Age-related structural and compositional variations of mouse skeletal muscle

Huijuan Wang

NANYANG TECHNOLOGICAL UNIVERSITY

AGE-RELATED STRUCTURAL AND COMPOSITIONAL VARIATIONS OF MOUSE SKELETAL MUSCLE

WANG HUIJUAN

SCHOOL OF CHEMICAL AND BIOMEDICAL ENGINEERING

2012

AGE-RELATED STRUCTURAL AND COMPOSITIONAL VARIATIONS OF MOUSE SKELETAL MUSCLE

WANG HUIJUAN

SCHOOL OF CHEMICAL AND BIOMEDICAL ENGINEERING

A thesis submitted to Nanyang Technological University

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy

2012

Acknowledgement

First and foremost I would like to express my sincere gratitude to my supervisor, Assistant Professor Lee Kijoon, for his initial encouragement when I was just transferred into his group, and further comprehensive guidance that made my Ph.D. experiences a better one. My postgraduate study is funded by Nanyang Technological University (NTU) research scholarship, thus I would like to take this opportunity to thank NTU, School of Chemical and Biomedical Engineering (SCBE) and all its excellent faculties, researchers and officers.

My heartfelt thanks go to my co-supervisor Dr. Listrat Anne (INRA, France) for her invaluable guidance and advice throughout the Ph.D. project. It is my great honor to have work with her. Besides the fancy techniques and scientific ideas, the enthusiasm for science that she endowed me would be my greatest treasure during the past three years. This appreciation also goes to Dr. Béchet Daniel. This Ph.D. project and thesis would not have been possible without his support in a number of ways, including his proficiency and overall sparking ideas.

I especially want to thank my ex-supervisor Dr. Goh Kheng Lim (Newcastle University, Singapore). Throughout my postgraduate life, his inspiring words

and perpetual energy have motivated me. Moreover, he has played an essential role in building and developing this project.

I wish to extend my sincere thanks to all the family like members of our lab: Ms. Dong Jing, Mr. Ho Jun Hui and Mr. Bi Ren Zhe, as well as Ms. Wee Yi Ting, the final-year-project student attached to me and has contributed to the study of image analysis. I am indebted to my many colleagues in INRA (France): Mr. Meunier for his extensive knowledge in image analysis and data mining; Ms. Dubost for her tremendous support and friendship; Ms. Coppellia for her effective computer skills; Mrs. Jurie, Mr. Montel, Mrs. Gentes, Mr. Chadeyron, Mrs Coudy-Gandilhon, Mrs Barboiron and Mr. Delavaud for their help on cellular and molecular biology experiments; Mrs. Picard, Mr. Micol and Mrs. Cassar-Malek for their enlightenment and support. As well, I thank Dr. Lethias (IBCP, France) for her critical opinions on histological study, and her generousness and help on the study with Confocal microscope.

This project was co-funded by Merlion Singapore (Joint Franco-Singaporean Research Project, °5.03.07). I would like to show my gratitude to French Embassy in Singapore and all its kind and patient officers: Mr. Mynard, Ms. Gerard and Ms. Tay. The administrative officers of INRA (France) and CROUS (France) have done much to assist this Merlion Singapore project, I am thus grateful to all of them.

Apart from academic affairs, I would like to thank my dear friends Qi Ya Ning, Su Kai, Yu Ting, Bai Jing, Wang Chun Ming, Shen Yi Qiang, Ji Qing, Zhuang Tao, Wei Li and Gong Cheng, My life in Singapore would be dull and unbearable without you. And Ren Li Bo, Chen Xiao Bin, Si Tu Yin, Sun Yuan, Guillemin Nicolas, Guo Xin Mei and Gao Xian Fu for their help and friendship in France.

Last but not least, I would like to dedicate this thesis to my parents for their endless love in the past twenty-seven years; and to my husband for his accompany and support. Without them, this thesis would never have been written.

Table of content

ACKNOWLEDGEMENT ..... **I**

TABLE OF CONTENT ..... **V**

LISTS OF FIGURE ..... **IX**

LISTS OF TABLE ..... **XIII**

ABBREVIATIONS ..... **XIV**

SUMMARY ..... **XVII**

CHAPTER 1 INTRODUCTION **1**

1.1 INTRODUCTION: AGING, A GENERAL THEORY 1

1.2 PRELIMINARIES: SKELETAL MUSCLE, FROM MYOFIBRE TO CONNECTIVE TISSUE 3

1.2.1 Myofibre *3*

1.2.1.1 Description 3

1.2.1.2 Typology 5

1.2.1.3 Molecular composition 8

1.2.1.4 Sarcolemma and costamere 11

1.2.2 Connective tissue *12*

1.2.2.1 Stromal cells 13

1.2.2.2 Extracellular matrix 13

1.3 MODEL AND OBJECTIVES 23

1.3.1 Model *23*

1.3.1.1 C57BL/6 mouse: popular model in aging research 23

1.3.1.2 *Gastrocnemius* muscle anatomy 24

1.3.1.3 Sample preparation 26

1.3.2 Objective *29*

CHAPTER 2 SIGNIFICANCE OF SARCOPENIA, ATROPHY, APOPTOSIS AND FIBROSIS DURING AGING **33**

2-1 INTRODUCTION 33

2-2 METHODS 35

2.2.1 Myofibre morphology *36*

2.2.2 Myofibre metabolic activity *37*

2.2.3 ECM structure *38*

2.2.4 ECM components localization and co-localization *39*

2.2.5 Nuclei distribution *44*

2.2.6. Collagen and crosslinking content *49*

2.2.7 Endorepellin *50*

2.2.8 Image analysis *52*

2.2.9 Statistical analyses *59*

2.3 RESULTS 60

2.3.1 Body composition *60*

2.3.2 Myofibre morphology and metabolism *61*

2.3.2.1 Myofibre atrophy 61

2.3.2.2 Myofibre morphology 62

2.3.3 ECM *64*

2.3.3.1 ECM structure 64

2.3.3.2 ECM component and co-localization 66

2.3.4 Muscular nuclei *69*

2.3.4.1 Nuclei distribution 69

2.3.4.2 Apoptosis 73

2.4 DISCUSSION 77

2.4.1 Aging is associated with profound alteration in myofibre morphology *78*

2.4.2 Aging is associated with less effective myonuclei *79*

2.4.3 Aging is associated with impaired satellite cells *80*

2.4.4 Apoptosis in stromal cells regulated by endorepellin *81*

2.4.5 Fibrosis in skeletal muscle *82*

SUMMARY 83

CHAPTER 3 DATA MINING STRATEGY FOR INVESTIGATING STRUCTURE AND COMPONENT OF MOUSE SKELETAL MUSCLE FROM MATURITY TO AGING 85

3.1 INTRODUCTION 85

3.1.1 PRINCIPAL COMPONENT ANALYSIS 85

3.1.2 Clustering 87

3.2 METHODS 91

3.2.1 Dataset 91

3.2.1 Data pre-processing 92

3.2.3 Clustering 94

3.3 RESULTS 97

3.3.1 Dimensionality reduction 97

3.3.2 Variable selection 99

3.3.3 Clustering 100

3.3.4 Clustering validation 104

3.4 DISCUSSION 104

SUMMARY 109

CHAPTER 4 EFFECT OF AGE ON THE TRANSCRIPTIONAL PROFILING OF SKELETAL MUSCLE DURING AGING 111

4.1 INTRODUCTION 111

4.1.1 Microarray 111

4.1.2 Altered expression of ECM components and its contribution to sarcopenia during aging 113

4.2 METHODS 114

4.2.1 RNA extraction 114

4.2.2 Exon array data collection 116

4.2.3 Transcriptomic data analysis *116*

4.3 RESULTS 119

4.3.1 Quality control *119*

4.3.2 Changes in genomic transcription in mouse GM at 2-, 11- and 25M-old *120*

4.3.4 Transcriptomic changes in extracellular space during aging *136*

4.4 DISCUSSION 140

SUMMARY 143

CHAPTER 5 CONCLUSION AND FUTURE WORK 145

5.1 CONCLUSION 145

5.2 FUTURE OUTLOOK 148

REFERENCE LIST 150

PUBLICATION LIST 163

LISTS OF FIGURE

Figure 1.1 schematic figure of skeletal muscle structure 4

Figure 1.2 Myofibre with component proteins 8

Figure 1.3 Molecular model of A band, I band and M line 10

Figure 1.4 Molecular model of costamere 11

Figure 1.5 Schematic diagrams of stromal cells in skeletal muscle CT..... 13

Figure 1.6 Schematic diagrams of collagen structure 14

Figure 1.7 Schematic diagrams of PG structure..... 17

Figure 1.8 (A) SLRP in skeletal muscle (B) Schematic model of perlecan protein core..... 19

Figure 1.9 Schematic representations of the interactions of ECM components in skeletal muscle 21

Figure 1.10 C57BL/6 male mouse 23

Figure 1.11 Mouse calf muscle anatomy 24

Figure 1.12 Morphology of GM in rat (longitudinal section)..... 25

Figure 1.13 GM dissection..... 27

Figure 2.1 Azorubin staining..... 36

Figure 2.2 Cox staining 37

Figure 2.3 Sirius red staining.. 38

Figure 2.4 Schematic diagram of the indirect method for carrying out IHC. 39

Figure 2.5 Flow-chart illustrating the general procedure for carrying out single immunolabeling..... 40

Figure 2.6 Flow-chart illustrating the general procedure for carrying out double immunolabeling..... 40

Figure 2.7 Flow-chart of the general procedure for detecting apoptotic nuclei. 46

Figure 2.8 Flow-chart of the competitive enzyme immunoassay procedure 49

Figure 2.9 Competitive enzyme immunoassay to measure pyridinium crosslinking 50

Figure 2.10 Endorepellin was detected by SDS-PAGE coupled with Western blotting 50

Figure 2.11 Optical microscopy for examining the structure of muscle 52

Figure 2.12 Quantifying the structure of myofibres from sirius red stained micrographs 54

Figure 2.13 Processing of images obtained from Hoechst staining. 57

Figure 2.14 Processing of image obtained from laminin 2α staining 58

Figure 2.15 Graphs of the (A) body weight of the C57BL/6 mouse and (B) GM weight versus age 60

Figure 2.16 Myofibre CSA (μm^2) versus age 61

Figure 2.17 Myofibre morphology 62

Figure 2.18 Myofibre metabolism 63

Figure 2.19 Quantifying ECM microstructure 64

Figure 2.20 Age-related variations of fibrillar collagen content parameterized by (A) OH-Pro content (μg / mg wt muscle) and (B) Pyridinolin crosslinking (nmol /mmol collagen) 65

Figure 2.21 Cross-section of mouse GM in 11M or 25M old age with different ECM components stained with specific antibodies 66

Figure 2.22 Pixel intensity of ECM components in endomysium of mouse GM in 2, 11, 22 and 25M old age 67

Figure 2.23 Co-localization of ECM components in mouse GM in 11M-old ... 68

Figure 2.24 Expression of endorepellin fragments 68

Figure 2.25 Age-related variation of myonuclei 69

Figure 2.26 Age-related variations of SCs 70

Figure 2.27 Age-related variations of capillary endothelial cells 71

Figure 2.28 Apoptosis 73

Figure 2.29 Apoptosis in capillary endothelial cells. 75

Figure 2.30 Apoptosis in myofibre and SCs 76

Figure 2.31 Schematic representation of ECM structure in (A) adult and (B) aged mice 82

Figure 3.1 *K*-means clustering approach. 88

Figure 3.2 Schematic representation of the distance between clusters by (A) single link; (B) complete link; (C) average link and (D) McQuitty's method. .. 89

Figure 3.3 Variance explained in PCs 97

Figure 3.4 Correlation coefficient between PCs (PC1 to PC6) and the categorical variable (age) 97

Figure 3.5 PCA result (individual data) 98

Figure 3.6 PCA result (variables data) 99

Figure 3.7 *K*-means clustering result 100

Figure 3.8 AHC results: five criteria were used here: (A) Ward's, (B) single, (C) complete, (D) average and (E) McQuitty's 102

Figure 3.9 Distinct grouping by performing Ward's method 103

Figure 3.10 Flowchart of the PCA-clustering strategy 105

Figure 3.11 Distinct grouping by performing Ward's method 107

Figure 4.1 Comparison of expression array 112

Figure 4.2 Alternative Splicing 113

Figure 4.3 RNA extraction. 115

Figure 4.4 Quality controls on sample 119

Figure 4.5 Altered transcripts between 11M and 25M 121

Figure 4.6 (A) Location and (B) type of the 407 molecules corresponding to the altered transcripts 122

Figure 4.7 Ingenuity Pathway Analysis (IPA) of biofunctions on mice at 11M and 25M 125

Figure 4.8 Ingenuity Pathway Analysis (IPA) of canonical pathways 126

Figure 4.9 Gene network for connective tissue disorder 140

LISTS OF TABLE

Table 1.1 Characteristics of different myofibres..... 7

Table 1.2 Overview of known collagen types in skeletal muscle. 16

Table 1.3 Proteoglycans identified in skeletal muscle 18

Table 1.4 Interactions between major ECM components in skeletal muscle 22

Table 1.5 Age equivalencies between C57BL/6 mice and humans 24

Table 1.6 Fifty-four C57BL/6 male mice served in the study. 26

Table 1.7 List of methods utilized in this study 29

Table 2.1 Single immunolabeling 42

Table 2.2 Double immunolabeling..... 43

Table 2.3 Immunostaining of muscular nuclei..... 45

Table 2.4 Immunostaining of apoptotic nuclei 48

Table 3.1 Twenty-eight characteristics of skeletal muscle structure and component 92

Table 4.1 IPA biofunction..... 127

Table 4.2 IPA canonical pathway 135

Table 4.3 IPA biological network 137

ABBREVIATIONS

ADAMTS	a disintegrin and metalloproteinase with thrombospondin Motifs
AHC	agglomerative hierarchical clustering
B-H FDR	Benjamini-Hochberg false discovery rate
BL	basal lamina
BMP-1	bone morphogenetic protein-1
C4S	chondroitin-4-sulfate
C6S	chondroitin-6-sulfate
CD31	platelet endothelial cell adhesion molecule-1
Col-I	type I collagen
Col-III	type III collagen
Col-IV	type IV collagen
Col-VI	type VI collagen
Col-XII	type XII collagen
Col-XIV	type XIV collagen
cox	cytochrome c oxidase
CS	chondroitin sulfate
CSA	cross-sectional area
CT	connective tissue
DAB	3-3’ diaminobenzidine
DPD	deoxypridinoline
DS	dermatan sulfate
ECM	extracellular matrix
FACIT	fibril associated collagens with interrupted triple helices
GAGs	glycosaminoglycans

GM	*gastrocnemius* muscle
GPs	glycoproteins
HS	heparan sulfate
HSD	honestly significant differences
IHC	immunohistochemistry
IPA	ingenuity pathway analysis
KS	keratin sulfate
M	month(s)
MND	myonuclear domain
MHC	myosin heavy chain
Mins	minutes
MM	mismatch
MMPs	matrix metalloproteinases
MTJ	myotendinous junction
MURF-1	muscle ring-finger protein-1
MyBP-C	myosin binding protein C
OD	optical density
OH-pro	hydroxyproline
p94	muscle-specific calpain homologue 3
Pax7	anti-paired box protein 7
PCA	principal component analysis
PCs	principal components
PGs	proteoglycans
PLAT	tissue plasminogen activator
PM	perfect match
POLECE	procollagen C-proteinase enhancer

PVDF	polyvinylidene fluoride
PYD	trifunctional pyridinium crosslinking
RL	reticular lamina
RMA	robust multiarray average
SCs	satellite cells
sec	second(s)
SDS	sodium dodecyl sulfate
SDS-PAGE	Sodium Dodecyl Sulfate PolyArcylamide Gel Electrophoresis
SLRP	small leucine rich PGs
SM	*soleus* muscle
TNC	tenascin C
TNX	tenascin X
TUNEL	terminal deoxynucleotidyl transferase dUTP nick end labeling
wk	week(s)
yr	years

SUMMARY

Skeletal muscles undergo loss of mass and strength during aging, known as sarcopenia. Sarcopenia impairs skeletal muscle function and is characterized by a series of structural and compositional variations. In this project, I first verified that a reduced myofibre volume, i.e. atrophy, was associated with modifications in myofibre morphology. I indicated that the quantity of myonuclei (per myofibre) maintained during aging. I also reported an increased myonuclear centralization, which suggested that in the old muscle, myofibres were still capable of being regenerated by satellite cells. Although it seems well known that myonuclei yielded from a balance between their regeneration and their apoptosis, I revealed that myonuclear apoptosis was a rare event even in the old muscle. I further investigated the distribution of muscle apoptosis and noticed that most apoptosis occurred in capillary endothelial cells. Such capillary apoptosis may trigger important molecular events, i.e. cleavage of endorepellin fragments, which have been reported to stimulate fibrosis.

Altered myofibre morphology and area were associated with the variations in ECM structure. The mechanism(s) of this process, known as fibrosis, is not yet elucidated. I found that age-dependent fibrosis was characterized by increased ECM area and thickening and abundant ECM fragments. I also found that fibrosis was associated with increased collagen content and crosslinking. I assumed that the accretion of ECM might be due to an altered balance between ECM synthesis and degradation. The down-

regulation of a majority of peptidases and proteinases encoding genes could provide support to explain such a reduced ECM degradation. However, collagens, the major ECM constituents, likely exhibited reduced synthesis rate because a majority of collagen genes were down-regulated during aging. To interpret the accumulation of ECM, I therefore assumed that there should be a stronger reduction in collagen degradation than in collagen synthesis. To verify this hypothesis, further investigations are required.

My study provided a thorough understanding of the association between altered muscle myofibres and altered ECM structures. Moreover, I noticed the necessity to carry out an integrative analysis on the skeletal muscle characteristics mentioned above. To best of my knowledge, no such attempts have been applied on mouse skeletal muscle. I developed a data mining strategy directing to the assembling of data collected from histological / immunehistochemical studies. I found that the apoptotic nuclei and myofibre morphology significantly correlated to overall aging. On the contrary, the ECM structure did not present such correlations.

CHAPTER 1 INTRODUCTION

1.1 Introduction: aging, a general theory

There has been and will continue to be a dramatic shift in the composition of population with regards to age. The proportion of elderly has increased steadily over the past centuries; and this growth will be accelerated in the next several decades. In the developed regions of the world, population of people older than 60 years old has been growing at 3.8% per year from 1988. By 2000, elderly people accounted for almost 20% and they will reach 33% of the total population by 2050. In developing regions, the ratio of elderly people is 8% at present but will rise to 20% by 2050 (data from Food and Agriculture Organization of the United Nations). This growth is known as global aging, which is attributed to the lengthening of life expectancy and to declining of fertility rate. Global aging further results in an irreversible social change on both public health and economy. For example, impact of aging on medical expenses will depend upon the future development of morbidity[1] [1]. A better understanding of the biological aging process, investing in basic and social research, and promoting "healthy aging" ("successful aging") to reduce the inter-generation tension would be possible ways to find lasting solutions to reduce the socioeconomic impact of aging population. However, a comprehensive understanding of aging is difficult because the aging process

[1] Morbidity: number of patient or case of disease in a given population at a given time (in one year is called incidence; or at a time is called prevalence).

accompanies an increased risk of disease [2]. Indeed, aging may be considered as a degenerative process that causes progressive deterioration and loss of physiological function and ultimately death. Currently, the biological origin of aging phenomenon remains largely obscure. It is attributed to intrinsic mechanisms, as explained by three key theories namely the physiological theory, apoptosis and the evolutionary theory. The physiological theory explains that aging is due to the poisoning of the body accompanied by a progressive accumulation of metabolites [3]. The apoptosis theory addresses how aging leads to the increased number of programmed cell death [4]. The evolutionary theory is connected with the genetic explanation of aging as well as effects arising from extrinsic causes that are clearly within the environment [5].

In general, aging of different organs and tissues is accompanied by a progressive loss of their ability to fulfill their biological functions [6, 7]. Diverse variations, i.e. tissue atrophy (reduced volume) and the increase in connective tissue gradually appear in most organs [8]. This thesis is concerned with **skeletal muscle**, which represents almost 40% of body proteins. Skeletal muscle is required for body locomotion and is the major reservoir of body proteins, and consequently of amino acids which are mobilized under pathologic conditions. Muscle is also essential for maintaining body temperature, and helps to maintain bone homeostasis. During aging, there is significant reduction of muscle protein mass. This disrupts the balance between protein synthesis/proteolysis and/or regeneration/cell death [9, 10]. Muscle wasting reduces the capacity of reaction and restoration of a good health in the

elderly. Therefore, a holistic understanding of the mechanisms leading to the muscle wasting during aging is crucial.

The muscle tissue forms, via its **connective tissue** (CT), a continuum with the tendons, which in turn connects muscle to bone. The changes undergone by the muscle during aging affect not only muscle cells but also the CT in which muscle cells are surrounded.

In the following sections of this chapter, I will lay out the preliminary information concerning the skeletal muscle, addressing the different levels, i.e. molecular, cellular and tissue levels (section 1.2), and finally, addresses the justifications for the animal model used in this study and also explain the objectives (section 1.3).

1.2 Preliminaries: skeletal muscle, from myofibre to connective tissue

1.2.1 Myofibre

The skeletal muscle represents approximately 40% of body weight [11]. Skeletal muscle consists of several types of tissue: muscle, connective, nervous, vascular and adipose tissue.

1.2.1.1 Description

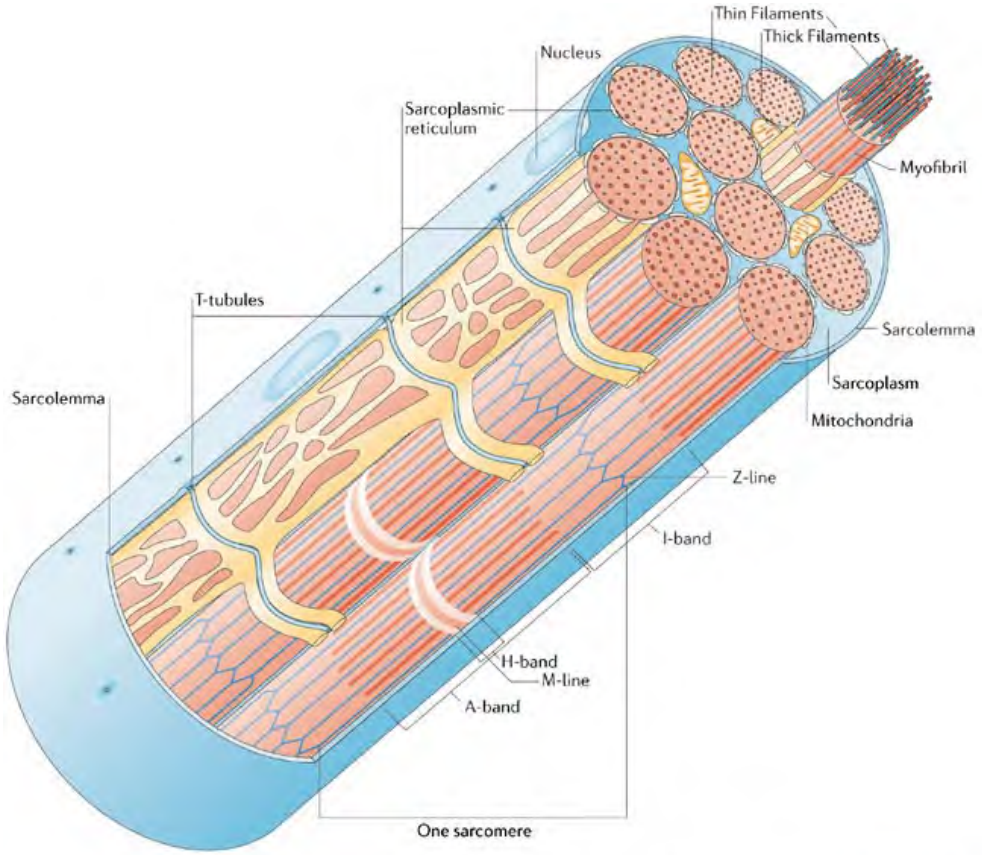


Figure 1.1 schematic figure of skeletal muscle structure, mainly myofibre and its surrounding ECM, belly splitting into various component parts, copyright 2006 by Nature Publishing Group

Basic organization of **skeletal muscle** structure is well documented in many test books. Figure 1.1 shows a schematic representation of the skeletal muscle. In particular, skeletal muscle connects to **bones** via **tendon;** the muscle-tendon junction is known as **myotendinous junction** (MTJ). Skeletal muscle comprises several types of muscle cells (**myofibres**) embedded within the **CT**. There are three separate levels of organization in the muscle CT: (1) the **epimysium**, which is the outer sheath of muscles, (2) the **perimysium**, which binds muscle fibre bundles (fascicle) together, and (3) the **endomysium**, which surrounds each individual myofibre and ensures continuity between the perimysium and the plasma membrane (**sarcolemma)** of myofibres. Endomysium consists of a reticular layer (0.2-1 microns) called the **reticular lamina** (RL) which is external as opposed to the **basal lamina** (BL), a thin basement membrane (about 50 nm) which is directly linked to the sarcolemma [12]. Between sarcolemma and BL, it is possible to observe myogenic cells, i.e. the **satellite cells** (SCs).

1.2.1.2 Typology

Myofibre is characterized by the metabolic activity and the speed of contraction. According to the metabolic pathway, myofibres are glycolytic or oxidative. The speed of contraction depends on myosin heavy chain that uses ATP energy for contraction. Several **myosin heavy chain** (MHC) isoforms have been identified based on their apparent molecular masses as indicated by migration rate (I > IIa > IIx) or antibody specificity. Those MHC isoforms are

used to determine the contractile types with specific antibodies. The above characteristics result in three main fibre types, as shown in Table 1.1.

Table 1.1 Characteristics of different myofibres, These characteristics fall under the following categories, twitch properties, metabolic properties, metabolic energy, physiology and morphology [13].

	I	IIa	IIx
Twitch properties	Slow	Fast	Fast
Metabolic properties	Oxidative	Oxidative-glycolytic	Glycolytic
Metabolic energy			
Glycogen level	+	+++	+++
Lipid level	+++	+	+
Myofibril ATPase	+	+++	+++
Hexokinase	+++	++	+
Phosphorylase	+	++	+++
Glycolytic enzyme	+	++	+++
Mitochondrial enzyme	+++	++	+
Physiology			
Moter unit	Slow fatigable resistant	Fast fatigable resistant	Fast fatigable
Contraction	Slow	Fast	Fast
Morphology			
Sarcoplasmic reticulum	+	+++	+++
Transverse tubules	+	+++	+++
Color	Red	Red	White
Myoglobin	+++	+++	+
Capillary density	+++	++	+
Mitochondrion	+++	+++	+
Collagen richness	+++	+++	++
Section area	+	+++	+++

Symbol '+' is used to indicate the level of involvement by the respective characteristics

1.2.1.3 Molecular composition

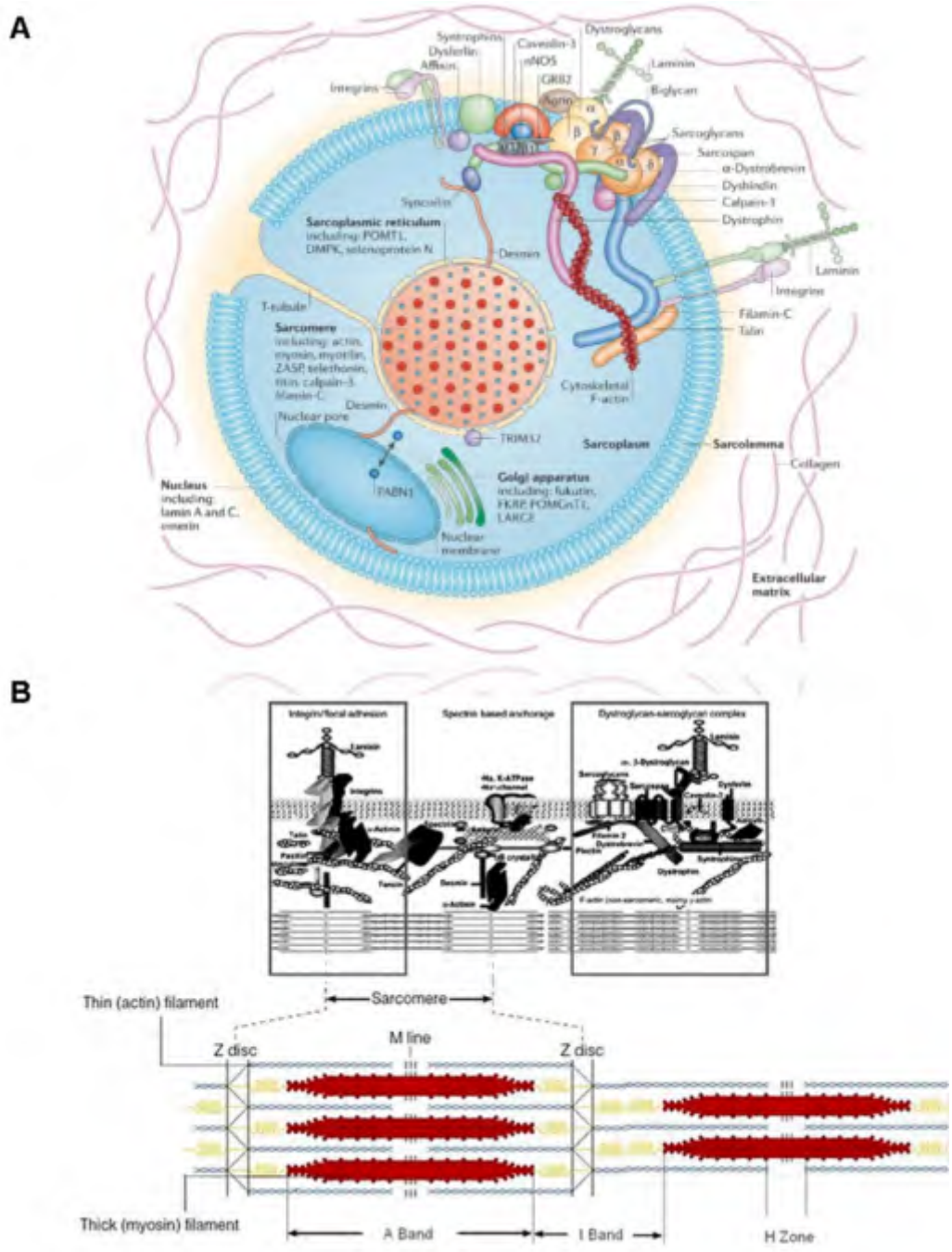


Figure 1.2 Myofibre with component proteins (A) Cross-section of a myofibre showing the approximate position of important component proteins. Here, only one myofibril and one T-tubule are presented. And the T-tubule has not been drawn the whole way through the myofibre. Copyright 2006 by Nature Publishing Group; (B) Myofibre partitioned with costamere

Muscle myofibre is one of the longest cells in the body. In human, it appears as long thin spindles with a diameter of 10 to 100 μm and a length of up to 30 cm. As multinucleated syncytia from fused cells, each myofibre contains several hundred nuclei located in a peripheral position (under the sarcolemma). Beneath the sarcolemma are narrow tubes (also known as **transverse** [**T] tubules**) that are continuous with the sarcolemma and that extend into the sarcoplasm. **Sarcoplasm** contains organelles for intracellular traffic system (**sarcoplasmic reticulum**, **Golgi apparatus, endosomes**), the **mitochondria** is important for energy production, and **ribosomes**, **lysosomes** and **proteasomes** is important for protein turnover (protein synthesis and degradation). The Golgi apparatus is located around the myonuclei close to the sarcolemma and sarcoplasmic reticulum is surrounding myofibrils (Figure 1.2A).

Myofibril is a cylindrical bundle of **myofilaments** (**myosin** thick and **actin** thin) within the skeletal myofibres. Longitudinally the dark anisotropic bands (**A band**) are bounded by two isotropic bands (**I band**) on each myofibril. The A band has a lighter area in its center (**H zone**) and each H zone is divided by a central darker line (**M line**). The I band has a darker area in the center (**Z line or Z disc**). From this point of view, the **sarcomere** is defined as the region between two successive Z lines (Figure 1.2B).

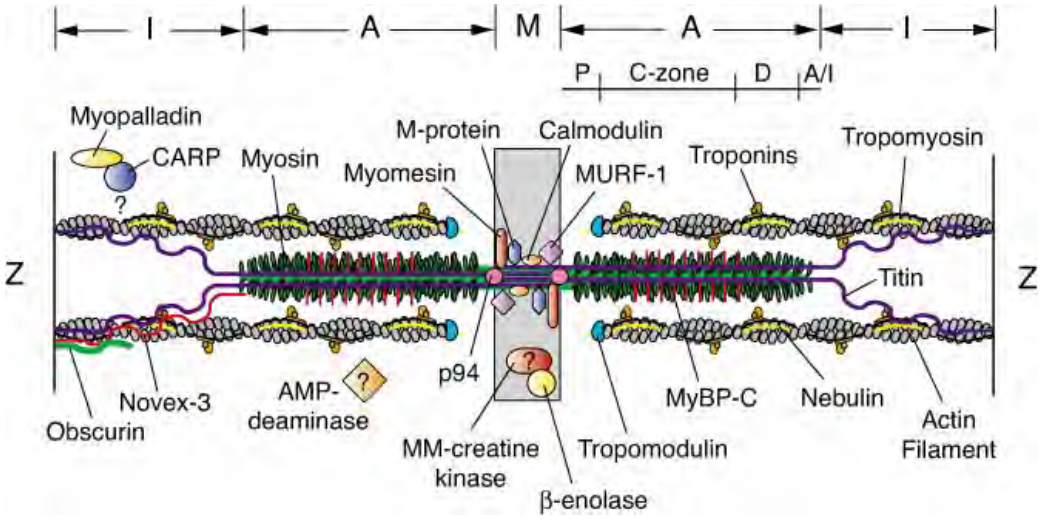


Figure 1.3 Molecular model of A band, I band and M line [14]

In the myofilaments, the I band contains mostly actin filaments and the A band mainly contains myosin filaments. The parts of the A band adjacent to the I-bands are occupied by the myosin and actin filaments. Except for actin and myosin, the thin and thick filament contains proteins such as **troponin** (C, I and T), **tropomyosin,** cardiac myosin binding protein C (**MyBP-C**) and titin. **Titin** is a giant protein that extends an entire half sarcomere. M-line proteins comprise **myomesin**, **M-protein**, muscle ring-finger protein-1 (**MURF-1*)*** and muscle-specific calpain homologue 3 (known as **p94). Calmodulin** on the M-line binds to **nebulin** and nebulin on the thin filament spans the A-I junction (Figure 1.3).

1.2.1.4 Sarcolemma and costamere

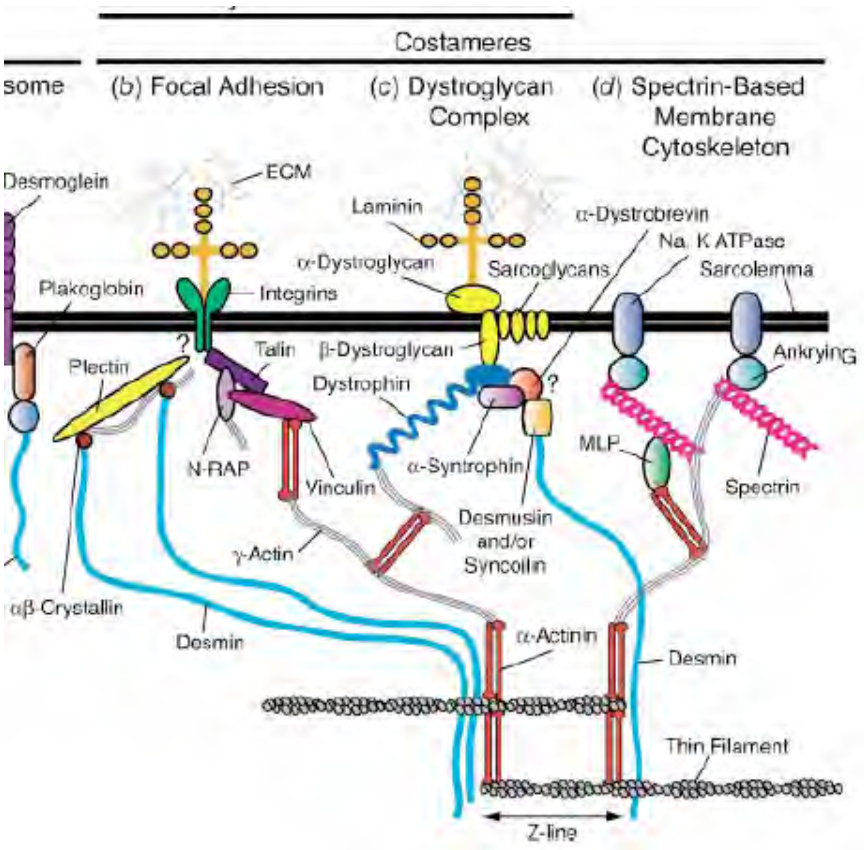


Figure 1.4 Molecular model of costamere [14]

Certain proteins located between the sarcomere and the sarcolemma. These proteins are responsible for forming a complex structure. This structure distributes peripherally to the myofibril and is known as **costamere**. The costamere aligns with the Z-line and coordinately transduces contractile forces and maintains the structural integrity of the membrane during contraction. Costamere contains three separated but interrelated networks: (1) integrin complex (2) dystroglycan complex and (3) spectrin complex (Figure 1.2B). **Integrins** are transmembrane proteins interacting with vinculin, actinin and actin (cytoskeletal molecules) via talin. The **integrin system** plays a major role

in signal transduction; it also responses to mechanical stress and induces cytoskeletal changes. The **dystroglycan complex** is responsible for linking the internal cytoskeletal system of the individual myofibre to the structural proteins within the extracellular matrix (ECM). It consists of dystrophin, dystroglycans, sarcoglycans, dystrobrevin and syntrophin. Among them, **dystrophin** is the major component and it binds actin to costamere. **Dystroglycans** and other molecules connect biglycan, laminin and some other ECM components. This complex has both mechanical and signaling roles in mediating interactions between the cytoskeleton, the sarcolemma, and the ECM. The **spectrin complex** contains spectrin, ankyrin and Na, K-ATPase and is important for anchoring the cytoskeleton to the sarcolemma (Figure 1.4).

1.2.2 Connective tissue

The CT in skeletal muscle is responsible for maintaining the muscle structure and for transmitting force via tendons to the bones. It contains (1) nerve cells embedded in muscle and tendon responsible for the production and regulation of strength; (2) intramuscular lipids, which represent a variable proportion of muscle mass (1 to 5% or more) and which regulate the metabolism of skeletal muscle; (3) stromal cells. These components, i.e. (1)-(3), are in turn surrounded by ECM. Since mechanical and metabolic properties are not in the scope of current study, I will introduce the stromal cells and ECM in the following section.

1.2.2.1 Stromal cells

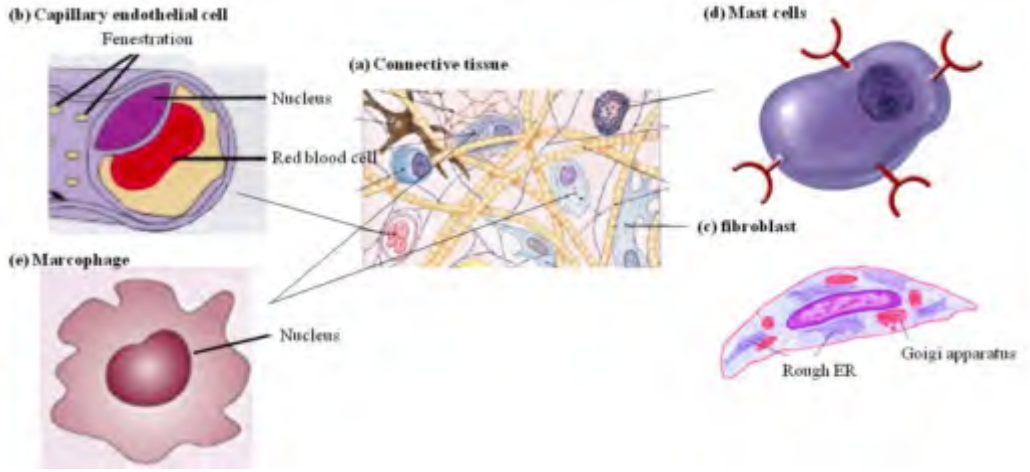


Figure 1.5 Schematic diagrams of stromal cells in skeletal muscle CT, depicting the (a) stromal cells in connective tissue. Copyright 2003 by Person education, (b) capillary endothelial cells; Copyright by University of Minnesota, (c) fibroblast; Copyright by Virginia-Maryland Regional College of Veterinary Medicine, (d) mast cells; Copyright by ADAM and (e) macrophage Copyright by Nature publication Group.

There are several types of **stromal cells** namely capillary cells, fibroblasts, mast cells, and macrophages (Figure 1.5a). Capillary is the smallest blood vessels and provides a larger surface area for the blood exchange in skeletal muscle. It is composed of simple squamous cells (**capillary endothelial cells**, Figure 1.5b). **Fibroblasts** (Figure 1.5c) are abundant CT cells; they produce and release molecular precursors of collagen fibrils. **Mast cells** (Figure 1.5d) facilitate the process of inflammation essential for proper healing. The CT also contains highly specialized **macrophages** (Figure 1.5e) that are capable of phagocytosis.

1.2.2.2 Extracellular matrix

The muscle ECM comprises several large families of molecules that interact with each other. Collagen and elastin locate at perimysium and

endomysium and are surrounded by proteoglycans (PGs) and glycoproteins (GPs). Some molecules are specific at perimysium (e.g. type XII and XIV collagen); some are specific at endomysium (e.g. laminin, type IV collagen and perlecan); some are transmembrane molecules (e.g. dystroglycan, sarcoglycan and associated molecule integrins (trans-sarcolemma molecules)). In the following paragraphs, I will discuss the function and availability of key biomacromolecules, namely collagen, proteoglycan, tenascin, fibronectin and laminin.

- Collagen

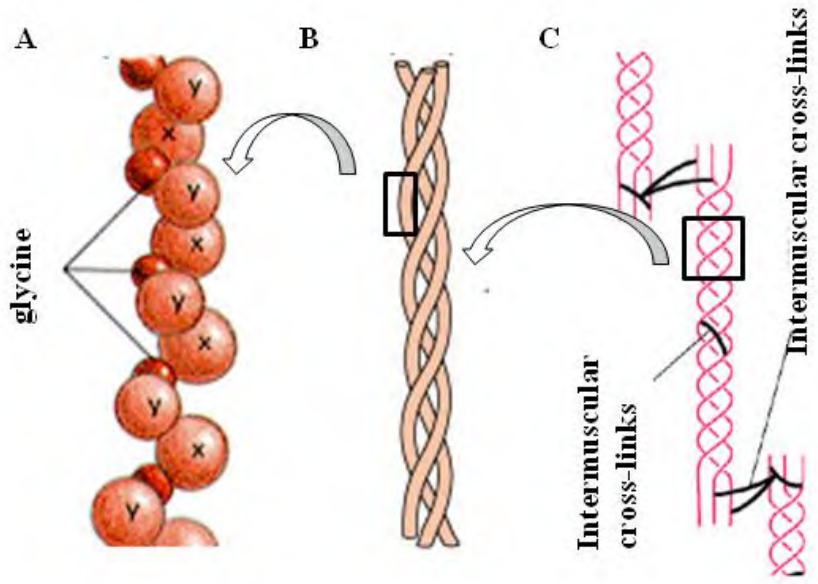


Figure 1.6 Schematic diagrams of collagen structure: (A) α chain composed by gly-x-y repeats; (B) collagen molecule. Copyright 2003 by Pearson education. (C) Intramuscular cross-link. Copyright molecular biology of cell, 4th edition.

Collagens in skeletal muscle consist of fibrillar and non-fibrillar collagens, and fibril-associated collagens with interrupted triple helices (FACIT). **Fibrillar collagens** refer to the collagen types being able to self-

assemble into fibrils. Fibrillar collagens comprise three α chains forming a left-handed triple helix. Each collagen α chain contains many gly-x-y repeats (x is often proline, y is often hydroxyproline). The inherent strength of collagen fibrils is derived from the extracellular formation of covalent intermolecular crosslinking [15]. In skeletal muscle, a **trifunctional pyridinium crosslinking** (PYD) is important to stabilize the mature collagen fibrils (Figure 1.6C). Fibrillar collagens include type I, III and V. Type I collagen (**Col-I**) represents the most abundant intramuscular collagen (reported from ~30% to 97% of total collagen) and it contributes to the structure backbone. Col-I is dominant in endo-, peri- and epimysium. Col-I co-polymerizes with type III (**Col-III**). **Non-fibrillar collagens** include type IV (Col-IV) and type VI collagens (Col-VI). **Col-IV** accounts for over 50% of BM and co-localizes with Col-I, Col-III and other components of BM [16, 17]. **Col-VI** is ubiquitous in the connective tissue (in both endomysium and perimysium). As a substrate for cell attachment and as an anchoring meshwork that connects collagen fibers to the surrounding matrix [18], **Col-VI** interacts with various matrix components such as fibrillar collagens, Col-IV and some PGs. **FACIT** in skeletal muscle comprises the collagen type XII (**Col-XII)** and XIV (**Col-XIV**). They are similar in structure and co-localize with Col-I, -III and -V. [19]. FACIT have important roles in modifying the structure and function of collagen fibrils [20].

Table 1.2 Overview of known collagen types in skeletal muscle [12].

