. Hence the passive tension properties of skeletal and cardiac muscles differ. Work by Freiburg and his team (2000) showed that cardiac muscles are stiffer than skeletal muscles. The N2B isoform of cardiac titin, which is a shorter spring, has also been shown to be stiffer than the N2BA isoform, which has a longer spring, and hence has increased passive stiffness relative to the N2BA isoform (Wu *et al*., 2002). The variability of passive stiffness adapted by different titin isoforms through alternative splicing has been suggested to be specific to the needs of the particular striated muscle. Some progress has been made in investigating the significance of titin in striated muscles but more work is needed in this field to determine and ascertain the specific functions of titin in different striated muscles. A few roles titin has been shown to play in the physiological function of muscles are:

- Functions as a molecular spring that generates passive tension (Linke and Granzier, 1998; Linke *et al*., 1999)
- Maintaining structural integrity of the sarcomere and setting the physiological sarcomere length range (Granzier and Labeit, 2004)

- Functions as an elastic framework to maintain the mechanical stability of muscles (Funatsu *et al*., 1990)

2.4 MUSCULAR DYSTROPHY

Muscular dystrophies simply refer to muscle diseases. The muscle diseases are associated with the skeletal muscles and results in the weakening of the musculoskeletal system which in turn negatively affects locomotion. Most muscular dystrophies are as a result of genetic mutations even though malnutrition has been implicated in some cases. The mutations result in the expression of misfolded proteins and hence they lose their biological functions. And since such mutations occur in the genes, muscular dystrophies tend to be inherited, but the patterns of inheritance differ from one muscular dystrophy to the other. One common example of a protein whose absence or misfolding is implicated in muscular dystrophies is dystrophin which causes Becker and Duchenne muscular dystrophies (Forrest *et al*., 1987; Koenig *et al*., 1987; Koenig *et al*, 1989). Titin has also been implicated in various muscular dystrophies. Distinct mutations that cause skeletal muscle diseases are classified as Titinopathy. Examples of titinopathies are tibial muscular dystrophy and limb-girdle muscular dystrophy 2J, both caused by mutations in the extreme C-terminus of the titin gene.

Common symptoms of muscular dystrophies are inability to walk, respiratory difficulties, joint contractures, muscle spasms, scoliosis, waddling gait, poor balance and muscle wasting. The signs and symptoms are often useful in the diagnosis of muscular dystrophies. Further diagnosis involves muscle biopsy, electromyopathy,

electrocardiography and DNA analysis. Muscular dystrophies have no known cure yet but procedures such as physical therapies and improved nutrition can be very useful when managing such diseases.

2.5 RNA SEQUENCING (SEQUENCING TECHNOLOGY)

Transcriptomics, also known as Expression Profiling, is the study of the whole transcriptome in a particular cell. This involves the examination of mRNA, rRNA, tRNA and other non-coding RNA transcribed in a given population of cells to identify genes in the genome that are being expressed at any given time. The transcriptome has been previously studied mainly by the use of High-throughput methods that depend on Microarray analysis. Microarray is a multiplex lab-on-a-chip described as a 2D array on a solid substrate that assays large amounts of biological materials using high-throughput screening miniaturized, multiplexed and parallel processing and detection methods. In transcriptomics, DNA microarray uses one or few probes to target particular parts of each gene to determine the expression levels of transcripts (Schena *et al*., 1995). Over the past years, Microarrays have contributed to the understanding of transcriptomics, however, it has some challenges that limit its use in fully characterizing and quantifying the diverse RNA molecules transcribed from genomes.

Currently, study of the transcriptome is widely being done by the use of Next-Generation Sequencing technology at the nucleotide level and this is known as RNA-Sequencing (RNA-seq) (Szabo, 2014). RNA-seq has replaced Microarray techniques due to its higher sensitivity, quantification and reproducibility of experiment. This high-

throughput technique allows RNA analysis by revealing the quality and quantity of RNA from a genome at a given time through large scale cDNA sequencing. RNA seq was introduced to the world through the publications of Nagalakshmi *et al*., Lister *et al*., Wilhelm *et al*., Cloonan *et al*. and Mortazavi *et al*., all in 2008. Early RNA-seq involved the use of reverse transcribed RNA into cDNA, amplification, fragmentation into short reads and then sequencing (Figure 2.6). Even though the RNA-seq technique is more advantageous than the Microarray technique in transcriptome studies, most of the steps involved in RNA seq served as avenues through which errors and biases were introduced into the sequenced data. For instance, reverse transcriptase used in RT-PCR unlike most DNA polymerases, has no proofreading ability, hence it has a high error rate which results in low sequence quality. Editing errors and errors due to DNA thermal damage also introduce bias to the data during the amplification procedure (Pienaar *et al*., 2006). Fragmentation into shorter reads increases the likelihood of incorrect mapping of transcripts with regions of repeated sequences. However, current developments in RNA-seq such as Direct RNA seq, Long Sequencing technology and Paired-End Seq overcome such biases by skipping the cDNA synthesis process, generating longer reads and sequencing both ends of a fragment respectively to generate high quality sequence data (Wang *et al*., 2009). Development of RNA-seq has tremendously improved understanding of transcriptomics in reference to alternative splicing events, gene fusion transcripts, sense and antisense transcripts, transcription initiation sites, unannotated exons, allele specific expressions, strand specific measurement and small RNA characterization (Ozsolak and Milos, 2011).

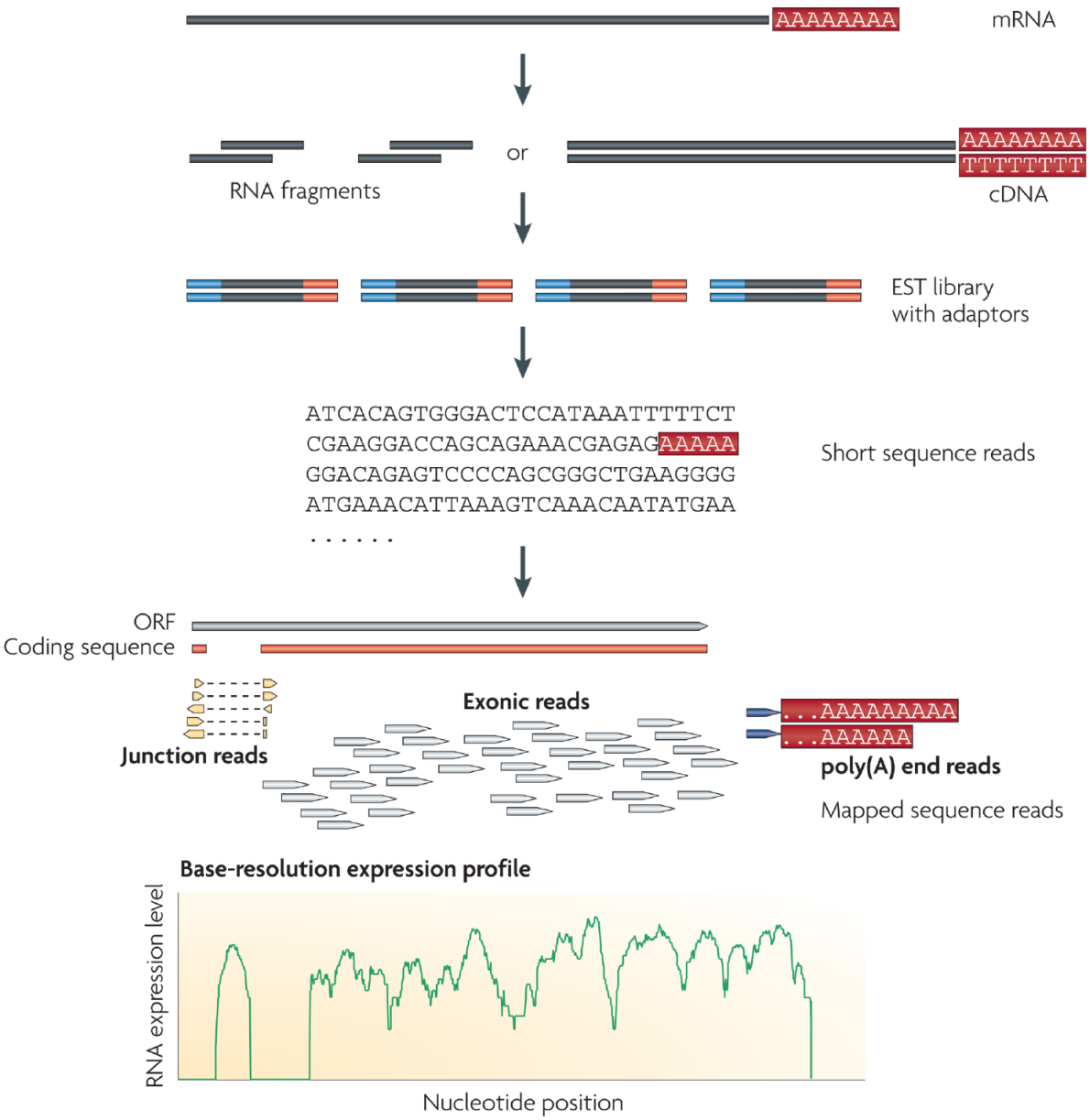


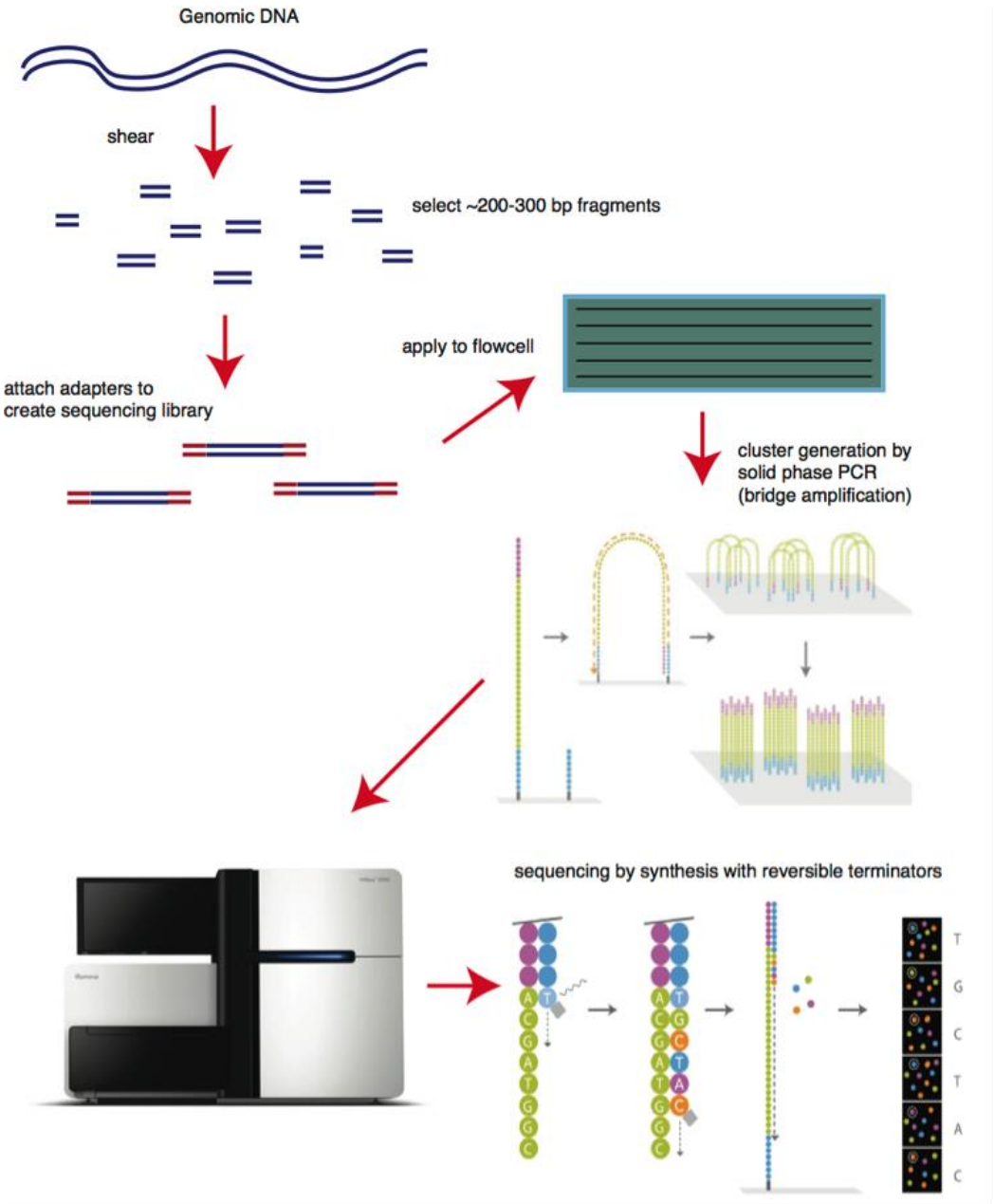
Figure 2.6: ***Summary of the RNA-Seq procedure. The mRNA is converted to cDNA and fragments from that library are used to generate short sequence reads. Those reads are assembled into contigs which may be mapped to reference sequences*** **(Wang *et al.*, 2009).**

2.5.1 Sequencing By Synthesis (SBS) – Illumina Sequencing Technology

Sequencing by synthesis is a next generation sequencing technology that is used by Illumina. This technology was initially invented by Shankar Balasubramanian and David Klenerman at the University of Cambridge. Illumina sequencing technology acquired the technology and has since improved on the original technology. SBS is currently used for whole-genome sequencing, DNA sequencing, Methylation analysis, De novo sequencing, RNA sequencing, Protein-Nucleic acid interaction analysis and candidate region targeted sequencing.

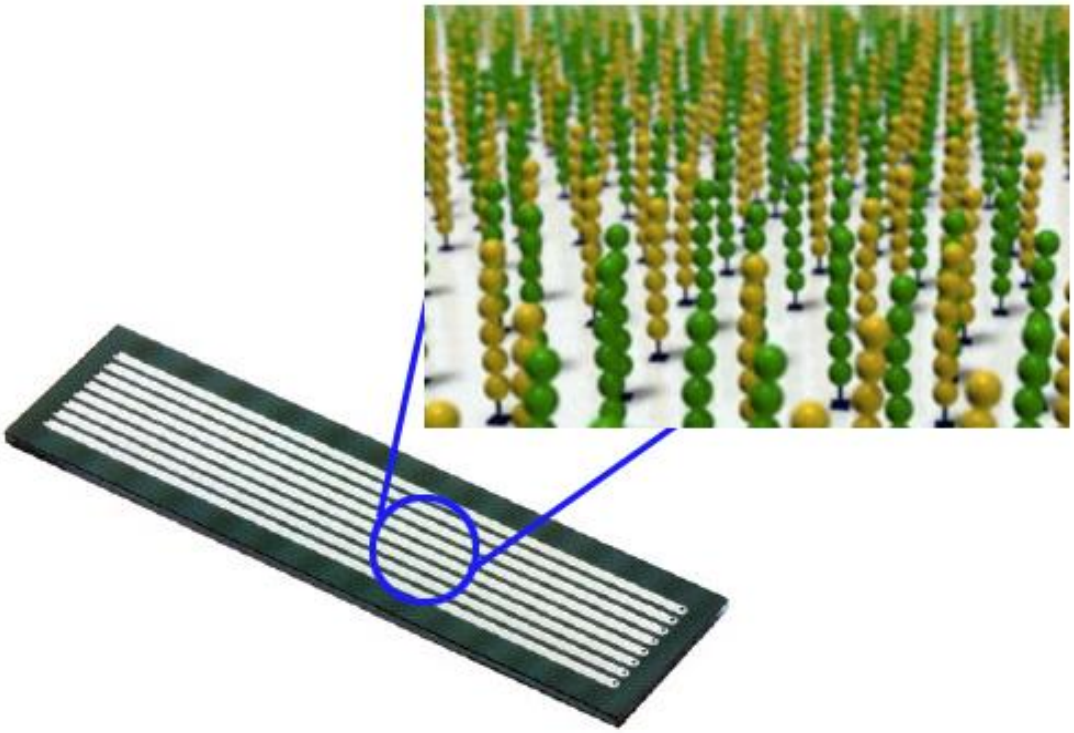
The SBS technology involves the use of fluorescent labeled terminator nucleotides. The nucleotide terminators (dNTPs) are reversible and hence only a single kind of nucleotide is allowed to bind at a time. The terminator is then removed and another cycle of dNTPs are added. The nucleotide fluoresces as extension of the DNA template proceeds and since all four dNTPs are present in each sequencing cycle, natural competition decreases bias and errors (Figure 2.7). This technology also minimizes bias associated with series of repeated nucleotides. The extension of the DNA fragment in SBS is carried out by DNA polymerase in terms of nucleotide addition and ligase in terms of oligonucleotide addition. The extension is done on a solid surface known as a flow cell (Figure 2.8). Adaptors that can stick to the surface of the flow cell are ligated to the DNA fragment and that enables the DNA fragment to position itself on the flow cell for nucleotide extension. Prior to the sequencing, the DNA fragments are amplified by polymerase chain reaction and this is described as Cluster Generation. Cluster generation turns the DNA libraries into clonal clusters on the flow cell. The clusters are generated through bridge amplification. In

bridge amplification, the DNA fragment to be amplified forms a bridge on the flow cell and is amplified. This process repeats until a cluster is formed.



