TBX5 Mechanism of Action in Skeletal Muscle Cell Proliferation and Differentiation

Massomeh Sheikh-Hassani

Thesis submitted to the University of Ottawa
in partial Fulfillment of the requirements for the
Degree of Doctor of Philosophy

Department of Biochemistry, Microbiology and Immunology
Faculty of Medicine
University of Ottawa

Table of Contents

Abstract

Skeletal muscle development and function is governed by a conserved set of Transcription Factors (TFs) that regulate gene expression. The TF gene regulation is stimulus driven and cell-type and time point specific. TBX5 is an essential dosage sensitive regulator of heart and limb development. In the skeletal system, TBX5 is expressed in early stages in the lateral plate mesoderm and gives rise to the forelimb. TBX5 is also involved in proliferation and differentiation and survival pathways in both heart and limb development. Mutations in *TBX5* gene lead to HOS which is characterized by various types of cardiac and musculoskeletal defects. TBX5 mechanism of action and its spatiotemporal function in skeletal muscle development has yet to be fully understood. TBX5 regulation is controlled through various factors such as alternative splicing, protein-protein interactions, Post-Translational Modifications (PTMs) and microRNAs. To date, many TBX5 interactors have been identified in cardiac cells however TBX5 protein interactors and target genes in skeletal muscle cells have not been studied. Understanding the protein interactome of TBX5 in skeletal muscle will enhance the current understanding of its mechanism of action. In this study we have characterized TBX5 with focus on its regulation, expression and biochemical properties in cardiac and skeletal muscle cells and moreover its mechanism of action specifically in skeletal muscle proliferation and differentiation. Chapter 1 discusses TBX5 regulation through alternative splicing leading to the existence of 5 distinct TBX5 isoforms with variable transcriptional activity, cardiac and limb expression pattern, biochemical properties and function. We show the pro-proliferation role of TBX5a in myoblasts while TBX5c shows to be pro-differentiation leading to the formation of myotubes in skeletal muscle C2C12 myoblasts. This opposing role of the two TBX5 isoforms lead us to studying

TBX5 mechanism of action in proliferation and differentiation of skeletal muscle cells. In this study using a mass spectrometry-based approach we have identified novel TBX5 interacting partners in skeletal muscle cells for the first time by using stably overexpressed 3xFlag TBX5 via retroviral transduction in C2C12 cell line. Nuclear protein extracts were immunoprecipitated and sent for HPLC-ESI-MS/MS to identify potential protein partners of TBX5 in skeletal muscle cells. Moreover, the same stable cell line was used to identify TBX5 downstream target genes in these cell types by sending RNA extracts for microarray analysis. Amongst the 200 protein interactors identified, MYBBP1a and TBX5 interaction was confirmed and studied. The microarray analysis identified over 1200 differentially expressed genes and potential downstream targets of TBX5a from which Myostatin (*Mstn*) and Cyclin D2 (*CcnD2*) were both significantly upregulated and further confirmed and studied in relation to proliferation and differentiation in skeletal muscle cells. Chapter 2 focuses on the cooperative interaction between TBX5a and MYBBP1a inhibiting muscle specific gene promoter, Myogenin (*MyoG*). TBX5a and TBX5c seem to both interact with MYBBP1a but result in variable transcriptional activity of both *MyoG* and *Mstn* gene promoters. We show that TBX5 is upstream of *Mstn*, it binds to the promoter on specific TBE sites, and is able to upregulate Mstn promoter activation. In vivo, we show that MDX mice limb skeletal muscle tissues show elevated levels of TBX5, MYBBP1a and MSTN expression which suggest that the TBX5 pathway is associated with and indicative of the onset of proliferation and regeneration in MDX skeletal muscle tissue. Chapter 3 discusses the role of TBX5 in proliferation and regeneration of skeletal muscle cells by identifying that TBX5 binds to *CcnD2* promoter and upregulates its activation which is a known cell cycle gene critical in cell proliferation and survival. Moreover, we identify GATA4 as a TBX5a

cofactor in myoblast proliferation and show synergistic activation of *Ccnd2* promoter by cooperative TBX5a and GATA4 action. We further show that *Tbx5* heterozygote mice exhibit decreased levels of CCND2 and other proliferation markers, as well as decreased expression of PAX7 (marker of satellite cells) compared to WT skeletal muscle tissues. We also show that the heterozygous loss of *Tbx5* impairs the process of regeneration in a cardiotoxin-induced injury model in mouse limb tissues. *Tbx5* heterozygote mice exhibit less proliferation and impaired regeneration 4 days after injury, followed by decreased formation of regenerated fibers by 7 days post-injury compared to the wildtype mice skeletal muscle tissues; suggesting that TBX5 function is important in maintaining adult muscle regenerative capacity. Together, this study has characterized TBX5 isoforms and identified novel TBX5 protein partners and targets in the skeletal muscle cells and sheds light on TBX5 regulatory mechanism in proliferation and differentiation of skeletal muscle cells and its potential implications in HOS and other muscular diseases.

Acknowledgements

This journey has not only been an academic and scientific experience for me, but it has impacted my life in a much larger scale personally; in ways that I could not imagine at the starting point. It has most definitely been the most challenging and extraordinary experience of my lifetime. I have had many essential and valuable people along my side to help me last and thrive through this roller coaster journey.

My dear supervisor, Dr. Mona Nemer, has been an amazing mentor, supporting me in tough times while always encouraging me to push my limits and aim high. She has been not only my mentor but also an amazing role model who has redefined 'the sky's limit' for me by exceling in her professional career while holding on to her passion and values, science and mentorship, and keeping them close and dear to her heart. I have learned from her the most precious lessons and that dedication, hard work and passion, make your wildest dreams come true. Mona, I am so thankful for your wisdom and efforts in advancing me as a scientist and a human.

I thank the members of my thesis advisory committee, Dr. Jean-François Couture and Dr. Marjorie Brand, for their helpful feedback and support throughout the years. Their inputs have always been very valuable and have helped me direct my research goals for the best outcome.

It has been such a pleasure working alongside the amazing people of the Nemer Lab. Everyone has been extremely helpful, friendly and fun to work with. Our wise and smart research associate, Dr. Hiba Komati, has always helped me go through the rough times by hearing out my experimental and non-experimental problems. Her scientific input has been

crucial to my work and I thank her for her continuous encouragement, consultations and support. I would like to thank our lab manager, Janie Beauregard, for being such a great support at all times with all the tasks that required help in the lab and at the same time really caring individually like a lab mom (as we call her). I would like to also thank Megan Fortier who always calmly helped me with all my animal work and all her flexibility and kindness in understanding my project deadlines. Thanks to Dr. Alice Lau who has always been super helpful in helping me figure out experiments and being patient in teaching me new things.

I would like to say a special thank you to Dr. Jamie Whitcomb, for being an amazing scientist which I have had the pleasure to work with but more importantly a superb and joyful human being, who has always lightened up my day at low times. You have been so willing and positive about sharing your knowledge (and chocolates) and helping me out, and I have learned so much from you in many aspects. You are smart, hardworking, friendly and caring and I have been lucky to work by your side.

I thank all the other past and present members of the Nemer lab for their continuous support throughout the years: Dr. Wael Maharsy, Dr. Abir Yamak, Dr. Lara Gharibeh, Dr. Georges Kanaan, Dr. Rami Darwich, Dr. Smail Messaoudi, Abir Mazloum, Mathieu Joyal, Laura Collins, Claudia Teran, Arianna Rostami, Alicia Jurado-Acosta.

I would also like to acknowledge the tremendous support and love, I have received from my lovely family members, who have whole heartedly stood by my side throughout these challenging years. To my dear husband, Saeed, you have offered me unfailing love, support and understanding during my pursuit of this PhD degree, that has made the completion of this thesis possible. You have witnessed my ups and downs and have accompanied me on every valley and hill with all your presence. I thank you and love you from a very special

place in my heart and will forever remember the sacrifices you have made to allow pursuit of my dreams and goals. To my loving, caring and supportive parents, Kambiz and Mahnaz, thank you for always believing in me and teaching me the value of education and hard work. You have from an early age encouraged and hugely supported me to go after my passion and education; you are the true reason I am here today. I love you dearly and appreciate everything you have done for me. To my little (not so little anymore) brother, Mohsen, thank you for all the love and support you have given me and for inspiring me with your academic success and making me a proud sister by evolving and advancing in your professional career. To my dear in-laws, Ali and Mali, I'm grateful and thankful to have had you in my life throughout my PhD years. Your continuous support and encouragement throughout the years has been truly heartwarming; especially in the month that you were witnesses to my thesis writing and helped me get through the stressful times by bringing warmth to my heart and a smile to my face.

Finally, I thank all my friends and other family members for all their support, encouragement and understanding throughout my PhD years.

List of Figures

List of Tables

List of Abbreviations

AA	Amino Acids
ACTH	Adrenocorticotropin
AER	Apical Ectodermal Ridge
AHR	Aromatic Hydrocarbon Receptor
ASD	Atrial Septal Defect
AV	Atrioventricular
bHLH	Basic Helix-Loop-Helix
BMD	Becker Muscular Dystrophy
Bmp	Bone Morphogenetic Protein
BMPR1a	Bone Morphogenetic Protein Receptor 1a
CCND2	Cyclin D2
CDH	Congenital Diaphragmatic Hernias
CDK	Cyclin-Dependent Kinase
CDKI	CDK Inhibitor
Chaf1a	Chromatin Assembly Factor 1 Subunit A
CHD	Congenital Heart Disease
Chek1	Checkpoint Kinase 1
CMD	Congenital Muscular Dystrophy
c-MET	Tyrosine-Protein Kinase Met
COPD	Chronic Obstructive Pulmonary Disease
Crabp1	Cellular Retinoic Acid Binding Protein 1
CREB	cAMP Response Element-Binding protein
CRM1	Chromosome Region Maintenance 1
CT	Computed Tomography
CTX	Cardiotoxin
CXCR4	Chemokine Receptor Type 4
DEXA	Dual Energy X-ray Absorptiometry
DKK2	Dickkopf WNT Signaling Pathway Inhibitor 2
DM	Myotonic Dystrophy

DMD	Duchenne Muscular Dystrophy
DMD	Dystrophin gene
DMRT2	Doublesex- and Mab-Related Transcription Factor 2
ECM	Extracellular Matrix
EDMD	Emery–Dreifuss Muscular Dystrophy
EYA1/2	Eyes Absent Homolog 1/2
FGF	Fibroblast Growth Factor
FL	Forelimb
FOXC1	Forkhead Box Protein Factor C1
FOXC2	Forkhead Box Protein Factor C2
FSHD	Facioscapulohumeral Muscular Dystrophy
GDNF	Glial Cell Derived Neurotropic Factor
GJA5	Gap Junction Protein Alpha 5
HAND2	Heart- and Neural Crest Derivatives-Expressed Protein 2
HDAC3	Histone Deacetylase 3
HGF	Hepatocyte Growth Factor
HGNC	HUGO Gene Nomenclature Committee
HL	Hindlimb
HOS	Holt-Oram Syndrome
HPLC-ESI-MS/MS	High Performance Liquid Chromatography, Electrospray Ionization Tandem Mass Spectrometry
IFN-γ	Interferon-gamma
IGF	Insulin-like Growth Factor
Igfbp3	Insulin-Like Factor Binding Protein 3
IL-6	Interleukin-6
ISL1	Insulin Gene Enhancer Protein
Klf4	Krupple-Like Factor
Lbh	Limb-bud-and-heart gene
LBX1	Ladybird Homeobox 1
LGMD	Limb Girdle Muscular Dystrophy

Mcm4	Mini-Chromosome Maintenance 4
MEF2	Myocyte Enhancer Factor 2
MEOX2	Homeobox Protein MOX-2
MHC	Myosin Heavy Chain
MHC-α	α-Cardiac Myosin Heavy Chain
miR	MicroRNA
MMP	Matrix Metalloproteinase
MPC	Myogenic Precursor Cells
MRF	Myogenic Regulatory Factor
MRI	Magnetic Resonance Imaging
MSTN	Myostatin (Gdf8)
MSX1	Msh Homeobox 1
MYBBP1a	Myb-Binding Protein (P160) 1a
MYF5	Myogenic Factor 5
MYOD	Myogenic/Myoblast Determination Protein 1
MYOG	Myogenin
NES	Nuclear Export Signal
NLS	Nuclear Localization Sequence
nNOS	Neuronal Nitric Oxide Synthase
NPPA	Natriuretic Peptide A
NRD	Negative Regulatory Domain
NTX	Notexin
NURD	Nucleosome Remodeling and Deacetylase
PAX7	Paired Box 7
PBX1	Pre-B-cell Leukemia Transcription Factor 1
PDGF	Platelet-Derived Growth Factor
PDLIM7	PDZ and LIM Domain Protein 7
PGC-1α	Proliferator-Activated Receptor Gamma Coactivator 1-alpha
PIAS1	Protein Inhibitor of Activated STAT1
PITX2	Pituitary Homeobox 2

POMC	Pro-Opiomelanocortin
PREP1	Pre-Sequence Protease 1
PTM	Post-Translational Modification
RA	Retinoic Acid
RA	Rheumatoid Arthritis
Rb	Retinoblastoma
RMS	Rhabdomyosarcoma
Scn5a	Voltage-Gated Channel Alpha Subunit 5
SDF-1	Stromal Derived Factor-1
Sema7a	Semaphorin 7a
Shh	Sonic Hedgehog
SIM2	Single-Minded Homolog 2
SIX1/4	Sine Oculis Homeobox 1/4
SRF	Serum Response Factor
TA	Tibialis Anterior
TAD	Transcriptional Activation Domain
TAZ	Tafazzin
TBE	T-box Binding Element
TBX20	T-box Transcription Factor 20
TBX5	T-box Transcription Factor 5
TF	Transcription Factor
TGF-β	Transforming Growth Factor beta
TNF	Tumor Necrosis Factor
VD	Variable Domain
VEGF	Vascular Endothelial Growth Factor
VSD	Ventricular Septal Defect
VSD	Ventricular Septal Defects
WNT	Wingless-Related Integration Site
YAP1	Yes-Associated Protein 1
3′-UTR	3′-Untranslated Region

1. General Introduction

1.1. Anatomy and function of skeletal muscle

One of the major tissue/organ systems of the body is comprised of the musculoskeletal system [1]. The musculoskeletal system consists of skeletal muscle which attaches to the bones by tendons, generating bodily movements [1]. Skeletal muscle, along with cardiac and smooth muscle, is one of the three major muscle types in vertebrates. Also known as voluntary muscle, it is one of the most dynamic tissues of the human body. The structure of skeletal muscle comprises of many muscle fibers (also called muscle cells or myofibers) that are arranged and wrapped together by sheaths of connective tissue [2]. Each individual myofiber is made of several myofibrils which contain some myofilaments such as actin and myosin filaments which form various bands on the skeletal muscle. The bundled myofibrils are structured in a striated pattern and form the principle contractile component of the skeletal muscle known as the sarcomeres [2]. The human skeletal muscle comprises about 40% of the total body weight and approximately 50-75% of proteins of the body. Numerous factors affecting muscle mass rely on the equilibrium between protein synthesis and degradation, and due to their substantial contribution to mobility, health, and exercise capacity, proteins involved in regulatory, structural and contractile functions of skeletal muscle have been extensively studied [3]. Skeletal muscle contributes to various functions of the body mechanically, metabolically and by serving as a reservoir. The conversion of chemical to mechanical energy is one of the main mechanical functions of skeletal muscle to generate force and produce movement while maintaining posture. Skeletal muscle also contributes, metabolically, to basal energy metabolism and act as storage for amino acids and

carbohydrates, and ultimately maintaining the core temperature by producing heat [3]. It also acts as an amino acid reservoir when required by other tissues for organ-specific protein synthesis [2]. The amino acid reservoir plays a role in starvation by maintaining blood glucose levels after their release. Thus, a reduced muscle mass negatively impacts most skeletal muscle functions explained above and impairs the ability of the body to respond to stress and disease.

Various techniques have been developed for muscle mass quantification and functional analysis such as Magnetic Resonance Imaging (MRI), Computed Tomography (CT), urinary creatinine concentration quantification, Dual Energy X-ray Absorptiometry (DEXA), magnetic resonance spectroscopy and many others [2,3]. In general, muscle mass is mainly determined by size and number of muscle fibers, which are the major component of muscle. The number and size of muscle fibers are determined by various environmental and genetic factors, which can influence myogenesis. Histological techniques are used to observe cross sections of muscle tissue to evaluate area of the myofibers and their functional properties. Although many diagnostic tools have been generated but the science behind the cellular and molecular regulation of skeletal muscle development is not fully understood even though it has been extensively studied over the years [4,5].

1.2.Skeletal muscle development

1.2.1.Skeletal muscle myogenesis in embryo and adult

Skeletal muscle development happens in similar and related pattern in all vertebrates. The foundation for most skeletal muscle of the embryonic body are somites (including trunk,

limb and some hand muscles) [6]. The paraxial mesoderm gives rise to somites (epithelial spheres) by a cranial to caudal migration from the neural tube. The non-somitic paraxial and prechordal head mesoderm are responsible for the remainder of the head muscles. Somites are further sorted into a dorsal epithelial dermamyotome and a ventral mesenchymal sclerotome (giving rise to axial skeleton). It is known that the dermamyotome compartment of the somite gives rise to skeletal muscle by its medially located cells migrating laterally to form the myotome (from E.08 in the mouse) [7]. A group of Basic Helix-Loop-Helix (bHLH) transcription factors known as the Myogenic Regulatory Factors (MRFs) play an essential role in regulating the skeletal muscle development program. The MRFs consist of (Myoblast Determination Protein 1) MYOD, Myogenin, MYF5, and MRF4, which have shown to be functionally redundant [8].

Myogenesis is the process of generating skeletal muscle cells, which primarily occurs during embryonic development but also happens as a result of injury to adult muscle. During embryonic development, myogenesis is initiated in somites by the migration of myogenic precursor cells to the limb buds, forming the early myoblasts (muscle precursor cells). The expression of c-MET (a tyrosine kinase receptor) regulate the migration of these cells and therefore formation of the myoblasts. *c-Met* activity is regulated by the transcription factor, Paired Box 3 (PAX3), which is expressed by the dorsal epithelial dermomyotome (from E10.5 in the mouse) [9]. With the expression of Myogenic Factor 5 (MYF5), theses early muscle cells are promoted to grow and proliferate, influenced by cell cycle associated genes and the downregulation of PAX3. Myoblast proliferation is regulated tightly regulated by MRFs, namely MYOD and MYF5. Myoblast fusion occurs with the expression of Myogenin and Myocyte Enhancer Factor 2 (MEF2), once enough myoblasts have been generated, by

exiting the cell cycle and beginning to fuse. At this stage the myoblasts will cease division and differentiate into myotubes. Initially, primary myotubes are formed by the fusion of embryonic myoblasts. The subsequent stage of differentiation involves the alignment and terminal differentiation of the myoblasts into myocytes and formation of myofibers (**Figure 1.1**).

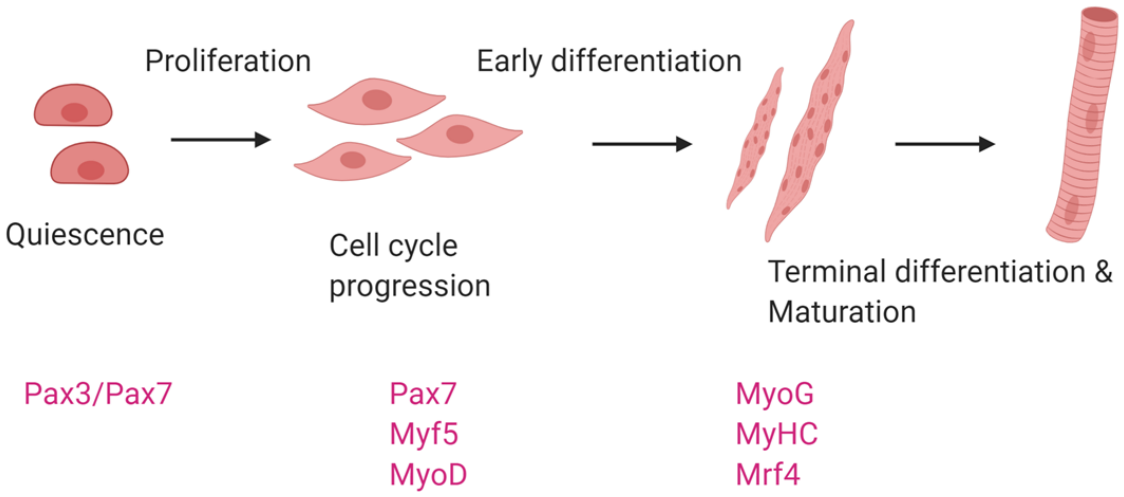


Figure 1.1 Stages of skeletal muscle myogenesis the major muscle-specific transcription factors involved. This image is created with BioRender.

There are various genetic studies which have identified different transcription factors regulating myogenesis, which will be discussed later [9–11]. In embryonic skeletal muscle development, after the muscle has matured, the myogenic progenitor cells enter quiescence and reside there as satellite cells. In adult muscle, myogenesis occurs upon injury to the muscle or in diseased muscle states. In both instances, muscle regeneration is dependent on

the activation of satellite cells, eventually differentiating to new fibers [12]. Most studies on myogenesis are linked to damage to the mature muscle, causing satellite cell activation and subsequently expansion and differentiation to repair the tissue and maintain its homeostasis [13]. Embryonic myogenesis and the myogenesis that occurs in adult muscle (known as regeneration) share many similarities including common signaling molecules and transcription factors regulating the different phases [14].

1.2.2. Genetic regulation of skeletal muscle development

Skeletal muscle development is a complex and multi-step process, which is regulated by many genetic factors. The various factors involve noncoding RNAs, epigenetic modifications, gene polymorphisms and transcription factors, all of which can impact its development at the level of the genome, epigenome, transcriptome and even proteome [15]. For the purpose of this study, the role of transcription factors in the formation of skeletal muscle will be elaborated. One of the most important factors that directs the gene regulatory network guiding muscle progenitor actions and regulate activation of myogenic determination genes is PAX3 [16]. PAX3 is co-expressed with Forkhead Box Protein Factor C2 (FOXC2) in the early somite along with Forkhead Box Protein Factor C1 (FOXC1) in the sclerotome, which is required for bone and cartilage formation [6]. Decrease in PAX3 levels favors non-muscle cell fate (namely vascular smooth muscle or endothelial cells). For a skeletal muscle fate, PAX3 associated myogenic progenitors dominate, whereas FOXC2 levels are reduced [17]. Amongst other roles of the PAX3 transcription factor is a pro-survival role in the somite and it controls the choice between self-renewal of progenitor cells or entering the muscle differentiation program via regulating Fibroblast Growth Factor (FGF)

signaling [18]. PAX3 also activates *c-met*, which is required for migration of myogenic progenitors to the limb. The migration stage also requires the Chemokine Receptor Type 4 (CXCR4) activated by Ladybird Homeobox 1 (LBX1) (downstream of PAX3) [19]. Further entry into the myogenic program is dependent on the myogenic determination factors MYOD, MYF5 (regulated by PAX3), and MRF4, and muscle cell differentiation primarily relies on Myogenin and MEF2 expression [20,21]. PAX3 also indirectly governs gene expression through its direct targets such as Doublesex- and Mab-Related Transcription Factor 2 (DMRT2), Wingless-Related Integration Site (WNT) signaling molecules and Pituitary Homeobox 2 (PITX2), collectively controlling MYF5 expression in the epaxial somite, maintaining somite integrity and acting in part of the hypaxial somite and the limb [22]. Msh Homeobox 1 (MSX1), Homeobox Protein MOX-2 (MEOX2) and Single-Minded Homolog 2 (SIM2) transcription factors play a role in preventing premature onset of myogenesis. Wnt (pro-myogenic), Sonic Hedgehog (Shh) (zone polarizing activity regulation in limb), Bone Morphogenetic Protein (Bmp) (inhibits MYF5 and MYOD expression) signaling pathways have major roles in controlling spatiotemporal activation of the myogenic determination genes [18]. PAX3 also cooperates with the Sine Oculis Homeobox 1/4 (SIX1/4) transcription factors, and their coactivators Eyes Absent Homolog 1/2 (EYA1/2), controlling hypaxial somite and limb expression of MYF5 and activate enhancer elements on the *MyoD* gene [23]. Six factors are known to activate Myogenin unlike PAX3 which inhibits differentiation [18]. Muscle gene transcription is also regulated by MEF2 and Serum Response Factor (SRF) families of MADS-box transcription factors by playing a positive role in skeletal muscle maturation and growth through recruitment of Myocardin [24,25]. Another *Pax3* related gene is *Pax7*, which is also essential in the myogenic program.

When *Pax3* is downregulated during fetal development, *Pax7* is co-expressed centrally in the dermomyotome and becomes the predominant Pax factor in myogenic progenitor cells. PAX7 also predominates the myogenic progenitors that allow adult muscle regeneration in the satellite cells in the perinatal/postnatal period [26]. Notch signaling is also essential in regulation of vertebrate myogenesis. Similar to Bmp signaling molecules, Notch acts by maintaining the pool of PAX3/7^{+} progenitors cells signaling to promote expansion of myogenic precursors, meanwhile inhibiting differentiation [27]. Overall, various signaling pathways and factors govern the onset and progression of myogenesis and formation of skeletal muscles [28].

1.3. Regulation of organogenesis by conserved transcription factor families

Organogenesis is governed by the expression of genes and transcription factors which orchestrate the early patterning of cell groups and ultimately confine organ-specific cell fates [29]. Transcription factors direct downstream signaling pathways by binding to specific DNA sequences to regulate the rate of DNA transcription into RNA. Unique and cell-specific gene expression is generated by the action of transcription factors [29]. The same transcription factors may exert different roles in the various organ system that they are committed to developing. During development of different organs, signaling pathways guide and direct the process in various stages of growth and differentiation in the same organ [30]. Epithelial morphogenesis occurs through signaling pathways allowing to shape developing tissue by cell polarity, bending and folding between the loose mesenchyme and epithelia. An example of this signaling between the two tissues is observed in kidney development WT1 (transcription factor Wilms tumor 1) in the mesenchyme causes Glial Cell Derived

Neurotropic Factor (GDNF) expression leading to epithelial ureteral bud outgrowth [31]. The functional networks of interacting transcription factor families, control and regulate spatiotemporal gene expression during embryogenesis. Another example is the identification of transcription factors required to specify cardiac progenitors in drosophila allowing formation of the primordial heart tube and a functional dorsal vessel, similar to the vertebrate heart [32]. GATA and T-box families of transcription factors have found to be the homologous genes in vertebrates carrying out different function during cardiogenesis [33,34]. Various stages of cardiac development such as cell specification, heart tube looping, chamber formation, differentiation and valvogenesis are shown to be controlled by GATA4, GATA5, GATA6, T-box Transcription Factor 5 (TBX5), and T-box Transcription Factor 20 (TBX20) (to name a few) transcription factors, shown to be implicated in animal models [35–38]. These factors are critical to normal cardiogenesis in humans thus specific mutations lead to various human syndromes and cardiac defects [39–42]. *Pax* genes are also major regulators of organogenesis during embryonic development in limb muscles, kidney, eye, ear, nose, brain and vertebral column, through determining spatiotemporal morphogenesis [43]. T-box transcription factors, specifically TBX4 and TBX5, also coordinate specification and determination of skeletal muscle and limb development through patterning signals in vertebrates [44]. In general, the robust and self-regulatory signaling systems are defined by certain interactions, interlinking multiple pathways, making it challenging to understand the genetic hierarchy amongst these genes. Studying the role of transcription factors in development will be insightful in understanding the complex networks governing normal and disease organ/tissue development.

1.4. The T-box gene family

1.4.1. T-box transcription factors (evolution, function and structure)

The T-box gene family are defined by a conserved domain, named the T-box, extending across a region of 180-190 amino acid residues located anywhere within the polypeptide, which was first discovered in the sequence of the mouse T locus, or Brachyury gene [45,46]. In 1927, the Brachyury locus was discovered by evidence that mutations in the locus lead to embryonic lethality in homozygotes and impaired tail development in heterozygotes [47]. Over the years, the T locus causing defects were studied and it was cloned and found to be a transcription factor. Decades after, a family of T-related genes was discovered in the mouse genome. They were named the T-box gene family and soon after its discovery, seized a lot of scientific attention, and became vastly studied and characterized in many species such as frogs, zebrafish, *C. elegans* and *homo sapiens* [48]. Studies of T-box gene mutations illustrated the importance and roles of this family in human developmental syndromes and their relevance to medical sciences [41,42].

The rise of the T-box genes has occurred over 600 million years ago from a common metazoan ancestor and were from a genome wide duplication during the early evolution of vertebrates [45]. There are 19 T-box family members listed by the HUGO Gene Nomenclature Committee (HGNC) (**Figure 1.2**). The T-box sequence is unique with no similarity found between DNA binding sequences of other transcription factors [49]. During embryogenesis, specifically in cardiac and limb development, essential roles have been implicated for several T-box genes through their specific expression pattern in mammalian tissues.

Overall, they are key to many aspects of organogenesis, embryonic cell fate and patterning, and the development of extraembryonic structures [50].

Structural analysis of the T-box proteins have shown the presence of a transcriptional activator/repressor domain in addition to the T-box domain (17-26 kDa), where their relative positions vary in the different members of the family but the order remains conserved in any one member and its orthologs [51]. All members of the family examined have shown to bind to the DNA consensus sequence TCACACCT, despite the T-box sequence variations among the members [49]. Studies show that the T-box proteins can act as both transcriptional activators and as repressors. The sequences located in the C-terminus of the protein are required for transcriptional regulation activity [52]. Regardless of the critical developmental roles of the T-box genes, so much more is yet be uncovered about their relevant genetic and biochemical pathways. Identifying upstream and downstream factors of the T-box genes is critical in understanding their role in development and function of various tissues/organs [53].

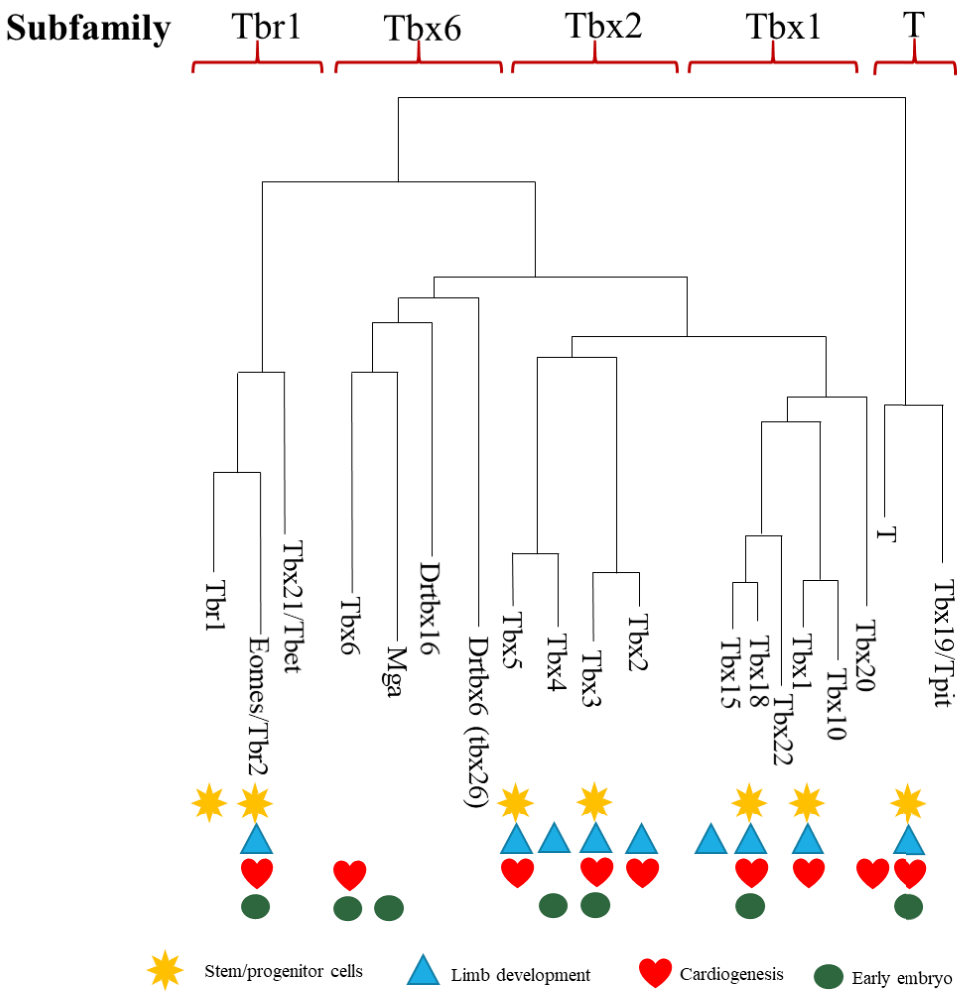


Figure 1.2 Members of the T-box gene family and subfamilies are represented in human and mouse (except Danio rerio (Dr) Tbx16, which is present in birds and frogs (VegT) but not mammals, and Drtbx6). Adapted and modified from Papaioannou (2014) [54].