Group	Type	Chains	Distribution/interaction
Fibrillar collagen	I	$[\alpha1(I)]_2$ $[\alpha1(I)]$	Most abundant in endo-, peri- and epi-, interact various components
	III	$[\alpha1\ (III)]_3$	Found in endo-, peri- and/or epi- and co-localized with Col-I; produce by young fibroblasts before Col-I is synthesized
	V	$[\alpha1(V)]$ $[\alpha2(V)]$ $[\alpha3(V)]$	Found in endo-, peri- and/or epi- and co-localized with Col-I
Non-fibrillar collagen *Network forming*	IV	$[\alpha1(IV)]_2[\alpha2(IV)]$	Major component in basal lamina
Beaded fibril forming collagen	VI	$[\alpha1(VI)]$ $[\alpha2(VI)]$ $[\alpha3(VI)]$	Interacts with Col-I and Col-IV
FACIT	XII	$[\alpha1(XII)]_3$	Interacts with Col-I fibrils, decorin, GAGs
	XIV	$[\alpha1(XIV)]_3$	Associate with mature collagen fibrils

The symbol ⟨ in the square brackets refer to the alpha helix; the numerical digit after the symbol is used to distinguish one helix from the other within the biomacromolecule.

- Proteoglycans

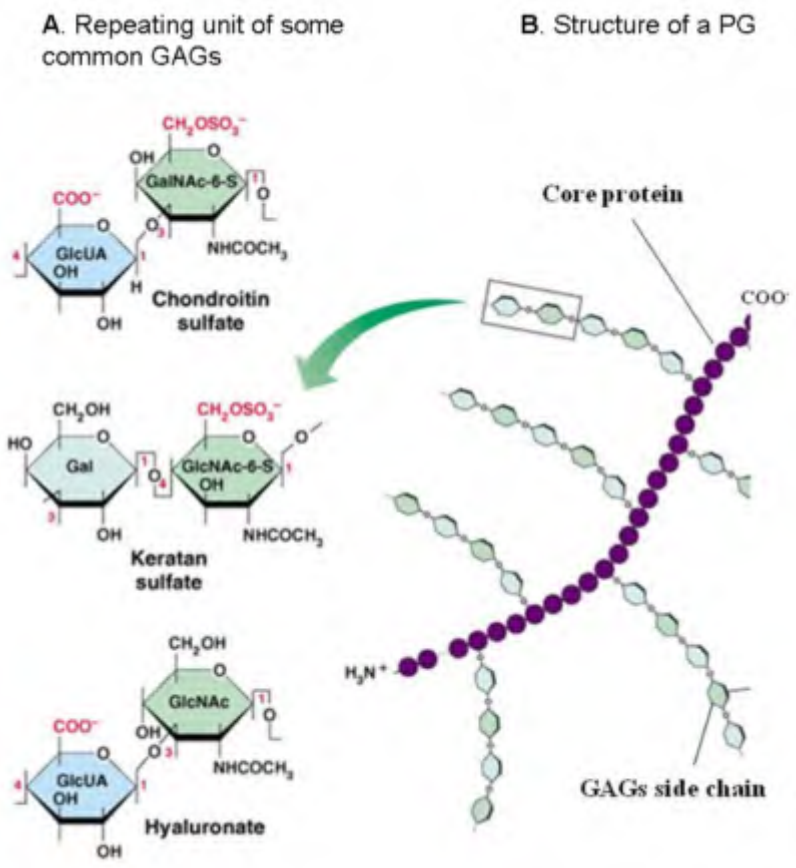


Figure 1.7 Schematic diagrams of PG structure (a) Repeating units of common glycosaminoglycans (GAG); (b) PG molecule composed of core protein and GAGs side chains, Copyright 2003 by the Pearson education.

PGs constitute the major components of the ECM. Their structural diversities underlie their various functions. The general structure of PG consists of a core protein and of **glycosaminoglycan** (GAGs) side chains (Figure 1.7B). Three different types of sulfated GAGs are described: **chondroitin** (CS) / **dermatan** (DS), **keratan** (KS) and **heparan sulfate** (HS) (Figure 1.7A). In skeletal muscle PGs are classified according to their type and number of associated GAGs [21] (Table 1.3).

Table 1.3 Proteoglycans identified in skeletal muscle

Group	Name	GAG	Interaction
Modular	Aggrecan	~100 CS/DS or ~60 KS chains	Hyaluronic acid
	Versican	15-20 CS/DS chains	Hyaluronic acid
SLRP	Decorin	1 CS/DS chain	Type I, III, V, VI, XII, XIV collagen, tenascin X
	Biglycan	2 CS/DS chains	Type I collagen
Basal lamina	Perlecan	3-4 HS or HS+DS chains on the amino terminal end	Type IV
Membrane	Syndecan	3-4 HS chains	Cytoskeleton protein

CS: chondroitin sulfate; *DS*: dermatan sulfate (chondroitin sulfate B), *HS*: heparin sulfate, SLRP: Small Leucine Rich Proteoglycans

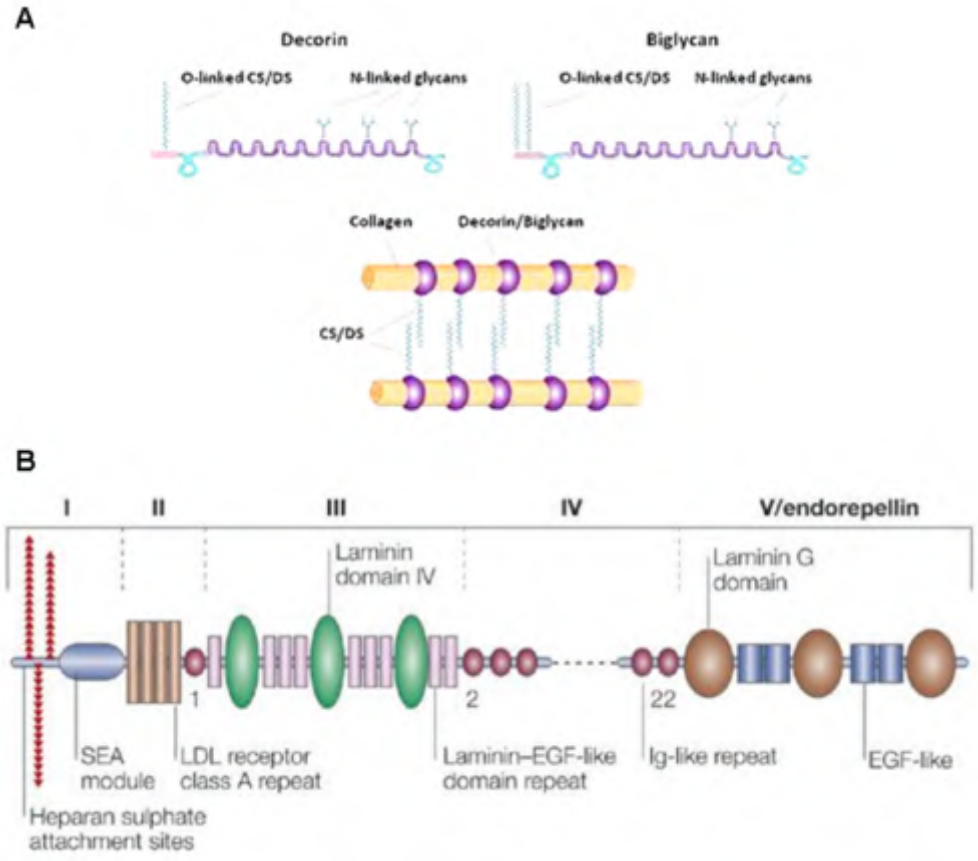


Figure 1.8 (A) SLRP in skeletal muscle: decorin and biglycan share a common core protein structure with ten leucine-rich repeats and at least one chain of GAG. Decorin has one DS chain (or CS chain), whereas biglycan has two CS/DS chains. Both decorin/biglycan connects collagen fibrils by GAG chain, copyright 2009 by Sigma-aldrich. **(B) Schematic model of perlecan protein core**. There are five domains named from the N terminus to the C terminus. Domain I contain and SES module and three triplets that are the binding sites for HS side chains (another potential HS attachment site in domain V is not shown). Domain II contains four repeats. Domain III has three globular domains that are homologous to the laminin domain IV. Domain IV contains 21 Ig-like repeats (only five are shown). Domain V (endorepellin) contains three globular domains that have homology to the laminin G domains and that are each separated by two sets of EGF-like repeats, copyright 2005 by the Nature Publishing Group.

Some PGs in Table 1.3 might have significant implications in aging research. **Decorin** and **biglycan** belong to **small leucine rich PGs** (**SLRP**). These biomacromolecules interact with the collagen fibre for assembling into the supermolecular structures (Figure 1.8A) [22]. **Perlecan** is a modular heparan sulfate proteoglycan and contains five domains (Figure 1.8B). The

perlecan core protein together with its covalently-bound HS chains spans a large 3-D space, which makes perlecan an appropriate link between ECM and cell surface receptors or complexes of receptors [23]. **Syndecan** family (syndecan 1 and 4) are transmembrane HSPG. **Syndecan-1** acts as a co-receptor for multiple growth factor receptors [24]. **Syndecan-3** and **syndecan-4**, known as satellite cell markers, are implicated in satellite cell maintenance and muscle regeneration [25].

- Tenascins

Tenascins are large glycoproteins and exist as 4 isoforms: tenascin-C, -R, -X and -W. All tenascins have a similar structure, consisting of a cysteine-rich N-terminal domain, followed by a series of EGF-like repeats, multiple fibronectin type-III repeats, and a C-terminal globular domain homologous to fibrinogen [26]. Their complex domain structure allows tenascins to interact with a variety of ECM proteins, notably with members of the integrin family. **Tenascin C** (TNC) and **tenascin X** (TNX) are shown to be expressed in skeletal muscle. TNC is found in the perimysium close to the MTJ while TNX is found in endo-, peri- and epimysium [26].

- Fibronectins and laminins

Fibronectins and laminins play a role in cell adhesion. **Fibronectins** are large glycoproteins that form an extensive network of elongated branching fibrils in the ECM [27]. Fibronectins also bind integrin to connect inside to outside of the myofibre. **Laminins** bind cell-surface receptors, basal lamina, i.e.

Col-IV and perlecan, and two transmembrane laminin receptors: dystrophin-glycoprotein complex and integrins.

- Connections between ECM components

A vast amount of reports have essentially studied the above components of ECM in skeletal muscle, but the regulations of those components during developmental and regenerative processes are not fully understood. Figure 1.9 presents the locations and well-established interactions between ECM components in skeletal muscle. Relationships between major components are also shown in Table 1.4.

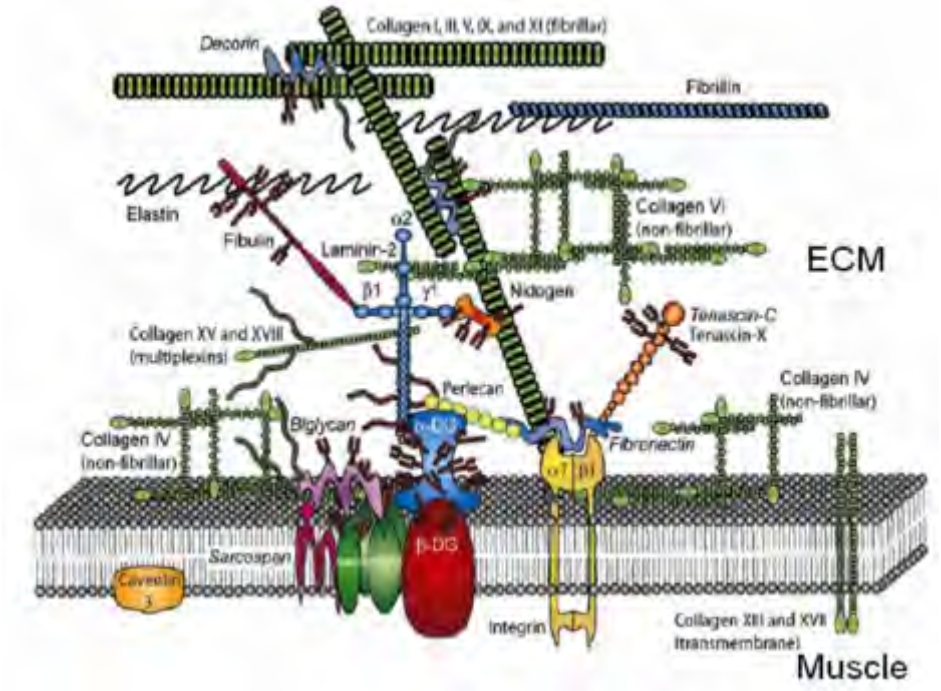


Figure 1.9 Schematic representations of the interactions of ECM components in skeletal muscle [12].

Table 1.4 Interactions between major ECM components in skeletal muscle

	Col I	Col III	Col IV	Col V	Col VI	Col XII	Col XIV	Decorin	Biglycan	Perlecan	Laminin 2α	TNC	TNX
Col I		+		+	+	+	+	+	+	+		+	+
Col III	+												+
Col IV					+					+	+		
Col V	+										+		+
Col VI	+		+										
Col XII	+												+
Col XIV	+												
Decorin	+												
Biglycan	+												
Perlecan	+		+										
Laminin 2α			+										
TNC	+												
TNX	+	+		+		+							

1.3 Model and objectives

1.3.1 Model

1.3.1.1 C57BL/6 mouse: popular model in aging research

The rodents (mouse and rat) are dominantly used to represent mammals for basic aging research. Specifically, long-lived strains of mouse and rat (C57BL/6 mouse, Fisher 344 and Brown Norway rats) are popular although they cost more to house [28]. The reason behind the selection of these strains is because they are close to human and may propose anti-aging intervention going to clinical trials. They also act in a similar way as humans do when investigating the muscle function [29]. For the preference of sexuality, male model is better to avoid potential confounding effects of ovarian hormonal changes that occur mid-life in female rodents [30].

Figure 1.10 C57BL/6 male mouse Copyright 2003 by JAX® Mice database

I have chosen to use the C57BL/6 male mouse (Figure 1.10). The life-span of this genetic strain is well known. Kaneko and colleague [31] believed puberty of C57BL strain almost ceases at 2 months (M). Flurkey and Currer [28] pointed out its survivability began to drop down while incidence of interfering age-related disease is still less than 10% till 22M. Tanaka *et al.* [32]

reported that the mortality is ~25% at around 23-26M but healthy-appearing animals could have been possibly affected by disease to bias the results after 24M. Mortality increases to ~50% at around 30-31M; average lifespan of this strain is 27.5±1M [33]. By scaling the life-span of the C57BL/6 mouse to that of human (given by Jackson Laboratory, US, Table 1.5), I have chosen to study the following age points, 2M, 11M, 22M and 25M. These age points correspond to human at 15 years (yr), 40yr, 65yr and 75yr old, and are categorized as young, adult, early old, and advanced old animals, respectively.

Table 1.5 Age equivalencies between C57BL/6 mice and humans

Equivalent ages				
Mouse	1M	6M	13M	24M
Human	12yr	30yr	45yr	70yr

1.3.1.2 *Gastrocnemius* muscle anatomy

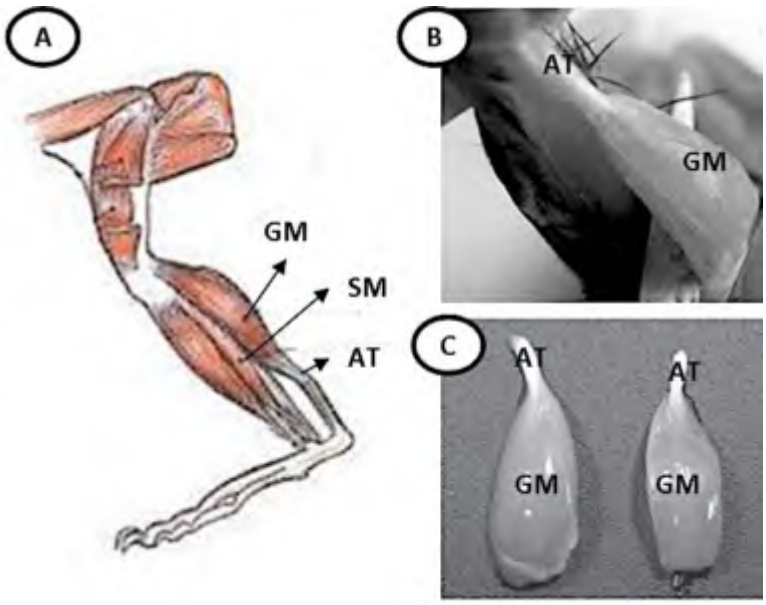


Figure 1.11 Mouse calf muscle anatomy. (A) Lateral view of mouse lower limb, GM: *gastrocnemius* muscle; SM: *soleus* muscle and AT: Achilles tendon. (B-C) *gastrocnemius* during and after dissection

Calf muscle locates at the dorsal lower limb. It is made up of ***gastrocnemius*** **muscle** (GM) that conjoins with ***soleus*** **muscle** (SM) by **Achilles tendon**. As shown in Figure 1.11A, GM arises from above the knee while SM arises from below the knee, but both muscles function the same to elevate the heel.

Although they are supposed to perform the same function from different position, composition of myofibre in GM is distinct from the one in SM. For GM, it comprises 30%, 62%, 8% in rat and 6%, 18%, 74% in mouse of type I, IIa and IIb myofibre, respectively. For SM, it comprises 87%, 13% in rat and 37%, 63% in mouse of type I and IIa myofibre, respectively [34]. In short, GM contains less type I but more type IIa than SM does. Therefore, the metabolism, physiology and morphology of the two muscles are significantly different.

Figure 1.12 Morphology of GM in rat (longitudinal section), the pennate structure of the GM can be seen. There exists three aponeuroses, enhanced by black-dot dash lines and the N-like configuration of the motor endplate bands (arrowheads) [35]**.**

The GM was selected for this project because the tissue features well-defined longitudinal pennate structure (Figure 1.12) similar to that of rat. In addition, three aponeuroses of the GM, i.e. sheets of pearly white fibrous tissue that take the place of tendon is sheet-like muscles having a wide area of

attachment, are not only visible (under the naked eye) but also yield heterogeneous myofibres in direction and density [35]. For quantifying the histological properties of GM, further protocol is required to ensure that the planes of the sectioned GM are perpendicular to the longitudinal axis of the aponeuroses.

1.3.1.3 Sample preparation

- Mice

Fifty-four C57BL/6 male mice were used in this study. They were obtained from the Laboratory Animal Centre at National University of Singapore at the age of 3-8 weeks (wk) and were housed in a temperature and humidity controlled facility (Laboratory Animal Facility at Nanyang Technological University) with a 12:12 h light-dark cycle and received food and water ad libitum.

Table 1.6 Fifty-four C57BL/6 male mice served in the study were assigned and slaughtered at 4 ages**.**

Age (M)	2	11	22	25
No.	15	15	9	15

- Mice termination and GM dissection

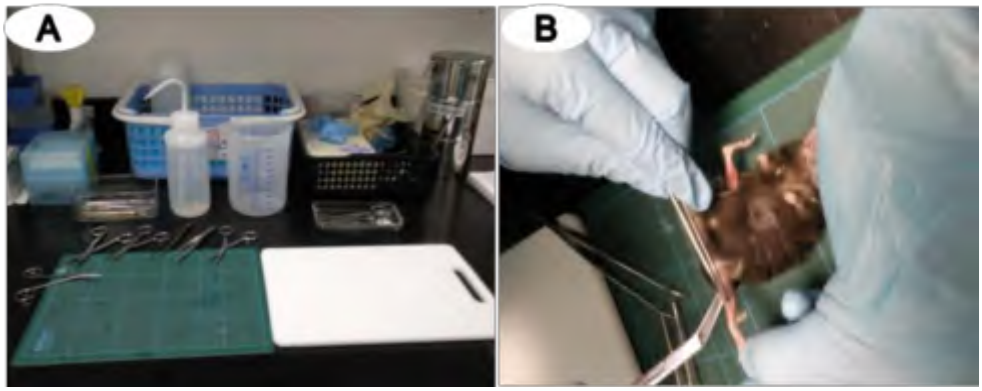


Figure 1.13 GM dissection. (A) Preparation: dissecting and trimming board, scissors, ultra-fine tweezers and forceps were used. Washing bottle contained 0.1% SDS. (B) Dissection.

Those mice were anesthetized by carbon dioxide and weighed immediately. Subsequently, the GM (from both left and right leg) was dissected within 10 minutes (mins) and weighed at room temperature (RT, around 20°C).

To minimize contamination, prior to dissection I sterilized the dissecting and trimming board, scissors, ultra-fine tweezers and forceps using 0.1% (v/v) sodium dodecyl sulfate (SDS) (Figure 1.13A). This sterilization was also performed before and after each operation during the whole processing of dissection.

Next, the entire legs, haunch and lower back of each mouse were shaved. Removing the integument (cuticle plus dermis) exposes the aponeurosis of GM (which appeared as a white translucent band beginning from the Achilles tendon in Figure 1.11A). Tissues outside the aponeurosis were carefully removed but small pieces of Achilles tendon near ankle were preserved to maintain the integrity of GM (Figure 1.13B).

Each GM block was weighed immediately and frozen. For thirty-two blocks from sixteen mice at ages: 2M, 11M, 22M and 25M (4 mice per age), the frozen medium (isopentane) was prepared by chilling it into liquid nitrogen for 6-8 mins. Muscle block was then fixed on cork pieces by OCT™ compound (Leica) and frozen in this medium (about -150°C) for 30 seconds (sec). The frozen block was then placed in a cryo-tube, labeled and preserved into -80°C for further analysis. For seventy-six GM blocks from 38 mice at ages of 2M, 11M, 22M and 25M (11, 11, 5 and 11 mice per age, respectively), the frozen medium was prepared in liquid nitrogen (-196°C). Muscle block was then directly frozen by this medium for 60-90 sec and grounded with a mortar cooled in liquid nitrogen. Frozen powder from each block was preserved in cryo-tubes at -80°C until further analysis.

- Section

The thirty-two GM blocks mentioned in above paragraph were employed for histological and immunohistochemical studies. Thus cryo-sections attached to the microscope slide were prepared. I have chosen the medial part of GM and prepared cryo-sections following the schematic representation in Table 1.7. Configuration of cryostat (MICROM, HM500M) was -25°C, 10/20µm (trim) and 40 (feed). Those microscopic slides were assigned in different experiments (Methods in Table 1.7) and freshly prepared before those experiments.

Table 1.7 List of methods utilized in this study

Purpose	Methods
Myofibre morphology	Azorubin (Protocol N°1)
ECM organization	Sirius red (Protocol N°2)
Myofibre metabolism	Cox (Protocol N°3)
ECM composition observation	IHC (Protocol N°4)
Complementarity	
ECM component co-localization	IHC (Protocol N°5)
Complementarity and control	
Muscular nuclei distribution (normal and apoptotic)	IHC (Protocol N°6 and N°7)
Complementarity and control	

Complementarity: substitute slides without assigned experiments at present. *: for some sections, respective fixation is required and detailed protocol will be described in section 2-2. Thickness for slide is 10 or 20μm

1.3.2 Objective

Global aging is taking place in majority of countries and regions across the world. To deal with its consequence, for instance, the increased morbidity and age-related frailty, scientists reach the same conclusion: perform new strategies to better understand the mechanism underlying aging. The progressive decline in muscle mass and function (**sarcopenia**) is an important feature of age-related frailty and inevitably contributes to an impaired functional autonomy and to loss of independence. Although sarcopenia has been intensively studied by numerous researchers focusing on the morphological, physiological and metabolic variations in myofibre, age-related variations in the

structure and composition of the CT during sarcopenia has not been well documented.

The CT provides the support and connections to myofibres, tendons and bones (section 1.2.2). Thus it can affect the muscle function via these specific interactions. In order to understand the importance of the modification in the connective tissue during aging, I chose the skeletal muscle from mice between 2 and 25 months old as the model, that covers the period of mature and old age (section 1.3.1.1). The modifications in the distribution and composition of connective tissue were investigated.

In the first part of this thesis, a review of the structure and composition of skeletal muscle, especially the muscle fibres and the CT was presented; followed by the objectives and models used in this project (Chapter 1). The next three chapters present this work in detail.

The objectives of this project have been implemented in several steps. It began with the characterization of the structural modifications in skeletal muscle during growth, maturation and aging. Following this a comprehensive study on the compositional modifications was conducted via two strategies. First, some candidate molecules in connective tissue were chosen based on bibliography. Because of the tiny amount of available sample, they were studied by histology and immunohistochemisty, associated with image analysis. My results highlighted a possible connection between connective tissue components with apoptosis. Being aware of the interaction between sarcopenia and apoptosis, I measured the apoptotic nuclei in different muscular components

including satellite cells (at the interface of myofibre and connective tissue). Those results were analyzed by classic statistical methods (Chapter 2). In chapter 3, I developed a novel method (data mining) to effectively estimate the importance of the individual characteristic in the contribution to the mechanism of aging, while the individual characteristic refers to the data derived from Chapter 2. Second, a global vision of age-related modifications in connective tissue was obtained by microarray. Microarray technology is a widely used tool for measuring gene expression and it generates a vast amount of expression data. Methods to process and analyze these data are well established, but in this project, I chose "exon array" (Affymetrix) which was little studied but capable of assessing the expression at the individual exon level. It exploits a much more comprehensive analysis of the transcriptome and enables the study of alternative splicing, a gene regulation mechanism being important for normal, disease and aging conditions. I performed this approach rather than a study of protein expression by proteomics because the connective tissue molecules are always insoluble and with high molecular weight, which makes the protein extraction extremely difficult (Chapter 4).

The thesis concludes with a summary of my findings and a discussion of future work (Chapter 5).

CHAPTER 2 SIGNIFICANCE OF SARCOPENIA, ATROPHY, APOPTOSIS AND FIBROSIS DURING AGING

2-1 Introduction

It is commonly believed that the body weight of mammals increases during growth and maturation and is then decreased in elderly individuals [36-38]. Body composition, including fat and lean tissue undergoes variations during these periods. For instance, fat mass progressively increases from growth to aging. However, skeletal muscle mass increases during growth and then declines during aging [39, 40], and this age-dependent loss of skeletal muscle is referred to as **sarcopenia**.

From a histological perspective, variation of myofibres is the chief reason of the muscle mass change, because myofibres are the major component of the skeletal muscle. During growth, the number and the volume of myofibres increase, and these processes are known as **hyperplasia** and **hypertrophy,** respectively. During aging in contrast, the number and the volume of myofibres decrease. The decreasing number of myofibres is ascribed to a selective loss of fast-twitch, type II myofibres [41]; but the decreasing volume of myofibres, also known as **atrophy**, is not well understood. The cross-sectional area (CSA) of a myofibre is characterized by the number of myonuclei and the amount of the cytoplasm per myonucleus, known as **myonuclear domain** (MND) [38]. It is

widely accepted that each myonucleus can only govern a finite volume of cytoplasm, thus a fix MND size is required and may be maintained during life time. However, Bruusgaard and co-workers reported variations in MND size in mouse [42] and rat muscle [43].

The majority of researches ascribe the loss of myonuclei to **apoptosis**, which is a programmed cell death involving DNA fragmentation. With respect to the fixed MND size, this elimination eventually leads to myofibre volume reduction, hence causing atrophy. Although it seems simple and elegant, it has not been unequivocally demonstrated that all the apoptotic nuclei are myonuclei. On the contrary, some researchers suggested that apoptosis does occur mainly in stromal cells and satellite cells rather than in myofibres [44].

Skeletal muscle hypertrophy that occurs during growth is partly associated with an increased number of myonuclei. This larger amount of myonuclei has been shown to result from the activation of **satellite cells** (SCs, section 1.2.1.1). SCs, lying on the surface of myofibre but beneath its basement membrane [45], can supply myonuclei during muscle development or in adult muscle [46]. First, SCs divide to produce satellite cell-derived myoblasts. These myoblasts further proliferate and fuse with existing myofibres or form new myotubes which then mature into myofibres [47]. During sarcopenia, SCs fulfill their roles in routine maintenance and repair of old skeletal muscle by self-renewal. In response to the loss or damage of myofibres, fusion of many SCs leads to the forming and repairing of myofibres. These newly regenerated myofibres are first marked by centrally aligned nuclei (**central myonuclei**);

later, those central myonuclei will migrate to assume a more peripheral location [48].

Muscle aging affects not only myofibres, but also surrounding cells and CT. Increased capillary, namely capillarization has been investigated in various conditions and models, and illustrated by different indexes [49]. In human, many researchers reported an increased capillary to myofibre ratio, which is associated with non-changed capillary density [49]. In mouse, the capillarization might improve, which was demonstrated by the increased capillary to myofibre ratio [50] but also the surrounding connective tissue. Age-related changes in stromal cells are little studied, but Cui and co-workers [51] reported the increased ratio of **capillary** per myofibre in aging muscle. It is associated with the increased proportion of type I myofibre because a rich capillary structure is required to provide demanded oxygen.

Muscle aging comes with **fibrosis**, which is essentially an excessive accumulation of ECM components [52, 53]. Fibrosis is characterized by the increased collagen content [54, 55]. It is associated with increased crosslinking [56] and comes with increased rigidity [52], decreased solubility and stabilization of collagen [57]. Although these results are exciting, still more data are required to better understand the modification of ECM composition.

2-2 Methods

In this chapter, cryo-sections (section 1.3.1.3) of the GM from GM in 2, 11, 22 and 25M old mice (4 mice per age) were prepared to study the (1) muscle

ECM and myofibre structure; (2) muscle ECM components and (3) muscular nuclei distribution. Each experiment was repeated on 5 sections per mouse to obtain optimal result. Frozen powder from the GM in 2, 11, 22 and 25M old age (5 mice per age, section 1.3.1.3) was employed to study some other ECM components (endorepellin, collagen content and its crosslinking).

First, histological techniques, including azorubin, cytochrome c oxidase and Sirius red staining provide a visual examination of the myofibre and ECM morphology and metabolism at microscopic level. In my study, these staining were applied on the GM cryo-sections (thickness ~ 10 μm).

2.2.1 Myofibre morphology

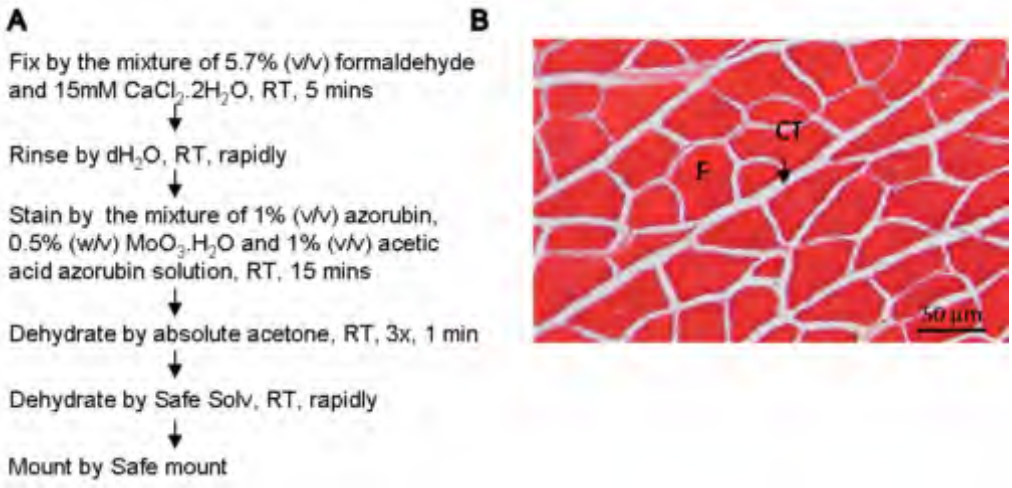


Figure 2.1 Azorubin staining. (A) Protocol (N° 1), RT: room temperature, formaldehyde 37% (Sigma), CaCl2.2H2O (Prolabo), azorubin (Serva), MoO3.H2O (Sigma), safe solv (Labonord), safe mount (Labonord). (B) Cross section of mouse GM in 11M-old with myofibre (F) stained in red and connective tissue (CT) remained in blank.

Azorubin stains myofibres in red and ECM in white. In my study, morphology of myofibres was examined by following the protocol N°1 [58]. A flow-chart of the protocol is illustrated in Figure 2.1.

2.2.2 Myofibre metabolic activity

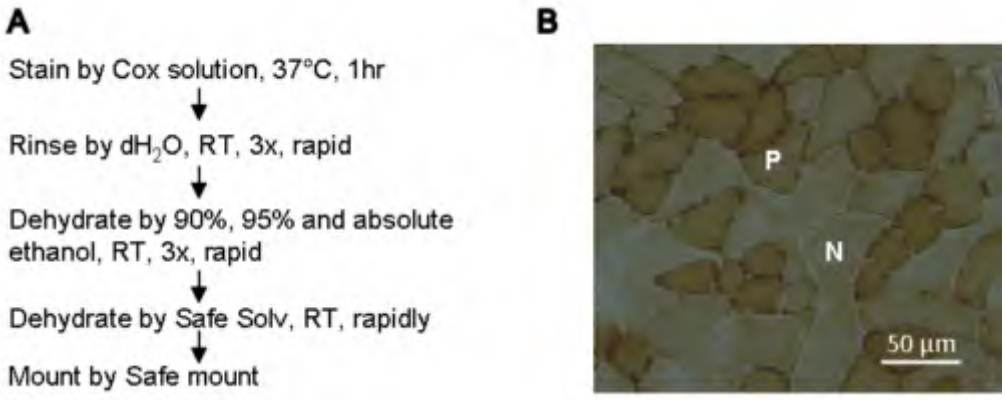


Figure 2.2 Cox staining, (A) Protocol (N° 2). Cox kit (Bio-optical, 30-30115LY). (B) Cross section of mouse GM in 11M-old mouse with Cox-positive (P) and Cox-negative (N) myofibres in dark brown and light brown, respectively.

Cytochrome c oxidase (cox), the terminal enzyme of the mitochondrial respiratory chain, comprises a variety of subunits and some of them presents large histological differences in skeletal muscle [59]. The use of 3-3' diaminobenzidine (DAB) results in a brown insoluble compound at the site of cytochrome c, and the metabolic activity of myofibre is then distinguished by darker (type I) and lighter (type II) color compound [60]. In this study, I performed cox staining by following the instruction of the manufacturer (protocol N°2). A flow-chart of the protocol is illustrated in Figure 2.2.

2.2.3 ECM structure

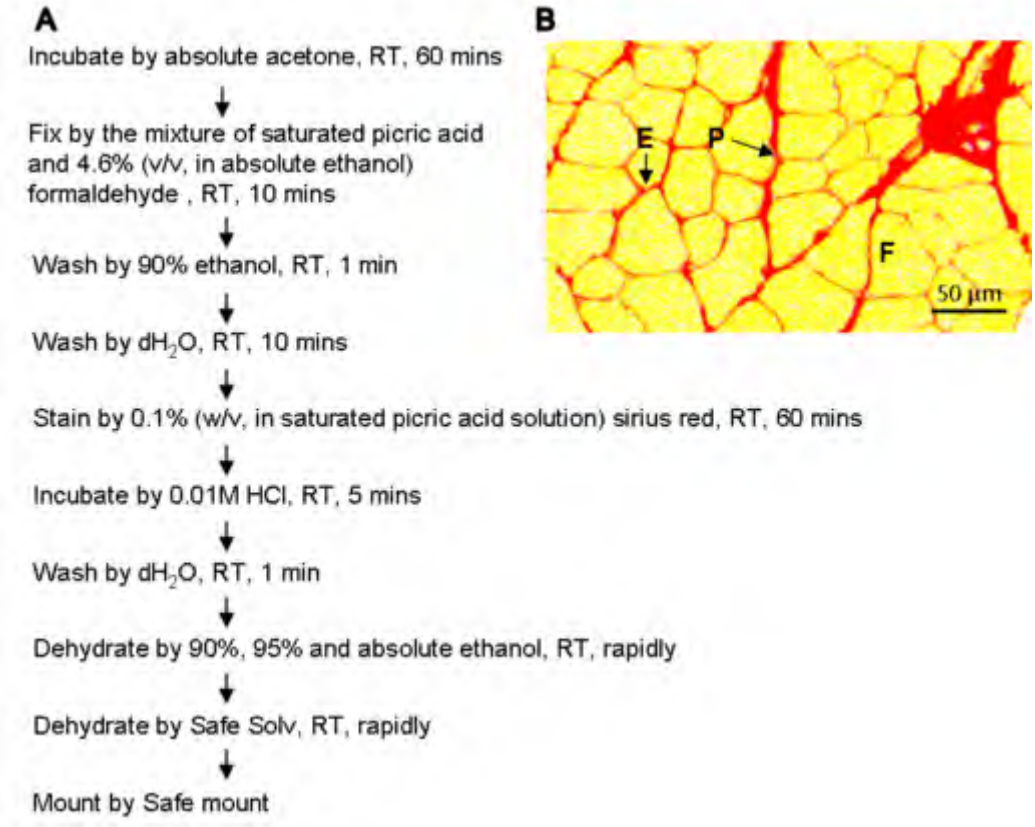


Figure 2.3 Sirius red staining. (A) protocol (N° 3), picric acid (prolabo), sirius red (BDH) (B) Cross section of mouse GM in 11M-old with endomysium (E) and perimysium (P) of ECM in red while myofibre (F) are in light yellow.

Sirius red is a strong anionic dye and reacts with the collagen molecule via its sulphonic acid group [61, 62]. It simultaneously colors Col-I and reticulin (Col-III) in red. As pointed out in section 1.2.2.2, collagen, especially fibrillar collagen, represents the major component of ECM and is present in both endo- and perimysium. In my study, I adopted the protocol N°3 (Figure 2.3). This protocol N°3 was first proposed by Flint and Pickering [63] and amended by Liu *et al*. [64].

2.2.4 ECM components localization and co-localization

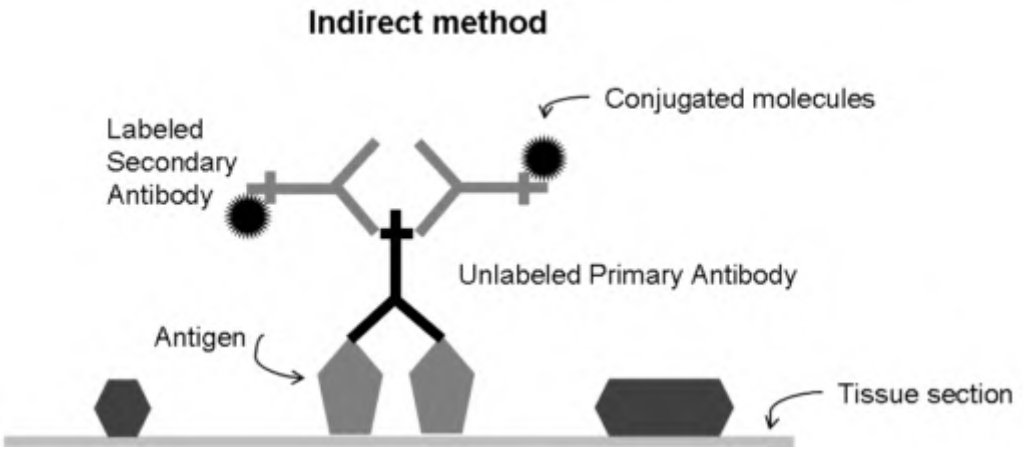


Figure 2.4 Schematic diagram of the indirect method for carrying out IHC.

Immunohistochemistry (IHC) is employed to identify a target tissue component (a protein) via its interaction with a specific antibody. While there are multiple approaches in IHC methodology, I chose to use the indirect method (Figure 2.4) to label and visualize the target protein in mouse GM. Here, an unlabeled primary antibody (first layer) first reacts with the tissue antigen, and then a labeled secondary antibody (second layer) reacts with the primary antibody. A standard protocol includes fixation (optional), enzyme action (optional) and blocking before reacting with the primary antibody. This method is more sensitive than direct immunohistochemical method. In addition, it is also economical since one labeled second antibody can be used with many primary antibodies (raised from the same animal species).

- Single immunolabeling

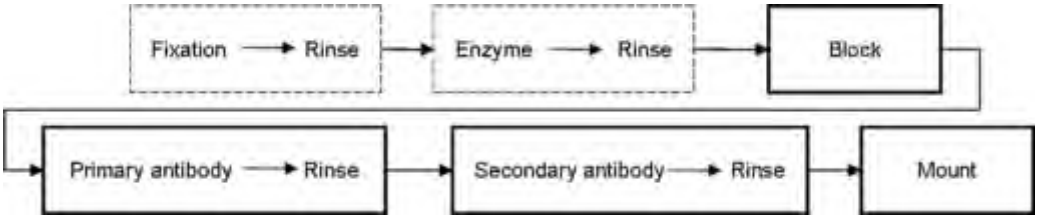


Figure 2.5 Flow-chart illustrating the general procedure for carrying out single immunolabeling. Dashed line box refers to the optional step.

I attempted to study the key ECM components (illustrated in Figure 1.4) as much as possible. However, not all antibodies work optimally. I studied the localization of ten kinds of ECM components (Table 2.1) by single immunolabeling on 10 μm thickness sections. I performed one negative control in each experiment by using PBS1x solution instead of primary antibody. False positive would result in the fallacious identification of target molecule.

- Double Immunolabeling

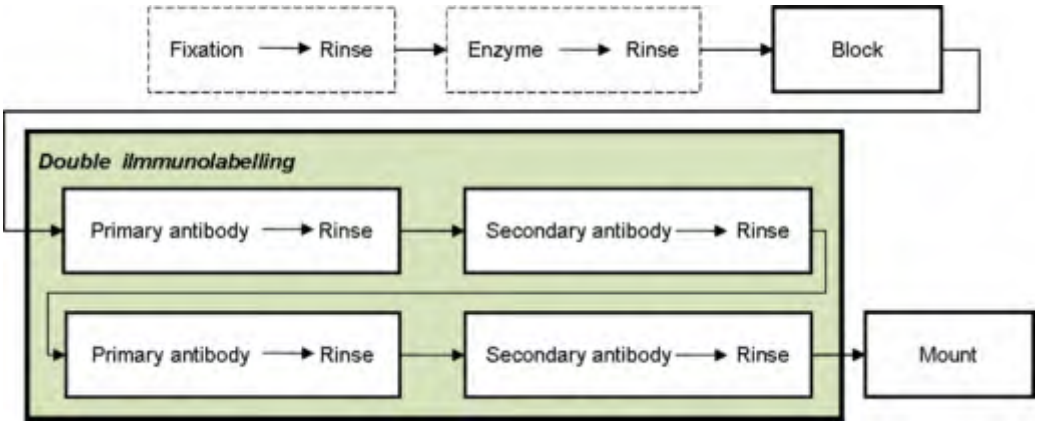


Figure 2.6 Flow-chart illustrating the general procedure for carrying out double immunolabeling. Dashed box refers to the optional step. Comparing to the single immunolabeling (Figure 2.5), here another pair of antigen-antibody reaction is added.

Co-localization of seven pairs of ECM components was studied by double immunolabeling applied on 20 μm thickness sections and followed the Protocol N°5. As shown in Table 2.2, it allows for the discrimination of two

ECM components by different secondary antibodies. Negative controls were not placed because they have been tested in the previous section (single immunolabeling). Cross-reactivity was examined. In brief, for each double labeling (antigen A and B), I examined the cross reactivity between primary antibody (anti-A) and secondary antibody which is used for another primary antibody (anti-B). Similarly, the cross reactivity between primary antibody (anti-B) and secondary antibody which is used for another primary antibody (anti-A) was taken into account. A strong cross-reactivity would lead to fallacious identification of target molecule.