Source: http://bitesizebio.com/13546/sequencing-by-synthesis-explaining-the-illumina-sequencing-technology/

Figure 2.7: ***Sequencing by synthesis. This involves the use of reversible terminators and fluorescently labeled dNTPs.***



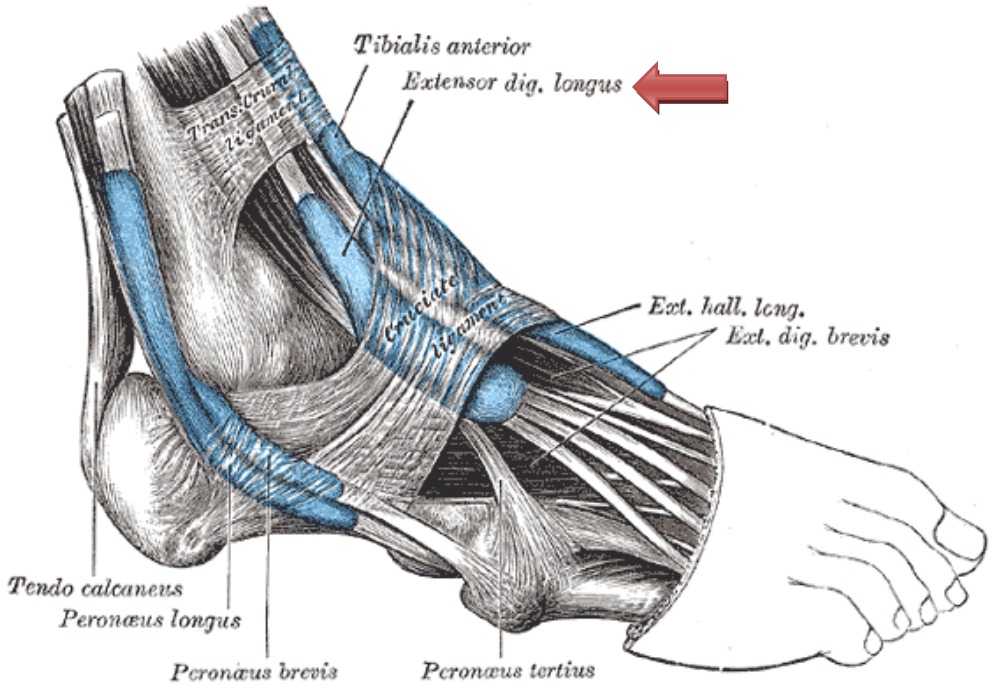
Source: http://www.giga.ulg.ac.be/upload/docs/image/jpeg/2009-04/illumina4.jpg

Figure 2.8: ***Amplified view of a flow cell. Embedded on the flow cell are adapters.***

2.6 DESCRIPTION OF DIFFERENT SKELETAL MUSCLE TYPES

2.6.1 *Elongus Digitorum Longus* (EDL)

The EDL is found in the outside of the lower leg, anterior part of the leg. It is a pinnate muscle that arises from the tibia, fibula, interosseous membrane fascia and the intermuscular septa. It splits into four parts that inserts into the four lesser toes as shown in Figure 2.9 (images depicting *Mus musculus* EDL were not readily available at the time of the study). The EDL aids in the flexion of the foot in the dorsal direction, movement of the sole of the foot away from the median plane and extension of the toes.

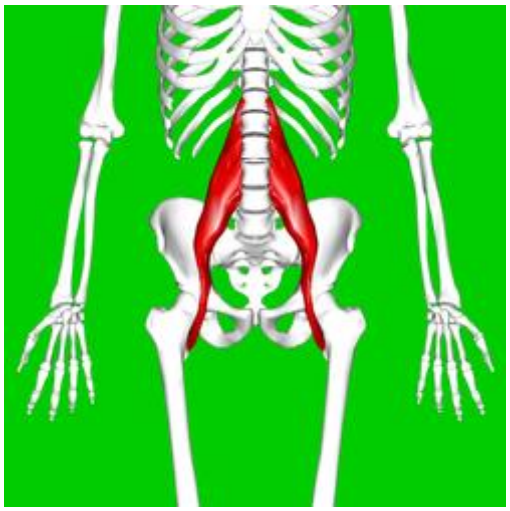


Source: http://en.wikipedia.org/wiki/Extensor_digitorum_longus_muscle

Figure 2.9: ***Position of the Extensor Digitorum Longus (EDL) in the Human Body.***

2.6.2 Psoas

The psoas muscle is located in the hips and it arises from the lumbar vertebrae as shown in Figure 2.10 (image depicting the *Mus musculus* psoas was not readily available at the time of the study). It merges with the Iliacus at their midpoint to form Iliopsoas. The psoas divides into two parts: the deep part, this is from the lumbar vertebrae I-V, and the superficial part, this is from the last thoracic vertebra, lumbar vertebrae I-IV and the intervertebral discs (Platzer, 2004). The psoas aids in hip joint flexion and hence is known as the most powerful flexor of the thigh at the hip joint. Together with other muscles, it contributes to the lifting of the upper leg towards the body when the body is fixed or vice versa. The psoas is a type II muscle in mice and it is fast twitching but it combines slow and fast twitching in fibers in humans (Nunes *et al.*, 1985; Arbanas *et al*., 2009).

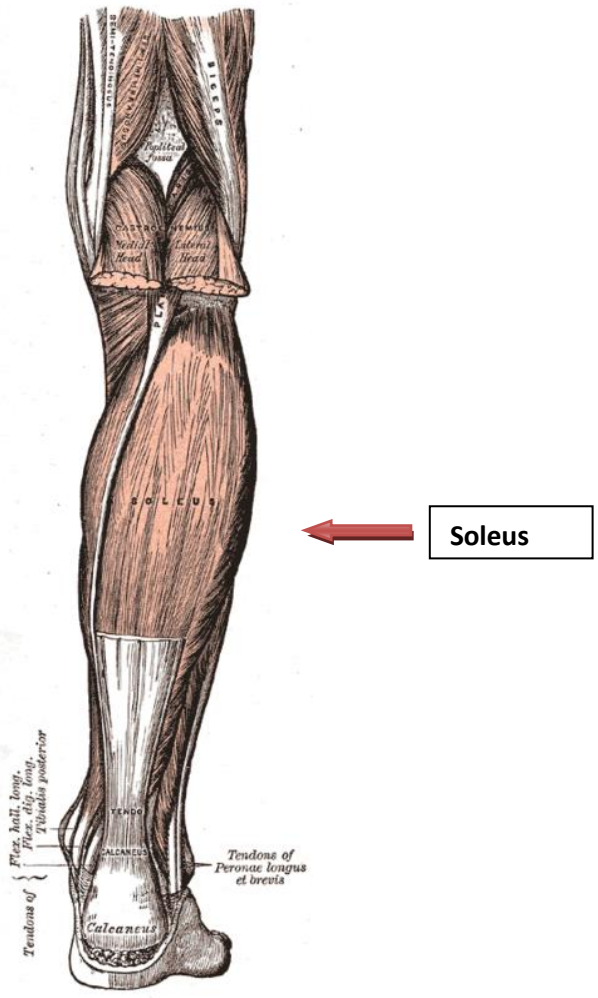


Source: http://en.wikipedia.org/wiki/Psoas_major_muscle

Figure 2.10: ***Position of the Psoas in the Human Body. Psoas is shown in red.***

2.6.3 Soleus

The soleus is situated in the calf; between the knee and the heel as depicted in Figure 2.11 (images depicting the *Mus musculus* soleus were not readily available at the time of the study). It is known to significantly aid actions such as walking, running and dancing. It also contributes to standing posture. The soleus also functions as a skeletal muscle pump that moves venous blood from the periphery to the heart (Botta *et al.*, 2001).



Source: http://en.wikipedia.org/wiki/Soleus_muscle

Figure 2.11: ***Position of Soleus in the Human Body.***

CHAPTER THREE

MATERIALS AND METHODS

3.1 MATERIALS

RNAlater RNA Stabilization reagent (*Catalog* # 76104) was purchased from Qiagen, USA.

RNeasy plus mini Kit (*Catalog* # 74134) was purchased from Qiagen, USA.

TruSeq Rapid SBS kit (*Catalog* # FC402-4001) was purchased from Illumina, USA.

Kapa Library Quantification kit (*Catalog* # KK4824) was purchased from Kapa Biosystems, USA.

High Sensitivity NGS Analysis Kit (*Catalog* # DNF-486) was purchased from Advanced Analytical Technologies, Inc., USA.

Quanti-iT RiboGreen RNA Assay Kit (*Catalog* # R11490) was purchased from ThermoFisher Scientific, USA.

Phusion High Fidelity DNA Polymerase (*Catalog* # M0530S) was purchased from New England Biolabs, USA.

Maxima H minus first strand cDNA Synthesis Kit (*Catalog* # K1681) was purchased from ThermoFisher Scientific, USA.

Diethylpyrocarbonate (DEPC) (*Catalog* # D5758) was purchased from Sigma-Aldrich, USA.

3.2 SAMPLE PREPARATION

Psoas, EDL and Soleus muscles excised from mice (*Mus musculus*) were provided by the Nishikawa laboratory in Northern Arizona University. The muscles were submerged in *RNAlater* immediately after excision and stored at -80 °C.

3.3 METHODS

The experimental strategy involved sequencing of the whole transcriptome and subsequent analysis of titin exons. Detailed descriptions of the experimental methods are outlined below.

3.3.1 RNA ISOLATION FROM SKELETAL MUSCLES

RNA was extracted using the Qiagen RNeasy plus mini kit. *RNAlater* stabilized tissue was placed into a cryotube that was half-filled with 1.00 mm Zirconia/ Silica beads. Six hundred microliters of RLT Plus buffer (lysis buffer) was added and compacted with 1.00 mm Zirconia/ Silica beads to the brim. The tissue was then disrupted and homogenized for 10 min. Needle-pointed pipettes were used to transfer the lysates into a micro centrifuge tube and centrifuged for 3 min at 10,000 rpm. The supernatant was carefully removed and 600 µl of 70% ethanol was added to the supernatant. It was then mixed immediately. Seven hundred microliters of the sample was transferred into RNeasy spin column in 2 ml collection tubes and centrifuged for 30 s at 10,000 rpm. The flow through was discarded and 700 µl buffer RWl was added to the RNeasy spin column. It was then centrifuged for 30 s at 10,000 rpm and the flow through discarded. Five hundred microliters of buffer RPE

was also added and centrifuged for 30 s at 10,000 rpm. The flow through was discarded. The centrifugation was repeated with 500 µl RPE for 2 min. The flow through was again discarded and the centrifugation repeated for 1 min. The flow through was discarded. The RNeasy column was then placed in a new 1.5 ml collection tube and 50 µl of RNase-free H_2O was directly added to the spin column membrane. It was then centrifuged for 1 min at 10,000 rpm to elute the RNA. The isolated RNA was then stored at -20 ℃.

3.3.2 RNA QUALITY ASSESSMENT

Qualitative analysis of RNA samples was carried out using an Advanced Analytics Fragment Analyzer *(High Sensitivity RNA Analysis Kit – Catalog # DNF-491/ User Guide DNF-491-2014AUG13).*

Aliquots of RNA were heat-denatured at 70 ℃ for 2 min and cooled immediately on ice to 4 ℃. An aliquot of 18 µl of the high sensitivity RNA Diluent Marker (DM) was added to each well. An aliquot of 2 µl of RNA was then added and mixed thoroughly. The plate was then centrifuged briefly to ensure that there was no air bubbles trapped at the bottom of the wells. RNA quality was then measured and analyzed using a bioanalyzer system at an absorbance ratio of 260 nm and 280 nm ($A_{260/280}$).

3.3.3 RNA QUANTIFICATION

Quantitative analysis of RNA samples was carried out using a RiboGreen quantification kit *(Quant-iT RiboGreen RNA Assay Kit – Catalog # R11490/ The Molecular Probes® Handbook, 11th Ed, Chapter 8 page 344)* as described below.

The RNA samples were diluted with Tris-EDTA (10 Mm Tris-HCl, 1 mM EDTA, pH 7.5) to a final volume of 1.0 ml. An aliquot of 1.0 ml of the aqueous working solution of the Quant-iT RiboGreen reagent was added to each sample and incubated for 5 min at room temperature away from light. The fluorescence was then measured and quantified.

Given satisfactory quality and quantity, RNA sequencing was carried out with the Illumina TruSeq RNA kit *(TruSeq RNA Sample Preparation Kit v2 – Catalog # RS-122-2001/ TruSeq RNA Sample Preparation Guide v2 15026495 Rev. F)*. All the sequencing was done by the genetics core laboratory at the University of Arizona.