1.4.2. T-box gene expression in different organs and their associated diseases

Several T-box genes have shown to have roles in mesoderm induction and development, and most of them are expressed in mesoderm or mesoderm precursors across all species around gastrulation time. Examples of T-box gene expression in different organs during embryonic and fetal development, and moreover in derivatives of all 3 embryonic germ layers and in both mesenchymal and epithelial cells, has been observed [55–57]. It is found that the T-box

genes are expressed in organs such as limbs, lungs, heart and kidney in highly tissue specific patterns. Mutational analysis studies will help further define their function and expression in later stages of organ development. Studies have reported the requirement of T-box genes in specific tissue types during later stages of development such as TBX2, TBX3, TBX4 and TBX5 in limb development [58]. There is high similarity and overlap in the expression sites of TBX2-5, but in contrast not much with TBX1, which is a reflection of their evolutionary divergence. The earliest factor to be expressed among TBX2-5 (TBX2 subfamily) is TBX3 in the blastocyst inner cells mass. TBX2 is shown to be expressed in mammary stroma, limb, heart, lung and kidney, regulating the organs cell fate, patterning, morphogenesis, and even coordinating development of the craniofacial structure. Overexpression of TBX2 has been related to various types of cancer such as breast, melanoma, liver, carcinoma, lung, pancreatic and bladder [59,60]. Similarly, TBX3 overexpression has been linked to various cancers [49,50,61]. Moreover, TBX3 is involved in bone development and cell specification and proliferation by acting as a transcriptional repressor [62]. The posterior expression of TBX3 is essential for the development of distal limb elements; its insufficiency in human causes ulnar-mammary syndrome – characterized by absence of the forearm bone (Ulna) and digits [48]. TBX2 and TBX3 expression in the anterior and posterior of forelimb and hindlimb and TBX4 and TBX5 (their close homologs) exclusive expression in the hindlimb and forelimb buds respectively, have been established [55]. TBX4 and TBX5 not only define the limb bud territories but also the type (forelimb *vs.* hindlimb, respectively) by directing the Fgf signaling pathway required for successful limb outgrowth [63]. Retroviral misexpression of TBX5 in the hindlimb or TBX4 in the forelimb mesenchyme of chick embryos has shown transformation of the tissue into forelimb or hindlimb, respectively [64,65]. Moreover,

mutations in TBX5 are known to cause Holt-Oram Syndrome (HOS), a congenital condition affecting both forelimb and heart development [39,41]. Missense mutations associated with the *TBX5* gene affecting its binding to the major groove of the target gene DNA are usually associated with cardiac defects, whereas alteration of residues contacting the minor groove results in the limb abnormalities [66]. Other member of the TBX2 subfamily, namely TBX1-5, TBX18 and TBX20 are also expressed in the vertebrate embryonic heart [67]. TBX1 and TBX18 play roles in myocardialization and cardiac tube elongation by maintaining proliferation of mesenchymal precursors [68]. Early heart tube and myocardial chamber formation programs require the expression and function of TBX5 and TBX20, whereas this program is repressed by TBX2 and TBX3 in favor of the conduction and valvuloseptal system development [67]. *TBX1* mutations have been linked to DiGeorge syndrome, which is characterized by congenital heart disease and outflow tract abnormalities [69,70]. *TBX2* mutations have been associated with cleft palate disorder where the paired palatal shelves fail to make contact to form the midline seam leading to feeding, dental, speech and hearing issues and various complications [71,72]. Studies have shown that *TBX19* is required for the expression of Pro-Opiomelanocortin (POMC) in the pituitary lineages where its absence results in Adrenocorticotropin (ACTH), which causes adrenal insufficiency [73,74]. Overall, T-box genes are known to play a significant role in various genetic human disorders and thus it is imperative to characterize the different factors and their tissue-specific regulatory network and targets.

1.5. TBX5: A key regulator of heart and limb development

The *TBX5* gene is located on the long arm of chromosome 12 and encodes a 518 amino acid protein, which consists of a 180 amino acid T-box domain (situated from amino acid residue 56 to 326) [39,41]. It has two Nuclear Localization Sequences (NLS) located in the T-box and the C-terminus (**Figure 1.3**). While each is sufficient to drive nuclear localization, the two NLS also work cooperatively [75,76]. It also contains three transactivation domains in both the N- and C-terminus of the protein. A Nuclear Export Signal (NES) is located between amino acids 152 - 160, through which subcellular localization of TBX5 can be regulated by the Chromosome Region Maintenance 1 (CRM1) and PDZ and LIM Domain Protein 7 (PDLIM7) export pathways [77,78]. Apart from these domains, TBX5 also contains protein-protein interaction domain, which will be discussed below.

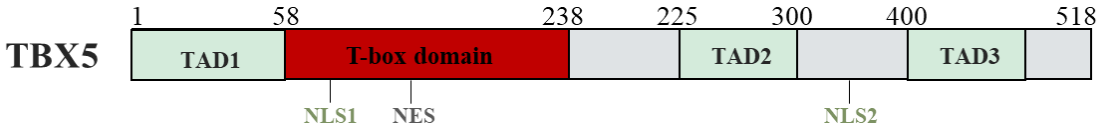


Figure 1.3 Schematic representation of the TBX5 protein structure. Nuclear Localization Sequences (NLS), Nuclear Export Signal (NES), and Trans-activating Domains (TAD) illustrated.

1.5.1. TBX5 expression in adult and embryonic heart

TBX5 is expressed throughout the human embryonic and adult heart in the endocardium, myocardium and epicardium and in the walls and septa of all four chambers. It is also specifically expressed in the embryonic Atrioventricular (AV) node. TBX5 cardiac

expression pattern is very comparable in frog, chick, mouse and human [79,80]. TBX5 expression has shown to be highest in the atrial appendages, after that the lungs, left ventricle, and the esophagus [81,82]. In mice, TBX5 is abundantly expressed as early as E8.0 in the cardiac crescent, which then is restricted to the posterior section leading to heart tube formation. TBX5 expression in the future left ventricles begins at E9.0. Atrial and ventricular TBX5 expression remains throughout embryonic development [55,83]. Genetic mappings have shown the first heart field gives rise to ventricular TBX5 expression, whereas the second heart field sources the atrial and atrial septum TBX5 expression [84]. This suggests that TBX5 might hold different roles in the heart fields during development. TBX5 is involved in these major cardiac developmental stages along with other essential cardiac transcription factors (**Figure 1.4**).

1.5.2. TBX5 expression during limb outgrowth and in the skeletal system

TBX5 is also expressed in many non-cardiac tissues, one of the major well studied expression domains being the developing forelimb. The transcription factor TBX5 is expressed in the lateral plate mesoderm prior to limb bud outgrowth and is crucial for the regulation of the inductive signaling interactions for forelimb bud formation [56,64,85]. In mouse embryonic development, TBX5 expression in the lateral plate mesoderm starts at E8.5 giving rise to the forelimb [55,56]. By E9.0, TBX5 is robustly expressed in the forelimb and its expression is sustained by E11.5 in the developing limb. At this stage, TBX5 expression is apparent in the mesenchyme surrounding the cartilage primordia of digits I and V. At E13.5, there is TBX5 expression evident in the perichondrium of both proximal and distal elements of the forelimb [56]. With forelimb bud initiation, TBX5 is broadly expressed in the forelimb

mesenchyme and is maintained and required for later stages of limb development shown by TBX5 misexpression experiment performed in chick where TBX5 function was knocked down by misexpression of dominant negative forms of TBX5 protein [86]. The genetic TBX5 loss of function results is absence of arms in mice and pectoral fins in zebrafish [85–87]. TBX5 has also shown to be essential in formation of all the appendicular skeleton elements in addition to the limb such as the clavicle and scapula of the pectoral girdle, established by TBX5 conditional knockout experiments in the limb mesenchyme [86]. TBX5 role in forelimb development, however, is regulated by a gene regulatory network and other essential transcription factors which will be discussed later (**Figure 1.4**).

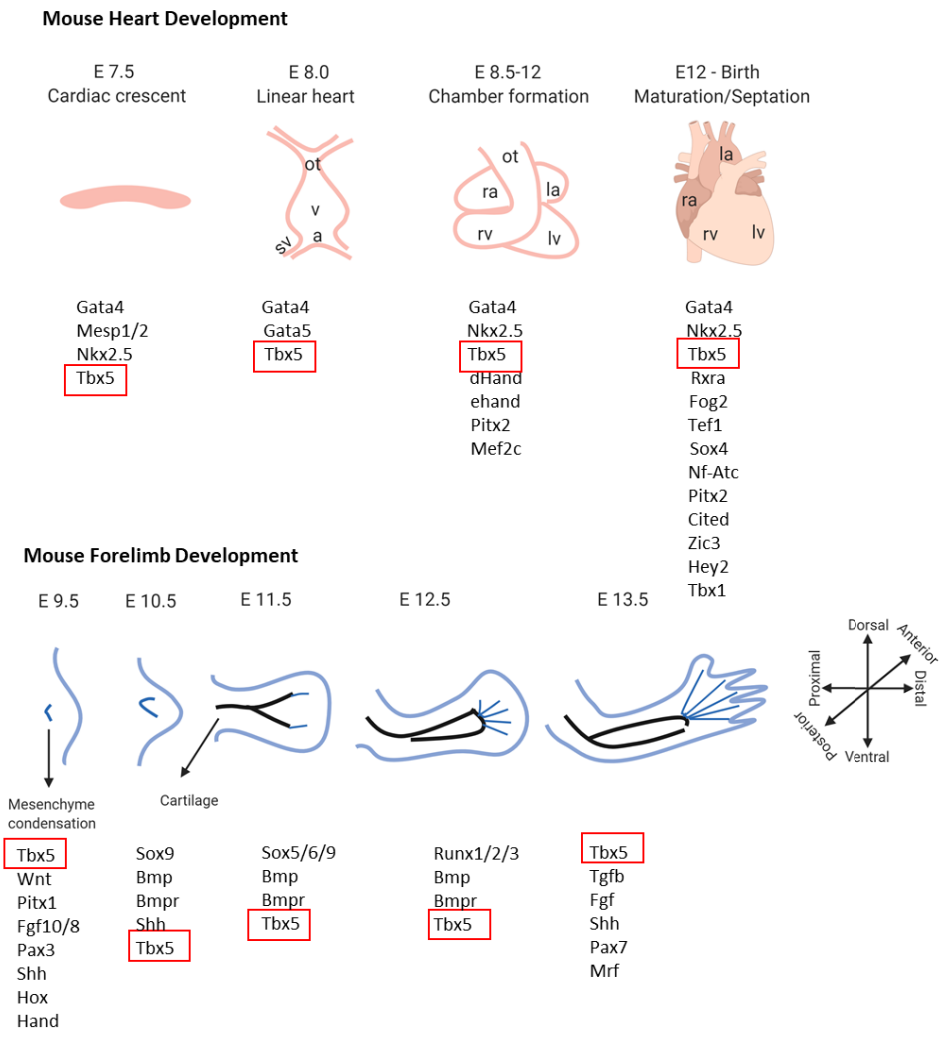


Figure 1.4 TBX5 in curial stages of mouse heart (top panel) and forelimb (bottom panel) development alongside other important transcription factors. Image created using BioRender.

1.5.3. *TBX5* haploinsufficiency: Holt-Oram syndrome and animal model

TBX5 mutations cause the Holt-Oram syndrome (HOS), which is an autosomal dominant disorder, characterized by forelimb musculoskeletal abnormalities and congenital heart defects. First reported in 1960, it is a rare syndrome occurring 1 in 100,000 live births [88]. The clinical cardiac and skeletal symptoms of the disease range from mild to severe with most

common cardiac anomalies including atrial and ventricular septal defects, mitral valve and conduction defects and atrial fibrillation [66,88]. The cardiac defects have mostly been associated with impaired interactions of the T-box domain with GATA4, NKX2.5 and MEF2c leading to major cardiac anomalies [89]. The skeletal defects are fully penetrant, while the cardiac malformations occur in 76% of HOS patients [90]. The clinical manifestations of the limb defects include impaired shoulder girdle and abnormalities in the hands, digits and bones of the lower arm [90,91]. Defect in the thumb and radius are important indicators in diagnosis of the disease [88,90,91]. Pathogenesis in the upper limb are non-classic phenotypes that involves interactions between mesodermal and ectodermal factors and TBX4, whereas the classic radial malformations are caused by disruptions between the SALL4, FGF10-FGF8 loop and TBX5 interactions [63]. HOS is caused by *TBX5* haploinsufficiency and genetic irregularities led by TBX5 coding and regulatory sequences diagnosed in 70% of the patients [92,93]. Over 150 mutations have been reported in the coding, splicing or regulatory sequences of *TBX5* including point mutations, gross deletions, nonsense/missense mutations, gene duplication, single base pair mutation, and mutations resulting in miss or alternative splicing that lead to no protein, truncated protein production or TBX5 decremented transcriptional activity [94]. The differential tissue defect severity in patients indicates a tissue specific role for the distinct domains of TBX5 [95]. This tissue dependent bias observed in the defects is suggested to be in part due to different binding partners and binding motif recognition, although no causing mechanism has been discovered [40,66,77]. On the other hand, in 30% of HOS patients, no mutations in *TBX5* have been found. This phenomenon opens up the possibility of mutations in the cis-regulatory elements or suggests the existence of alternative HOS-causing TBX5 isoforms [96].

The animal models in mouse, chick, frog and zebrafish have helped in investigating the role of TBX5. The mouse model heterozygous for a *Tbx5* knockout allele is the most well characterized and utilized HOS model [93]. The *Tbx5*$^{tm1.1Jse}$ mouse is generated by the germline deletion of exon 3, giving rise to truncated TBX5 transcripts. The heterozygous mice exhibit the HOS phenotype with defective forelimb and cardiac function [97]. The mouse model has been pivotal in studying HOS and TBX5's role in heart and limb development.

1.5.4. *TBX5* gene regulation influencing its expression and function

Many *TBX5* mutations result in altered protein production, in some cases, causing modifications in DNA binding, subcellular localization, transcriptional activation or disruption of functional interactions with other transcription factors. Although the phenotype-genotype correlation in HOS patients is yet unclear, since mutations causing modified cardiac protein interactions do not associate with the limb defects. Understanding TBX5 mechanism of action and regulation in both heart and limb will enhance the current view on the pathogenesis of *TBX5* associated diseases.

- **Alternative splicing**

About 90% of human genes are affected by alternative splicing, which results in different protein products from one gene, named isoforms, which play a role in the increased genome complexity and function [98–100]. Various studies have demonstrated the critical role of alternative splicing in TBX protein activities, effectively affecting expression levels, subcellular localization and functional characteristics. The spliceosome proteins and the pre-mRNA sequence are factors regulating alternative splicing in a tissue specific manner [101].

Three developmentally critical TBX transcription factors, TBX3, TBX5 and TBX20, have been shown to have alternatively spliced isoforms affecting their function [102–104]. Examining isoform specific expression during development will contribute to understanding cell type specific function and ultimately shed light on their complex disease mechanisms in cardiogenesis, cancer and limb development.

In HOS patients, intragenic mutation have been found in addition to mutations in the coding sequence, which could affect RNA splicing and result in production of less protein or altered splice isoforms [105,106]. In 30-35% of HOS sporadic or familial cases, where no *TBX5* mutation in the coding sequence is observed, the unscreened introns and regulatory sequence mutations may justify the low detection rate [107]. The existence of splice variants in the human *TBX5* locus, truncated at amino acid 350 in the C-terminus, has previously been reported [39]. Our lab has reported that a TBX5 isoform, TBX5b, is generated by an alternate exon 8 producing a c-terminal frame shift and thus a shorter protein [104]. An additional exon 7b in splicing results in addition of four amino acids (RPLL) and a stop codon downstream of the T-box. The study shows that the two isoforms TBX5a (full 518 amino acid protein) and TBX5b (251+RPLL) have distinct biochemical properties and their production is growth factor regulated [104]. These finding provide evidence that the generation of functionally distinct isoforms may regulate TBX5 dosage during development and will aid us further in understanding pathogenesis of HOS and other TBX5-associated disorders.

- **Post-translational modifications**

Many cellular processes are mediated by post-translational modifications (PTMs), which are chemical and covalent conjugations of small proteins, including cell proliferation, differentiation, apoptosis and epigenetics in both health and diseased physiological

conditions [108]. It is predicted that TBX proteins have several sites for modification, allowing for methylation, phosphorylation, sumoylation and sulphanation, which may act to regulate their activities or protein partner interactions [109]. There are two conserved sumoylation motifs reported within the T-box regions required for H3K4-methyltransferase and H3K27-demethylase activities [110]. Sumoylation has been reported to enhance the activity of TBX5 [108]. Physical interaction of TBX5 and Protein Inhibitor of Activated STAT1 (PIAS1) leads to full sumoylation of TBX5, enhancing the functional synergy of TBX5 and its protein interactors and transcriptional activity [111]. Analysis of endogenous TBX5 protein in TBX5 expressing cells, has shown the existence of other immunoreactive bands co-migrating with TBX5a and TBX5b isoforms which are suggestive of PTMs [104]. Bioinformatic analysis, has revealed that TBX5 carries numerous phosphorylation sites in the c-terminus of the protein, as well as sumoylation sites which could indicate generation of higher molecular weight isoforms. The study shows that the TBX5 isoforms are present in distinct DNA binding fractions, suggesting that these modifications may affect protein-protein interactions [104]. Moreover, Histone Deacetylase 3 (HDAC3) is shown to modulate TBX5 acetylation, repressing TBX5-directed activation of cardiomyocyte specific genes, hence playing an essential role in cardiogenesis [112]. The identification of TBX5 PTMs and specification of their role in TBX5 expression and function in heart and limb development will expand the current understanding of TBX5 mechanism of action. TBX5 tissue specific PTMs, in particular in limb development have yet to be uncovered.

- **MicroRNAs**

MicroRNAs (miRNA or miR) are 18-25 nucleotide short non-coding RNAs which have been recently discovered to play major roles in developing organs at the transcriptional and

translational levels. Potentially targeting 60-90% of mRNA, miRNA affects normal biological development and even more importantly, are known to be involved in disease pathogenesis [113]. miRNAs bind to the 3′-Untranslated Region (3′-UTR) of target genes, modulating their gene expression at the post-transcriptional level, to regulate developmental signaling pathways [114]. In mouse models, the dysregulation of cardiac factors Heart- and Neural Crest Derivatives-Expressed Protein 2 (HAND2) and Serum response factor (SRF) by haploinsufficiency of miR-1 and miR-133a causes Ventricular Septal Defects (VSD), suggesting maintenance of their expression levels is required in cardiac development [115,116]. Moreover, alteration of TBX5 expression by some miRNAs have been reported; namely, miR-218 in zebrafish is found to mediate TBX5 expression [117]. Another study has demonstrated high expression of miR-10a and miR-10b in Congenital Heart Disease (CHD) tissues by negatively effecting TBX5 expression at the translational level [118]. Also, miR-10a-5p is shown to regulate arthritis joint inflammation in relation with TBX5 which can be implemented in Rheumatoid Arthritis (RA) treatments [119]. There are also miRNAs known to regulate limb development and the signaling molecules such as Shh and the Hox family [120], but miRNAs effecting *TBX5* regulation during limb development are yet to be uncovered.

- **TBX5 protein-protein interactions and direct targets in heart and limb development**

The majority of studies examining TBX5 interactors and targets have been in the heart. TBX5 is primarily known for its role in targeting gene transcription regulation, where it acts as a positive regulator of transcriptional activity. However, it has been recently suggested to also cause transcriptional repression. TBX5 regulation of target gene activity is known to be

modified by its various protein interactors identified to date (**Table 1.1**). In the heart, it is known that TBX5, NKX2.5 and GATA4 cooperate in cardiomyocytes to activate essential genes [53,121]. Although in *Caenorhabditis elegans*, homologues of the T-box and GATA factors regulate different developmental gene profiles in mesoderm and endoderm differentiation which suggests evolution of the interactions in various developmental systems [109]. NKX2.5, a tinman transcription factor, was identified to be the first protein interactor of TBX5, allowing for synergistic activation of Natriuretic Peptide A (*Nppa*) (encoding ANF- Atrial Natriuretic Factor) [97,122]. The synergistic activities resulting from TBX5 and NKX2.5 interactions are regulated at specific cis-regulatory elements and in the absence of either factor causes improper localization and activation of non-cardiac genes by the other present factor [123]. Other major cardiac transcription factor that interact with, TBX5 are GATA4 [40], GATA6 [124], Myocardin [125], TBX20 [126], and MEF2c [127]. *GATA4* mutations diminish its interaction with TBX5, causing various heart defects exhibited in CHDs and HOS [40]. TBX5 interactions with GATA4 and MEF2c are essential for activation of α-Cardiac Myosin Heavy Chain (MHC-α) (encoded by MYH6) [124,127]. Interestingly, TBX5 and GATA6 interactions are not required for this activation, indicating that interactions are tissue and context-specific [124,127]. Furthermore, it has been reported that TBX5 interacts with members of the SWI/SNF family, which are involved in chromatin remodeling, BAF60 and BRG1, to determine the cardiomyocyte fate from mesodermal cells [128]. *Tbx5* haploinsufficient mice have shown loss of chromatin remodeling complexes at the promoters of *Tbx5* cardiac gene targets [129].

Protein	Result of interaction	Reference
NKX2.5	Myocyte diversification and differentiation	Bruneau et al., Cell. 2001.
GATA4/5/6	Cardiac morphogenesis	Garg et al., Nature. 2003. Nadeau et al., Proc Natl Acad Sci U S A. 2010.
TBX20	Chamber specification	Brown et al., Development. 2005.
BAF60c	Transcriptional Synergy	Ghallagher et al., Mol Cell Biol. 2012.
TAZ	Acetyltransferase recruitment	Murakami et al., Proc Natl Acad Sci U S A. 2005.
MEF2C	Cardiomyocyte differentiation	Chosh et al., Mol Cell Biol. 2009.
LMP4	facilitates cytoplasmic localization of the Tbx5 via LMP4	Kulisz and Simon, Mol Cell Biol. 2008. Krause et al., Dev Biol. 2004
SMAD1,5	Mediates Tgfβ signaling	Singh et al., Trends Cardiovasc Med. 2010.
Myocardin	smooth muscle and cardiac gene transcription	Wang et al., PLoS One. 2011.
GLI1	atrioventricular septation	Hoffman et al., PLoS Genet. 2014.
NuRD Complex	Transcriptional repression	Waldron et al., Dev Cell. 2016.

Table 1.1 Some of the known TBX5 interactors in heart cells.

Although TBX5 is primarily known to be a positive regulator of cardiac gene regulatory networks, but recently it has been discovered that it can also act as a direct transcriptional repressor to inhibit improper gene expression. Transcriptional repression has also been documented in TBX2, TBX3 and TBX20 in cardiac development [130–132]. It has been reported that TBX5 interacts with the Nucleosome Remodeling and Deacetylase (NURD) complex, inhibiting non-cardiac gene regulatory programs during embryonic development (similar to the inhibitory role of TBX20) [133]. Disruption of the domain where TBX5 and NURD complex interact can lead to HOS [134]. Many of these protein interactions result in critical gene expression. Direct targets of TBX5 in the heart are the most characterized and functionally confirmed, which are exclusively implicated in cardiac proliferation,

maturation, and function including *Nppa*, Gap Junction Protein Alpha 5 (*Gja5*) (encoding Connexin40) and Sodium Voltage-Gated Channel Alpha Subunit 5 (*Scn5a*) [135–138] (**Table 1.2**). In recent years, ChIP-seq data sets have been generated to identify direct targets of TBX5 binding and regulation [123,139]. However, the direct TBX5 targets mediating morphogenesis are yet to be discovered and will be key to understanding the role of TBX5 in cardiac, limb and other developmental processes.

	Cardiac function	Cardiac expression	Cellular function	Reference
Nppa	Chamber myocardium marker	Myocardial marker	Hormone	(Claycomb, 1988)
Cx40	Gap junction protein involved in cardiac conduction and marker of conduction system	Cardiac muscle cells, vascular smooth muscle, endothelial cells	Involved in gap junctions and in electrical signal transduction	(Pizard et al., 2005)
Cx43	normal activation of ventricular cardiomyocytes mediated by Cx43 gap junction channels in adult mouse heart	Cardiac muscle cells, vascular smooth muscle, endothelial cells	integrating the smooth muscle cell function	(Filipczyk et al., 2007)
Fgf10	essential for cardiac fibroblast development and growth of the myocardium	Cardiac myocytes	mediator of mesenchymal–epithelial signaling during vertebrate organogenesis	(Takeuchi et al., 2003)
Wnt2b	specification of cardiac cell types from the early mesoderm	Cardiac myocytes	controls retinal cell differentiation at the ciliary marginal zone	(Takeuchi et al., 2003)
Bclxl	*Bclxl* appears to be a target of a TBX5/GATA4 transcription complex	Cardiac myocytes	An anti-apoptotic gene (cell survival)	(Nadeau et al., 2010)
eNOS	regulates mobilization and function of endothelial progenitor cells (EPCs), key regulators of vascular repair	Endothelial marker	cell-specific gene regulation	(Nadeau et al., 2010)
Vegf	cardiac function, angiogenesis, and remodeling	Myocardial and endocardial marker	to create new blood vessels during embryonic development	(Nadeau et al., 2010)
Id2	Cardiac Conduction System Development	Expressed in endocardial cushions (EC) and endothelium	a dominant-negative factor for the cell cycle control	(Nadeau et al., 2010) (Moskowitz et al., 2007)
Myh6	gene encodes the alpha heavy chain subunit of cardiac myosin, role in regulating cardiac growth	cardiac myosin heavy-chain (MHC) gene in cardiomyocytes	myocytes influences both cell function and cell viability	(Ghosh et al., 2009)
Srf	regulating genes responsible for maintenance of cardiac function, key role in the modulation of cardiac fibrosis	Cardiomyocytes	participates in cell cycle regulation	(Barron et al., 2005)
Hey2	Ventricular spetation, functions in heart and vessel development	Mouse embryonic ventricular expression	mediator of the Notch signaling	(Kokubo et al., 1999)

Table 1.2 Some of the known gene targets of TBX5a in heart cells.

In limb development, TBX5 is known to interact with SALL4 resulting in morphogenesis and patterning of the anterior limb [140]. More specifically, this interaction is critical in the patterning and growth of the thumb (digit 1) in combination with anterior forelimb morphogenesis through Shh and Bmp signaling (counteracting TBX2 and TBX3) [140]. The study suggests that SALL4 is downstream of TBX5 in the forelimbs, influencing FGF10 expression during its development.

In zebrafish and chick embryos, mutational analysis show that TBX5 regulates *Fgf10* (downstream of Wnt signaling) maintaining TBX5 expression during limb outgrowth and governing forelimb identity. Moreover, TBX5 and WNT2b function together in limb development, initiating and specifying forelimb outgrowth and identity, by targeting the Fgf signaling network [141]. Characterizing the T-box, Wnt and Fgf interactions have been key to understanding not only limb development, but also outgrowth of other tissues and organs. TBX5 is also known to be regulated by interactions with PDZ-LIM domain (cytoskeleton associated) in chicken, modifying TBX5 expression, distribution and binding specificity in developing limb and heart [142]. Nonetheless, TBX5 partners and interactors in skeletal muscle cells are not as well identified as in the heart and thus is one of the main goals of our study.

1.6.Common factors and pathways affected in cardiac and limb diseases

The identification of genes involved in patterning organs and body axes embryogenesis has long been extensively studied. The common molecules contributing to the patterning of distinct organs such as the heart and limb has become evident [143]. Spatiotemporal expression of transcription factors plays a key role in organ patterning at different stages. A common

gene controlling pattern is seen heart and limb morphogenesis, although the formation of the two organs occurs distinctively [143]. Studies have revealed that there is high association and overlap of heart and limb defects in hundreds of Mendelian disorders. In 1 out of 5000 live births, human congenital diseases occur presenting both heart and limb defects [143]. Analysis of fetal autopsies have shown that heart and limb anomalies overlap in 81% of partial aneuploidies and that 57% of limb defects happened to occur with heart defects [144]. Identification of HOS, heart-limb syndrome, has led to the discovery of common genetic pathways affecting both heart and limb development. Studies in HOS cases have shown that Connexin40 (gap junction protein) in both heart and limb is a downstream target of TBX5, contributing to cardiac and skeletal muscle defects [66]. The same trend is seen with another transcription factor, TBX3, which is expressed both during heart and limb development and is known to be associated with heart and limb association when mutated in UMS [145]. Moreover, it has been reported that Insulin Gene Enhancer Protein (ISL1) (a LIM-homeodomain protein) is expressed in both heart and hindlimb progenitors, marking the earlier steps of their generation, suggestive of a common pathway downstream of ISL1 potentially contributing to syndromes associated with both heart and limb [146]. Additionally, Bmp signaling is established to play a role by regulating ISL1 domain in both heart and limb. *TBX3*, is also a downstream target of Bmp signaling through Bone Morphogenetic Protein Receptor 1a (BMPR1a), and is affected in heart-hand syndromes [146]. Some other shared genes involved in both tissue formations are such as *FGF8*, which regulates looping and patterning of the heart as well induction of distal limb patterning [147–149]. Furthermore, Retinoic Acid (RA) is also essential to certain aspects of both limb and heart patterning and formation, influencing limb Apical Ectodermal Ridge (AER) formation and heart looping

patterning [150,151]. TBX5 and dHAND also play dual roles in the development of both organs at a transcriptional level [39,152]. It is also reported that a mouse gene, the Limb-bud-and-heart gene (*Lbh*), expressed in the AER in the ventral limb ectoderm in early stages of limb outgrowth and in the myocardium during cardiogenesis, acts as a transcriptional coactivator in molecular pathways patterning the heart and limb [153].

1.7.Skeletal muscle regeneration

Tissue regenerative capacity is high in lower vertebrates but in higher vertebrates including mammals this ability is diminished [12]. However, in mammals, skeletal muscle is one of the few tissues that retains its ability to rapidly and extensively regenerate and maintain homeostasis in response to severe damage (such as that caused by trauma or inborn genetic defects). To regenerate post-injury, skeletal muscle reactivates the embryonic developmental program to rebuild muscle tissue; thus, regeneration and muscle development share common features [154]. Due to this similarity, regenerative medicine has become an essential learning tool to study developmental biology. Typically, muscle regeneration is characterized by a degenerative phase, where necrosis of muscle fibers occurs (as a result of an inborn or acquired injury to the muscle), and a regenerative phase in which undifferentiated progenitor cells are recruited to the site of injury [12]. Muscle regeneration is activated and maintained by satellite cells which divide asymmetrically to promote the myogenic progeny and for self-maintenance, in a pattern similar to embryonic myogenesis [13]. Proliferating satellite cells and their consequent progeny are named the adult myoblasts or often referred to as the Myogenic Precursor Cells (MPC) [155]. Satellite cells are able to divide and differentiate into

various cell types such as brown fat, bone and muscle, thus can be considered stem cells. The stem cells highly rely on intrinsic programming compared to somatic cells which are highly dependent on extrinsic factors [156]. The regulatory input controlling myogenesis in embryonic development and more so adult muscle repair have been essential in understanding skeletal muscle disease pathogenesis. To this end, many proteome and genome wide association studies have been performed in recent years facilitated by technological advances that have improved our understanding of the underling regulatory network influencing skeletal muscle regeneration [157]. The ultimate goal is to understand the behavior and function of muscle stem cells *in vivo* and *in vitro* so that we can manipulate it in favor of pathogenic muscle regeneration.