Table 2.1 (Protocol N°4) Single immunolabeling

Antibody	Ref	Specificity	Dilution	Protocol
M-IgG1-rat anti TNC	Abcam ab6346	Binds to the recombinant fragments of mouse **tenascin C**	1:500	**Fix** in cold acetone for 10mins at 4°C ⇒ **Rinse** by PBS 1x for 3x 5mins ⇒ **Block** by 5% (w/v) BSA for 30mins at RT⇒**React** with TNC for 2hr at RT ⇒**Rinse** by PBS 1x for 3x 5mins ⇒**React** with Dylight 488 donkey anti rat (1:400) for 45mins in dark room⇒**Rinse** by PBS 1x for 3x 5mins⇒**Mount**
M-rat anti perlecan	Abcam ab17848	Epitope on **perlecan** not determined	N/A	**Incubate** by 440U/ml hyaluronidase for 30mins at 37°C⇒ **Rinse** by PBS 1x for 3x 5mins⇒**Fix** in cold acetone for 10mins at 4°C ⇒**Rinse** by PBS 1x for 3x 5mins ⇒ **Block** by 5% (w/v) BSA for 30mins at RT⇒**React** with perlecan for 1hr at RT ⇒**Rinse** by PBS 1x for 3x 5mins ⇒**React** with Dylight 488 donkey anti rat (1:400) for 45mins in dark room⇒**Rinse** by PBS 1x for 3x 5mins⇒**Mount**
M-IgG1-rat anti laminin 2α	Abcam ab11576	Reacts with the P1 fragment of **laminin**	1:200	**Fix** in cold acetone for 10mins at 4°C ⇒**Rinse** by PBS 1x for 3x 5mins ⇒ **Block** by 5% (w/v) BSA for 30mins at RT⇒**React** with Laminin 2a for 1hr at RT ⇒**Rinse** by PBS 1x for 3x 5mins ⇒**React** with Dylight 549 donkey anti rat (1:400) for 45mins in dark room⇒**Rinse** by PBS 1x for 3x 5mins⇒**Mount**
M-IgG1-mouse anti C4S	Cardiff University	Reacts with **chondroitin-4-sulfate**	1:50	**Incubate** by 0.5U/ml Chondroitinase ABC and Keratanase2 for 1hr at 37°C⇒**Rinse** by PBS 1x for 3x 5mins ⇒ **Block** by 5% (w/v) BSA for 30mins at RT⇒**Rinse** by PBS 1x for 3x 5mins ⇒**Block** by 10% (v/v, in PBS 1x) unconjugated affinipure Fab fragment goat anti mouse IgG (H+L) for 1hr at RT⇒**Rinse** by PBS 1x for 3x 5mins ⇒**React** with C4S for 1hr at RT ⇒**Rinse** by PBS 1x for 3x 5mins ⇒React with Dylight 488 donkey anti mouse(1:400) for 45mins in dark room⇒**Rinse** by PBS 1x for 3x 5mins⇒**Mount**
M-IgM-mouse anti C6S	Cardiff University	Reacts with **chondroitin-6-sulfate**.	1:50	**Incubate** by 0.5U/ml Chondroitinase ABC and Keratanase2 for 1hr at 37°C⇒**Rinse** by PBS 1x for 3x 5mins ⇒ **Block** by 5% (w/v) BSA for 30mins at RT⇒**Rinse** by PBS 1x for 3x 5mins ⇒**Block** by 10% (v/v, in PBS 1x) unconjugated affinipure Fab fragment goat anti mouse IgG (H+L) for 1hr at RT⇒**Rinse** by PBS 1x for 3x 5mins ⇒**React** with C4S for 1hr at RT ⇒**Rinse** by PBS 1x for 3x 5mins ⇒React with FITC IgM goat anti mouse(1:40) for 45mins in dark room⇒**Rinse** by PBS 1x for 3x 5mins⇒**Mount**

P: polyclonal; M: monoclonal; BSA: bovine serum (Sigma). Paraformadehyde (EMS); Chondroitinase ABC and Keratanase (Sigma); Hyaluronidase (Bio-optical); Mount media (Biomeda); Dylight-conjugated secondary antibodies (712-505-150, 712-486-150, 715-505-151, 715-486-150, 711-486-152) are from Jackson immunoresearch laboratory. Alexa-conjugated secondary antibody (A11074) is from invitrogen. FITC-conjugated secondary antibody (115-095-020) is from Jackson immunoresearch laboratory. Primary and secondary antibodies are diluted by PBS 1x

Table 2.2 (Protocol N°5) Double immunolabeling

Antigens	Protocol
Col-I + TNX	**Block** by 5% (w/v) BSA for 30mins at RT⇒**React** with TNX (1:100) for 1hr at RT ⇒**Rinse** by PBS 1x for 3x 5mins ⇒React with Alexa 546 goat anti guniea pig (1:500) for 45mins in dark room⇒**Rinse** by PBS 1x for 3x 5mins⇒**React** with Col-I (1:20) for 1hr at RT ⇒**Rinse** by PBS 1x for 3x 5mins ⇒**React** with dylight 488 donkey anti rabbit (1:400) for 45mins in dark room⇒**Rinse** by PBS 1x for 3x 5mins⇒**Mount**
Col-I + TNC	**Fix** in cold acetone for 10mins at 4°C ⇒ **Rinse** by PBS 1x for 3x 5mins ⇒ **Block** by 5% (w/v) BSA for 30mins at RT⇒**React** with TNC (1:500) for 2hr at RT ⇒**Rinse** by PBS 1x for 3x 5mins ⇒**React** with Dylight 549 donkey anti rat (1:400) for 45mins in dark room⇒**Rinse** by PBS 1x for 3x 5mins⇒**React** with Col-I (1:20) for 1hr at RT ⇒**Rinse** by PBS 1x for 3x 5mins ⇒**React** with dylight 488 donkey anti rabbit (1:400) for 45mins in dark room⇒**Rinse** by PBS 1x for 3x 5mins⇒**Mount**
TNC + TNX	**Fix** in cold acetone for 10mins at 4°C ⇒ **Rinse** by PBS 1x for 3x 5mins ⇒ **Block** by 5% (w/v) BSA for 30mins at RT⇒**React** with TNC (1:500) for 2hr at RT ⇒**Rinse** by PBS 1x for 3x 5mins ⇒**React** with Dylight 549 donkey anti rat (1:400) for 45mins in dark room⇒**Rinse** by PBS 1x for 3x 5mins⇒**React** with TNX (1:100) for 1hr at RT ⇒**Rinse** by PBS 1x for 3x 5mins ⇒**React** with Alexa 546 goat anti guniea pig (1:500) for 45mins in dark room⇒**Rinse** by PBS 1x for 3x 5mins⇒**Mount**
Col-IV + Perlecan	**Incubate** by 440U/ml hyaluronidase for 30mins at 37°C⇒**Rinse** by PBS 1x for 3x 5mins ⇒**Fix** in cold acetone for 10mins at 4°C ⇒**Rinse** by PBS 1x for 3x 5mins ⇒ **Block** by 5% (w/v) BSA for 30mins at RT⇒**React** with perlecan (1:1) for 1hr at RT ⇒**Rinse** by PBS 1x for 3x 5mins ⇒**React** with Dylight 549 donkey anti rat (1:400) for 45mins in dark room⇒**Rinse** by PBS 1x for 3x 5mins⇒**React** with Col-IV (1:20) for 1hr at RT ⇒**Rinse** by PBS 1x for 3x 5mins ⇒**React** with dylight 488 donkey anti rabbit (1:400) for 45mins in dark room⇒**Mount**
Col-IV + Laminin 2α	**Fix** in cold acetone for 10mins at 4°C ⇒**Rinse** by PBS 1x for 3x 5mins ⇒ **Block** by 5% (w/v) BSA for 30mins at RT⇒**React** with Laminin 2a (1:200) for 1hr at RT ⇒**Rinse** by PBS 1x for 3x 5mins ⇒**React** with Dylight 549 donkey anti rat (1:400) for 45mins in dark room⇒**Rinse** by PBS 1x for 3x 5mins⇒**React** with Col-IV (1:20) for 1hr at RT ⇒**Rinse** by PBS 1x for 3x 5mins ⇒**React** with dylight 488 donkey anti rabbit (1:400) for 45mins in dark room⇒**Mount**

2.2.5 Nuclei distribution

- Normal nuclei

Muscular nuclei were detected by Hoechst 33258 because it offers high specificity and sensitivity for dsDNA quantification. Muscular nuclei include stromal nuclei (including capillary endothelial nuclei), SCs and myonuclei. To distinguish them, I used laminin 2α to show the basement membrane; used CD31 to show the capillary endothelial cells; used Pax7 to show the quiescent and activated SCs; and used dystrophin to show the sarcolemma. By following Protocol N°6 (Table 2.3), SCs were identified by the double-labelling of Pax7 and Hoechst 33258; capillary endothelial cells were identified by the double-labelling of CD31 and Hoechst 33258. Myonuclei were identified by the labelling of Hoechst 33258 beneath the dystrophin labelling. And stromal nuclei were identified by the labelling of Hoechst 33258 external the laminin 2α labelling. I also performed one negative control in each experiment by using PBS1x solution instead of primary antibody. False positive would result in the fallacious identification of target molecule. In addition, I measured the total muscular nuclei by accounting Hoechst 33258 in each experiment

Table 2.3 (Protocol N°6) Immunostaining of muscular nuclei

Targets	Primary antibody	Ref	Protocol
Satellite cells	M-mouse anti-Pax7	DSHB	**Block** by 5% (w/v) BSA for 30mins at RT⇒**Rinse** by PBS 1x for 3x 5mins⇒**Block** by 10% (v/v, in PBS 1x) unconjugated affinipure Fab fragment goat anti mouse IgG (H+L) for 1hr at RT⇒**Rinse** by PBS 1x for 3x 5mins⇒**React** with Pax7 (1:50) for 1hr at RT⇒**Rinse** by PBS 1x for 3x 5mins⇒**React** with dylight 549 donkey anti mouse (1:400) for 45mins in dark room⇒**Rinse** by PBS 1x for 3x 5mins⇒**React** with 1% Hoechst (v/v in PBS 1x) for 10 secs at RT⇒**Rinse** by PBS 1x rapidly⇒**Mount**
Capillary endothelial nuclei	M-rat anti - CD31	Abcam ab7388	**Fix** in 4% (v/v) paraformadehyde for 10mins at RT ⇒**Rinse** by PBS 1x for 30mins ⇒**Block** by 5% (w/v) BSA for 30mins at RT⇒**React** with CD31 (1:100) for 1hr at RT ⇒**Rinse** by PBS 1x for 3x 5mins ⇒React with Dylight 549 donkey anti rat (1:400) for 45mins in dark room⇒**Rinse** by PBS 1x for 3x 5mins⇒**React** with 1% Hoechst (v/v in PBS 1x) for 10 secs at RT⇒**Rinse** by PBS 1x rapidly⇒**Mount**
Myonuclei	P-rabbit anti-dystrophin	Abcam ab15277	**Fix** in 4% (v/v) paraformadehyde for 10mins at RT ⇒**Rinse** by PBS 1x for 30mins ⇒**Block** by 5% (w/v) BSA for 30mins at RT⇒**React** with dystrophin (1:500) for 1hr at RT ⇒**Rinse** by PBS 1x for 3x 5mins ⇒React with Dylight 488 donkey anti rabbit (1:400) for 45mins in dark room⇒**Rinse** by PBS 1x for 3x 5mins⇒**React** with 1% Hoechst (v/v in PBS 1x) for 10 secs at RT⇒**Rinse** by PBS 1x rapidly⇒**Mount**
Stromal nuclei	M-rat anti-laminin 2a	Abcam ab11576	**Fix** in cold acetone for 10mins at 4°C ⇒**Rinse** by PBS 1x for 3x 5mins ⇒ **Block** by 5% (w/v) BSA for 30mins at RT⇒**React** with Laminin 2a (1:200) for 1hr at RT ⇒**Rinse** by PBS 1x for 3x 5mins ⇒**React** with Dylight 549 donkey anti rat (1:400) for 45mins in dark room⇒**Rinse** by PBS 1x for 3x 5mins⇒**React** with 1% Hoechst (v/v in PBS 1x) for 10 secs at RT⇒**Rinse** by PBS 1x rapidly⇒**Mount**

CD31: platelet endothelial cell adhesion molecule-1; Pax7: anti-paired box protein 7; DSHB: Developmental Studies Hybridoma Bank (USA)

- Apoptotic nuclei

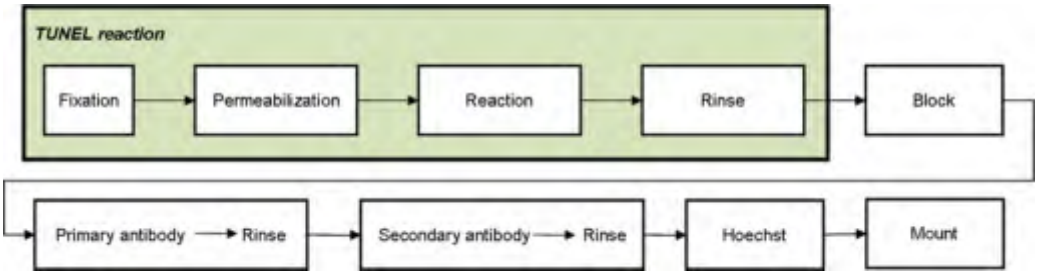


Figure 2.7 Flow-chart of the general procedure for detecting apoptotic nuclei. Comparing to the single immunolabeling (Figure 2.6), here the TUNEL reaction has been introduced

To measure the apoptotic nuclei, terminal deoxynucleotidyl transferase dUTP nick end labeling (TUNEL) was introduced because this reaction identifies the DNA strand breaks generated during apoptosis (instruction of manufacture). Although this technique has been widely used, Pulkkanen and co-worker [65] observed the false-positive result. To improve the accuracy of apoptotic measurement, apoptotic nuclei was confirmed by combining the labels of TUNEL and Hoechst 33258 in my study. To locate those apoptotic nuclei, I added Pax7, CD31, dystrophin and laminin immunostaining to distinguish apoptotic SCs (overlapped Pax7, Hoechst 33258 and TUNEL), apoptotic capillary endothelial cells (overlapped CD31, Hoechst 33258 and TUNEL), apoptotic myonuclei (overlapped Hoechst 33258 and TUNEL beneath dystrophin) and apoptotic stromal nuclei (overlapped Hoechst 33258 and TUNEL external laminin 2α) as shown in Protocol N°7 (Table 2.4). For TUNEL reaction, one negative and one positive control were performed in each experiment. For negative control, I used label solution (provided by the kit) instead of the TUNEL reaction solution; false positive results would present

DNA cleavage due to the extensive DNA fragmentation. For positive control, I incubated the slides with DNase I recombinant (Roche, 04-536-282-001 10U/ml) at 25°C for 10mins before TUNEL reaction; false negative results would not detect any DNA cleavage because it may be absent or incomplete in some forms of apoptotic cell death (instruction of manufacture). Here the DNase I stock solution was prepared according to the manufacture's instruction: DNase I (10000U/ml) was diluted by the mixture of 20mM Tris-HCl, 50mM NaCl, 100 μg /ml BSA, 1mM dithiothreitol and 50% (v/v, in glycocel). Before assay, this DNase I stock solution (5000U/ml) was further diluted to 500U/ml with 50mM Tris-HCl containing 1mg/ml BSA.

Table 2.4 (Protocol N°7) Immunostaining of apoptotic nuclei

Target	Protocol
Satellite cells nuclei	**Fix** in 4% (v/v) paraformadehyde for 20mins at RT ⇒ **Rinse** by PBS 1x for 30mins ⇒ **Permeabilize** by permeabilisation solution for 2mins at 4°C ⇒ **React** with TUNEL reacting solution for 1hr at 37°C ⇒ **Rinse** by PBS 1x for 30mins ⇒ **Block** by 5% (w/v) BSA for 30mins at RT ⇒ **Block** by 10% (v/v, in PBS 1x) unconjugated affinipure Fab fragment goat anti mouse IgG (H+L) for 1hr at RT ⇒ **Rinse** by PBS 1x for 3x 5mins ⇒ **React** with Pax7 for 1hr at RT ⇒ **Rinse** by PBS 1x for 3x 5mins ⇒ **React** with dylight 549 donkey anti mouse for 45mins in dark room ⇒ **Rinse** by PBS 1x for 3x 5mins ⇒ **React** with 1% Hoechst (v/v in PBS 1x) for 10 secs at RT ⇒ **Rinse** by PBS 1x rapidly ⇒ **Mount**
Capillary endothelial cells nuclei	**Fix** in 4% (v/v) paraformadehyde for 20mins at RT ⇒ **Rinse** by PBS 1x for 30mins ⇒ **Permeabilize** by permeabilisation solution for 2mins at 4°C ⇒ **React** with TUNEL reacting solution for 1hr at 37°C ⇒ **Rinse** by PBS 1x for 30mins ⇒ **Block** by 5% (w/v) BSA for 30mins at RT ⇒ **React** with CD31 for 1hr at RT ⇒ **Rinse** by PBS 1x for 3x 5mins ⇒ **React** with dylight 549 donkey anti rat for 45mins in dark room ⇒ **Rinse** by PBS 1x for 3x 5mins ⇒ **React** with 1% Hoechst (v/v in PBS 1x) for 10 secs at RT ⇒ **Rinse** by PBS 1x rapidly ⇒ **Mount**
Myonuclei	**Fix** in 4% (v/v) paraformadehyde for 20mins at RT ⇒ **Rinse** by PBS 1x for 30mins ⇒ **Permeabilize** by permeabilisation solution for 2mins at 4°C ⇒ **React** with TUNEL reacting solution for 1hr at 37°C ⇒ **Rinse** by PBS 1x for 30mins ⇒ **Block** by 5% (w/v) BSA for 30mins at RT ⇒ **React** with dystrophin for 1hr at RT ⇒ **Rinse** by PBS 1x for 3x 5mins ⇒ **React** with dylight 549 donkey anti rabbit for 45mins in dark room ⇒ **Rinse** by PBS 1x for 3x 5mins ⇒ **React** with 1% Hoechst (v/v in PBS 1x) for 10 secs at RT ⇒ **Rinse** by PBS 1x rapidly ⇒ **Mount**
Stromal cell nuclei	**Fix** in 4% (v/v) paraformadehyde for 20mins at RT ⇒ **Rinse** by PBS 1x for 30mins ⇒ **Permeabilize** by permeabilisation solution for 2mins at 4°C ⇒ **React** with TUNEL reacting solution for 1hr at 37°C ⇒ **Rinse** by PBS 1x for 30mins ⇒ **Block** by 5% (w/v) BSA for 30mins at RT ⇒ **React** with laminin 2a for 1hr at RT ⇒ **Rinse** by PBS 1x for 3x 5mins ⇒ **React** with dylight 549 donkey anti rat for 45mins in dark room ⇒ **Rinse** by PBS 1x for 3x 5mins ⇒ **React** with 1% Hoechst (v/v in PBS 1x) for 10 secs at RT ⇒ **Rinse** by PBS 1x rapidly ⇒ **Mount**

Permeabilisation solution is freshly prepared by mixing 0.1% (w/v) sodium citrate and 0.1% (v/v) Triton X-100. TUNEL reaction solution is prepared according to the manufacture instruction

2.2.6. Collagen and crosslinking content

Total collagen amount was estimated by hydroxyproline (OH-Pro) content [66]. Four-hydroxyproline is only found in collagen and interestingly 14% amino acid in collagen is OH-Pro [67]. In this study, frozen GM powder (100 mg per mouse) was hydrolyzed in 2 ml HCL 6N, overnight at 105°C, incubated with activated charcoal (Norit A, Sigma), 1 ml was diluted with 4 vol H2O. OH-Pro content was then determined according to the procedure of Woessner [68], by measuring optical density, at 557 nm.

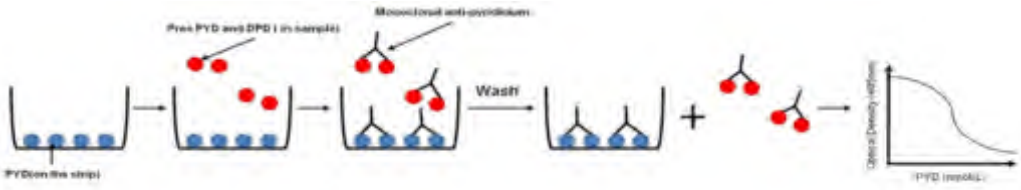


Figure 2.8 Flow-chart of the competitive enzyme immunoassay procedure

Total pyridinoline crosslinking contents (PYD) and deoxypridinoline (DPD) in skeletal collagen were measured by a competitive enzyme immunoassay in a microtiter strip format utilizing a monoclonal anti-pyridinium crosslinking antibody (Metra PYD, Quidel). This product is developed for urines and its principle is introduced in Figure 2-8. Briefly, the PYD and DPD in the sample compete for the antibody with PYD coated on the strip. The reaction is then detected with a pNPP substrate (p-Nitrophenyl phosphate) and analyzed using standard curve. In this study, I adapted the manufacture instruction to skeletal muscle study, as shown in protocol N°8 (Figure 2.9).

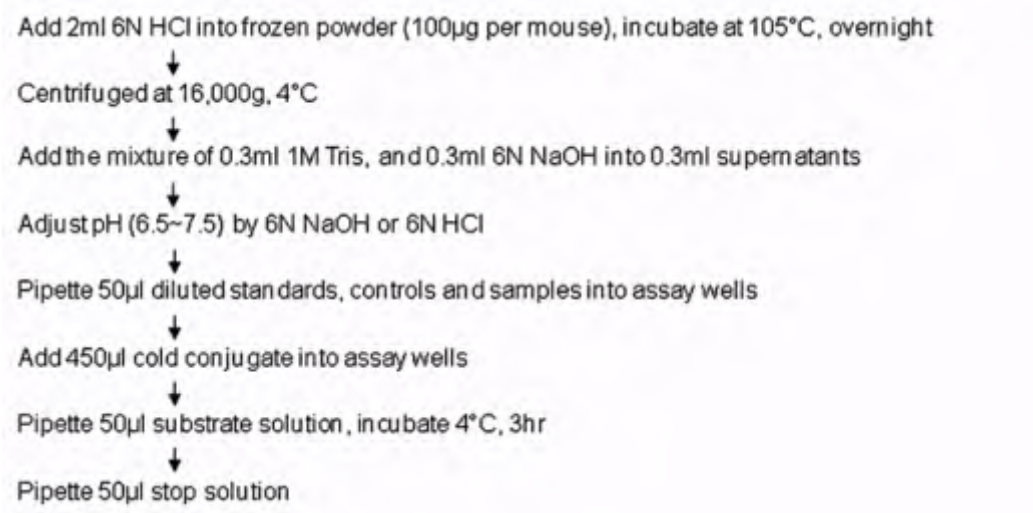


Figure 2.9 (Protocol N°9) Competitive enzyme immunoassay to measure pyridinium crosslinking

2.2.7 Endorepellin

Protein extraction:

- Mix frozen powder (40mg per mouse) with 25v extraction buffer (8M Urea, 2M Thiourea, 2% Chaps , 1% DTT)
- Homogenize hydrolysate with a Tissue Ruptor (Qiagen) at full speed for 15 s and centrifuge at 10,000g, 4°C, 30mins
- Determine the protein concentration using Bradford method*

SDS-PAGE gel preparation and electrophoresis:

- Prepare SDS-PAGE gel (10%)
- Load prestained molecular weighted standards markers
- Load protein sample (30µg per mouse)
- Migrate samples using a Mini-protean® 3 Cell system (BioRad) at 170V, 75mins
- Perform electrotransfer to PVDF membranes at 100V, 90mins

Immunodetection:

- Block membrane with 5% milk/TBS solution, 1hr
- Wash membrane with TTBS 1x, 3x5mins
- Incubate with primary anti-endorepellin (1:200, diluted with TTBS 1X and 5% milk), 4°C , overnight
- Wash with TPBS 1x and 5% milk, 3x5mins
- Incubate with secondary anti-mouse conjugated to HRP (1:5000) 1hr
- Wash with PBS 1x, 2x10mins

Figure 2.10 (Protocol N°10) Endorepellin was detected by SDS-PAGE coupled with Western blotting. UREA, thiourea, chaps, "DTT": dithiothreitol, SDS-PAGE gels, "PVDF membrane": polyvinylidene fluoride membrane, "TBS": tris buffered saline, "TTBS": tween / tris buffered salt solution, "TPBS": BupH Phosphate Buffered Saline.

The presence and relative abundance of skeletal endorepellin is determined by sodium dodecyl sulfate polyarcylamide gel electrophoresis

(SDS-PAGE) coupled with Western blotting. When protein samples and standard protein with known size are electrophoresed on an SDS-PAGE, they are largely separated on the basis of distinct molecular weight. After electrotransfer onto polyvinylidene fluoride (PVDF) membrane, the protein is detected using specific primary antibody and secondary enzyme labelled antibody and substrate. In this study, protein from frozen powder (40mg per mouse) at 11, 22 and 25M were extracted and then separated by SDS-PAGE. After electrotransferring, PVDF membranes were probed with monoclonal anti-endorepellin (Abcam, A74), and with secondary anti-mouse conjugated to HRP. Visualization of the immune complexes utilized ECL detection system (Amersham, Arlington Heights, IL) and subsequent exposure to X-ray film (Amersham, Arlington Heights, IL) (Figure 2.10) [69].

2.2.8 Image analysis

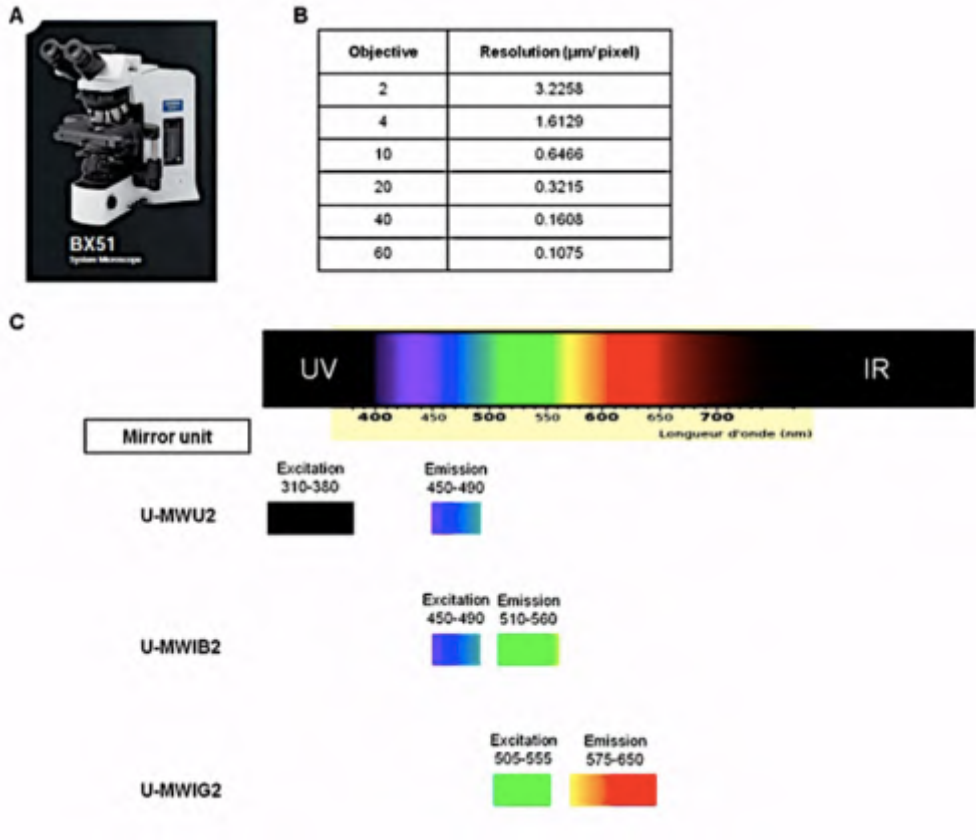


Figure 2.11 Optical microscopy for examining the structure of muscle, (A) Olympus BX-51 microscope (B) Resolution and (C) Excitation and Emission filter

Images were captured with an Olympus DP-72 camera coupled to an Olympus BX-51 microscope (Figure 2.11A) at a resolution of 0.64 or 0.32 µm/pixel according to the distinct objective (Figure 2.11B). For sections treated with Sirius red and azorubin, five TIFF-RGB color images per mouse (~1000 myofibres) were acquired under identical conditions (exposure time; 4ms) in bright fields. For section treated by immunostaining, namely single immunelabeling (section 2.2.4) to study ECM component and multiple immunolabeling to study nuclei distribution (Hoechst 33258 with laminin 2α, dystrophin, CD31, Pax7, and Hoechst 33258 coupled with TUNEL), ten images

per mouse were captured under optimal condition (exposure time) through adequate filters (Figure 2.12C): The blue one (excitation at 345nm, emission at 450-490nm) was used for Hoechst 33258, the green one (excitation at 470nm, emission at 510-560nm) was used for Dylight 488 and FITC-conjugated secondary antibodies and the red one (excitation at 530nm, emission at 575-650nm) was used for Dylight 549 and Alexa 546-conjugated secondary antibodies.

For Confocal microscopy, three to five images per mouse were captured using a Leica SP5 spectral detection Confocal microscope at PLATIM (Plateau Technique d'Imagerie / Microscopie, IFR128, Lyon Gerland). For Dylight 488 observations, laser at 488 nm (30% power, 40 % AOTF) was used for excitation, and PMT detection was performed between 500 to 541 nm. For Dylight 549 and Alexa 546, He/Ne laser was adjusted to 561 nm (100% power) and emission was analyzed from 578 to 649 nm.

- Images stained by azorubin

Images stained by azorubin were processed through homemade visual basic program developed under Visilog 6.7 software (Noesis, Gif sur Yvette, France). For azorubin stain, the original RGB image was converted to a grayscale luminance image. After threshold, the watershed algorithm was applied on the binarized image to segment each fibre. Next, the **CSA** (**A_Fib,** Table 3.1) and **perimeter** (**P_Fib,** Table 3.1) were measured directly, which was visualized by random color in the final image. The **number of myofibre**

per unit area (**N_Fib,** Table 3.1) was calculated. In each step, manual editing was introduced to improve the fibre segmentation.

- Images stained by Sirius red

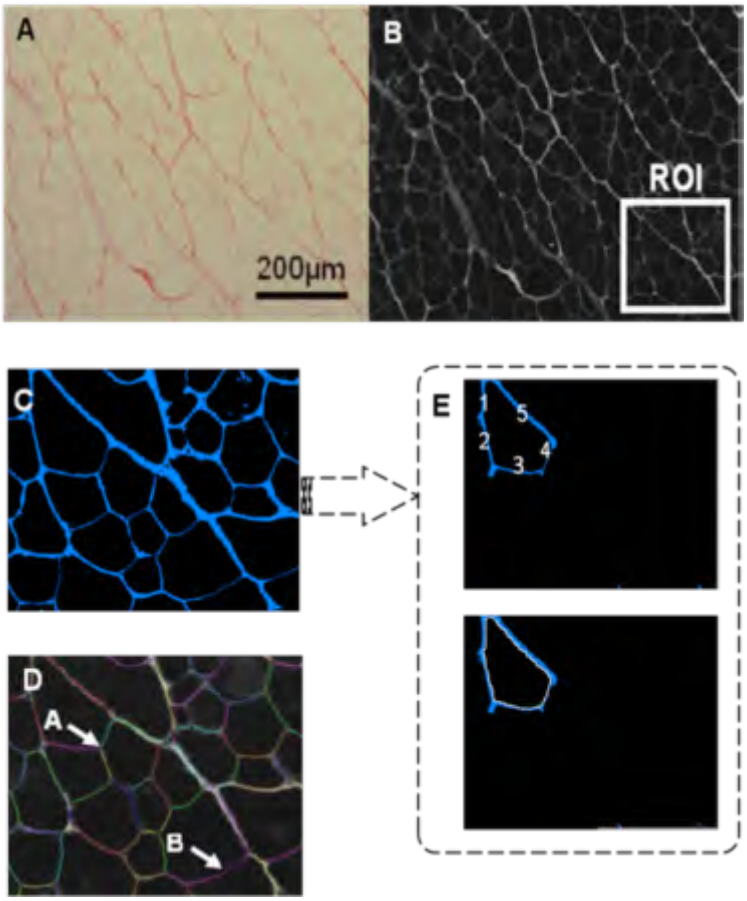


Figure 2.12 Quantifying the structure of myofibres from sirius red stained micrographs: (A) Original TIFF-RGB images, (B) Inverted optical density (OD) grayscale images and a representative region of interest (ROI). (C) Binarized images showing ECM in blue (equal to 1 in the ROI). This image was obtained from (B) by removing background and threshold. (D) images generated by masking (B) and (C). Connection points (rather use arrow) and ECM segments (rather use arrow head) were automatically detected. (E) Binarized myofibre image was reconstructed to calculate the number of adjacent ECM segment per myofibre (1 to 5), as shown in above image.

Images from Sirius red stained sections were also processed through a homemade visual basic program developed under Visilog 6.7. Firstly, green component of original TIFF-RGB color images (Figure 2.12A) was selected and

converted into inverted optical density (OD) grayscale image (Figure 2.12B). Background of this grayscale image was subtracted before threshold and generated the binarized image (Figure 2.12C) where ECM (in blue) referred to as 1 and background (in black) referred to as 0. This image was used to mask OD images, as shown in Figure 2.12D.

For image processing, several parameters of ECM structure were defined and calculated (see Table 3.1). On Figure 2.12C, **ECM area** (**A_ECM**) was measured as the area of blue part. Image erosion[2] was carried out to extract the skeleton of ECM. Thus on Figure 2.12C, the ECM pattern became thinner (width=1 pixel), here, the **ECM length** (**L_ECM**) was measured. On Figure 2.12D, **ECM volume** (**V_ECM**) was determined by accounting the total pixel value. This parameter was used to characterize the ECM content on the cross-section. Then the **ECM intensity** (**I_ECM**) was determined by dividing the V_ECM by A_ECM. It was used to evaluate the concentration of ECM on the cross-section. Common links of skeleton (≥ 3) was marked and represented as **connection points** (**CP**, arrow in Figure 2.12D). The number of connection

---

[2] Erosion: a basic morphological operations during image processing, referring to remove the pixels on object boundaries.

[i] Vignetting correction: due to the optical defect, I obtain the images in which the borders are darker than the center. Hence, I introduced the model proposed by Kang and Weiss to correct the vignetting

[ii] *Sobel* filter: Sobel is a discrete differentiation operator that is used in computation of the gradient of image intensity at each point. By using Sobel filter, edges in the result image will highlight the part which has a difference in intensity

[iii] *Canny* filter: an edge detector based on the first derivative of a Gaussian

point per myofibre (**N_CP**) was measured accordingly. ECM skeleton segments (arrow head in Figure 2.12D) linking two CP were then identified. The quantity and mean length of ECM segments were calculated and named **N_Seg** and **L_Seg**, respectively.

In addition to ECM structure, the morphology of myofibres was studied by processing images stained with Sirius red. As an example Figure 2.12E shows (from top to bottom) how each myofibre and its surrounding ECM segments were identified and reconstructed in a new frame. During the reconstruction, the **number of segments per myofibre** (**X3Seg to X7Seg,** Table 3.1) was recorded and used to characterize the shape of myofibre. In addition, the area of each myofibre was estimated by carrying out erosion with fixed value (3 pixels) (Figure 2.13E). The information was collected for 600-1100 individual myofibres per mouse and for 4-5 mice per age.

- Images stained by laminin 2α and Hoechst 33258

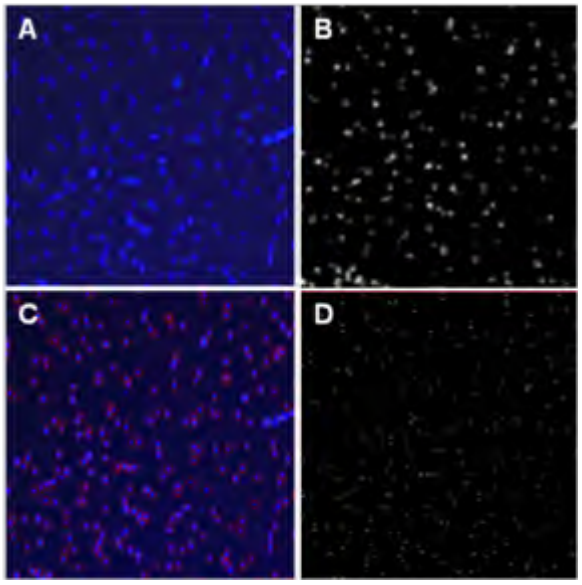


Figure 2.13 Processing of images obtained from Hoechst staining. (A) Original RGB-TIFF image. (B) Grayscale image, nuclei were segmented; (C) Image was generated by masking (A) and (B). Nuclei were detected and mapped. (D) Image was generated by limiting nuclei area in (B), which would be discussed in the later sections.

Images from sections stained for laminin 2α and with Hoechst 33258 were processed with Matlab version 2007b (Mathworks Inc) to distinguish muscular nuclei. For the nuclear images stained with Hoechst 33258 (Figure 2.13A): I performed Vignetting correction[i] on the grayscale images as a pre-processing step (Figure 2.13B). Next, *Sobel* filter[ii] was used to detect the nuclear edges (Figure 2.13C). Erosion was done as a post-processing step to remove the connected components. Artifacts on the eroded image were further removed by limiting area condition (area ≤1 unit2). The final nuclear image was a result of mapping the segmented regions onto the original images displayed in Figure 2.13D in which nuclear region was given a value "1".

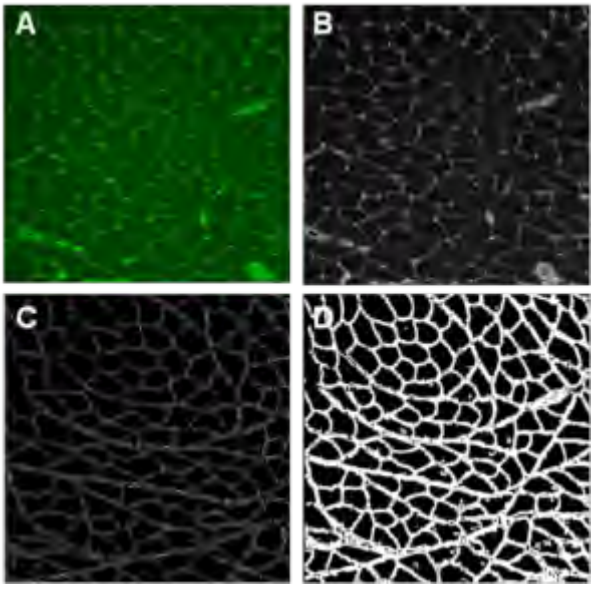


Figure 2.14 Processing of image obtained from laminin 2α staining. (A) Original RGB-TIFF image. (B) Grayscale image, with vignetting corrected. (c) Grayscale image, edge of ECM was detected by *canny* filter. (D) Grayscale image, ECM edges were dilated.

For the ECM images stained for laminin 2α: I converted the original images (Figure 2.14A) into grayscale and before applying Vignetting correction (Figure 2.14B). *Canny* filter[iii] was then applied to detect the edge of ECM (Figure 2.14C). Next, a dilation algorithm was applied to stitch up the connected points (Figure 2.14D). In the resulting image, ECM region was given a value "2".

Thereafter, the processed nuclear (Figure 2.13D) and ECM (Figure 2.14D) images were mapped, leading to the result images carrying integer value from 0 to 3: background was presented as a value of "0"; central nuclei were recognized with a value of "1"; ECM surrounding each myofibre were highlighted by a value of "2" and other nuclei (myonuclei, satellite cells and stromal nuclei) as a value of "3". Hence, the number of **total muscular nuclei**

per unit area (**N_Nu,** Table 3.1) and the **number of central nuclei** per unit area (**N_CenNu,** Table 3.1) were calculated.

- Images stained by antibodies labeling ECM components

Images from sections stained by immunohistochemistry to investigate the variation of ECM components were processed with ImageJ version 1.42q (NIH). For each image, pixel intensity was randomly collected in 10 endomysial regions and 10 perimysial regions.

2.2.9 Statistical analyses

Statistical analyses of each dependent variable were carried out using one-way ANOVA for data that satisfied the criteria of normality and homogeneity otherwise non-parametric Kruskal-Wallis one-way analysis of variance was employed. Multiple comparisons of the honestly significant differences (HSD) were accessed using the Turkey test with statistical significance accepted at $p<0.05$.

2.3 Results

2.3.1 Body composition

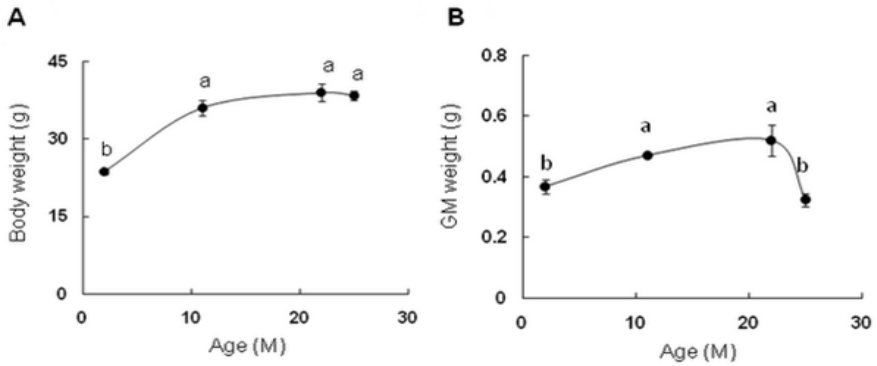


Figure 2.15 Graphs of the (A) body weight of the C57BL/6 mouse and (B) GM weight versus age. For the quantity of mice per age see Table 1.6. The magnitudes of the body and GM weight are represented by mean ± SEM. Different letters indicate significant difference (p<0.05) between ages.

It is observed that body and GM weight varied significantly with age. As shown in Figure 2.15A, body weight increased between 2M and 11M and stabilizes from 11M to 25M. GM weight (Figure 2.15B) similarly rose between 2M and 11M, was maintained between 11M and 22M but strongly decreased from 22M to 25M.

2.3.2 Myofibre morphology and metabolism

2.3.2.1 Myofibre atrophy

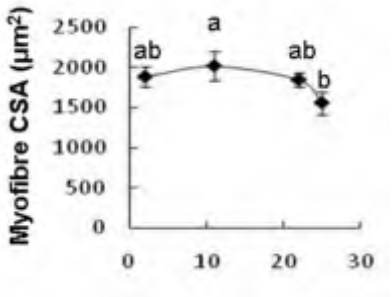


Figure 2.16 Myofibre CSA (µm^2) versus age and the data was obtained from C57BL/6 mice at 2, 11, 22 and 25M (4 mice per age). The magnitudes of the myofibre cross sectional are represented as mean ± SEM (vertical bars). Different letters indicate significant difference (p<0.05) between ages.

Age-dependent muscle atrophy was associated with modifications in myofibre morphology. Mean value of the CSA for all GM myofibres is observed to decrease (but not appreciably) by 22% between 11M and 25M muscle (Figure 2.16), although not significantly.

2.3.2.2 Myofibre morphology

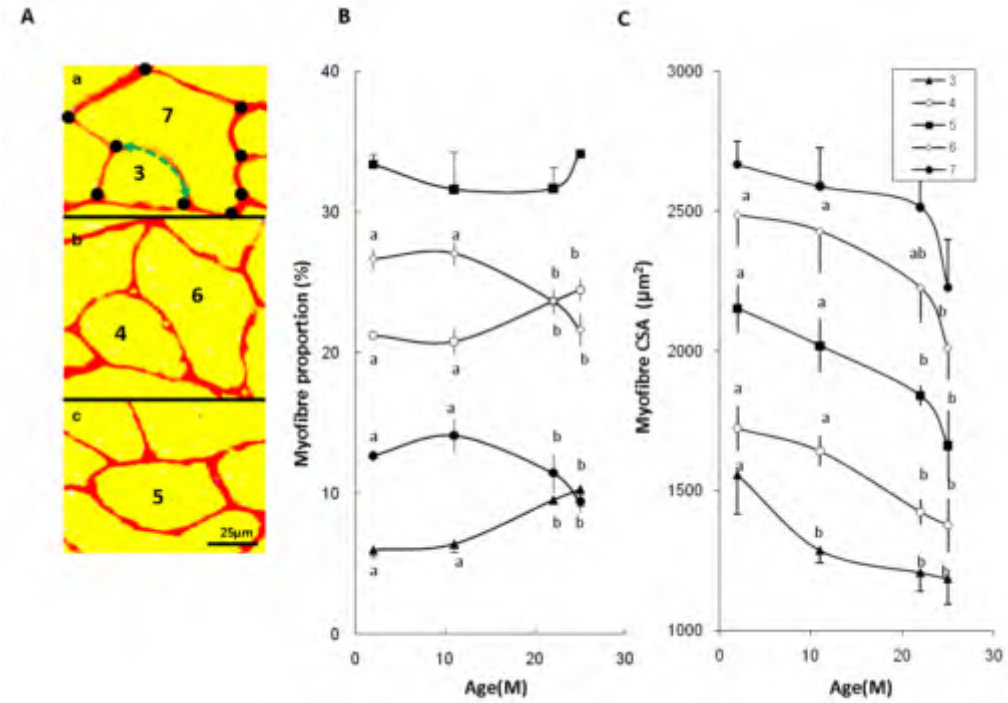


Figure 2.17 Myofibre morphology: (A) Cross section of GM in 11M-old mouse with ECM stained by sirius red. Connection points (dots) and one example of ECM segment (green arrow) were labeled in the top image. Numbers "3" to "7" refers to the number of adjacent ECM segments for the current myofibre. (B) Age-related variation of myofibre area with different number of segments (3-7 in Figure 2.17A). (C) Cross-sectional area of myofibres with different number of segments (3-7 in Figure 2.17A). Values are means ± SEM. Different letters indicate significant difference (p<0.05) between ages.

To further improve the analysis, I classified the myofibres according to the number of adjacent ECM segments (i.e. the number of neighborhood myofibres) (Figure 2.17A). As expected, larger myofibres is associated with higher number of segments, and as shown in Figure 2.17C, each class of myofibre exhibited age-dependent decrease in the CSA. I have observed that myofibres with 5 segments, so called regular myofibres, were the most abundant (31-34%, Figure 2.17B) in all the age groups. However during aging (11M to 25M), the proportion of 6- and 7-segment myofibres, so called

polygonal myofibres, decreased, while symmetrically the 4- and 3-segment myofibres, so called angular myofibres, increased, indicating that myofibres became more angular.

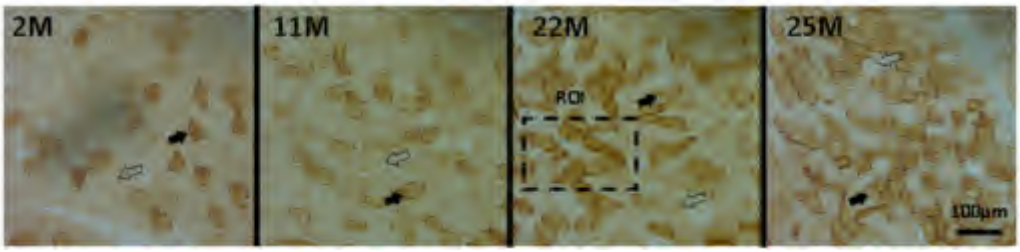


Figure 2.18 Myofibre metabolism: Cross section of mouse GM in 2, 11, 22 and 25M old with oxidative type I myofibre (solid arrow) stained in dark brown while glycolytic type II myofibre (hollow arrow) stained in light brown. Region of interest (ROI) from section of GM in 22M-old highlights the approaching of cox-positive myofibres (oxidative, type I).