WORKFLOW FOR RNA SEQUENCING

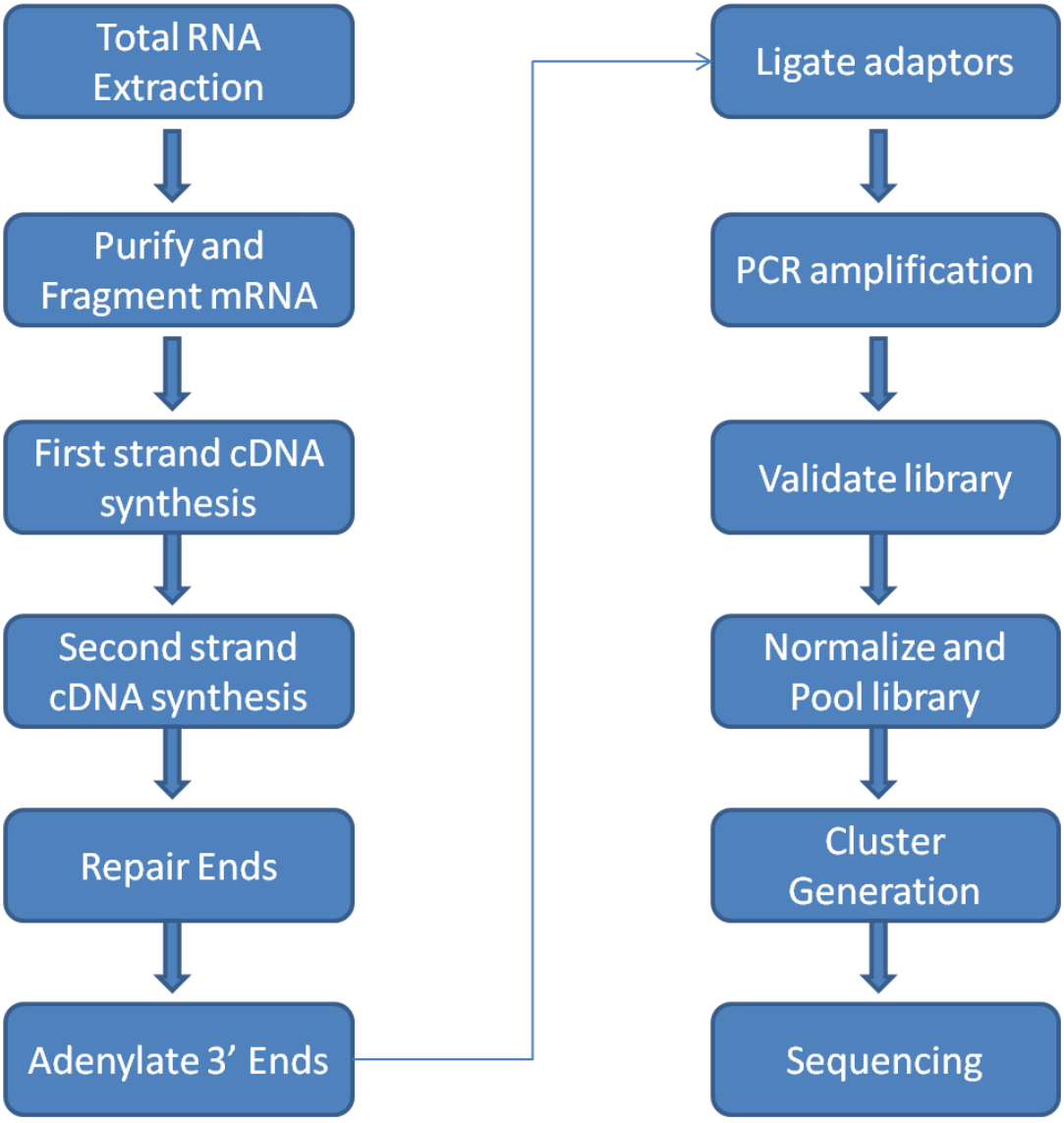


Figure 3.1: ***Workflow for RNA Sequencing***

3.3.4 POLYA mRNA ISOLATION AND FRAGMENTATION

A mass of 0.6 ng of the total RNA was diluted with nuclease-free ultra pure water to make a final volume of 50 µl in a 96-well 0.3 ml PCR plate labeled RBP. RNA purification beads (oligo-dT beads) were vortexed vigorously to resuspend. An aliquot of 50 µl of the beads was mixed with the diluted total RNA to enable the oligo-dT beads to bind to the polyA RNA. The mixture was pipetted up and down 6 times to mix thoroughly. The RBP plate was then sealed with a Microseal 'B' adhesive and placed in a thermo cycler programmed at 65 °C for 5 min and 4 °C for hold. This temperature denatures the RNA and facilitates binding of the polyA RNA to the beads. The RBP plate was then incubated at room temperature for 5 min to allow binding of the RNA to the beads. The adhesive seal was removed and the RBP plate placed on a magnetic stand at room temperature for 5 min to separate the polyA RNA bound beads from the solution. The supernatant was then pipetted out and discarded. The plate was then removed from the magnetic stand and the beads were washed by adding 200 µl Bead Washing buffer to remove unbound RNA. Washing was done by pipetting the mixture up and down 6 times to mix thoroughly. The RBP plate was again left on the magnetic stand at room temperature for 5 min and the supernatant removed and discarded. The supernatant mainly contains most of the ribosomal and other non-messenger RNA. The plate was removed from the magnetic stand, 50 µl of elution buffer was added to each well and mixed thoroughly. The plate was then placed in a pre-programmed thermal cycler: 80 °C for 2 min and 25 °C hold. This was to elute the mRNA and any other contaminant rRNA from the beads. After, the adhesive seal was removed and 50 µl of bead building buffer was added to each well to allow the mRNA to specifically rebind the beads while reducing the amount of rRNA that non-specifically binds. The plate

was then allowed to stand for 5 min at room temperature and another 5 min on the magnetic stand. The supernatant was then removed and discarded. The plate was removed from the magnetic stand and 200 µl of bead washing buffer was added to each well to wash the beads. The plate was again placed on the magnetic stand for 5 min at room temperature and all supernatant was removed and discarded. The plate was then removed from the magnetic stand and 19.5 µl of Elute, Prime, Fragment Mix was added to each well and mixed thoroughly. The Elute, Prime, Fragment mix contains random hexamers for RT priming and served as the first strand cDNA synthesis reaction buffer. The RBP plate was then sealed with a Microseal 'B' adhesive and placed in a pre-programmed thermal cycler: 94 °C for 8 min and 4 °C hold. The RBP plate was then centrifuged briefly at 600 × g for 5 seconds.

3.3.5 FIRST STRAND cDNA SYNTHESIS

The RBP plate was placed on a magnetic stand for 5 min at room temperature and 17 µl of each supernatant was transferred into a new 0.3 ml PCR plate labeled CDP. The CDP plate was then removed from the magnetic stand. To each 9 µl of the first strand master mix, 1 µl SuperScript II was added, mixed gently and centrifuged at 600 × g for 5 s. An aliquot of 8 µl of the first strand master mix and superscript II mix was added to each well of the CDP plate and mixed thoroughly. The CDP plate was then sealed and centrifuged at 600 × g for 5 s. The sealed CDP plate was then placed in a pre-programmed thermal cycler to run the first strand synthesis: Pre-heated lid at 100 °C, 25 °C for 10 min, 42 °C for 50 min, 70 °C for 15 min and Hold at 4°C.

3.3.6 SECOND STRAND cDNA SYNTHESIS

Twenty five microliters of the second strand master mix was added to each well of the CDP plate and mixed thoroughly. The plate was then sealed with a microseal 'B' adhesive and placed in a pre-heated thermal cycler to incubate for 1 h at 16 °C. After, the CDP plate was placed on a bench to attain room temperature and 90 µl of well-mixed AMPure beads was added to 50 µl ds cDNA in each well. The entire volume was mixed gently but thoroughly by pipetting up and down 10 times. The plate was then incubated at room temperature for 15 min and on a magnetic stand for 5 min. An aliquot of 135 µl of supernatant was removed and discarded from each well. With the CDP plate on the magnetic stand, 200 µl of freshly prepared 80% EtOH was added to each well without disturbing the beads and incubated at room temperature for 30 s. The supernatant was removed and discarded. The washing was repeated and the plate was allowed to dry at room temperature. The plate was then removed from the magnetic stand, 52.5 µl of resuspension buffer was added and mixed thoroughly. The plate was then incubated at room temperature for 2 min and placed on the magnetic stand for 5 min. An aliquot of 50 µl of the supernatant (ds cDNA) was then transferred from the CDP plate to a new 96-well 0.3 ml PCR plate labeled IMP.

3.3.7 END REPAIR OF DOUBLE STRANDED (ds) cDNA

Ten microliters of resuspension buffer were added to 50 µl ds cDNA in each well of the IMP plate. An aliquot of 40 µl end repair mix was then added and mixed thoroughly. The IMP plate was then sealed with a Microseal 'B' adhesive and placed in a pre-heated thermal cycler for incubation at 30 °C for 30 min. After the incubation, 160 µl of well-

mixed AMPure XP beads was added to each well and mixed thoroughly. The plate was then incubated at room temperature for 15 min and placed on the magnetic stand for 5 min. An aliquot of 127.5 µl supernatant was then removed and discarded from each well. This was repeated. With the IMP plate on the magnetic stand, 200 µl freshly prepared 80% EtOH was added to each well without disturbing the beads and incubated at room temperature for 30 s. The supernatant was removed and discarded without disturbing the beads and the washing repeated. The IMP plate was then allowed to stand at room temperature for 15 min to dry and then removed from the magnetic stand. An aliquot of 17.5 µl resuspension buffer was added to each well and mixed thoroughly. The plate was incubated at room temperature for 2 min on the bench and 5 min on the magnetic stand. An aliquot of 15 µl of the supernatant was transferred into a new 0.3 ml PCR plate labeled ALP.

3.3.8 ADENYLATING THE 3' ENDS

An aliquot of 2.5 µl of resuspension buffer was added to each well of the ALP plate, 12.5 µl of A-Tailing mix was also added and mixed thoroughly. The plate was then sealed with a Microseal 'B' adhesive and placed in a pre-programmed thermal cycler: Pre-heated lid 100 °C, 37 °C for 30 min, 70 °C for 5 min and hold at 4 °C.

3.3.9 LIGATING ADAPTERS TO SAMPLES

To each well of the ALP plate, 2.5 µl of resuspension buffer and 2.5 µl of ligation mix were added and mixed thoroughly. An aliquot of 2.5 µl of RNA adapter index was also added, mixed thoroughly and the plate was sealed. The plate was then incubated at 30 °C for 10 min. An aliquot of 5 µl of stop ligation buffer was added and mixed thoroughly to

inactivate the ligation. An aliquot of 42 µl mixed AMPure XP beads was added and mixed thoroughly. The ALP plate was then incubated at room temperature for 15 min on a bench and 5 min on a magnetic stand. An aliquot of 79.5 µl of the supernatant was removed and discarded from each well without disturbing the beads. With the ALP plate on the magnetic stand, 200 µl of freshly prepared 80% EtOH was added to each well without disturbing the beads. The plate was incubated at room temperature for 30 s and all the supernatant was removed and discarded without disturbing the beads. The washing was repeated and the plate left on the magnetic stand to air-dry at room temperature. The plate was removed from the magnetic stand and 52.5 µl of resuspension buffer was added to each well and mixed thoroughly. The plate was then incubated at room temperature for 2 min and on a magnetic stand for 5 min. An aliquot of 50 µl of the supernatant was then transferred to a new 0.3 ml PCR plate labeled with a CAP barcode without disturbing the beads. An aliquot of 50 µl of mixed AMPure XP beads was then added to each well of the CAP plate and mixed thoroughly for a second cleanup. The CAP plate was then incubated at room temperature for 15 min and on the magnetic plate for 5 min. An aliquot of 95 µl of supernatant from each well of the CAP plate was removed and discarded without disturbing the beads. With the plate on the magnetic stand, 200 µl of freshly prepared 80% EtOH was added to each well without disturbing the beads. The plate was then incubated at room temperature for 30 s and all the supernatant was removed and discarded. The washing was repeated. With the CAP plate on the magnetic stand, the plates were air-dried at room temperature for 15 min and then removed from the magnetic stand. An aliquot of 22.5 µl resuspension buffer was added to each well of the plate and mixed thoroughly. The plate was then incubated at room temperature for 2 min and on a magnetic stand for 5 min. An

aliquot of 20 µl of the supernatant from each well was then transferred into a new 0.3 ml PCR plate labeled with PCR barcode without disturbing the beads.

3.3.10 ENRICHMENT OF DNA FRAGMENTS (PCR)

Five microliters of PCR prime cocktail were added to each well of the PCR plate and 25 µl of PCR Master Mix was also added. They were mixed thoroughly and the plate was sealed with a Microseal 'B' adhesive. The plate was then put in a pre-programmed thermal cycler: Pre-heated lid at 100 °C, 98 °C for 30 s, 15 cycles of (98 °C for 10 s, 60 °C for 30 s and 72 °C for 30 s), 72 °C for 15 min and 10 °C hold.

After the PCR, the samples were cleaned up by adding 50 µl of mixed AMPure XP beads to each well and mixed thoroughly. The plate was incubated at room temperature for 15 min and on a magnetic stand for 5 min. An aliquot of 95 µl of the supernatant was removed and discarded while the plate was still on the magnetic stand. An aliquot of 200 µl of freshly prepared 80% EtOH was added to each well and incubated at room temperature for 30 s. All the supernatant were removed, discarded and the washing repeated. The samples were then air-dried at room temperature for 15 min and then the plate was removed from the magnetic stand. The dried pellet in each well was resuspended in 32.5 µl resuspension buffer and mixed thoroughly. The plate was incubated at room temperature for 2 min and on a magnetic stand for 5 min. An aliquot of 30 µl of clear supernatant was then transferred from each well of the PCR plate into a new 0.3 ml PCR plate labeled with TSP1 barcode.

3.3.11 ASSESSMENT OF LIBRARY QUALITY

Upon library build completion, samples had quality and average fragment size assessed with the Advanced Analytics Fragment Analyzer *(High Sensitivity NGS Analysis Kit – Catalog # DNF-486 / User Guide DNF-486-2014MAR10)* as described below.

One microliter of resuspeneded construct was diluted with 1 µl RSB and loaded on an advanced fragment Analyzer using Standard sensitivity NGS kit.

3.3.12 QUANTIFICATION OF LIBRARY

The quantity was assessed with an Illumina Universal-Adaptor-specific qPCR kit from Kapa Biosystems *(Kapa Library Quantification kit for Illumina NGS – Catalog # KK4824 / KAPA Library Quantification Technical Guide - AUG2014)* as described below.

Serial dilutions of 1:1000 to 1:8000 of the purified library were prepared using a Library Dilution buffer. Aliquots of 12 µl of KAPA SYBR FAST qPCR Master mix containing Primer premix, 4 µl of PCR-grade water and 4 µl of diluted library DNA were loaded into the qPCR plate and sealed. The plate was then centrifuged at 600 × g for 5 s to collect all components in the bottom of the wells. The qPCR plate was then put in a pre-programmed real-time thermocycler: Initial activation/ Denaturation 95 °C for 5 min and 15 cycles each of Denaturation 95 °C for 30 s and Annealing/ Extension/Data acquisition 60 °C for 45 s.

3.3.13 NORMALIZING AND POOLING LIBRARIES

Ten microliters of sample library from the TSP1 plate was transferred into a new MIDI plate labeled with a DCT barcode. The concentration of the sample library in each

well of the DCT plate was normalized to 10 nM using Tris-HCl 10 mM, pH 8.5 with 0.1% tween 20 and sealed with a Microseal 'B' adhesive. An aliquot of 10 µl of each normalized sample library from the DCT plate was transferred into a new 0.3 ml PCR plate labeled with PDP barcode. The total volume in each well of the PDP plate was 10× the number of combined sample libraries.

3.3.14 SEQUENCING RUN ON HiSEQ2500

After the samples were equimolar-pooled, they were clustered for sequencing on the HiSeq2500 machine. The sequencing run on the HiSeq2500 machine was performed using Illumina Rapid-Run SBS chemistry *(TruSeq Rapid SBS Kit – Catalog # FC-402-4001 / HiSeq Rapid SBS Kit v2 Reagent Prep Guide – 15058772A)* and the data were sent to the University of Arizona Genetic Core Biocomputing Group for further analysis.

3.4 DATA ANALYSIS

Sequencing was performed at the University of Arizona Genetic Core. The data analysis was carried out in the Caporaso laboratory (situated in the Center of Microbial Genetics and Genomics) at Northern Arizona University. The sequence data were obtained through irods data management system (Rajasekar *et al*., 2010). Sample identifiers were assigned to reads using QIIME 1.9.0 (Caporaso *et al*., 2010). Reads with a Phred quality score less than 3 were discarded. Reads were then aligned to reference *Ttn* exons (ExonMine) on a per sample basis using nucleotide blat (Kent, 2002). The reference *Ttn* exons were made up of 322 exons, starting from exon 3234_Ttn_36665 to 3234_Ttn_36987. Reads were not included in subsequent analysis if the e-value was greater than $1e^{-10}$, if the minimum percent identity was less than 0.75 or if the fraction of the query sequence in the alignment was less than 0.35. Plots of coverage of exons were generated based on QIIME output where sequences that aligned to more than one exon were included in the plots for both exons.

CHAPTER FOUR

RESULTS

4.1 Transcriptome Sequencing and Titin Exons Analysis

The purpose of the study was to determine whether different skeletal muscles in *Mus musculus* express different isoforms of titin and to investigate the exon skipping patterns involved in titin splicing. To investigate this, the entire mouse muscle transcriptome was sequenced and analyzed for the exons of the titin transcript using 322 reference titin exons. The total mRNA was isolated from each of the muscle types, fragmented and sequenced using the next generation sequencing method (RNA-seq). The transcriptome generated from sequencing the total mRNA was mapped to 322 reference titin exons from *Mus musculus*. Sequences from the transcriptome that matched the reference titin exons were recorded as hits. Out of the 322 exons, 7 exons showed no hits in all the tissue samples. In cases where no hits were recorded, it implied that the sequences of those exons did not have any matching sequences from the transcriptome, further implying that they were excluded during transcription. And most importantly, this further implied that the exons were not involved in the expression of titin in the EDL, Psoas and Soleus muscle tissues. The 7 exons were 3234_Ttn_36771, 3234_Ttn_36822, 3234_Ttn_36823, 3234_Ttn_36824, 3234_Ttn_36825, 3234_Ttn_36826 and 3234_Ttn_36829. The exons have a total translation capacity of 135 amino acids (Table 4.1).