1.7.1. Stages of regeneration (quiescence, activation, proliferation, differentiation, fusion, return to quiescence)

Skeletal muscle regeneration is activated by various factors after injury. Before this, the population of skeletal muscle satellite cells are mainly quiescent, maintaining a very low rate of cell division. Rather than being an inactive basal state, quiescence is an important state required for muscle homeostasis that is maintained via direct transcriptional control [158]. However, studying satellite cell quiescence has been difficult since their isolation leads to their activation and that they are a small population of cells. Use of reversibly arrested muscle cells, G0 myoblasts, could potentially help understand the biology of quiescent satellite cells [158]. The paired-box transcriptional factor PAX7 plays a key role in both quiescent and activated satellite cells as a conserved specific marker in satellite cell lineage development [159]. Satellite cells are activated as a result of injury or disease through a

multistep process. Exit from the quiescent state and transition from G0 to G1 stage of cell cycle occurs by the detection of signals received by the quiescent satellite cells. Unlike their quiescent counterparts, these activated cells will now express both PAX3 and PAX7 and hence enter a highly proliferative stage and rapidly expand the progenitor cell pool [160]. At this stage, the undifferentiated progenitor cells myoblasts that are now capable of undergoing myogenic differentiation and fusion. The transition of quiescent satellite cells to an activated proliferative state has been studied extensively *in vivo* and *in vitro*. The proliferative phase allows expansion of the progenitor population along a myogenic lineage and is strongly regulated by the Notch signaling pathway. Decline in Notch and activation of Wnt signaling persuades progression of progenitor cells to the myoblast stage [161]. Proliferation of myoblasts, their differentiation into myotubes and their alignment into myofibers, are stages governed by expression of PAX7, MYOD, and Myogenin, some of the major muscle-specific transcription factors with regards to satellite cell differentiation into fully functional muscle tissue. Upon satellite cell activation, MyoD is upregulated and transitions the quiescent cells (PAX7+MYOD−) to proliferating myoblasts (PAX7+MYOD+). As myoblasts proliferate, some of the PAX7+MYOD+ myoblasts withdraw from cell cycle and return to quiescence by downregulation of *MyoD*. While some other myoblasts, with the start of Myogenin expression, downregulate *Pax7* and prepare to enter the differentiation program (PAX7−MYOD+Myogenin+). Upregulation of MYOD, Myogenin and other MRFs activate the Myosin Heavy Chain (MHC), creatine kinases and other genes essential for terminal myogenic differentiation and formation of mature muscle fibers. After alignment of myocytes, fusing of myoblasts occurs with withdrawal of certain growth factors (discussed below), which is known to be an irreversible process [162]. After mature fiber formation, to

maintain muscle homeostasis, the remaining satellite cells return to a quiescent phase until the need to be activated is met. In normal adult muscle, a sufficient reservoir of satellite cells exists due to the satellite cell number remaining constant throughout injury and regeneration phases [155]. In diseased muscle state, this ability is hindered. Studies in regenerative medicine, have attempted at modulating satellite cell activation and expansion ability, through the use of appropriate biochemical marker expression and facilitation of following regenerative responses. This requires a deep understanding of factors mediating and regulating each stage of regeneration.

1.7.2 Factors influencing proliferation and differentiation of muscle cells

Both during embryonic development and postnatal skeletal muscle regeneration, muscle satellite cells are activated to proliferate and differentiate into mature muscle cells [163]. Two major events of cell proliferation and differentiation occur that are controlled by various muscle specific genes, transcription factors, growth factors and hormones governing myogenesis [164]. The transcription factors discussed above are amongst the most well-established regulators of muscle development, but several other factors have also been implicated in regulating these events. Studies have shown the involvement of members of the Fibroblast Growth Factor (FGF) family in stimulating proliferation of muscle progenitor cells and decreasing their differentiation. Members of the Insulin-like Growth Factor (IGF) also stimulate myoblast proliferation. Other factors such as Hepatocyte Growth Factor (HGF), Tumor Necrosis Factor (TNF-α), Interleukin-6 (IL-6) family of cytokines and the Transforming Growth Factor beta (TGF-β) family collectively aid in maintaining a balance between growth and differentiation of muscle progenitor cells [12,164]. TGF-β is a known

inhibitor of myoblast differentiation and has been extensively studies in various muscle disorders. Many genes have been identified to influence different stages of myogenesis through interacting with the crucial transcription factors, cell cycle associated genes and growth factors involved in skeletal muscle development. Below some factors are highlighted.

1.7.3 Implication of cyclins in proliferation and differentiation of skeletal muscle cells

The Cyclin family have shown to be essential in regulating cell growth and proliferation in collaboration with Cyclin-Dependent Kinases (CDKs). The D-type Cyclins, namely D1, D2 and D3, are family members which have been studied in skeletal muscle development along with their kinase partners, CDK4 and CDK6, that promote cell cycle progression [165]. These are also named the G1 Cyclins based on the proliferative role they have in the G1 stage of cell cycle. Moreover, Cyclin D1 and D2 are characterized as protooncogenes with high expression in human and animal malignancies [166]. Studies have revealed that Cyclin D3 plays an essential role in both cell proliferation and differentiation in skeletal muscle cells [167]. During embryonic skeletal muscle development or in adult muscle regeneration, high Cyclin D3 along with Cyclin D1 have been reported in the G1/S so called the proliferating zone where the progenitor cell population is amplifying [165]. Moreover, Cyclin D3 is also thought to be involved in induction of terminal differentiation of myotubes but not in their maintenance, indicating a dual role for this cyclin in myogenesis [167]. Moreover, expression of Cyclin D1 and the Cyclin-CDK complex has shown to inhibit *MyoD* mediated activation of muscle-specific genes in proliferating myoblasts maintaining a proliferative state and inhibiting myogenic differentiation [168]. Irreversible interruption of the cell cycle allows the

myogenic cell to commit for skeletal muscle differentiation. Dysregulation of the cell cycle antagonizes the function of myogenic factors such as MyoD and Myogenin, required for differentiation, causing uncontrolled proliferation and lack of differentiation-specific gene expression in dividing cells [169]. Cyclin D2 has also been identified a key regulator of skeletal muscle myoblast proliferation and differentiation serving as a great target for muscle disorder treatments that are shown to have impaired regeneration and loss of muscle mass [170]. CDK Inhibitors (CDKIs) have been utilized in disrupting the cell cycle and promoting muscle differentiation. Specially the group of CDKIs which bind Cdk4 and Cdk6 preventing association with the D-type cyclins [171]. CDKIs contribute to maintaining cell cycle dormancy with high expression levels in non-proliferating cells [172]. The expression of other Cyclins has also been observed during myogenesis such as Cyclin A (G2 Cyclin) which is required for the S phase and mitosis during cell cycle and is usually downregulated in differentiated cells [173]. As well, Cyclin B and Cyclin E also are involved and upregulated in later stages during differentiation [173]. Understanding the interactions of the signaling pathways regulating the determining stages of fetal and adult myogenesis is crucial in taking therapeutic approaches in muscle degeneration and regeneration in various muscle disorders such as muscular dystrophies, muscle cancers and other types of diseases which lead to muscle wasting.

1.7.4 Role of growth factors in regulation of myogenesis

During both embryonic and adult muscle development, regulation of muscle specific genes is largely dependent on growth factor-activated intracellular signaling pathways [12]. In post-natal skeletal muscle cells, growth factor production happens through activated immune cells after upon injury, by the vasculature, through activated stem cells and by motor neurons [174].

Some growth factor production occurs through the Extracellular Matrix (ECM) directed by Matrix Metalloproteinases 2 and 9 (MMP2 and 9) as a result of satellite cell activation after injury or during embryonic development [175]. Various growth factors play major role in myogenic proliferation and differentiation such as hepatocyte growth factor (HGF), Fibroblast Growth Factor (FGF), Insulin growth factor (IGF), Vascular Endothelial Growth Factor (VEGF), Stromal Derived Factor-1 (SDF-1), and Platelet-Derived Growth Factor-AA and -BB (PDGF-AA and -BB) [161]. IGF-1 and IGF-2 promote proliferation and differentiation of myoblasts derived from stem cells, and direct injection of IGF-1 in mice improves regeneration [176]. HGF and VEGF both can also improve muscle healing by inducing stem cell proliferation and stimulating angiogenesis, but HGF has an inhibitory effect on myotube formation and muscle differentiation [177]. The muscle progenitor cells are also regulated through collaborative action of PDGF and VEGF [161]. Moreover, SDF1 acts as a chemoattractant to enhance skeletal muscle proliferation [178]. FGF2 and FGF6 along with HGF synergistically play roles in promoting satellite cell proliferation [179]. In addition to growth stimulatory factors, there are also major inhibitory factors involved in skeletal muscle regeneration and development. The TGF-β superfamily are the principle regulatory factors with inhibitory effects on both muscle development and postnatal regeneration of skeletal muscle. Myostatin (MSTN), Transforming Growth Factor-α and -β1 (TGF-α and -β1), and Bmps are members of this family known to negatively regulate skeletal muscle development. Myostatin and TGF-β1 regulate satellite cell and myoblast proliferation and differentiation by downregulating *MyoD* expression through binding the Smad protein (signal transducers for receptors of the transforming growth factor beta), SMAD3, and inhibiting myoblast differentiation [180]. Lack of myostatin results in muscle hypertrophy due

to dysregulation of proliferation and differentiation in muscle development [181]. TGF-β1 also stimulates fibroblasts to induce remodeling of the ECM and the formation scar tissue and muscle fibrosis (this act is inhibited by Decorin improving regeneration) [182]. BMP also negatively regulates cell proliferation, which is counteracted by an inhibitor of Bmp signaling, Noggin [183]. Extensive studies are required to identify all the partners and their function and networks and downstream pathways in skeletal muscle development and regeneration.

1.7.5 Implication of non-muscle restricted transcription factors in skeletal muscle myogenesis

Muscle-specific genes such as *Pax3, Pax7, Myf5, MyoD, Mrf4,* Myogenin have critical roles throughout essential stages of myogenesis in both fetal and adult skeletal muscle development. As well, various non-muscle restricted genes have been reported to regulate and modify skeletal muscle gene expression throughout the major stages of myogenesis [24]. The Serum Response Factor (SRF) family, through interaction with the Myocardin family, regulates muscle gene expression, where the lack of SRF expression leads to smaller multinucleated muscle fibers, hinting their importance in muscle maturation [25]. The Myocyte Enhancer Factor 2 (MEF2) isoforms, MEF2C and MEF2D, also regulate muscle gene expression by associating with HDACs, promoting their sumoylation and cooperating with myogenic bHLH [184]. Moreover, members of the Wnt family have been shown to regulate the expression of muscle specific genes (*Pax3, Myf5* and *MyoD*) through association with cAMP Response Element-Binding protein (CREB), required for activation of the myogenic program [185]. Another factor involved in skeletal muscle myogenesis is the Retinoblastoma (Rb) tumor suppressor protein by promoting skeletal muscle differentiation. Rb is

specifically required for differentiation of myoblasts into myotubes and its continuous expression in myotubes is required for optimum muscle transcription [186].

TBX5 has also been shown to be involved in proliferation and differentiation of skeletal muscle cells. Our group has previously demonstrated that C2C12 myoblasts express higher levels of the TBX5a isoform compared to TBX5b. The distribution of TBX5a and TBX5b isoforms in C2C12 myoblasts versus myotubes suggests that TBX5a is the predominant isoform in proliferating myoblasts, whereas TBX5b is more highly expressed in the differentiated myotubes [104]. The study showed that TBX5b misexpression in proliferating C2C12 myoblasts leads to growth arrest and apoptosis. Moreover, TBX5 has been reported to play a promoting role in limb bud outgrowth thought *Fgf10* expression and has also been implicated in forelimb regeneration [85,86,187,188]. TBX5 in the heart has also shown to have proliferative role where TBX5 depletion in *Xenopus laevis* interferes with cell cycle progression in cardiomyocytes leading to decreased cardiac cell numbers [136]. Another study in zebrafish shows that *Tbx5* deficiency caused apoptosis and dysmorphogenesis in both heart and limb development (by lacking markers of early differentiation in both not developing fin buds and failing in heart looping resulting in cardiac defects) [189]. With accumulating evidence that TBX5a has a growth promoting effect in not only heart but also in the skeletal system, it is important to identify its involvement in skeletal muscle tissue and its specific partners and targets therein.

Another non-muscle specific family of transcription factors that are known to interact with TBX5 in heart development are the GATA family; which are also speculated to have roles in limb and skeletal muscle development. GATA4 and GATA6, members of the family, have been found to be expressed in an anterior to posterior gradient in the early limb bud. GATA6

has been mainly associated with hindlimb expression and loss of *GATA6* has shown to result in polydactyly (presence of additional digits) in the hindlimbs and while its excessive expression represses Shh signaling giving rise to hypomorphic limbs [190]. Conversely, loss of *Gata6* did not seem to cause abnormal digits in the forelimb suggesting factors such as GATA4 might play a role with GATA6 to block polydactyly in the forelimb [190]. *GATA6* is reported to be expressed in Rhabdomyosarcoma (malignant skeletal muscle cancer) cell lines where its downregulation has shown to reduce cellular proliferation suggesting it might potentially be valuable in therapeutic drug development [191]. Moreover, the diaphragm is also a part of the skeletal muscle system, and defects in its development lead to Congenital Diaphragmatic Hernias (CDH). A study has shown that the heterozygous deletion of *Gata4* in mice leads to CDH development in 14% and diaphragm defects in 29% of the mice [192]. Another cardiac essential transcription factor is NKX2.5 which has also been detected in cranial skeletal muscles [193]. Studies have shown that increased NKX2.5 expression in C2C12 myoblasts inhibits the expression of Myogenin and key players of myogenic differentiation that are essential for differentiation and formation of myotubes [194]. Furthermore, microRNAs such as miR-1 are also able to regulate myogenic differentiation by targeting *Hdac4* (a transcriptional repressor of muscle gene expression), whereas, miR-133 by repressing SRF enhances myoblast proliferation [195]. NKX2.5 is reported to effect *miR-17-92* cluster expression modulating the myogenic program [196].

1.8 Skeletal muscle disorders and animal models: Dystrophies, skeletal muscle cancers, sarcopenia and ageing

Skeletal muscle is a complex and highly specialized tissue which is required to maintain suitable contractility and mobility function by enduring mechanical and physiological stress [197]. Muscle disease or physical trauma leads to progressive weakening of the muscle and disabilities seen in over hundred different human disorders. There are various causes of muscle injury which can influence muscle structure (loss of fiber integrity) and ultimately function. External traumas affecting skeletal muscle include, exposure of muscle to extreme temperatures, ischemia, punctures and sharp trauma, diseases such as muscle dystrophies, muscle contraction and introduction of myotoxic agents [198]. Some of the inherited muscle degenerative diseases include: Myotonic Dystrophy (DM), Limb Girdle Muscular Dystrophy (LGMD), Emery–Dreifuss Muscular Dystrophy (EDMD), Congenital Muscular Dystrophy (CMD), Facioscapulohumeral Muscular Dystrophy (FSHD), Duchenne Muscular Dystrophy (DMD) and Becker Muscular Dystrophy (BMD). All these conditions have different genetic causative lesions leading to impaired gene expression regulation, deficient repair potential and modified stability in muscle fibers [199]. Most of these disorders are accompanied by progressive loss of skeletal muscle mass and strength referred to as Muscle wasting or atrophy [199]. Another muscle condition which results in muscle wasting is Sarcopenia, which occurs with advancing age and is a worldwide health concern; affecting 25% of those older than 70 and 40% of those older than 80 [200]. The molecular mechanism of muscle wasting has thought to be associated with chronic inflammation in the muscle tissue, metabolic changes, oxidative stress and mitochondrial impairments, motor neuron loss and satellite cell deficits, all of which affect the normal muscle regenerative response [200]. Furthermore, muscle wasting

due to cancer, Cachexia, is a muscle wasting syndrome with a multifactorial and complex pathogenicity characterized by weight loss, anemia and anorexia and contributes to one third of cancer deaths [201]. Studies have shown that muscle wasting is mediated by tumor factor TNF-α and cytokine Interferon-gamma (IFN-γ) which suppress muscle gene products such as myosin expression in myotubes and ultimately muscle loss [201]. Likewise, skeletal muscle can also be directly impacted by cancer. Rhabdomyosarcoma (RMS) is a relatively solid tumor derived from mesenchymal cells from the skeletal muscle lineage [202]. It's development has been attributed to the molecular impairment and dysregulation of the p53 and Rb tumor suppressor pathway along with upregulation of protooncogenes and activation of the Ras pathway [203]. The current therapeutic include surgery, chemotherapy and radiation therapy, although there yet remain challenges in tacking the disease. As discussed above, to understand muscle structure and function, studies have aimed to investigate the cellular and molecular responses to muscle damage and the consequent regeneration by using various animal models.

Currently, in most degenerative muscle diseases, the primary therapeutic approaches are to treat the disease symptoms; while disease prevention approaches are extensively being studied [197]. An improved understanding of the regulatory gene network that is involved in muscle development and regeneration will help identify potential pharmacological targets to develop newer and more efficient therapies.

1.8.1 Duchenne Muscular Dystrophy and the MDX mouse model

Among the most significant diseases and injuries targeting skeletal muscle, Duchenne muscular dystrophy (DMD) is the most common X-linked genetic human disease. It is

mutations in the Dystrophin gene (*DMD*) (required for cytoskeletal stability) that result in reduced or nonexistent protein expression [204]. Becker Muscular Dystrophy (BMD) patients show a milder phenotype than DMD because the *DMD* gene mutation still produced a partially functional dystrophin protein [205]. DMD is characterized by progressive weakening of the muscle beginning in early years of life in addition to cardiovascular and respiratory complications that usually lead to premature death. In dystrophic patients, muscle fibers are susceptible to contraction induced injury whereas mechanical stress activates proteases causing membrane destruction and finally necrosis of the myofiber [197]. As a result of muscle damage, efforts to compensate are made with recurring attempts at regeneration which would normally be supported by the satellite cells. With progression of the disease, the mitotic potential of satellite cells is exhausted, decreasing the satellite cell pool, and reducing their telomere length that collectively reduce the regenerative capacity of muscle [206]. As such, the muscle tissue continues to weaken and due to the absent muscle regenerative response, muscle fibers become replaced with fat and fibrotic tissue [207]. Moreover, damage in dystrophic muscle is also induced by mis-localization and reduced expression of Neuronal Nitric Oxide Synthase (nNOS) in the sarcolemma, also due to loss of functional Dystrophin, impairing blood supply and causing muscle ischemia [208].

Given the severity of the DMD phenotype, its prevalence (affects 250,000 individuals in the United States) and the common consequential muscle wasting complication it shares with other muscle degenerative diseases; there has been a considerable effort by the research community to better understand the pathology and molecular mechanisms underlying this disease [209]. To do so, there are various mouse models generated to understand the pathogenesis of muscular dystrophies. The Dmdmdx (MDX) mouse model mouse is the

most common animal model which recapitulates DMD and most widely used to study the disease. The mouse line was identified spontaneously during a red blood cell defect screening in a C57BL/10 colony with high serum creatine kinase levels and Dystrophin deficiency [210]. The MDX mice recapitulate the human disease by displaying active myofiber necrosis and cellular infiltration in their skeletal muscle cells and as a result, they have various sizes of regenerating myofibers. The diaphragm of the MDX mice is the most affected skeletal muscle with progressive degeneration and muscle loss occurring resulting in weakening [211]. They do also experience forelimb and hindlimb muscle degeneration, with lowered grip and strength, respectively [212]. However, the pathogenic progression is not exactly the same as human DMD despite the common absence of cardiac and skeletal dystrophin [207]. The MDX mice exhibit hindlimb necrosis and regeneration peaks that plateau by the age of 3 to 4 weeks. By 2 years of age (if they survive), severe muscle loss and weakness and fat and fibrosis increase appears [207]. Cardiac defects are not observed in young adult MDX mice, unlike the human patients. However, some immune infiltrations in the heart suggest that matrix remodeling occurs this age. Phenotypically, in both mice and humans, the MDX muscle tissue seem to be in a proliferative stage early on, due to activation of the satellite cells as a result of muscle degeneration, but the myoblasts are less able to differentiate and form mature myofibers [207]. Furthermore, as discussed above, with ongoing degeneration in the muscle cells and with increasing age the proliferative capacity of the satellite cells is reduced exhausting their pool and reducing generation of new myofibers to replace the degenerated fibers and as a result muscle wasting and weakening happens [210]. As mentioned, there are many mouse models generated (mdx/utr−/−, mdx/dtna−/−, mdx/myod1−/−, mdx/α7−/− and other knock out models utilized for better

humanizing the disease) recapitulating different aspects of DMD resulting in a different range of phenotypes [207]. The choice of animal model utilized in studies is dependent on the goals and perspectives of the specific research question.

1.8.2 Cardiotoxin injury model in mice

Alternatively, the process of skeletal muscle regeneration can be studied using animal models of muscle injury. Myotoxins have long been used as muscle injury inducing agents and include Bupivacaine (Marcaine), Cardiotoxin (CTX), and Notexin (NTX). Each of the Myotoxins have different biological activities [12,213]. Marcaine is known to produce muscle injury by effecting intracellular calcium levels of muscle. NTX is a neurotoxin that blocks neuromuscular transmission and is extracted from snake venoms [214]. CTX is a kinase C specific inhibitor, from *Naja nigricollis* snake venom, causing depolarization and contraction of muscular cells disrupting and lysing membrane organization and cells. The extent of muscle injury and the animal model determine the kinetics and amplitude of each regenerative phase among different muscle types. Typically for studies in animal models, 25ul of 10mM CTX is injected in the hindlimb (specifically the Tibialis Anterior (TA)) muscle to induce muscle degeneration with cell infiltration within 1 day. From day 1-4 after injection, the inflammatory response peaks and mononuclear cell proliferation occurs. From day 5-7 post injection, the proliferated myoblasts start to differentiate and formation of myotubes occur. The overall muscle structure is restored by 10 days after injection, where myofibers are smaller with the nuclei centrally located. By 3-4 weeks post injection, the mature muscle formation is seen with normal muscle morphology [12,198] (**Figure 1.5**). Even though CTX injury is a highly reproducible method of inducing muscle regeneration, its

toxic effects on various muscle cells such as satellite cells remains unknown and is a topic under investigation. An alternative model to study muscle regeneration is wound inflection by crushing or freezing of the muscle or transplantation of a single muscle causing the muscle denervation and devascularization [215]. As previously mentioned, the MDX mice also are used to study muscle repair and regeneration in addition to individually studying proliferation and differentiation of skeletal muscle cells. However, in the MDX mice the mouse strain influences the process of regeneration making it harder to reproduce results [216]. Moreover, they provide the platform to study all the genes and regulatory pathways involved in muscle degeneration and regeneration processes.

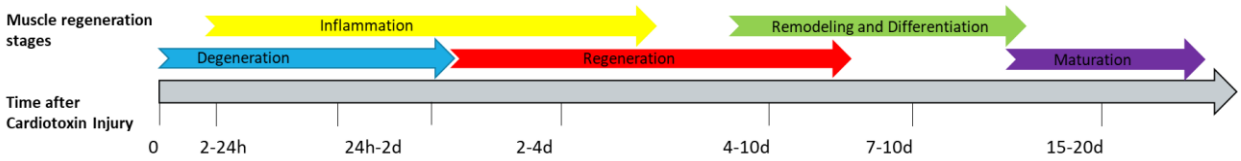


Figure 1.5 Stages of skeletal muscle regeneration after cardiotoxin injury. Representing a timeline of up to 20 days post Cardiotoxin injury.

1.8.3 Current limitations in muscle disorder treatments

Currently research and studies have led to development of treatments controlling DMD disease symptoms, manifestations and progression. however, within the recent years promising experimental strategies are emerging such as the use of gene therapy via various

viral vectors such as adeno-associated viral, lentiviral and adenoviral vectors [217]. But also, non-viral methods such as plasmid DNA gene therapy, cell-based therapies and, exon skipping have shown promise for the regeneration potential of dystrophic muscle [217]. These advances have been made possible by the various animal models for muscular dystrophies and regeneration. The viral vector mediated gene therapies have shown to be amongst the most promising of approaches by specific tissue targeting, high and prolonged efficiency [218]. Also, transgene promoter and enhancer modulation are utilized to enhance gene transfer approaches by improving muscle specific expression [219]. Optimization of gene delivery and immune suppression are dilemmas that are heavily studied and provide a lot of hope for enhancing current clinical interventions [218]. However, a very good understanding of the modulators, modifiers, partners and targets of the gene regulatory network involved in skeletal muscle development is required to fully understand muscle development, regeneration and disease. This will aid in generating treatments and combination therapies which can have specific targets for each stage of muscle development and regeneration to have better control on different types of skeletal muscle pathogenesis.

1.9 Objectives and Hypothesis

Given the information discussed above, mutations in *TBX5* lead to a large spectrum of complex heart and limb malformations and thus it is important to understand the mechanism of pathogenesis and the molecular basis for the variable disease expressivity. The objective of this project is to identify novel TBX5 protein partners and downstream gene targets involved skeletal muscle development. I hypothesize that TBX5a plays a role in activation of pro-proliferative signals in C2C12 mouse myoblast cells and therefore enables myoblast proliferation and inhibits myotube differentiation and that the cell specific roles of TBX5 are caused by yet undiscovered protein-protein interactions which regulate distinct target genes and pathways in both cardiac and skeletal muscle development. Also, in addition to HOS, *TBX5* is associated with other skeletal muscle disorders such as muscular dystrophies and the skeletal muscle regenerative process. The overall objective of this study is to characterize TBX5 and understand the mechanism of action of TBX5 in development (proliferation and differentiation) of skeletal muscle cells and its role in disease pathogenesis. The objective of this study has been presented as follows:

1. Identification of novel exons in the *Tbx5* gene locus generate protein isoforms with distinct expression domains and function.
2. TBX5a role in inhibiting differentiation of skeletal muscle cells by interacting with MYBBP1a and regulating Myostatin.
3. TBX5a role in myoblast proliferation and skeletal muscle regenerative capacity.

1.10 References

1. Dave, H. D. & Varacallo, M. Anatomy, Skeletal Muscle. in *StatPearls* (StatPearls Publishing, 2020).
2. Frontera, W. R. & Ochala, J. Skeletal Muscle: A Brief Review of Structure and Function. *Calcif Tissue Int* **96**, 183–195 (2015).
3. Mcnally, E. M., Lapidos, K. A. & Wheeler, M. T. Skeletal Muscle Structure and Function. in *Principles of Molecular Medicine* (eds. Runge, M. S. & Patterson, C.) 674–681 (Humana Press, 2006). doi:10.1007/978-1-59259-963-9_67.
4. Rehfeldt, C., Fiedler, I., Dietl, G. & Ender, K. Myogenesis and postnatal skeletal muscle cell growth as influenced by selection. *Livestock Production Science* **66**, 177–188 (2000).
5. Peters, S. E. Structure and Function in Vertebrate Skeletal Muscle. *Am Zool* **29**, 221–234 (1989).
6. Lagha, M. *et al.* Pax3:Foxc2 reciprocal repression in the somite modulates muscular versus vascular cell fate choice in multipotent progenitors. *Dev. Cell* **17**, 892–899 (2009).
7. Kablar, B. & Rudnicki, M. A. Skeletal muscle development in the mouse embryo. *Histol. Histopathol.* **15**, 649–656 (2000).
8. Megeney, L. A. & Rudnicki, M. A. Determination versus differentiation and the MyoD family of transcription factors. *Biochem. Cell Biol.* **73**, 723–732 (1995).
9. Chal, J. & Pourquié, O. Making muscle: skeletal myogenesis *in vivo* and *in vitro*. *Development* **144**, 2104–2122 (2017).
10. Buckingham, M. Myogenic progenitor cells and skeletal myogenesis in vertebrates. *Current Opinion in Genetics & Development* **16**, 525–532 (2006).
11. Wallace, M. A., Hughes, D. C. & Baar, K. Chapter 3 - mTORC1 in the Control of Myogenesis and Adult Skeletal Muscle Mass. in *Molecules to Medicine with mTOR* (ed. Maiese, K.) 37–56 (Academic Press, 2016). doi:10.1016/B978-0-12-802733-2.00025-6.
12. Chargé, S. B. P. & Rudnicki, M. A. Cellular and Molecular Regulation of Muscle Regeneration. *Physiological Reviews* **84**, 209–238 (2004).
13. Bentzinger, C. F., Wang, Y. X. & Rudnicki, M. A. Building Muscle: Molecular Regulation of Myogenesis. *Cold Spring Harb Perspect Biol* **4**, (2012).
14. Tajbakhsh, S. Skeletal muscle stem cells in developmental versus regenerative myogenesis. *J. Intern. Med.* **266**, 372–389 (2009).
15. The genetic regulation of skeletal muscle development: insights from chicken studies. http://journal.hep.com.cn/fase/article/2017/2095-7505/19904.
16. Buckingham, M. & Relaix, F. PAX3 and PAX7 as upstream regulators of myogenesis. *Semin. Cell Dev. Biol.* **44**, 115–125 (2015).
17. Mayeuf-Louchart, A. *et al.* Notch regulation of myogenic versus endothelial fates of cells that migrate from the somite to the limb. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 8844–8849 (2014).
18. Buckingham, M. Gene regulatory networks and cell lineages that underlie the formation of skeletal muscle. *PNAS* **114**, 5830–5837 (2017).
19. Daubas, P. *et al.* Fine-tuning the onset of myogenesis by homeobox proteins that interact with the Myf5 limb enhancer. *Biol Open* **4**, 1614–1624 (2015).
20. Rudnicki, M. A. & Jaenisch, R. The MyoD family of transcription factors and skeletal myogenesis. *BioEssays* **17**, 203–209 (1995).
21. Wang, Y., Schnegelsberg, P. N. J., Dausman, J. & Jaenisch, R. Functional redundancy of the muscle-specific transcription factors Myf5 and myogenin. *Nature* **379**, 823–825 (1996).
22. Buckingham, M. & Vincent, S. D. Distinct and dynamic myogenic populations in the vertebrate embryo. *Curr. Opin. Genet. Dev.* **19**, 444–453 (2009).