Cox, which is a marker for oxidative energy metabolism characterizes slow contracting type I myofibres [73]. In GM of mice, Cox preferentially labeled the small 3- to 5-segment myofibres for all the ages. As shown in Figure 2.18, Cox-labeled myofibres yielded a checkerboard -like distribution in the cross-sections of 2M and 11M GM, while approaching of cox-positive myofibres was clearly visible in 22M and 25M muscles.

2.3.3 ECM

2.3.3.1 ECM structure

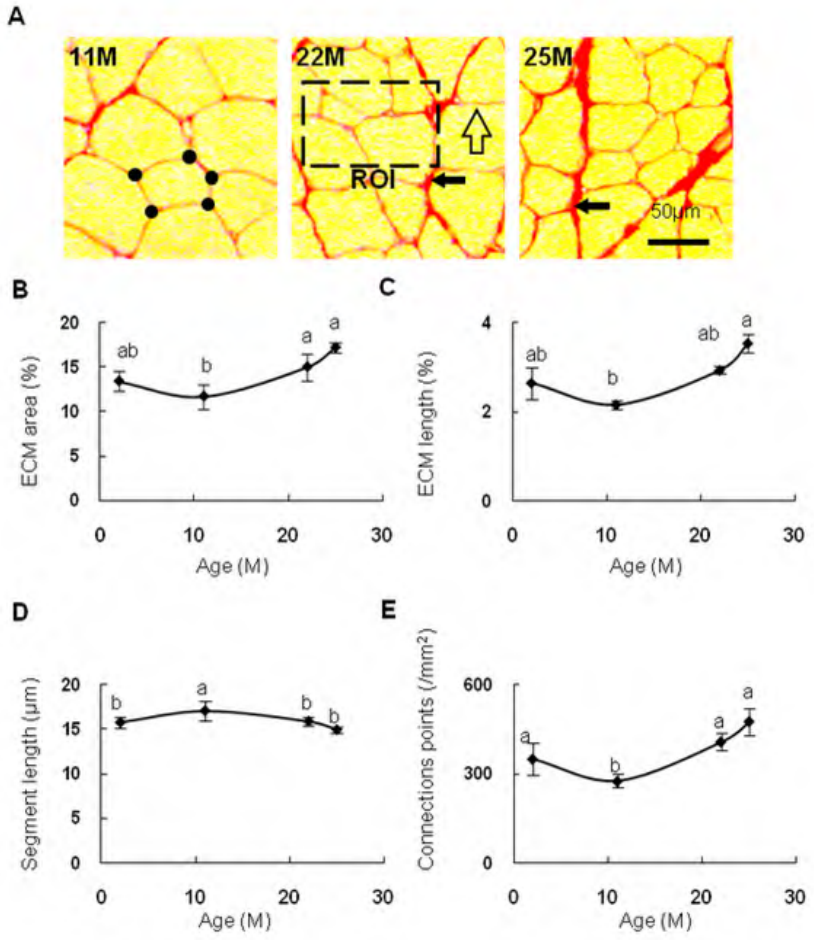


Figure 2.19 Quantifying ECM microstructure (A) Cross-section of GM (11M) showing the stained ECM using sirius red. The dots represent connection points. Region of interest (ROI) in section from the GM (22M) encloses a myofibre with irregular morphology. Thin (hollow arrow) and thick (solid arrow) ECM segments are observed in the GM cross-sections (22Mand 25M). Age-related variations in the (B) ECM area (%), (C) ECM length (%), (D) mean value of ECM segment length (⌈m) and (E) number of connection points per unit area (mm^2). Values are means ± SEM. Different letters indicate significant difference (p<0.05) between ages.

Alterations in myofibre size and morphology suggest a profound remodeling in the connective tissue embedding myofibres. In order to characterize ECM modifications, image analysis was performed with muscle cross-sections stained by Sirius red (Figure 2.19A). The total area (Figure

2.19B) and total length (Figure 2.19C) of the ECM skeleton both showed similar U-shaped evolutions ($p<0.05$) with a minimum value at 11M, and a maximum at 25M.

In order to elucidate the enlarged ECM area and length, I analyzed the segment between pair of myofibres and computed the number of CP. It is observed that the average length of the segments slightly decreased during early aging (Figure 2.19D), while the number of CP (where ≥ 3 myofibres interact) expanded during late adulthood (22M) and aging (25M) (Figure 2.19E). Segment is defined by the ECM skeleton between 2 neighbors CPs; therefore in the old muscle, longer ECM was characterized by shorter but more abundant segments.

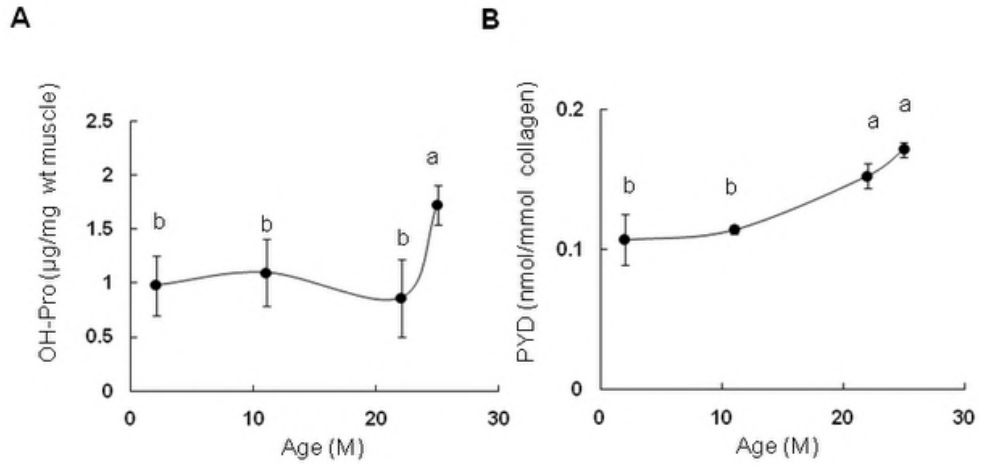


Figure 2.20 Age-related variations of fibrillar collagen content parameterized by (A) OH-Pro content (µg / mg wt muscle) and (B) Pyridinolin crosslinking (nmol /mmol collagen). The magnitudes of the parameter on the vertical axes represent mean ± SEM. Different letters indicate significant difference ($p<0.05$) between ages.

The muscle content of OH-Pro (Figure 2.20A) fluctuated from 2M to 22M; thereafter a dramatic increase is observed at 25M. This could be attributed

to ECM fibrosis in old muscle. It is further observed that the dramatic increase in OH-Pro corresponds to an increase in non-reducible collagen crosslinking (Figure 2.20B). An increase in non-reducible collagen crosslinking could influence that macro-tissue mechanical property, resulting in an increase in the stiffness of tissue [74].

2.3.3.2 ECM component and co-localization

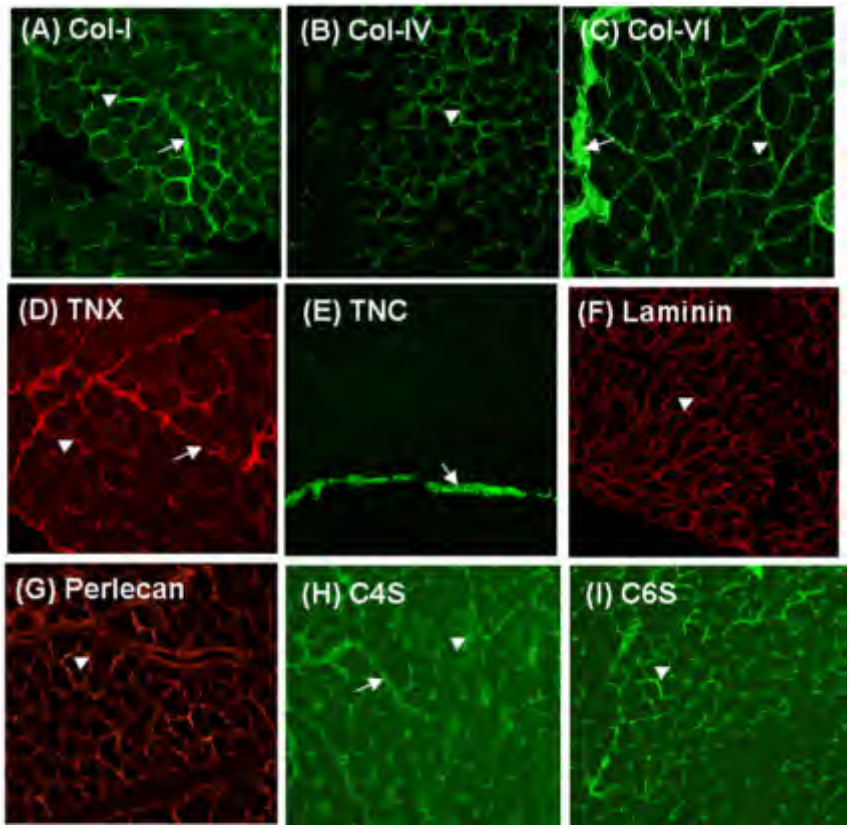


Figure 2.21 Cross-section of mouse GM in 11M or 25M old age with different ECM components stained with specific antibodies. (A) type I collagen (Col-I), (B) type IV collagen (Col-IV), (C) type VI collagen (Col-VI), (D) Tenascin X (TNX), (E) Tenascin C (TNC), (F) Laminin 2a, (G) Perlecan, (H) chondroitin-4-sulfate (C4S) and (I) chondroitin-6-sulfate (C6S). Arrow refers to perimysium and arrowhead refers to endomysium.

Eleven ECM components were investigated using the single immunolabeling technique and complemented by semi-quantitative measurement (Figure 2.21). It is observed that (1) Col-I, Col-III, Col-VI (Table

1.2), TNX (section 1.2.2.2) and C4S (Table 1.3) were presented in both endomysium and perimysium; (2) Col-IV (Table 1.2), laminin 2α (section 1.2.2.2), perlecan (Table 1.3) and C6S (Table 1.3) were present only in endomysium; (3) TNC (section 1.2.2.2) was present in major perimysium and epimysium. It is also observed that the average pixel intensity of all the eleven ECM components in perimysium remained unchanged for all the age groups. It is further observed that the average pixel intensity of three components: TNX, Col-VI and perlecan, in endomysium varied at different ages (Figure 2.22), while those of the other eight components remained unchanged.

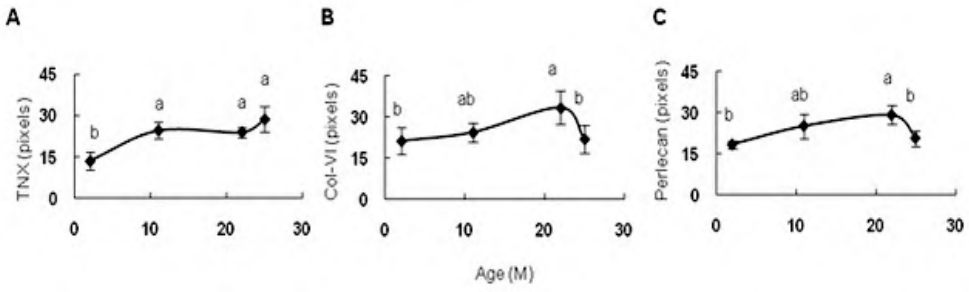


Figure 2.22 Pixel intensity of ECM components in endomysium of mouse GM in 2, 11, 22 and 25M old age. Age related variations are represented by the number of pixels counted for (A) TNX, (B) Col-VI and (C) Perlecan, values are presented as mean ± SEM. Different letters indicate significant difference (p<0.05) between ages.

Results from semi-quantitative immunohistochemistry revealed that age-related modifications for TNX, Col-VI and perlecan immunolabeling in the endomysium but with distinct timings. In particular, TNX increased between 2M and 11M (Figure 2.22A) and remained stable thereafter. In contrast, Col-VI (Figure 2.22B) and perlecan (Figure 2.22C) progressively rose to reach a maximum at 22M, and significantly dropped down at 25M.

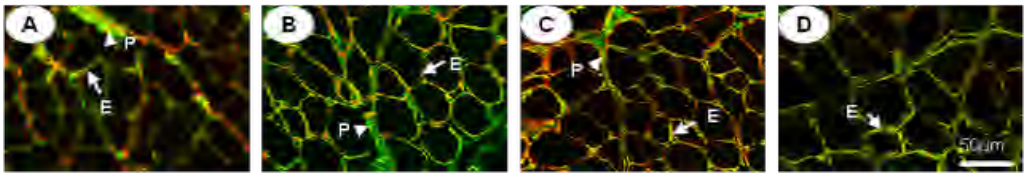


Figure 2.23 Co-localization of ECM components in mouse GM in 11M-old. Cross sections of (A) Col-I (green) with TNX (red); (B) Col-VI (green) with Perlecan (red); (C) Col-VI (green) with Laminin 2α (red); (D) Col-IV (green) with Laminin 2α (red). Endomysium (E) and perimysium (P) was marked by solid arrow and arrowhead respectively. Co-localizations were visualized as yellow signal.

For investigating the changes in the ECM at the molecular level, several ECM components were assessed for their co-localization. Col-IV, perlecan and laminin 2α are known to be exclusively located in the endomysial basement membrane, while Col-VI and TNX are localized both in the endomysium and the perimysium [12]. Previous studies also demonstrated specific interactions between ECM components. Accordingly using Confocal microscopy, I confirmed the co-localizations of Col-VI and perlecan, Col-VI and laminin 2α, and laminin 2α and Col-IV in the endomysium of mice muscles (Figure 2.23). During aging, these co-localizations did not vary.

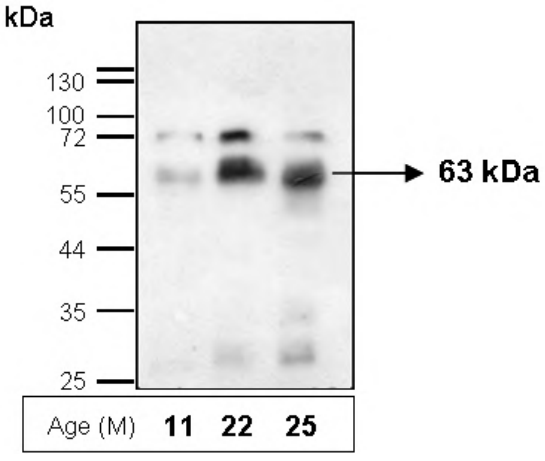


Figure 2.24 Expression of endorepellin fragments in mouse GM at 11, 22 and 25M

Unfortunately, my attempts to use western blots of total muscle extracts with a specific antibody to Col-VI or TNX were unsuccessful. Interestingly, my attempts at using western blots of the total muscle extracts with a specific antibody to perlecan domain V (endorepellin) yield results indicating a 63 kDa fragment (Figure 2.24) corresponding to endorepellin LG1-LG2 fragment [75]. Interestingly these experiments revealed that GM aging was associated with a sharp increase in this endorepellin 63 kDa fragment.

2.3.4 Muscular nuclei

Skeletal muscles comprise different cells: multinucleated myofibres, SCs and stromal cells of the connective tissue. SCs are located between the myofibre sarcolemma and the basal lamina (section 1.2.1), while stromal cells are located outside the basal lamina. Further studies were then performed to establish the influence of aging on the specific cellular population. The results are presented in the following sub-sections.

2.3.4.1 Nuclei distribution

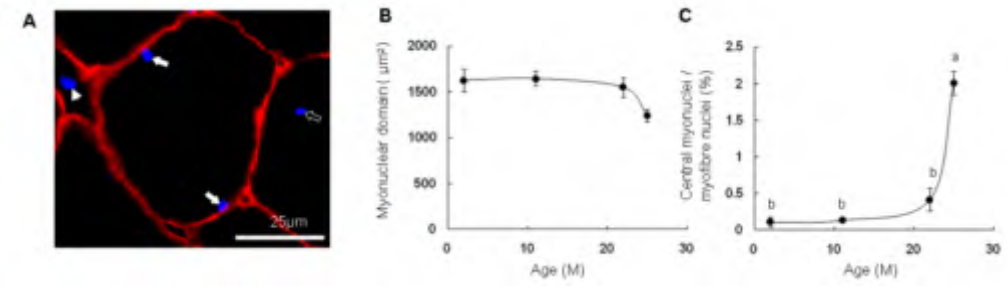


Figure 2.25 Age-related variation of myonuclei. (A) Cross section of GM in 25M-old mouse with muscular nuclei stained by Hoechst 33258 (blue). Laminin 2α (red) shows the basement membrane of myofibre. Myofibre nuclei (solid arrow), central myonuclei (hollow arrow) and nuclei in stromal cells (arrowhead) are marked. (B) Age-related variation of myonuclear domain size (μm^2). (C) Age-related variation in the proportion of central myonunclei in myofibre (%).

Values are means ± SEM. Different letters indicate significant difference (P<0.05) between ages.

Co-staining of basement membrane (by anti-laminin 2α) and nuclei (by Hoechst 33258) was used to distinguish stromal nuclei from myofibre nuclei, which contain a majority (>99%) of myonuclei and a minority (<1%) of satellite cell nuclei (see below) (Figure 2.25A). This analysis indicated that during aging myofibres maintained a similar content of myonuclei (data not shown). However, because of smaller cross-section area, the myonuclear domain (the myofibre volume controlled per myonucleus) decreased (-20%) in 25M muscles (Figure 2.25B). Although myonuclei are typically located beneath the plasmalemma of young and adult myofibres, I also observed a dramatic increase in the centrally located myonuclei in GM from older mice (Figure 2.25C), which is a recognized marker of muscle regeneration [48].

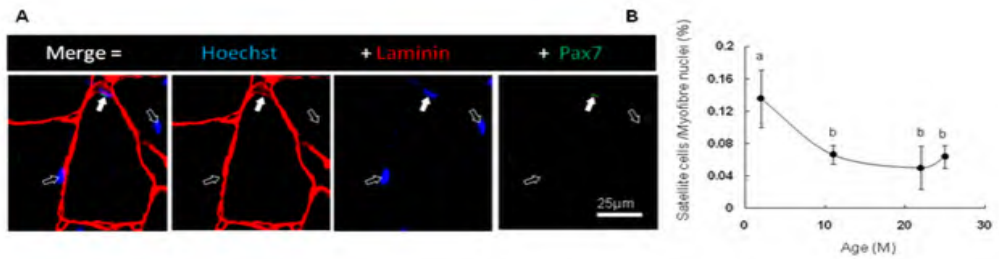


Figure 2.26 Age-related variations of SCs. (A) Cross section of GM in 2M-old mouse with nuclei stained by Hoechst (blue), BM stained by laminin 2α (red). SCs (solid arrow) were confirmed by both Pax7 (green) and Hoechst (blue) staining. Thus myonuclei (beneath BM but not labelled by Pax7) were marked by hollow arrow. (B) age-related variation of SCs in proportion of myofibre nuclei (%). Values are means ± SEM. Different letters indicate significant difference (p<0.05) between ages.

Anti-Pax7 labels both quiescent and activated satellite cells [76]. After counting the number of Pax7+ satellite cells, I found that Pax7+ satellite cells represented only a small proportion (<0.15%) of myofibre myonuclei in the GM

of young mice. This proportion of satellite cells strongly decreased appreciably (-51%) during growth (2M to 11M) and remained at a low level thereafter in adult and old muscles (Figure 2.26B).

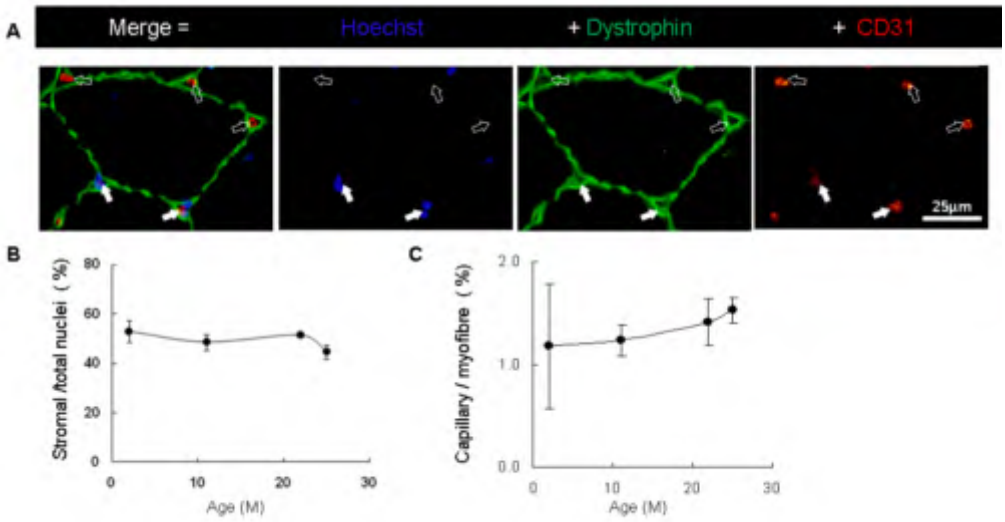


Figure 2.27 Age-related variations of capillary endothelial cells. (A) Cross section of GM in 11M-old mouse with nuclei stained by Hoechst (blue), BM stained by Dystrophin (green) and capillary endothelial cells stained by CD31 (red). Capillary endothelial nucleus (solid arrow) was confirmed by both CD31 (green) and Hoechst (blue) staining. Thus capillary endothelial cell with nucleus out of the current layer was labeled by hollow arrow. (B) Age-related variation of stromal nuclei in proportion of total muscular nuclei (%). (C) Age-related variation of capillary endothelial number per each myofibre. Values are means ± SEM. Different letters indicate significant difference ($p<0.05$) between ages.

Hoechst staining of total muscle nuclei and co-staining of the basal lamina further indicated that stromal nuclei represented a significant proportion (45-55%) of total muscle nuclei, and this proportion of stromal nuclei remained unchanged during aging (Figure 2.27B). The ECM nonetheless contains different types of stromal cells, including fibroblasts, immune cells, adipocytes and capillary cells. Anti-CD31 was introduced to label capillary endothelial cells [77]; co-labelling with anti-dystrophin confirmed that CD31 labelling exclusively occurred in the ECM (Figure 2.27A). CD31 studies revealed that

the density of capillaries around each myofibres fluctuates about a constant value during aging (Figure 2.27C).

2.3.4.2 Apoptosis

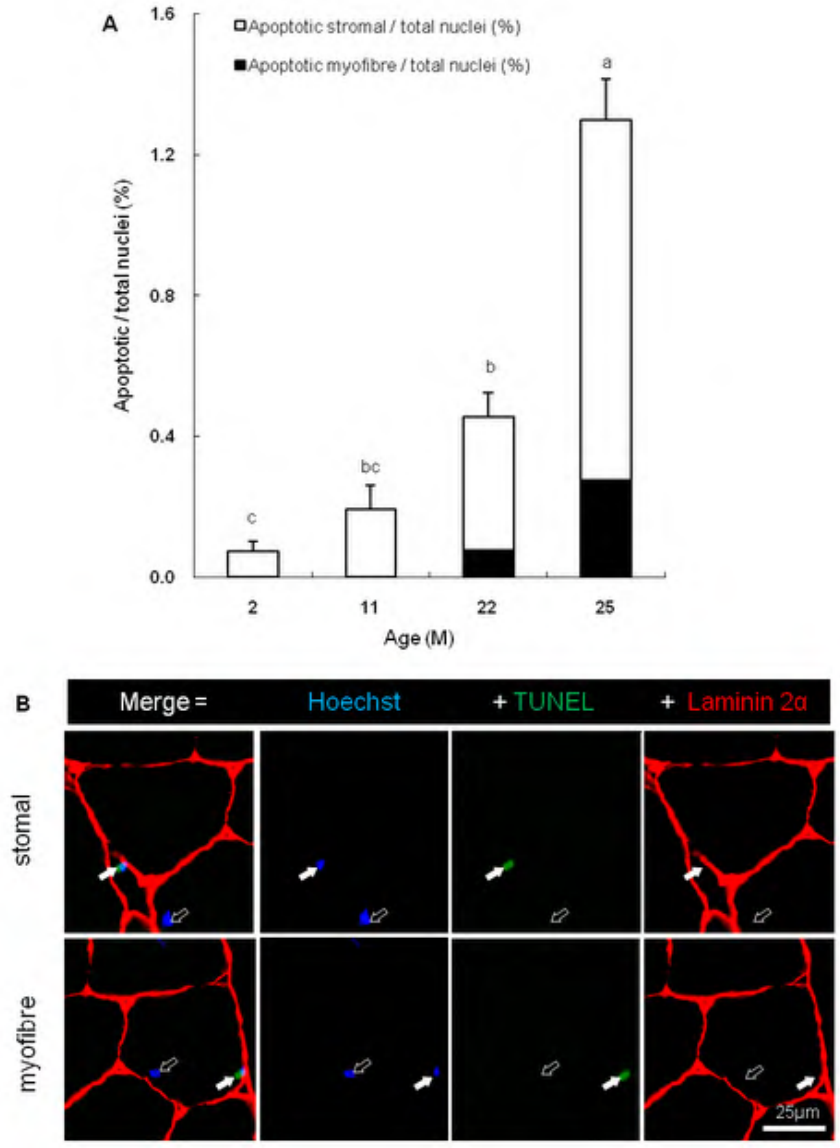


Figure 2.28 Apoptosis. (A) Age-related variations of apoptotic nuclei in stromal cells and myofibre in proportion of total muscular nuclei (%). Values are mean ± SEM and different letters indicate significant difference (p<0.05) between ages. (B) Cross section of mouse GM in 25M-old with nuclei stained by Hoechst (blue), BM stained by laminin 2a (red). Apoptotic nucleus was confirmed by co-staining of TUNEL (green) and Hoechst. Apoptotic nuclei in stromal cells (top row images) and in myofibre (bottom row images) are indicated by solid arrows; normal nuclei are marked by blank arrows.

Since the modifications in the populations of muscle nuclei that occur during aging could be related to apoptosis, further analysis using the TUNEL technique was carried out. As shown in Figure 2.28A, TUNEL-positive nuclei strongly increased in old muscles. However, the increase is less than 1.3% of total muscle nuclei. To further distinguish whether apoptosis occurred in myofibres or in stromal cells, laminin staining was performed together with TUNEL and Hoechst (Figure 2.28B). Interestingly in 2M to 11M mice, apoptotic nuclei exclusively belonged to stromal cells in the connective tissue. Moreover, the age-dependent rise in apoptotic nuclei was mostly attributed to stromal cells (Figure 2.28A).

- Apoptosis in capillary endothelial cells

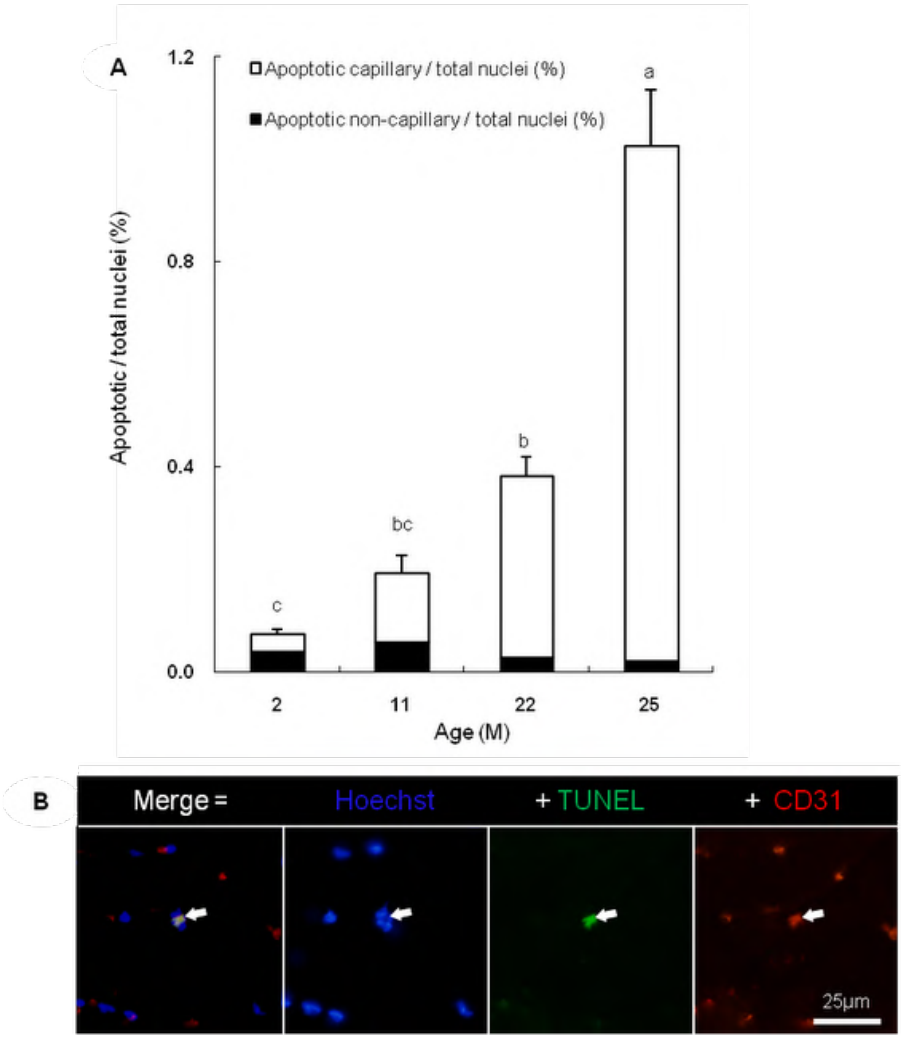


Figure 2.29 Apoptosis in capillary endothelial cells. (A) Age-related variations of apoptotic nuclei in capillary endothelial cells and other stromal cells in proportion of total muscular nuclei (%). Values are mean ± SEM and different letters indicate significant different (p<0.05). (B) Cross section of mouse GM in 25M-old with nuclei stained by Hoechst (blue), BM stained by laminin 2a (red). Apoptotic nucleus was confirmed by co-staining of TUNEL (green) and Hoechst. Apoptotic nucleus in capillary endothelial cells (solid arrow) was marked.

Further analysis of apoptosis in stromal cells was carried out using triple labelling of muscle cross-sections with TUNEL, Hoechst and CD31 to label apoptotic nuclei, total nuclei and capillary endothelial cells, respectively. These

analyses indicated that the age-dependent rise in stromal cell death was mostly due to apoptosis of capillary endothelial cells (Figure 2.29).

- Apoptosis in myofibre and SCs

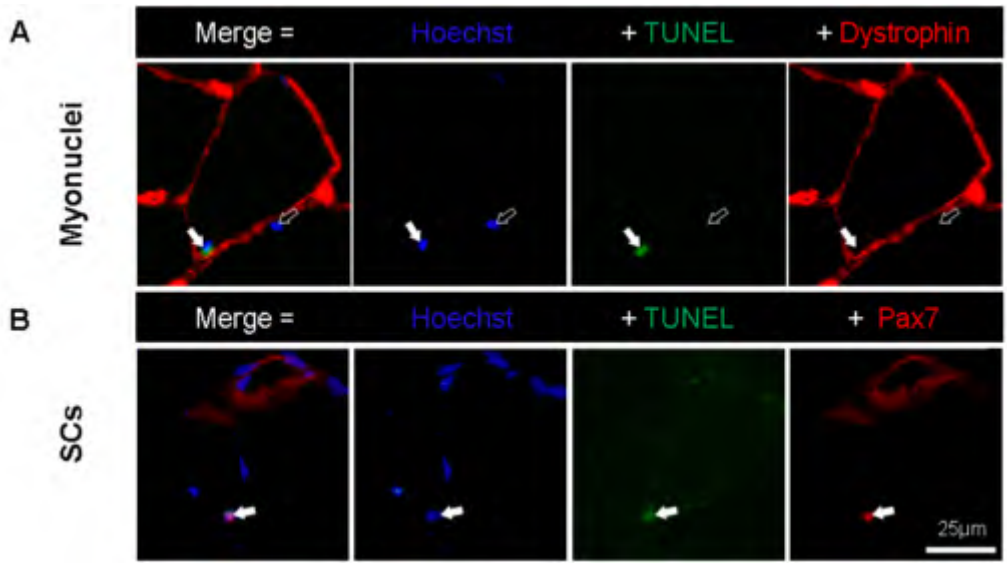


Figure 2.30 Apoptosis in myofibre and SCs. Cross section of GM (25M) showing nuclei stained by Hoechst (blue), sarcolemma stained by dystrophin (red). (A) Apoptotic myonucleus (solid arrow) was confirmed by co-staining of TUNEL (green) and Hoechst, and it located beneath the sarcolemma layer. (B) Apoptotic nuclei in SCs (solid arrow) are confirmed by co-staining of TUNEL (green), Pax7 (red) and Hoechst.

TUNEL studies coupled with laminin immunostaining also indicated that apoptosis occurred for myofibre-associated nuclei (myonuclei and/or satellite nuclei) but only in old muscles where it represented less than 20% total apoptosis (Figure 2.28A). Further studies were then performed with old muscles (25M) to specify whether myofibre-associated apoptosis was due to myonuclei and/or to satellite cells. Muscles cross-sections were immuno-labelled for dystrophin to specifically localize the sarcolemma and to distinguish myonuclei (inside the sarcolemma), from satellite and stromal nuclei (outside the sarcolemma). Triple labelling with TUNEL, Hoechst 33258 and anti-dystrophin

of GM cross-sections was carried out for 4 old mice (Figure 2.30A). More than 600 myonuclei were counted, but these data indicated that apoptosis affected only a limited proportion (0.24 + 0.05 %, n = 4) of the total myofibre myonuclei.

In order to identify apoptotic satellite cells, cross-sections of 25M muscle were triple labelled with TUNEL, Hoechst 33258 and anti-Pax7 (Figure 2.30B). These studies revealed the existence of TUNEL-positive satellite cells, and strikingly that during aging apoptosis did occur for a significant proportion (46 + 7 %, n = 4) of the satellite cells (Nuclei were counted as TUNEL positive only if co-localization of the Hoechst dye and TUNEL staining was observed)

2.4 Discussion

Aging study is a fast-moving and dynamic field. However, many studies only accessed 2 age groups, i.e. young and adult or young and old animals, to investigate the age-related variations [38, 78, 79]. The alterations that occur in skeletal muscle during different stages of aging, i.e. early aging, advanced aging and senescence, therefore are not fully understood. Although some researchers [80, 81] characterized the effect of age at more than 3 key points across life span, few attempts have been applied on skeletal muscle [69]. Therefore I studied C57BL/6 male mice from 2, 11, 22 and 25M in this study, which provided a greater insight into the biological events during growth, early and advanced aging. Age-related variations in myofibre have been widely reported; however previous researchers specifically investigated one or fewer targets, i.e. myofibre type [82], atrophy [81], apoptosis [83]. In addition, ECM of skeletal

muscle is poorly studied. In this study, I examined the similar properties of myofibre to confirm or challenge the previous hypotheses. I also applied further analyses on muscle ECM, which makes my study as a unique and integrative investigation.

2.4.1 Aging is associated with profound alteration in myofibre morphology

Age-dependent atrophy of the skeletal muscle has been well studied. The diminution in myofibre CSA and the apparition of myofibre-type grouping have been demonstrated in the skeletal muscle [41, 84]. However, it remains to be clarified if the reduced area is associated with the modifications in shape [85]. Ilyina-Kakueva described the myofibres as atrophied triangle-shaped, gigantic round-shaped and split, but they only quantified the myofibre shape in the old animals [86]. On the other hand, the image analysis algorithm implemented in my study is shown to be effective for determining the shape of each myofibre according to the number of adjacent ECM segments. I first found that the number of ECM segment and the CSA are positively correlated, i.e., the smaller myofibre have less ECM segments and are prone to be angular. I also noticed that the proportion of myofibres with fewer numbers of ECM segments, i.e., with 3 or 4 ECM segments increased at 22M (early aging). This resulted in a decreased mean value of myofibre CSA. In addition, myofibres with the same number of ECM segment underwent area reduction in early aging and this reduction accentuated during advanced aging (at 25M). These evidences imply

that the muscle aging is characterized by an increased population of angular fiber and the thorough reduction of myofibre CSA.

2.4.2 Aging is associated with less effective myonuclei

An important consensus on the structure of myofibre emphasizes that a myonucleus regulates a constant volume of cytoplasm, so called MND size, has been challenged. Across fiber type, no correlation was found between fiber CSA and the number of myonuclei in rat diaphragm muscle [87], while a weaker correlation was found in mouse tibialis anterior [88]. By sorting myofibres into fiber type, a robust correlation between CSA and the number of myonuclei was observed in young rat plantaris muscle [89] and similarly in limb muscle of young mice [90, 91]. Therefore, the MND size was suggested to be "myofibre type-specific". Aging influences this "type-specific" MND size differently. MND size of rat SM that comprises mainly type I fiber, significantly decreased in the old age. On the contrary, MND size of rat plantaris muscle which mainly comprises type II fiber did not change [92]. In my study, MND size decreases 20% during aging, possibly because the proportion of type I fiber increases in old GM [93].

Previous investigations about cell death in old skeletal muscles suggested that apoptotic nuclei were mostly myonuclei, although the identification of the nuclei has been ambiguous. My study in mice indicates that myonuclear apoptosis could not be detected in young and adult mice GM and is thereafter a rare event in aged mice. Nonetheless, apoptotic myonuclei appear during aging, although they do not account for more than 20% of total apoptotic

nuclei in the old GM. Myonuclear apoptosis is therefore a rare event in old mice muscle (0.3% of total myonuclei) which partly explained why aging myofibres maintained a similar content of myonuclei. With respect to the global reduced volume and MND size, I thus suggest that the existed myonuclei in old GM became less effective to govern the cytoplasm than they did previously.

2.4.3 Aging is associated with impaired satellite cells

The number of myonuclei remained stable in old GM and myonuclei underwent minor apoptosis (section 2.4.2). I would expect the regeneration of myonuclei, which might be supplied by the fusion of SCs [94]. Zammit [46] has proposed the model of SCs myogenesis. After being activated, SCs proliferated and fused into myotubes that eventually matured into myofibre. In my study, those regenerated myofibres were suggested by the centralization of myonuclei (2% of total myonuclei).

Sarcopenia is thought to partly result from an impaired regenerative potential of skeletal muscle. SCs function as myogenic progenitors in adult muscles and an age-linked decline in satellite cell regenerative potential might limit repair of old muscles. The capacity of SCs to support muscle maintenance depends on their abundance, on their myogenic potential and on their local environment [47, 95]. In mouse GM, the major decrease of SCs abundance occurs at 11M old. Previous studies performed on C57BL/6 mouse similarly showed that in EDL muscle SCs number decreases before 1 year of age [96]. Although SCs number remains at a low level in old muscles, I also provide evidence that apoptosis occurs for a significant proportion of old SCs [97].

Previous studies have demonstrated increased apoptotic SCs in response to exercise [98] and denervation [42] in limb muscles, but the importance of aging for SCs apoptosis was only acknowledged in thyroarytenoid muscle for laryngeal function [99]. My observation in GM suggests that increased SCs apoptosis may impair the regenerative response and contribute to age-related loss of GM. This is in agreement with the indications that SCs derived from old rats demonstrate an increased susceptibility to apoptosis in vitro [97], although alterations in SCs local environmental niche may also be involved [100].

2.4.4 Apoptosis in stromal cells regulated by endorepellin

Besides myonuclei and SCs, stromal cells of the CT account for a significant part of total nuclei in skeletal muscle. Similar estimations of non-myonuclei population were provided for rat before and after mechanical ventilation [101]. The increased density in capillary endothelial cells that I report during aging is likely related to the well-known predominance of oxidative myofibres characterizing old muscles [51]. My study also reveals that the low levels of apoptosis detected in young and adult mice muscle are mostly due to CD31+ capillary endothelial cells. Furthermore, while aging was associated with a detectable raise of apoptosis in satellite and myonuclei, the age-dependent increase in muscle apoptosis was essentially attributable to capillary endothelial cells. Capillary apoptosis indeed explained more than 75% of apoptosis in old muscle.

Since capillary are maintained and capillary endothelial cells are more apoptotic during aging, I suggest an enhanced turnover of capillary endothelial

cells in old muscle, although capillary density per myofibre was maintained during aging.

2.4.5 Fibrosis in skeletal muscle

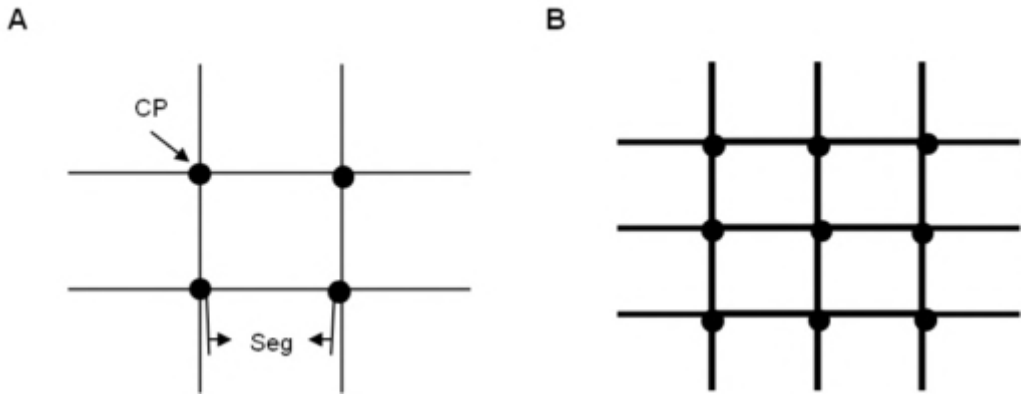


Figure 2.31 Schematic representation of ECM structure in (A) adult and (B) aged mice. Here CP denotes connection point and Seg denotes ECM segments. This diagram is intended to illustrate how the number of connection points and thickening of segments increase in the GM from adult to old mice.

Fibrosis of the ECM was previously mentioned to occur in skeletal muscle of rat [55] and human [102] during aging, but this was not specifically quantified in mouse skeletal muscle. In my study, fibrosis was characterized as increase in collagen content and crosslinking, which is in agreement with previous data reported in rat and human [103]. On the other hand, result from image analyses showed the increased ECM area on the cross-sectional view. I also provided evidence that larger area was associated with increased node points, i.e, connection points, and thickening of ECM segments, which could be schematically represented in Figure 2.31.

In addition to the accumulation of collagen, I observed alteration in TNX and Col-VI level. TNX interacts with fibrillar collagen (especially Col-I)

in vitro [104] and increases collagen gel stiffness *in vitro* [105]. Col-VI extensively locates in muscle ECM and interacts with various ECM components, including Col-I [106]. In this study, increased intensity of endomysial TNX may suggest increased tensile strength of collagen [107]. With the increased intensity of endomysial Col-VI during early aging, I suggest both TNX and Col-VI may play a role in fibrillogenesis and increased fibrosis, respectively [108].

Summary

A profound understanding of structural and compositional variations underlying sarcopenia in skeletal muscle is required. To elucidate the morphological variation of myofibre, histology coupled with image analysis were introduced. Myofibres were described as angular, regular and polygonal (or subround) according due to the number of segment, i.e. 3~4, 5 and 6~7 respectively. Interestingly, myofibre with more segments were statistically larger. Age-dependent muscle atrophy was characterized by a global reduction in myofibre cross-sectional area, and was also associated with increased proportion of angular myofibres and decreased proportion of polygonal myofibres. Myonuclear domain size was concurrently decreased with the occurrence of atrophy, which was suggested by the counter-balanced myonuclei and muscle atrophy. The constant quantity of myonuclei was due to the impaired myofibre regeneration and minor apoptosis in myofibre. The former was proved by the altered satellite cells and the latter was verified by the examination of apoptosis. My result about apoptosis argues again the hypothesis

that apoptosis mainly occurred in myofibre. On the contrary, 75% of apoptosis took place in capillary endothelial cells. And this apoptosis might trigger the liberation of LG3 in endorepellin (domain V of perlecan).

CHAPTER 3 DATA MINING STRATEGY FOR INVESTIGATING STRUCTURE AND COMPONENT OF MOUSE SKELETAL MUSCLE FROM MATURITY TO AGING

3.1 Introduction

In Chapter 2, structural and compositional properties of skeletal muscle at different age groups have been investigated from a histological perspective. I measured and analyzed nearly 30 characteristics covering the whole gamut of organ, cellular and molecular level. I noted that studies of age-related variations, i.e. the result in Chapter 2 and investigations by a large number of researchers focus on specific biological levels rather than a multitude of levels simultaneously. And interestingly, the key to establish a comprehensive understanding linking the different areas (from organ to cellular) is directed at integrating data [109]. Therefore, an objective approach to investigate this integrative data is necessary. Fortunately, there are several important analyzing tools developed that will facilitate this approach.

3.1.1 Principal component analysis

As the mainstay, principal component analysis (PCA) is a data reduction technique for simplifying relevant information from a complex dataset. PCA was first described and named in 1901 [110], but to date, it is still one of the most used multivariate technique in data-preprocessing, pattern recognition

[111], neuro-computing [112] and other fields. The central idea of PCA is to reduce the dimensionality of original dataset while retaining as much as possible of the variations present in the generated dataset. This is achieved by a linear transformation to a new set of variables, the principal components (PCs), which are uncorrelated and ordered so that the first few PCs remain most of the variation present in all of the original variables [113].