Table 4.1: Exons that were excluded in all Three Muscle Tissues

EXONS (3234_Ttn_)	NUCLEOTIDE SIZE	AMINO ACID
36710	2646	882
36771	27	9
36802	354	118
36808	291	97
36821	84	28
36822	84	28
36823	50	16
36824	66	22
36825	84	28
36826	29	9
36829	71	23
Total	**3786**	**1260**

The table shows the eleven exons that were skipped in all three titin isoforms in the different skeletal muscles and their corresponding nucleotide lengths. The amino acid translation capacity of each of the exons is also shown. The translation capacity refers to how many amino acids can be coded from each of the exons. Hence a deletion of an exon is synonymous to deletion of the corresponding amino acids it codes for. Exon 3234_Ttn_36710 was the longest in nucleotide length and exon 3234_Ttn_36771 was the shortest in nucleotide length.

Even though the hits showed the number of reads that mapped to sequences of the reference exons, the reads might not completely align to the whole length of the exon sequence. This most likely implied that such reads were only observed as hits because they had sequence similarity with the exon sequence even though they were not involved in the coding of the titin exon. The possibility of this occurring was not ruled out because the reads that were being used to map the reference exons were RNA transcripts from the mouse transcriptome. The transcriptome, as explained earlier, is the set of all RNA molecules transcribed in a population of cells, implying that there is an enormous pool of sequences which might likely align with sequences of some titin exons even though they are not titin coding sequences. Therefore, in order to ensure that the reads observed mapped completely to their corresponding titin exons, further analysis was carried out using a base-resolution profile. The base-resolution profile was generated using QIIME. This software aligned the reads obtained for each exon to its corresponding reference exon to determine whether the reads mapped the complete sequence length of each reference exon. Through the base-resolution profile analysis, the confidence level of the presence or absence of a particular titin exon in the transcriptome was increased. The base-resolution analysis increased the confidence level from number of reads that had similar sequences with the reference exons to reads whose sequences mapped the entire length of the whole exon sequence, which indicated that a particular titin exon was present or absent in the transcriptome. An example of reads that covered the entire length of an exon is depicted in Figure 4.1. The base-resolution profile confirmed the exclusion of the 7 exons mentioned earlier. An example of the confirmatory data is shown in Figure 4.2. Analysis through the base-resolution profile identified four more exons that were excluded in the titin transcript.

The four exons received hits but the hit sequences did not have complete coverage of the entire sequences of the exons. An example is depicted in Figure 4.3. This indicated that the hit sequences that were registered as hits to the individual exons might be sequences that were involved in the expression of other proteins but not that of titin. The four exons were 3234_Ttn_36710, 3234_Ttn_36802, 3234_Ttn_36808 and 3234_Ttn_36821. These exons comprised a total of 3375 nucleotides which had a translation capacity of 1125 amino acid. In all, 11 exons were excluded in all three samples. The 11 exons comprised a total of 3786 nucleotides with a translation capacity of 1260 amino acids (Table 4.1). The fact that titin from the three tissues had the 11 exons deleted, implied that titin from the three tissues were shortened by 1260 residues.

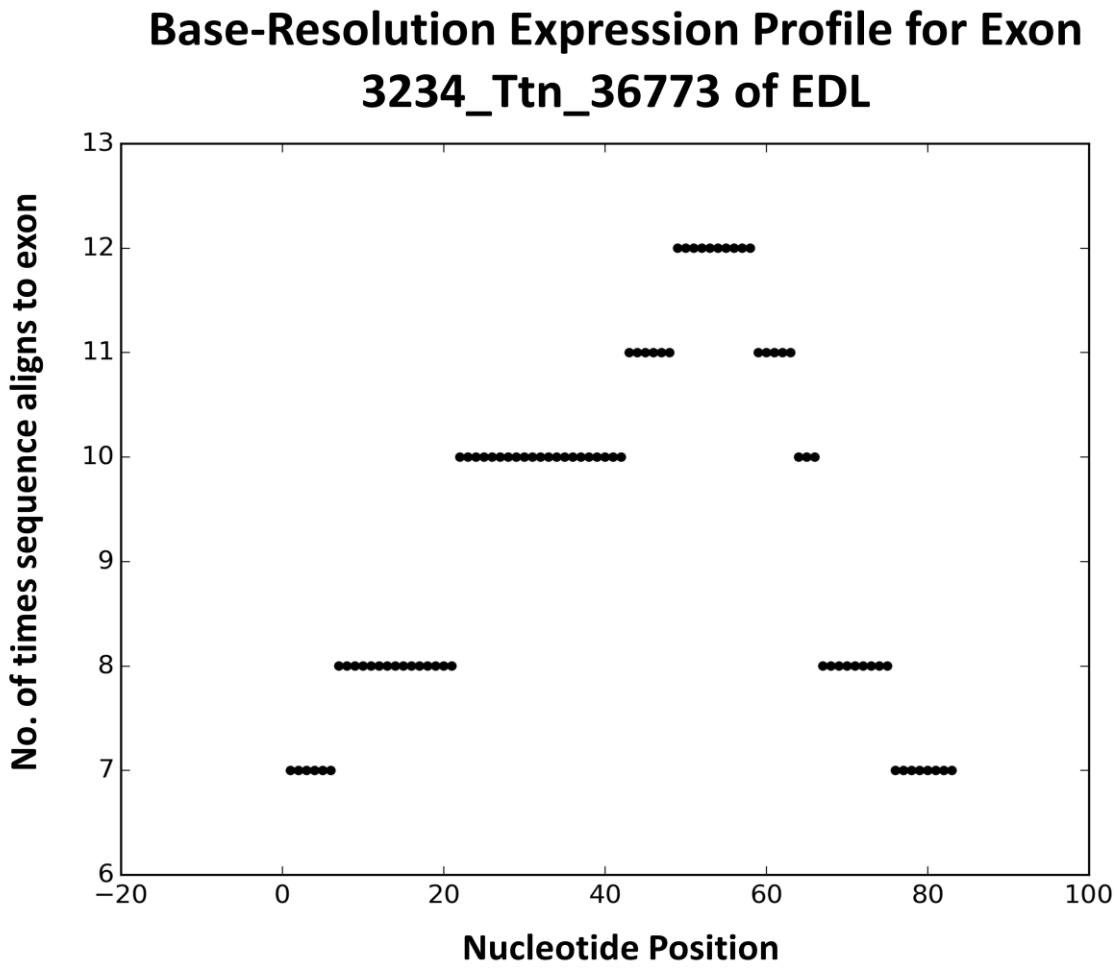


Figure 4.1: ***Base-Resolution Profile showing Sequence coverage for exon 3234_Ttn_36773 for EDL titin. The black bands indicated sequence similarity and the y-axis represent how many times the sequences aligned to the exons. Exon 3234_Ttn_36773 was an example of the exons that had complete coverage of their sequence in the transcriptome. This meant that this exon was present since its full sequence was accounted for in the transcriptome.***

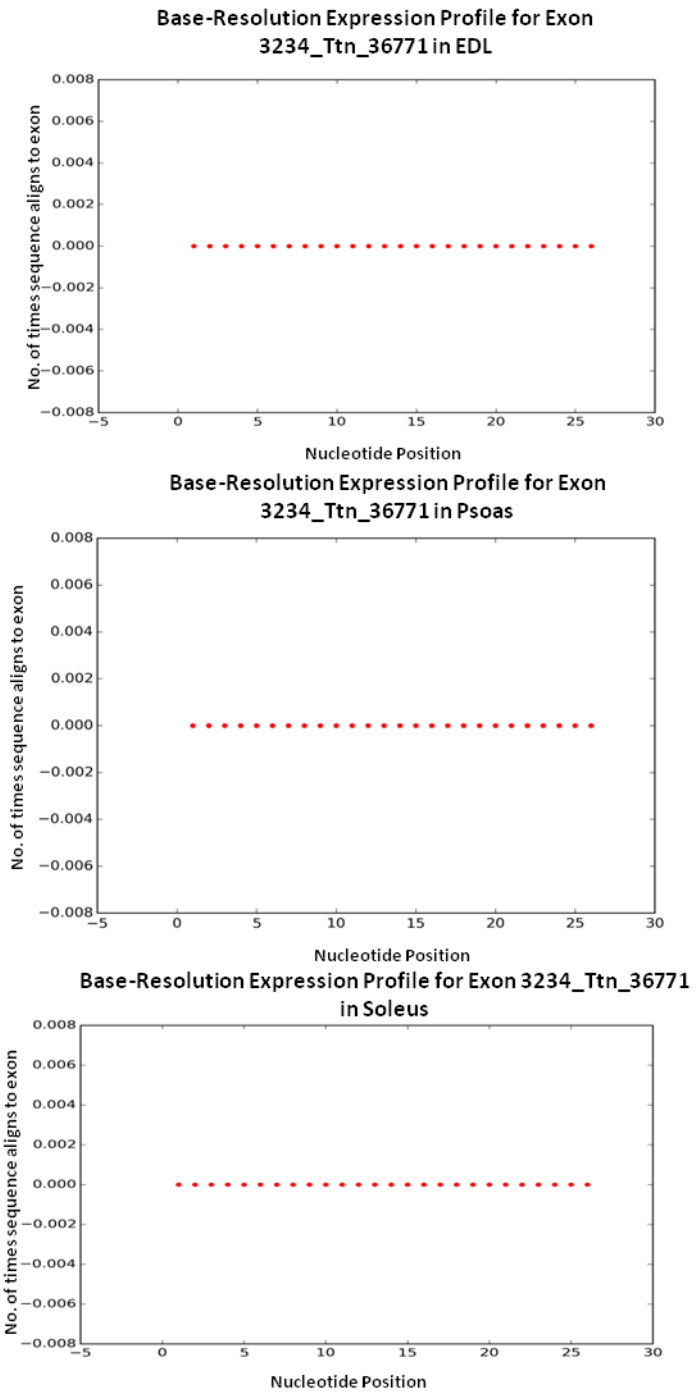


Figure 4.2: ***Base-Resolution Profile showing sequence coverage for exon 3234_Ttn_36771 for EDL, psoas and soleus titins. There was no sequence coverage for titin of each of the muscle types. Exon 3234_Ttn_36771 is an example of the 7 exons that received no hits in all three samples.***

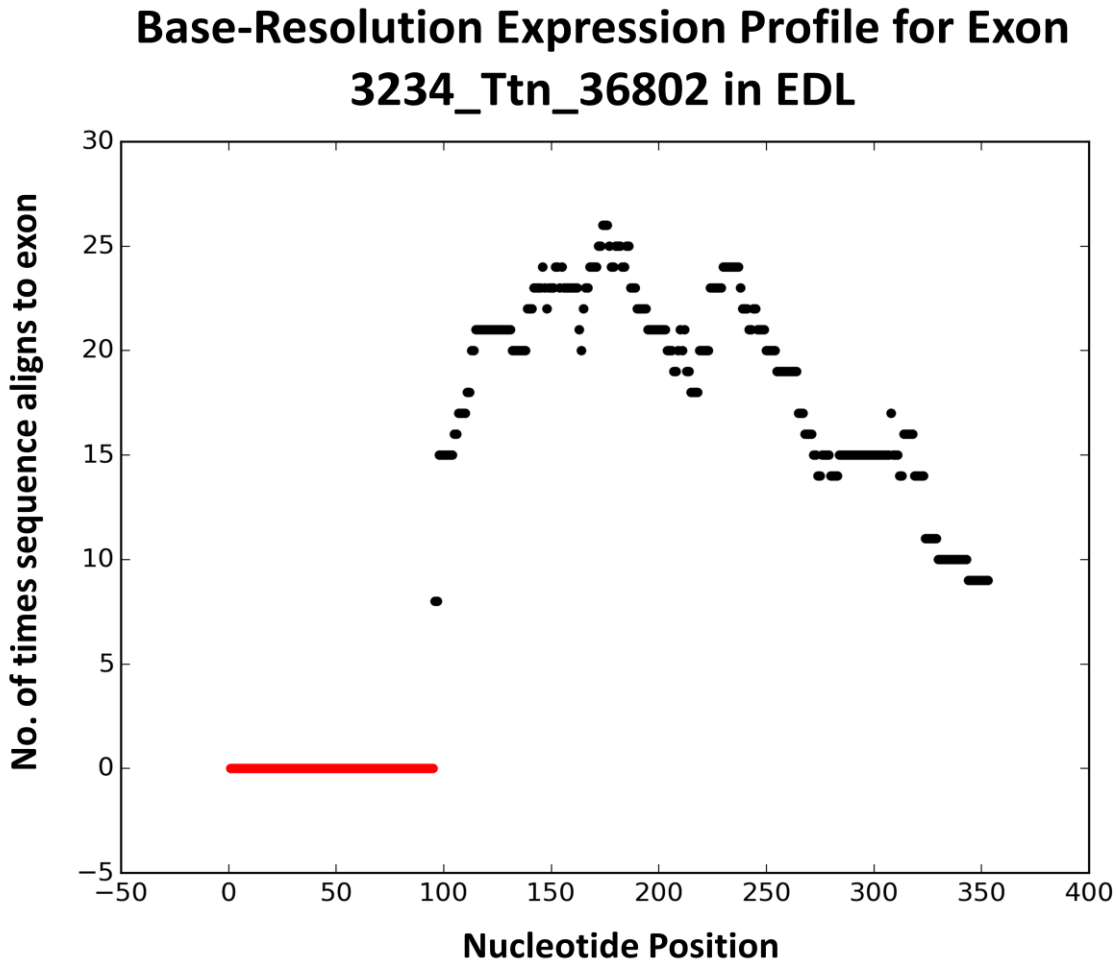


Figure 4.3: ***Base-Resolution Profile showing sequence coverage for exon 3234_Ttn_36802 for EDL titin. The black bands indicated sequence similarity whereas red band indicated that there was no read sequence similar to that region of the exon. Exon 3234_Ttn_36802 was an example of the exons that had partial coverage of their sequence in the transcriptome.***

4.2 Splicing Pattern in Soleus Titin

In addition to the 11 exons skipped in all the three tissues studied, specific exons were also excluded in the expression of titin in the soleus muscle. The total number of exons that showed full coverage of the hit sequences to the corresponding exons in the soleus titin was 309, indicating that those exons were included in the expression of titin whereas 13 exons were skipped (Figure 4.4). This was the highest number of exon inclusion in the expression of titin amongst the tissue samples. The exon deletion in the soleus titin occurred mainly in the PEVK domain. Exons 3234_Ttn_36831 and 3234_Ttn_36832 were the additional exons that were skipped in addition to the 11 mentioned earlier. These two exons were of the same sequence length, 84 nucleotides. Exons 3234_Ttn_36710, 3234_Ttn_36771, 3234_Ttn_36802, 3234_Ttn_36808 and 3234_Ttn_36829 were all skipped singly whereas exons 3234_Ttn_36821 through to 3234_Ttn_36826 and 3234_Ttn_36831 through to 3234_Ttn_36832 were excluded as groups (Figure 4.5). Being skipped singly implied that the exon was deleted alone whereas in a group implied that the exon was deleted together with other exons as a segment. The 13 exons were made up of 3954 nucleotides (Tables 4.1 and 4.2) that had a translation capacity of 1318 amino acids (Figure 4.10) which implied that the soleus titin was shortened by a deletion of 1318 amino acids. This was the least number of exons skipped amongst the three tissue samples.