23. Buckingham, M. & Rigby, P. W. J. Gene regulatory networks and transcriptional mechanisms that control myogenesis. *Dev. Cell* **28**, 225–238 (2014).
24. Sartorelli, V. & Caretti, G. Mechanisms underlying the transcriptional regulation of skeletal myogenesis. *Current Opinion in Genetics & Development* **15**, 528–535 (2005).
25. Li, S. *et al.* Requirement for serum response factor for skeletal muscle growth and maturation revealed by tissue-specific gene deletion in mice. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 1082–1087 (2005).
26. Relaix, F., Rocancourt, D., Mansouri, A. & Buckingham, M. A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. *Nature* **435**, 948–953 (2005).
27. Vasyutina, E. *et al.* RBP-J (Rbpsuh) is essential to maintain muscle progenitor cells and to generate satellite cells. *Proceedings of the National Academy of Sciences* **104**, 4443–4448 (2007).
28. Hernández-Hernández, J. M., García-González, E. G., Brun, C. E. & Rudnicki, M. A. The myogenic regulatory factors, determinants of muscle development, cell identity and regeneration. *Seminars in Cell & Developmental Biology* **72**, 10–18 (2017).
29. Andrew, D. J. & Yelon, D. Editorial overview: Developmental mechanisms, patterning and organogenesis. *Curr Opin Genet Dev* **32**, v–viii (2015).
30. Kato, S. & Sekine, K. FGF-FGFR signaling in vertebrate organogenesis. *Cell Mol Biol (Noisy-le-grand)* **45**, 631–638 (1999).
31. Hammerman, M. R. Organogenesis of Kidney and Endocrine Pancreas. *Organogenesis* **3**, 59–66 (2007).
32. Tao, Y. & Schulz, R. A. Heart development in Drosophila. *Semin. Cell Dev. Biol.* **18**, 3–15 (2007).
33. Olson, E. N. Gene regulatory networks in the evolution and development of the heart. *Science* **313**, 1922–1927 (2006).
34. Rikin, A. & Evans, T. The tbx/bHLH transcription factor mga regulates gata4 and organogenesis. *Developmental Dynamics* **239**, 535–547 (2010).
35. Plageman, T. F. & Yutzey, K. E. Differential Expression and Function of Tbx5 and Tbx20 in Cardiac Development. *J. Biol. Chem.* **279**, 19026–19034 (2004).
36. Singh, M. K. *et al.* Tbx20 is essential for cardiac chamber differentiation and repression of Tbx2. *Development* **132**, 2697–2707 (2005).
37. Holtzinger, A. & Evans, T. Gata4 regulates the formation of multiple organs. *Development* **132**, 4005–4014 (2005).
38. Holtzinger, A. & Evans, T. Gata5 and Gata6 are functionally redundant in zebrafish for specification of cardiomyocytes. *Dev. Biol.* **312**, 613–622 (2007).
39. Basson, C. T. *et al.* Mutations in human cause limb and cardiac malformation in Holt-Oram syndrome. *Nat Genet* **15**, 30–35 (1997).
40. Garg, V. *et al.* GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. *Nature* **424**, 443–447 (2003).
41. Li, Q. Y. *et al.* Holt-Oram syndrome is caused by mutations in TBX5, a member of the Brachyury (T) gene family. *Nat. Genet.* **15**, 21–29 (1997).
42. Kirk, E. P. *et al.* Mutations in cardiac T-box factor gene TBX20 are associated with diverse cardiac pathologies, including defects of septation and valvulogenesis and cardiomyopathy. *Am. J. Hum. Genet.* **81**, 280–291 (2007).
43. Dahl, E., Koseki, H. & Balling, R. Pax genes and organogenesis. *BioEssays* **19**, 755–765 (1997).
44. Zeller, R., López-Ríos, J. & Zuniga, A. Vertebrate limb bud development: moving towards integrative analysis of organogenesis. *Nature Reviews Genetics* **10**, 845–858 (2009).
45. Agulnik, S. I. *et al.* Evolution of mouse T-box genes by tandem duplication and cluster dispersion. *Genetics* **144**, 249–254 (1996).

46. Herrmann, B. G., Labeit, S., Poustka, A., King, T. R. & Lehrach, H. Cloning of the T gene required in mesoderm formation in the mouse. *Nature* **343**, 617–622 (1990).
47. Papaioannou, V. E. & Silver, L. M. The T-box gene family. *BioEssays* **20**, 9–19 (1998).
48. Bamshad, M. *et al.* Mutations in human TBX3 alter limb, apocrine and genital development in ulnar-mammary syndrome. *Nature Genetics* **16**, 311–315 (1997).
49. Wilson, V. & Conlon, F. L. The T-box family. *Genome Biol.* **3**, REVIEWS3008 (2002).
50. Naiche, L. A., Harrelson, Z., Kelly, R. G. & Papaioannou, V. E. T-box genes in vertebrate development. *Annu. Rev. Genet.* **39**, 219–239 (2005).
51. Smith, J. T-box genes: what they do and how they do it. *Trends Genet.* **15**, 154–158 (1999).
52. Conlon, F. L., Fairclough, L., Price, B. M., Casey, E. S. & Smith, J. C. Determinants of T box protein specificity. *Development* **128**, 3749–3758 (2001).
53. Tada, M. & Smith, J. C. T-targets: clues to understanding the functions of T-box proteins. *Dev. Growth Differ.* **43**, 1–11 (2001).
54. Papaioannou, V. E. The T-box gene family: emerging roles in development, stem cells and cancer. *Development* **141**, 3819–3833 (2014).
55. Chapman, D. L. *et al.* Expression of the T-box family genes, Tbx1-Tbx5, during early mouse development. *Dev. Dyn.* **206**, 379–390 (1996).
56. Gibson-Brown, J. J. *et al.* Evidence of a role for T-box genes in the evolution of limb morphogenesis and the specification of forelimb/hindlimb identity. *Mech. Dev.* **56**, 93–101 (1996).
57. Chapman, D. L., Agulnik, I., Hancock, S., Silver, L. M. & Papaioannou, V. E. Tbx6, a mouse T-Box gene implicated in paraxial mesoderm formation at gastrulation. *Dev. Biol.* **180**, 534–542 (1996).
58. Simon, H.-G. T-box genes and the formation of vertebrate forelimb- and hindlimb specific pattern. *Cell Tissue Res* **296**, 57–66 (1999).
59. Abrahams, A., Parker, M. I. & Prince, S. The T-box transcription factor Tbx2: its role in development and possible implication in cancer. *IUBMB Life* **62**, 92–102 (2010).
60. Wansleben, S., Peres, J., Hare, S., Goding, C. R. & Prince, S. T-box transcription factors in cancer biology. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* **1846**, 380–391 (2014).
61. Fan, W., Huang, X., Chen, C., Gray, J. & Huang, T. TBX3 and its isoform TBX3+2a are functionally distinctive in inhibition of senescence and are overexpressed in a subset of breast cancer cell lines. *Cancer Res.* **64**, 5132–5139 (2004).
62. Govoni, K. E. *et al.* Whole genome microarray analysis of growth hormone-induced gene expression in bone: T-box3, a novel transcription factor, regulates osteoblast proliferation. *Am. J. Physiol. Endocrinol. Metab.* **291**, E128-136 (2006).
63. Zhu, T. *et al.* T-box family of transcription factor-TBX5, insights in development and disease. *Am J Transl Res* **9**, 442–453 (2017).
64. Isaac, A. *et al.* Tbx genes and limb identity in chick embryo development. *Development* **125**, 1867–1875 (1998).
65. Rodriguez-Esteban, C. *et al.* The T-box genes Tbx4 and Tbx5 regulate limb outgrowth and identity. *Nature* **398**, 814–818 (1999).
66. Basson, C. T. *et al.* Different TBX5 interactions in heart and limb defined by Holt-Oram syndrome mutations. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 2919–2924 (1999).
67. Greulich, F., Rudat, C. & Kispert, A. Mechanisms of T-box gene function in the developing heart. *Cardiovasc Res* **91**, 212–222 (2011).
68. Packham, E. A. & Brook, J. D. T-box genes in human disorders. *Hum Mol Genet* **12**, R37–R44 (2003).
69. Scambler, P. J. The 22q11 deletion syndromes. *Hum. Mol. Genet.* **9**, 2421–2426 (2000).
70. Jerome, L. A. & Papaioannou, V. E. DiGeorge syndrome phenotype in mice mutant for the T-box gene, Tbx1. *Nat. Genet.* **27**, 286–291 (2001).

71. Schutte, B. C. & Murray, J. C. The Many Faces and Factors of Orofacial Clefts. *Hum Mol Genet* **8**, 1853–1859 (1999).
72. Wilkie, A. O. & Morriss-Kay, G. M. Genetics of craniofacial development and malformation. *Nat. Rev. Genet.* **2**, 458–468 (2001).
73. Lamolet, B. *et al.* A pituitary cell-restricted T box factor, Tpit, activates POMC transcription in cooperation with Pitx homeoproteins. *Cell* **104**, 849–859 (2001).
74. Yi, C. H. *et al.* Identification, mapping, and phylogenomic analysis of four new human members of the T-box gene family: EOMES, TBX6, TBX18, and TBX19. *Genomics* **55**, 10–20 (1999).
75. Collavoli, A. *et al.* TBX5 nuclear localization is mediated by dual cooperative intramolecular signals. *Journal of Molecular and Cellular Cardiology* **35**, 1191–1195 (2003).
76. Zaragoza, M. V. *et al.* Identification of the TBX5 transactivating domain and the nuclear localization signal. *Gene* **330**, 9–18 (2004).
77. Camarata, T. *et al.* LMP4 regulates Tbx5 protein subcellular localization and activity. *The Journal of Cell Biology* **174**, 339–348 (2006).
78. Kulisz, A. & Simon, H.-G. An evolutionarily conserved nuclear export signal facilitates cytoplasmic localization of the Tbx5 transcription factor. *Mol. Cell. Biol.* **28**, 1553–1564 (2008).
79. Hatcher, C. J. & Basson, C. T. Holt-Oram Syndrome and the TBX5 Transcription Factor in Cardiogenesis. in *Molecular Genetics of Cardiac Electrophysiology* (eds. Berul, C. I. & Towbin, J. A.) 297–315 (Springer US, 2000). doi:10.1007/978-1-4615-4517-0_19.
80. Hatcher, C. J., Goldstein, M. M., Mah, C. S., Delia, C. S. & Basson, C. T. Identification and localization of TBX5 transcription factor during human cardiac morphogenesis. *Dev. Dyn.* **219**, 90–95 (2000).
81. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013).
82. Melé, M. *et al.* Human genomics. The human transcriptome across tissues and individuals. *Science* **348**, 660–665 (2015).
83. Bruneau, B. G. *et al.* Chamber-specific cardiac expression of Tbx5 and heart defects in Holt-Oram syndrome. *Dev. Biol.* **211**, 100–108 (1999).
84. Mori, A. D. *et al.* Tbx5-dependent rheostatic control of cardiac gene expression and morphogenesis. *Developmental Biology* **297**, 566–586 (2006).
85. Agarwal, P. *et al.* Tbx5 is essential for forelimb bud initiation following patterning of the limb field in the mouse embryo. *Development* **130**, 623–633 (2003).
86. Rallis, C. *et al.* Tbx5 is required for forelimb bud formation and continued outgrowth. *Development* **130**, 2741–2751 (2003).
87. Ahn, D., Kourakis, M. J., Rohde, L. A., Silver, L. M. & Ho, R. K. T-box gene tbx5 is essential for formation of the pectoral limb bud. *Nature* **417**, 754–758 (2002).
88. Holt, M. & Oram, S. FAMILIAL HEART DISEASE WITH SKELETAL MALFORMATIONS. *Br Heart J* **22**, 236–242 (1960).
89. Linhares, V. L. F. *et al.* Transcriptional regulation of the murine Connexin40 promoter by cardiac factors Nkx2-5, GATA4 and Tbx5. *Cardiovasc. Res.* **64**, 402–411 (2004).
90. Newbury-Ecob, R. A., Leanage, R., Raeburn, J. A. & Young, I. D. Holt-Oram syndrome: a clinical genetic study. *J. Med. Genet.* **33**, 300–307 (1996).
91. Basson, C. T. *et al.* The clinical and genetic spectrum of the Holt-Oram syndrome (heart-hand syndrome). *N. Engl. J. Med.* **330**, 885–891 (1994).
92. McDermott, D. A. *et al.* TBX5 genetic testing validates strict clinical criteria for Holt-Oram syndrome. *Pediatr. Res.* **58**, 981–986 (2005).
93. Debeer, P., Race, V., Gewillig, M., Devriendt, K. & Frijns, J.-P. Novel TBX5 mutations in patients with Holt-Oram syndrome. *Clin. Orthop. Relat. Res.* **462**, 20–26 (2007).

94. Stenson, P. D. *et al.* The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine. *Hum. Genet.* **133**, 1–9 (2014).
95. Isphording, D., Leylek, A. M., Yeung, J., Mischel, A. & Simon, H.-G. T-box genes and congenital heart/limb malformations. *Clin. Genet.* **66**, 253–264 (2004).
96. Smemo, S. *et al.* Regulatory variation in a TBX5 enhancer leads to isolated congenital heart disease. *Hum. Mol. Genet.* **21**, 3255–3263 (2012).
97. Bruneau, B. G. *et al.* A Murine Model of Holt-Oram Syndrome Defines Roles of the T-Box Transcription Factor Tbx5 in Cardiogenesis and Disease. *Cell* **106**, 709–721 (2001).
98. DeBenedittis, P. & Jiao, K. Alternative splicing of T-box transcription factor genes. *Biochem Biophys Res Commun* **412**, 513–517 (2011).
99. Matlin, A. J., Clark, F. & Smith, C. W. J. Understanding alternative splicing: towards a cellular code. *Nat. Rev. Mol. Cell Biol.* **6**, 386–398 (2005).
100. Wang, E. T. *et al.* Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**, 470–476 (2008).
101. David, C. J. & Manley, J. L. The search for alternative splicing regulators: new approaches offer a path to a splicing code. *Genes Dev* **22**, 279–285 (2008).
102. Stennard, F. A. & Harvey, R. P. T-box transcription factors and their roles in regulatory hierarchies in the developing heart. *Development* **132**, 4897–4910 (2005).
103. Bamshad, M. *et al.* The spectrum of mutations in TBX3: Genotype/Phenotype relationship in ulnar-mammary syndrome. *Am J Hum Genet* **64**, 1550–1562 (1999).
104. Georges, R., Nemer, G., Morin, M., Lefebvre, C. & Nemer, M. Distinct Expression and Function of Alternatively Spliced Tbx5 Isoforms in Cell Growth and Differentiation. *Mol Cell Biol* **28**, 4052–4067 (2008).
105. Fan, C., Liu, M. & Wang, Q. Functional analysis of TBX5 missense mutations associated with Holt-Oram syndrome. *J. Biol. Chem.* **278**, 8780–8785 (2003).
106. Heinritz, W., Shou, L., Moschik, A. & Froster, U. G. The human TBX5 gene mutation database. *Human Mutation* **26**, 397–397 (2005).
107. Borozdin, W. *et al.* Expanding the spectrum ofTBX5 mutations in Holt-Oram syndrome: detection of two intragenic deletions by quantitative real time PCR, and report of eight novel point mutations. *Hum. Mutat.* **27**, 975–976 (2006).
108. Wang, J. & Schwartz, R. J. Post-translational Modification. in *Congenital Heart Diseases: The Broken Heart: Clinical Features, Human Genetics and Molecular Pathways* (eds. Rickert-Sperling, S., Kelly, R. G. & Driscoll, D. J.) 173–202 (Springer, 2016). doi:10.1007/978-3-7091-1883-2_14.
109. Miller, S. A. & Weinmann, A. S. Common themes emerge in the transcriptional control of T helper and developmental cell fate decisions regulated by the T-box, GATA and ROR families. *Immunology* **126**, 306–315 (2009).
110. Miller, S. A., Huang, A. C., Miazgowicz, M. M., Brassil, M. M. & Weinmann, A. S. Coordinated but physically separable interaction with H3K27-demethylase and H3K4-methyltransferase activities are required for T-box protein-mediated activation of developmental gene expression. *Genes Dev.* **22**, 2980–2993 (2008).
111. Beketaev, I. *et al.* Potentiation of Tbx5-mediated transactivation by SUMO conjugation and protein inhibitor of activated STAT 1 (PIAS1). *The International Journal of Biochemistry & Cell Biology* **50**, 82–92 (2014).
112. Lewandowski, S. L. *et al.* Histone deacetylase 3 modulates Tbx5 activity to regulate early cardiogenesis. *Hum Mol Genet* **23**, 3801–3809 (2014).
113. Hassani, M. S. & Green, J. R. Active Learning for microRNA Prediction. in *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* 5–12 (2018). doi:10.1109/BIBM.2018.8621144.

114. Latronico, M. V. G. & Condorelli, G. MicroRNAs and cardiac pathology. *Nat Rev Cardiol* **6**, 419–429 (2009).
115. Zhao, Y. *et al.* Dysregulation of cardiogenesis, cardiac conduction, and cell cycle in mice lacking miRNA-1-2. *Cell* **129**, 303–317 (2007).
116. Liu, N. *et al.* microRNA-133a regulates cardiomyocyte proliferation and suppresses smooth muscle gene expression in the heart. *Genes Dev.* **22**, 3242–3254 (2008).
117. Chiavacci, E. *et al.* MicroRNA 218 mediates the effects of Tbx5a over-expression on zebrafish heart development. *PLoS ONE* **7**, e50536 (2012).
118. Wang, F. *et al.* MiR-10a and MiR-10b Target the 3′-Untranslated Region of TBX5 to Repress Its Expression. *Pediatr Cardiol* **35**, 1072–1079 (2014).
119. Hussain, N. *et al.* Down-regulation of miR-10a-5p in synoviocytes contributes to TBX5-controlled joint inflammation. *Journal of Cellular and Molecular Medicine* **22**, 241–250 (2018).
120. Hornstein, E. *et al.* The microRNA miR-196 acts upstream of Hoxb8 and Shh in limb development. *Nature* **438**, 671–674 (2005).
121. Viger, R. S., Guittot, S. M., Anttonen, M., Wilson, D. B. & Heikinheimo, M. Role of the GATA Family of Transcription Factors in Endocrine Development, Function, and Disease. *Mol Endocrinol* **22**, 781–798 (2008).
122. Hiroi, Y. *et al.* Tbx5 associates with Nkx2-5 and synergistically promotes cardiomyocyte differentiation. *Nature Genetics* **28**, 276 (2001).
123. Luna-Zurita, L. *et al.* Complex Interdependence Regulates Heterotypic Transcription Factor Distribution and Coordinates Cardiogenesis. *Cell* **164**, 999–1014 (2016).
124. Maitra, M. *et al.* Interaction of Gata4 and Gata6 with Tbx5 is critical for normal cardiac development. *Dev. Biol.* **326**, 368–377 (2009).
125. Wang, C., Cao, D., Wang, Q. & Wang, D.-Z. Synergistic activation of cardiac genes by myocardin and Tbx5. *PLoS ONE* **6**, e24242 (2011).
126. Brown, D. D. *et al.* Tbx5 and Tbx20 act synergistically to control vertebrate heart morphogenesis. *Development* **132**, 553–563 (2005).
127. Ghosh, T. K. *et al.* Physical interaction between TBX5 and MEF2C is required for early heart development. *Mol. Cell. Biol.* **29**, 2205–2218 (2009).
128. Lickert, H. *et al.* Baf60c is essential for function of BAF chromatin remodelling complexes in heart development. *Nature* **432**, 107–112 (2004).
129. Takeuchi, J. K. & Bruneau, B. G. Directed transdifferentiation of mouse mesoderm to heart tissue by defined factors. *Nature* **459**, 708–711 (2009).
130. Christoffels, V. M. *et al.* T-box transcription factor Tbx2 represses differentiation and formation of the cardiac chambers. *Dev. Dyn.* **229**, 763–770 (2004).
131. He, M. l, Wen, L., Campbell, C. E., Wu, J. Y. & Rao, Y. Transcription repression by Xenopus ET and its human ortholog TBX3, a gene involved in ulnar-mammary syndrome. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10212–10217 (1999).
132. Lingbeek, M. E., Jacobs, J. J. L. & van Lohuizen, M. The T-box repressors TBX2 and TBX3 specifically regulate the tumor suppressor gene p14ARF via a variant T-site in the initiator. *J. Biol. Chem.* **277**, 26120–26127 (2002).
133. Waldron, L. *et al.* The cardiac TBX5 interactome reveals a chromatin remodeling network essential for cardiac septation. *Dev Cell* **36**, 262–275 (2016).
134. Brassington, A.-M. E. *et al.* Expressivity of Holt-Oram syndrome is not predicted by TBX5 genotype. *Am. J. Hum. Genet.* **73**, 74–85 (2003).
135. Arnolds, D. E. *et al.* TBX5 drives Scn5a expression to regulate cardiac conduction system function. *J. Clin. Invest.* **122**, 2509–2518 (2012).
136. Goetz, S. C., Brown, D. D. & Conlon, F. L. TBX5 is required for embryonic cardiac cell cycle progression. *Development* **133**, 2575–2584 (2006).

137. Hatcher, C. J. *et al.* TBX5 transcription factor regulates cell proliferation during cardiogenesis. *Dev. Biol.* **230**, 177–188 (2001).
138. Puskaric, S. *et al.* Shox2 mediates Tbx5 activity by regulating Bmp4 in the pacemaker region of the developing heart. *Hum. Mol. Genet.* **19**, 4625–4633 (2010).
139. He, A., Kong, S. W., Ma, Q. & Pu, W. T. Co-occupancy by multiple cardiac transcription factors identifies transcriptional enhancers active in heart. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 5632–5637 (2011).
140. Koshiba-Takeuchi, K. *et al.* Cooperative and antagonistic interactions between Sall4 and Tbx5 pattern the mouse limb and heart. *Nature Genetics* **38**, 175–183 (2006).
141. Ng, J. K. *et al.* The limb identity gene Tbx5 promotes limb initiation by interacting with Wnt2b and Fgf10. *Development* **129**, 5161–5170 (2002).
142. Krause, A. *et al.* Tbx5 and Tbx4 transcription factors interact with a new chicken PDZ-LIM protein in limb and heart development. *Developmental Biology* **273**, 106–120 (2004).
143. Wilson, G. N. Correlated heart/limb anomalies in Mendelian syndromes provide evidence for a cardiomelic developmental field. *American Journal of Medical Genetics* **76**, 297–305 (1998).
144. Barr, M., Blackburn, W. R. & Cooley, N. R. Human fetal somatic and visceral morphometrics. *Teratology* **49**, 487–496 (1994).
145. Davenport, T. G., Jerome-Majewska, L. A. & Papaioannou, V. E. Mammary gland, limb and yolk sac defects in mice lacking Tbx3, the gene mutated in human ulnar mammary syndrome. *Development* **130**, 2263–2273 (2003).
146. Yang, L. Isl1Cre reveals a common Bmp pathway in heart and limb development. *Development* **133**, 1575–1585 (2006).
147. Crossley, P. H., Minowada, G., MacArthur, C. A. & Martin, G. R. Roles for FGF8 in the Induction, Initiation, and Maintenance of Chick Limb Development. *Cell* **84**, 127–136 (1996).
148. Meyers, E. N. & Martin, G. R. Differences in Left-Right Axis Pathways in Mouse and Chick: Functions of FGF8 and SHH. *Science* **285**, 403–406 (1999).
149. Reifers, F., Walsh, E. C., Leger, S., Stainier, D. Y. & Brand, M. Induction and differentiation of the zebrafish heart requires fibroblast growth factor 8 (fgf8/acerebellar). *Development* **127**, 225–235 (2000).
150. Stratford, T., Logan, C., Zile, M. & Maden, M. Abnormal anteroposterior and dorsoventral patterning of the limb bud in the absence of retinoids. *Mechanisms of Development* **81**, 115–125 (1999).
151. Zile, M. H. *et al.* Retinoid Signaling Is Required to Complete the Vertebrate Cardiac Left/Right Asymmetry Pathway. *Developmental Biology* **223**, 323–338 (2000).
152. Charite, J., McFadden, D. G. & Olson, E. N. The bHLH transcription factor dHAND controls Sonic hedgehog expression and establishment of the zone of polarizing activity during limb development. *Development* **127**, 2461–2470 (2000).
153. Briegel, K. J. & Joyner, A. L. Identification and Characterization of Lbh, a Novel Conserved Nuclear Protein Expressed during Early Limb and Heart Development. *Developmental Biology* **233**, 291–304 (2001).
154. Musarò, A. The Basis of Muscle Regeneration. *Advances in Biology* **2014**, 1–16 (2014).
155. Yin, H., Price, F. & Rudnicki, M. A. Satellite cells and the muscle stem cell niche. *Physiol. Rev.* **93**, 23–67 (2013).
156. Jones, D. L. & Wagers, A. J. No place like home: anatomy and function of the stem cell niche. *Nature Reviews Molecular Cell Biology* **9**, 11–21 (2008).
157. Doherty, M. K. & Whitfield, P. D. Proteomics moves from expression to turnover: update and future perspective. *Expert Review of Proteomics* **8**, 325–334 (2011).
158. Dhawan, J. & Rando, T. A. Stem cells in postnatal myogenesis: molecular mechanisms of satellite cell quiescence, activation and replenishment. *Trends Cell Biol.* **15**, 666–673 (2005).

159. Wozniak, A. C., Kong, J., Bock, E., Pilipowicz, O. & Anderson, J. E. Signaling satellite-cell activation in skeletal muscle: Markers, models, stretch, and potential alternate pathways. *Muscle Nerve* **31**, 283–300 (2005).
160. Kuang, S., Gillespie, M. A. & Rudnicki, M. A. Niche Regulation of Muscle Satellite Cell Self-Renewal and Differentiation. *Cell Stem Cell* **2**, 22–31 (2008).
161. Ten Broek, R. W., Grefte, S. & Von den Hoff, J. W. Regulatory factors and cell populations involved in skeletal muscle regeneration. *J. Cell. Physiol.* n/a-n/a (2010) doi:10.1002/jcp.22127.
162. Wakelam, M. J. The fusion of myoblasts. *Biochemical Journal* **228**, 1–12 (1985).
163. Allen, R. E., Merkel, R. A. & Young, R. B. Cellular aspects of muscle growth: myogenic cell proliferation. *J. Anim. Sci.* **49**, 115–127 (1979).
164. Allen, R. E. & Boxhorn, L. K. Regulation of skeletal muscle satellite cell proliferation and differentiation by transforming growth factor-beta, insulin-like growth factor I, and fibroblast growth factor. *J. Cell. Physiol.* **138**, 311–315 (1989).
165. Luca, G., Ferretti, R., Bruschi, M., Mezzaroma, E. & Caruso, M. Cyclin D3 critically regulates the balance between self-renewal and differentiation in skeletal muscle stem cells. *Stem Cells* **31**, 2478–2491 (2013).
166. Rodriguez-Puebla, M. L. *et al.* Cyclin D2 Overexpression in Transgenic Mice Induces Thymic and Epidermal Hyperplasia whereas Cyclin D3 Expression Results Only in Epidermal Hyperplasia. *Am J Pathol* **157**, 1039–1050 (2000).
167. Bartkova, J., Lukas, J., Strauss, M. & Bartek, J. Cyclin D3: requirement for G1/S transition and high abundance in quiescent tissues suggest a dual role in proliferation and di€ erentiation. 11.
168. Skapek, S. X., Rhee, J., Spicer, D. B. & Lassar, A. B. Inhibition of myogenic differentiation in proliferating myoblasts by cyclin D1-dependent kinase. *Science* **267**, 1022–1024 (1995).
169. Olson, E. N. Interplay between proliferation and differentiation within the myogenic lineage. *Developmental Biology* **154**, 261–272 (1992).
170. Khanjyan, M. V., Yang, J., Kayali, R., Caldwell, T. & Bertoni, C. A high-content, high-throughput siRNA screen identifies cyclin D2 as a potent regulator of muscle progenitor cell fusion and a target to enhance muscle regeneration. *Hum Mol Genet* **22**, 3283–3295 (2013).
171. Saab, R. *et al.* Pharmacologic inhibition of cyclin-dependent kinase 4/6 activity arrests proliferation in myoblasts and rhabdomyosarcoma-derived cells. *Mol Cancer Ther* **5**, 1299–1308 (2006).
172. Biferi, M. G. *et al.* Proliferation of Multiple Cell Types in the Skeletal Muscle Tissue Elicited by Acute p21 Suppression. *Molecular Therapy* **23**, 885–895 (2015).
173. Jahn, L., Sadoshima, J. & Izumo, S. Cyclins and Cyclin-Dependent Kinases Are Differentially Regulated during Terminal Differentiation of C2C12 Muscle Cells. *Experimental Cell Research* **212**, 297–307 (1994).
174. Cannon, J. G. & St Pierre, B. A. Cytokines in exertion-induced skeletal muscle injury. *Mol. Cell. Biochem.* **179**, 159–167 (1998).
175. Kovanen, V. Intramuscular Extracellular Matrix: Complex Environment of Muscle Cells. *Exercise and Sport Sciences Reviews* **30**, 20–25 (2002).
176. Doumit, M. E., Cook, D. R. & Merkel, R. A. Fibroblast growth factor, epidermal growth factor, insulin-like growth factors, and platelet-derived growth factor-BB stimulate proliferation of clonally derived porcine myogenic satellite cells. *Journal of Cellular Physiology* **157**, 326–332 (1993).
177. Hayashi, S. *et al.* Sequence of IGF-I, IGF-II, and HGF expression in regenerating skeletal muscle. *Histochem Cell Biol* **122**, 427–434 (2000).
178. Ratajczak, M. Z. *et al.* Expression of Functional CXCR4 by Muscle Satellite Cells and Secretion of SDF-1 by Muscle-Derived Fibroblasts is Associated with the Presence of Both Muscle Progenitors in Bone Marrow and Hematopoietic Stem/Progenitor Cells in Muscles. *STEM CELLS* **21**, 363–371 (2003).

179. Sheehan, S. M. & Allen, R. E. Skeletal muscle satellite cell proliferation in response to members of the fibroblast growth factor family and hepatocyte growth factor. *J. Cell. Physiol.* **181**, 499–506 (1999).
180. Liu, D., Black, B. L. & Derynck, R. TGF-β inhibits muscle differentiation through functional repression of myogenic transcription factors by Smad3. *Genes Dev* **15**, 2950–2966 (2001).
181. Lee, S. J. & McPherron, A. C. Myostatin and the control of skeletal muscle mass. *Curr. Opin. Genet. Dev.* **9**, 604–607 (1999).
182. Sato, K. *et al.* Improvement of muscle healing through enhancement of muscle regeneration and prevention of fibrosis. *Muscle & Nerve* **28**, 365–372 (2003).
183. Moore, K. A. & Lemischka, I. R. Stem Cells and Their Niches. *Science* **311**, 1880–1885 (2006).
184. Grégoire, S. & Yang, X.-J. Association with class IIa histone deacetylases upregulates the sumoylation of MEF2 transcription factors. *Mol. Cell. Biol.* **25**, 2273–2287 (2005).
185. Roth, J.-F. *et al.* Differential role of p300 and CBP acetyltransferase during myogenesis: p300 acts upstream of MyoD and Myf5. *EMBO J.* **22**, 5186–5196 (2003).
186. Huh, M. S., Parker, M. H., Scimè, A., Parks, R. & Rudnicki, M. A. Rb is required for progression through myogenic differentiation but not maintenance of terminal differentiation. *J Cell Biol* **166**, 865–876 (2004).
187. Khan, P., Linkhart, B. & Simon, H.-G. Different regulation of T-box genes Tbx4 and Tbx5 during limb development and limb regeneration. *Dev. Biol.* **250**, 383–392 (2002).
188. Minguillon, C., Del Buono, J. & Logan, M. P. Tbx5 and Tbx4 are not sufficient to determine limb-specific morphologies but have common roles in initiating limb outgrowth. *Dev. Cell* **8**, 75–84 (2005).
189. Lu, J. *et al.* Induction of apoptosis and inhibition of cell growth by tbx5 knockdown contribute to dysmorphogenesis in Zebrafish embryos. *Journal of Biomedical Science* **18**, 73 (2011).
190. Kozhemyakina, E., Ionescu, A. & Lassar, A. B. GATA6 Is a Crucial Regulator of Shh in the Limb Bud. *PLOS Genetics* **10**, e1004072 (2014).
191. Khoury, M. F. E. A Novel Role Of GATA6 in Rhabdomyosarcoma. (c2016). in (2016). doi:10.26756/th.2016.28.
192. Merrell, A. J. & Kardon, G. Development of the diaphragm -- a skeletal muscle essential for mammalian respiration. *FEBS J.* **280**, 4026–4035 (2013).
193. Kasahara, H., Bartunkova, S., Schinke, M., Tanaka, M. & Izumo, S. Cardiac and extracardiac expression of Csx/Nkx2.5 homeodomain protein. *Circ. Res.* **82**, 936–946 (1998).
194. Gladman, J. T. *et al.* NKX2-5, a modifier of skeletal muscle pathology due to RNA toxicity. *Hum Mol Genet* **24**, 251–264 (2015).
195. Chen, J.-F. *et al.* The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nat. Genet.* **38**, 228–233 (2006).
196. Chen Jinghai *et al.* mir-17–92 Cluster Is Required for and Sufficient to Induce Cardiomyocyte Proliferation in Postnatal and Adult Hearts. *Circulation Research* **112**, 1557–1566 (2013).
197. Tabebordbar, M., Wang, E. T. & Wagers, A. J. Skeletal Muscle Degenerative Diseases and Strategies for Therapeutic Muscle Repair. *Annual Review of Pathology: Mechanisms of Disease* **8**, 441–475 (2013).
198. Karalaki, M., Fili, S., Philippou, A. & Koutsilieris, M. Muscle Regeneration: Cellular and Molecular Events. *In Vivo* **23**, 779–796 (2009).
199. Bonaldo, P. & Sandri, M. Cellular and molecular mechanisms of muscle atrophy. *Disease Models & Mechanisms* **6**, 25–39 (2013).
200. Jang, Y. C., Sinha, M., Cerletti, M., Dall'Osso, C. & Wagers, A. J. Skeletal Muscle Stem Cells: Effects of Aging and Metabolism on Muscle Regenerative Function. *Cold Spring Harbor Symposia on Quantitative Biology* **76**, 101–111 (2011).
201. Acharyya, S. *et al.* Cancer cachexia is regulated by selective targeting of skeletal muscle gene products. *J Clin Invest* **114**, 370–378 (2004).