If dataset X has n rows (individuals) and p columns (variables), written as $X(n \times p)$, I denote the rows of X by $x_1^{'}, x_2^{'}....x_n^{'}$; and denote the columns of X by $x_{(1)}, x_{(2)}....x_{(p)}$. Dataset X is shown as:

$$X = \begin{bmatrix} x_1^{'} \\ x_2^{'} \\ \vdots \\ x_n^{'} \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^{'} = \begin{bmatrix} x_{(1)} & x_{(2)} & \cdots & x_{(p)} \end{bmatrix}$$

Derivation of PCA is described by following schematic representation [113]:

- Statistics

(1) Mean of j th variable

$$\bar{x}_j = \frac{1}{n} \sum_{r=1}^{n} x_{rj} \qquad \text{Eq. 1}$$

(2) Variance of j th variable

$$s_{jj} = s_j^2 = \frac{1}{n} \sum_{r=1}^{n} (x_{rj} - \bar{x}_j)^2 \qquad \text{Eq. 2}$$

(3) Covariance between i th and j th variable

$$s_{ij} = \frac{1}{n}\sum_{r=1}^{n}(x_{rj} - \bar{x}_j)(x_{rj} - \bar{x}_i) \quad \text{Eq. 3}$$

(4)Correlation coefficient between i th and j th variable

$$r_{ij} = s_{ij} / (s_i s_j) \quad \text{Eq. 4}$$

(5) Covariance matrix:

$$S = (s_{ij}) (\text{size } p \times p) \quad \text{Eq. 5}$$

(6) Correlation matrix:

$$R = (r_{ij}) (\text{size } p \times p) \quad \text{Eq. 6}$$

(7) *Eigenvalue* (λ) of covariance matrix S if $Sx = \lambda x$, x is called the corresponding *eigenvector*. Then covariance matrix S size$(p \times p)$ can be written as

$$S = \Gamma \Lambda \Gamma' \quad \text{Eq. 7}$$

Where Γ is an orthogonal matrix whose columns consist of the standardized eigenvectors corresponding to the eigenvalues $\lambda_1 \geq \lambda_1 \geq ... \geq \lambda_p$; Λ is a diagonal matrix of eigenvalues of S

Next, for the given $X = (X_1 \cdots X_p)'$ with covariance matrix S, principal components (PC) transformation: $X \to Y = \Gamma'(X - \mu)$ is found. Thus PCs are determined by the eigenvectors associated with corresponding eigenvalues.

3.1.2 Clustering

During clustering, individuals are placed into homogeneous groups or clusters on the basis of their similarity. To quantify the similarity, distance

between homogeneous groups is measured by Euclidean Distance, which is the length of a straight line between two groups.

Numerous clustering algorithms have been proposed [114], yet no single algorithm is capable of identifying all kinds of groups and structures that are encountered in practice. Basically, all clustering techniques can be divided into non-hierarchical and hierarchical algorithms. Each of them has their own approaches to determine the number of clusters [115], impose distinct cluster structures [114] and validate the results [116]. Considering the importance and variety of performing clustering techniques, it is desirable to apply several clustering algorithms to the given data and then determine the optimal procedure.

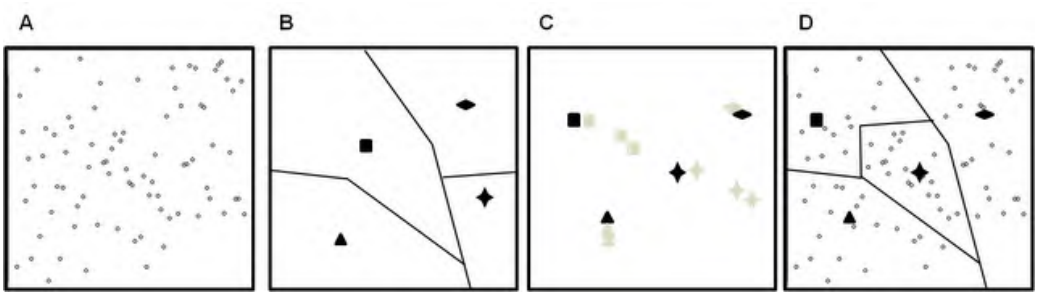


Figure 3.1 *K*-means clustering approach. (A) 2-D randomly plotted data to be clustered. (B) Solid lines divide the space into 4 clusters with random centroid staring positions. Each centroid is denoted by an individual marker. (C) *K*-means approach to find 4 stable clusters. (D) The final clustering is shown.

Non-hierarchical clustering is heuristic in nature which requires the number of cluster in advance. *K*-means algorithm is the most popular non-hierarchical technique to deal with large dataset. It starts with specifying *k* cluster to randomly create *k* partitions. After calculating the centroid for each cluster, distance (Euclidian Distance) between each data points and their

centroid is determined. In order to minimize the overall sum of distance, the procedure is repeated, as shown in Figure 3.1 and the formulated overall sum of distance is given.

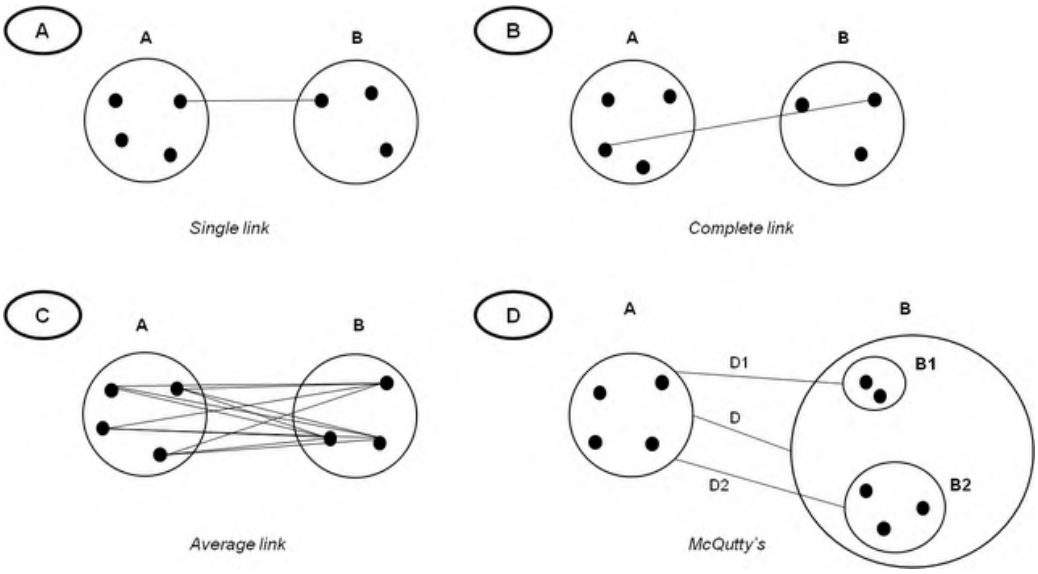


Figure 3.2 Schematic representation of the distance between clusters by (A) single link; (B) complete link; (C) average link and (D) McQuitty's method.

Non-hierarchical methods require less computation than the hierarchical methods. However, the hierarchical methods results in a dendrogram to show the pairwise coupling in the dataset, and the value of the similarity function (level) at which each fusion occurred. The agglomerative hierarchical clustering (AHC) [117], starts with each data point representing a individual cluster. Then the algorithm employed different criteria to search the two clusters that provide the smallest joining cost, merge them to form one new cluster and does so until all data points are agglomerated into one cluster. There are several widely used criteria serve to AHC:

(1) Single link: dissimilarity between cluster A and B is represented by the minimum of all possible distances between the data in cluster A or B (Figure 3.2A).

(2) Complete link: dissimilarity between cluster A and B is represented by the maximum of all possible distances between the data in cluster A or B (Figure 3.2B).

(3) Average link: dissimilarity between cluster A and B is represented by the average of all possible distances between the data in cluster A or B (Figure 3.2C).

(4) McQuitty's: also known as Weighed Pair Group Method with Arithmetic Mean (WPGMA), dissimilarity between new cluster (B) and old ones (A) are computed based on the distance of the two clusters that were merged (B1 and B2). As shown in Figure 3.2D, distance D is defined as:

$$D = (\text{size}(B1) \times D1 + \text{size}(B2) \times D2) \div (\text{size}(B1) + \text{size}(B2)) \quad \text{Eq. 8}$$

(5) Wards's: dissimilarity between cluster A and B is represented by the loss of information from joining the two clusters: merging of cluster A and B results in the minimum increase in the total within-group error sum of square, on the basis of Euclidean distance between the centroids of cluster A and B.

There are another two AHC: centroid and median algorithms. For both methods, each cluster is represented by the coordinate of a group centroid. And the merging occurs when two clusters have similar mean centroid. Difference between the two methods is the centroids of the two merging clusters are

weighed or not respectively, corresponding to the size of the clusters. However, a lot of changes throughout the hierarchy may result from updates. Therefore these methods are not used in this chapter.

After testing and running the cluster algorithms, I need objective criteria to evaluate the quality of clusters, i.e. (1) how well the clusters match the data, (2) the stability of cluster structure resulting from different runs. The final algorithm has to be stable against repeated applications and should stay valid under the variations of the dataset.

3.2 Methods

3.2.1 Dataset

I collected a large set of data comprising twenty-eight different characteristics describing the age-related structural and compositional variance of mouse GM (Table 3.1) and set up the data matrix X with twenty-eight columns (variables) by sixteen rows (individuals):

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1j} \\ \vdots & \ddots & \vdots \\ x_{i1} & \cdots & x_{ij} \end{bmatrix}$$

where subscripts i = 16 and j = 28

In addition, I have also ‘padded’ an extra column, containing the age of the respective individual, to this matrix to denote the categorical variable.

Table 3.1 Twenty-eight characteristics of skeletal muscle structure and component

ID		Unit
Weight		
1	Body	g
2	*Gastrocnemius muscle*	g
Myofibre		
3	Area	μm^2
4	Perimeter	µm
5	Number density	$/mm^2$
6	Myofibre with 3 segments	%
7	Myofibre with 4 segments	%
8	Myofibre with 5 segments	%
9	Myofibre with 6 segments	%
10	Myofibre with 7 segments	%
ECM		
11	Length of segments	µm
12	Number density of segments	$/mm^2$
13	Number density of connection points	$/mm^2$
14	Volume	number of pixels $/mm^2$
15	Area	μm^2
16	Intensity	counts in the pixel
17	Length	µm
Nucleus		
18	Myonucleus	%
19	Myonuclear domain	μm^2
20	Stromal nucleus	%
21	Endothelial nucleus	%
22	Satellite cell	%
23	Apoptotic nucleus	%
24	Apoptotic stromal nucleus	%
25	Apoptotic endothelial nucleus	%
26	Apoptotic satellite cell	%
Collagen		
27	Quantity (hydroxyproline)	µg OH-Pro/mg muscle
28	Crosslink (Pyridinoline crosslinking)	nM PYD/mg muscle

3.2.1 Data pre-processing

Prior to PCA, I normalized the value of the respective entry in the matrix using the following equation:

$$x'_{ij} = (x_{ij} - \mu_i) \div \delta_i \qquad \text{Eq. 9}$$

where μ_i represents the mean value and δ_i represents the standard deviation of j^{th} column in ***X***.

PCA was implemented using the FactoMineR package [118] available in the R-Statistical program (version 2.12.2). Eigenvectors (PCi, i = 1 to 28) of the covariance matrix of X determine the characteristics matrix and act as its basis. Natural rule to reduce the dimension is to keep the first *K* largest eigenvalues (λ_i) according the following criterion [119, 120].

$$\frac{\sum_{i=1}^{K} \lambda_i}{\sum_{i=1}^{N} \lambda_i} > 0.9 \qquad \text{Eq. 10}$$

where N=28 in my study. Further reduction of the dimensionality was carried out by selecting the PCi (1<i<*K*) that correlates to the categorical variable at 1% level via factor analysis. Those PCi's composed the final characteristics matrix and span the lower-dimensional space, in such a way that individual data points from different categories are distributed along the axes with maximum variations. To visualize this, individual data are plotted in 3-D space which was made up of a combination from the first *K* PCs that best correlated with age. Once I confirmed that the individuals were best profiled in the two dimensional space composed of PCs from final characteristic matrix, I calculated the coordinates of the centroid of each category.

As a next step, quantitative variables are projected onto the lower-dimensional space (in my case, it happens to be a 2D plane made up of PC1 and PC3). Quality of projection was represented by the squared cosine of angles

between variables and their projections: higher value of squared cosine amounts to better quality of projection. Correlation between each variable and PC1 (PC3) was represented by the correlation coefficient between its projection and PC1 (PC3).

3.2.3 Clustering

Individual data were studied by clustering. Among the various algorithms, *k*-means and AHC were tested by using stats, cluster and fpc packages available in the R-statistical program (version 2.8.1). These two methods are representing non-hierarchical and hierarchical clustering, respectively.

- *K*-means

As one of the simplest and fastest algorithms, *k*-means clustering [121] aims to minimize the sum of squared distances between all individual points and the cluster centre. I chose Euclidean distance since it is more intuitive and applicable, especially in resulting 2-D space for visualization. Then I computed the intra-cluster error sum, also known as within-group error sum [122] and the inter-cluster error sum, also known as between-cluster error sum by supposing there is a global centroid.

It is clear that the singleton clusters have the minimum intra-cluster error sum. On the other hand, where all the entities belong to one cluster, the intra-cluster error sum is the maximum. Interestingly, the inter-cluster error sum shows the opposite trend to that of the intra-cluster error sum. By using these

conflicting error sums, I worked out a measure for optimal cluster configuration.

A key issue in obtaining a set of cluster is the determination of the number of clusters. Non-hierarchical techniques require the number to be specified to run the analysis. It begins with computing the optimal number of cluster based on Eq. 13, where ideal number of cluster (N_I) is suggested by the minimum *clustering gain* (Δ_i) [123].

$$\Delta_i = \gamma_i - \varepsilon_i \qquad \text{Eq. 11}$$

where γ_i refers to the decreased inter-cluster error sum, and ε_i refers to the increased intra-cluster error sum. In this N_I clusters, each individual belongs to the cluster with the nearest mean distance.

- Agglomerative hierarchical clustering

AHC approach begins with considering each individual as one cluster. I measured the euclidean distance between any two clusters and merged the closest two. Once the cluster comprises more than one sample, euclidean distance between clusters were determined by the (1) single link, (2) complete link and (3) average link (4) McQuitty's and (5) Ward's methods. These 5 criteria emphasize the connectedness, compactness, averaged similarity, unevenness and minimum variance of patterns in a cluster respectively [124-126]. And then I merged the clusters repeatedly with minimum euclidean distance to build up the clustering hierarchy, until it eventually led to one cluster.

To select the optimal criterion among the various AHC, I measured the *agglomerative coefficient* which demonstrates the strength of a cluster structure [127]. Then the optimal AHC algorithm was suggested by the largest *agglomerative coefficient*.

This optimal AHC algorithm provides information that allows the analyst to decide on the number by examining the tabular or graphical output to identify the gaps that define logical clusters, which means measuring the *gap distance* (***l***) between the newly formed and previous clusters (as illustrated in Figure 3.9). The optimal number of cluster, denoted as ***N2***, corresponded to the longest ***l*** [114].

To evaluate the performances between *k*-means and the optimal AHC, I first compared N_1, N_2 to 4 (number of category): If $N_1 \neq N_2$, the N_i (i = 1,2) closer to 4 associated with the better algorithm; If $N_1 = N_2$, *cluster quality* (τ_i, Eq. 13) was calculated and larger value referred to a better quality; if the same τ_i were obtained, I examined a number of distance based statistics, denoted validity [128] and also considered the interpretability.

$$\tau_i = \varepsilon_i / (\varepsilon_i + \gamma_i) \qquad \text{Eq. 12}$$

3.3 Results

3.3.1 Dimensionality reduction

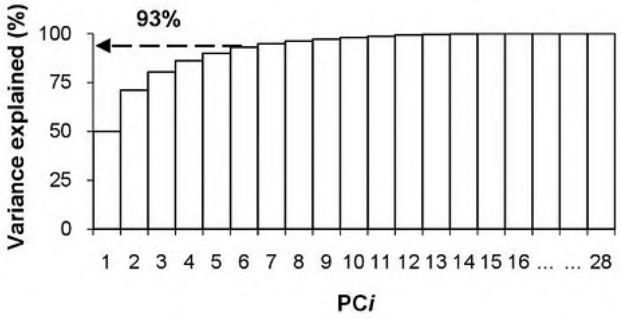


Figure 3.3 Variance explained in PCs. PC1 to PC6 explained 93% variance

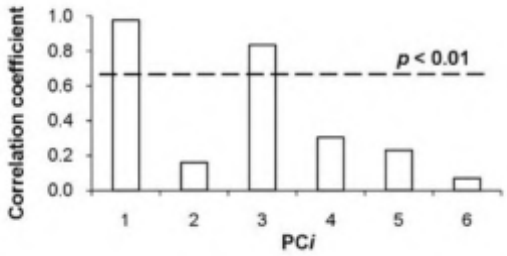


Figure 3.4 Correlation coefficient between PCs (PC1 to PC6) and the categorical variable (age). Only PC1 and PC3 correlated to age at 1% level.

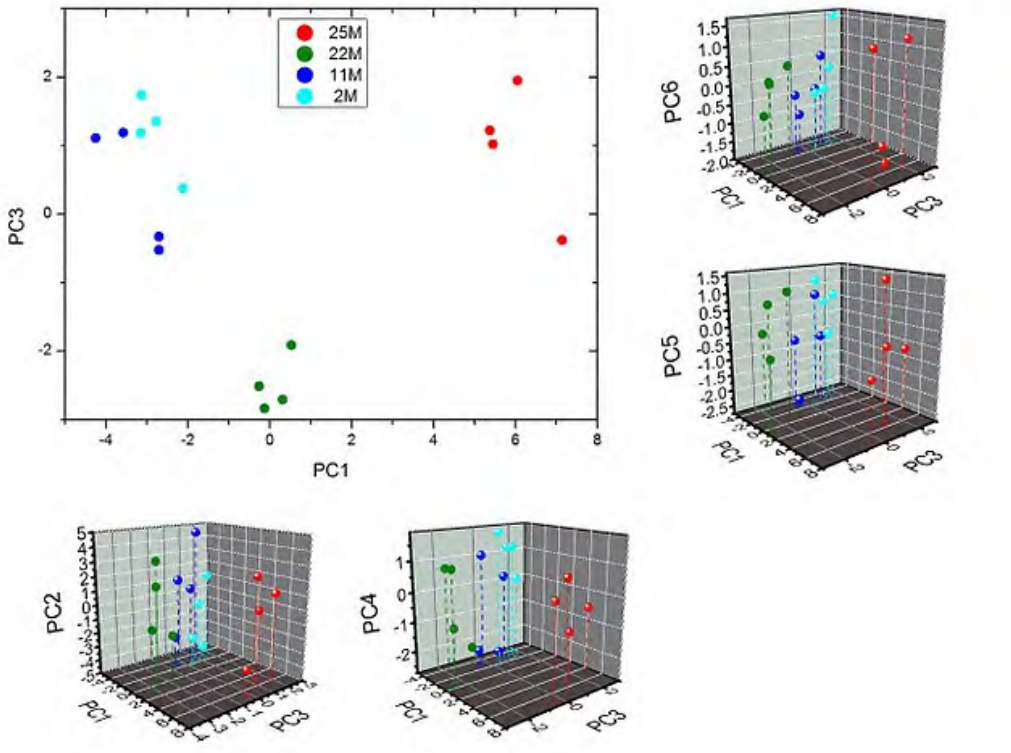


Figure 3.5 PCA result (individual data). Individual data points plotted (A) on PC1-PC3 plane; (B-E) in three-dimensional space comprising PC1, PC3 and PCi (i = 2, 4, 5 and 6). PCi (i = 1 to 6) explained 50.02%, 21.18%, 9.20%, 5.80%, 3.75% and 3.10% variance respectively. Individual from 2M, 11M, 22M and 25M was colored by sky blue, dark blue, green and red respectively.

In order to remove the noise of the dataset, PC1 to PC6 from the twenty-eight PCs explaining 93% variance were preserved (Figure 3.3). PC1 and PC3 were then singled out because both yielded significant correlation to age ($p<0.01$) (Figure 3.4). To visualize this in a comprehensive manner, I plotted individuals in 3-D space comprising PC1, PC3 and PCi (i = 2, 4, 5 and 6) as x-, y- and z-axis, respectively. As shown in Figure 3.5, only the PC1-PC3 plane was available to enable a distinguishable pattern of individuals from different ages. To be more specific, 2M, 11M and 25M locates at the extremity of PC1, and the same goes for 22M along PC3 (Figure 3A). For these reasons, PC1 and

PC3 that explained 50.02% and 9.20% variance respectively were informative enough to represent age-associated variations and were selected to build up the optimal lower-dimensional space (2-D).

3.3.2 Variable selection

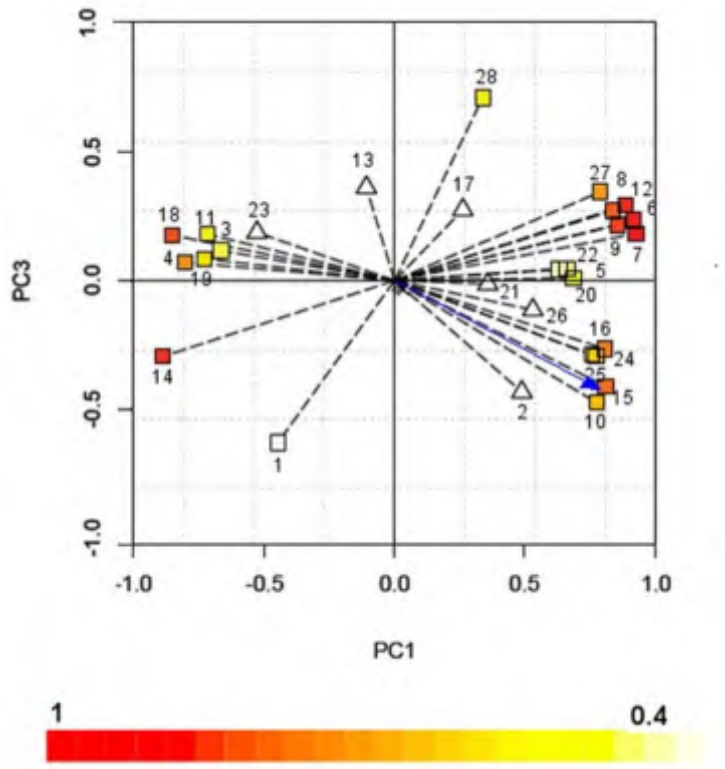


Figure 3.6 PCA result (variables). Variables were plotted to show the correlations between twenty- nine variables (twenty-eight quantitative variables plus one categorical variable presented by blue dash line) and PC1-PC3 plane. Quantitative variables were marked by IDs shown in Table 1. The twenty-two out of twenty-eight variables were correlated to PC1-PC3 plane at 1% level and were presented by vectors ending with rectangles. The other six out of twenty-eight variables were presented by vectors ending with triangle. For the twenty-two variables, quality of projection (cos^2) was presented by the varying colors shown in color bar.

PCA is applied to find the relationships between informative variables and PCs. As illustrated in Figure 3.6, I first found twenty-two out of twenty-eight variables were correlated to PC1 or PC3 at 1% level. For each of those

variables, the quality of its projection from higher-dimensional space onto the PC1-PC3 plane was examined and the correlation coefficient to PC1 (PC3) was calculated. Interestingly, decreased quality of projection and correlation to PC1 was observed from the distribution of apoptotic muscular nuclei (6-9 in Table 3.1), myofibre morphology (15, 16, 18 and 19 in Table 3.1) and size (3, 5 and 11 in Table 3.1), the distribution of normal muscular nuclei (10, 12-14 in Table 1) to ECM structure (20-26 in Table 3.1).

3.3.3 Clustering

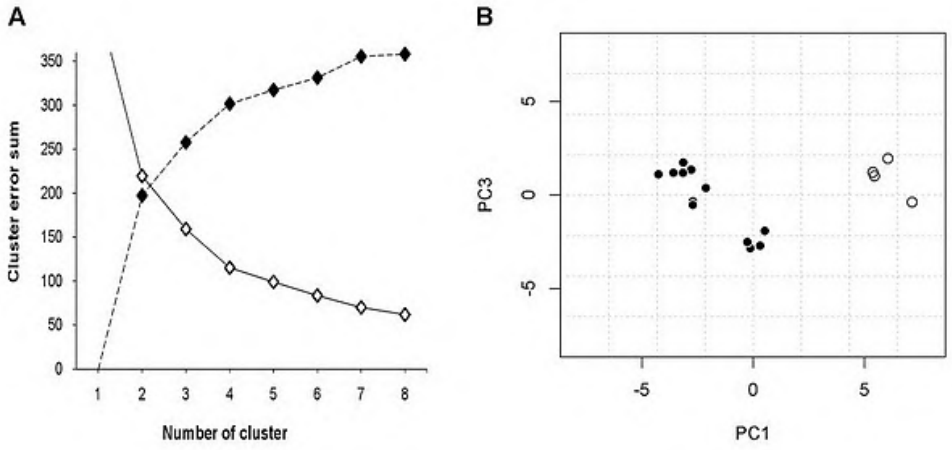


Figure 3.7 *K*-means clustering result. (A) Optimal number of cluster (*N1*) was determined by the decreased *inter-cluster error sum*, marked by hollow diamond and increased *intra-cluster error sum*, marked by solid diamond, which reached the minimum difference when *K*=2. (B) Structure of *k*-means clustering (*K*=2) was presented onto PC1-PC3 plane and two clusters were nicely labeled by black circles and white circles.

After performing PCA, primary age-related information was preserved. Then I applied k-means and AHC on PC1-PC3 plane to seek for the optimal clustering algorithm. For *k*-means clustering, I first found the minimum clustering gain when *K*=2 (Figure 3.7A). *K*-means clustering was whereupon applied with the optimal number (N1=2) and I defined the bigger cluster as

cluster 1 including individuals from 2-22M, the smaller cluster as cluster 2 including individuals from 25M (Figure 3.7B).

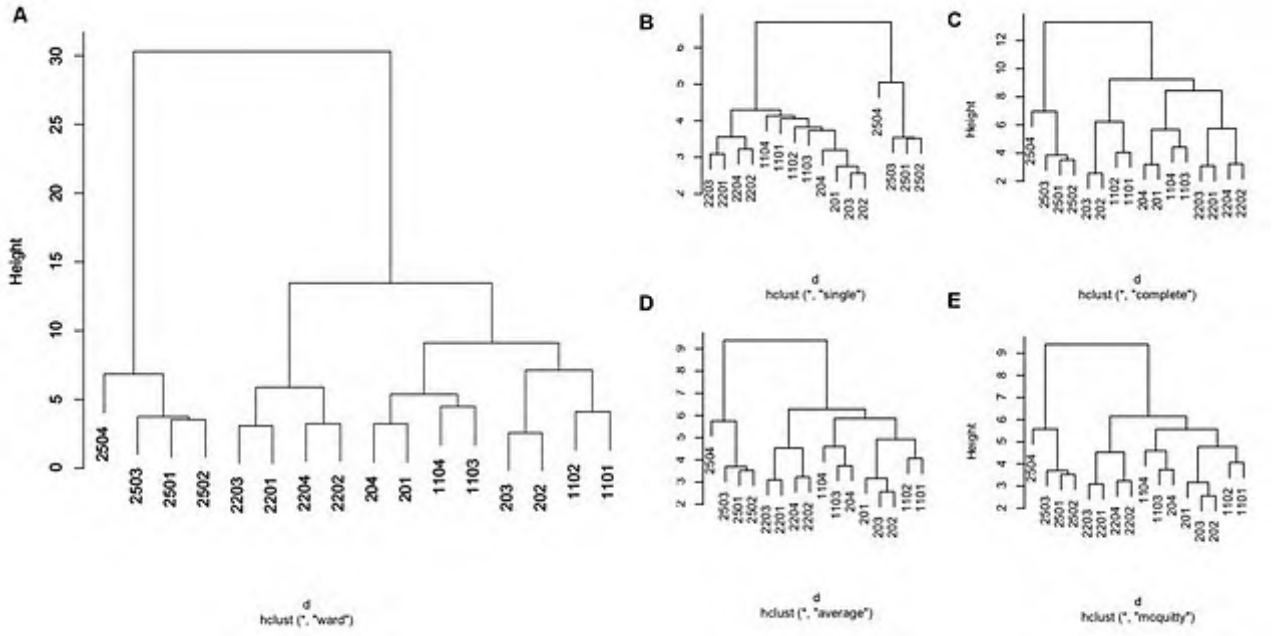


Figure 3.8 AHC results: five criteria were used here: (A) Ward's, (B) single, (C) complete, (D) average and (E) McQuitty's. Height (y-axis) refers to the distances between merging clusters at the successive stages. 201-204, 1101-1104, 2201-2204, 2501-2504 refer to mice from 2M, 11M, 22M and 25M respectively.

For AHC, I singled out Ward's as the optimal algorithm because of the maximum agglomerative coefficient ($0.88_{\text{Ward's}}$ > 0.72_{complete} >$0.62_{\text{McQuitty's}}$ > 0.62_{average} > 0.49_{single}, subscripts refer to the criterion). To visualize this, dendrogram were plotted (Figure 3.8) and the one derived from Ward's criterion (Figure 3.8A) presented the maximum height when forming the last hierarchical tree.

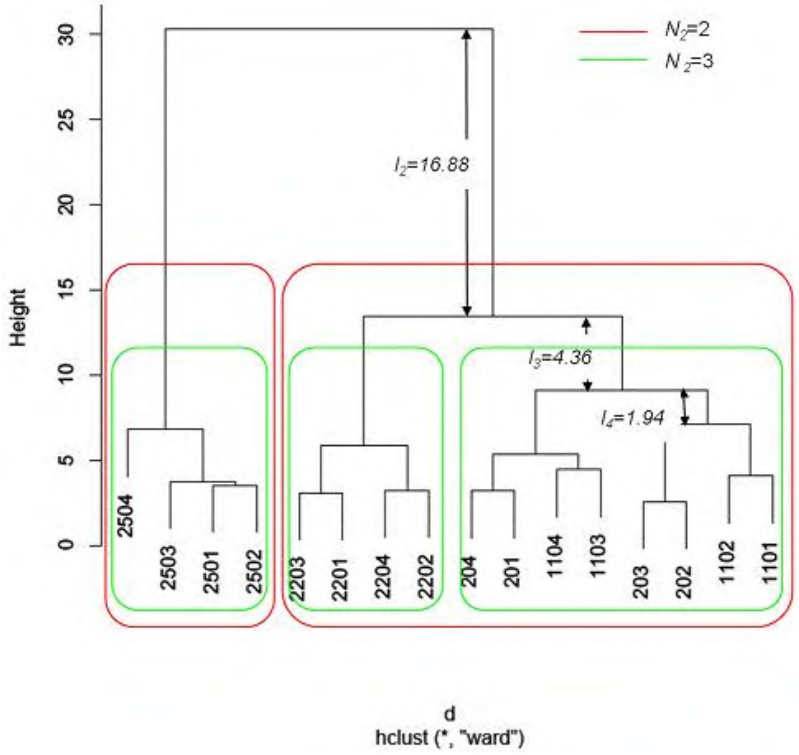


Figure 3.9 Distinct grouping by performing Ward's method. Two and three groups were labeled by red and green rectangles respectively. 201-204, 1101-1104, 2201-2204, 2501-2504 refer to mice from 2M, 11M, 22M and 25M respectively. I defined gap distance (*l*) as the distinct height between the newly formed and previous clusters. Example (*l2* to *l4*) was shown.

The longest gap distance (***l***) leads to the identification of N2 cluster. In my study, dendrogram was cut when N2=2 because ***l***2>**l**3>***l***4 (Figure 3.9). Then individual data points from 25M (*n*=4) belonged to cluster 1 while the others (*n*=12) belonged to cluster 2 (red rectangle in Figure 3.9). Importantly, one

advantage of AHC is the flexibility to group data, in other words, cut the dendrogram at arbitrary level. I noticed if cut the tree with 3 clusters (green rectangle in Figure 3.9), the threshold was still significant. This empirical clustering also resulted in a clear structure and classified individuals of 25M, 22M and 2-11M (n= 4, 4, 8) to cluster 1-3 respectively.

3.3.4 Clustering validation

Since $N1$=$N2$=2, I measured the cluster quality of the results from k-means and Ward's algorithms at 2 clusters and obtained the same value (52.68%). I then examined the validity of the two algorithms and also obtained the same values (cluster diameter, average and median distance, separation, etc., data not shown). I eventually used Ward's hierarchical clustering algorithm as the concluding clustering since the result of Ward's are just more interpretable in terms of how the individuals are grouped on the clustering, i.e. dendrogram with arbitrary tree-cut.

3.4 Discussion

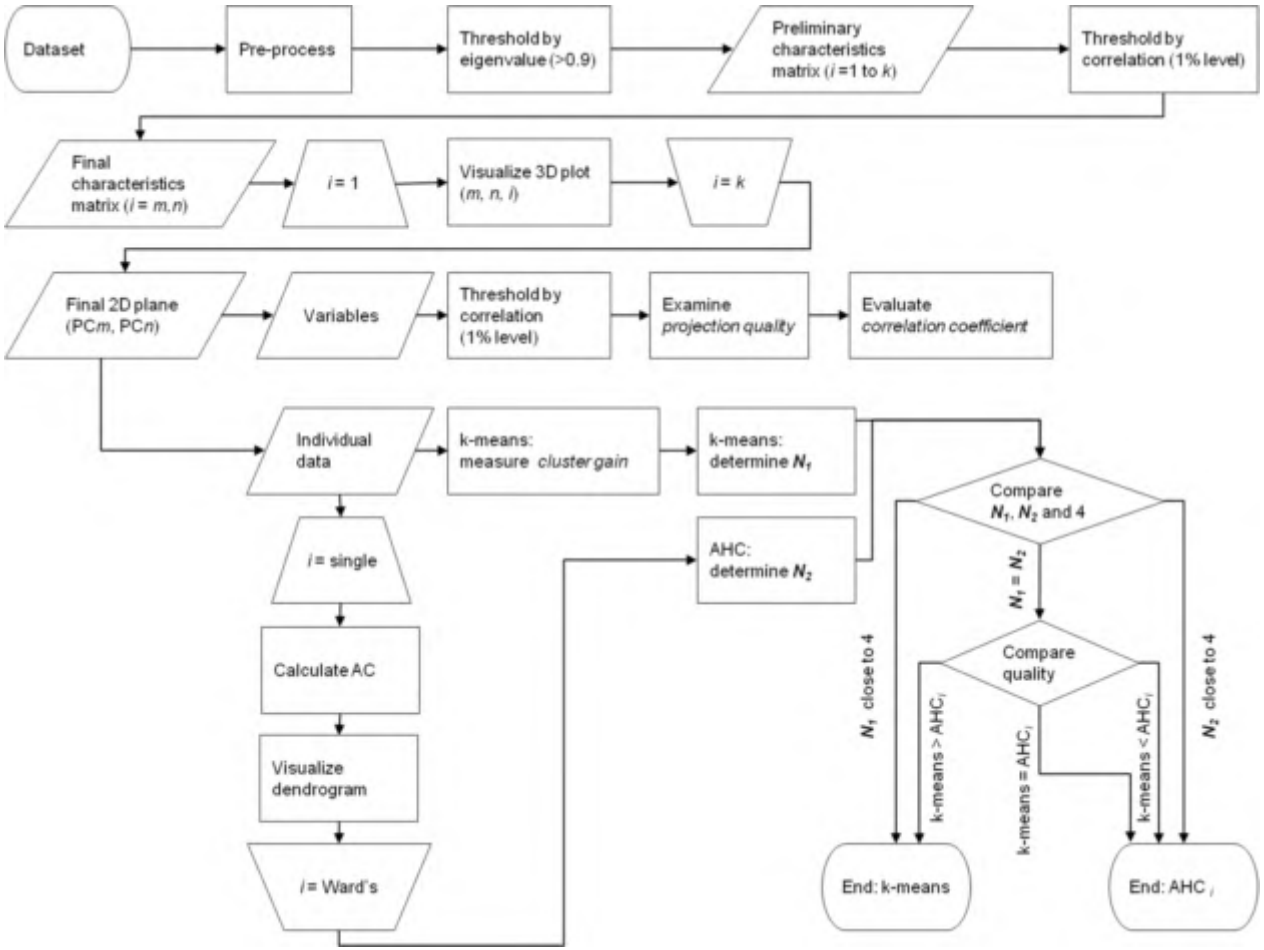


Figure 3.10 Flowchart of the PCA-clustering strategy

I developed an effective protocol applying data mining approaches on a complex dataset to result in a pattern showing age-related information (Figure 3.10). I perform PCA at the beginning, not only to eliminate noise, which has been reported in numerous studies [129-132]; but also to select age-related information. Yeung and Ruzzo have not recommend PCA before clustering on their sporulation data because they demonstrated the first few PCs may not contain cluster information [133]. They suggested to investigate the effectiveness of PCA as a pre-processing step to cluster analysis but did not propose any protocol. In this study, I chose the PC1-PC3 instead of first two PCs according to the correlation between PCs and age groups. This efficient selection is seen to be appropriate since different age groups separate very well in PC1-PC3 plane and not very well on planes of other combinations. I also showed the degraded cluster quality if applying the cluster algorithms (take ward's method as an example) on the original dataset (Figure 3.11A) instead of dataset pre-treated by PCA (Figure 3.11B). It is rather unusual to draw biplots with PC1-PC3 omitting PC2, but it is important to note that PC1 and PC3 are age-related PCs. While principle of PCA allows investigators to extract patterns explaining other variance in the data, for the future it will be interesting to extract, for instance, disease-specific pattern and search for the underlying covariations.

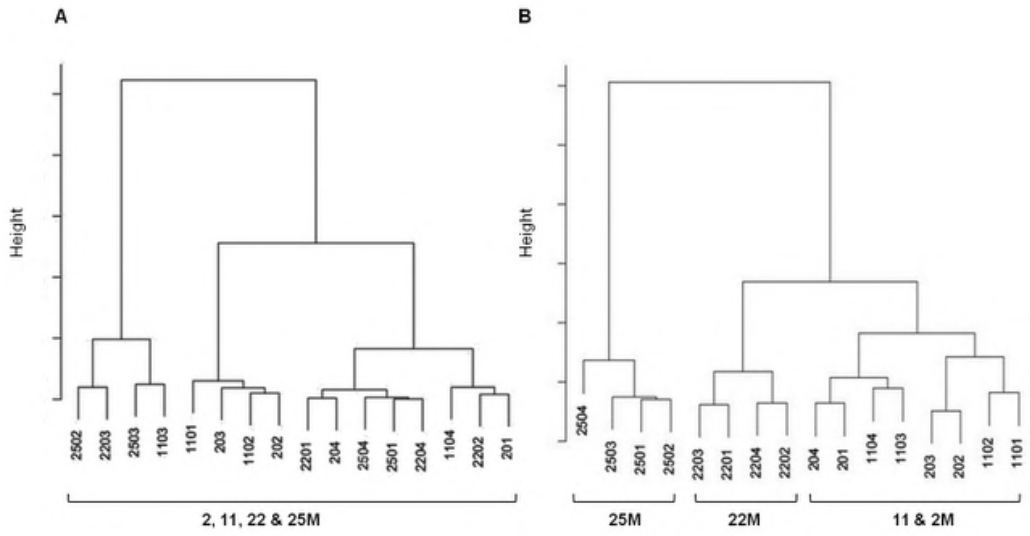


Figure 3.11 Distinct grouping by performing Ward's method on (A) original dataset and (B) dataset pre-treated by PCA. 201-204, 1101-1104, 2201-2204, 2501-2504 refer to mice from 2M, 11M, 22M and 25M respectively.

To further study the partition among individuals, *k*-means clustering and AHC with 5 different methods were tested independently. For *k*-means algorithm, an investigator normally has to test with different numbers of cluster (random or based on prior knowledge) to evaluate the resulting cluster structure. It is time consuming if one has to test a dataset with large range of appropriate number of cluster. I shortened this procedure by introducing *clustering gain*, suitable number of cluster that was suggested and associated with significant cluster structure.

For AHC algorithms, I proved Ward's method was the best within the 5 AHC criteria by showing the largest *agglomerative coefficient.* But it is difficult to compare Ward's and *k*-means because both profiled the same result, quality and validity. This similarity originates from the same measure of adequacy, i.e. both clustering behaviors can be interpreted as a procedure seeking the

minimum dissimilarity. Therefore, Ward's algorithm can be regarded as the hierarchical analogue of *k*-means. Nevertheless, for a dataset comprising 28 x 16 entries, I selected Ward's algorithm as the most suitable clustering algorithm because it elucidated the hierarchy by dendrogram. However, if dealing with a large dataset, AHC is not suited due to its time complexity: for *n* objects, there are *n* steps and at each step, the size *n*2 distance matrix must be updated and searched [134].

PC1 illustrated the most important information because it not only shows most abundant variance (50.02%) but also explains overall aging, i.e. 2M, 11M and 25M. I thus presume the correlations between variables and PC1 can reflect the correlation between variables and overall aging. My result shows that muscular nuclei (normal and apoptosis) distribution, myofibre size and morphology more strongly correlate to PC1 than ECM structure does. And present findings suggest muscle atrophy, type transformation of myofibre and type alteration of stromal cells lead to the varied distribution of normal and apoptotic nuclei [51, 135]. According to my on-going study, it is also expectable that muscle atrophy is accompanied by altered myofibre morphology. Therefore I had an idea that myofibre and muscular nuclei might strongly correlate to overall aging. On the contrary, a weaker correlation is assumed between ECM structure and overall aging. These different correlation between ECM and myofibre to aging is first proposed and compared.

Except for PC1, PC3 is crucial because it helps separate 22M from other age groups and it is attributed to muscle mass and collagen content. On the

variable biplot, GM mass and collagen content locate on the opposite ends. On the individual biplot, individuals from 22M locate along negative PC3, similar to the GM mass. Therefore, if studying the two biplot conjointly [118], I notice individual at 22M are specific with maximum GM mass and minimum collagen content.

Result of clustering evaluates the global difference between categories. Individuals at 2M and 11M are hardly distinguishable and it suggests, by and large, the stability of mice skeletal muscle at the organ- cellular-molecular levels. Individuals at 22M are marked by an empirical grouping, which implies the less significant difference between 22M and 25M, and also between 22M and 2-11M. Individuals at 25M are singled out with the maximum cluster difference. I thus conclude the most significant age-related difference between 25M and the other age groups. In addition, the overall age-associated variations start after 11M and accentuates after 22M.

Summary

In summary, I apply this PCA-clustering strategy on structural and compositional data obtained from mouse skeletal muscle across a wide age range. To find out the age-related PCs, I first considered a subset of PCs that cover 90% of total variance, and then considered the significance of correlation between each PC and age. These PCs explains different stages of life and describes the correlations between the skeletal muscle characteristics and ages. To distinguish individual data points at different ages, I tested several clustering algorithms and subsequently proved Ward's hierarchical clustering to yield the

best individual partition in terms of profile, quality, and validity. Result of clustering describes a minimal difference between muscle at 2M and 11M, commencement and accentuation of age-related variations at 22M and 25M respectively. The proposed strategy provides us with a more integrative understanding of the age-related variations across the whole gamut of organ-molecular-cellular scale studies.

CHAPTER 4 EFFECT OF AGE ON THE TRANSCRIPTIONAL PROFILING OF SKELETAL MUSCLE DURING AGING

4.1 Introduction

4.1.1 Microarray

Microarray technology is a widely used high-throughput tool for profiling gene expression [136, 137]. Expression microarrays use probes targeting specific genes on the basis of nucleotide sequence complementarity, in order to quantitatively measure mRNA levels for a vast amount of genes. To date, there are several popular platforms: spotted cDNA arrays, Affymetrix GeneChip arrays, Agilent ink-jet arrays and Illumina long-oligonucleotide bead-based arrays [138]. These platforms are different in their probe design, hybridization protocol, labeling and production methods [138]. Despite their difference, these traditional gene expression microarray platforms computes the gene-level expression by measuring hybridization intensity from multiple perfect match (PM), which measures the sequence-specific hybridization, and mismatch (MM) probes, which measure the non-specific hybridization to provide a correction of PM intensity for the chemical background (Figure 4.1A-C).

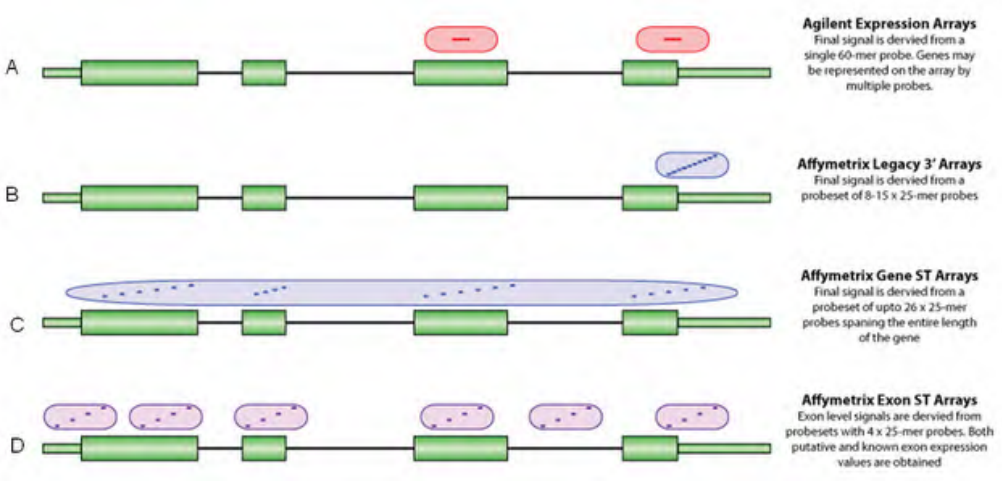


Figure 4.1 Comparison of expression array

In brief, those traditional microarrays consider transcription at the level of individual genes thus they are unable to distinguish different isoforms and able to miss certain transcripts entirely. Affymetrix Exon arrays differ from them on in the number and placement of oligonucleotide probes and in the control probes for background correction. Exon array has no MM probes and instead includes a set of PM probes (reduced to 4) designed to detect hybridization due to pure background (Figure 4.1D). Based on Affymetrix's web, mouse exon array contains approximately 5.5 million probes, which forms about 1.2 million probeset that are together used to separately interrogate 1 million known and predicted exons. Comparing to another widely used array: human genome U133 Plus 2.0 array, exon arrays represents an over four-fold increase in probe density and an eight-fold increase in the number of PM target. This dramatic increase offers a more fine-grained view of gene expression and with the aim to support global inferences about gene expression at the level of individual isoforms and exons.