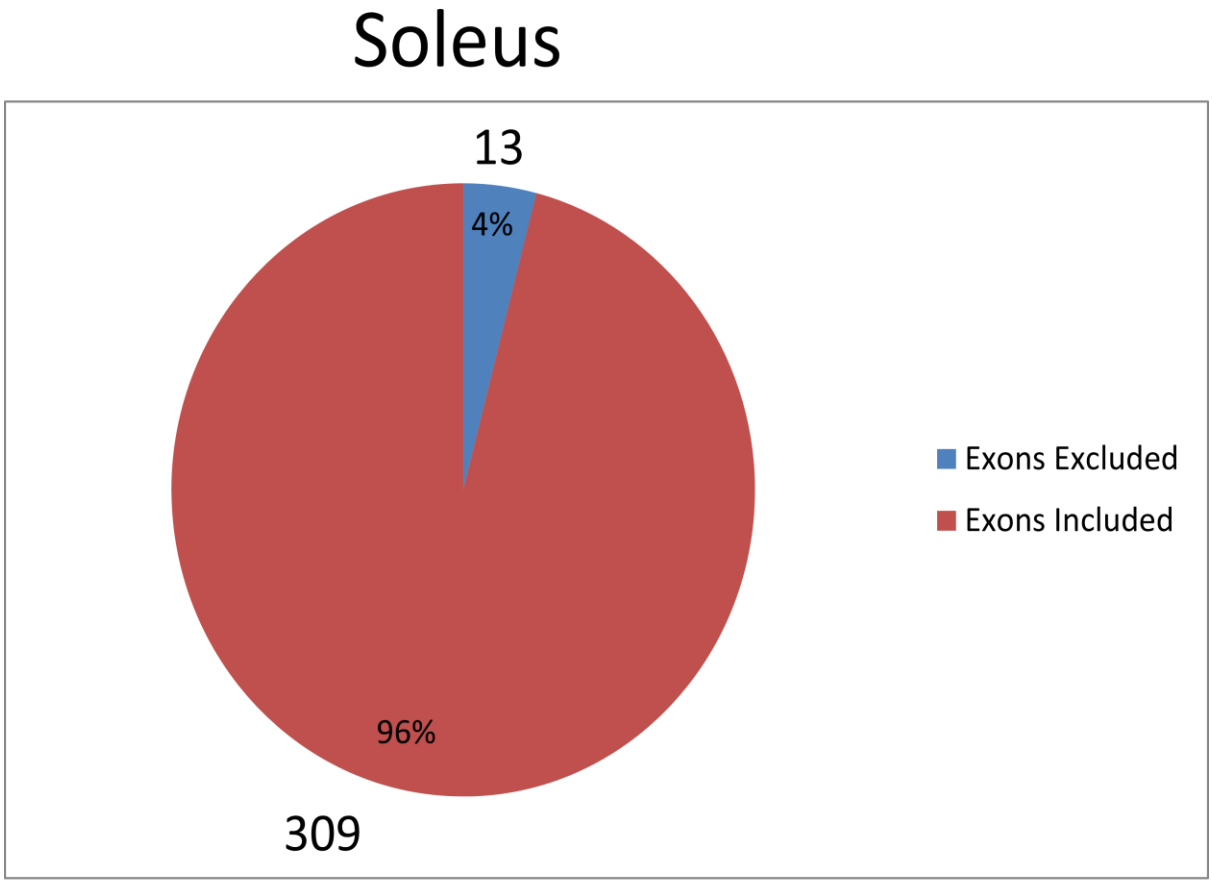


Figure 4.4: ***Ratio of Exon Deletion to Exon Inclusion in the Soleus Titin. The figure shows percentage of exons that were deleted in the expression of titin in the soleus. The number of exons that was skipped is 13, which represents 4% of the total exons, 322. This is shown in the blue pie of the chart. The red region represents the remaining 96%, 309, which was included in the expression.***

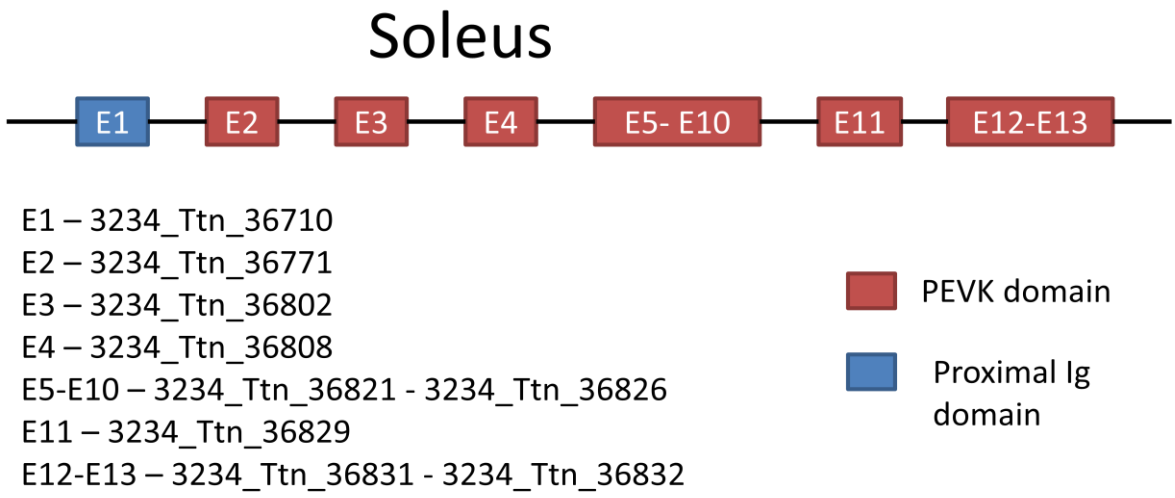


Figure 4.5: ***Exon Deletion Pattern in the Soleus Titin. The red represents the PEVK domain. The blue represents deletion of exon in the proximal tandem Ig domain. Each E-number represents a skipped exon. A range of numbers, such as E5-E10, represents a group of exons that were skipped as a segment. In all, 13 exons were skipped.***

4.3 Splicing Pattern in Psoas Titin

In the psoas, 20 exons were skipped in the expression of titin (Figure 4.6). The 20 exons that were skipped comprised 4518 nucleotides (Tables 4.1 and 4.2) which had a translation capacity of 1506 amino acid (Figure 4.10), implying that the psoas titin was shortened by a deletion of 1506 residues. The exon deletions in the psoas titin also occurred mainly in the PEVK domain as seen in the soleus titin. Exons 3234_Ttn_36710, 3234_Ttn_36771, 3234_Ttn_36802, 3234_Ttn_36808, 3234_Ttn_36817, 3234_Ttn_36834 and 3234_Ttn_36836 were skipped singly whereas 3234_Ttn_36819 - 3234_Ttn_36827 and 3234_Ttn_36829 - 3234_Ttn_36832 were skipped as groups (Figure 4.7). All the exons skipped in the soleus titin were also skipped in the psoas but the pattern of exclusion differed, for instance exon 3234_Ttn_36829 which was skipped in a group, 3234_Ttn_36829 - 3234_Ttn_36832, in psoas titin was skipped singly in soleus titin. Another exclusion of note observed in the psoas was the skipping of 3234_Ttn_36817, 3234_Ttn_36819, 3234_Ttn_36820, 3234_Ttn_36827, 3234_Ttn_36830, 3234_Ttn_36834 and 3234_Ttn_36836 which were not skipped in the soleus titin (Figures 4.5 and 4.7).

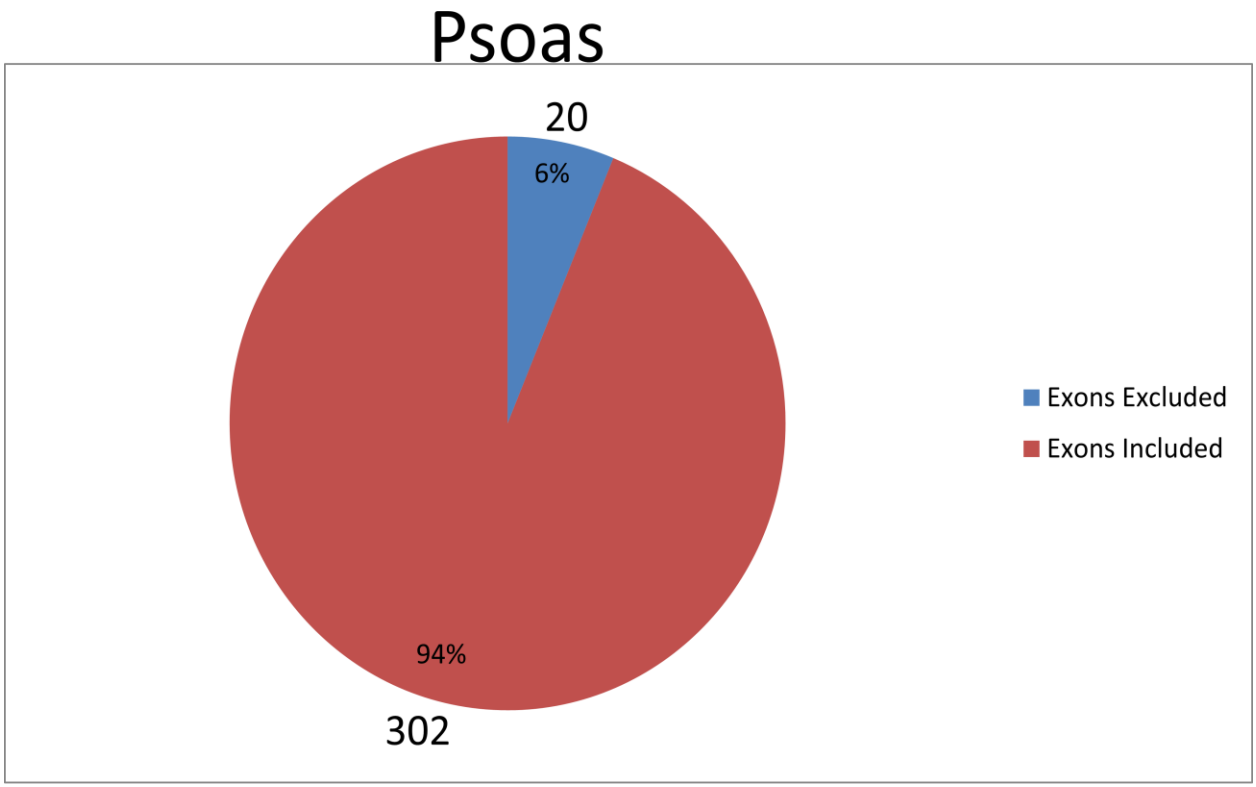


Figure 4.6: ***Ratio of Exon Deletion to Exon Inclusion in Psoas Titin. The pie chart shows percentage of exons that were skipped in the expression of titin in the psoas. The number of exons that was skipped is 20, which represents 6% of the total exons, 322. This is shown in the blue pie of the chart. The red region represents the remaining 94% included in the expression of the psoas titin.***

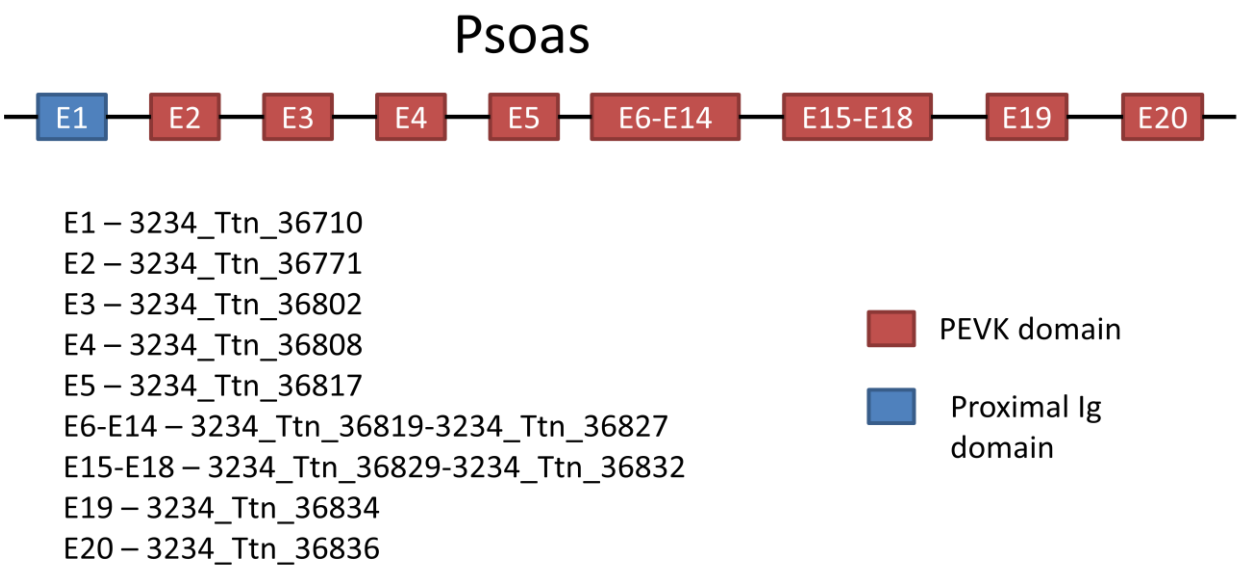


Figure 4.7: ***Exon Deletion Pattern in the Psoas Titin. The red boxes represent exons in the PEVK domain. The blue box represents exons in the proximal tandem Ig domain. Each number represents a deleted exon. A range of numbers, such as E6-E14, represents a group of exons that were skipped as a segment. The exon represented by each E code is also shown. In all, 20 exons were skipped. Seven of the 20 exons were skipped singly, that is alone whereas the remaining was skipped in two groups, which were E6-E14 and E15-E18.***

4.4 Splicing Pattern in EDL Titin

For the EDL, 14 exons were excluded, leaving 308 exons being involved in the expression of titin (Figure 4.8). The 14 exons skipped comprised a total of 4035 nucleotides (Tables 4.1 and 4.2) which had a translation capacity of 1345 amino acid, implying that the EDL titin was shortened by a deletion of 1345 residues (Figure 4.10). The deletions in the EDL titin also occurred mainly in the PEVK domain. The pattern of exclusion of the exons in the EDL titin was similar to that of the soleus titin except that the exons 3234_Ttn_36820 - 3234_Ttn_36826 and 3234_Ttn_36829 - 3234_Ttn_36831 were skipped as group exclusions whereas in the soleus titin it was 3234_Ttn_36821 - 3234_Ttn_36826 and 3234_Ttn_36831 - 3234_Ttn_36832 as groups and 3234_Ttn_36829 as single exon exclusion (Figures 4.5 and 4.9). The EDL titin differed more from the psoas titin as compared to the soleus titin. Several exons were skipped in the psoas titin which were not skipped in the EDL titin. The exons skipped were 3234_Ttn_36817, 3234_Ttn_819, 3234_Ttn_36827, 3234_Ttn_36832, 3234_Ttn_36834 and 3234_Ttn_36836 (Figures 4.7 and 4.9). All the exons skipped in both the EDL and soleus titins were also skipped in the psoas titin but the number of exons skipped in the psoas titin was 2% more than what was skipped in the EDL and soleus titins. The differential number of exons involved in the expression of titin in the EDL, psoas and soleus implied that titin in each of the skeletal muscles was a different isoform even though they were all expressed from a single gene. Also the translational capacities of the skipped exons in each of the isoforms indicated that psoas titin was the shortest in length with an amino acid deletion of 1506, followed by EDL titin with an amino acid deletion of 1345 and then soleus titin which was the longest amongst the three isoforms with an amino acid deletion of 1318

(Figure 4.10). The patterns of exclusion in the three tissue samples differed from each other and the exons excluded were of varying nucleotide lengths, implying that titin was expressed as different isoforms in the EDL, psoas and soleus of *Mus musculus* (Figure 4.10).