202. Dagher, R. & Helman, L. Rhabdomyosarcoma: An Overview. *The Oncologist* **4**, 34–44 (1999).
203. Linardic, C. M., Downie, D. L., Qualman, S., Bentley, R. C. & Counter, C. M. Genetic Modeling of Human Rhabdomyosarcoma. *Cancer Res* **65**, 4490–4495 (2005).
204. Burghes, A. H. M. *et al.* A eDNA clone from the Duchennel Becker muscular dystrophy gene. 4 (1987).
205. Chaturvedi, L. S., Mukherjee, M., Srivastava, S., Mittal, R. D. & Mittal, B. Point mutation and polymorphism in Duchenne/Becker Muscular Dystrophy (D/BMD) patients. *Exp Mol Med* **33**, 251–256 (2001).
206. Sacco, A. *et al.* Short Telomeres and Stem Cell Exhaustion Model Duchenne Muscular Dystrophy in mdx/mTR Mice. *Cell* **143**, 1059–1071 (2010).
207. Yucel, N., Chang, A. C., Day, J. W., Rosenthal, N. & Blau, H. M. Humanizing the mdx mouse model of DMD: the long and the short of it. *npj Regenerative Medicine* **3**, 4 (2018).
208. Brenman, J. E. *et al.* Interaction of Nitric Oxide Synthase with the Postsynaptic Density Protein PSD-95 and ⎵1-Syntrophin Mediated by PDZ Domains. 11.
209. Duchenne Muscular Dystrophy. *NORD (National Organization for Rare Disorders)* https://rarediseases.org/rare-diseases/duchenne-muscular-dystrophy/.
210. Watchko, J. F., O'Day, T. L. & Hoffman, E. P. Functional characteristics of dystrophic skeletal muscle: insights from animal models. *Journal of Applied Physiology* **93**, 407–417 (2002).
211. Stedman, H. H. *et al.* The mdx mouse diaphragm reproduces the degenerative changes of Duchenne muscular dystrophy. *Nature* **352**, 536–539 (1991).
212. Takeshita, H. *et al.* Modified forelimb grip strength test detects aging-associated physiological decline in skeletal muscle function in male mice. *Scientific Reports* **7**, 1–9 (2017).
213. d'ALBIS, A., Couteaux, R., Janmot, C., Roulet, A. & Mira, J.-C. Regeneration after cardiotoxin injury of innervated and denervated slow and fast muscles of mammals. Myosin isoform analysis. *Eur J Biochem* **174**, 103–110 (1988).
214. Harris, J. B. & Johnson, M. A. FURTHER OBSERVATIONS ON THE PATHOLOGICAL RESPONSES OF RAT SKELETAL MUSCLE TO TOXINS ISOLATED FROM THE VENOM OF THE AUSTRALIAN TIGER SNAKE, NOTECHIS SCUTATUS SCUTATUS. *Clin Exp Pharmacol Physiol* **5**, 587–600 (1978).
215. Fisher, B. D. & Rathgaber, M. An Overview of Muscle Regeneration Following Acute Injury. *J Phys Ther Sci* **18**, 57–66 (2006).
216. Grounds, MirandaD. & McGeachie, JohnK. A comparison of muscle precursor replication in crush-injured skeletal muscle of Swiss and BALBc mice. *Cell Tissue Res.* **255**, (1989).
217. Muir, L. A. & Chamberlain, J. S. Emerging strategies for cell and gene therapy of the muscular dystrophies. *Expert Rev. Mol. Med.* **11**, e18 (2009).
218. Schultz, B. R. & Chamberlain, J. S. Recombinant Adeno-associated Virus Transduction and Integration. *Mol Ther* **16**, 1189–1199 (2008).
219. Zaldumbide, A. & Hoeben, R. C. How not to be seen: immune-evasion strategies in gene therapy. *Gene Therapy* **15**, 239–246 (2008).

2. Chapter I: Novel Exons in the *Tbx5* Gene Locus Generate Protein Isoforms with Distinct Expression Domains and Function

Abir Yamak[1,*], Romain O. Georges[2,*], **Massomeh Sheikh-Hassani**[1], Martin Morin[2], Hiba Komati[1] and Mona Nemer[1,2†]

1- Laboratory of Molecular Genetics and Cardiac Regeneration, Department of Biochemistry, Microbiology, and Immunology, University of Ottawa, Ottawa (Ontario), K1N 6N5 Canada

2- Graduate Program in Molecular Biology, Institut de Recherches Cliniques de Montréal (IRCM), Université de Montréal, Montréal (Québec), H2W 1R7 Canada

*Co-first authors

† To whom correspondence should be addressed: Professor Mona Nemer, University of Ottawa, 550 Cumberland, Room 246, Ottawa, Ontario K1N 6N5, Tel: 613 562 5270/Fax 613 562 5271, Email: mona.nemer@uottawa.ca

2.1 Statement of the Manuscript

The manuscript "Novel exons in the Tbx5 gene locus generate protein isoforms with distinct expression domains and function" was published in the journal of biological chemistry (JBC) on January 25, 2015, pubmed ID 25623069.

2.2 Contribution Statement

In this manuscript, MSH performed experiments (Figures 2.5F, 2.6, 2.7, 2.8 and 2.9), analyzed/interpreted results and helped with writing the manuscript. AY performed experiments and wrote manuscript (Figures 2.1-2.5E and 2.10). HK helped perform experiments (Figure 2.7 and 2.8) and analyzed and interpreted results. MM and RG helped with some experiments (Figure 2.1). MN conceived the project, designed experiments, interpreted results and helped write the manuscript.

2.3 Acknowledgements

We thank Drs Benoit Bruneau and Jacques Drouin for gifts of plasmids, Chantal Lefebvre and Nathalie Bouchard for technical assistance and Hélène Touchette for secretarial help.

2.4 Sources of Funding

This work was supported by a grant from the Canadian Heart and Stroke Foundation and the McCain Foundation.

2.5 Disclosures

None

2.6 Abstract

TBX5 is the gene mutated in Holt-Oram Syndrome, an autosomal dominant disorder with complex heart and limb deformities. Its protein product is a member of the T-box family of transcription factors and an evolutionary conserved dosage sensitive regulator of heart and limb development. Understanding *TBX5* regulation is therefore of paramount importance. Here we uncover the existence of novel exons and provide evidence that TBX5 activity may be extensively regulated through alternative splicing to produce protein isoforms with differing N- and C-terminal domains. These isoforms are also present in human heart indicative of an evolutionary conserved regulatory mechanism. The newly identified isoforms have different transcriptional properties and can antagonize TBX5a target gene activation. Droplet digital PCR as well as immunohistochemistry with isoform specific antibodies reveal differential as well as overlapping expression domains. In particular, we find that the predominant isoform in skeletal myoblasts is Tbx5c and we show that it is dramatically upregulated in differentiating myotubes and is essential for myotube formation. Mechanistically, TBX5c antagonizes TBX5a activation of pro-proliferative signals such as IGF-1, FGF-10 and BMP4. The results provide new insight into *Tbx5* regulation and function that will further our understanding of its role in health and disease. The finding of new exons in the *Tbx5* locus may also be relevant to mutational screening especially in the 30% of Holt-Oram Syndrome patients with no mutations in the known *TBX5a* exons.

2.7 Introduction

Holt-Oram syndrome (HOS) is an autosomal dominant disorder characterized by upper limb and cardiac defects (1,2). The most common structural heart abnormalities include Atrial Septal Defects (ASDs) and Ventricular Septal Defects (VSDs). Conduction defects have also been commonly seen and they mostly involve atrioventricular blocks. Hypoplastic left ventricle, mitral valve problem and endocardial cushion defects have also been reported in HOS patients (1,2). Genetic linkage analyses have mapped the disease to the chromosomal locus where *TBX5* is located and mutations in *TBX5* have been found in patients with HOS. Moreover, *Tbx5* expression pattern in the upper limb, atria and left ventricle along with mouse genetics studies have strengthened the causative link between *TBX5* and HOS (3). Over 70 mutations in the *TBX5* locus have been identified so far in HOS patients (4). Many result in no protein production or in truncated proteins. Other more subtle mutations generate functionally impaired proteins with altered subcellular localisation, DNA binding, transcriptional activity, and/or interaction with cofactors (5-7). These findings led to the suggestion that haploinsufficiency may be the mechanism of pathogenesis but this remains uncertain in many cases. Interestingly, in about 30-35% of HOS patients no mutations in *TBX5* coding sequences or intron-exon junctions are detected (8) which has raised the controversial suggestion of the existence of another as yet unidentified HOS-causing locus. An alternative explanation could be that unscreened mutations within presumed untranscribed regions of *TBX5* account for this low detection rate. Consistent with this, we recently reported the existence of a new exon downstream of the T-box whose alternative splicing results in a TBX5 isoform lacking the entire C-terminal which contains several functional domains (9).

TBX5 is a member of the large family of T-box transcription factors critical for early cellular commitment, differentiation and organ development (10). T-box or Tbx-proteins bind specific DNA motifs, called TBE (T-box Binding Element) to activate or repress target promoters. TBX5 appears to act essentially as a transcriptional activator and cooperates with other transcription factors such as GATA4 and NKX2.5 to synergistically regulate downstream targets (3,6,11). As such, TBX5 activity can be modulated at the DNA binding level and through protein-protein interactions. In addition to transcriptional regulators, TBX5 was shown to interact with the cytoskeletal associated LIM protein LMP4 which represses its transcriptional activity, possibly by stimulating its cytoplasmic redistribution (12). $TBX5_{1-518}$ (referred to thereafter as TBX5a) resides largely if not exclusively in the nucleus and two nuclear localisation signals (NLS) have been identified, one within the T-box DNA binding domain, and another, between amino acids (AA) 325-340 outside the T-box (13). A putative nuclear export signal within the T-box has also been suggested to mediate Crml-dependent nuclear export of TBX5 (14) but this has been challenged based on the crystal structure of the T-box domain of TBX5 in DNA bound and unbound forms (15). The crystal structure also identified the T-box residues that contact DNA as those towards the C-terminal of the T-box. Interactions between TBX5 and other transcriptional regulators also require the T-box (3,9). In addition to DNA binding, transcriptional activation by TBX5 depends on sequences outside the T-box. Deletion analysis showed that removal of the N-terminal 50AA decreases TBX5 transcriptional activity albeit not as severely as removal of the C-terminal 100AA. Another domain that contributes to transcriptional activation was localised between AA255-316 just C-terminal of the T-box. Interestingly, the 3 domains are differentially required for physical and functional interaction with GATA4 and NKX2.5 (9). Gal4-TBX5 chimera, confirmed the

presence of a potent Transcriptional Activation Domain (TAD) in the last 250AA of TBX5a (16); an autonomous TAD was also mapped between AA339 and 379 (13).

In the present work, we report the existence of new *Tbx5* exons and additional alternatively spliced TBX5 isoforms that differ from TBX5 in the N- or C-terminal domains. We show that the new isoforms are expressed in distinct domains that sometimes overlap with Tbx5a. We show that one of the novel isoforms plays a unique role in skeletal muscle differentiation where it suppresses proliferation signals and induces differentiation. The results provide important information on the *Tbx5* locus and novel insight into *Tbx5* regulation and function.

2.8 Materials and Methods

Cloning of novel *Tbx5* isoforms. *Tbx5* transcripts with variable 3'end (*Tbx5c* and *d*) were obtained using nested cDNA amplification of mRNA from embryonic hearts and limbs. cDNAs were amplified by reverse transcription (RT-PCR) using a 5' oligonucleotide spanning the first codons of the T-box and at the 3' end, an oligo dT. The resulting products were cloned into the bluescript plasmid and individual clones were sequenced. Full length clones containing the N-terminal sequences were then amplified using specific primers, and subcloned in the pcDNA3 expression vector in phase with Kozak-triple flag epitope. To isolate isoforms with variable 5' end (*Tbx5e*), primer extension amplification was used with the primer sequence overlapping the 5' end of the T-box sequences. Oligos used were: 5'-GGAGGTACCGCCGATACAGATGAGGGCTTTG-3' (forward for *Tbx5a, Tbx5c, Tbx5d* and *Tbx5b*); 5'- GCAGGTACCGAAGGAATCAAGGTGTTTCTTCATG-3' forward for *Tbx5e*; 5'- CCGGAATTCTTAGCTATTCTCACTCCACTCTGG-3'reverse for *Tbx5a* and *Tbx5e*; 5'- GCACTCGAGCTAATGAAAGGATGGTGAGAGAG-3' reverse for *Tbx5c*; 5'-

CGCGAATTCCTATATTTCTGTGCCACTTACTT-3' reverse for *Tbx5d*; 5'-CCCCTCGAGCTAAAGCAGAGGCCTTTGCATCCGAG-3' reverse for *Tbx5b*. GATA4 and NKX2.5 expression vectors as well as the luciferase reporters were previously described (9,17-19).

Cell culture and transfections. Cell lines were maintained in culture and transfected as described previously (9,20) C2C12 were maintained in 10% FBS for 24h and the media was switched to 1% BSA or 2% HS for differentiation when indicated. For co-transfection assays, the total amount of DNA was maintained constant by adding the appropriate amount of empty DNA vector. *Fgf10*-luc reporter construct was a kind gift from Dr. Benoit Bruneau from the Gladstone Institute of Cardiovascular Disease and was previously described (21). *Nppa-luc*, *Bclx-luc*, *Ccnd1-luc* and *Bmp4-luc* reporters were previously described (19,20,22). GAL4 fusion proteins were designed as previously described (23).

Electrophoretic mobility shift assay (EMSA). Nuclear extracts were prepared from 293T cells. Binding reactions were done at room temperature using 1ug of poly (dI-dC). The TBE probe used was from the *Nppa* promoter and previously described (9).

In vitro pull-down assay: Production of the GST-TBX5a and MBP-NKX2.5 constructs and the pull down assays were done as previously described (24).

Protein analysis. Western blots were done on nuclear extracts from 293T cells overexpressing the relevant *Tbx5* constructs and Flag or TBX5 antibody as previously described (9). Endogenous TBX5 proteins were analyzed in 2 relevant mouse lines, the TC13 cardiogenic cell line and in C2C12 myoblasts as well as in 30 d old mouse hearts. Protein extraction and analysis was as reported in Georges *et al* (9). The N-terminal TBX5a antibody was previously described. The C-terminal antibody TBX5a as well as TBX5c and TBX5d antibodies were

similarly raised in rabbits against the unique C-terminal domains: AA331-425 for TBX5a, AA 327-404 for TBX5c and AA 327-376 for TBX5d.

Immunocytochemistry was done on 293T cells overexpressing the relevant Flag-TBX5 constructs as previously described (9). Mouse anti-Flag M2 was purchased from Sigma (F1804). Immunofluorescence was performed on C2C12 cells overexpressing the indicated HA-TBX5 constructs. Myosin Heavy Chain and Myogenin antibodies were both obtained from mouse hybridoma cells (MF20 and F5D respectively). HA antibody was purchased from Santa Cruz (SC-805). Immunohistochemistry was done on mouse tissues at different embryonic stages as previously described (25).

RNA analysis. Droplet digital PCR was carried out on the QX200 Biorad Droplet Digital PCR system using cDNA from embryonic and postnatal tissues as well as cardiac and muscle cells using sequence specific primers as indicated and according to manufacturer's protocol. Primer sequences are available on request. QPCR analysis was performed on C2C12 cells overexpressing the indicated TBX5 constructs as previously described (25). SiRNAs were obtained from Sigma and transfected in C2C12 cells using the Hiperfect transfection reagent from Qiagen (Cat. # 301705). SiRNA sequences are available upon request.

2.9 Results

Identification of novel *Tbx5* exons

To analyze endogenous TBX5 protein, we developed a specific antibody against the N-terminal TBX5 region; in Western blots, this antibody detected, in addition to the expected TBX5a band, immunoreactive bands that co-migrated with a TBX5 protein truncated of the last 118 amino acids (9). To determine if these bands represent novel TBX5 isoforms, we used

a PCR amplification strategy to isolate *Tbx5* cDNAs from mouse heart and limb RNA. A 5' oligonucleotide primer spanning the first codons of the T-box and a 3' oligo dT primer allowed amplification of *Tbx5* transcripts containing the T-domain together with any variable 3' sequences. Isolation and sequencing of several independent cDNA clones revealed the presence of 4 distinct transcripts that may result in 4 different TBX5 proteins: in addition to sequences corresponding to *Tbx5a* and the previously described short isoform, two new sequences were identified which would encode novel TBX5 isoforms of 404 and 376AA, termed *Tbx5c* and *Tbx5d* respectively (**Figure 2.1 A**). These two isoforms result from the use of two alternative exons (exon 9a or exon 10), which are mutually exclusive with exon 11. *Tbx5c* is generated from the splicing of exon 9 to exon 10 instead of exon 11, whereas *Tbx5d* originates from the usage of an alternative exon 9, named exon 9a, which emanates from the retention of additional intronic sequences 3' of exon 9 (**Figure 2.1 C**). The result in both cases is the addition, after position 327 of a 77 AA (TBX5c) or 49AA (TBX5d) divergent region (**Figure 2.1 B and D**). Once the additional exons were sequenced, we cloned the entire cDNA coding sequence of the novel isoforms from mouse embryonic atria (*Tbx5d*) and forelimb (*Tbx5c*) using primers spanning the 1st codon (ATG) and the stop codon of the different isoforms. Sequence analysis of the genomic *TBX5* locus in existing data bases (NCBI, Ensembl) revealed the possibility of a fifth TBX5 isoform, TBX5e (**Figure 2.1 A**). This isoform would result from the splicing of an alternative exon 1 in mice (termed exon1b) with exon 3; thus, skipping exon 2 leading to an N-terminal truncation of 50AA (**Figure 2.1 C and D**). Exon 1b is highly homologous to the human exon 1. To confirm the *in vivo* existence of such *Tbx5e* transcripts, a primer overlapping exon 1b and 3 was used in RT-PCR amplification together with a reverse primer spanning the stop codon of exon 11; this resulted in the isolation

of *Tbx5e* cDNA clones from mouse heart RNA. Thus, a N-terminally truncated TBX5 isoform is present in both human and rodents.

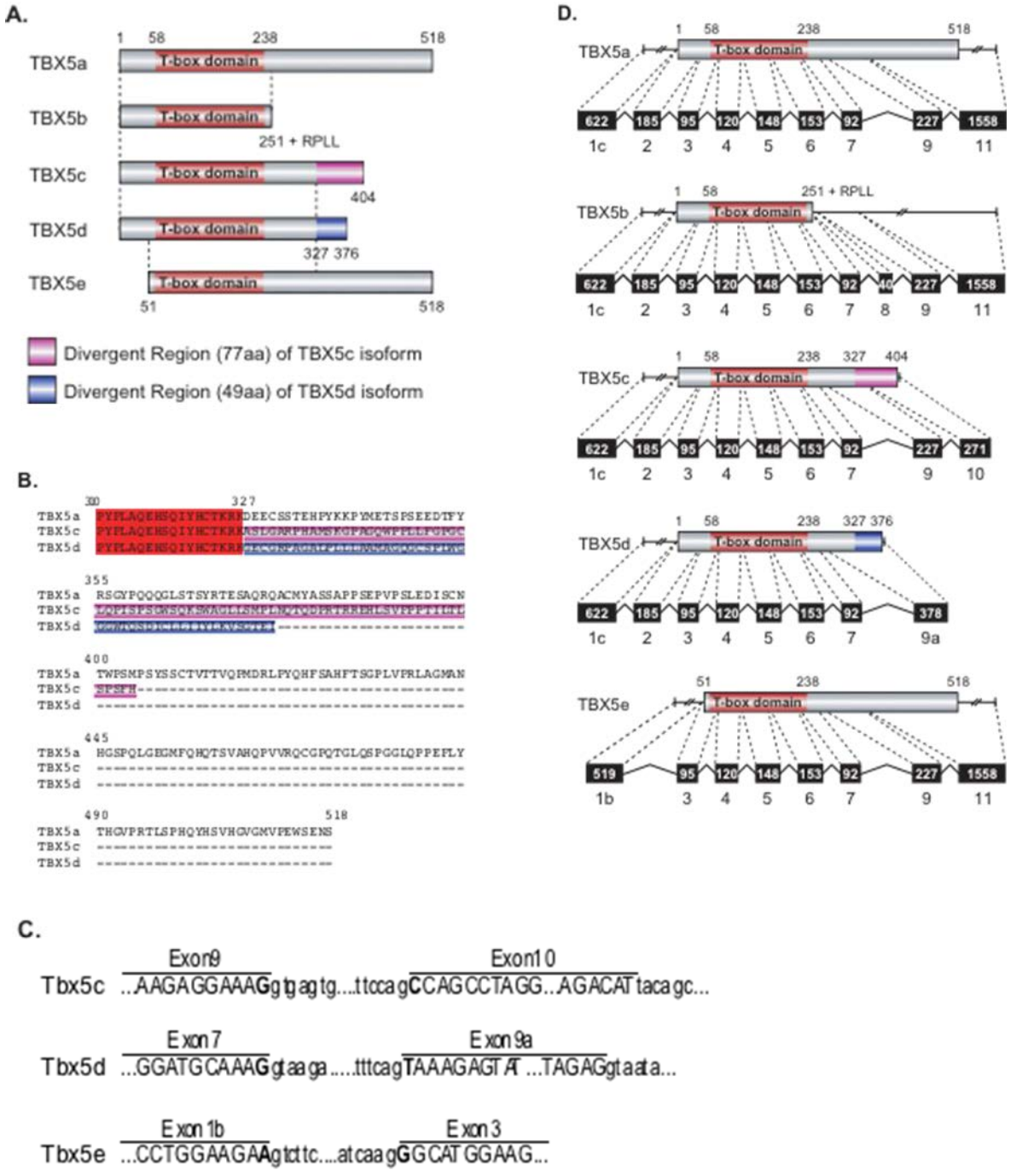


Figure 2.1 Representation of Mus musculus TBX5 isoforms. A. Schematic representation of Mus musculus Tbx5 isoforms. At least 5 different isoforms are translated from the different splice variants. The red box is the DNA-binding T-box domain of the protein. The purple and blue boxes are the regions that are divergent in the *Tbx5c* and *Tbx5d*, respectively. **B.** Multiple sequence alignment of

Tbx5a, *Tbx5c* and *Tbx5d* showing the divergence in the C-terminal region of the three *Mus musculus Tbx5* isoforms. **C.** Schematic representation of the genomic sequences showing the splice junctions. **D.** Schematic representation of the exons alternatively spliced to produce the different protein isoforms in mice.

Next, we analyzed the tissue distribution of *Tbx5c* and *5d*, which contain novel exons using droplet digital PCR. In case of *Tbx5a* and *Tbx5d,* a forward primer corresponding to sequences in the common exon 7 was used along with a reverse primer specific to exon 11 (*Tbx5a*) or exon 9a (*Tbx5d*). Forward and reverse primers were chosen in exon 10 in case of *Tbx5c*. Using this strategy, we detected all isoforms in e11.5 atria and forelimbs. E11.5 ventricles expressed mostly *Tbx5a* and lower levels of the other isoforms. *Tbx5a and Tbx5c* transcripts were also detected in the hindlimbs albeit at much lower levels (**Figure 2.2 A**). Similar expression pattern was seen in the postnatal heart where *Tbx5a* was always the predominant isoform (**Figure 2.2 B**). We also verified the presence of these isoforms in cardiogenic cell lines - notably the atrial like HL-1 cells and the endocardial progenitors TC13 cells - as well as in skeletal muscle progenitors like C2C12 myoblasts. The *Tbx5* expression pattern in HL1 resembled that of the e11.5 heart with the predominance of *Tbx5a* (**Figure 2.2 C**). *Tbx5a* was the only isoform present in TC13 cells but not in C2C12 myoblasts where *Tbx5c* was the predominant isoform and was dramatically upregulated after differentiation.

We then used the N-terminal TBX5 antibody (9) to analyze the profile of TBX5 proteins in cardiac cells. Western blot analysis of nuclear extracts from TC13 cells revealed the existence of multiple immunoreactive bands (**Figure 2.2 D**) co-migrating with TBX5a, 5b, 5c, 5d; TBX5a and 5c were also detected in C2C12 cells (**Figure 2.2 D**) although the TBX5a band was much weaker. To verify the presence of TBX5e, a similar blot was incubated with an antibody raised against a C- terminal epitope encoded by exon 11 (aa 331-425); this antibody can detect TBX5a and 5e but not 5b, 5c and 5d. Consistent with the digital PCR results, a

strong band co-migrating with TBX5a was detected in TC13 nuclear extracts as well as a weaker band co-migrating with TBX5e. A 50 kDa band that may be TBX5e was also present in C2C12 cells (**Figure 2.2 E**). Next, we analyzed TBX5 complexes in postnatal hearts. Nuclear extracts were obtained using 300mM NaCl extraction, dialyzed to 100mM salt and chromatin binding proteins were enriched on a phosphocellulose column; chromatin bound proteins were eluted at 300mM salt and size fractionated under non-denaturing conditions using a gel filtration Superose 6 column. Western blots on the resulting fractions were carried out using the N-terminal TBX5 antibody. Bands corresponding to TBX5a (~64-80 kDa) were evident in several high MW fractions. An immunoreactive band around 48 kDa was also present and would correspond to TBX5d (**Figure 2.2 F**). These results confirm the presence in the heart of TBX5a and 5d and suggest the existence of different protein complexes containing distinct TBX5 isoforms in Tbx5 expressing organs.

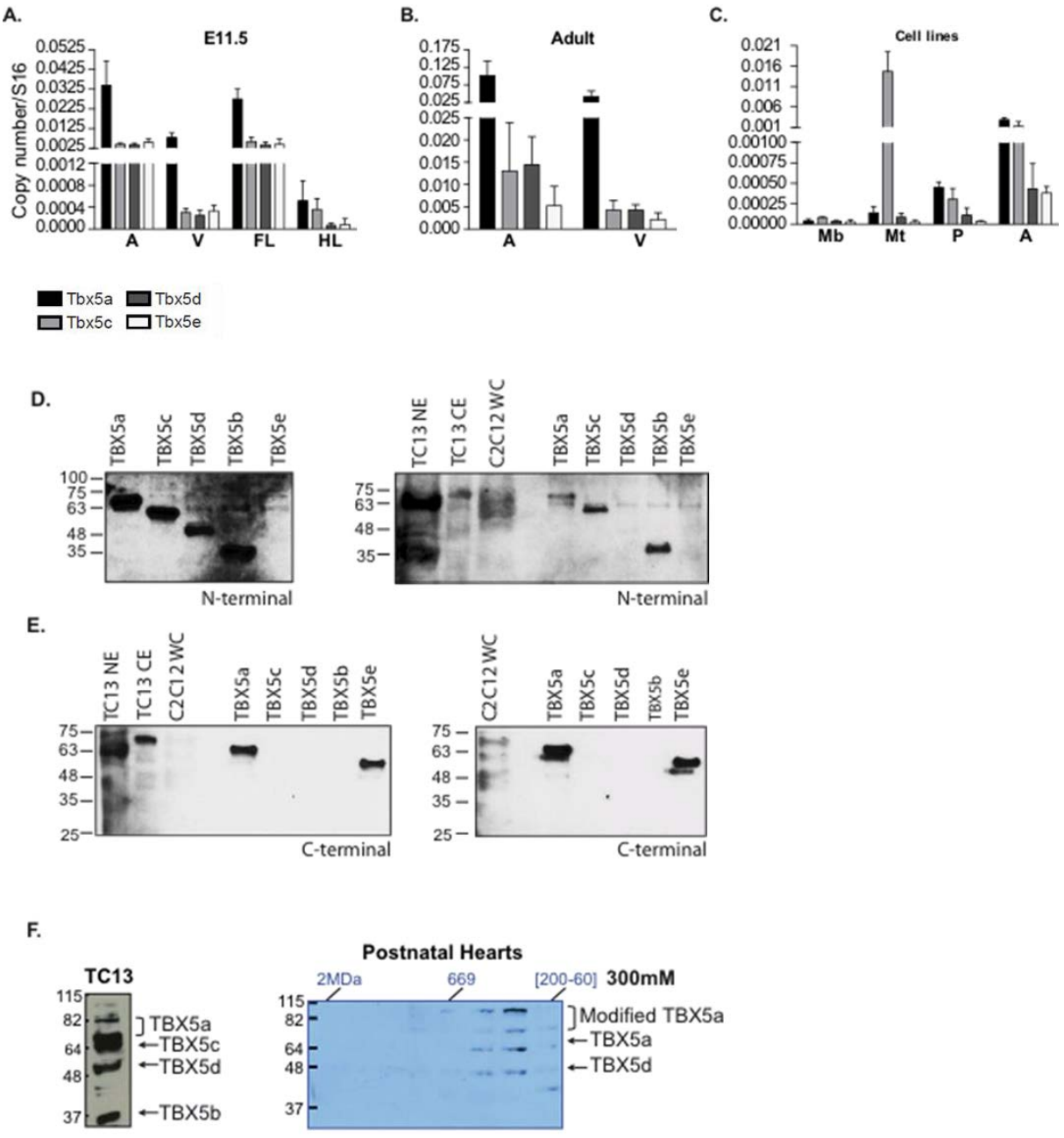


Figure 2.2 Detection of novel TBX5 isoforms in mouse tissues and cell lines. A. Differential expression of *Tbx5* isoforms in heart and limbs of mouse E11.5 embryos. Droplet Digital PCR performed using a forward primer in the common exon 7 and with a reverse primer specific to exon 11 (*Tbx5a*) or exon 9a (*Tbx5d*). Forward and reverse primers were chosen in exon 10 in case of *Tbx5c*. **B**. Same Droplet Digital PCR performed on mRNA from adult atria and ventricles. Note that *Tbx5a* is the predominant isoform in the mouse heart at all stages. **C.** Droplet Digital PCR on mRNA from the indicated cell lines. Note the presence of the four isoforms in HL1 cells (atrial like). *Tbx5a and Tbx5c* are the only isoforms in TC13 cells (endocardial progenitors). *Tbx5c* is the predominant isoform in C2C12 myoblasts and its expression is greatly increased in myotubes. Mb: myoblasts; Mt: myotubes; P: progenitors; A: atria; V: ventricular; FL: forelimb; HL: hindlimb. In all cases, RS16 was used as a

house-keeping gene. The values are a ratio of the gene copy number to RS16 copy number and are the average of biological replicates with standard deviation. **D-F: Western blot analysis of endogenous TBX5 in expressing cells. D.** Western blot using the N-terminal TBX5 antibody. The left panel shows the position of recombinant flag-TBX5 isoforms expressed in 293T cell. The right panel contains the same extracts as well as nuclear (NE) cytoplasmic (CE) or whole cell (WCE) extracts from TC13 and C2C12 cells. Note that this antibody does not detect TBX5e that lacks the N-terminal domain. **E.** A parallel blot incubated with the C-terminal TBX5a antibody. The right panel represents a longer exposure necessary to detect TBX5a and 5e in C2C12. Note the much higher abundance of TBX5a in TC13 cells. Note that this antibody detects only TBX5a and TBX5e that share the same C-terminal domain. **F**. TBX5 complexes in postnatal mouse heart. Nuclear extracts were obtained with 300 mM NaCl, Chromatin binding proteins were enriched with phosphocellulose IEX, and size fractionated over Superose 6 column as described in the text. Western blot analysis was then performed on the different fractions. On the left, unfractionated nuclear TC13 extracts prepared the same way.