The development of exon array is also motivated by the inability of studying alternative splicing on other generation of arrays. Alternative splicing is an essential and prevalent forms of transcript change [139]. Alternative splicing is due to the altered pre-mRNA. It refers to the production of multiple transcript isoforms from a single gene [140]. It has been shown by the genome-analyses of expressed sequences that 40%-60% human genes are multiple spliced [141] (Figure 4.2).

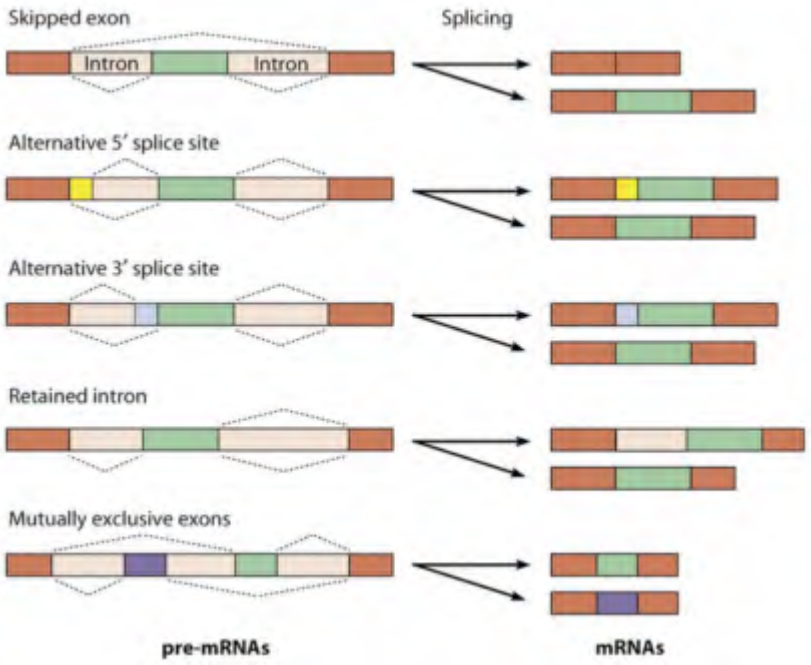


Figure 4.2 Alternative Splicing. In these graphics, exons are represented by boxes and introns are represented by lines. Exon regions included in the messages by alternative splicing are colored while constitutive exons are shown in light orange.

4.1.2 Altered expression of ECM components and its contribution to sarcopenia during aging

Fibrosis of the CT leads to tissue dysfunction and is characterized by an excessive accumulation of ECM components, especially fibrillar collagens

(section 1.3.2.2). This phenomenon may include enhanced ECM protein synthesis or decreased ECM protein degradation, or both. Several growth factors have been implicated in fibrosis: TGF beta 1 and TGF beta 3 induce fibrosis and are involved in ECM protein synthesis, which ultimately lead to CT fibrosis in immortalized human uterine leiomyoma cells [142], in cardiac fibroblasts [143], in liver [144] and in rat glomerular mesangial cells [145]. Baggio *et al*. suggested that polyunsaturated fatty acids interfered with ECM protein synthesis [146].With regard to ECM protein degradation, Roeb *et al*. discussed the possible involvement of matrix metalloproteinases (MMPs)-1 in liver fibrosis under pathogenetic condition [147]. Batlle *et al.* examined MMP-2 expression in pathological fibrosis and suggested that MMPs are responsible for ECM protein degradation [148]. However, the underlying cellular mechanisms causing ECM accumulation is not well understood. In addition, net rate of ECM turnover may play a much crucial role than the absolute rate of ECM protein synthesis in studying ECM alternation during aging. Thus I performed the transcriptomic studies to investigate the turnover of ECM components and the implication of regulating mechanism(s) to fibrosis during muscle aging.

4.2 Methods

4.2.1 RNA extraction

For each age group, independent RNA isolations were carried out for each mouse and used in independent hybridization reactions to generate data sets. Integrity of RNA samples was verified using an Agilent Bioanalyser with the RNA 6000 Nano labchip® kit (Agilent Technologies). Concentration and

purity was determined with a Nanodrop ND-1000 spectrophotometer (Nanodrop Technologies Inc). Samples with a 28S/18S ratio <1.0 and RNA integrity number (RIN) <7 were excluded (Figure 4.3).

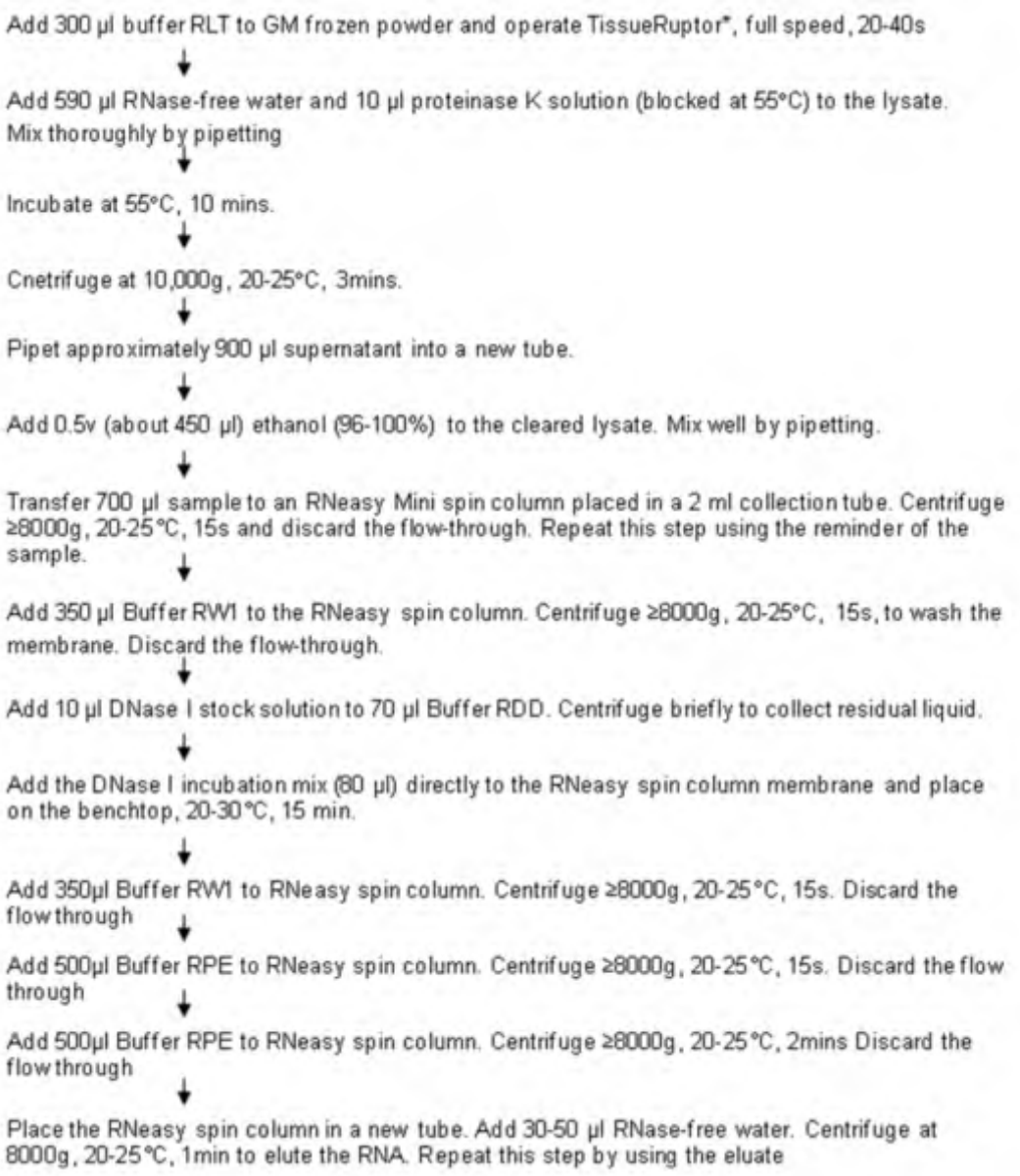


Figure 4.3 (Protocol N° 11) RNA extraction.

4.2.2 Exon array data collection

I used GeneChip Mouse Exon 1.0 ST arrays (Affymetrix Inc, Santa Clara, CA, USA) and for each age group six mouse RNA were processed and run on 6 separate chip. Two microgram of total RNA was labeled according to the GeneChip Whole Transcript (WT) Sense Target Labeling Assay as provided by the manufacturer (Affymetrix) and hybridized to Human Exon 1.0 ST Arrays (Affymetrix) overnight. Chips were washed and stained on a fluidics station 450 (Affymetrix UK Ltd, High Wycombe, UK) and scanned using a GeneChip® Scanner 3000 7G according to the GeneChip Expression Analysis Technical Manual to produce 18 .CEL intensity files (2, 11 and 25M-old, 6 per ages) and 18 .ARR sample attribute files.

4.2.3 Transcriptomic data analysis

Raw data was processed using Genespring GX 11 (Agilent Technologies). I first loaded the .CEL and .ARR files and set the experiment: after normalization. I chose robust multiarray average (RMA, [149]) as summarization algorithm and processed baseline to median of all samples, i.e., for each probe the median of the log summarized values from all the samples is calculated and subtracted from each of the samples. I used the list so-called "core" from metaprobeset list (sourced from Expression Console by Affymetrix), which comprises 17,800 transcript clusters from RefSeq and full-length GenBank mRNAs. I also added parameter "age" to help define the grouping and replicate structure of the experiment.

Before running analyses, I performed quality control to examine the data and remove the ambiguous samples. It comprises hybridization control to depict hybridization quality and 3D-PCA to displays the correlation of expression. Hybridization control used 8 internal probes called AFFX-BioC, AFFX-BioB, AFF-BioDn (genes in the biotin synthesis pathway of E. coli), AFFX-cre (recombines gene from P1 bacteriophage), AFFX-r2-Ec-bioB (All), AFFX-r2-Ec-bioC (All), AFFX-r2-Ec-bioD (All) and AFFX-r2-P1-cre (overall signals for above spikes). Those targets were spiked before the labeling step with different concentration. In addition, BioB is at the level of sensitivity; thus it should be present at least 50% of time. In 3D-PCA, all the array data were projected on 3 PCs (principle in Chapter 3.1.2, in this section, it happened to be PC1-PC3). And this 3D view provides a rich, interactive and dynamic set of visualization.

Filtering step was then employed by selecting 80% entities with higher signal intensity (in either 2, 11 or 25M age groups). This step allowed us to remove very low signal values or those that have reached saturation. I then performed one-way ANOVA and run Tukey HSD posthoc test to find differentially expressed entities. P-value was calculated asymptotically and Benjamini Hochberg (B-H) false discovery rate (FDR) was selected to correct the p-value. The corrected p-value < 0.5 was used to determine statistically significant differential expression of genes between mice at 2, 11 or 25M-old.

After identification of the differentially expressed genes, a new list including the IDs of corresponding genes and their fold change was generated. As next step, I performed comparative analysis on biofunction, canonical

network and pathway of the identified differentially expressed genes by employing Ingenuity IPA software (Ingenuity System, Redwood City, CA, USA). For these analyses, I chose ingenuity knowledge base as the reference set and used B-H multiple testing to calculate the *p* value.

4.3 Results

4.3.1 Quality control

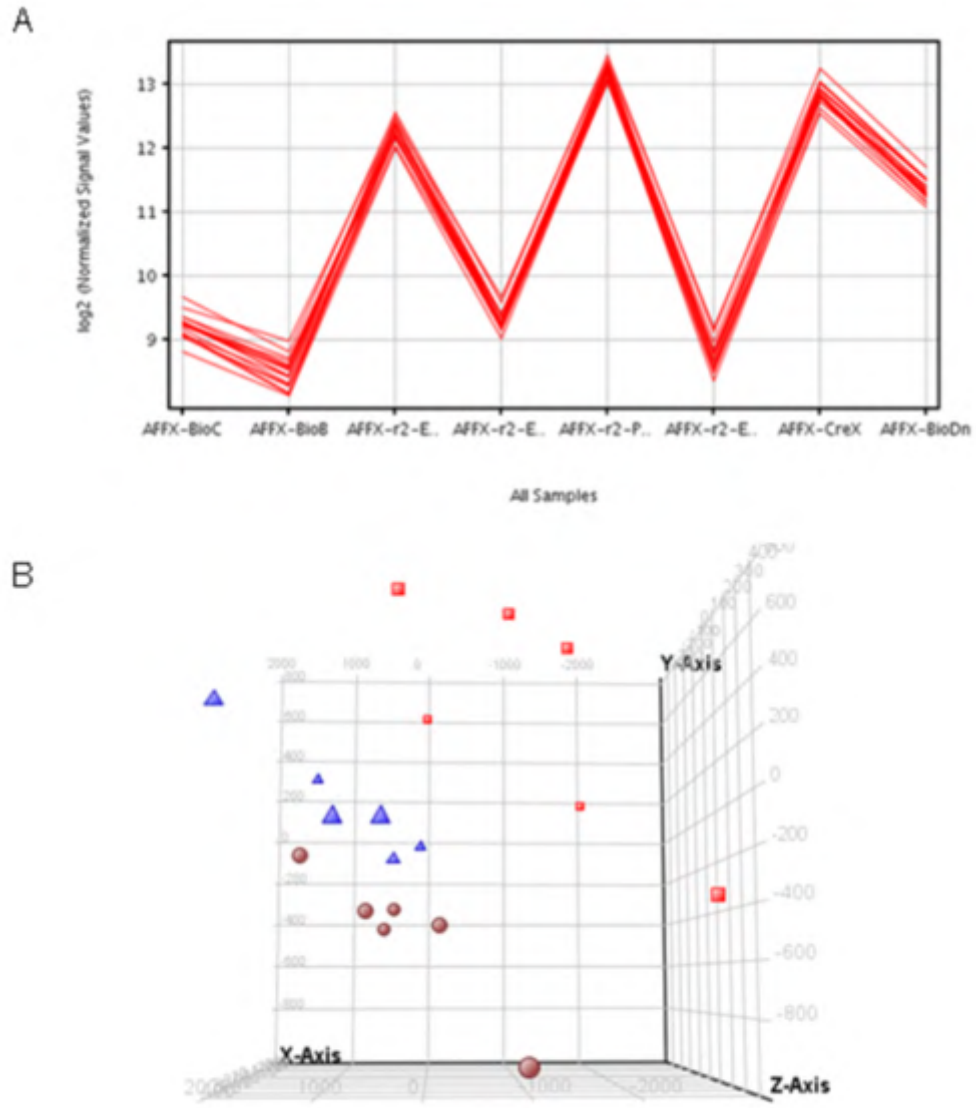


Figure 4.4 Quality controls on sample. (A) Spike-in hybridization controls: 8 targets are spiked before the labeling step. From left to right, they are AFFX-BioC, AFFX-BioB, AFFX-r2-Ec-bioD (All), AFFX-r2-Ec-bioC (All), AFFX-r2-P1-cre, AFFX-r2-Ec-bioB (All), AFFX-CreX and AFF-BioDn. (B) 18 arrays data were projected in 3D spaces comprising PC1, PC2 and PC3 as x-, y- and z-axis. Red cubes, blue triangles and brown balls are used to represent array data from mice at 2, 11 and 25M respectively.

I examined the hybridization quality by plotting hybridization controls view. As shown in Figure 4.4A, the pattern satisfied two criteria: (1) increased concentration from AFFX-BioB, AFFX-r2-Ec-bioB (All), AFFX-BioC, AFFX-r2-Ec-bioC (All), AFF-BioDn, AFFX-r2-Ec-bioD (All), AFFX-CreX to AFFX-r2-P1-cre probesets. (2) Abundant presence of BioB (>50%). Therefore, this pattern was signed to be a good hybridization.

Eighteen array data were projected on 3D space comprising PC1, PC2 and PC3. These 3 PCs explained 54.31%, 20.92% and 13.08% of variance, respectively. As shown in Figure 4.4B, array data from mice at 2M were not well clustered. For the left array data, even PC2 (Y-axis) projected outlier array 25M-03 (ID not shown), PC1 (X-axis) gave interesting result of this array that were not spoiled by the outlier on PC2. I thus kept all the 6 array data from mice at 25M.

4.3.2 Changes in genomic transcription in mouse GM at 2-, 11- and 25M-old

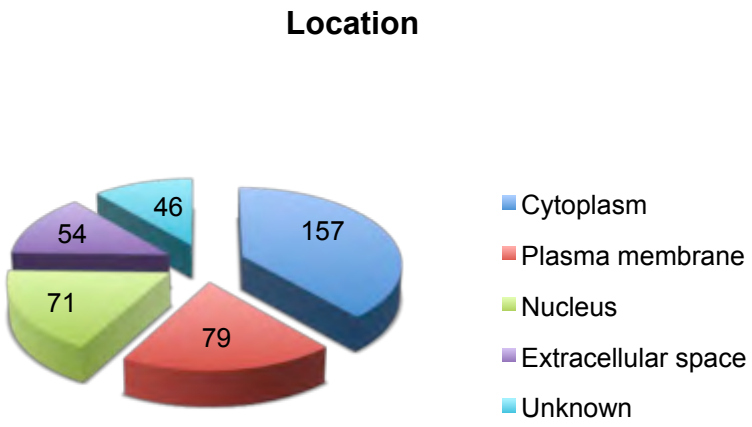
A distinct transcriptional profile was found mice at different ages. In all the 14051 probeset, a total of 2801 were identified as being significantly altered ($p<0.05$, corrected by B-H methods). Of the 2801 differently expressed probesets, 140 were expected to be false positives, i.e., FDR was 4.9%. Among these 2801 differentially expressed probesets, 2434 were identified as being significantly altered between 2M and 25M; 2349 being significantly altered between 2M and 11M; and 466 being significantly altered between 11M and 25M.

Further examining of the altered expressed probesets between 11M and 25M showed 158 out of 466 were more highly expressed in mice at 25M than mice at 11M (up-regulated), and 308 out of 466 showed the opposite pattern (down-regulated) (Figure 4.5).

Figure 4.5 Altered transcripts between 11M and 25M. Approximately 1% of the over 14051 transcripts surveyed by Affymetrix mouse ST exon array were identified to be significantly up-regulated, and 3% to be down-regulated.

Of the 466 altered transcripts, I found out 407 corresponding molecules. As shown in Figure 4.6A, 38.6% were located in cytoplasm; 19.4% and 17.4% were located in plasma membrane and nucleus, respectively; another 13.3% located in extracellular spaces. Among these 407 molecules, 19.2% were enzyme; 7.4%, 6.6% and 6.4% were transporter, kinase and transcriptionregulator, respectively. I also found some minor composition: peptidase (3.9%), ionchannel (2.7%), transmembranereceptor (2.2%), phosphatase (1.7%), growth factor (1.2%), G-proteincoupledreceptor (1.0%) and cytokine (0.5%) (Figure 4.6B).

A



B

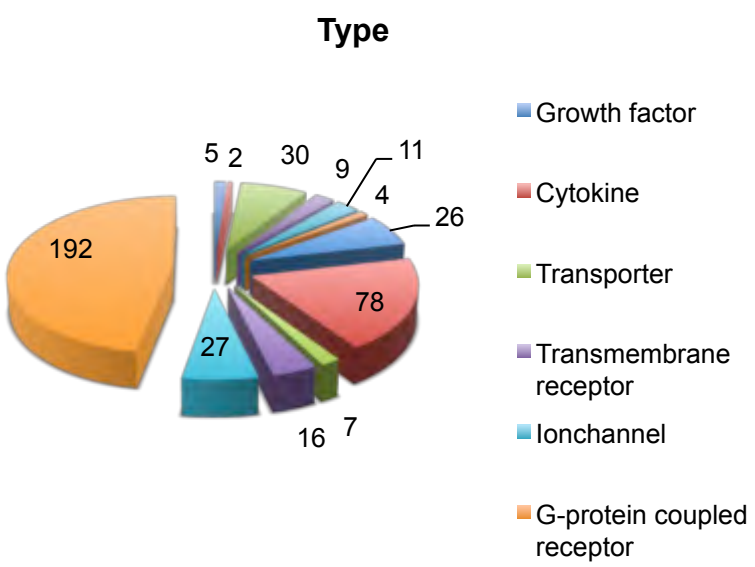


Figure 4.6 (A) Location and (B) type of the 407 molecules corresponding to the altered transcripts

- Biofunction

I performed biofunction analysis using ingenuity pathway analysis (IPA). A total of 73 biofunctions were significantly enriched between mice at 11M and 25M old (maximum p-value <0.05, corrected by B-H methods) (Figure 4.7A). Of these 73 biofunctions, I only focused on the one associated with certain amount of genes (>10). Therefore, 45 biofunctions were preserved and assigned to 3 major categories: diseases and disorders, molecular and cellular functions and physiological system development and functions (Figure 4.7B-D).


A
Bio-functions
Protein Folding
Post-Translational Modification
Organ Morphology
Lymphoid Tissue Structure and Development
Humoral Immune Response
Cell-mediated Immune Response
Hair and Skin Development and Function
Gene Expression
Hepatic System Development and Function
Digestive System Development and Function
Organ Development
Cellular Compromise
Nutritional Disease
Renal and Urological System Development and Function
Hematopoiesis
Infection Mechanism
Infectious Disease
Embryonic Development
Ophthalmic Disease
Protein Synthesis
Protein Degradation
Visual System Development and Function
Reproductive System Development and Function
Lipid Metabolism
Nucleic Acid Metabolism
DNA Replication, Recombination, and Repair
Endocrine System Disorders
Amino Acid Metabolism
Small Molecule Biochemistry
Skeletal and Muscular System Development and...
Behavior
Nervous System Development and Function
Respiratory Disease
Hematological System Development and Function
Immune Cell Trafficking
Inflammatory Response
Molecular Transport
Cell Signaling
Tissue Development
Organismal Survival
Organismal Injury and Abnormalities
Inflammatory Disease
Tissue Morphology
Carbohydrate Metabolism
Connective Tissue Development and Function
Psychological Disorders
Immunological Disease
Renal and Urological Disease
Cell Cycle
Tumor Morphology
Cell-To-Cell Signaling and Interaction
Cellular Function and Maintenance
Cellular Assembly and Organization
Developmental Disorder
Cell Morphology
Cell Death
Gastrointestinal Disease
Hepatic System Disease
Dermatological Diseases and Conditions
Reproductive System Disease
Cardiovascular Disease
Cardiovascular System Development and Function
Organismal Development
Hematological Disease
Connective Tissue Disorders
Cellular Development
Genetic Disorder
Metabolic Disease
Cellular Movement
Cellular Growth and Proliferation
Skeletal and Muscular Disorders
Cancer
Neurological Disease
Threshold
0
1
2
3
4
5
6
7
8
-log (p-value)

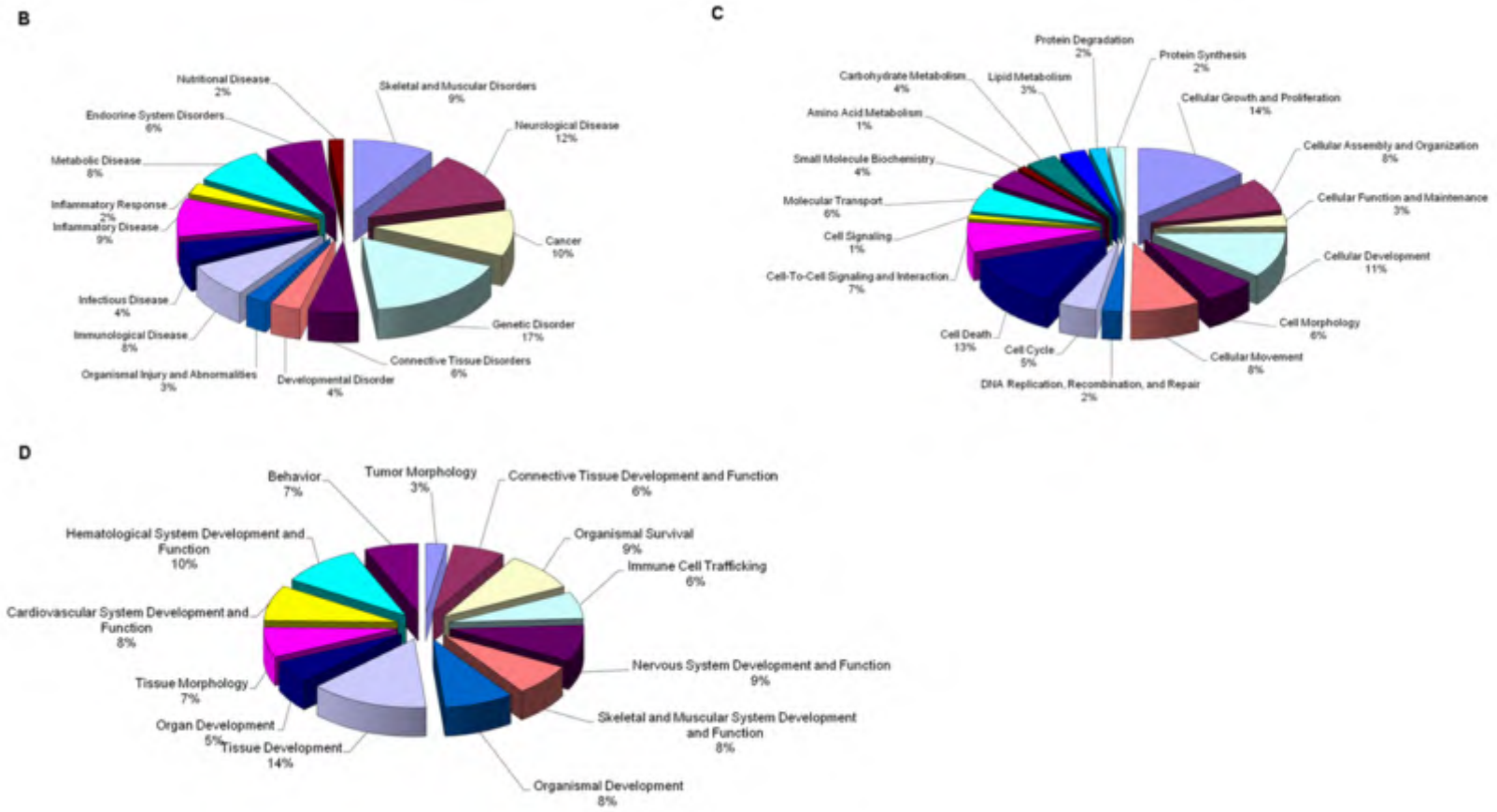


Figure 4.7 Ingenuity Pathway Analysis (IPA) of biofunctions on mice at 11M and 25M. (A) Bar chart shows 73 biofunctions were considered to be statistically different (max *p*-value <0.05, corrected by B-H method). (B-D) 45 of the 73 biofunctions were associated with more than 10 genes an assigned to 3 categories : (B) diseases and disorders ; (C) molecular and cellular functions ; and (D) physiological system development and functions.

- Canonical pathways

IPA analysis was also used to determine biological pathways that were significantly enriched with those differentially expressed genes in mouse at 11M compared with mice at 25M. In all the 240 pathways, only 13 were considered to be statistically different (maximum p-value <0.05 corrected by B-H methods) between mice 11M and 25M old (Figure 4.8).

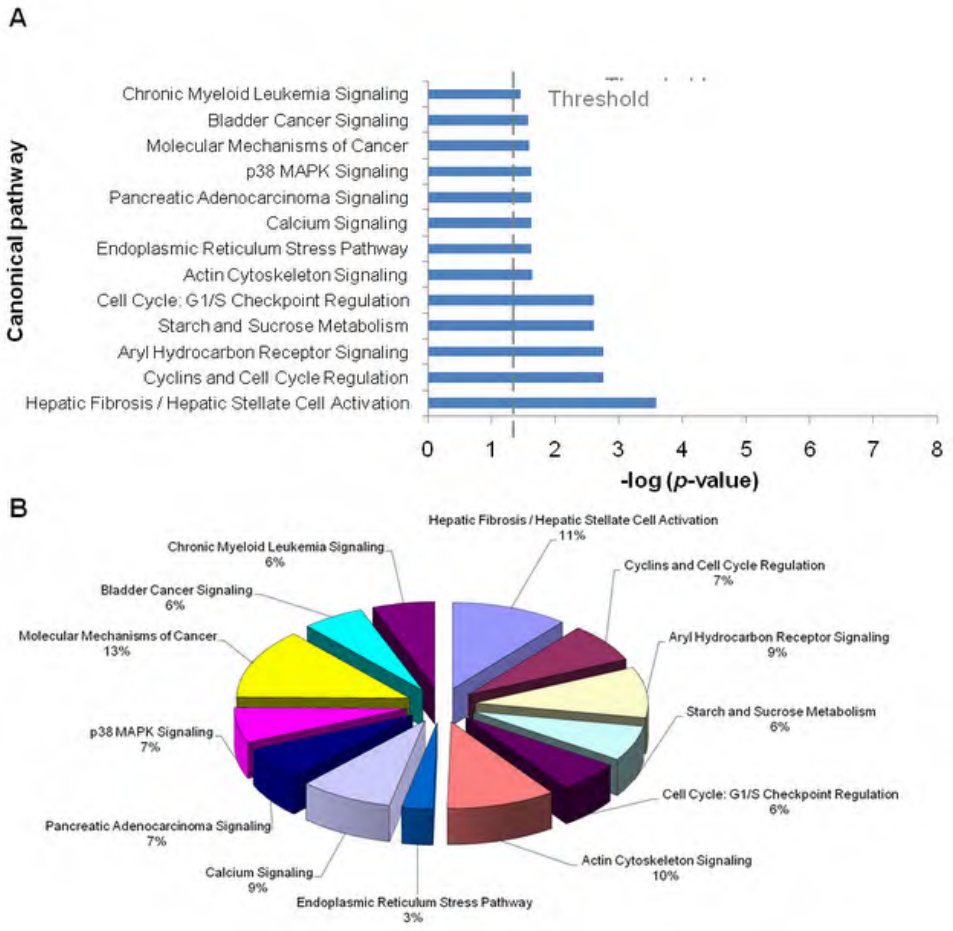


Figure 4.8 Ingenuity Pathway Analysis (IPA) of canonical pathways. (A) Bar chart showed 13 canonical pathways were considered to be statistically different (maximum p-value <0.05 corrected by B-H methods) between mice 11M and 25M old. (B) Pie chart showed the number of genes associated with these 13 canonical pathways (in percentage). And the names of those genes were listed in Table 4.2.

Table 4.1 IPA biofunction

IPA BioFunction	Associated Molecules
A) Diseases and Disorders	
Skeletal and Muscular Disorders (135)	*GPD1, GOLIM4, SNTA1, CA5B, CHRNB1, TNXB, ADAMTS2, VEGFA, RB1, BCL7A, PAK1, LPIN1, IL1RL2, PPBP, CDH13, CHRNA3, SEMA5A, IGFBP5, MMP2, ANXA2, ELOVL1, PLOD1, SLC2A4, UBE2G2, TSPAN14, ANKH, RGL2, GNAL, COL3A1, FYN, FCGRT, KCNN3, FN1, COL4A2, CORO2B, PRKCZ, UCHL1, NOTCH2, MAP4, GADD45A, LBP, CAMK2B, BMP1, PLAT, ABCB1, CTSK, EPHB4, MAPKAPK3, TNFRSF1A, ACTB, ACTN2, SDCCAG8, MLF1, SIN3A, MYO5A, F11R, HP, TUBA1A, C1orf93, CDKN1A, FKBP4, SGCD, CHRNA1, TEKT1, NDRG2, ENO3, MAP1B, DHODH, HSPA5, PDE4D, FAM173A, FDFT1, KCTD20, CYFIP2, KCNAB1, TMTC2, AMPH, NOD1, S100A10, PGK1, TIAM1, HDAC4, COL4A1, APLNR, MICAL2, PF4, SLC1A1, GSN, LBX1, APC, PFKM, COL6A3, NPR1, MBTPS1, LOXL1, FBN1, SYN3, CDK2, NOS1, SCD, DYNLT1, NFIX, BHMT, ANTXR1, EMX2, COL1A2, COL6A1, KCNMA1, ANXA1, ZNF440/ZNF808, CHEK2, ATP13A2, AGT, ATP4B, DPP4, ADCY2, SOD1, MTHFR, RBM38, OPN1LW, RYR2, HDAC1, PYGL, PLEKHB1, SYT9, TEC, COL1A1, NTRK2, B4GALT1, ASTN2, UBAC2, SPARC, CDKN1B, CPEB1, C5orf13*
Neurological Disease (169)	*COL14A1, GPD1, CHRNB1, CA5B, SNTA1, TNXB, LPHN1, VEGFA, LAMC1, BCL7A, MPZL1, RB1, PAK1, SEC24D, MADD, LPIN1, IL1RL2, EHD2, CHRNA3, CDH13, PDE2A, ORM1/ORM2, STK36, TKT, SEMA5A, MMP2, IGFBP5, ANXA2, NFKB2, ELOVL1, SLC2A4, ALG9, KCNB1, PLA2G6, PIP5K1C, HAS2, ME2, ECM1, GNAL, FCGRT, FYN, PGM2, KCNN3, IGFN1, COL4A2, SEPT11, CORO2B, PRKCZ, PHYHD1, UCHL1, NOTCH2, MAP4, GADD45A, PDLIM5, PLAT, CAMK2B, ABCB1, EPHB4, MLLT11, MAPKAPK3, TNFRSF1A, DISC1, ACTN2, ACTB, SERPINF1, XBP1, MLF1, SIN3A, FADD, MYO5A, F11R, HP, TUBA1A, FAM114A1, ANXA3, PROCR, FKBP4, AMY2A, SGCD, NCAM1, CHRNA1, DGKD, NDRG2, ENO3, MAP1B, DHODH, HSPA5, PDE4D, FAM173A, CTSD, FDFT1, CYFIP2, KCNAB1, TMTC2, CDON, AMPH, CARKD, S100A10, HUWE1, PGK1, IGFBP6, CRISPLD2, TIAM1, SSTR4, HDAC4, SESN3, COL4A1, APLNR, MICAL2, PF4, CD93, BACE1, SLC1A1, SLC39A3, GSN, APC, PFKM, NPC1, FLOT2, COL6A3, MBTPS1, LOXL1, SLC22A5, SYN3, RABL5, JAG1, AK2, CDK2, NOS1, SCD, RAB3B, DYNLT1, NFIX, EGF, EMX2, DIAPH1, KCNMA1, FXN, ANXA1, DUSP10, NEDD8, KCNC1, ZNF440/ZNF808, LASP1, ATP13A2, AGT, ATP4B, DPP4, ADCY2, IRF4, TRPC1, SOD1, MTHFR, RYR2, OPN1LW, CANX, PYGL, LOXL3, TEC, COL1A1, NTRK2, B4GALT1, OLFML2B, ASTN2, SPARC, CDKN1B, CPEB1, GVIN1, C5orf13*

Cancer (149)	*CHRNB1, TNXB, LPHN1, ADAMTS2, PCOLCE, VEGFA, RB1, NCOA7, PAK1, FBLN1, FIGF, CHRNA3, CDH13, PDE2A, ORM1/ORM2, KERA, MFAP5, SEMA5A, MMP2, IGFBP5, ANXA2, NFKB2, CLCN5, MYL9, PLA2G6, RUNX1T1, FAM198B, PRELP, HAS2, RGL2, ECM1, STEAP3, COL3A1, SPTBN1, FYN, FN1, COL4A2, PRKCZ, UCHL1, NOTCH2, ARRDC3, GADD45A, DPT, ANTXR2, DAB2, LBP, BMP1, PLAT, CAMK2B, ABCB1, EPHB4, CTSK, TNFRSF1A, ACTB, DAB2IP, XBP1, OLFML3, SERPINF1, MLF1IP, MLF1, FADD, HP, NOV, TUBA1A, CHFR, CDKN1A, ANXA3, PROCR, COL15A1, AMY2A, PTGR1, NCAM1, Tnfrsf22/Tnfrsf23, CHRNA1, NDRG2, ENO3, HDAC10, DHODH, HSPA5, PDE4D, CEBPG, TMEFF2, CTSD, FDFT1, PGK1, HUWE1, IGFBP6, SSTR4, TIAM1, COL4A1, HDAC4, PF4, SLC1A1, GSN, LBX1, BAG4, APC, NPC1, COL6A3, NPR1, GCK, LOXL1, SLC22A5, BAP1, FBN1, CD248, JAG1, AK2, CDK2, NOS1, AOC3, SCD, HIST2H2BE (includes others), BHMT, CRY2, ANTXR1, EGF, EMX2, LOXL2, FKBP10, COL1A2, DIAPH1, COL6A1, UGDH, KCNMA1, ANXA1, SLC40A1, PLIN3, LASP1, CHEK2, AGT, PEX11A, DPP4, ATP4B, ADCY2, IRF4, SOD1, MTHFR, RYR2, HDAC1, RALY, CANX, PLEKHB1, PISD, COL1A1, NTRK2, B4GALT1, SPARC, CDKN1B*

IPA BioFunction	**Molecules**
Genetic Disorder (240)	*COL14A1, CA5B, CHRNB1, PLEKHA3, RB1, MPZL1, MADD, GYS1, FBLN1, EHD2, AMICA1, MFAP5, SEMA5A, ANXA2, IGFBP5, NFKB2, ALG9, MYL9, RUNX1T1, ANKH, HAS2, ECM1, ERI3, KCNN3, PGM2, PTPN14, MID1, SEPT11, SMARCA4, NOTCH2, MAP4, GADD45A, ANTXR2, DAB2, ABCB1, EPHB4, MLLT11, MAPKAPK3, DISC1, TNFRSF1A, DAB2IP, SERPINF1, MLF1, MYO5A, HP, TUBA1A, MYOC, FAM114A1, CDKN1A, PROCR, MAN2B1, SGCD, DGKD, TEKT1, NDRG2, ENO3, HDAC10, HSPA5, PDE4D, TMEFF2, FDFT1, CTSD, SNRPN, KCTD20, CYFIP2, KCNAB1, TMTC2, CARKD, HUWE1, PGK1, AQP7, TIAM1, HDAC4, SESN3, MICAL2, PF4, BACE1, APC, PFKM, NPC1, NUP210, GCK, CCNY, SYN3, ALDH18A1, NOS1, SCD, RAB3B, DYNLT1, EMX2, DIAPH1, COL6A1, UGDH, NEDD8, KCNC1, PLIN3, ATP13A2, LASP1, AGT, DPP4, SOD1, MTHFR, OPN1LW, CANX, PYGL, LOXL3, TEC, NTRK2, ASTN2, SPARC, CDKN1B, C5orf13, GVIN1, PMVK, GPD1, TNXB, LPHN1, ADAMTS2, VEGFA, BCL7A, PAK1, NCOA7, SEC24D, LPIN1, IL1RL2, PPBP, FIGF, CDH13, CHRNA3, RGL1, PELI2, PDE2A, ORM1/ORM2, KERA, TKT, MMP2, ELOVL1, SLC2A4, PLOD1, UBE2G2, CLCN5, PLA2G6, KCNB1, PIP5K1C, ME2, GNAL, COL3A1, SPTBN1, FYN, FN1, IGFN1, COL4A2, RFX5, H2AFZ, CORO2B, PRKCZ, UCHL1, PHYHD1, FSTL1, PDLIM5, LBP, CAMK2B, PLAT, CTSK, ACTB, ACTN2, XBP1, MLF1IP, MFAP4, RAMP1, SIN3A, FADD, F11R, C8orf42, RHPN2, CHFR, ANXA3, FKBP4, COL15A1, AMY2A, NCAM1, CHRNA1, MAP1B, DHODH, FAM173A, AMPH, NOD1, CRISPLD2, IGFBP6, SSTR4, APLNR, COL4A1, SLC1A1, SLC39A3, GSN, FLOT2, PYGM, COL6A3, MBTPS1, NPR1, LOXL1, SLC37A4, SLC22A5, FBN1, RABL5, AK2, JAG1, CDK2, AOC3, NFIX, CRY2, ANTXR1, MMP15, EGF, FMO5, COL1A2, FXN, KCNMA1, ANXA1, DUSP10, ZNF440/ZNF808, SLC40A1, OGDH, HMGCS1, CHEK2, ATP4B, ADCY2, IRF4, RYR2, HDAC1, QSOX1, PLEKHB1, SYT9, COL1A1, B4GALT1, OLFML2B, UBAC2, CPEB1, OSBPL3, KCNG4*
Connective Tissue Disorders (83)	*COL14A1, TEKT1, GOLIM4, DHODH, HSPA5, TNXB, PDE4D, ADAMTS2, TMEFF2, VEGFA, RB1, KCTD20, LPIN1, AMPH, NOD1, CDH13, S100A10, PGK1, SSTR4, TIAM1, COL4A1, MICAL2, MMP2, SLC1A1, IGFBP5, GSN, PLOD1, UBE2G2, COL6A3, MBTPS1, TSPAN14, ANKH, RGL2, FBN1, SYN3, CDK2, COL3A1, FCGRT, FYN, FN1, NFIX, BHMT, MMP15, ANTXR1, COL4A2, PRKCZ, COL1A2, COL6A1, KCNMA1, ANXA1, ANTXR2, LBP, ATP13A2, CHEK2, PLAT, BMP1, ATP4B, ABCB1, DPP4, ADCY2, TNFRSF1A, RBM38, HDAC1, RYR2, SDCCAG8, PLEKHB1, SYT9, TEC, COL1A1, F11R, HP, TUBA1A, NTRK2, C1orf93, B4GALT1, ASTN2, CDKN1A, SPARC, UBAC2, MAN2B1, CDKN1B, COL15A1, SGCD*
Developmental Disorder (51)	*CA5B, MAP1B, SNTA1, VEGFA, LAMC1, FDFT1, SNRPN, CTSD, RB1, PAK1, FBLN1, FIGF, ATF4, TAB1, S100A10, HUWE1, SSTR4, HDAC4, BACE1, SLC1A1, MMP2, NFKB2, GSN, SLC2A4, NPR1, MBTPS1, ANKH, FBN1, ECM1, NOS1, SCD, NFIX, MID1, EGF, H2AFZ, COL1A2, NOTCH2, AGT, ATP4B, DPP4, CTSK, TRPC1, SOD1, XBP1, MFAP4, COL1A1, NTRK2, TUBA1A, CDKN1A, CDKN1B, SGCD*

Organismal Injury and Abnormalities (37)	*NOS1, FYN, SCD, CHRNA1, CHRNB1, EGF, PDE4D, VEGFA, COL1A2, RB1, FDFT1, SMPX, GADD45A, HMGCS1, AGT, BMP1, PLAT, DPP4, TIAM1, SSTR4, SOD1, TNFRSF1A, PF4, CD93, XBP1, MMP2, GSN, COL1A1, NPC1, HP, B4GALT1, NPR1, PROCR, CDKN1B, SGCD, CDK2, COL3A1*