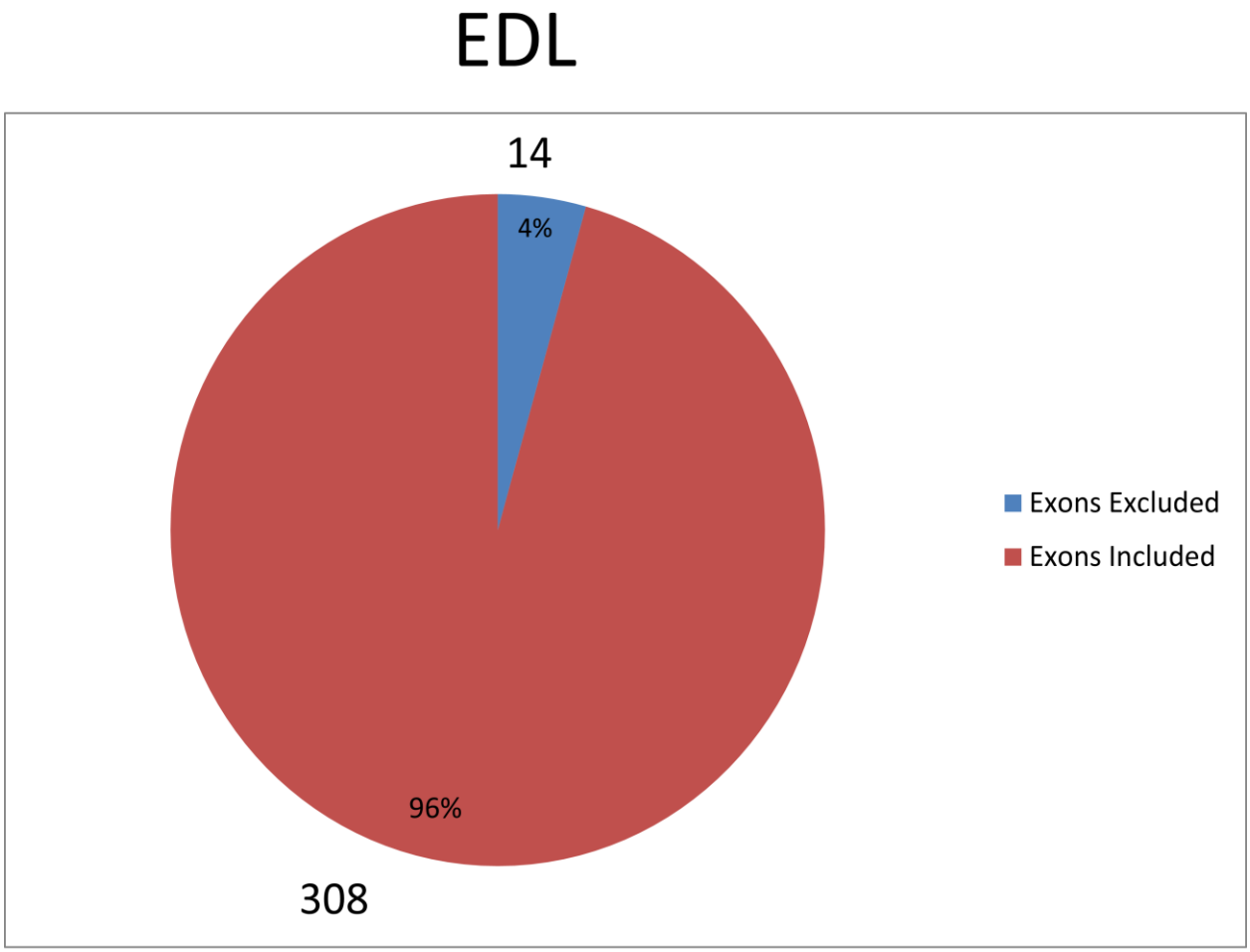


Figure 4.8: ***Ratio of Exon Deletion to Exon Inclusion in the EDL Titin. The figure shows percentage of exons that were skipped in the expression of titin in EDL. The number of exons that was skipped is 14, which represents 4% of the total exons, 322. This is shown in the blue pie of the chart. The red region represents the remaining 96%, 308, included in the expression of titin.***

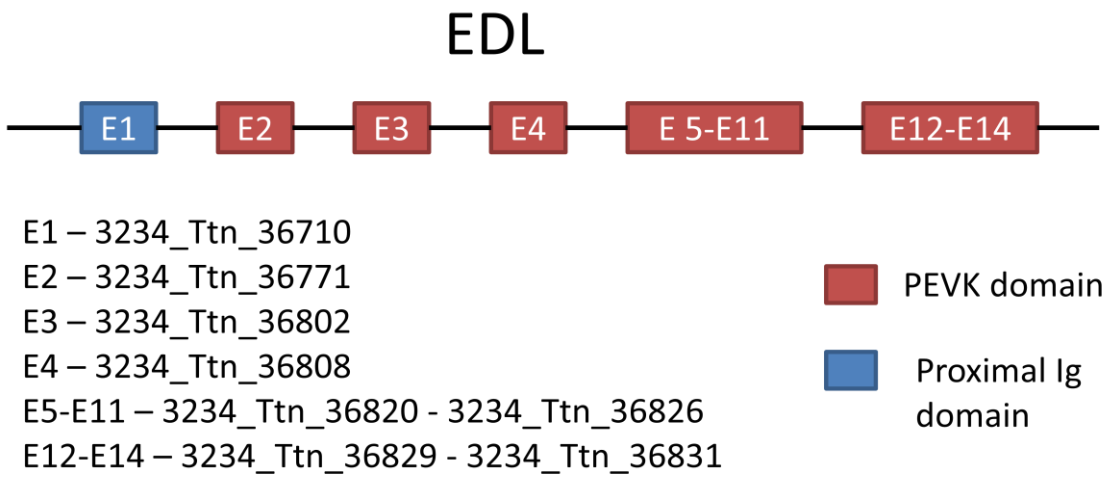


Figure 4.9: ***Exon Deletion Pattern in the EDL Titin. The red boxes represent exons that were involved in the expression of the PEVK domain. The blue box represents the proximal tandem Ig domain. Each E-number represents a skipped exon. A range of numbers, such as E5-E11, represents a group of exons that were skipped as a segment. In all, 14 exons were skipped. Four of the 14 exons were skipped singly, that is alone whereas the remaining was skipped in two groups, which were E5-E11 and E12-E14.***

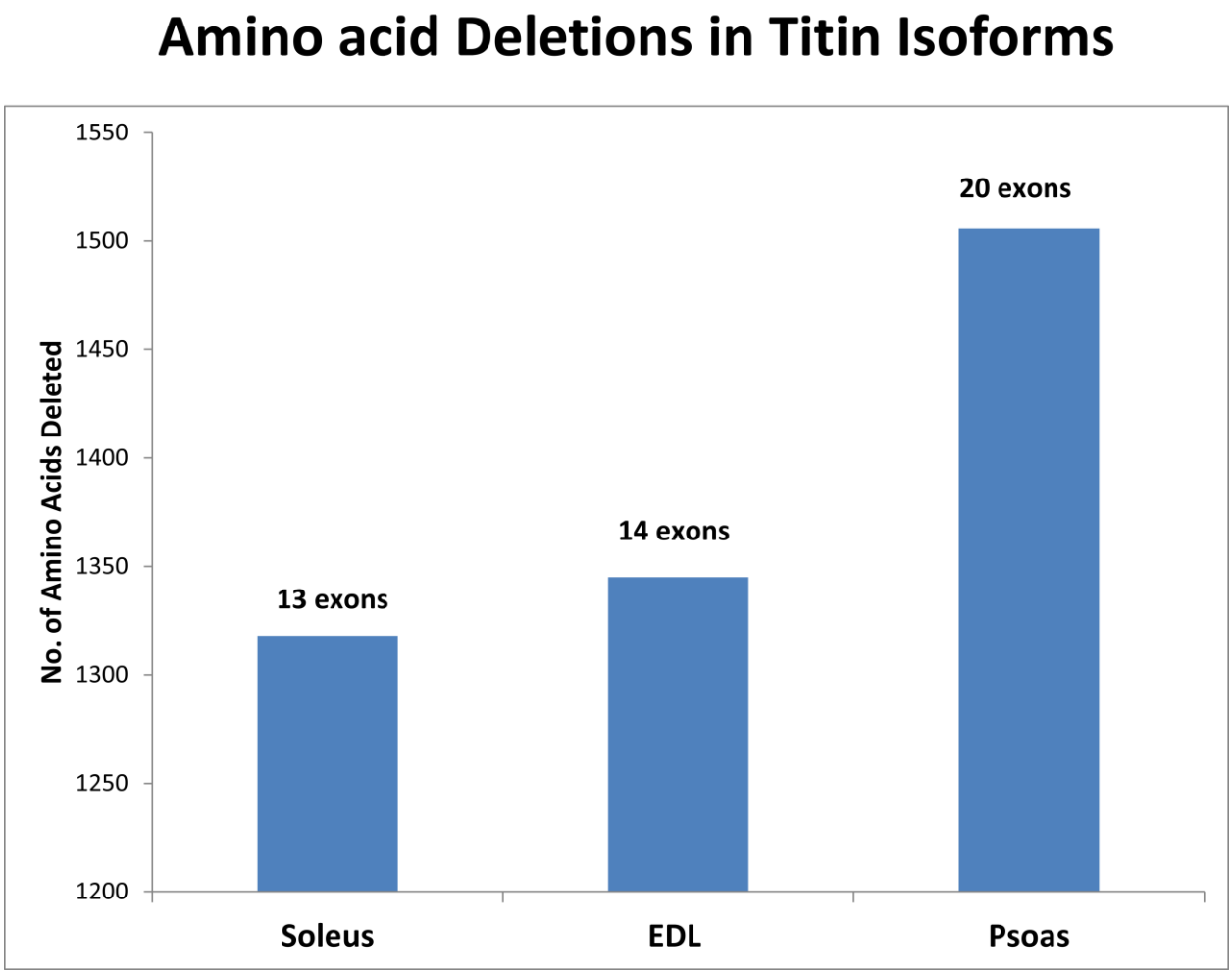


Figure 4.10: ***Number of Amino Acids Skipped in each Titin Isoform. This graph shows the translational capacity of the exons skipped in each of the titin isoforms. Skipping of the exons implied that the titin isoform was shortened by the corresponding amino acids that the exons coded for. In the soleus titin, 13 exons were skipped, indicating that 1318 amino acid residues were excluded in the expression of titin. In the EDL titin, 14 exons were skipped, representing an exclusion of 1345 amino acid residues. In the psoas titin, 20 exons were skipped, indicating an exclusion of 1506 amino acid residues.***

Table 4.2: Exons skipped in each muscle type and their corresponding nucleotide length as well as possible number of amino acid translations

SAMPLE	EXONS SKIPPED (3234_Ttn_no.)	NUCLEOTIDE SIZE	AMINO ACIDS
Soleus	36831, 36832	84, 84	56
Psoas	36817, 36819, 36820, 36827, 36830, 36831, 36832, 36834, 36836	78, 75, 81, 84, 84, 84, 84, 84, 78	295
EDL	36820, 36830, 36831	81, 84, 84	83

Each 3234_Ttn_no. represents a titin exon that was skipped and hence was absent in the transcriptome. The exon exclusions in this table are specific to each tissue sample. The sum of the exon exclusions in this table and the 11 exon exclusions (Table 4.1) that were common in all tissue samples gives the total exon exclusion in each tissue sample.

CHAPTER FIVE

DISCUSSION

The roles of titin in muscle elasticity and sarcomere assembly make it an indispensable muscle protein. Disease-causing mutations that lead to the development of muscular dystrophy have been associated with titin, yet little is known about titin splicing mechanisms in skeletal muscles. Therefore the aim of this study was to generate information on titin splicing in skeletal muscles of *Mus musculus*. The interest in mice stems from the fact that the titin orthologue of the mouse is similar to that of the human titin. Also, there are a good number of mutant mice that serve as good experimental models to study the significance of titin in the sarcomere.

Titin expression in the cardiac muscle of humans, dogs and mice has been associated with multiple splicing patterns to generate different isoforms based on the tissue type. Similarly, multiple titin isoforms have been identified in different skeletal muscles of rabbits (Freiburg *et al*., 2000). However, little is known about the expression and splicing patterns of titin isoforms in the skeletal muscles of *Mus musculus*. Three types of skeletal muscles; Soleus, EDL and Psoas, were investigated to elucidate the splicing pattern of titin in those tissues.

The findings of the study provide evidence that titin transcripts are processed by distinct splice routes in the Soleus, EDL and Psoas and the differential routes lead to the expression of distinct isoforms from the same gene. The soleus muscle expresses the largest titin isoform with only 1318 amino acid residues spliced out, followed by the EDL titin with 1345 amino acids spliced out and then the psoas titin with 1506 amino acids

spliced out. This is in line with literature where it was reported that titin from soleus of rabbits had a higher molecular weight, 3.7 MDa, relative to a smaller titin from psoas of rabbits, 3.35 MDa (Freiburg *et al*., 2000). In the same study by Freiburg and co-workers, they observed that the difference in the molecular weights of the soleus and psoas titin was due to a shorter proximal tandem Ig and PEVK segments in the psoas. However this current study suggests that the change in the PEVK segment is the cause of the differential lengths in the titin isoforms. This is because differential exon deletion was mainly observed in the PEVK segment of titin and the deletion observed in the tandem Ig region was observed in all three isoforms. The reason for observing shorter PEVK only and not shorter proximal tandem Ig might be due to the experimental procedure used in the study. The experimental design only focused on identifying the presence of titin exons from the transcriptome of different skeletal tissues. However, Ig domain is ubiquitous and encoded by many genes, for instance the Ig domain in humans is encoded by more than 750 genes. Hence any changes in the expression of the Ig domain in titin might be masked by the presence of Ig domains from other transcripts in the transcriptome. So even though there might be deletions in the titin Ig domain, the presence of other Ig domains might result in a false positive observation. Further work is therefore required to overcome any possibility of false positive observations by using an experimental procedure that would lead to the sequencing of only the titin transcript. This would ensure that the Ig domain sequenced was really from titin only and not from Ig domains from other proteins in the muscle. Such an approach would first involve a separation procedure such as gel electrophoresis followed by sequencing of the isolated titin.

Titin elasticity has been reported by Labeit and Kolmerer (1995) to be associated with both the PEVK and the Ig domains of the I-band region of the protein. However, from the present study, differential elasticity of skeletal muscles appears to result from exon skipping in the PEVK domain only. As stated earlier, no differential exon skipping was observed in the Ig-domain in the present study. Nevertheless this observation is subject to further investigation. Thus, per the current data, difference in elasticity is suggested to be due to differences in the sequences of the PEVK domain only. The present findings therefore support the views of Linke *et al.* (1996) and Gautel and Goulding (1996) who reported that the PEVK domain is the main determinant of titin extensibility in skeletal muscles. The work by Linke *et al.* demonstrated that at moderate to extreme stretch, the major extensible region was the PEVK domain and the unfolding of the domain correlated with a steady passive tension increase. On the other hand, they also demonstrated that the Ig domain lengthened during small stretch but the lengthening did not generate significant passive tension. Similarly, the work by Gautel and Goulding also demonstrated that the PEVK domain extended beyond 2.6 µm sarcomere length whereas the Ig domain only extended between 2 and 2.8 µm sarcomere lengths. The conclusion made from this current study is based on the differential expression of the PEVK domain of titin observed through RNA sequencing techniques. Linke *et al*. made conclusion based on data generated using antibody epitopes and immunofluorescent labeling and Gautel and Goulding used immunoelectron microscopy with sequence-assigned monoclonal antibodies in their investigation. The different experimental approaches used to conclude that the PEVK domain is the main determinant of titin elasticity makes such a conclusion more credible. In addition to titin elasticity, studies show novel functions of the PEVK region in the

sarcomere. An example was by Linke *et al.* (2002) where cardiac PEVK was shown to interact weakly with actin filaments and was also described as an important modulator of contractile properties. It is therefore reasonable to speculate that the differential expression of the PEVK domain in the different muscle types is tailored to enhance the functions of titin in the sarcomere based on the biological requirements of each tissue. The expression of the distinct PEVK domains with unique characteristics is therefore very critical to titin elasticity and sarcomere assembly in the skeletal muscles.