Next, we wanted to verify whether similar splice isoforms exist in human. *In silico* analysis of the mouse and human *TBX5* genomic sequences revealed the presence of similar alternate splice site junctions for Tbx5c and 5d. For Tbx5d, this predicted the existence of alternate splicing within exon 9 to produce the splice variant shown in **Figure 2.3 A** which turned out to be identical to the reported transcribed sequence MN_181486 which would produce a 350AA isoform whose last 22AA are divergent from TBX5a (**Figure 2.3 B**). The presence of this cDNA sequence was also reported by Basson *et al.* (1) in the initial identification of TBX5 as the HOS gene. Interestingly, only the first 3 additional AA are conserved between mouse and human suggesting that protein conformation not primary sequence may need to be conserved. Another transcript that would encode the human equivalent of TBX5e can also be found in the Ensembl data base (NM080717). Based on mouse-human sequence alignment, we were able to predict the human sequence of *TBX5c* which seems to be more conserved among mouse and human than the TBX5d specific region (**Figure 2.3 C**). We designed primers specific for human *TBX5a*, *c* and *d* by targeting the unique exons of each isoform. RT-PCR analysis revealed the presence in human heart tissues of the expected size transcripts (**Figure 2.3 D**) indicating that similar isoforms are present in human and rodent hearts.

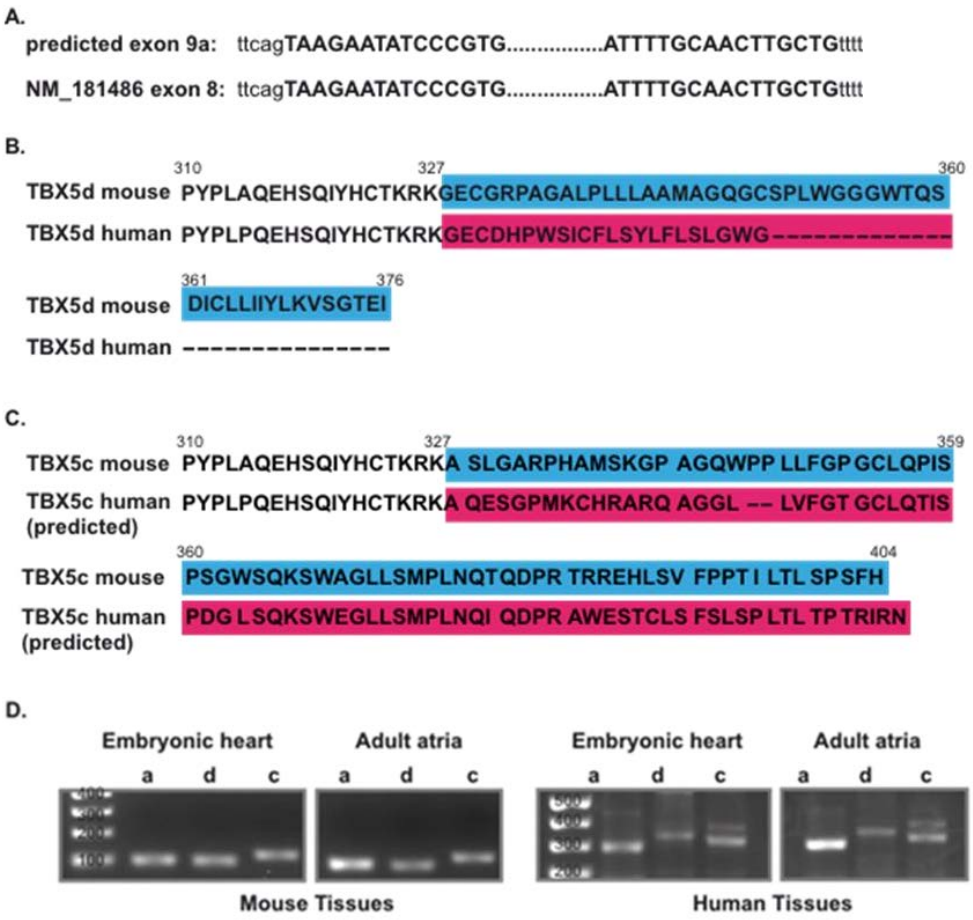


Figure 2.3 Identification of human TBX5 isoforms. A. Alignment of the predicted novel exon9b sequence with exon8 of NM_181486 sequence. Note that the two sequences are 100% identical. **B. & C.** Alignment of the variant regions of TBX5d and TBX5c of mouse (blue) and human (magenta) sequences, respectively. **D.** Detection of the novel exons in human hearts. RT-PCR was performed on human and mouse embryonic and atrial mRNA. Both the known *Tbx5* (*Tbx5a*) and the novel isoforms were detected.

Differential expression of distinct TBX5 isoforms

In order to study the expression pattern of the different isoforms at different mouse embryonic stages, we generated two new antibodies raised against the C-terminal epitopes of TBX5c (aa 327-404) and TBX5d (aa 327-376), respectively. The antibody specificity was confirmed by immunofluorescence done on C2C12 cells transfected with HA-TBX5a, HA-TBX5c or HA-TBX5d using HA antibody co-stained with C-terminal TBX5a, TBX5c or TBX5d antibodies (data not shown). Immunohistochemistry was then performed using the C-terminal antibodies

specific to TBX5a, TBX5c and TBX5d. The N-terminal directed antibody that detects all TBX5 isoforms containing this domain was also used on mouse embryos (E11.5, E15.5) and neonate (P1.5) tissue sections (**Figure 2.4 A** and **2.4 B**). All isoforms were detected in the myocardium as well as the endocardium of E11.5 mouse hearts. At E15.5, in addition to their expression in the atria, the isoforms were also detected in the atrioventricular valves where TBX5c and TBX5d where more predominant. TBX5a and 5c but not 5d were also detected in the forelimbs at E15.5. Expression of the isoforms in the AV valves persisted in the postnatal mouse heart (P1.5) where TBX5c and TBX5d continued to be the more abundant isoforms in the mitral and tricuspid valves. In contrast, TBX5a was the most abundant isoform in the atrioventricular bundle (AVB). The differential expression of TBX5 isoforms in the limbs and in different heart cells raises intriguing questions regarding their specific functions in organ development.

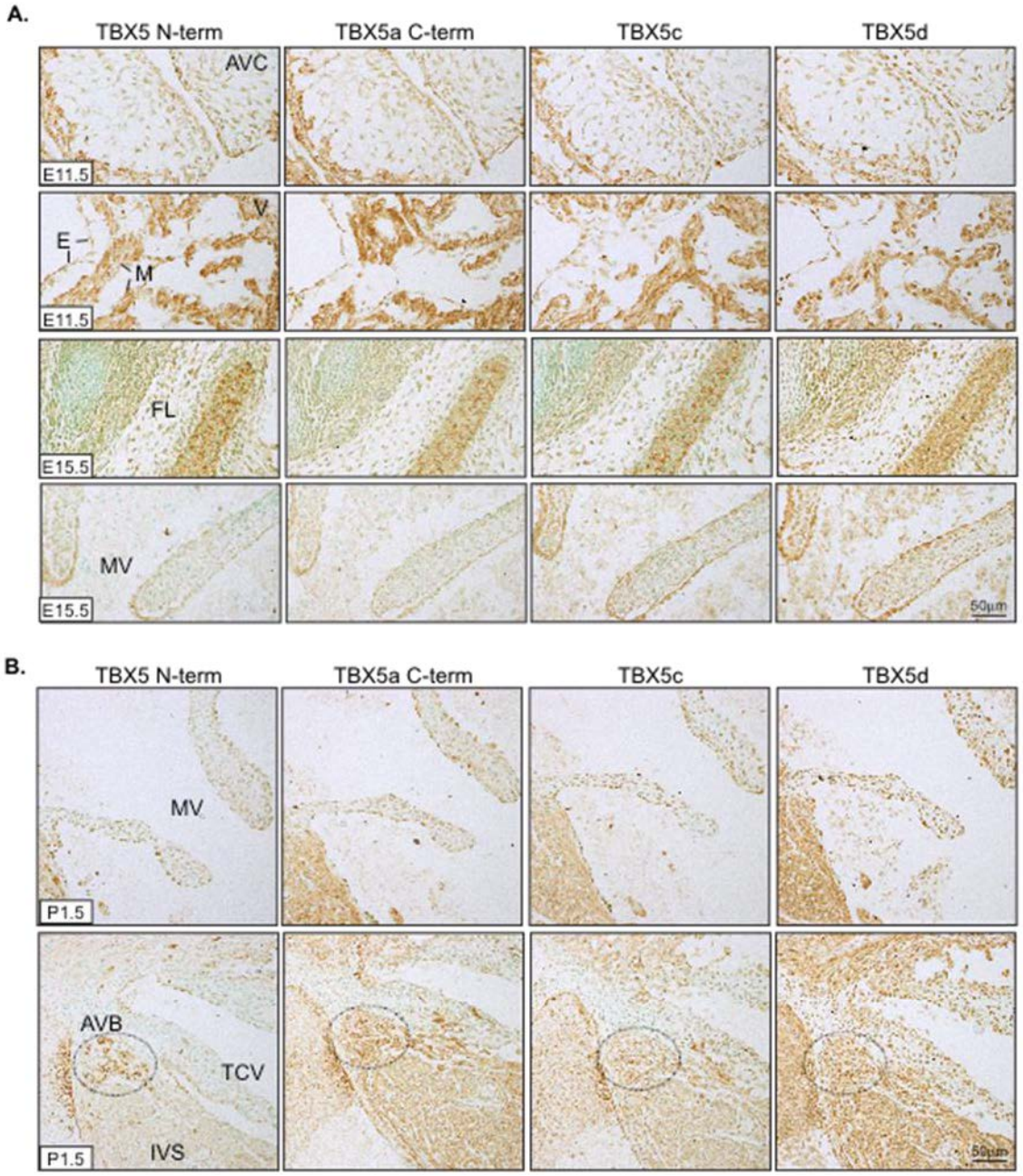


Figure 2.4 Differential expression of the different TBX5 isoforms in mouse tissues at different embryonic stages E11.5 and E15.5. (A) and in postnatal hearts P1.5 (B) using the specific TBX5 antibodies. Note the expression of the different isoforms in the myocardium and endocardium at E11.5. At E15.5, TBX5c and TBX5d expression was higher in the AV valves. TBX5a and 5c but not 5d staining was also seen in the forelimbs. At P1.5, expression of the isoforms persisted in the AV valves where TBX5c and 5d remained more abundant. Also note expression of the isoforms in the AVB where TBX5a is the most abundant isoform. AVC: atrioventricular cushion; E: endocardium; M: myocardium; V: ventricle; FL: forelimb; MV: mitral valve; RA: right atrium; AVB: atrioventricular bundle; IVS: interventricular septum; TCV: tricuspid valve.

TBX5 isoforms possess distinct biochemical properties

We investigated the biochemical properties of the newly identified TBX5 isoforms. Western blot analysis revealed that all N-terminally Flag-tagged isoforms were well expressed in transfected 293T cells (**Figure 2.5 A**). Immunocytochemistry revealed distinct subcellular localization with TBX5c and TBX5d detected in both the nucleus and the cytoplasm indicating that TBX5 sequences between AA327-518 promote nuclear localization (**Figure 2.5 B**). This is in line with the reported presence of a nuclear localization signal between AA 325-340 (9,26). Electrophoretic mobility shift assays (EMSA) revealed that all isoforms were able to bind the cognate TBE site of the *Nppa* promoter with good affinity (**Figure 2.5 C**). Because the various newly identified transcripts appear to be co-expressed with the long *Tbx5a,* we tested the consequences of their presence on TBX5a binding to its cognate site (**Figure 2.5 D**). When present in equimolar ratio with HA-TBX5a, all isoforms tested (Flag-TBX5c,d and e) appeared to form heterodimers (the lower asterix corresponding to the isoform homodimer and the upper asterix to the heterodimer) that were supershifted with both the Flag and HA antibodies (last 2 right lanes of each panel). This result is in line with the fact that homo and hetero-dimerization of TBX proteins is mediated by the T-box, which is identical for all TBX5 isoforms tested.

We also tested the ability of the various isoforms to form a ternary complex with NKX2.5 on the composite TBE-NKE site. As shown in **Figure 2.5 E**, when co-expressed with NKX2.5, TBX5a, c, d and e efficiently formed a slower migrating ternary complex (indicated by asterix). The protein composition of the complex was confirmed by the addition of TBX5 (N-terminal) and NKX2.5 antibodies which both resulted in supershifting the ternary complex.

The TBX5 heterodimers as well as the TBX5-NKX2.5 complexes were also confirmed by *in vitro* pull down assay as shown in **Figure 2.5 F**. Interestingly, TBX5c and 5d as well as TBX5e bound NKX2.5 and TBX5a with a higher affinity than TBX5a; this raises the possibility that they can compete for TBX5a binding to NKX2.5 and confirms their ability to form DNA binding heterodimers with TBX5a.

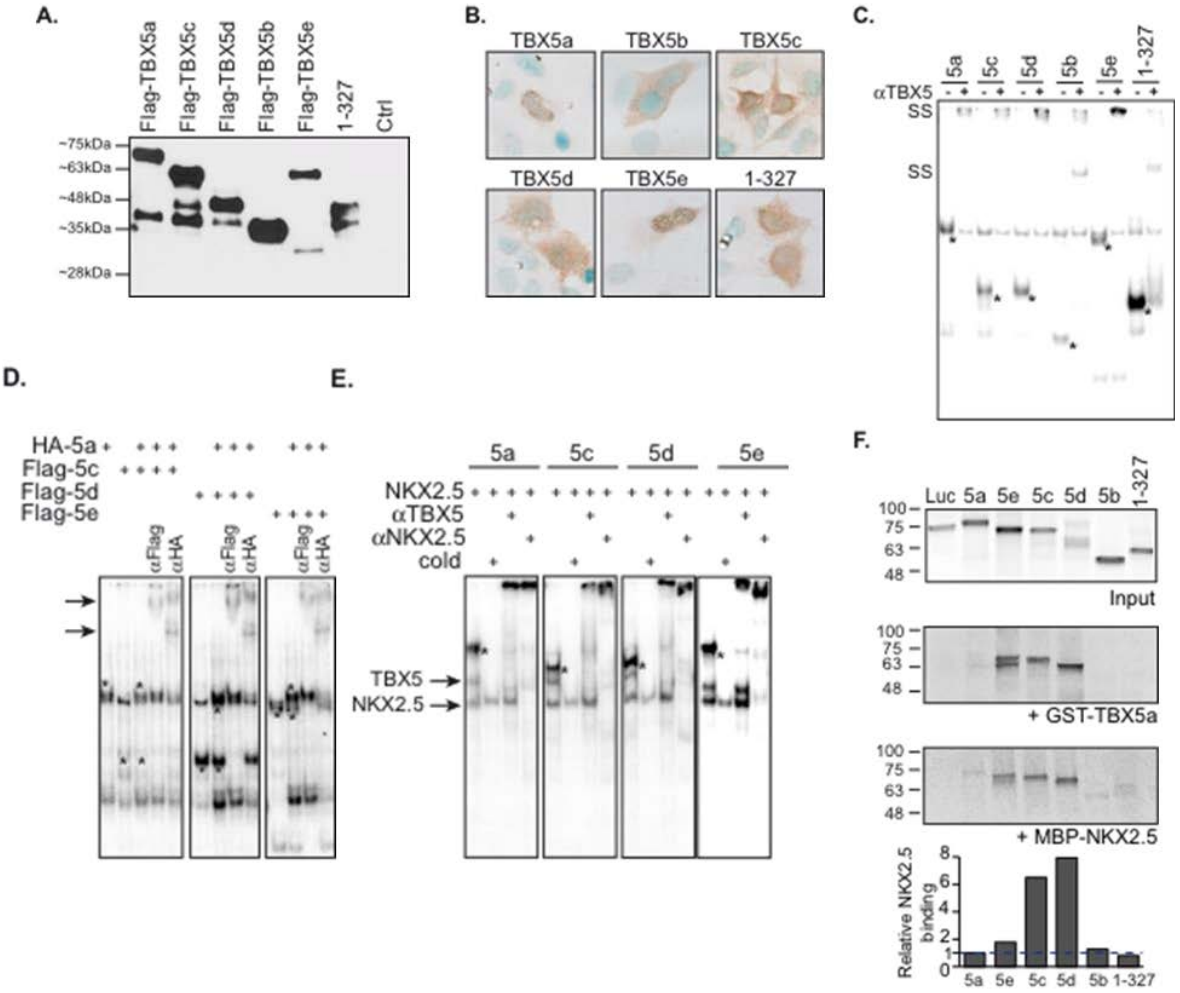


Figure 2.5 TBX5 isoforms possess distinct biochemical properties. A. Expression of recombinant TBX5 isoforms in whole cell extracts using western blot analysis of 293T cells overexpressing the indicated TBX5 isoform and the Flag-M2 antibody. Note how all isoforms have similar expression levels. **B.** Immunocytochemical localization of the same TBX5 isoforms expressed in 293T cells using the Flag-M2 antibody. Note the complete nuclear localization of TBX5a and TBX5e, the mostly cytoplasmic localization of TBX5b and the nuclear and cytoplasmic distribution of TBX5c and TBX5d. **C.** DNA binding properties of TBX5 isoforms. Electrophoretic mobility shift assay was performed using nuclear extracts from 293T cells overexpressing the indicated TBX5 isoform. The exogenous TBX5 binding is indicated by an asterix. Supershift/blocking was done with

homemade anti-TBX5 antibody (N-terminal TBX5 in case of TBX5a,c,d,b and 1-327; C-terminal TBX5 in case of TBX5e) SS: supershift.. **D.** Ability of novel TBX5 isoforms to form heterodimers with TBX5a. EMSA was performed as above. 1:1 ratio of HA-TBX5a and the indicated Flag-TBX5 isoform were used as indicated. The asterix marks the binding of the individual isoform or the complex when applicable. Note that all isoforms are capable of forming a heterodimer with TBX5a that is supershifted with both Flag and HA antibodies. **E.** Formation of a TBX5/NKX2.5 complex. EMSA was carried out as above. The asterix marks the binding of the complex. Supershift was done with anti-TBX5 or anti-NKX2.5 antibody. 100X self-cold probe was used as a competitor. **F.** *In vitro* translated radiolabeled TBX5 proteins (or luciferase protein as a negative control) were incubated with glutathione Sepharose beads containing GST-TBX5a fusion protein (middle panel) or amylose beads containing MBP-NKX2.5 fusion protein (lower panel). Note that the isoforms form heterodimers with TBX5a and ternary complex with NKX2.5. The autoradiograph shown is representative of two independent experiments. Quantitation of the NKX2.5-TBX5 binding relative to the NKX2.5-TBX5a binding is shown in the bottom graph. Values are corrected over the corresponding TBX5 inputs.

Next, we tested their transcriptional effects on target promoters. As stated earlier and schematically shown in **Figure 2.6 A**, TBX5a harbors three transactivation domains, one in the N-terminal (TAD1) and two in the C-terminal (TAD2 and TAD3). TAD1 is absent in TBX5e whereas TBX5c/d lack TAD3 (**Figure 2.6 A**). Co- transfection experiments were carried out using three target promoters: *Nppa (*a differentiation marker (3), *BCLX* (a survival marker (18), and Ccnd1 (cyclin D1, a proliferation marker). The results revealed unexpected promoter specific differences in the transcriptional activity of TBX5 isoforms. Compared to TBX5a, TBX5e and TBX5a$_{1\text{-}327}$ were weaker activators of the *Nppa* promoter; modest, if any, activation was observed with TBX5c and TBX5d on this promoter. These results are consistent with our previous structure-function study (9) and suggest that maximal *Nppa* activation requires all three TADs; they also suggest that the additional C-terminal sequences do not contribute activation functions (**Figure 2.6 B**, left panel). When tested on the *BCLX* promoter, the activity of TBX5a and 5e were almost similar; TBX5c was able to activate transcription albeit to a lower extent whereas TBX5d was inactive. This result suggests a strong dependence on the C-terminal TADs and reveals distinct behaviors for TBX5c and 5d. (**Figure 2.6 B** middle panel). In contrast, activation of the Ccnd1 promoter was independent of TAD3 as

evidenced by the ability of TBX5a$_{1-327}$ to maximally activate the promoter; however, absence of the N-terminal domain reduced activation by half (**Figure 2.6 B**, right panel). Interestingly, as in the case of *Nppa*, TBX5c and 5d were less active than TBX5a1-327 suggesting that the additional C-terminal domain interferes with existing TADs.

Next we tested the consequences of co-expressing the new isoforms on TBX5a transcriptional activity. Co-transfections of the *Nppa*-luciferase reporter with TBX5a was carried out in the presence of increasing concentrations of the TBX5c, 5d or 5e; the concentrations used resulting in a 1:1, 3:1 and 10:1 ratio of isoform:TBX5a. As shown in **Figure 2.6 C**, all isoforms resulted in a dose dependent attenuation of TBX5a activation. These results, together with the gel shifts data in **Figure 2.5 C and D** likely reflect the presence on the promoter of heterodimers with weaker activating properties than a TBX5a homodimers. Next, we tested the transcriptional properties of the TBX5-NKX2.5 complex. Neither TBX5c nor TBX5d was able to synergize with NKX2.5 (**Figure 2.6 D**) whereas TBX5e was able to cooperate with NKX2.5 to a nearly similar extent as TBX5a. Given that all isoforms formed stable ternary complexes with NKX2.5, these results suggest that some TBX5 isoforms might attenuate NKX2.5 function on target promoters. A similar profile was observed when the isoforms were tested with GATA4: TBX5e retained the ability to enhance GATA4 activation but to a lower level than TBX5a, TBX5c showed modest activation and TBX5d was unable to support synergy (**Figure 2.6 E**). These results are consistent with our previous findings that the T-box is essential for GATA4-TBX5 interaction and that sequences in the N- and C-terminal contribute to the overall synergy (9). To test if the TBX5c and TBX5d specific domains have a transcriptional activity on their own, reporter gene transactivation assays were performed with expression plasmids encoding fusion proteins of these domains with the DNA-binding

domain of Gal4 (Gal4-DBD). As shown in **Figure 2.6 F**, TBX5a TAD3 is able to activate transcription on its own (10X). This activity is greatly enhanced in presence of TAD2 (30X) although TAD2 has minimal activity on its own (2.2X). Neither TBX5c nor TBX5d specific domains seemed to possess autonomous transcriptional activity. Moreover, they had no effect (positive or negative) on TAD2 transcriptional activity (**Figure 2.6 F**).

Together the results reveal that the different TBX5 isoforms have distinct transcriptional properties.

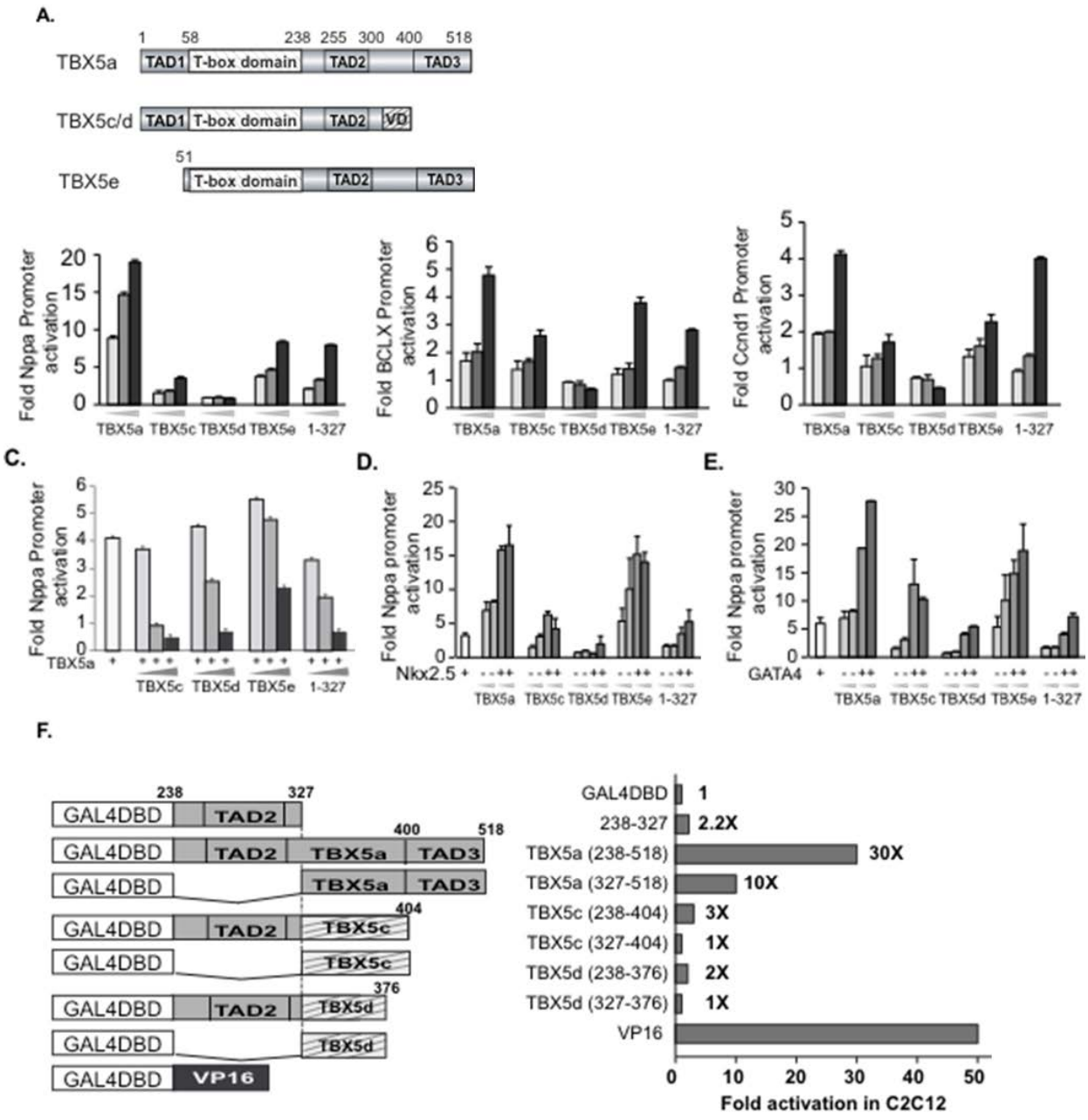


Figure 2.6 Differential transcriptional activities of the TBX5 isoforms. A. Schematic representation of the different TBX5 isoforms with the various transactivation (TAD) domains indicated. Note that TBX5c and TBX5d lack TAD3; TBX5e lacks TAD1. **B.** Transcriptional activity of TBX5 isoforms was determined in NIH3T3 cells using three different TBX5 target reporters and increasing doses of the indicated TBX5 proteins. **C.** Effect of coexpression of the novel isoforms on TBX5a activity. The 3 bars in each case represent increasing ratios of the relevant isoform vs TBX5a as follows: 1:1, 3:1, 10:1. **D&E.** Synergistic activity of the various TBX5 isoforms with NKX2.5 (**D**) or GATA4 (**E**) on the Nppa promoter. Note the weak/absent synergy of TBX5c and TBX5d with NKX2.5 and GATA4. The results shown are from one representative experiment out of at least three carried out in duplicates. **F.** The variant regions of TBX5c and TBX5d do not possess autonomous

transcriptional activity. C2C12 cells were transfected with the indicated GAL4-DBD fusion protein together with $(UAS)_5$-TK reporter plasmid. Note that TBX5c and TBX5d fusion proteins do not activate transcription (TBX5c 327-404 and TBX5d 327-376) nor do they repress it (TBX5c 238-404 and TBX5d 238-376 compared to 238-327). The results are one representative experiment of three.

Effect of TBX5 isoforms on skeletal muscle differentiation

As stated above, TBX5c is the predominant isoform in skeletal myoblasts and its expression is greatly increased in differentiated myotubes. In order to determine the effect of TBX5c in these cells, we carried out gain and loss of function studies. First, we transfected myoblasts with TBX5a or TBX5c and monitored their fate in quiescent (1% Bovine Serum Albumin) or differentiating media (2% Horse Serum). As shown in **Figure 2.7 A**, TBX5c but not TBX5a was able to induce differentiation in quiescent cells as evident by the myogenic markers Myosin Heavy Chain (MF20) (left panel) and Myogenin (right panel). TBX5c was able to accelerate myotube formation when cells were switched to the differentiation medium, in contrast to TBX5a which inhibited differentiation (**Figure 2.7 B**).

We then tested the effect of downregulation of Tbx5a or Tbx5c on myocyte differentiation. C2C12 myoblasts were transfected with a universal control SiRNA and with 2 different Tbx5a specific or Tbx5c specific SiRNAs. Cells were examined after 2 and 4 days in differentiation medium. Down regulation of Tbx5a was able to remove the inhibitory effect of TBX5a on myotube formation at D2 and D4 whereas down regulation of Tbx5c but not Tbx5a reduced myotube formation (**Figure 2.8 A** and **2.8 B**) indicating that TBX5c plays an important role in skeletal muscle differentiation.

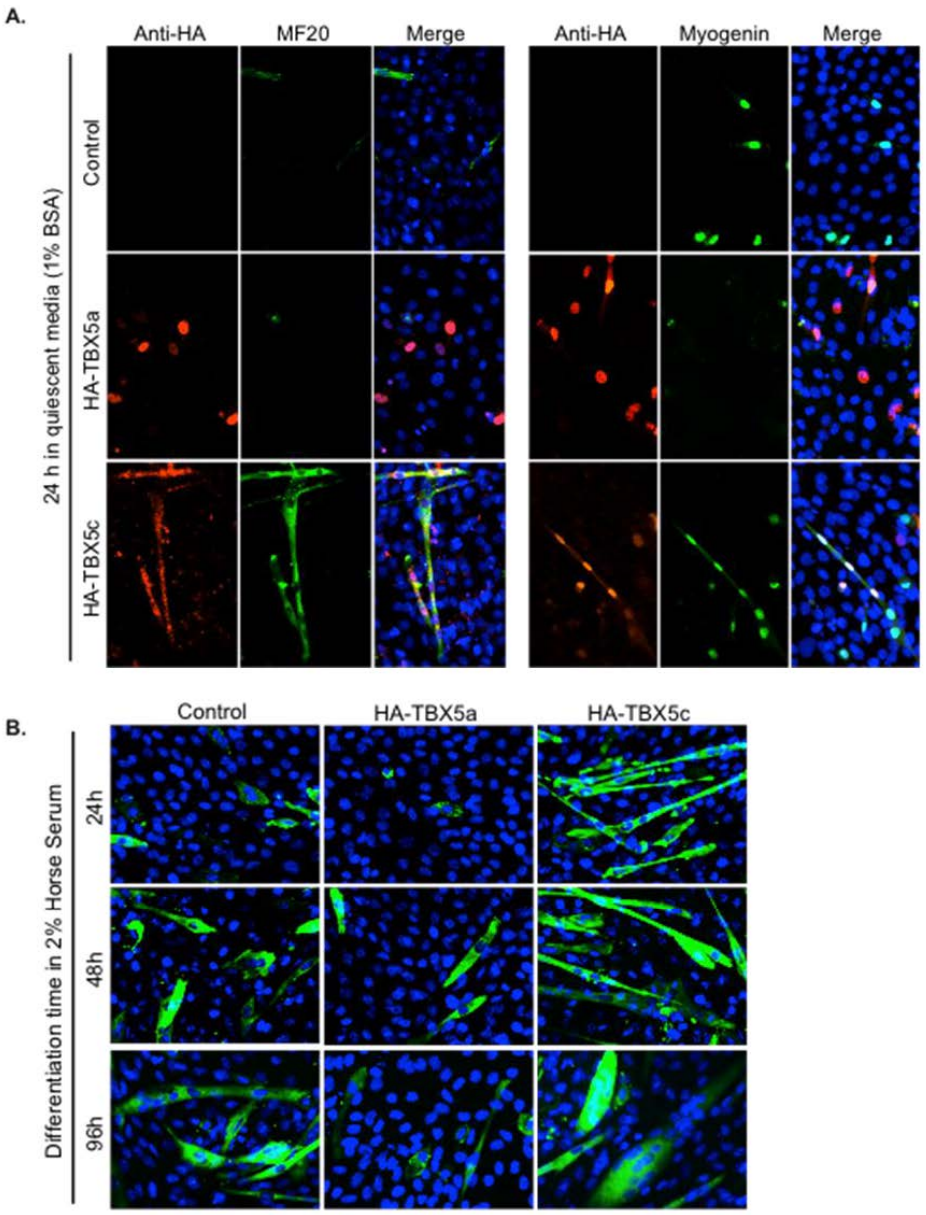


Figure 2.7 TBX5c enhances myocyte differentiation. C2C12 myoblasts were transfected with HA-TBX5a or HA-TBX5c (or pCGN as a negative control) and incubated for (A) 24h in quiescent media (1% BSA) or (B) 24h, 48h, or 96h in differentiating media (2% horse serum). MF20 and F5D (anti-Myogenin) are in green. Anti-HA is labelled red and Dapi in blue. Note that TBX5c is able to induce myotube formation in quiescent and differentiation media where as TBX5a inhibits myotube formation when cells are incubated in differentiation media. MF20 staining was used in B. BSA: bovine serum albumin; HS: horse serum.