IPA BioFunction	Molecules
Immunological Disease (111)	*COL14A1, GPD1, GOLIM4, CHRNB1, TNXB, ADAMTS2, VEGFA, RB1, PPBP, AMICA1, CDH13, RGL1, PELI2, ORM1/ORM2, SEMA5A, TKT, IGFBP5, MMP2, NFKB2, ALG9, RUNX1T1, TSPAN14, CRYBG3, RGL2, GNAL, COL3A1, FYN, FCGRT, FN1, PTPN14, COL4A2, RFX5, H2AFZ, NOTCH2, GADD45A, PDLIM5, LBP, BMP1, PLAT, ABCB1, MAPKAPK3, TNFRSF1A, ACTB, SDCCAG8, MYO5A, F11R, HP, TUBA1A, TRAFD1, C1orf93, ANXA3, CDKN1A, FKBP4, SGCD, NCAM1, CHRNA1, TEKT1, DHODH, HSPA5, PDE4D, SNRPN, KCTD20, AMPH, S100A10, PGK1, IGFBP6, TIAM1, COL4A1, GABPB2, MICAL2, PF4, SLC6A17, GSN, ERP29, FBN1, ALDH18A1, SYN3, JAG1, CDK2, HIST2H2BE (includes others), BHMT, ANTXR1, EGF, COL1A2, DIAPH1, KCNMA1, DUSP10, ANXA1, CHEK2, ATP13A2, AGT, DPP4, ATP4B, ADCY2, SOD1, IRF4, TRPC1, RBM38, HDAC1, RYR2, PLEKHB1, SYT9, TEC, COL1A1, NTRK2, B4GALT1, CCDC101, ASTN2, UBAC2, CDKN1B, ABHD16A*
Infectious Disease (63)	*COL14A1, AFG3L1, CA5B, LIMK2, DHODH, PLEKHA3, TNXB, HSPA5, PDE4D, VEGFA, FDFT1, PAK1, PPBP, CLEC4M, AMPH, NOD1, HUWE1, SSTR4, GABPB2, PF4, TKT, SEMA5A, CD93, SLC6A17, ANXA2, MMP2, PFKM, FLOT2, PIP5K1C, MBTPS1, GCK, ALDH18A1, GNAL, COL3A1, SPTBN1, ERI3, FN1, HIST2H2BE (includes others), EGF, TRPT1, TMEM63A, MAP4, DIAPH1, ANTXR2, LBP, HMGCS1, BMP1, PLAT, ATP4B, DPP4, TRPC1, TNFRSF1A, FGD6, HDAC1, PYGL, FADD, HP, TUBA1A, TRAFD1, PROCR, SPARC, OSBPL3, AMY2A*
Inflammatory Disease (123)	*GPD1, GOLIM4, CA5B, TNXB, ADAMTS2, VEGFA, RB1, LPIN1, FBLN1, CHRNA3, CDH13, PDE2A, STK36, TKT, IGFBP5, MMP2, NFKB2, UBE2G2, ANKH, TSPAN14, HAS2, RGL2, COL3A1, FCGRT, FYN, FN1, Gp49a/Lilrb4, COL4A2, PRKCZ, MAP4, NOTCH2, GADD45A, LBP, PLAT, BMP1, ABCB1, EPHB4, DISC1, TNFRSF1A, ACTB, SDCCAG8, XBP1, RAMP1, F11R, RHPN2, HP, TUBA1A, C1orf93, FAM114A1, CDKN1A, ANXA3, FKBP4, PROCR, SGCD, AMY2A, DGKD, TEKT1, HDAC10, DHODH, HSPA5, PDE4D, TMEFF2, FDFT1, CTSD, SNRPN, KCTD20, KCNAB1, TMTC2, AMPH, NOD1, S100A10, PGK1, IGFBP6, SSTR4, TIAM1, HDAC4, COL4A1, MICAL2, PF4, SLC1A1, GSN, NPC1, SLC22A5, CCNY, FBN1, SYN3, JAG1, CDK2, NOS1, SCD, NFIX, BHMT, ANTXR1, EGF, FMO5, COL6A1, KCNMA1, ANXA1, DUSP10, ATP13A2, CHEK2, AGT, ATP4B, DPP4, ADCY2, SOD1, IRF4, RBM38, MTHFR, HDAC1, RYR2, PLEKHB1, SYT9, LOXL3, TEC, COL1A1, NTRK2, B4GALT1, ASTN2, UBAC2, SPARC, CDKN1B, OSBPL3*
Inflammatory Response (36)	*AOC3, FYN, FN1, PRDX5, EGF, PDE4D, PRKCZ, VEGFA, DIAPH1, ANXA1, CLEC4M, PPBP, NOD1, FIGF, LBP, AGT, S100A10, PLAT, TIAM1, HDAC4, ORM1/ORM2, COL4A1, TNFRSF1A, PF4, SERPINF1, XBP1, ANXA2, NFKB2, GSN, PLA2G6, F11R, COL1A1, HP, B4GALT1, ANXA3, SLC37A4*
Metabolic Disease (120)	*GPD1, CA5B, CHRNB1, PLEKHA3, TNXB, ADAMTS2, VEGFA, RB1, NCOA7, GYS1, LPIN1, MADD, AMICA1, CHRNA3, CDH13, RGL1, PELI2, SEMA5A, IGFBP5, MMP2, PLOD1, SLC2A4, ALG9, CLCN5, UBE2G2, KCNB1, PLA2G6, CRYBG3, COL3A1, SPTBN1, ERI3, FYN, KCNN3, FN1, PTPN14, COL4A2, SMARCA4, UCHL1, NOTCH2, PDLIM5, LBP, TNFRSF1A, ACTN2, MLF1IP, SERPINF1, MLF1, FADD, C8orf42, TUBA1A, TRAFD1, FAM114A1, CHFR, FKBP4, MAN2B1, AMY2A, SGCD, NCAM1, CHRNA1, PDE4D, TMEFF2, CTSD, FDFT1, SNRPN, KCNAB1, AMPH, S100A10, AQP7, SSTR4, TIAM1, COL4A1, MICAL2, GABPB2, SLC1A1, GSN, PFKM, NPC1, PYGM, ERP29, SLC37A4, GCK, NUP210, SLC22A5, FBN1, SYN3, ALDH18A1, JAG1, CDK2, NOS1, AOC3, SCD, HIST2H2BE (includes others), CRY2, EGF, ARFRP1, COL1A2, FXN, ANXA1, SLC40A1, ATP13A2, CHEK2, AGT, ATP4B, DPP4, ADCY2, SOD1, TRPC1, MTHFR, RYR2, PRKAG3, TEC, COL1A1, NTRK2, B4GALT1, CCDC101, ASTN2, UBAC2, CDKN1B, OSBPL3, KCNG4, ABHD16A*

IPA BioFunction	Molecules
Endocrine System Disorders (94)	*CHRNA1, GPD1, CHRNB1, CA5B, TNXB, PLEKHA3, PDE4D, TMEFF2, VEGFA, FDFT1, RB1, SNRPN, NCOA7, MADD, KCNAB1, AMPH, AMICA1, CDH13, CHRNA3, RGL1, PELI2, S100A10, SSTR4, TIAM1, COL4A1, GABPB2, MICAL2, SEMA5A, MMP2, SLC2A4, UBE2G2, PLA2G6, KCNB1, ERP29, CRYBG3, SLC22A5, NUP210, GCK, FBN1, ALDH18A1, SYN3, JAG1, CDK2, AOC3, ERI3, SPTBN1, FYN, KCNN3, HIST2H2BE (includes others), FN1, CRY2, PTPN14, EGF, COL4A2, SMARCA4, COL1A2, UCHL1, NOTCH2, FXN, PDLIM5, ANXA1, LBP, ATP13A2, CHEK2, AGT, DPP4, ADCY2, TRPC1, SOD1, TNFRSF1A, ACTN2, RYR2, SERPINF1, MLF1IP, MLF1, FADD, TEC, C8orf42, TUBA1A, NTRK2, FAM114A1, TRAFD1, CCDC101, ASTN2, CHFR, FKBP4, UBAC2, CDKN1B, SGCD, OSBPL3, AMY2A, NCAM1, ABHD16A, KCNG4*
Nutritional Disease (25)	*NOS1, SCD, CHRNA1, FN1, CRY2, CA5B, CHRNB1, PDE4D, VEGFA, RB1, GYS1, LPIN1, CHRNA3, ATP4B, DPP4, AQP7, SSTR4, SOD1, HDAC4, TNFRSF1A, TKT, NTRK2, GCK, CDKN1B, AMY2A*
B) Molecular and Cellular Functions	
Cellular Growth and Proliferation (135)	*COL14A1, ADAMTS2, SKAP2, VEGFA, LAMC1, RB1, PAK1, LPIN1, MADD, PPBP, FBLN1, FIGF, ATF4, CDH13, IGFBP5, MMP2, ANXA2, NFKB2, MYL9, PLA2G6, RUNX1T1, PIP5K1C, RGL2, HAS2, ECM1, STEAP3, SPTBN1, FYN, FN1, PTPN14, SEMA6A, Gp49a/Lilrb4, COL4A2, PRKCZ, SMARCA4, UCHL1, NOTCH2, ARRDC3, GADD45A, DPT, DAB2, DPH1, BMP1, PLAT, FBXW4, ABCB1, EPHB4, TNFRSF1A, DISC1, ACTB, DAB2IP, XBP1, SERPINF1, SIN3A, FADD, F11R, NOV, UXT, CDK20, CHFR, CDKN1A, FKBP4, PROCR, NCAM1, CHRNA1, NDRG2, HDAC10, HSPA5, PDE4D, TMEFF2, CTSD, FDFT1, CDON, CD84, HUWE1, PGK1, IGFBP6, AQP7, TIAM1, SSTR4, COL4A1, APLNR, HDAC4, GABPB2, PF4, GSN, LBX1, BAG4, APC, ANKRD1, FLOT2, COL6A3, NPR1, LOXL1, BAP1, FBN1, CD248, RPS15A, AK2, JAG1, CDK2, NOS1, NFIX, ACER2, ANTXR1, MMP15, EGF, EMX2, COL1A2, DIAPH1, COL6A1, FXN, ANXA1, DUSP10, PLIN3, LASP1, CHEK2, AGT, DPP4, IRF4, SOD1, MTHFR, RBM38, HDAC1, RYR2, QSOX1, TEC, COL1A1, NTRK2, B4GALT1, TSPYL2, SPARC, MEF2C, CDKN1B, C5orf13*
Cellular Assembly and Organization (75)	*CHRNA1, PPP2R5B, MAP1B, SNTA1, LIMK2, CDC42EP2, HSPA5, TNXB, LPHN1, ADAMTS2, PCOLCE, VEGFA, LAMC1, PAK1, AMPH, CDH13, TIAM1, PEX6, STK36, COL4A1, SEMA5A, BACE1, MMP2, ANXA2, GSN, PLOD1, APC, NPC1, PIP5K1C, NPR1, SLC22A5, HAS2, SYN3, JAG1, CDK2, COL3A1, FYN, DYNLT1, RAB3B, NFIX, FN1, PTPN14, MID1, EGF, SMARCA4, COL1A2, MAP4, DIAPH1, FXN, DPT, DAB2, LASP1, AGT, PLAT, PEX11A, Pdlim3, SOD1, TNFRSF1A, DISC1, ACTN2, SERPINF1, CANX, SYT9, MYO5A, COL1A1, TUBA1A, NTRK2, UXT, B4GALT1, CDKN1A, FKBP4, SPARC, CDKN1B, C11orf73, NCAM1*
Cellular Function and Maintenance (27)	*CHRNA1, RAB3B, FN1, EGF, TNXB, LPHN1, ADAMTS2, PRKCZ, COL1A2, PAK1, DPT, AMPH, DAB2, Pdlim3, SOD1, TNFRSF1A, ANXA2, SYT9, APC, PFKM, COL1A1, B4GALT1, PIP5K1C, HAS2, CDKN1B, COL3A1, NCAM1*

IPA BioFunction	Molecules
Cellular Development (104)	*COL14A1, CHRNB1, LAMC1, VEGFA, RB1, PAK1, LPIN1, MADD, FIGF, ATF4, TMBIM1, CDH13, ADIG, ANXA2, IGFBP5, MMP2, NFKB2, SLC2A4, MYL9, PTBP1, PLA2G6, RUNX1T1, HAS2, RGL2, STEAP3, FYN, FN1, PTPN14, COL4A2, PRKCZ, SMARCA4, UCHL1, NOTCH2, ARRDC3, GADD45A, DAB2, DPH1, BMP1, ABCB1, EPHB4, CTSK, TNFRSF1A, XBP1, SERPINF1, MLF1, FADD, MYO5A, F11R, NOV, MYOC, TUBA1A, UXT, CHFR, CDKN1A, MYOG, NCAM1, ADAM22, MAP1B, LIMK2, HSPA5, PDE4D, CTSD, CDON, HUWE1, IGFBP6, TIAM1, APLNR, COL4A1, HDAC4, PF4, GSN, LBX1, APC, BAG4, COL6A3, SLC37A4, FBN1, JAG1, CDK2, SCD, SACS, ANTXR1, EGF, EMX2, LOXL2, DIAPH1, FXN, KCNMA1, ANXA1, AGT, PEX11A, DPP4, IRF4, RBM38, HDAC1, TEC, COL1A1, NTRK2, B4GALT1, TSPYL2, SPARC, MEF2C, CDKN1B, C5orf13*
Cell Morphology (60)	*CHRNA1, MAP1B, CHRNB1, CDC42EP2, PCOLCE, VEGFA, RB1, PAK1, FIGF, CHRNA3, AQP7, TIAM1, ANXA2, NFKB2, GSN, SLC2A4, PLOD1, APC, PTBP1, PLA2G6, NPR1, FBN1, JAG1, CDK2, FYN, DYNLT1, FN1, ANTXR1, PTPN14, EGF, SEPT11, COL4A2, PRKCZ, SMARCA4, DIAPH1, ARRDC3, GADD45A, KCNMA1, FXN, ANXA1, DAB2, CHEK2, AGT, PLAT, EPHB4, SOD1, DISC1, ACTB, XBP1, FADD, COL1A1, F11R, MYOC, NTRK2, CDKN1A, SPARC, MYOG, CDKN1B, COL15A1, NCAM1*
Cellular Movement (81)	*MAP1B, LIMK2, PDE4D, VEGFA, LAMC1, CTSD, PAK1, FBLN1, CLEC4M, PPBP, FIGF, CDH13, TAB1, S100A10, IGFBP6, PDE2A, TIAM1, APLNR, COL4A1, ORM1/ORM2, SEMA5A, PF4, CD93, IGFBP5, ANXA2, MMP2, NFKB2, GSN, LBX1, APC, NPC1, PLA2G6, PIP5K1C, SLC37A4, HAS2, JAG1, AOC3, FYN, FN1, NFIX, MMP15, PTPN14, EGF, COL4A2, PRKCZ, FSTL1, DIAPH1, ARRDC3, GADD45A, ANXA1, DAB2, LBP, LASP1, PLAT, AGT, DPP4, CTSK, EPHB4, SOD1, TNFRSF1A, DISC1, ACTB, HDAC1, SERPINF1, FADD, COL1A1, F11R, HP, NTRK2, NOV, TUBA1A, B4GALT1, CHFR, ANXA3, CDKN1A, PROCR, FKBP4, SPARC, CDKN1B, C5orf13, NCAM1*
DNA Replication, Recombination, and Repair (23)	*NOS1, IGFBP6, PDE2A, COL14A1, ADCY2, FN1, TNFRSF1A, HDAC1, EGF, IGFBP5, MMP2, PDE4D, SMARCA4, PRKCZ, VEGFA, RB1, NOV, CDKN1A, SPARC, CDKN1B, CDK2, CHEK2, AGT*
Cell Cycle (44)	*FYN, FN1, PPP2R5B, EGF, EMX2, COL4A2, SEPT11, PRKCZ, SMARCA4, VEGFA, RB1, NOTCH2, MAP4, GADD45A, FIGF, CDH13, TAB1, CHEK2, AGT, DPP4, ADCY2, SSTR4, SOD1, HDAC4, TNFRSF1A, HDAC1, XBP1, IGFBP5, MLF1, SSRP1, APC, FADD, COL1A1, RUNX1T1, NTRK2, CHFR, CDKN1A, BAP1, HAS2, SPARC, RGL2, MYOG, CDKN1B, CDK2*
Cell Death (120)	*MIF4GD, VEGFA, RB1, PAK1, MADD, PPBP, FBLN1, ATF4, MFAP5, MMP2, IGFBP5, NFKB2, PLA2G6, RUNX1T1, HAS2, STEAP3, DRAM1, FYN, FN1, PTPN9, COL4A2, SMARCA4, PRKCZ, UCHL1, FSTL1, NOTCH2, MAP4, GADD45A, ANTXR2, DAB2, LBP, CAMK2B, BMP1, PLAT, ABCB1, EPHB4, TNFRSF1A, ACTB, DAB2IP, SERPINF1, XBP1, SIN3A, FADD, HP, TUBA1A, UXT, MYOC, CDK20, CHFR, CDKN1A, PROCR, FKBP4, SGCD, PTGR1, NCAM1, CHRNA1, Tnfrsf22/Tnfrsf23, MAP1B, DPP7, DHODH, HSPA5, PDE4D, CTSD, FDFT1, CYFIP2, NOD1, S100A10, HUWE1, AQP7, IGFBP6, TIAM1, HDAC4, PF4, BACE1, SLC1A1, GSN, SSRP1, BAG4, APC, PFKM, NPC1, ANKRD1, MBTPS1, NUP210, SLC37A4, BAP1, FBN1, JAG1, CDK2, NOS1, SCD, PRDX5, ACER2, ANTXR1, EGF, LOXL2, FXN, KCNMA1, ANXA1, DUSP10, NEDD8, SLC40A1, PLIN3, CCBL1, CHEK2, AGT, DPP4, SOD1, IRF4, TRPC1, PHLDA3, HDAC1, RYR2, AIFM2, TEC, NTRK2, B4GALT1, SPARC, MEF2C, CDKN1B*

IPA BioFunction	Molecules
Cell-To-Cell Signaling and Interaction (68)	*CHRNA1, CHRNB1, SNTA1, HSPA5, TNXB, VEGFA, LAMC1, PAK1, CYFIP2, CDON, PPBP, CLEC4M, FIGF, AMICA1, CD84, CDH13, CHRNA3, S100A10, TIAM1, COL4A1, ORM1/ORM2, HDAC4, SEMA5A, PF4, CD93, IGFBP5, ANXA2, MMP2, GSN, APC, PLA2G6, PIP5K1C, HAS2, SYN3, CD248, JAG1, CDK2, AOC3, FYN, FN1, PTPN14, MMP15, EGF, LOXL2, DIAPH1, NOTCH2, MAP4, ANXA1, DAB2, LBP, CAMK2B, PLAT, AGT, DPP4, DISC1, ACTN2, HDAC1, SERPINF1, RAMP1, MYO5A, F11R, MYOC, NTRK2, B4GALT1, PROCR, SPARC, IGHA1, NCAM1*
Cell Signaling (10)	*PDE2A, RAB3B, B4GALT1, PIP5K1C, AMPH, SYT9, PDE4D, LPHN1, AGT, NCAM1*
Molecular Transport (56)	*CHRNB1, GBP5, LPHN1, CTSD, FDFT1, LPIN1, PPBP, AMPH, CHRNA3, RGL1, AQP7, SLC39A3, GSN, SLC2A4, NPC1, PLA2G6, KCNB1, PIP5K1C, MBTPS1, NPR1, SLC37A4, GCK, SLC22A5, STEAP3, CDK2, AOC3, NOS1, SCD, RAB3B, KCNN3, BHMT, EGF, PRKCZ, KCNMA1, SLC40A1, LBP, AGT, CAMK2B, ATP4B, ABCB1, TRPC1, SOD1, TNFRSF1A, MTHFR, RYR2, DAB2IP, XBP1, PRKAG3, SYT9, FADD, HP, B4GALT1, FKBP4, MYOG, CDKN1B, NCAM1*
Small Molecule Biochemistry (40)	*SCD, BHMT, EGF, GBP5, PDE4D, CTSD, FDFT1, LPIN1, KCNMA1, PPBP, LBP, AGT, ABCB1, AQP7, ADCY2, PDE2A, TRPC1, SOD1, MTHFR, TNFRSF1A, DAB2IP, XBP1, SLC1A1, GSN, SLC2A4, FADD, NPC1, PLA2G6, HP, B4GALT1, PIP5K1C, MBTPS1, NPR1, CDKN1A, GCK, FKBP4, HAS2, ME2, ALDH18A1, CDKN1B*
Amino Acid Metabolism (10)	*TNFRSF1A, MTHFR, BHMT, CDKN1A, EGF, ME2, ALDH18A1, SLC1A1, CDKN1B, AGT*
Carbohydrate Metabolism (37)	*AOC3, SCD, EGF, GBP5, PRKCZ, VEGFA, UGDH, GYS1, PPBP, LBP, RGL1, AGT, AQP7, ABCB1, ADCY2, TRPC1, SOD1, TNFRSF1A, TKT, PF4, DAB2IP, XBP1, PRKAG3, PYGL, PISD, GSN, SLC2A4, ALG9, PLA2G6, PYGM, B4GALT1, PIP5K1C, SLC37A4, GCK, HAS2, MAN2B1, MYOG*
Lipid Metabolism (30)	*SCD, BHMT, EGF, GBP5, CTSD, FDFT1, LPIN1, KCNMA1, PPBP, LBP, AGT, AQP7, ABCB1, TRPC1, SOD1, TNFRSF1A, DAB2IP, XBP1, GSN, SLC2A4, FADD, NPC1, PLA2G6, HP, B4GALT1, PIP5K1C, MBTPS1, NPR1, GCK, FKBP4*
Protein Degradation (19)	*FYN, DPP4, CTSK, DPP7, EGF, BACE1, MMP2, APC, PCOLCE, UBE2G2, UCHL1, CTSD, DIAPH1, CHFR, NEDD8, CDKN1B, PLAT, AGT, BMP1*
Protein Synthesis (19)	*FYN, DPP4, CTSK, DPP7, EGF, BACE1, MMP2, APC, PCOLCE, UBE2G2, UCHL1, CTSD, DIAPH1, CHFR, NEDD8, CDKN1B, PLAT, AGT, BMP1*
C) Physiological System Development and Function	
Tumor Morphology (13)	*SERPINF1, COL4A2, APC, VEGFA, COL1A2, COL1A1, RB1, GADD45A, ANXA3, CDKN1A, FIGF, CDKN1B, CDK2*

IPA BioFunction	Molecules
Connective Tissue Development and Function (33)	*COL14A1, FN1, EGF, SMARCA4, PCOLCE, VEGFA, RB1, MAP4, LPIN1, GADD45A, PPBP, ATF4, CHEK2, AGT, PLAT, PGK1, TNFRSF1A, PF4, HDAC1, IGFBP5, MMP2, NFKB2, GSN, APC, SIN3A, NOV, CDKN1A, SPARC, FBN1, CDKN1B, MYOG, JAG1, CDK2*
Organismal Survival (47)	*FYN, NFIX, COL4A2, CEBPG, VEGFA, FDFT1, RB1, NOTCH2, MADD, GADD45A, FXN, ANXA1, AMPH, FBLN1, ANTXR2, TAB1, PLAT, PEX11A, AGT, COL4A1, SOD1, TNFRSF1A, TKT, HDAC1, RYR2, XBP1, BACE1, IGFBP5, NFKB2, PISD, PLOD1, SIN3A, APC, FADD, HP, NTRK2, PIP5K1C, MBTPS1, NPR1, CDKN1A, PROCR, FBN1, HAS2, MEF2C, CDKN1B, CDK2, GNAL*
Immune Cell Trafficking (33)	*AOC3, FYN, FN1, EGF, PDE4D, PRKCZ, VEGFA, DIAPH1, PAK1, ANXA1, PPBP, CLEC4M, FIGF, LBP, PLAT, S100A10, AGT, TIAM1, COL4A1, TNFRSF1A, PF4, CD93, MMP2, NFKB2, GSN, PLA2G6, F11R, COL1A1, HP, B4GALT1, PIP5K1C, SLC37A4, PROCR*
Nervous System Development and Function (45)	*FYN, CHRNA1, NFIX, FN1, ADAM22, SNTA1, MAP1B, CHRNB1, EGF, EMX2, VEGFA, LAMC1, RB1, PAK1, KCNMA1, CDON, CHRNA3, AGT, PLAT, CAMK2B, TIAM1, APLNR, COL4A1, DISC1, SEMA5A, HDAC1, SERPINF1, CANX, BACE1, IGFBP5, GSN, LBX1, APC, MYO5A, NPC1, NTRK2, TUBA1A, PIP5K1C, MEF2C, SYN3, CDKN1B, JAG1, C5orf13, CDK2, NCAM1*
Skeletal and Muscular System Development and Function (39)	*NOS1, CHRNA1, FN1, NDRG2, SNTA1, CHRNB1, EGF, PDE4D, SMARCA4, PCOLCE, LAMC1, VEGFA, RB1, PAK1, SMPX, KCNMA1, CDH13, CHRNA3, PLAT, AGT, SOD1, COL4A1, ACTN2, RYR2, MMP2, IGFBP5, GSN, LBX1, MYL9, F11R, NOV, STBD1, NPR1, CDKN1A, SPARC, FBN1, MYOG, CDKN1B, CDK2*
Organismal Development (44)	*SPTBN1, AOC3, NOS1, FN1, EGF, COL4A2, HSPA5, PDE4D, SMARCA4, COL1A2, VEGFA, LAMC1, FDFT1, RB1, FIGF, AGT, PLAT, PGK1, EPHB4, PDE2A, ORM1/ORM2, STK36, COL4A1, TNFRSF1A, TKT, PF4, SEMA5A, HDAC1, SERPINF1, IGFBP5, ANXA2, MMP2, F11R, COL1A1, NTRK2, B4GALT1, NPR1, ANXA3, SPARC, MEF2C, CDKN1B, JAG1, ECM1, COL3A1*
Tissue Development (75)	*CHRNA1, ADAM22, MAP1B, TNXB, LAMC1, VEGFA, FDFT1, RB1, CYFIP2, CDON, PPBP, FBLN1, FIGF, AMICA1, CD84, CDH13, CRISPLD2, TIAM1, SEMA5A, PF4, CD93, BACE1, MMP2, IGFBP5, NFKB2, GSN, APC, PLA2G6, COL6A3, PIP5K1C, NPR1, MBTPS1, JAG1, MUSTN1, ECM1, NOS1, AOC3, FYN, NFIX, FN1, MMP15, PTPN14, EGF, LOXL2, SMARCA4, DIAPH1, NOTCH2, ANXA1, DAB2, AGT, ABI3BP, PLAT, BMP1, FBXW4, CTSK, EPHB4, TNFRSF1A, ACTN2, HDAC1, SERPINF1, MYO5A, F11R, MYOC, NTRK2, B4GALT1, CDKN1A, OLFML2B, PROCR, FKBP4, SPARC, MEF2C, IGHA1, MYOG, SGCD, NCAM1*

IPA BioFunction	Molecules
Organ Development (28)	*SPTBN1, FN1, SMARCA4, VEGFA, GYS1, DPT, ATF4, FIGF, TAB1, AGT, Pdlim3, TNFRSF1A, XBP1, SERPINF1, IGFBP5, NFKB2, LBX1, MYO5A, FADD, B4GALT1, CDKN1A, SLC22A5, HAS2, FBN1, MEF2C, CDKN1B, JAG1, COL3A1*
Tissue Morphology (37)	*NOS1, CHRNA1, NFIX, FN1, CHRNB1, SNTA1, EGF, COL4A2, PDE4D, PRKCZ, PCOLCE, VEGFA, RB1, SMPX, KCNMA1, FIGF, CHRNA3, AGT, PLAT, COL4A1, SOD1, ACTN2, RYR2, MMP2, IGFBP5, GSN, LBX1, APC, MYL9, STBD1, CDKN1A, TSPYL2, SPARC, HAS2, CDKN1B, MYOG, CDK2*
Cardiovascular System Development and Function (44)	*NOS1, AOC3, FN1, EGF, COL4A2, ADAMTS2, SMARCA4, VEGFA, LAMC1, COL1A2, RB1, KCNMA1, ANXA1, FIGF, CDH13, PLAT, AGT, PGK1, EPHB4, PDE2A, ORM1/ORM2, COL4A1, MTHFR, TNFRSF1A, PF4, SEMA5A, HDAC1, SERPINF1, MMP2, ANXA2, BAG4, FADD, COL1A1, F11R, NTRK2, B4GALT1, NPR1, ANXA3, PROCR, SPARC, MEF2C, JAG1, ECM1, COL3A1*
Hematological System Development and Function (51)	*CHRNA1, HSPA5, PDE4D, VEGFA, RB1, LPIN1, PPBP, FIGF, CD84, S100A10, TIAM1, COL4A1, GABPB2, PF4, CD93, MMP2, NFKB2, GSN, PLA2G6, RUNX1T1, PIP5K1C, SLC37A4, JAG1, CDK2, AOC3, FYN, FN1, EGF, Gp49a/Lilrb4, PRKCZ, SMARCA4, DIAPH1, GADD45A, ANXA1, DUSP10, LBP, PLAT, AGT, DPP4, IRF4, TNFRSF1A, MLF1, FADD, TEC, F11R, COL1A1, HP, B4GALT1, CDKN1A, MEF2C, CDKN1B*
Behavior (34)	*AOC3, NOS1, FYN, ADAM22, CRY2, EGF, VEGFA, UCHL1, SEC24D, KCNAB1, KCNMA1, FXN, AMPH, KCNC1, CHRNA3, AGT, S100A10, PLAT, ATP4B, SOD1, NTAN1, TNFRSF1A, BACE1, SLC1A1, PLOD1, MYO5A, NPC1, NTRK2, NOV, SLC22A5, GCK, MAN2B1, GNAL, NCAM1*

Table 4.2 IPA canonical pathway

Ingenuity Canonical Pathways	Associated Molecules
Hepatic Fibrosis / Hepatic Stellate Cell Activation (15)	*FN1, TNFRSF1A, EGF, IGFBP5, MMP2, NFKB2, COL1A2, VEGFA, MYL9, COL1A1, IL1RL2, FIGF, LBP, COL3A1, AGT*
Cyclins and Cell Cycle Regulation (10)	*RB1, HDAC4, PPM1J, CDKN1A, HDAC1, PPP2R5B, HDAC10, CDKN1B, SIN3A, CDK2*
Aryl Hydrocarbon Receptor Signaling (13)	*NFIX, NFKB2, SMARCA4, CTSD, RB1, NCOA7, ALDH1L2, NEDD8, CDKN1A, ALDH18A1, CDKN1B, CHEK2, CDK2*
Starch and Sucrose Metabolism (9)	*UCHL1, PGM2, UGDH, GYS1, PYGM, GCK, PYGL, AMY2A, AMY1A (includes others)*
Cell Cycle: G1/S Checkpoint Regulation (8)	*RB1, HDAC4, CDKN1A, HDAC1, HDAC10, CDKN1B, SIN3A, CDK2*
Actin Cytoskeleton Signaling (14)	*TIAM1, FN1, ACTN2, ACTB, EGF, LIMK2, GSN, APC, MYL9, DIAPH1, PAK1, CYFIP2, PIP5K1C, LBP*
Endoplasmic Reticulum Stress Pathway (4)	*MBTPS1, XBP1, ATF4, HSPA5*
Calcium Signaling (12)	*MYL9, CHRNA1, TRPC1, HDAC4, RYR2, CHRNB1, HDAC1, HDAC10, ATF4, MEF2C, CHRNA3, CAMK2B*
Pancreatic Adenocarcinoma Signaling (9)	*VEGFA, RB1, CDKN1A, FIGF, EGF, NFKB2, CDKN1B, SIN3A, CDK2*
p38 MAPK Signaling (9)	*FADD, PLA2G6, IL1RL2, MAPKAPK3, TNFRSF1A, DUSP10, ATF4, MEF2C, TAB1*
Molecular Mechanisms of Cancer (18)	*FYN, ADCY2, STK36, NFKB2, PRKCZ, APC, SIN3A, FADD, RB1, PAK1, CDKN1A, CDKN1B, GNAL, CHEK2, CDK2, TAB1, BMP1, CAMK2B*
Bladder Cancer Signaling (8)	*VEGFA, RB1, MMP15, CDKN1A, FIGF, EGF, MMP2, SIN3A*
Chronic Myeloid Leukemia Signaling (8)	*RB1, HDAC4, CDKN1A, HDAC1, HDAC10, NFKB2, CDKN1B, SIN3A*

- Network

In order to elucidate how genes are involved in aging in mice GM, 25 gene networks were constructed for these genes and their partners by IPA analysis (Table 4.3). Among them, gene network 1 and 14 highlighted the functions of connective tissue disorder, genetic disorder and cellular assembly and organization.

Overall, the IPA analyses identified some key biological functions, pathways and networks that appear to be altered in aging mice through gene up- or down- regulation. These differentially expressed genes alter the protein production while some proteins are connective tissue components.

4.3.4 Transcriptomic changes in extracellular space during aging

Of the 407 significantly altered molecules, 54 were located in extracellular spaces. As shown in Table 4.1, 9 were collagen; 6 were peptidase, 5 were growth factor and 6 were enzyme. I found some ECM proteins which have been studied in chapter 2, i.e.: tenascin X (encoded by tenascin XB). Of those molecules, 6 were up-regulated between mice at 11M and 25M; the other 48 were down-regulated. I also noted these 54 molecules were involved in 20 networks but mainly in network 1 (27.8%) and network 7 (24.1%).

Table 4.3 IPA biological network

ID	Top Functions	Molecules in Network
1	Connective Tissue Disorders, Genetic Disorder, Dermatological Diseases and Conditions	*ANTXR1, ANTXR2, BMP1, COL15A1, COL1A2, COL3A1, COL4A1, COL4A2, COL6A1, COL6A3, collagen, Collagen type III, Collagen type IV, CTSK, ERK1/2, F11R, FIGF, FXN, Gelatinase, KCNMA1, KCNN1, KCNN3, LAMC1, MPZL1, NADPH oxidase, ORM1/ORM2, PCOLCE, Pdgf Ab, Plasminogen Activator, PLAT, Potassium Channel, PPBP, SERPINF1, TMEFF2, TSPYL2*
2	Cell Signaling, Nucleic Acid Metabolism, Cellular Development	*ACTN2, ADCY, ADCY2, Alpha tubulin, AOC3, ATF7IP, DISC1, EPHB4, FSH, G protein beta gamma, GNAL, hCG, Insulin, Lh, ME2, MYL9, NDRG2, NFkB (family), NNAT, NPR1, NTRK2, PDE4D, Pdlim3, PEX6, PGK1, PLIN3, PTPN9, Ras, SACS, SPTBN1, SSRP1, STBD1, SYT9, TUBA1A, VEGFA*
3	Cardiovascular Disease, Genetic Disorder, Immunological Disease	*AChR, ANKRD1, B4GALT1, BACE1, BAG4, Calcineurin protein(s), CANX, Cbp, CHRNA1, CHRNA3, CHRNB1, Cyclin D, FCGRT, GPD1, HAS2, Hdac, HDAC4, HDAC10, IL1RL2, MEF2, MEF2C, MYOG, MYPN, NCOA7, NFAT (complex), NFkB (complex), Nicotinic acetylcholine receptor, Notch, NOTCH2, NRG, PYGM, RGL1, RIC3, SNTA1, TRAFD1*
4	Cell Morphology, Cellular Development, Cellular Assembly and Organization	*ACTB, Alpha Actinin, Alpha catenin, ANXA1, ANXA2, AQP7, ARFRP1, Arp2/3, atypical protein kinase C, CAMK2B, CDC42EP2, CORO2B, Cpla2, CPNE2, CYFIP2, DIAPH1, F Actin, G-Actin, GSN, IRS, KCNAB1, LASP1, LRP, MADD, MYO5A, PI3K (complex), PLA2G6, Pld, PRKCZ, Profilin, RAB3B, S100A10, SLC40A1, TIAM1, TRP*
5	Cardiovascular Disease, Cardiovascular System Development and Function, Drug Metabolism	*14-3-3, AGT, ANXA3, Calpain, Cathepsin, COL14A1, Cytochrome c, DPH1, EHD2, ERP29, FDFT1, FKBP10, Hsp27, HSPA5, JINK1/2, KDELR3, LDL, MHC Class I (complex), Nfat (family), OGDH, PF4, Pkc(s), PLC gamma, PTBP1, Ras homolog, SEC24D, SGCD, SLC2A4, Smad2/3, SOD1, TAB1, TCR, TEC, VAV, XBP1*
6	Cellular Assembly and Organization, Cell Cycle, Hair and Skin Development and Function	*ADAM22, APC, APLNR, Beta Arrestin, Calbindin, Caveolin, CD84, CHEMOKINE, DYNLT1, EGF, FLOT2, Focal adhesion kinase, FYN, G protein, Gpcr, Gsk3, IGHA1, IKK (complex), KCNB1, KCNG4, LPHN1, MBTPS1, Metalloprotease, Mmp, NCKAP5, OPN1LW, PLC, PROCR, Shc, SKAP2, SLC37A4, SRC, SSTR4, STAT, ZFP106*
7	Cardiovascular Disease, Cardiovascular System Development and Function, Embryonic Development	*7S NGF, Alp, ANKH, COL1A1, Collagen Alpha1, Collagen type I, Collagen(s), DPP4, DPT, ECM1, Elastase, FBLN1, FBN1, Fibrin, FN1, Growth hormone, Igf, Igfbp, IGFBP5, IGFBP6, Importin alpha, Integrin, JAG1, Laminin, Mapk, MAPKAPK3, MFAP5, MMP2, MMP15, MYOC, NOV, PRELP, RFX5, SMOOTH MUSCLE ACTIN, Tgf beta*
8	Carbohydrate Metabolism, Lipid Metabolism, Small Molecule Biochemistry	*Akt, AMPK, CaMKII, CDH13, CDON, CHFR, CLCN5, Cofilin, CPEB1, GADD45A, GCK, Glycogen synthase, GYS1, HMG CoA synthase, HMGCS1, LPIN1, MAP2K1/2, Mlc, N-cor, NMDA Receptor, NPC1, Nr1h, p70 S6k, Pak, Pdgf (complex), PDGF BB, PP1 protein complex group, PP2A, PPM1J, PPP2R5B, SCD, SLC1A1, SPARC, Tnfrsf22/Tnfrsf23, UGDH*
9	Cell Cycle, Connective Tissue Development and Function, Cancer	*ACER2, Caspase, Cdc2, Cdk, CDK2, CDK20, CDKN1B, CHEK2, Ctbp, Cyclin A, Cyclin B, Cyclin E, DAB2IP, DUSP10, E2f, Estrogen Receptor, FSTL1, Hat, HDAC1, Jnk, LIMK2, MAP4, Mek, MLF1, NEDD8, NFIX, NOD1, Nuclear factor 1, Rb, RB1, Sos, Tenascin, TNXB, TSH, UXT*

ID	Top Functions	Molecules in Network
10	Cellular Compromise, Infection Mechanism, Infectious Disease	*Actin, ATF4, BCL7A, CCDC101, CD93, CEBPG, ENO3, GABPB2, GABRR1, Hist2h4 (includes others), Histone h3, Histone h4, IRF4, LOXL2, MAP1B, Mi2, MID1, mir-9, NFATC2IP, PAK1, PARP10, PLA2, PLEKHG2, Rac, RNA polymerase II, RPL12, RUNX1T1, SIN3A, SMARCA4, TACSTD2, Tdg (includes others), Vegf, ZBTB5, ZNF83, ZNF238*
11	Cancer, Hepatic System Disease, Carbohydrate Metabolism	*26s Proteasome, ABCB1, ARRDC3, C5orf13, CDKN1A, Ck2, CRISPLD2, CTSD, Fibrinogen, H2AFZ, Hemoglobin, HIST2H2BE (includes others), HISTONE, HP, HUWE1, IFN Beta, IgG, IL1, IL12 (complex), Immunoglobulin, Interferon alpha, LBP, MARS, MICAL2, NFKB2, P38 MAPK, p85 (pik3r), Pka, Pro-inflammatory Cytokine, RPS15A, Tlr, Trypsin, UBE2G2, Ubiquitin, UCHL1*
12	Cell Death, Cellular Function and Maintenance, Cell Cycle	*ADH1C, AIFM2, AMY1A (includes others), BAP1, BMF, C11orf83, DNMT3L, DRAM1, FAM173A, FASTKD2, FBXO11, LPIN1, MED13L, MRPS16, MYOD1, NMT1, NR3C1, PARP2, PHLDA3, PRKAB1, PRKAG3, RBM38, RDH11, SF3B14, SNRPN, SON, STEAP3, TKT, TP53, TRIM28, TSPYL2, UCHL1, UNC45A, ZNF180, ZNF346*
13	Cell Death, Cell Signaling, Cellular Assembly and Organization	*Adaptor protein 2, AMPH, Ap1, Ap2 alpha, Calmodulin, Clathrin, Creb, DAB2, ERK, FADD, FKBP4, GABPB2, Hsp70, Hsp90, KCNC1, LOXL1, NCAM1, Nos, NOS1, PFK, PFKM, Pias, PIP5K1C, Raf, Rock, RYR2, Sapk, SESN3, Smad, Sod, TNFRSF1A, TRPC1, Tubulin, Vla-4, ZNF440/ZNF808*
14	Cellular Assembly and Organization, Connective Tissue Disorders, Genetic Disorder	*ABHD11, CKAP4, COL4A1, COL5A2, DEGS1, Dgk, DGKD, EGFR, EMX2, F2, FBN1, FKBP9, GAK, GLI1, GPER, IER2, IFRD2, MEOX1, mir-1, MPND, MRC2, MYCN, PGM2, PGM2L1, PLOD1, POGK, RPS2, STK36, SYNE1, TAGLN2, TRIM2, TUFM, UCHL1, USP30, USP36*
15	Gene Expression, Lipid Metabolism, Nucleic Acid Metabolism	*ACOX1, ACOX3, AFG3L1, C11orf10, C11orf73, C19orf43, C3orf26, CDC5L, GOLIM4, GSTK1, HNF1A, HNF4A, HTRA2, KPNB1, MLF1IP, MRPL2, ONECUT1, OPA1, PPARA, PPARD, PPARG, RXRA, RXRB, SART3, SF3A2, SLC38A4, SMURF2, SNRPA, TMEM63A, TMTC4, TSPAN14, TXNDC12, WBP4*
16	Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Cellular Function and Maintenance	*ADAMTS2, BGN, BSG, C1q, CASP12, CCBL1, CD248, Cg, CIRBP, COL1A2, CYP11A1, DPP7, EFEMP1, ENTPD2, ERI3, FAM198B, GAS6, GBP5, GVIN1, HLA-DQA1, HP1BP3, IFNG, ITGA1, KLF10, LGALS3BP, MAN2B1, Mannosidase Alpha, NCAM1, PRNP, SMAD7, Sod, SUMO2, TAB1, TGIF1, UBA2*
17	Cell Morphology, Cellular Development, Organ Morphology	*ADIG, ANKS1A, BEND7, CCNY, CDK14, CLDN4, CPT2, DECR1, EMD, ERBB2, FGD6, GIT2, HTRA1, IGF1R, KIF2A, KLHDC3, LMO7, NAA20, NPTN, NUP210, OSBPL3, PISD, PLEKHB1, PRDX5, RHPN2, Rpl10a, RPS10, TGFBI, TGFBR1, UBXN6, VCP, VHL, YWHAG, ZAK, ZFP36L2*
18	Antigen Presentation, Cell-To-Cell Signaling and Interaction, Inflammatory Response	*AK2, C1q, CD209, CDX1, DDX58, DMBT1, FRZB, HAS2, HMGB1L1, HSPC159, LGALS8, Lilrb3 (includes others), LY96, MAP4K5, MFAP4, MURC, OLFML2B, PANK4, PDE2A, PDE4B, PDE8A, PEA15, PEX11A, PLEKHA3, PRTN3, PYCARD, RTP3, SFTPD, SPARC, TNF, TOR1AIP2, WNT3A, WNT7A, ZFP36, ZFP36L1*
19	Cell-To-Cell Signaling and Interaction, Cellular Assembly and Organization, Nervous System Development and Function	*ABI3BP, AMY2A, ASB12, ATP4B, ATP5O, CDKL5, DLG4, DLGAP3, DLGAP4, GAB3, GRB2, KERA, KRT8, MARCH10, MKL1, MYLK, MYOCD, MYRIP, NADK, NCKAP5, NOS1AP, OLFML3, PABPN1, PHACTR1, PPP3CB, PYGO1, RBMXL1, RIMS2, SEMA6A, SNX8, SYN1, SYN2, SYN3, TOX4, USP32*

ID	Top Functions	Molecules in Network
20	Cell-mediated Immune Response, Cellular Movement, Hematological System Development and Function	*AMICA1, ATP13A2, CARD11, CHAD, DHODH, EEF1A2, ELOVL1, EPPK1, GARS, GART, IKBKG, IL2, Integrin alpha 2 beta 1, ITGB1, KCTD20, LAMC1, MAP3K8, MARS, MAST2, mir-124, NIK, PELI3, PHF7, PRKCQ, PTPN7, SDCCAG8, SEC11A, SEPT9, SEPT11, SLC22A5, STX2, TIFA, TMBIM1, TRAF6, TUBB6*
21	Cell Death, Cell-To-Cell Signaling and Interaction, Skeletal and Muscular System Development and Function	*ABL1, ADD3, BCL2, CALCA, CALCRL, CGNL1, CRCP, CTNNB1, ECSIT, EIF2D, EPHB4, FBXW4, JUN, KIF26A, LIMA1, LOXL3, LRBA, MAP3K1, MEPE, MTHFR, NDRG2, PTH, PTPN14, RAD51, RAMP1, RGL2, SEMA5A, SFRP2, SLC1A5, SMPX, SNX3, SPNS1, STAT5a/b, WIF1, ZDHHC16*
22	Carbohydrate Metabolism, Small Molecule Biochemistry, Cellular Growth and Proliferation	*ABHD16A, ADH5, ALDH, ALDH16A1, ALDH18A1, ALDH1L2, ALDH3A1, ALDH6A1, ALT, CHD1, CHD3, CLEC4M, FBP2, FMO5, GATAD2A, GATAD2B, IGFBP6, ITGAM, LBX1, LEP, MBD3 (includes EG:17192), MBD3L1, MLLT11, MTA3, NuRD, OPLAH, OSM, PDLIM5, POU5F1, PYGL, RALY, Timp, TSPAN6, UCP1, ZGPAT*
23	Lipid Metabolism, Small Molecule Biochemistry, Hematological Disease	*AK3, Aox3, ARNTL, ARRB1, ATP6V1F, ATP6V1H, BHMT, CRY2, Cryptochrome, CYP4F2, CYP8B1, EBP, HDGFRP2, HNF4A, IL5, IL13, NAA50, NAT8, PEMT, PGM1, PMVK, PSAT1, QSOX1, RDH11, RORA, Scd2, SERPINB3, SLC26A6, SLC41A3, SPCS2, SREBF2, TPM4, TRIP11, TTC1, YKT6*
24	Antigen Presentation, Cell-To-Cell Signaling and Interaction, Hematological System Development and Function	*BAG6, CCDC107, CDKL2, COPS5, CPSF1, CSTF2, DNAJA1, EEF1G, EIF3E, EIF4G1, EIF4G3, GET4, Gp49a/Lilrb4, Hnrpfl, Igh-1a, IGHM, IMMT, KIAA1377, LHPP, MALT1, MIF4GD, PELI2, PRRC2A, QSOX1, SIRPA, TNFRSF14, TNFRSF13B, TNFSF13, TRAF2, TRAF3, TRAF5, TRPT1, UBAC2, UBQLN4, VPREB3*
25	Reproductive System Development and Function, Organismal Development, Endocrine System Development and Function	*FAM114A1, SCHIP1*

4.4 Discussion

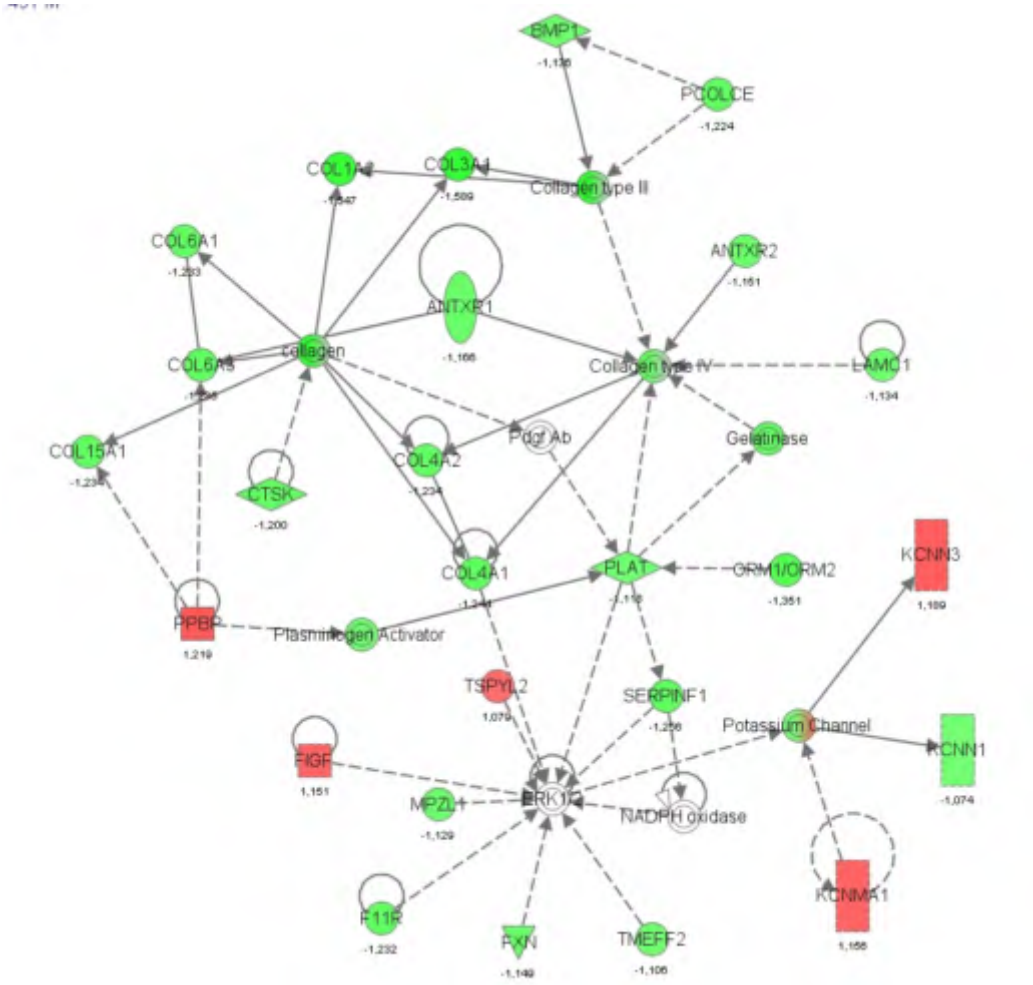


Figure 4.9 Gene network for connective tissue disorder, visual representation of the principal network (score: 35) generated containing 32 molecules. Each node represents a gene and its association with other genes is represented by a line (solid line: bind to; solid line with solid arrowhead: directly act on; solid line with hollow arrowhead: directly translate to; dotted line with solid arrow head: indirectly act on). Node has different shapes to represent different molecule types (horizontal diamonds=peptidases, cube=cytokine, dotted cube=growth factor, dotted rectangle=ion channel, vertical ellipse= transmembrane receptor, circle=other, concentric circles=complex or groups). The 32 molecules in this network have a color background: 5 molecules in red are up-regulated; 26 molecules in green are down-regulated; 1 molecule in red-green are not significantly varied. Molecules without background color were undetected in this study but have been inserted by IPA to produce a highly connected.