The PEVK region of titin in the soleus was found in this study to be the longest, followed by that of the EDL titin and then that of the psoas titin. This finding is consistent with work by Freiburg *et al.* (2000) who reported that the length of the PEVK region of titin in the soleus was longer than the PEVK region of titin in the psoas of rabbits. They also reported that the shorter length of the PEVK of titin in the psoas generated a higher passive tension relative to that of the longer PEVK in the soleus titin. It is therefore proposed that for a given sarcomere length, the relative passive tension generated by stretched PEVK domain of soleus titin from mice would be lower than that of the psoas titin. It would also be reasonable to propose a lower passive tension in the PEVK domain of EDL titin than that of the psoas titin due to the longer PEVK region of the EDL titin relative to the PEVK domain of the psoas titin. Once again, the differential elasticity requirement by different skeletal muscles based on their specific functions is demonstrated by the differential expression of the PEVK domain, the main contributor of titin elasticity, through splicing. The differential lengths of the PEVK domain observed in the EDL, psoas and soleus titins also support the claim by Labeit and Kolmerer (1995) that the PEVK domain assumes differential lengths depending on the tissue type.

The splicing events in each of the muscle tissues led to the deletion of particular titin exons. Based on the findings of the present study, the deletion of exons in the expression of *TTN* can be grouped into two: exon deletions that were common to all three muscle types and tissue-specific exon deletions. The specificity of exons deleted in *TTN* depending on tissue type forms the basis for titin diversity in different skeletal muscles. The tissue-specific exon deletions occurred mainly in the PEVK domain of the I-band of titin. This is consistent with similar observations by Freiburg *et al*. (2000) who reported that extensive exon shuffling in the I-band region generated titin isoform diversity in rabbit striated muscles. An example of the diversity observed in this study was the deletion of exons 3234_Ttn_36817, 3234_Ttn_36834 and 3234_Ttn_36836 from psoas titin only but not the EDL titin and the soleus titin. Another was the deletion of exon 3234_Ttn_36830 in the titins of the psoas and EDL only but not the soleus titin. The diversity generated through exon skipping events in the expression of *TTN* suggests that tissue-specific titin might be required for specific functions in muscle elasticity, protein-protein interactions, protein–nucleic interactions as well as novel roles. In terms of unique muscle elasticity for instance, the three skeletal muscles used in this study contribute to movement, however each muscle type plays a different role in aid of locomotion of the mouse. The EDL aids in the bending of the foot, the psoas aids in the lifting of the upper limb towards the body and the soleus, in addition to aiding in walking and running, serves as a skeletal muscle pump that moves venous blood from the periphery to the heart. All the specific roles played by the skeletal muscles are linked by the fact that each role of the skeletal muscles involves stretching of the muscle cells. But the degree to which the muscle cells would stretch in response to the biological function of the muscle type would differ and hence unique

elasticity is required to ensure efficient functioning of the sarcomere. In terms of specific protein-protein interactions, the skeletal muscle isoform of titin has been shown to bind F-actin differentially. The poly-E fragments of the PEVK domain of titin for instance have been shown to have a stronger apparent actin binding (Nagy *et al*., 2004). Therefore skeletal muscles whose biological functions depend on stronger binding of titin to F-actin would require the expression of a titin isoform with higher poly-E fragments content. It is therefore proposed that splicing allows functional adaptation of titin in skeletal muscles through rearrangement of the titin exons. Titin diversity generated from the same gene through alternative splicing also helps explains the complexity associated with the higher number of protein isoforms in a cell relative to the fewer number of genes that can express those proteins.

Alternative splicing is largely influenced by the activation or synthesis of splicing factors. The splicing factor 2/alternative splicing factor (SF2/ASF) has been reported to control splicing enhancers to cause the skipping of exon 11 in the expression of a constitutively active Ron (Ghigna *et al*., 2005). Another splicing factor, RNA-binding motif protein (RBM20), has been reported to mediate exon skipping by repressing the splicing of some regions of *TTN* and the action of this splicing factor differs by tissue type (Li *et al*., 2013). In a similar manner, the findings of the present study points to the involvement of common splicing factors and tissue specific splicing factors in *TTN* splicing even though investigating the precise splicing factors involved in the exon skipping was beyond the scope of the study. Therefore it is speculated that some of the splicing patterns observed were controlled by the common splicing factors where some specific exons were excluded and their exclusion occurred along a similar pattern in all three tissue types. For

example, exons 3234_Ttn_36710, 3234_Ttn_36802 and 3234_Ttn_36808 were all observed to be excluded in all the isoforms and the exclusions occurred singly, that is not in a group as represented in Figures 4.5, 4.7 and 4.9. Such common splicing factors are possibly ubiquitous and most likely recognize the same splice sites. Hence cumulatively, such observations indicate (1) that the splice sites involved are independently recognized and (2) the splice factors involved in the skipping of those exons act independently of the tissue in which the titin is expressed. However, further work is required to uncover and characterize such factors. This would enhance detailed knowledge on titin splicing mechanisms. Contrary to the suggested common regulatory factors causing similar exons to be excluded along a common route, some exons were excluded in only one or two of the isoforms and their exclusion did not follow a common pattern as seen earlier. Such splicing events are suggested to be regulated by more flexible regulatory factors. For example, exons 3234_Ttn_36821 - 3234_Ttn_36826 were excluded in the soleus titin whereas 3234_Ttn_36820 - 3234_Ttn_36826 and 3234_Ttn_36819 - 3234_Ttn_36827 were rather excluded in the EDL titin and the Psoas titin respectively. Another example of note was the exclusion of exon 3234_Ttn_36829 as a single spliced exon in the soleus titin whereas this exon was spliced together with other exons in a group in the EDL titin, 3234_Ttn_36829 – 3234_Ttn_36831 and for the Psoas titin, 3234_Ttn_36829 – 3234_Ttn_36832. The flexible regulatory factors are possibly unique to each skeletal muscle tissue since they lead to the generation of isoforms with distinct structural sequences. The differential structural sequences of the isoforms suggest that there might be differences in the biological functions of the isoforms as well, indicating how important the tissue specific factors are in the expression of tissue specific titin. Therefore, it can be

suggested that the expression of the flexible splicing factors are based on the biological functions of the tissue in which they are expressed.

According to Nakano *et al.*, (2012) a mutation in the *Srrm*4 gene led to the skipping of exons which resulted in impaired balance, deafness and defective hair-cell development in mutant mice. Similarly, single-nucleotide polymorphisms (SNPs) which may seem to be silent during translation could even have tremendous impact on alternative splicing. For example, silent SNPs in *Multidrug Resistance 1* (*MDR1*) gene have been shown to affect the *in vivo* folding and functioning of P-glycoprotein (Kimchi-Sarfaty *et al.*, 2007). Such observations indicate that mutations influence alternative splicing events. Therefore the exon skipping events observed in this study is most likely subject to change under the influence of a mutation. It would therefore be interesting to carry out further work to investigate exon skipping events in a mutant form such as the Muscular Dystrophy with Myositis (MDM) mouse which has a mutation in its titin gene. This MDM mutant is the mouse form of the human tibial muscular dystrophy. The knowledge of the exon skipping events of titin obtained from the present study, paralleled with investigations in the mutant form would broaden understanding of how titin-associated mutations result in muscular dystrophy. The knowledge obtained can then be employed in the treatment of muscular dystrophy in humans especially the human tibial muscular dystrophy.

The ultimate goal of the project is to understand the mechanisms involved in the alternative splicing events associated with titin expression in different skeletal muscles. The present study has identified exons that are excluded and those that are included in the expression of titin in the soleus, psoas and EDL of mice. The study sets a platform on which the splicing mechanism in mice skeletal muscles can be elucidated. More work is

therefore required to understand the detailed mechanisms involved in the regulation of titin splicing events.

CHAPTER SIX

CONCLUSION

This study has provided evidence that titin is expressed in different isoforms in different skeletal muscles of *Mus musculus*. The outcome is consistent with reports obtained by studies on other species that say that titin exists in different isoforms through alternative splicing. The number of exons deleted in each isoform has been identified in this study, however further work is required in studying titin expression in order to fully elucidate the mechanism by which titin is expressed in different isoforms. The findings of the present study have identified the exons that are involved and those that are skipped in the expression of titin. This is a step in elucidating the mechanistic details associated with titin alternative splicing in skeletal muscles. To fully achieve the ultimate goal of elucidating the splicing mechanism, understanding of the roles of splicing factors as well as interactions of titin with other molecules is very important and hence further research is proposed along such lines. From the study, splicing factors that regulate the deletion of common exons and tissue-specific exons in different muscle types have been suggested to be involved in the splicing mechanism. Knowledge of these factors would give details on how the exon skipping in titin expression occurs. Information generated in the study can be useful in the quest of finding treatments for muscular dystrophy in humans. This would have a tremendous positive impact on public health.

The PEVK domain was identified to undergo numerous differential splicing events relative to the N2A and Ig domains. This reinforces the view that the PEVK domain is the main spring element that is involved in the generation of significant passive force in titin.

Differences in exon skipping events in the PEVK domain generate tissue-specific titin isoforms. And the tissue-specific titin is suggested to be expressed depending on the biological functions of the tissue. However, the roles of the different isoforms in the functions of the specific tissues remain understudied. Therefore, it is recommended that further work should be carried out to investigate novel roles of the diverse isoforms of titin in different skeletal tissues.

REFERENCES

Arbanas, J., Starcevic Klasan, G., Nikolic, M., Jerkovic, R., Miljanovic, I. and Malnar, D. (2009). "Fibre type composition of the human psoas major muscle with regard to the level of its origin". *Journal of Anatomy* **215 (6)**, 636–641.

Bang, M. L., Centner, T., Fornoff, F., Geach, A. J., Gotthardt, M., McNabb, M., Witt, C. C., Labeit, D., Gregorio, C. C., Granzier, H. and Labeit, S. (2001). The complete gene sequence of titin, expression of an unusual approximately 700-kDa titin isoform, and its interaction with obscurin identify a novel Z-line to I-band linking system. *Circulation Research* **89,** 1065–1072.

Baron, M. and Norman, D. G. (1991). I.D Campbell.Protein modules.*Trends in Biochemical Science* **16**, 13–17.

Beadle, G. W. and Tatum, E. L. (1941). Genetic Control of Biochemical Reactions in Neurospora. *Proceedings of the National Academy of Science USA* **27 (11)**, 499–506.

Bennett, P. M., Hodkin, T. E. and Hawkins, C. (1997). Evidence that the tandem Ig domains near the end of the muscle thick filament form an inelastic part of the I-band titin. *Journal of Structural Biology* **120 (1)**, 93-104.

Berg, J. M., Tymoczko, J. L. and Stryer, L. (2002). The Immunoglobulin Fold Consists of a Beta-Sandwich Framework with Hypervariable Loops. *Biochemistry*, **5th edition**., Section 33.2, New York: W H Freeman.

Berget, S. M, Moore, C. and Sharp P. A. (1977). "Spliced segments at 5' terminus of adenovirus 2 late messenger-RNA". *Proceedings of the National Academy of Science USA* **74 (8)**, 3171–3175.

Berry, S. D., Howard, R. D. and Akers, R. M. (2003). Mammary localization and abundance of laminin, fibronectin, and collagen IV proteins in prepubertal heifers. *Journal of Dairy Science* **86 (9)**, 2864-2874.

Black, D. L. (2003). Mechanisms of alternative pre-messenger RNA splicing. *Annual Review of Biochemistry* **72**, 291-336.

Bork, P. and Doolittle, R. F. (1992). Proposed acquisition of an animal protein domain by bacteria. *Proceedings of the National Academy of Science USA* **89**, 8990–8994.

Bork, P., Holm, L. and Sander, C. (1994). "The immunoglobulin fold. Structural classification, sequence patterns and common core". *Journal of Molecular Bioliogy.* **242 (4)**, 309–320.

Botta, G., Piccinetti, A., Giontella, M. and Mancini, S. (2001). Strengthening of venous pump activity of the sural tricipital in orthopaedics and traumatology by means of a new equipment for vascular exercise. *Giornale Italiano di Ortopedia e Traumatologia* **27**, 84–88.

Breitbart, R. E., Andreadis, A. and Nadal-Ginard, B. (1987). Alternative Splicing: A ubiquitous mechanism for the generation of multiple protein isoforms from single genes. *Annual of Review Biochemistry.* **56**, 467-495.

Brown R. H. Jr. and Mendell, J. R. (2005). Harrison's Principles of Internal Medicine. p. 2527. doi:10.1036/0071402357.

Brümmendorf, T. and Rathjen, F. G. (1995). "Cell adhesion molecules 1: immunoglobulin superfamily". *Protein Profile* **2 (9)**, 963–1108.

Caporaso, J. G. *et al*. (2010). "QIIME Allows Analysis of High-Throughput Community Sequencing Data." *Nature methods* **7 (5)**, 335–336.

Centner, T., Fougerousse, F., Freiburg, A., Witt, C., Beckmann, J. S., Granzier, H., Trombitas, K., Gregorio, C. C. and Labeit, S. (2000). Molecular tools for the study of titin's differential expression. *Elastic Filament Cell* **481**, *35-49.*

Chauveau, C., Rowell, J. and Ferreiro, A. (2014). A Rising Titan: TTN Review and Mutation Update. *Human Mutation* **35 (9)**, 1046-1059.

Cloonan, N., Forrest A. R., Kolle, G., Gardiner B. B., Faulkner G. J., Brown M. K., Taylor D. F., Steptoe A. L., Wani, S., Bethel G. *et al*. (2008). Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nature Methods* **5 (7)**, 613-619.

Fackenthal, J. and Godley, L. (2008). "Aberrant RNA splicing and its functional consequences in cancer cells". *Disease Models & Mechanisms* **1 (1)**, 37–42.

Forrest, S. M., Smith, T. J., Cross, G. S., Read, A. P., Thomas, N. S. T, Mountford, R. C., Harper, P. S., Geirsson, R. T. and Davies K. E. (1987). Effective strategy for prenatal prediction of Duchenne and Becker muscular dystrophies. *Lancet* **2**, 1294-1297.

Freiburg, A., Trombitas, K., Hell, W., Cazorla, O., Fougerousse, F., Centner, T., Kolmerer, B., Witt, C., Beckmann, J. S., Gregorio, C. C., Granzier, H. and Labeit, S. (2000). Series of Exon-Skipping Events in the Elastic Spring Region of Titin as the Structural Basis for Myofibrillar Elastic Diversity. *Circulation Research* **86**, 1114-1121.

Funatsu, T., Higuchi, H and Ishiwata, S. (1990). Elastic Filaments in Skeletal Muscle Revealed by Selective Removal of Thin Filaments with Plasma Gelsolin. *Journal of Cell Biology* **110**, 53-62.

Funatsu, T., Kono, E., Higuchi, H., Kimura, S., Ishiwata, S., Yoshioka, T., Maruyama, K. and Tsukita, S. (1993). Elastic filaments in situ in cardiac muscle: deepetch replica analysis in combination with selective removal of actin and myosin filaments. *Journal of Cell Biology* **120**, 711–724.

Furst, D. O., Osborn, M., Nave, R., Weber, K. (1988). The organization of titin filaments in the half-sarcomere revealed by monoclonal antibodies in immunoelectron microscopy: a map of ten nonrepetitive epitopes starting at the Z line extends close to the M line. *Journal of Cell Biology* **106**, 1563–1572.

Gautel, M. and Goulding, G. (1996). A molecular map of titin/connectin elasticity reveals two different mechanisms acting in series. *Federation of European Biochemical Societies Letters* **385**, 11–14.

Ghigna, C., Giordano, S., Shen, H., Benvenuto, F., Castiglioni, F., Comoglio, P. M., Green, M. R., Riva, S. and Biamonti, G. (2005). Cell Motility Is Controlled by SF2/ASF through Alternative Splicing of the *Ron* Protooncogene. *Molecular Cell* **20 (6)**, 881 – 890.

Granzier, H. L. and Labeit, S. (2004). The Giant Protein Titin: A Major Player in Myocardial Mechanics, Signaling, and Disease. *Circulation Research* **94**, 284-295.

Granzier, H. L., Hutchinson, L. K., Tonino, P., Methawasin, M., Li, F. W., Slater, R. E., Bull, M. M., Saripalli, C., Pappas, C. T., Gregorio, C. C. and Smith III, J. E. (2014). Deleting titin's I-band/A-band junction reveals critical roles for titin in biomechanical sensing and cardiac function. *Proceedings of the National Academy of Science USA* **111 (40)**, 14589–14594.