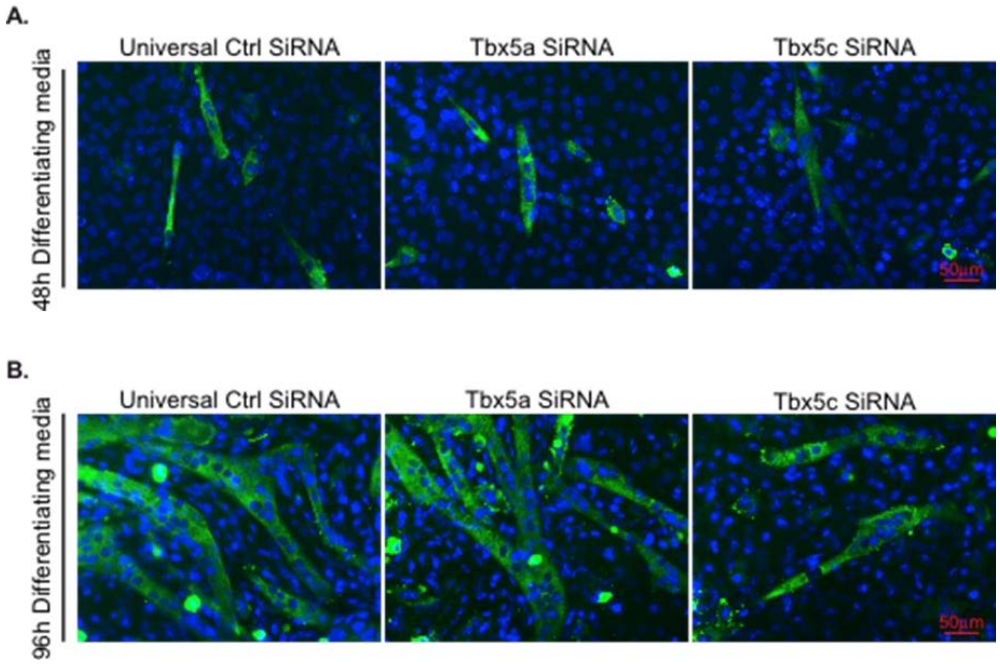


Figure 2.8 TBX5c downregulation inhibits myotube formation. C2C12 myoblasts were transfected with a universal control SiRNA or SiRNA specific to Tbx5a or Tbx5c and incubated in 2%HS differentiating media for 48h (A) or 96h (B). Note that downregulation of Tbx5a allows myotube formation whereas downregulation of Tbx5c greatly reduces myotube formation. The results are one representative experiment of two.

To get insight into the mechanisms underlying TBX5a and TBX5c effects on muscle progenitors, we monitored changes in expression of several genes linked to myocyte proliferation or differentiation using QPCR analysis. As shown in **Figure 2.9 A**, TBX5a induces *Igf1* and *Bmp4* expression whereas TBX5c inhibits *Igf1* and *Fgf10* transcript levels. *Igf1* is involved in C2C12 proliferation and differentiation (27). *Bmp4* was previously shown to inhibit C2C12 myotube formation (28). TBX5a was previously shown to activate *Fgf10* promoter (21) and its induction is important for myocyte proliferation (29,30). As previously reported, TBX5a- but not TBX5c-activated transcription from the *Fgf10* promoter (21).

TBX5c inhibited TBX5a activation of both Fgf10 and Bmp4 promoters. (**Figure 2.9 B and 2.9 C**).

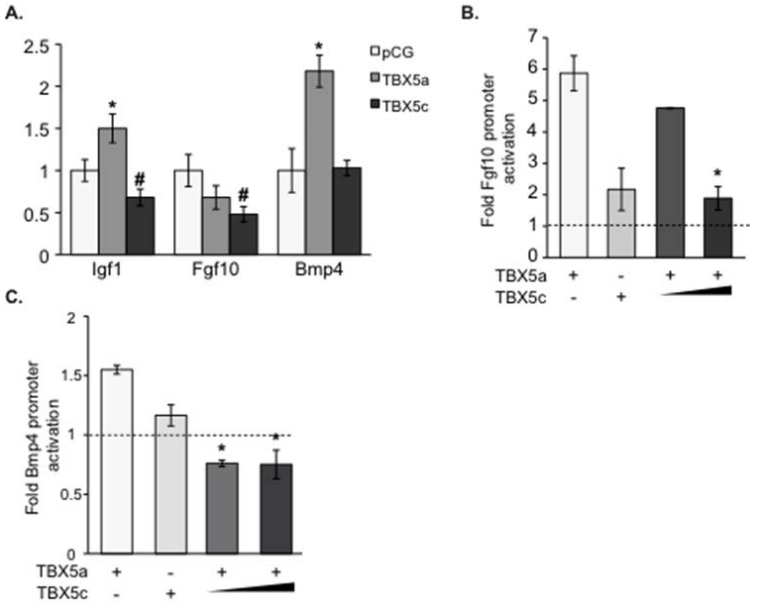


Figure 2.9 Mechanisms underlying TBX5a and TBX5c effects on muscle progenitors. A. QPCR analysis on C2C12 myoblasts transfected with pCG, TBX5a or TBX5c expression vectors. Note that TBX5a increases *Igf1* and *Bmp4* transcript levels where as TBX5c inhibits *Igf1* and *Fgf10* expression. * $p < 0.05$ TBX5a vs. pCG; # $p < 0.05$ TBX5c vs. pCG. **B. & C.** NIH3T3 cells transfected with TBX5a, TBX5c or TBX5a with increasing doses of TBX5c and *Fgf10* (B) or *Bmp4* (C) reporter constructs. Note that TBX5c inhibits TBX5a activation of both promoters. * $p < 0.0005$ vs. TBX5a.

These results support a role of *Tbx5c* in myocyte differentiation and indicate that *Tbx5a* and *Tbx5c* have opposing functions on skeletal muscle progenitor expansion and differentiation.

2.10 Discussion

TBX5 is a dosage-sensitive regulator of heart and limb development (31) and is the causative gene of Holt-Oram syndrome (HOS). Little is known about the mechanisms regulating TBX5 levels and activity. In this report, we provide further evidence that *Tbx5* is highly regulated through alternative splicing and identify the existence of novel exons including two that result in TBX5 proteins containing new C-terminal domains. All isoforms retain the ability to bind DNA as homodimers, as well as heterodimers with TBX5a or NKX2.5. Nonetheless, they display distinctive transcriptional properties that can be target gene specific. Using isoform specific antibodies, we confirm the in vivo presence of TBX5c and 5d isoforms in developing mouse heart and forelimbs where they display both overlapping but also specific expression domains relative to TBX5a. TBX5c which is the dominant isoform in skeletal muscle progenitors, enhances their differentiation and its downregulation blocks myotube formation. The data provide for the first time insight into TBX5 protein distribution in Tbx5 expressing cells and uncover a critical role for a novel isoform in muscle differentiation.

In addition to helping understand the role of TBX5 in health and disease, the data may be significant for elucidating genotype-phenotype correlations in patients with mutations in the *TBX5* gene.

Genotype-phenotype correlations are a major clinical challenge in congenital heart disease in general, and in HOS patients in particular mainly due to the wide range of mutations and the variability in disease-expressivity within family members. As shown in **Figure 2.10**, many of the known mutations in the N- and C-terminal region of TBX5 would affect TBX5a but not other isoforms; for example, mutations in Exon2 do not alter TBX5e while mutations in Exon11 affect TBX5e but not TBX5c or 5d. This is noteworthy given the findings presented

in this paper which suggest that subgroups of TBX5 target genes will be differentially impacted by such mutations. Moreover, decreased TBX5a activity resulting from mutations in Exon11 which harbor TAD3 (and is specific to TBX5a) might not only decrease progenitor expansion but might favor premature differentiation given the altered ratio of TBX5a and TBX5c activity in these cells. Coordinated cell proliferation, apoptosis and differentiation is critical for heart and limb morphogenesis. Through alternative splicing that generate protein isoforms that can promote or antagonize each one of the processes, Tbx5 might play a distinctive role at various stages of heart and limb development. Such paradigm may help explain the prevalence of certain phenotypes in patients harboring specific *TBX5* mutations and deserves to be further explored. Lastly, for nearly 30% of HOS patients, no mutation has been found in the known *TBX5* exons. Our results provide additional sequences within the *TBX5* locus for mutational screening of patients with HOS and other heart and/or limb abnormalities. Moreover, intragenic mutations-such as those in intron 2,7,8,9- whose mechanism of pathogenesis is still uncertain may be affecting expression of the isoforms reported in this study. Regulation of *TBX5* alternative splicing would affect the level at which specific isoforms are produced and could contribute to the variable phenotype expressivity. Consequently, our findings provide novel insights that will help elucidate the mechanisms of action of TBX5 and shed more light on genotype-phenotype correlations in HOS and possibly other human congenital heart disease involving *TBX5*.

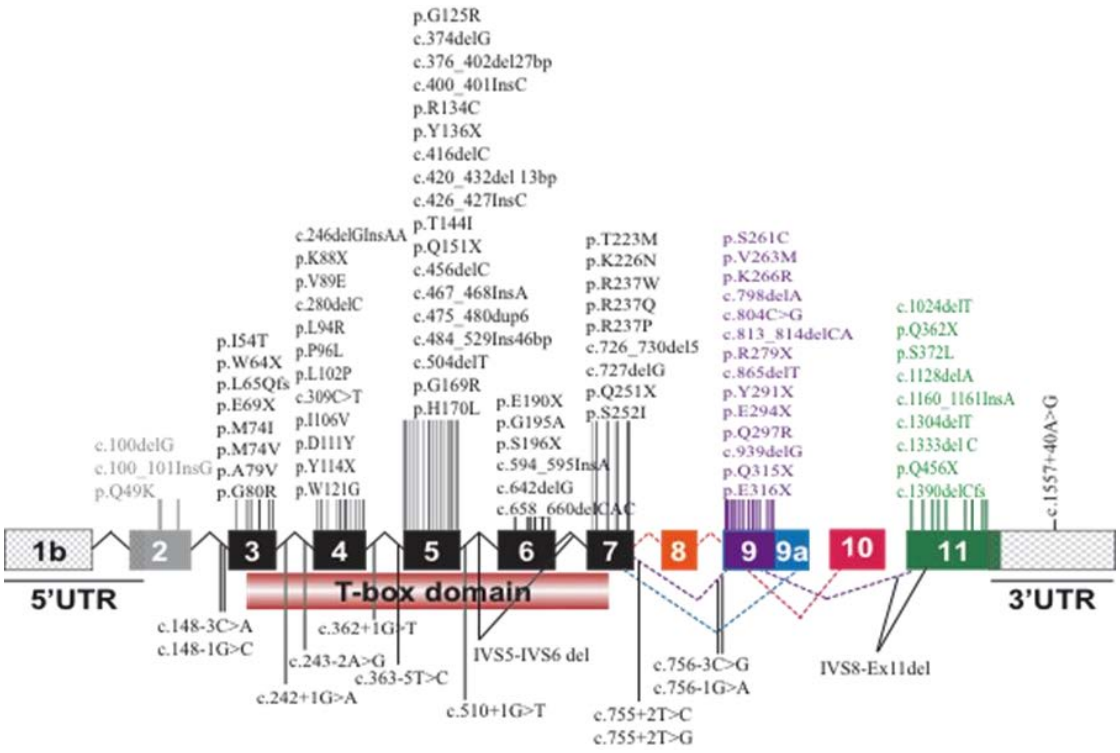


Figure 2.10 Representation of the *TBX5* gene showing the known as well as the newly identified exons with the intronic regions. Non-sense and missense mutations are represented at the protein level (p.) while frameshift and splice mutations are represented at the cDNA level (c.). Note that the mutations would affect only the isoforms containing the color-matching exons.

2.11 References

1. Basson, C. T., Bachinsky, D. R., Lin, R. C., Levi, T., Elkins, J. A., Soults, J., Grayzel, D., Kroumpouzou, E., Traill, T. A., Leblanc-Straceski, J., et al. (1997) Mutations in human TBX5 [corrected] cause limb and cardiac malformation in Holt-Oram syndrome. *Nat.Genet.*, **15**, 30-35.

2. Li, Q. Y., Newbury-Ecob, R. A., Terrett, J. A., Wilson, D. I., Curtis, A. R., Yi, C. H., Gebuhr, T., Bullen, P. J., Robson, S. C., Strachan, T., et al. (1997) Holt-Oram syndrome is caused by mutations in TBX5, a member of the Brachyury (T) gene family. *Nat.Genet.*, **15**, 21-29.

3. Bruneau, B. G., Nemer, G., Schmitt, J.P., Charron, F., Robitaille, L., Caron, S., Conner, D. A., Gessler, M., Nemer, M., et al. (2001) A murine model of Holt-Oram syndrome defines roles of the T-box transcription factor Tbx5 in cardiogenesis and disease. *Cell*, **106**, 709-721.

4. Heinritz, W., Shou, L., Moschik, A., and Froster, U. G. (2005) The human TBX5 gene mutation database. *Hum.Mutat.*, **26**, 397.

5. Fan, C., Liu, M., and Wang, Q. (7-3-2003) Functional analysis of TBX5 missense mutations associated with Holt-Oram syndrome. *J.Biol.Chem.*, **278**, 8780-8785.

6. Garg, V., Kathiriya, I. S., Barnes, R., Schluterman, M. K., King, I. N., Butler, C. A., Rothrock, C. R., Eapen, R. S., Hirayama-Yamada, K., Joo, K., et al. (24-7-2003) GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. *Nature*, **424**, 443-447.

7. Ghosh, T. K., Packham, E. A., Bonser, A. J., Robinson, T. E., Cross, S. J., and Brook, J. D. (1-9-2001) Characterization of the TBX5 binding site and analysis of mutations that cause Holt-Oram syndrome. *Hum.Mol.Genet.*, **10**, 1983-1994.

8. Borozdin, W., Bravo Ferrer Acosta, A. M., Bamshad, M. J., Botzenhart, E. M., Froster, U. G., Lemke, J., Schinzel, A., Spranger, S., McGaughran, J., Wand, D., et al. (2006) Expanding the spectrum of TBX5 mutations in Holt-Oram syndrome: detection of two intragenic deletions by quantitative real time PCR, and report of eight novel point mutations. *Hum.Mutat.*, **27**, 975-976.

9. Georges, R., Nemer, G., Morin, M., Lefebvre, C., and Nemer, M. (2008) Distinct expression and function of alternatively spliced Tbx5 isoforms in cell growth and differentiation. *Mol.Cell.Biol.*, **28**, 4052-4067.

10. Hariri, F., Nemer, M., and Nemer, G. (2012) T-box factors: insights into the evolutionary emergence of the complex heart. *Ann.Med.*, **44**, 680-693.

11. Hiroi, Y., Kudoh, S., Monzen, K., Ikeda, Y., Yazaki, Y., Nagai, R., and Komuro, I. (2001) Tbx5 associates with Nkx2-5 and synergistically promotes cardiomyocyte differentiation. *Nat.Genet.*, **28**, 276-280.

12. Camarata, T., Bimber, B., Kulisz, A., Chew, T. L., Yeung, J., and Simon, H. G. (31-7-2006) LMP4 regulates Tbx5 protein subcellular localization and activity. *J Cell Biol*, **174**, 339-348.

13. Zaragoza, M. V., Lewis, L. E., Sun, G., Wang, E., Li, L., Said-Salman, I., Feucht, L., and Huang, T. (14-4-2004) Identification of the TBX5 transactivating domain and the nuclear localization signal. *Gene*, **330**, 9-18.

14. Kulisz, A. and Simon, H. G. (2008) An evolutionarily conserved nuclear export signal facilitates cytoplasmic localization of the Tbx5 transcription factor. *Mol.Cell Biol.*, **28**, 1553-1564.

15. Stirnimann, C. U., Ptchelkine, D., Grimm, C., and Muller, C. W. (2-7-2010) Structural basis of TBX5-DNA recognition: the T-box domain in its DNA-bound and -unbound form. *J Mol Biol.*, **400**, 71-81.

16. Plageman, T. F., Jr. and Yutzey, K. E. (30-4-2004) Differential expression and function of Tbx5 and Tbx20 in cardiac development. *J.Biol.Chem.*, **279**, 19026-19034.

17. Durocher, D., Charron, F., Warren, R., Schwartz, R. J., and Nemer, M. (1997) The cardiac transcription factors Nkx2-5 and GATA-4 are mutual cofactors. *EMBO J.*, **16**, 5687-5696.

18. Nadeau, M., Georges, R. O., Laforest, B., Yamak, A., Lefebvre, C., Beauregard, J., Paradis, P., Bruneau, B. G., Andelfinger, G., and Nemer, M. (9-11-2010) An endocardial pathway involving Tbx5, Gata4, and Nos3 required for atrial septum formation. *Proc.Natl.Acad.Sci.U.S.A*, **107**, 19356-19361.

19. Lavallee, G., Andelfinger, G., Nadeau, M., Lefebvre, C., Nemer, G., Horb, M. E., and Nemer, M. (1-11-2006) The Kruppel-like transcription factor KLF13 is a novel regulator of heart development. *EMBO J*, **25**, 5201-5213.

20. Aries, A., Paradis, P., Lefebvre, C., Schwartz, R. J., and Nemer, M. (2004) Essential role of GATA-4 in cell survival and drug-induced cardiotoxicity. *Proc.Natl.Acad.Sci.USA*, **101**, 6975-6980.

21. Agarwal, P., Wylie, J. N., Galceran, J., Arkhitko, O., Li, C., Deng, C., Grosschedl, R., and Bruneau, B. G. (2003) Tbx5 is essential for forelimb bud initiation following patterning of the limb field in the mouse embryo. *Development*, **130**, 623-633.

22. Nemer, G. and Nemer, M. (1-10-2003) Transcriptional activation of BMP-4 and regulation of mammalian organogenesis by GATA-4 and -6. *Dev.Biol.*, **254**, 131-148.

23. Charron, F., Tsimiklis, G., Arcand, M., Robitaille, L., Liang, Q., Molkentin, J. D., Meloche, S., and Nemer, M. (2001) Tissue-specific GATA factors are transcriptional effectors of the small GTPase RhoA. *Genes Dev.*, **15**, 2702-2719.

24. Yamak, A., Latinkic, B. V., Dali, R., Temsah, R., and Nemer, M. (28-1-2014) Cyclin D2 is a GATA4 cofactor in cardiogenesis. *Proc Natl.Acad Sci U.S A*, **111**, 1415-1420.

25. Yamak, A., Temsah, R., Maharsy, W., Caron, S., Paradis, P., Aries, A., and Nemer, M. (15-10-2012) Cyclin D2 rescues size and function of GATA4 haplo-insufficient hearts. *Am.J.Physiol Heart Circ.Physiol*, **303**, H1057-H1066.

26. Collavoli, A., Hatcher, C. J., He, J., Okin, D., Deo, R., and Basson, C. T. (2003) TBX5 nuclear localization is mediated by dual cooperative intramolecular signals. *J.Mol.Cell.Cardiol.*, **35**, 1191-1195.

27. Sato, M., Ito, A., Kawabe, Y., Nagamori, E., and Kamihira, M. (2011) Enhanced contractile force generation by artificial skeletal muscle tissues using IGF-I gene-engineered myoblast cells. *J Biosci.Bioeng.*, **112**, 273-278.

28. Dahlqvist, C., Blokzijl, A., Chapman, G., Falk, A., Dannaeus, K., Ibanez, C. F., and Lendahl, U. (2003) Functional Notch signaling is required for BMP4-induced inhibition of myogenic differentiation. *Development*, **130**, 6089-6099.

29. Hosokawa, R., Oka, K., Yamaza, T., Iwata, J., Urata, M., Xu, X., Bringas, P., Jr., Nonaka, K., and Chai, Y. (1-5-2010) TGF-beta mediated FGF10 signaling in cranial neural crest cells controls development of myogenic progenitor cells through tissue-tissue interactions during tongue morphogenesis. *Dev.Biol.*, **341**, 186-195.

30. Rochais, F., Sturny, R., Chao, C. M., Mesbah, K., Bennett, M., Mohun, T. J., Bellusci, S., and Kelly, R. G. (1-12-2014) FGF10 promotes regional foetal cardiomyocyte proliferation and adult cardiomyocyte cell-cycle re-entry. *Cardiovasc.Res.*, **104**, 432-442.

31. Mori, A. D., Zhu, Y., Vahora, I., Nieman, B., Koshiba-Takeuchi, K., Davidson, L., Pizard, A., Seidman, J. G., Seidman, C. E., Chen, X. J., et al. (15-9-2006) Tbx5-dependent rheostatic control of cardiac gene expression and morphogenesis. *Dev.Biol.*, **297**, 566-586.

3. Chapter II: TBX5a Interacts with MYBBP1a to Regulate Myostatin and Skeletal Muscle Differentiation

Massomeh Sheikh Hassani, Hiba Komati and Mona Nemer*.

Molecular Genetics and Cardiac Regeneration Laboratory, Department of Biochemistry, Microbiology, and Immunology, University of Ottawa, Ottawa, Ontario, Canada.

*Corresponding author:
Dr. Mona Nemer
University of Ottawa
451 Smyth Road, Room 4522
Ottawa, Ontario K1H 8M5
Tel: 613 562 5800 ext. 3995
Email: mnemer@uottawa.ca

3.1 Statement of the Manuscript

The manuscript “TBX5a interacts with MYBBP1a and regulates Myostatin to inhibit differentiation of skeletal muscle cells” is ready to be submitted (January 2020).

3.2 Contribution Statement

In this manuscript, MSH performed all experiments, analyzed/interpreted results and wrote the draft manuscript. HK helped analyze and interpret results. MN designed experiments, interpreted results and edited the manuscript.

3.3 Acknowledgements

The authors recognize the contributions of Megan Fortier and Janie Beauregard for assistance with experiments, all members of the Nemer lab for helpful discussions, the uOttawa Histology Core Facility for in vivo staining and the Ottawa Hospital Research Institute for completion of the microarray and mass spectrometry experiments.

3.4 Sources of Funding

This work was made possible by grants provided by the Canadian Institutes for Health Research awarded to MN and to uOttawa tuition fee scholarships awarded to MSH.

3.5 Disclosures

None

3.6 Abstract

TBX5 is the causative gene for Holt-Oram Syndrome, an autosomal dominant disease characterized by heart and upper limb abnormalities. TBX5 is present within several cardiac cell types where it acts as a dosage sensitive regulator of heart morphogenesis and cardiac conduction. The mechanisms that link *TBX5* gene mutations and TBX5 activity to limb defects remain incompletely understood. We previously reported the presence in skeletal myoblasts of several TBX5 splice isoforms whose expression varies during skeletal muscle proliferation and differentiation. TBX5a, the most abundant isoform in C2C12 myoblasts, inhibits their differentiation into myotubes. This property is specific to TBX5a and not shared with other TBX5 isoforms. To elucidate the molecular mechanisms underlying the function of TBX5a in skeletal myoblasts, we have analyzed TBX5a dependent changes in C2C12 genetic program and identified several novel protein interactors, including the Myb-binding protein, MYBBP1a, a negative regulator of skeletal myogenesis. We show that TBX5a and MYBBP1a physically interact and regulate transcription of several genes involved in myogenesis. Notably, TBX5a and MYBBP1a synergistically activate Myostatin (*Mstn*), a negative regulator of muscle differentiation and inhibit Myogenin a key inducer of skeletal myogenesis. Interestingly, TBX5c also interacts with MYBBP1a resulting in opposite effects at the level of target genes, with these targets leading to promotion of differentiation. Consistent with a role in promoting myoblast proliferation and inhibiting differentiation, we find significantly elevated levels of *Tbx5a*, *Mybbp1a* and *Mstn* in the proliferative skeletal muscle of MDX mice. Together, the results identify TBX5 as an important regulator of skeletal myogenesis and provide molecular insight that may be

relevant to understanding congenital and acquired skeletal muscle diseases such as muscle wasting.

3.7 Introduction

TBX5 belongs to a family of transcription factors characterized by a highly conserved 180 AA DNA binding domain known as the T-box. TBX5 can regulate gene transcription on its own or in collaboration with other transcription factors [1]. To date, more than 80 mutations in the *Tbx5* locus have been identified in Holt Oram Syndrome (HOS) patients [2], which result in either no protein production, or the formation of a truncated or functionally impaired protein. HOS is an autosomal dominant condition characterized by cardiac abnormalities and upper limb malformations [3]. The most common cardiac abnormalities are atrial septal defects (ASD) and ventricular septal defects (VSD), Conduction defects, hypoplastic left ventricle, mitral valve problems and endocardial cushion defects have also been reported as a consequence of the syndrome [4,5]. The skeletal abnormalities range from preaxial radial ray limb deformities and malformations of the carpal bones to thumb abnormalities (triphalangism or aplasia) and phocomelia [4]. TBX5 is highly regulated through alternative splicing and generates at least five distinct TBX5 isoforms that share the T-box domain but differ in the N- and C- terminal domains [6]. Our previous studies have shown amongst the five known TBX5 isoforms, TBX5a is the predominant isoform in e11.5 atria, ventricles, forelimbs and hind limbs as well as in cardiogenic cell lines. TBX5a is also abundantly expressed in proliferating C2C12 myogenic cells where its upregulation inhibits myotube formation; whereas the TBX5c isoform induces myotube formation. These results suggest

that TBX5 proteins can promote proliferation and differentiation of myoblasts [6]. It is now generally believed that combinatorial interactions of transcription factors lead to the complexity of gene expression during development. The differential affinity of TBX5 for cofactors might be important for its cell specific function. Some TBX5 cofactors have been identified to date mostly interacting within different cardiac cell types. They include cell specific DNA binding proteins NKX2.5, GATA4, MEF2c, SAL4, as well as general transcriptional regulators including TIP60 (an acetyl transferase), ASH21 (a histone methyl transferase) and BAF60 (a chromatin remodeling factor) BAF60c/SMARCD3 [7] as well as the NURD complex (CHD5, MTA1, HDAC1, RBBP7, GATAD2a, GATAD2, MTA3, RBBP4, CHD4, MTA2) [8]. There is little knowledge about TBX5 interactors in skeletal muscle cells, although some of the identified interactors in cardiac cells such as MEF2c, SAL4 and BAF60c are involved in skeletal muscle development. In fact,TBX5 role in skeletal myogenesis is not fully elucidated in contrast to its extensively studied function in cardiomyogenesis[7]. TBX5 role in limb development has also received attention; whereas its role in limb bud growth is established, its contribution to limb identity remains unsettled [9–11]. Identification of TBX5 cofactors and target genes in skeletal muscle cells will provide insight into its precise role and mechanism of action therein.

We have used the C2C12 myoblasts to overexpressing a Flag-TBX5a construct and identify TBX5a interactors as well as downstream targets in skeletal muscle progenitors. We found that in these cells TBX5 interacts with several proteins involved in general as well as muscle specific transcription to regulate C2C12 cell proliferation and differentiation. Importantly, we identified MYBBP1a as a novel TBX5a co-factor and show that both proteins synergistically activate Myostatin while inhibiting MyoD-mediated activation of Myogenin.

The results suggest that TBX5 is an upstream regulator of several pathways that control skeletal myogenesis. This insight may have clinical implications in inherited as well as acquired skeletal muscle diseases.

3.8 Material and Methods

Plasmids

The TBX5a-pcgn, TBX5a-pCMV, TBX5c-pcgn, TBX5c-pCMV and GST-TBX5-constructs were previously described [1,6]. GST-MyoD plasmid was provided from Ptglab-ag13208. To produce the control pLNCX2 plasmid, the 3xFlag sequence was subcloned into the pLNCX2 vector (Clontech, 631503) at the BglII/BamHI sites. A V5His tag was also subcloned between the BamHI and HindIII sites. To produce 3xFlag TBX5a plasmid, the complete mouse TBX5a sequences were subcloned from the TBX5a-pcgn plasmid and subcloned into the BamHI in frame with the 3xFlag and V5-His constructs. The pcDNA-p160MBP (full length MYBBP1a), pcDNA-p67MBP (N-terminal), pcDNA-p160C (C-terminal) available from addgene (https://www.addgene.org/41/). MyoG-pXP2 and MyoD-pCMV plasmids were kind gifts from Dr. Blais [12]. The MSTN promoter was amplified from mouse gDNA between -1100 and -1 bp from the transcription start site using the primers and subcloned into the pXP2 plasmid at the BamHI site. All constructs were confirmed by sequencing. Primers are available upon request.

Immunoprecipitation and Mass Spectrometry Analysis

To obtain 3xFlag TBX5a-coupled complexes, nuclear extracts were prepared from the 3xFlag TBX5a and pLNCX2 C2C12 cell lines and stored at -80◦C until use.

Immunoprecipitations were completed by incubating 25 mg of nuclear extracts in IP Buffer (120 mM NaCl, 1 mM EDTA, 50 mM Tris-HCl pH 8, 0.5% NP-40) with 200 ul of flag or IgG coupled protein A/G magnetic beads (Sigma M8823, Abcam ab18413, Millipore LSKMAGAG10) overnight at 4◦C with rotation. Beads were washed 3 times in IP buffer and 10 ul were reserved for verification by western blot. Beads were boiled for 5 minutes in western blot sample buffer (ensuring that it does not contain beta-mercaptoethanol) and separated by SDS-PAGE on a 12% polyacrylamide gel. Gels were silver stained, bands were excised and sent for HPLC-ESI-MS/MS analysis by the Proteomics Resource Centre at the Ottawa Hospital Research Institute (OHRI).

Co-immunoprecipitation

Co-immunoprecipitations were completed as previously described [13]. Briefly, flag-coupled complexes were immunoprecipitated from nuclear extracts of NIH 3T3 cells transfected with TBX5a-pCMV, TBX5c-pCMV and c-Myc tagged P160-pCMV (MYBBP1a) or relevant controls. Western blots were revealed with anti-Flag (Sigma F1804) and anti-c-Myc (sc-789) antibodies.

In vitro Pull-Down Assays

Pull down assays were completed as previously described [6]. Briefly, the T7 Luciferase, TBX5a-pCMV, P160-pCMV, P67-pCMV and P160C-pCMV plasmids were in vitro translated and labelled with 35S-Met using the TnT Quick Coupled Transcription/Translation System (Promega) and incubated with agarose beads with bound GST alone, GST-MyoD, GST-TBX5a 518 aa, GST-TBX5a 58-518 aa, GST-TBX5a 327-518 aa, GST-TBX5c 51-404 aa or GST-TBX5c 327-404 aa. Bound protein was resolved on 12% SDS-PAGE gels, dried and exposed with film.

Cell Culture, Transfections, Transductions and Differentiations

NIH 3T3 and C2C12 cells were maintained in DMEM supplemented with 10% fetal bovine serum and penicillin/streptomycin. 24 hours prior to calcium phosphate transfection, cells were plated at a confluency of 50,000 cells/well in 12-well dishes and transfected as previously described [6]. Luciferase assays were completed as previously described [6]. To produce the C2C12 pLNCX2, 3xFlag TBX5a stable cell lines, the 3xFlag pLNCX2, 3xFlag TBX5a pLNCX2 plasmids were transfected via calcium phosphate into AD293 cells. 48 hours post-transfection, media containing retroviral particles was collected and added to 80% confluent C2C12 cells in the presence of 10 uM hexadimethrine bromide (Sigma, TR-1003-G). 48 hours post-transduction, media was replaced with DMEM supplemented with 500 ug/ml G418 and was replaced every 48 hours (Sigma G8168). Overexpression of the 3xFlag-tagged construct was confirmed one-week post-transduction by QPCR, western blot and immunofluorescence.