In order to elucidate how genes involved in connective tissue disorder might be implicated during muscle aging, a network was constructed for ECM

genes and their partners. Most ECM partners with altered gene expression were found to be down-regulated during the 11M-25M period. Collagens were key players of this network, but peptidases and proteases involved in procollagen processing and in collagen degradation were also differentially expressed.

For one thing, I noticed the down-regulation of several collagen genes between 11M and 25M. They encompassed type I (COL1A2), type III (COL3A1), type IV (COL4A1, COL4A2), type VI (COL6A1, COL6A3) and type XV (COL15A1) collagens.

Fibrillar collagens are synthesized as procollagens, i.e. a precursor molecules [150], and I noticed the down-regulation of several enzymes implicated in procollagen processing. Procollagen processing involves the removal of the N-propeptide by aminoprocollagen peptidases ADAMTS (a disintegrin and metalloproteinas with thrombospondin), and the removal of C-propeptide by BMP-1 (bone morphogenetic protein-1) [151], which is stimulated by PCOLCE (procollagen C-proteinase enhancer) [152, 153]. Procollagen processing also involves the oxidative determination of lysine residues and the resulting crosslinking are essential for CT integrity. BMP-1 also exhibits lysyl oxidase activity and catalyzes collagen lysine deamination [154]. The down-regulation of several COL and of ADAMTS-2, BMP-1 and PCOLCE therefore suggest potential alternation in collagen synthesis and processing during muscle aging.

On the other hand, my data on gene expression also emphasize profound alterations in the ECM degrading machinery. Increasing attention has recently

been focused on cathepsin K (encoded by CTSK genes) because it is the potent ECM hydrolase, i.e. representing 98% of the total cysteine protease activity of collagenases [155]. Cathepsin K mRNA has been detected in a variety of tissue including skeletal muscle [155], but its functional role for skeletal muscle has not been reported. In human inflammation site, cathepsin K increase the degradation of human collagen(s) proteins [156]. In my model, if the down-regulation CTSK gene causes low cathepsin K level, collagen degradation can be expected to be reduced (Figure 4.9).

Tissue plasminogen activator (PLAT) activates several collagenases (I and IV) [157]. In this model, if the down-regulation of the PLAT gene leads to low levels of PLAT protein, this will decrease collagenase activities and similarly reduce the degradation of ECM components (Figure 4.9).

Gelatinases include MMP2 and MMP9. Both are associated with ECM fibrosis and degrade basement membrane collagen (Col-IV) [158-161]. I noted the down-regulation of gelatinases in the old mouse muscle, which is also in agreement with an increase in basement membrane collagen (Col-IV) and in ECM protein accumulation.

In sum, the down-regulation of PLAT, CTSK, MMP-2 and MMP-9 genes may attenuate the ECM degradation and thus be associated with fibrosis.

In addition, collagen(s) syntheses were studied by transcriptomic analyses. Down-regulated collagens at 25M implicated that collagen(s) syntheses was also attenuated. Taking into account the accumulation of ECM

that I observe at the protein level, a stronger reduction in ECM degradation than ECM synthesis is expected. However, current study can not assess this comparison since the mediators and controllers affecting the protein expression are manifold: translational rates, post-translational modifications, secretion etc. Therefore, further studies are required.

Summary

The molecular events causing ECM accumulation (fibrosis) during aging in vivo are poorly understood. Here, *gastrocnemius muscle* from mouse at young (2M), adult (11M) and advanced old (25M) were studied by transcriptomic analysis and their gene expression pattern was evaluated. I noted that profound alterations occurred between 11M and 25M and that the majority of them are down-regulated. I used IPA to analyze the biofunctions, canonical pathways and gene network. The principle finding was the altered balance between ECM protein synthesis and degradation which eventually leads to the accumulation of ECM protein. Some down-regulated genes corresponding to the degradation enzyme, thus decreased degradation was expected. Meanwhile, genes of collagens were down-regulated, which might refer to a decreased synthesis of ECM collagen. Since ECM protein accumulated during aging, I suggest that the stronger reduction takes place in ECM degradation than synthesis.

CHAPTER 5 CONCLUSION AND FUTURE WORK

5.1 Conclusion

Skeletal muscles become less well adapted to their function during aging. Not only do myofibres undergo atrophy but connective tissue also changes. However, the age-related change in connective tissue is poorly studied, to the best of my knowledge. In this project, I have selected *gastrocnemius* muscle from young (2M), adult (11M), early old (22M) and advanced old (25M) C57BL/6 mice. I have utilized histological approaches coupled with image analyses, biochemical approaches, bioinformatics microarray analyses to investigate the age-related structural and compositional variations and their regulating mechanism.

Firstly, structural and compositional variations of mouse skeletal muscle were collected and analyzed independently. Study of myofibre morphology offered interesting insight into the relationship between the cross-sectional area and shape of myofibre. With increased area, I found that myofibres are prone to be angular, regular and polygonal. With increased age, a larger proportion of angular myofibre and a global reduction of CSA result in myofibre atrophy. Study of myonuclei during aging argued that myonuclear domain size is unalterable during aging. On the contrary, myonuclear domain size decreases, which is due to the reduced CSA associated with changeless myonuclei. Changeless myonuclei might be adapted to the nuclei loss (apoptosis) and / or

regenerations. I thus investigate the apoptosis and muscle regeneration in skeletal muscle. I found that a significant proportion of satellite cells undergo apoptosis, which may be critical for routine myofibre maintenance and regeneration. I further noticed that the majority of apoptosis occurred in capillary endothelial cells. It may trigger the liberation of endorepellin fragments. In my study, an abundant LG1+LG2 fragment was detected in the old muscle.

Additional alterations in connective tissue include the fibrosis associated with the structural variation. My result of ECM structure firstly underlines the increased connection points and thickened segments of ECM, which offers a visual interpretation and profound understanding of altered structure during fibrosis. I studied the distribution and co-localization of some ECM component. The semi-quantitative result implies the potential roles of TNX and Col-VI in fibrillogenesis and fibrosis.

Secondly, an integrative approach was proposed to analyze those structural and compositional variations. I noted that aging studies often focus on specific biological events rather than on a multitude of levels simultaneously. I hereby developed a data mining strategy to illuminate the sensitivity of the individual structural or compositional property to age. I also provided a visualization representation to evaluate the global difference between mice from different age groups. My results suggest apoptosis play a crucial role in distinguishing mice from different age groups. I also showed that the overall age-associated variations start after 11M and accentuates after 22M.

Thirdly, transcriptomic analysis was employed to investigate the global variation in mouse skeletal muscle. This attempt to define gene expression pattern in response to increased age presents significant variations in 3 major biofunctions, 13 canonical pathways and 25 gene networks between mice at 11M and 25M.

I have reported the fibrosis in the old mice muscle, and I hypothesize that fibrosis could result from an accentuated protein synthesis and/or from an attenuated protein degradation. My results from transcriptomic data analyses indicated the bone morphogenetic protein-1, procollagen C-proteinase enhancer and peptidases (matrix metalloproteinases, cathepsin K, plasminogen activator and a disintegrin and metalloproteinases with thrombospondin type 1 motif 2), involved in ECM degradation, were down-regulated. Thus if the less expressed proteins were caused by the down-regulated corresponding genes, attenuated degradation of ECM was suggested. In addition, I found that the major component of ECM components, i.e. collagens etc, are down-regulated. If attenuated collagen synthesis has resulted from these down-regulations, age-dependent increased fibrosis would be derived from a stronger reduction in degradation than in synthesis. Nevertheless, I can not compare the activity of synthesis and degradation based on the current analyses. But I noted another possibility is that mRNA levels do not perfectly represent protein levels, which also depend on translational rates, post-translational modifications and secretion. Additional studies are therefore needed to specify the molecular mechanism(s) that determine fibrosis in the old skeletal muscle.

5.2 Future outlook

With respect to the disadvantage of this on-going study, continued efforts are expected on the transcriptomic studies. First, some genes are capable of coding for several proteins and being interpreted by the alternative splicing among exon and intron. Therefore, studies of alternative splicing will be conducted by Genespring GX in the near future. Significantly altered exon(s) will be countered and grouped into different genes.

Next, to access the quality of micro-array measurement, RT-PCR validation will be carried out on a basis of a random selection of transcripts among my list. To verify the biological events identified above, RT-PCR validation will be carried out on specific sets. To analyze the alternative splicing events, transcriptomic data will be further analyzed.

With mouse exon array, I was allowed to gather more probe level data which determine the overall level of expression. However, the genome is only a source of information. In order to produce protein, the transcription of genes is followed by the translation of mRNA. Therefore, study of the complete set of proteins produced by genome, so called proteome, is expected. Proteome is complex because each protein can be chemically modified in different way after synthesis. Proteome is also dynamic because it varies considerably in differing circumstances due to different patterns of gene expression and different patterns of proteins modification. A large-scale analysis of protein expression can help to identify the main protein found in the mouse muscle. Therefore technologies such as 2D-PAGE and mass spectrometry will be introduced.

In chapter 2, I used semi-quantitative method to estimate the variations of ECM components because of the tiny sample amount and the difficulties in extracting ECM molecules. However, the semi-quantitative method based on the basis of image processing is not “real quantification” thus I need further study to verify my observation(s). There are some techniques available: flow cytometry (accounting and examining the microscopic particles), western-blot (electrotransfer and immunoblot the target molecules), enzyme-linked immunosorbent assay (ELISA, can determine protein concentration in cell/tissue extracts), analyses by mass spectrometry of muscle trypsic hydrolysats. Then I will test and develop optimal protocol(s) for quantifying the ECM molecule(s) of interest.

REFERENCE LIST

1. Dormont, B., M. Grignon, and H. Huber, *Health expenditure growth: reassessing the threat of ageing.* Health Economics, 2006. **15**(9): p. 947-963.
2. Harman, D., *The aging process: major risk factor for disease and death.* Proc Natl Acad Sci USA, 1991. **88**(12): p. 5360-3.
3. Yu, B.P., *Aging and oxidative stress: modulation by dietary restriction.* Free Radic Biol Med, 1996. **21**(5): p. 651-68.
4. Dirks, A. and C. Leeuwenburgh, *Apoptosis in skeletal muscle with aging.* Am J Physiol Regul Integr Comp Physiol, 2002. **282**(2): p. R519-27.
5. Franceschi, C., *Cell proliferation, cell death and aging.* Aging (Milano), 1989. **1**(1): p. 3-15.
6. Ståhlberg, D., B. Angelin, and K. Einarsson, *Age-related changes in the metabolism of cholesterol in rat liver microsomes.* Lipids, 1991. **26**(5): p. 349-52.
7. Marino, M., et al., *Age-dependent accumulation of dolichol in rat liver: is tissue dolichol a biomarker of aging?* J Gerontol A Biol Sci Med Sci, 1998. **53**(2): p. B87-93.
8. Serrano, A.L. and P. Muñoz-Cánoves, *Regulation and dysregulation of fibrosis in skeletal muscle.* Experimental Cell Research, 2010: p. 1-9.
9. Combaret, L., et al., *Skeletal muscle proteolysis in aging.* Curr Opin Clin Nutr Metab Care, 2009. **12**(1): p. 37-41.
10. Narici, M.V. and N. Maffulli, *Sarcopenia: characteristics, mechanisms and functional significance.* British medical bulletin, 2010: p. 1-21.
11. Yan, Z., et al., *Regulatory elements governing transcription in specialized myofiber subtypes.* J Biol Chem, 2001. **276**(20): p. 17361-6.
12. Voermans, N., et al., *Clinical and molecular overlap between myopathies and inherited connective tissue diseases.* Neuromuscular Disorders, 2008. **18**(11): p. 843-856.

13. Piec, I., *Etude protéomique du vieillissement du muscle squelettique chez le rat LOU/c/jall*, in *Ecole Doctorale des Sciences de la Vie et de la Sante*2005, UNIVERSITE BLAISE PASCAL: St-Genès-Champanelle. p. 1-308.
14. Clark, K., et al., *S TRIATEDM USCLEC YTOARCHITECTURE: An Intricate Web of Form and Function.* Annu. Rev. Cell Dev. Biol., 2002. **18**(1): p. 637-706.
15. Bailey, A.J., S.P. Robins, and G. Balian, *Biological significance of the intermolecular crosslinks of collagen.* Nature, 1974. **251**(5471): p. 105-9.
16. Light, N. and A.E. Champion, *Characterization of muscle epimysium, perimysium and endomysium collagens.* Biochem J, 1984. **219**(3): p. 1017-26.
17. LeBleu, V.S., B. MacDonald, and R. Kalluri, *Structure and function of basement membranes.* Experimental Biology and Medicine, 2007.
18. Bonaldo, P., et al., *Structural and Functional Features of the a3 Chain Indicate a Bridging Role for*

Chicken Collagen VI in Connective Tissues. Biochemistry, 1990. **29**: p. 1245-1254.

19. Gelse, K., E. Pöschl, and T. Aigner, *Collagens--structure, function, and biosynthesis.* Adv Drug Deliv Rev, 2003. **55**(12): p. 1531-46.
20. Zhang, G., et al., *Development of tendon structure and function: regulation of collagen fibrillogenesis.* J Musculoskelet Neuronal Interact, 2005. **5**(1): p. 5-21.
21. Hannesson, K.O., et al., *An immunological study of glycosaminoglycans in the connective tissue of bovine and cod skeletal muscle.* Comp Biochem Physiol B, Biochem Mol Biol, 2007. **146**(4): p. 512-20.
22. Zanotti, S., et al., *Decorin and biglycan expression is differentially altered in several muscular dystrophies.* Brain, 2005. **128**(Pt 11): p. 2546-55.
23. Iozzo, R., J. Zoeller, and A. Nyström, *Basement membrane proteoglycans: modulators par excellence of cancer growth and angiogenesis.* Molecules and cells, 2009. **27**(5): p. 503-513.

24. Iozzo, R. and K. Danielson, *Transcriptional and Posttranscriptional Regulation of Proteoglycan Gene Expression.* Progress in Nucleic Acid Research and Molecular Biology, 1998. **62**: p. 19-53.
25. Cornelison, D.D.W., M.S. Filla, and H.M. Stanley, *Syndecan-3 and syndecan-4 specifically mark skeletal muscle satellite cells and are ….* Developmental Biology, 2001.
26. Jones, F.S. and P.L. Jones, *The tenascin family of ECM glycoproteins: structure, function, and regulation during embryonic development and tissue remodeling.* Dev Dyn, 2000. **218**(2): p. 235-59.
27. Hynes, R., *Molecular biology of fibronectin.* Annual review of cell biology, 1985.
28. Flurkey, K. and J.M. Currer, *Pitfalls of animal model systems in ageing research.* est Practice & Research Clinical Endocrinology & Metabolism, 2004. **18**(3): p. 407-421.
29. Brooks, S.V. and J.A. Faulkner, *Contractile properties of skeletal muscles from young, adult and aged mice.* The Journal of Physiology, 1988. **404**: p. 71-82.
30. Moran, A.L., G.L. Warren, and D.A. Lowe, *Soleus and EDL muscle contractility across the lifespan of female C57BL/6 mice.* Exp Gerontol, 2005. **40**(12): p. 966-75.
31. KANEKO, N., et al., *Puberty acceleration in mice. II. Evidence that the vomeronasal organ is a receptor for the primer pheromone in male mouse urine.* Biology of Reproduction, 1980. **22**(4): p. 873-878.
32. Tanaka, S., et al., *Establishment of an Aging Fram of F344/N Rats and C57BL/6 Mice at National Institute for Longevity Sciences (NILS).* Archives of Gerontology and Geriatrics, 2000. **30**: p. 215-223.
33. Rowlatt, C., F.C. Chesterman, and M.U. Sheriff, *Lifespan, age changes and tumour incidence in an ageing C57BL mouse colony.* Lab Anim, 1976. **10**(10): p. 419-42.

34. Taillandier, D., et al., *Role of protein intake on protein synthesis and fiber distribution in the unweighted soleus muscle.* J Appl Physiol, 1993. **75**(3): p. 1226-32.
35. Prodanov, D., et al., *Three-dimensional topography of the motor endplates of the rat gastrocnemius muscle.* Muscle & nerve, 2005.
36. Cosmas, C., et al., *Age and diet alter skeletal muscle tubular aggregates.* Age, 1995. **18**: p. 69-78.
37. Restriction, C., *Interventions in Aging and Age-Related Diseases: The Present and the Future.* Age, 2006. **28**: p. 1-75.
38. Cristea, A., et al., *Effects of aging and gender on the spatial organization of nuclei in single human skeletal muscle cells.* Aging Cell, 2010. **9**(5): p. 685-97.
39. Leeuwenburgh, C., et al., *Age-related differences in apoptosis with disuse atrophy in soleus muscle.* Am J Physiol Regul Integr Comp Physiol, 2005. **288**(5): p. R1288-96.
40. Evans, W.J. and W.W. Campbell, *Sarcopenia and age-related changes in body composition and functional capacity.* The Journal of nutrition, 1993.
41. Lexell, J., *Human aging, muscle mass, and fiber type composition.* J Gerontol A Biol Sci Med Sci, 1995. **50 Spec No**: p. 11-6.
42. Bruusgaard, J.C. and K. Gundersen, *In vivo time-lapse microscopy reveals no loss of murine myonuclei during weeks of muscle atrophy.* J Clin Invest, 2008. **118**(4): p. 1450-7.
43. Brooks, N., M. Schuenke, and R. Hikida, *Ageing influences myonuclear domain size differently in fast and slow skeletal muscle of rats.* Acta Physiologica, 2009. **197**(1): p. 55-63.
44. Gundersen, K. and J.C. Bruusgaard, *Nuclear domains during muscle atrophy: nuclei lost or paradigm lost?* The Journal of Physiology, 2008.
45. Mauro, A., *Satellite cell of skeletal muscle fibers.* Journal of Cell Biology, 1961.

46. Zammit, P.S., T.A. Partridge, and Z. Yablonka-Reuveni, *The skeletal muscle satellite cell: the stem cell that came in from the cold.* J Histochem Cytochem, 2006. **54**(11): p. 1177-91.
47. Snijders, T., L.B. Verdijk, and L.J. van Loon, *The impact of sarcopenia and exercise training on skeletal muscle satellite cells.* Ageing Res Rev, 2009. **8**(4): p. 328-38.
48. Yablonka-Reuveni, Z. and J.E. Anderson, *Satellite cells from dystrophic (mdx) mice display accelerated differentiation in primary cultures and in isolated myofibers.* Dev Dyn, 2006. **235**(1): p. 203-12.
49. Harris, B.A., *The influence of endurance and resistance exercise on muscle capillarization in the elderly: a review.* Acta physiologica Scandinavica, 2005. **185**(2): p. 89-97.
50. Davidson, Y.S., et al., *The effect of aging on skeletal muscle capillarization in a murine model.* The journals of gerontology. Series A, Biological sciences and medical sciences, 1999. **54**(10): p. B448-51.
51. Cui, L., et al., *Arteriolar and venular capillary distribution in skeletal muscles of old rats.* J Gerontol A Biol Sci Med Sci, 2008. **63**(9): p. 928-35.
52. Grounds, M.D., L. Sorokin, and J. White, *Strength at the extracellular matrix-muscle interface.* Scand J Med Sci Sports, 2005. **15**(6): p. 381-91.
53. Kääriäinen, M., et al., *Relation between myofibers and connective tissue during muscle injury repair.* Scandinavian Journal of Medicine & Science in Sports, 2000. **10**(6): p. 332-7.
54. Eyre, D.R., M.A. Paz, and P.M. Gallop, *Cross-linking in collagen and elastin.* Annu Rev Biochem, 1984. **53**: p. 717-48.
55. Zimmerman, S.D., et al., *Age and training alter collagen characteristics in fast- and slow-twitch rat limb muscle.* J Appl Physiol, 1993. **75**(4): p. 1670-4.
56. Eyre, D., M. Weis, and J. Wu, *Advances in collagen cross-link analysis.* Methods, 2008. **45**(1): p. 65-74.
57. Ricard-Blum, S. and G. Ville, *Collagen cross-linking.* Int J Biochem, 1989. **21**(11): p. 1185-9.

58. Brandstetter, F. and E. Schuller, *Azorubin clearance in normal pregnancy.* Z Geburtshilfe Gynakol, 1954. **142**(1): p. 37-44.
59. Romero, N., et al., *Immunohistochemical demonstration of fibre type-specific isozymes of cytochrome c oxidase in human skeletal muscle.* Histochemistry, 1990. **94**(2): p. 211-5.
60. Blume, G., et al., *Polymyositis with cytochrome oxidase negative muscle fibres. Early quadriceps weakness and poor response to immunosuppressive therapy.* Brain, 1997. **120 (Pt 1)**: p. 39-45.
61. James, J., et al., *Histophotometric estimation of volume density of collagen as an indication of fibrosis in rat liver.* Histochemistry, 1986. **85**(2): p. 129-33.
62. James, J., et al., *Sirius red histophotometry and spectrophotometry of sections in the assessment of the collagen content of liver tissue and its application in growing rat liver.* Liver, 1990. **10**(1): p. 1-5.
63. Flint, F.O. and K. Pickering, *Demonstration of collagen in meat products by an improved picro-Sirius Red polarisation method.* The Analyst, 1984. **109**(11): p. 1505.
64. Liu, A., T. Nishimura, and K. Takahashi, *Relationship between structural properties of intramuscular connective tissue and toughness of various chicken skeletal muscles.* Meat Science, 1996. **43**(1): p. 43-49.
65. Pulkkanen, K.J., et al., *False-positive apoptosis signal in mouse kidney and liver detected with TUNEL assay.* Apoptosis, 2000. **5**(4): p. 329-33.
66. Mohan, S. and E. Radha, *Hydroxyproline excretion and collagen catabolism in rats of different age groups.* Biochem Med, 1980. **24**(1): p. 1-5.
67. Etherington, D.J., D. Pugh, and I.A. Silver, *Collagen degradation in an experimental inflammatory lesion: studies on the role of the macrophage.* Acta Biol Med Ger, 1981. **40**(10-11): p. 1625-36.
68. WOESSNER, J.F., *The determination of hydroxyproline in tissue and protein samples containing small proportions of this imino acid.* Arch Biochem Biophys, 1961. **93**: p. 440-7.

69. Piec, I., et al., *Differential proteome analysis of aging in rat skeletal muscle.* The FASEB Journal, 2005. **19**(9): p. 1143-1145.
70. Yu, W. and Y. Chung…, *Vignetting distortion correction method for high quality digital imaging.* Pattern Recognition, 2004.
71. Schau, H.C., *Statistical filter for image feature extraction.* Appl Opt, 1980. **19**(13): p. 2182-90.
72. van der Zwet, P.M., et al., *Derivation of optimal filters for the detection of coronary arteries.* IEEE Trans Med Imaging, 1998. **17**(1): p. 108-20.
73. Gosker, H.R., et al., *Skeletal muscle fibre-type shifting and metabolic profile in patients with chronic obstructive pulmonary disease.* Eur Respir J, 2002. **19**(4): p. 617-25.
74. Lin, J., et al., *Age-related cardiac muscle sarcopenia: Combining experimental and mathematical modeling to identify mechanisms.* Exp Gerontol, 2008. **43**(4): p. 296-306.
75. Mongiat, M., et al., *Endorepellin, a novel inhibitor of angiogenesis derived from the C terminus of perlecan.* J Biol Chem, 2003. **278**(6): p. 4238-49.
76. Mouly, V. and J. Beauchamp, *Qu est-ce qu'une cellule muscularie satellite?* Medecine Sciences, 2003. **19**: p. 696.
77. de Resende, M.M., et al., *Role of endothelial cell apoptosis in regulation of skeletal muscle angiogenesis during high and low salt intake.* Physiol Genomics, 2006. **25**(2): p. 325-35.
78. White, K.K. and D.W. Vaughan, *Age effects on cytochrome oxidase activities during denervation and recovery of three muscle fiber types.* Anat Rec, 1991. **230**(4): p. 460-7.
79. Kirkeby, S. and C. Garbarsch, *Aging affects different human muscles in various ways. An image analysis of the histomorphometric characteristics of fiber types in human masseter and vastus lateralis muscles from young adults and the very old.* Histol Histopathol, 2000. **15**(1): p. 61-71.

80. Kwong, L.K. and R.S. Sohal, *Age-related changes in activities of mitochondrial electron transport complexes in various tissues of the mouse.* Arch Biochem Biophys, 2000. **373**(1): p. 16-22.
81. Yarovaya, N.O., et al., *Age-related atrophy of rat soleus muscle is accompanied by changes in fibre type composition, bioenergy decline and mtDNA rearrangements.* Biogerontology, 2002. **3**(1-2): p. 25-7.
82. Bianospino, E., et al., *Age-related changes in muscle fiber type frequencies and cross-sectional areas in straightbred and crossbred rabbits.* animal, 2008. **2**(11): p. 1627-1632.
83. Siu, P.M., et al., *Aging influences cellular and molecular responses of apoptosis to skeletal muscle unloading.* Am J Physiol, Cell Physiol, 2005. **288**(2): p. C338-49.
84. Lushaj, E.B., et al., *Sarcopenia accelerates at advanced ages in Fisher 344xBrown Norway rats.* J Gerontol A Biol Sci Med Sci, 2008. **63**(9): p. 921-7.
85. Judge, L.M. and J.S. Chamberlain, *Gene therapy for Duchenne muscular dystrophy: AAV leads the way.* Acta Myol, 2005. **24**(3): p. 184-93.
86. Ilyina-Kakueva, E., *Morphological Investigations of Skeletal Muscles of SLS-2 (SLS2 BSP3)*, 1993, NASA. p. 1-2.
87. Aravamudan, B., et al., *Denervation effects on myonuclear domain size of rat diaphragm fibers.* J Appl Physiol, 2006. **100**(5): p. 1617-22.
88. Brack, A.S., H. Bildsoe, and S.M. Hughes, *Evidence that satellite cell decrement contributes to preferential decline in nuclear number from large fibres during murine age-related muscle atrophy.* Journal of cell science, 2005. **118**(Pt 20): p. 4813-21.
89. Allen, D.L., R.R. Roy, and V.R. Edgerton, *Myonuclear domains in muscle adaptation and disease.* Muscle Nerve, 1999.
90. Bruusgaard, J., et al., *Number and spatial distribution of nuclei in the muscle fibres of normal mice studied in vivo.* The Journal of Physiology, 2003. **551**(2): p. 467-478.

91. Bruusgaard, J.C., K. Liestol, and K. Gundersen, *Distribution of myonuclei and microtubules in live muscle fibers of young, middle-aged, and ....* Journal of Applied Physiology, 2006.
92. Brooks, N.E., M.D. Schuenke, and R.S. Hikida, *No change in skeletal muscle satellite cells in young and aging rat soleus muscle.* J Physiol Sci, 2009. **59**(6): p. 465-71.
93. Hansen, G., et al., *Effects of spaceflight on myosin heavy-chain content, fibre morphology and succinate dehydrogenase activity in rat diaphragm.* Pflugers Arch, 2004. **448**(2): p. 239-47.
94. Conboy, I.M. and T.A. Rando, *Aging, stem cells and tissue regeneration: lessons from muscle.* Cell Cycle, 2005. **4**(3): p. 407-10.
95. Hughes, S.M. and S. Schiaffino, *Control of muscle fibre size: a crucial factor in ageing.* Acta Physiol Scand, 1999. **167**(4): p. 307-12.
96. Shefer, G., et al., *Satellite-cell pool size does matter: defining the myogenic potency of aging skeletal muscle.* Developmental Biology, 2006. **294**(1): p. 50-66.
97. Jejurikar, S.S., et al., *Aging increases the susceptibility of skeletal muscle derived satellite cells to apoptosis.* Experimental gerontology, 2006. **41**(9): p. 828-36.
98. Podhorska-Okolow, M., et al., *Apoptosis of myofibres and satellite cells: exercise-induced damage in skeletal muscle of the mouse.* Neuropathol Appl Neurobiol, 1998. **24**(6): p. 518-31.
99. Malmgren, L.T., C.E. Jones, and L.M. Bookman, *Muscle fiber and satellite cell apoptosis in the aging human thyroarytenoid muscle: a ....* Otolaryngology-Head and Neck Surgery, 2001.
100. Gopinath, S.D. and T.A. Rando, *Stem cell review series: aging of the skeletal muscle stem cell niche.* Aging Cell, 2008. **7**(4): p. 590-8.
101. McClung, J.M., et al., *Caspase-3 regulation of diaphragm myonuclear domain during mechanical ventilation-induced atrophy.* Am J Respir Crit Care Med, 2007. **175**(2): p. 150-9.

102. Haus, J., et al., *Collagen, cross-linking, and advanced glycation end products in aging human skeletal muscle.* Journal of Applied Physiology, 2007. **103**(6): p. 2068.
103. Mohan, S. and E. Radha, *Age related changes in muscle connective tissue: acid mucopolysaccharides and structural glycoprotein.* Exp Gerontol, 1981. **16**(5): p. 385-92.
104. Lethias, C., et al., *A model of tenascin-X integration within the collagenous network.* FEBS Lett, 2006. **580**(26): p. 6281-5.
105. Elefteriou, F., et al., *Binding of tenascin-X to decorin.* FEBS Lett, 2001. **495**(1-2): p. 44-7.
106. Wiberg, C., et al., *Biglycan and decorin bind close to the n-terminal region of the collagen VI triple helix.* J Biol Chem, 2001. **276**(22): p. 18947-52.
107. Bristow, J., et al., *Tenascin-X, collagen, elastin, and the Ehlers-Danlos syndrome.* Am J Med Genet C Semin Med Genet, 2005. **139C**(1): p. 24-30.
108. Minamitani, T., H. Ariga, and K. Matsumoto, *Deficiency of tenascin-X causes a decrease in the level of expression of type VI collagen.* Exp Cell Res, 2004. **297**(1): p. 49-60.
109. Hunter, P.J. and T.K. Borg, *Integration from proteins to organs: the Physiome Project.* Nat Rev Mol Cell Biol, 2003. **4**(3): p. 237-43.
110. Pearson, R.A., *Sectionim social and economic science.* Science, 1901. **14**(363): p. 912-26.
111. Jain, A.K., *Statistical pattern recognition: a review.* 2003: p. 1-34.
112. Ham, F.M., I.N. Kostanic, and G.M. Cohen…, *Determination of glucose concentrations in an aqueous matrix from NIR spectra using optimal time-domain filtering and partial least-squares regression.* … Engineering, 1997.
113. T. Jolliffe, I., *Principal component analysis.* Springer, 2002: p. 487.
114. Jain, A.K., M.N. Murty, and P.J. Flynn, *Data clustering: a review.* ACM Comput. Surv., 1999. **31**(3): p. 264-323.
115. Kothari, R. and D. Pitts, *On finding the number of clusters.* Pattern Recognition Letters, 1999. **20**(4): p. 405-416.

116. Bailey Jr, T.A. and R. Dubes, *Cluster validity profiles.* Pattern Recognition, 1982. **15**(2): p. 61-83.
117. Day, W.H.E. and Edelsbrunner, H, *Efficient algorithms for agglomerative hierarchical clustering methods.* Journal of classification, 1984. **1**: p.7-24.
118. Lê, S., J. Josse, and F. Husson, *FactoMineR: An R package for multivariate analysis.* Journal of Statistical Software, 2008. **25**: p. 1-18.
119. Jolliffe, I., *Principal Component Analysis*. Encyclopedia of Statistics in Behavioral Science2005: John Wiley & Sons, Ltd.
120. Gardner, J.W., *Detection of vapours and odours from a multisensor array using pattern recognition Part 1. Principal component and cluster analysis.* Sensors and Actuators B: Chemical, 1991. **4**(1-2): p. 109-115.
121. Hartigan, J.A. and M.A. Wong, *A K-Means Clustering Algorithm.* Journal of the Royal Statistical Society. Series C (Applied Statistics), 1979. **28**(1): p. 100-108.
122. K. Jain, A. and R. C. Dubes, *Algorithms for clustering data.* 1988: p. 320.
123. Jung, Y., et al., *A Decision Criterion for the Optimal Number of Clusters in Hierarchical Clustering.* Journal of Global Optimization, 2003. **25**(1): p. 91-111.
124. Fred, A.L. and A.K. Jain, *Combining multiple clusterings using evidence accumulation.* IEEE Trans Pattern Anal Mach Intell, 2005. **27**(6): p. 835-50.
125. Lorr, M., ed. *Cluster analysis for social scientists: Techniques for analyzing and simplifying complex blocks of data*. 1st ed. 1983, Jossey-Bass: San Francisco. xvi, 233.
126. Ward, J., *Hierarchical grouping to optimize an objective function.* Journal of the American Statistical Association, 1963. **58**(301): p. 236-244.
127. Kaufman, L. and P.J. Rousseeuw, *Agglomerative Nesting (Program AGNES)*. Finding Groups in Data2008: John Wiley & Sons, Inc. 199-252.
128. Halkidi, M., Y. Batistakis, and M. Vazirgiannis, *Cluster validity methods: part I.* SIGMOD Rec., 2002. **31**(2): p. 40-45.

129. De Haan, J.R., et al., *Integrating gene expression and GO classification for PCA by preclustering.* BMC Bioinformatics, 2010. **11**: p. 158.
130. Hsieh, K.L. and I.C. Yang, *Incorporating PCA and fuzzy-ART techniques into achieve organism classification based on codon usage consideration.* Comput Biol Med, 2008. **38**(8): p. 886-93.
131. Zuendorf, G., et al., *Efficient principal component analysis for multivariate 3D voxel-based mapping of brain functional imaging data sets as applied to FDG-PET and normal aging.* Hum Brain Mapp, 2003. **18**(1): p. 13-21.
132. Feng, J., et al., *Principal component analysis reveals age-related and muscle-type-related differences in protein carbonyl profiles of muscle mitochondria.* J Gerontol A Biol Sci Med Sci, 2008. **63**(12): p. 1277-88.
133. Yeung, K.Y. and W.L. Ruzzo, *Principal component analysis for clustering gene expression data.* Bioinformatics, 2001. **17**(9): p. 763-74.
134. Murtagh, F., *A Survey of Recent Advances in Hierarchical Clustering Algorithms.* The Computer Journal, 1983. **26**(4): p. 354-359.
135. Marzetti, E. and C. Leeuwenburgh, *Skeletal muscle apoptosis, sarcopenia and frailty at old age.* Exp Gerontol, 2006. **41**(12): p. 1234-8.
136. Allison, D.B., et al., *Microarray data analysis: from disarray to consolidation and consensus.* Nat Rev Genet, 2006. **7**(1): p. 55-65.
137. Schena, M., et al., *Quantitative monitoring of gene expression patterns with a complementary DNA microarray.* Science, 1995. **270**(5235): p. 467-70.
138. Stoughton, R.B., *Applications of DNA microarrays in biology.* Annu Rev Biochem, 2005. **74**: p. 53-82.
139. Blencowe, B.J., *Alternative splicing: new insights from global analyses.* Cell, 2006. **126**(1): p. 37-47.
140. Black, D.L. and P.J. Grabowski, *Alternative pre-mRNA splicing and neuronal function.* Prog Mol Subcell Biol, 2003. **31**: p. 187-216.
141. Modrek, B. and C. Lee, *A genomic view of alternative splicing.* Nat Genet, 2002. **30**(1): p. 13-9.

142. Halder, S.K., J.S. Goodwin, and A. Al-Hendy, *1,25-Dihydroxyvitamin D3 reduces TGF-beta3-induced fibrosis-related gene expression in human uterine leiomyoma cells.* J Clin Endocrinol Metab, 2011. **96**(4): p. E754-62.
143. Dobaczewski, M., et al., *Smad3 signaling critically regulates fibroblast phenotype and function in healing myocardial infarction.* Circ Res, 2010. **107**(3): p. 418-28.
144. Kovalszky, I., et al., *Experimental and human liver fibrogenesis.* Scand J Gastroenterol Suppl, 1998. **228**: p. 51-5.
145. Chin, B.Y., et al., *Stimulation of pro-alpha(1)(I) collagen by TGF-beta(1) in mesangial cells: role of the p38 MAPK pathway.* Am J Physiol Renal Physiol, 2001. **280**(3): p. F495-504.
146. Baggio, B., E. Musacchio, and G. Priante, *Polyunsaturated fatty acids and renal fibrosis: pathophysiologic link and potential clinical implications.* J Nephrol, 2005. **18**(4): p. 362-7.
147. Roeb, E., et al., *Regulation of tissue inhibitor of metalloproteinases-1 gene expression by cytokines and dexamethasone in rat hepatocyte primary cultures.* Hepatology, 1993. **18**(6): p. 1437-42.
148. Batlle, M., et al., *Down-regulation of matrix metalloproteinase-9 (MMP-9) expression in the myocardium of congestive heart failure patients.* Transplant Proc, 2007. **39**(7): p. 2344-6.
149. Irizarry, R.A., et al., *Summaries of Affymetrix GeneChip probe level data.* Nucleic Acids Research, 2003. **31**(4): p. e15.
150. D. Hay, E., *Cell biology of extracellular matrix.* 1991: p. 468.
151. Colige, A., et al., *Domains and maturation processes that regulate the activity of ADAMTS-2, a metalloproteinase cleaving the aminopropeptide of fibrillar procollagens types I-III and V.* J Biol Chem, 2005. **280**(41): p. 34397-408.
152. Moschcovich, L., et al., *Folding and activity of recombinant human procollagen C-proteinase enhancer.* Eur J Biochem, 2001. **268**(10): p. 2991-6.

153. Moali, C., et al., *Substrate-specific modulation of a multisubstrate proteinase. C-terminal processing of fibrillar procollagens is the only BMP-1-dependent activity to be enhanced by PCPE-1.* J Biol Chem, 2005. **280**(25): p. 24188-94.

154. Borel, A., et al., *Lysyl oxidase-like protein from bovine aorta. Isolation and maturation to an active form by bone morphogenetic protein-1.* J Biol Chem, 2001. **276**(52): p. 48944-9.

155. Brömme, D. and K. Okamoto, *Human cathepsin O2, a novel cysteine protease highly expressed in osteoclastomas and ovary molecular cloning, sequencing and tissue distribution.* Biol Chem Hoppe-Seyler, 1995. **376**(6): p. 379-84.

156. Serveau-Avesque, C., et al., *Active cathepsins B, H, K, L and S in human inflammatory bronchoalveolar lavage fluids.* Biol Cell, 2006. **98**(1): p. 15-22.

157. Palotie, A., et al., *Types I and IV collagenolytic and plasminogen activator activities in preovulatory ovarian follicles.* J Cell Biochem, 1987. **34**(2): p. 101-12.

158. Sakalihasan, N., et al., *Activated forms of MMP2 and MMP9 in abdominal aortic aneurysms.* J Vasc Surg, 1996. **24**(1): p. 127-33.

159. Rougier, J.P., et al., *Hyperosmolality suppresses but TGF beta 1 increases MMP9 in human peritoneal mesothelial cells.* Kidney Int, 1997. **51**(1): p. 337-47.

160. Zhao, W., Y. O'Malley, and M.E. Robbins, *Irradiation of rat mesangial cells alters the expression of gene products associated with the development of renal fibrosis.* Radiat Res, 1999. **152**(2): p. 160-9.

161. Colin, C., et al., *High expression of cathepsin B and plasminogen activator inhibitor type-1 are strong predictors of survival in glioblastomas.* Acta Neuropathol, 2009. **118**(6): p. 745-54.

---

[i] 70. Yu, W. and Y. Chun., *Vignetting distortion correction method for high quality digital imaging.* Pattern Recognition, 2004.

[ii] 71. Schau, H.C., *Statistical filter for image feature extraction.* Appl Opt, 1980. **19**(13): p. 2182-90.

[iii] 72. van der Zwet, P.M., et al., *Derivation of optimal filters for the detection of coronary arteries.* IEEE Trans Med Imaging, 1998. **17**(1): p. 108-20.

PUBLICATION LIST

1. **Wang, H.J**., Meunier, B., Goh, K.L., Béchet, D., Listrat, A. and Lee, K., *Age-related variations in skeletal muscle structure and components: data mining approach.* Journal of Medical Imaging and Health Informatics, submitted.
2. **Wang, H.J.**, Listrat A., Béchet D., Meunier B., Lee K. and Goh K.L., *Age variations in the bimodal distributions of muscle fibre cross-sectional area and possible influence on active force generation*, Journal of Biomechanics, submitted
3. **Wang, H.J.**, Béchet, D., Meunier, B., Goh, K.L., Listrat, A. and Lee, K., *Effect of aging in the structure and composition of mouse skeletal muscle,* Aging Cell, in preparation
4. Yeo, Y.L., Goh, K.L., Kin, L., **Wang, H.J.**, Listrat, A. and Béchet, D., *Structure-Property Relationship of Burn Collagen Reinforcing Musculo-Skeletal Tissues,* Key Engineering Materials, 2011, **478**, p.87-92
5. **Wang, H.J.**, Meunier, B., Béchet, D., Lee, K., Labas, R., Goh, K.L., Léonard, C. and Listrat, A., *Image Analysis of the Intramuscular Connective Tissue*, 55th International Congress of Meat Science and Technology (ICoMST), Copenhagen, Denmark, 16th-21st Aug, 2009