Granzier, H. L., Radke, M. H., Peng, J., Westermann, D., Nelson, L. O., Rost, K., King, N. M., Yu, P. Q., Tscho¨pe, C., McNabb, M., Larson, D. F., Labeit, S. and Gotthardt, M. (2009). Truncation of Titin's Elastic PEVK Region Leads to Cardiomyopathy with Diastolic Dysfunction. *Circulation Research* **105**, 557-564.

Graveley, B. R. (2001). Alternative Splicing: Increasing diversity in the proteomic world. *Trends in Genetics* **17 (2)**, 100-107.

Greaser, M. (2001). Identification of new repeating motifs in titin. *Proteins* **43,** 145–149.

Greaser, M. L., Berri, M., Warren, C. M. and Mozdziak, P. E. (2002). Species variations in cDNA sequence and exon splicing patterns in the extensible I-band

region of cardiac titin: relation to passive tension. *Journal of Muscle Research and Cell Motility* **23 (5-6)**, 473-482.

Gregorio, C. C., Granzier, H., Sorimachi, H. and Labeit, S. (1999). Muscle assembly: a titanic achievement. *Current Opinion in Cell Biology* **11**, 18–25.

Gregorio, C. C., Perry, C. N. and Mcelhinny, A. S. (2005). Functional properties of the titin/connectin-associated proteins, the muscle-specific RING finger proteins (MURFs), in striated muscle. *Journal of Muscle Research and Cell Motility* **26**, 389–400.

Gutierrez-Cruz, G., van Heerden, A. H. and Wang, K. (2001). Modular motif, structural folds and affinity profiles of the PEVK segment of human fetal skeletal muscle titin. *Journal of Biological Chemistry* **276,** 7442–7449.

Hayashi, C., Ono, Y., Doi, N., Kitamura, F., Tagami, M., Mineki, R., Arai, T., Taguchi, H., Yanagida, M., Hirner, S., Labeit, D., Labeit, S. and Sorimachi, H. (2008). Multiple Molecular Interactions Implicate the Connectin/Titin N2A Region as a Modulating Scaffold for p94/Calpain 3 Activity in Skeletal Muscle. *Journal of Biological Chemistry* **283 (21)**, 14801-14814.

He, C., Zhou, F., Zuo, Z., Cheng, H. and Zhou, R. (2009). "A Global View of Cancer-Specific Transcript Variants by Subtractive Transcriptome-Wide Analysis". *Public Library of Science* **4 (3),** e4732. doi: 10.1371/journal.pone.0004732.

Kent, W. J. (2002). "BLAT—The BLAST-Like Alignment Tool." *Genome Research* **12 (4)**, 656–664.

Kimchi-Sarfaty, C., Oh, J. M., Kim, I. W., Sauna, Z. E., Calcagno, A. M., AMbudkar, S. V. and Gottesman M. M. (2007). A 'Silent' Polymorphism in the *MDR1* Gene Changes Substrate Specificity. *Science* **315**, 525.

Koenig, M., Beggs, A. H., Moyer, M., Scherpf, Q. S., Heindrich, T. K., Bettecken, T., Meng, T. G., Mullerj, T. C. R., Lindlof, M. *et al*. (1989). The Molecular Basis for Duchenne versus Becker Muscular Dystrophy: Correlation of Severity with Type of Deletion. *American Journal of Human Genetics* **45**, 498-506.

Koenig, M., Hoffman, E. P., Bertelson, C. J., Monaco, A. P., Feener, C., Kunkel, L. M. (1987). Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. *Cell* **50**, 509-517.

Kulke, M., Fujita-Becker, S., Rostkova, E., Neagoe, C., Labeit, D., Manstein, D. J., Gautel, M. and Linke, W. A. (2001). Interaction between PEVK-titin and actin filaments: origin of a viscous force component in cardiac myofibrils. *Circulation Research* **89**, 874–881.

Labeit, D., Watanabe, K., Witt, C., Fujita, H., Wu, Y., Lahmers, S., Funck, T., Labeit, S. and Granzier, H. (2003). Calcium-dependent molecular spring elements in the giant protein titin. *Proceedings of the National Academy of Science U S A.* **100**, 13716 –13721.

Labeit, S. and Kolmerer, B. (1995). Titins: Giant proteins in charge of muscle ultrastructure and elasticity. *Science* **270**, 293 -296.

Labeit, S., Granzier, H., Cazorla, O., Freiburg, A., Helmes, M., Centner, T., McNabb, M., Wu, Y., Trombitás, K. (2000). Stiffness Differential Expression of Cardiac Titin Isoforms and Modulation of Cellular. *Circulation Research* **86**, 59-67.

Labeit, S., Kolmerer, B. and Linke, W. A. (1997). "The giant protein titin: emerging roles in physiology and pathophysiology," *Circulation Research* **80 (2)**, 290–294.

LeWinter, M. M., Wu, Y., Labeit, S. and Granzier, H. (2007). Cardiac titin: Structure, functions and role in disease. *Clinica Chimica Acta* **375**, 1–9.

Li, H., Oberhauser, A. F., Redick, S. D., Carrion-Vazquez, M., Erickson, H. P. and Fernandez, J. M. (2001). Multiple conformations of PEVK proteins detected by single-molecule techniques. *Proceedings of the National Academy of Science USA* **98,** 10682–10686.

Li, S., Guo, W., Dewey, C. N. and Greaser, M. L. (2013). Rbm20 regulates titin alternative splicing as a splicing repressor. *Nucleic Acids Research* **41 (4)**, 2659-2672.

Lim, K. H., Ferraris, L., Filloux, M. E., Raphael, B. J. and Fairbrother, W. G. (2011). "Using positional distribution to identify splicing elements and predict pre-mRNA processing defects in human genes". *Proceedings of the National Academy of Science USA* **108 (27)**, 11093–11098.

Linke, W. A. and Granzier, H. (1998). A spring tale: new facts on titin elasticity. *Biophysical Journal* **75**, 2613 -2614.

Linke, W. A., Ivemeyer, M., Labeit, S., Hinssen, H., Ruegg, J. C. and Gautel, M. (1997). Actin-titin interaction in cardiac myofibrils: Probing a physiological role. *Biophysical Journal* **73**, 905–919.

Linke, W. A., Ivemeyer, M., Mundel, P., Stockmeier, M. R. and Kolmerer, B. (1998). Nature of PEVK–titin elasticity in skeletal muscle. *Proceedings of the National Academy of Science USA* **95**, 8052–8057.

Linke, W. A., Ivemeyer, M., Olivieri, N., Kolmerer, B., Ruegg, J. C. and Labeit, S. (1996). Towards a molecular understanding of the elasticity of titin. *Journal of Molecular Biology* **261**, 62 -71.

Linke, W. A., Kulke, M., Li, H., Fujita-Becker, S., Neagoe, C., Manstein, D. J., Gautel, M. and Fernandez, J. M. (2002). PEVK Domain of Titin: An Entropic Spring with Actin-Binding Properties. *Journal of Structural Biology* **137**, 194–205.

Linke, W. A., Rudy, D. E., Centner, T., Gautel, M., Witt, C., Labeit, S. and Gregorio, C. C. (1999). I-band titin in cardiac muscle is a three element molecular spring and is critical for maintaining thin filament structure. *Journal of Cell Biology* **146**, 631 -644.

Lister, R., O'Malley, R. C., Tonti-Filippini, J., Gregory, B. D., Berry, C. C., Millar, A. H. and Ecker, J. R. (2008). Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell* **133**, 523-536.

Maruyama, K. (1994). Connectin, an elastic protein of striated muscle. *Biophysical Chemistry* **50**, 73-85.

Maruyama, K., Matsubara, S., Natori, R., Nonomura, Y. and Kimura, S. (1977). "Connectin, an elastic protein of muscle. Characterization and Function". *Journal of Biochemistry* **82 (2)**, 317–327.

Miller, M. K., Bang, M. L., Witt, C. C., Labeit, D., Trombitas, C., Watanabe, K., Granzier, H., McElhinny, A. S., Gregorio, C. C. and Labeit, S. (2003). The muscle ankyrin repeat proteins: CARP, ankrd2/Arpp and DARP as a family of titin filament-based stress response molecules. *Journal of Molecular Biology* **333**, 951–964.

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* **5**, 621-628.

Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M. and Snyder, M. (2008). The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **320**, 1344-1349.

Nagy, A., Cacciafesta, L., Grama, L., Kengyel, A., Malnasi-Csizmadia, A. and Kellermayer, M. S. Z. (2004). Differential actin binding along the PEVK domain of skeletal muscle titin. *Journal of Cell Science* **117 (24)**, 5781-5789.

Nakano, Y., Jahan, I., Bonde, G., Sun, X., Hildebrand, M. S., Engelhardt, J. F., Smith, R. J. H., Cornell, R. A., Fritzsch, B. and Banfi, B. (2012). A Mutation in the Srrm4 Gene Causes Alternative Splicing Defects and Deafness in the Bronx Waltzer Mouse. *Public Library of Science Genetics* **8 (10)**, e1002966.

Nilsen, T. W. and Graveley, B. R. (2010). Expansion of the eukaryotic proteome by alternative splicing. *Nature*. **463 (7280)**, 457–463.

Nishikawa, K. C., Monroy, J. A., Uyeno, T. E., Yeo, S. H., Pai, D. K. and Lindstedt, S. L. (2012). Is titin a 'winding filament'? A new twist on muscle contraction. *Proceedings of the Royal Society B* **279**, 981–990.

Nunes, M. T., Bianco, A. C., Migala, A., Agostini, B. and Hasselbach, W. (1985). "Thyroxine induced transformation in sarcoplasmic reticulum of rabbit soleus and psoas muscles". *Zeitschrift fur Naturforschung. Section C: Biosciences* **40 (9–10)**, 726–734.

Obermann, W. M. J., Gautel, M., Steiner, F. P., van der Ven, F. M., Weber, K. and Furst, D. O. (1996). "The structure of the sarcomeric M band: localization of defined domains of myomesin, M protein, and the 250-kD carboxy-terminal region of titin by immunoelectron microscopy," *Journal of Cell Biology* **134 (6)**, 1441–1453.

Olson, S., Blanchette, M., Park, J., Savva, Y., Yeo, G. W., Yeakley, J. M., Rio, D. C. and Graveley, B. R. (2007). A regulator of *Dscam* mutually exclusive splicing fidelity. *Nature Structural & Molecular Biology* **14**, 1134 – 1140.

Ozsolak, F. and Milos, P. M. (2011). RNA sequencing: advances, challenges and opportunities. *Nature Reviews Genetics* **12**, 87-98.

Petersen, T. E., Thøgersen, H. C., Skorstengaard, K., Vibe-Pedersen, K., Sahl, P., Sottrup-Jensen, L. and Magnusson, S. (1983). Partial primary structure of bovine plasma fibronectin: three types of internal homology. *Proceedings of the National Academy of Sciences of the USA* **80 (1)**, 137-141.

Pienaar, E., Theron, M., Nelson, M. and Viljoen, H. J. (2006). A quantitative model of error accumulation during pcr amplification. *Computational Biology and Chemistry* **30 (2)**, 102–111.

Platzer, W. (2004). Color Atlas and Textbook of Human Anatomy: Locomotor system. North Yorkshire: Thieme Medical Publishers.

Punetha, J. and Hoffman, E. P. (2013). Short read (next-generation) sequencing: a tutorial with cardiomyopathy diagnostics as an exemplar. *Circulation: Cardiovascular Genetics* **6 (4)**, 427-434.

Rajasekar, A. *et al.* (2010). "iRODS Primer: Integrated Rule-Oriented Data System.". Synthesis Lectures on Information Concepts, Retrieval, and Services **2 (1)**, 1–143.

Schena, M., Shalon, D., Davis, R. W. and Brown, O. P. (1995). Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray. *Science* **270**, 467-470.

Skotheim, R. I. and Nees, M. (2007). "Alternative splicing in cancer: noise, functional, or systematic?" *The International Journal Of Biochemistry & Cell Biology* **39 (7-8)**, 1432–1449.

Szabo, D. (2014). Transcriptomic biomarkers in safety and risk assessment of chemicals. Biomarkers in Toxicology (ed. R. Gupta), pp. 1033–1038. Ramesh Gupta: Oxford: Academic Press.

Tiffany, H. N. (2014). Structural characterization of the unique 'insertion sequence' region of titin's N2A domain. Northern Arizona University. ProQuest 1563886. (http://gradworks.umi.com/15/63/1563886.html)

Udd, B., Vihola, A., Sarparanta, J., Richard, I. and Hackman, P. (2005). "Titinopathies and extension of the M-line mutation phenotype beyond distal myopathy and LGMD2J". *Neurology* **64 (4)**, 636–42.

Wang, Z., Gerstein, M. and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics* **10**, 57-63.

Westermann D., Rutschow, S., Jager, S., Linderer, A., Anker, S., Riad, A., Unger, T., Schultheiss, H. P., Pauschinger, M. and Tschope, C. (2007). Contributions of inflammation and cardiac matrix metalloproteinase activity to cardiac failure in diabetic cardiomyopathy: the role of angiotensin type 1 receptor antagonism. *Diabetes* **56**, 641-646.

Wilhelm, B. T., Marguerat, S., Watt, S., Schubert, F., Wood, V., Goodhead, I., Penkett, C. J., Rogers, J. and Bähler, J. (2008). Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature* **453**, 1239-1243.

Wu, Y., Bell, S. P., Trombitas, K., Witt, C. C., Labeit, S., LeWinter, M. M. and Granzier, H. (2002). Changes in Titin Isoform Expression in Pacing-Induced Cardiac Failure Give Rise to Increased Passive Muscle Stiffness. *Circulation* **106**, 1384-1389.

Yamasaki, R., Berri, M., Wu, Y., Trombitas, K., McNabb, M., Kellermayer, M. S., Witt, C., Labeit D., Labeit, S., Greaser, M. and Granzier, H. (2001). Titin-actin interaction in mouse myocardium: passive tension modulation and its regulation by calcium/S100A1. *Biophysical Journal* **81**, 2297–2313.

National Human Genome Research Institute - http://www.genome.gov/Images/EdKit/bio2j_large.gif

http://bitesizebio.com/13546/sequencing-by-synthesis-explaining-the-illumina-sequencing-technology/

http://www.giga.ulg.ac.be/upload/docs/image/jpeg/2009-04/illumina4.jpg

http://en.wikipedia.org/wiki/Extensor_digitorum_longus_muscle

http://en.wikipedia.org/wiki/Soleus_muscle

http://en.wikipedia.org/wiki/Psoas_major_muscle

http://www.imm.fm.ul.pt/exonmine/index.php

APPENDIX

A.1 PREPARATION OF RNase-FREE WATER USING DIETHYLPYROCARBONATE (DEPC)

An aliquot of 0.1 ml of DEPC was added to 100 ml H_2O and shaken vigorously to bring the DEPC into solution. The solution was incubated overnight at 37 °C and autoclaved for 15 min to remove any traces of DEPC.

A.2 PREPARATION OF 1× TAE BUFFER

A mass of 242 g of tris base was weighed into 750 ml of RNase-water. An aliquot of 57.1 ml glacial acetic was added. An aliquot of 100 ml of 500 mM EDTA (pH 8) was then added and the final volume was brought to 1000 ml. The 1× TAE was then prepared from the 50× TAE by diluting 1× TAE with 49 ml of RNase-free water.

A.3 PREPARATION OF 1% AGAROSE GEL

One gram of agarose was weighed into 100ml 1× TAE into an Erlenmeyer flask. The solution was microwaved on high for 2 min till the agarose was fully dissolved in the 1× TAE. The solution was then allowed to cool down, but not too cool to solidify the gel. An aliquot of 3 µl of SYBR Safe was then added to the solution. The liquefied gel was then poured into a gel cassette and a comb was inserted into the notches on the cassette. The solidified gel was then used to run samples in a gel apparatus with 1× TAE.