Western Blot

Western blots were completed as previously described [6]. Antibodies were used at the following concentrations: Anti-Flag (Sigma F1804) 1/1000, anti-c-Myc (Santa Cruz sc-789) 1/200, and anti-Nucleolin (Cell Signaling 14574) 1/2000. Homemade anti-TBX5a antibody was used at a concentration of 1/1000 [1,6].

Microarray

Microarray analysis of the 3xFlag TBX5a and respective control cell line was completed as previously described [14]. Functional analysis was completed using Ingenuity Pathway Analysis (IPA) by Qiagen (www.qiagen.com/ingenuity).

Chromatin Immunoprecipitation

ChIP assays were completed as previously described [15,16]. The primers used are sense 5'-TCCTCCCCATCTGTGTCATC-3' and antisense 5'-GGATCCATCACCATCAATAACC-3' for the Chromosome 15 Gene Desert negative control, sense 5'-AGCTTGCCCTCGACTGTAAC-3', and antisense 5'-CTGCTCCACAATGAATCTCG-3'for Proximal MyoD E-box site (-44 to -55 bp from TSS), sense 5'-AGCTTGCCCTCGACTGTAAC -3' and antisense 5'- CTGCTCCACAATGAATCTCG -3'for the proximal Mstn promoter TBE binding site (-76 to -164 bp from TSS), sense 5'-CAGCCCTGGAAGTCTGAGTC-3' and antisense 5'-AGGCCAGGCTCTGTACTTAGG-3'for the proximal Mstn promoter TBE binding site (-753 to -844 bp from TSS), sense 5'-GGCTCCATTCTCCATTCTCC-3' and antisense 5'-GTGCTGTCTGCTGCTTTCTG-3'for the distal Mstn promoter TBE binding site (-2281 to -2414 bp from TSS), sense 5'-TGGTCTAAAGGCCTCGACAG-3' and antisense 5'-ACACGTAGCTCACCTCACC-3'for the distal Mstn promoter TBE binding site (-2772 to -2862 bp from TSS). 3 ug of antibody was used per 25 ug of chromatin. The antibodies used are anti-IgG mouse (Abcam ab18413) and anti-TBX5 (Proteintech 13178-1-AP).

QPCR and RNA analysis

Total RNA was isolated from cells using Trizol reagent (Invitrogen). QPCR was carried out as previously described [6].Quantitative PCR analysis was performed on C2C12 cells overexpressing the indicated TBX5 constructs as described previously [6]. siRNAs were obtained from Sigma and transfected in C2C12 cells using the HiPerFect transfection reagent from Qiagen (catalog number 301705). Primers are available upon request.

Immunohistochemistry and Immunofluorescence

Immunohistochemistry was done on mouse tissues at different embryonic stages as described previously [6]. Homemade anti-TBX5a antibody was used at a concentration of 1/2300. For immunofluorescence, tissues were fixed in 4% PFA for 24 hours. Paraffin sections were stained with primary antibodies against TBX5a, PAX7, KI67, MYBBP1a and MSTN followed by fluorescent secondary antibodies. For eMHC-stained slides, panoramic pictures were acquired with a Carl Zeiss AxioVision Observer Z1.

In Vivo Studies

All animal experiments were completed in accordance with the University of Ottawa guidelines for animal care and were approved by the Institutional Animal Care and Use Committee (IACUC). 6 to 9 weeks old Wild type and *MDX mice* (C57BL/10ScSn-Dmdmdx from Jackson Laboratories) were kindly provided by Dr. Bernard Jasmin at uOttawa. At end points, mice were perfused in 4% paraformaldehyde, embedded and sectioned for immunofluorescence.

3.9 Results

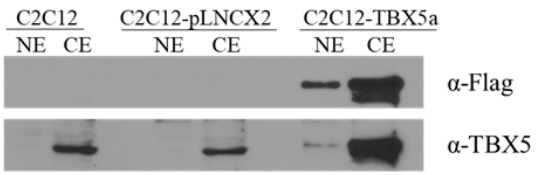
MYBBP1a is a novel TBX5 interacting partner in skeletal myoblasts

To identify novel protein partners that modify TBX5a function in skeletal muscle cells, we employed an unbiased immunoprecipitation and mass-spectrometry approach using the C2C12 skeletal muscle cell line. To facilitate the identification of TBX5a interacting partners in this cell line, TBX5a was stably overexpressed via retroviral transduction of a 3xFlag TBX5a construct and its expression was validated using Western blot and QPCR experiments (**Figure 3.1 A-C**). The resulting Flag-tagged complexes were

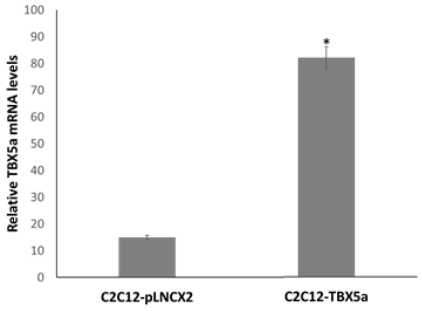
immunoprecipitated from nuclear protein extracts and identified using high performance liquid chromatography electrospray ionization tandem mass spectrometry (HPLC-ESI-MS/MS) (**Figure 3.1 D-E**). Using this approach, we have identified over 180 potential TBX5a interacting partners in the skeletal muscle cells including other transcription factors, enzymes, factors involved in cell cycle, differentiation, focal adhesion and cytoskeleton pathways (**Figure 3.1 F** and **Supplementary Table 3.1**). Some of the TBX5a associated proteins were previously described, including the NURD complex which to some extent served as a positive control [17]. Additionally, several novel interactors were identified (**Figure 3.1 G**).

A

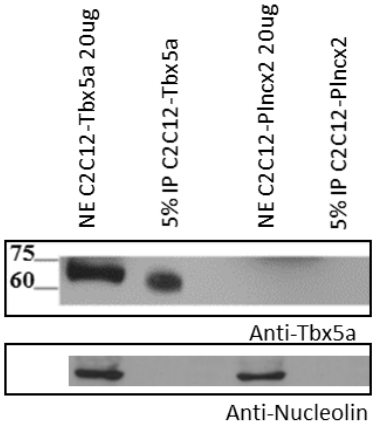
B



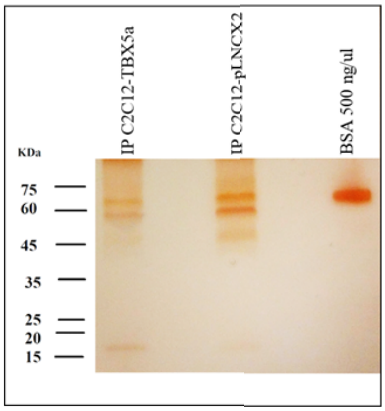
C



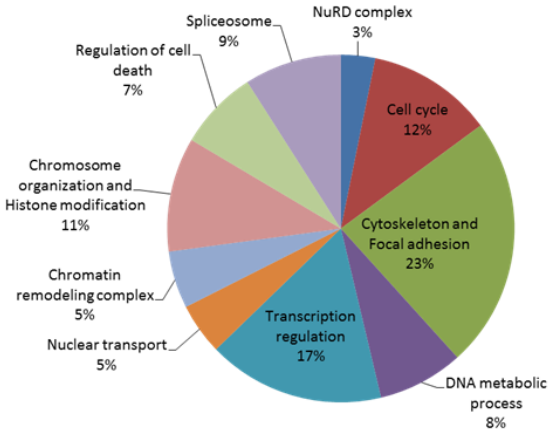
D



E



F



G

Selected functional pathways of Tbx5a protein interactors	Protein involved in pathway
NurD complex	HDAC1, RBBP7, GATAD2B, RBBP4, CHD4, MTA3
Chromosome organisation and Histone modification	HDAC1/2,PRMT5, SMARCD2/3, BRG1
Regulation of transcription	NONO,PSPC1,ILF3,TEAD1,RBPJ
Nuclear transport	NUP54, MYBBP1A, KPNA3
Regulation of cell death	BCLAF1, API5, HSPD1,CTNNBL1,CLU

Figure 3.1 Identification of TBX5a protein interactors in skeletal muscle cells. A. Structure of retrovirus construct used to infect C2C12 cells to generate the stable cell line overexpressing TBX5a. **B.** Western blots of nuclear and cytoplasmic extracts of the C2C12, C2C12-pLNCX (empty vector) and C2C12-TBX5a cells revealed with Flag and TBX5a antibody showing overexpression of TBX5a at protein level. **C.** qPCR analysis of C2C12-pLNCX (empty vector) and C2C12-TBX5a cell lines showing levels of mRNA fold change of TBX5a expression normalized to Gapdh in the two stable cell lines. *P ≤ 0.05. **D.** Western blot of nuclear extracts (NE) and 5% of IP product of control C2C12-PLNCX2 and C2C12-TBX5a for validation revealed by anti-TBX5a and anti-Nucleolin. **E.** Silver stain gel of C2C12-TBX5a stable line and C2C12-PLNCX2 after 3xFlag immunoprecipitation. TBX5a ~70 kDa and C2C12-PLNCX2 serves as a control (Empty vector). **F.** Representation of some of the functional pathways of the interacting proteins of C2C12-TBX5a from MS/MS analysis **G.** List of selected functional pathways of TBX5a protein interactors and selected proteins of each pathway from the MS/MS results.

Of this group, the most consistently identified protein with a high abundance of peptides was Myb-Binding Protein (P160) 1a (MYBBP1a). MYBBP1a is a 160 kDa protein, expressed in all tissues and predominantly located in the nucleolus, and contains a Nuclear Export Signal (NES), a leucine zipper domain and four LXXLL motifs in the N-terminal and a Nuclear and Nucleolar Localization Sequences (NLS), another leucine zipper domain and basic and acidic regions in the C-terminal of the protein (**Figure 3.2 A**). The LXXLL motifs of MYBBP1a often mediate interactions between nuclear receptors and their cofactors and participate in many protein–protein interactions associated with different aspects of transcriptional regulation [18]. Mybbp1a was originally identified as a protein binding to the proto-oncogene c-MYB which occurs through the leucine zipper motifs present in the N-terminal region of MYBBP1a and the Negative Regulatory Domain (NRD) of c-MYB [19]. Subsequently, it has been shown to interact with and regulate several transcription factors. It has been shown to interact with the peroxisome Proliferator-Activated Receptor Gamma Coactivator 1-alpha (PGC-1α) through both the N- and C-terminal domains repressing its ability to stimulate mitochondrial respiration [20]. MYBBP1a also binds to an represses Pre-

Sequence Protease 1 (PREP1) through the N-terminal and competes with Pre-B-cell Leukemia Transcription Factor 1 (PBX1) binding to PREP1 regulating development and organogenesis [21]. MYBBP1a is also known to compete with P300 in binding to NF-KB and repressing its function [22]. On the other hand, MYBBP1a also acts an activator and positive regulator of some transcription factors such as the Aromatic Hydrocarbon Receptor (AHR) where it enhances its ability to activate transcription in response to certain hydrophobic ligands [23]. Overall, MYBBP1a functions by regulating the activity of various transcription factors and is important in essential biological processes such as proliferation, differentiation, cell division and apoptosis [24].

Interestingly, MYBBP1a was shown to suppress C2C12 cell differentiation by inhibiting MyoD, thus preventing Myogenin activation [25]. Given that TBX5a inhibits differentiation of C2C12 myoblasts into myotubes [6], we further characterized the interaction between TBX5a and MYBBP1a. To confirm that MYBBP1a can interact directly with TBX5a, pull down assays were performed. Moreover, its ability to interact with the other major TBX5 isoform in C2C12, TBX5c was tested. As mentioned earlier, the two TBX5 isoforms have opposing effects in these cells, with TBX5c promoting differentiation into myotubes [6]. This difference may be the result of structural variations in the C-terminal of the two isoforms resulting in a novel Variable Domain (VD) in TBX5c C-terminus causing the loss of trans-activation domain 3 (TAD3). This difference might produce variable protein-protein interactions and different transcriptional regulation of their downstream targets (**Figure 3.2 E**).

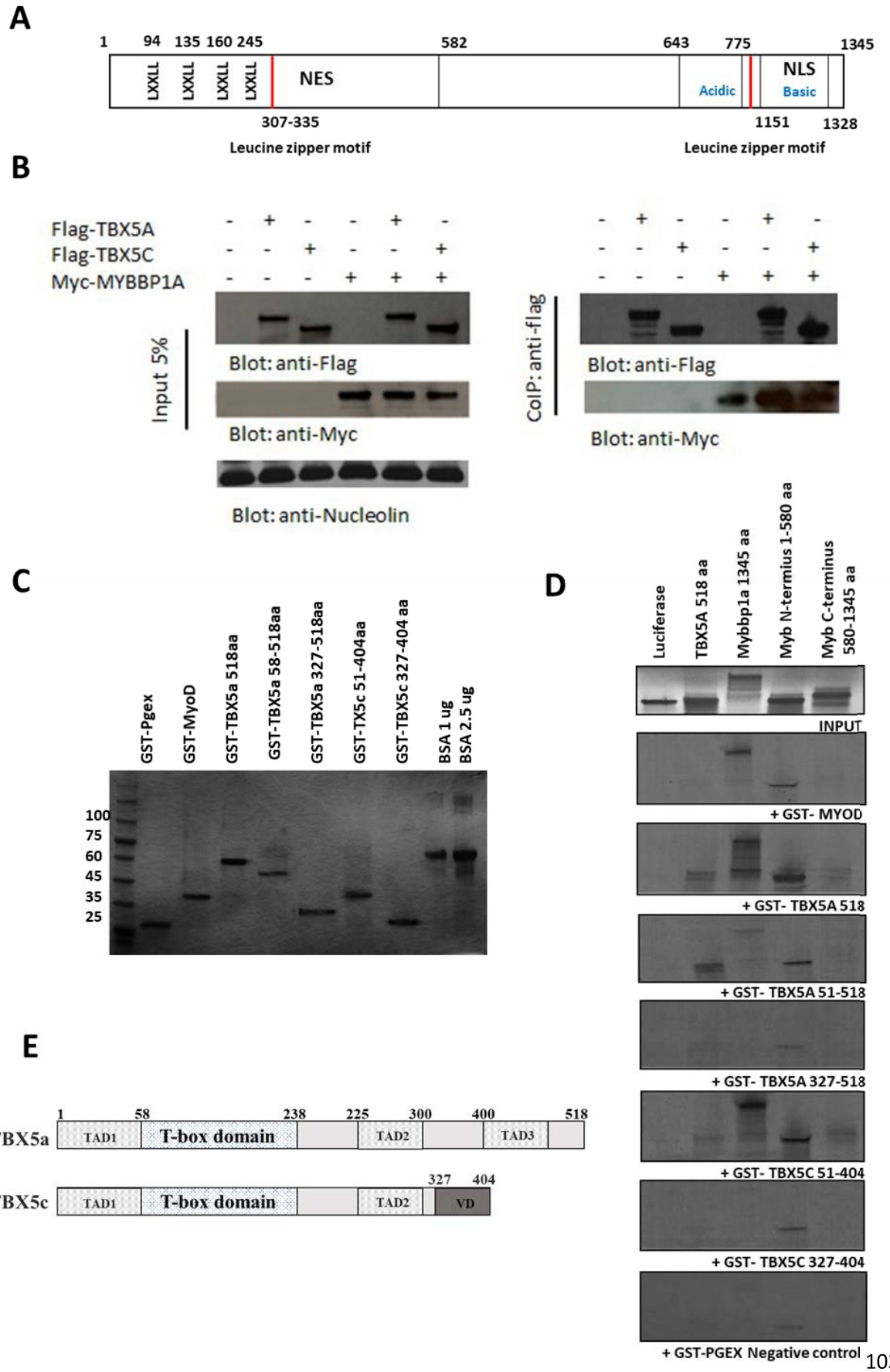

A
1
94
135
160
245
582
643
775
1345
LXXLL
LXXLL
LXXLL
LXXLL
NES
Acidic
NLS
Basic
307-335
1151
1328
Leucine zipper motif
Leucine zipper motif
B
Flag-TBX5A
Flag-TBX5C
Myc-MYBBP1A
Input 5%
Blot: anti-Flag
Blot: anti-Myc
Blot: anti-Nucleolin
CoIP: anti-flag
Blot: anti-Flag
Blot: anti-Myc
C
GST-Pgex
GST-MyoD
GST-TBX5a 518aa
GST-TBX5a 58-518aa
GST-TBX5a 327-518aa
GST-TX5c 51-404aa
GST-TBX5c 327-404 aa
BSA 1 ug
BSA 2.5 ug
100
75
60
45
35
25
D
Luciferase
TBX5A 518 aa
Mybbp1a 1345 aa
Myb N-termius 1-580 aa
Myb C-terminus 580-1345 aa
INPUT
+ GST- MYOD
+ GST- TBX5A 518
+ GST- TBX5A 51-518
+ GST- TBX5A 327-518
+ GST- TBX5C 51-404
+ GST- TBX5C 327-404
+ GST-PGEX Negative control
E
TBX5a
1
58
238
225
300
400
518
TAD1
T-box domain
TAD2
TAD3
TBX5c
327
404
TAD1
T-box domain
TAD2
VD

Figure 3.2 MYBBP1a is a protein interactor of TBX5 in skeletal muscle cells. A. schematic of MYBBP1a shown with NLS, NES, LXXLL motifs, Leucine zipper motifs (red) and acidic and basic region (blue). **B.** In vitro CoIP experiments revealed with flag mouse Antibody- which shows TBX5a, TBX5c bands; and Myc Goat Antibody detects MYBBP1a. **C.** GST Fusion proteins used for pulldown were ran on gel and stained with Coomassie blue and BSA loaded as quantitative measure. **D.** Pulldown of MYBBP1a and TBX5a and TBX5c: a) in vitro translated proteins (INPUT 5%): T7 Luciferase -control, TBX5a +control, MYBBP1a 1328 aa, P67 (N-terminal of MYBBP1a), P160C (C-terminal of MYBBP1a). b) were incubated with GST-MYOD fusion protein as positive control, c) GST-TBX5a 518 aa fusion protein, d) GST-TBX5a 58-518 aa fusion protein, e) GST-TBX5a 327-518 aa fusion protein (no T-box), f) GST-TBX5c 51-404 aa fusion protein, g) GST-TBX5c 327-404 aa (only Variable domain) fusion protein, h) GST-PGEX (Negative control) fusion protein. **E.** Schematic illustration of protein structure of TBX5a and TBX5c isoforms showing T-box domain and trans-activation domains (TAD). Note: TBX5c has a variable domain (VD) which results in loss of TAD3.

Interactions between MYBBP1a and both TBX5a and TBX5c were confirmed initially by co-IP from nuclear extracts (**Figure 3.2 B**). Furthermore, TBX5-GST fusion proteins were used to study the direct interaction of MYBBP1a and TBX5. The results showed that the two proteins physically interact (**Figure 3.2 C**). This interaction is comparable to the known interaction of MYBBP1a with MyoD and is mediated by the T-box domain of TBX5, which is required for most protein-protein interactions of TBX5 (**Figure 3.2 D**). The pulldown experiments also demonstrated that both TBX5 isoforms interact with the N-terminal of MYBBP1a which contains four LXXLL motifs and a leucine zipper domain. These results show for the first time that TBX5 protein binds to a leucine zipper domain and that MYBBP1a is a TBX5 interactor.

TBX5 modulates MYBBP1a transcriptional activity on muscle specific promoter

MYBBP1a has been shown to act as a transcriptional co-repressor in proliferating muscle progenitor cells by exerting a negative effect on MYOD transcriptional activity [25]. Myogenic/Myoblast differentiation 1 (MYOD) transcription factor belongs to the basic

helix-loop-helix (bHLH) family proteins that are key regulators of myogenesis during embryogenesis [26,27]. The family includes Myogenin, a muscle specific gene essential for muscle differentiation which is transcriptionally silenced in the undifferentiated cells by multiple factors including MYBBP1a. We tested whether TBX5a or TBX5c can modify MYBBP1a function by either interfering with or enhancing its ability to inhibit MYOD-dependent Myogenin promoter. MYBBP1a, MYOD, TBX5a or TBX5c mammalian expression plasmids were cotransfected with the Myogenin promoter reporter plasmid into C2C12 cells. As expected, MYOD activated the Myogenin promoter and addition of MYBBP1a blocked this activation. The addition of TBX5a potentiated MYBBP1a inhibition (**Figure 3.3 A**). Interestingly, TBX5c had an opposite effect alleviating MYBBP1a inhibitory effect and synergistically enhanced MyoD mediated activation of the Myogenin promoter (**Figure 3.3 B**).

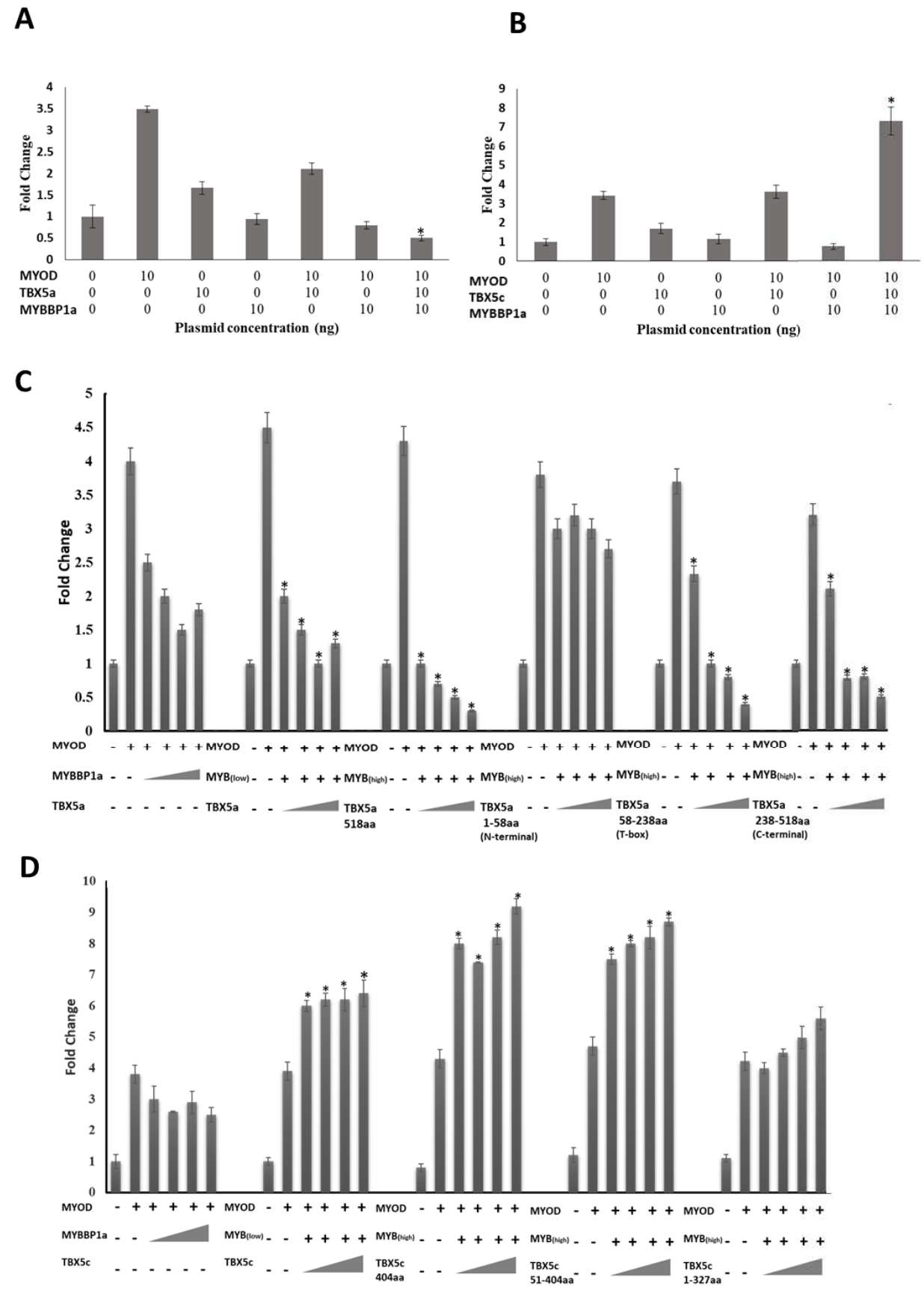

A
Fold Change
MYOD 0 10 0 0 10 10 10
TBX5a 0 0 10 0 10 0 10
MYBBP1a 0 0 0 10 0 10 10
Plasmid concentration (ng)
B
Fold Change
MYOD 0 10 0 0 10 10 10
TBX5c 0 0 10 0 10 0 10
MYBBP1a 0 0 0 10 0 10 10
Plasmid concentration (ng)
C
Fold Change
MYOD
MYBBP1a
MYB(low)
MYB(high)
TBX5a
TBX5a 518aa
TBX5a 1-58aa (N-terminal)
TBX5a 58-238aa (T-box)
TBX5a 238-518aa (C-terminal)
D
Fold Change
MYOD
MYBBP1a
MYB(low)
MYB(high)
TBX5c
TBX5c 404aa
TBX5c 51-404aa
TBX5c 1-327aa

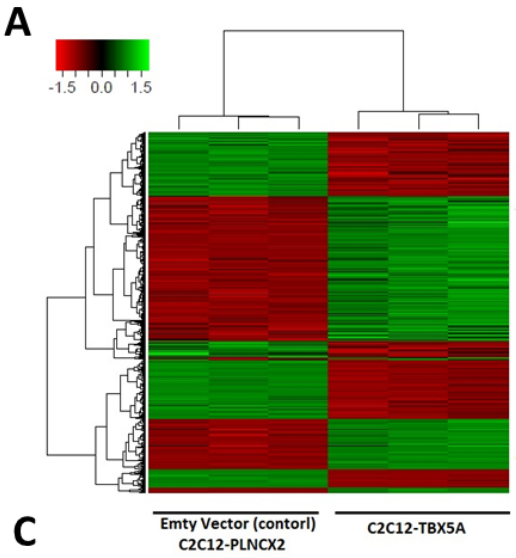
Figure 3.3 MYBBP1a and TBX5a inhibit differentiation by downregulating MyoG while TBX5c and MYBBP1a activate MyoG. A, B. C2C12 cells transfected with MyoD, TBX5a or TBX5c, MYBBP1a or triple transfection of all three plasmids and Myogenin reporter constructs. Note that TBX5a and MYBBP1a inhibit Myogenin promoter activation by MyoD, Whereas TBX5c and MYBBP1a enhance MyoD mediated activation of Myogenin promoter. **C.** Structure function analysis of TBX5a and MYBBP1a interaction**.** C2C12 cells transfected with MyoD (10 ng), increasing doses of MYBBP1a (10, 25, 50, 100), or low dose of MYBBP1a (10 ng) versus high dose (100 ng) and increasing doses (10, 25, 50, 100) of TBX5a (518 aa full protein) or TBX5a N-terminal (1-58aa), or TBX5a T-box region (58-238aa), or TBX5a C-terminal (238-518aa), and the Myogenin reporter constructs. **D.** Structure function analysis of TBX5c and MYBBP1a interaction**.** Same doses as explained above where TBX5a is replaced with TBX5c (full protein 404aa), or TBX5c 51-404aa (including T-box and C-terminal), or TBX5c 1-327aa (common region with TBX5a). Note: all representative of three experiments, * $p < 0.0005$ vs. *MyoD*.

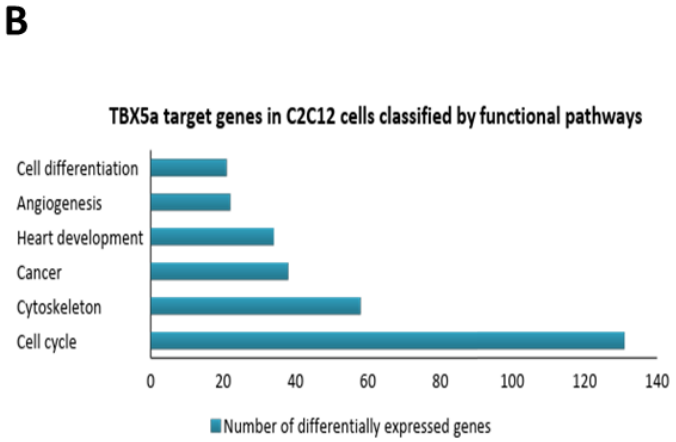
As both TBX5a and TBX5c physically interact with MYBBP1a but affect its function differently, we carried out a structure-function study to determine which regions of TBX5a/TBX5c are responsible for transcriptional regulation of the MyoG promoter. TBX5a or TBX5a deletion constructs (N-terminus, T-box or T-box+C-terminus) or TBX5c and its deletion constructs (N-terminus+T-box or T-box+C-terminus) were cotransfected with the Myogenin promoter reporter plasmid into C2C12 cells together with MYOD and MYBBP1a. TBX5a potentiated MYBBP1a inhibition of Myogenin promoter activity at both low and high doses; this action was recapitulated by a truncated TBX5a protein containing the T-box region (58-238 aa), where MYBBP1a and TBX5a physically interact (**Figure 3.3 C**). In contrast, increasing doses of TBX5c counters MYBBP1a inhibition and increases MYOD-dependent activation of Myogenin promoter (**Figure 3.3 D**). These results are consistent with our previous findings on the opposing roles of TBX5a and TBX5c in skeletal muscle differentiation [6]; TBX5a enhances MYBBP1a downregulation of Myogenin, a key differentiation factor during myogenesis whereas TBX5c pro-differentiation effect is consistent with its activation of Myogenin.

TBX5a regulates multiple pathways involved in negative regulation of myogenesis including Myostatin

To further elucidate the mechanism of action of TBX5a, we identified its potential downstream gene targets using transcriptomic analysis of the C2C12 cells overexpressing or not a 3xFlag TBX5a. Using the Affymetrix GeneChip Mouse Gene 2.0 ST Array (**Figure 3.4 A**). We identified 1200 genes with altered expression between the C2C12 pLNCX2 and C2C12 3xFlag TBX5a cell lines (760 increased and 590 decreased, Log2 fold change ≥0.6 or ≤-0.6 and P≤0.05, Mann–Whitney test with P value adjusted—Benjamini and Hochberg—for multiple comparisons) (**Supplementary Table 3.2 and Supplementary Figure 3.1**). Analysis of the data using the Ingenuity Pathway Analysis (IPA) Software (QIAGEN) identified multiple pathways enriched by TBX5a overexpression functionally linked to processes involved in skeletal muscle development (**Figure 3.4 B&C**). Enrichment of several genes in these pathways, many not previously known to be targets of TBX5a, was confirmed by qPCR and was consistent with our microarray results (**Figure 3.4 D**). interestingly, the Myostatin gene (*Mstn*) was significantly upregulated by TBX5a as revealed in both microarray results (by 9 folds) and qPCR validations. Myostatin is known to be a negative regulator of myogenesis shown to function by controlling proliferation of myoblasts. In a previous study, it was show that Myostatin is an inhibitor of myoblast differentiation and that the inhibition is mediated through Smad3 (a key mediator of TGF-β inhibition of myogenesis) interference with MyoD activity and expression [28]. Animal models show loss of *Mstn* results in muscle hypertrophy and the pharmacological inhibition of it positively affects disease progression in dystrophic animals [29,30]. Furthermore, *MSTN* has shown to be elevated with advanced age, patients with cancer cachexia and chronic diseases

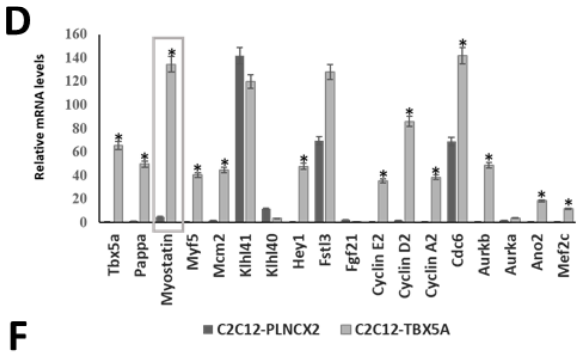
such has kidney disease, heart failure and Chronic Obstructive Pulmonary Disease (COPD) [30–33].

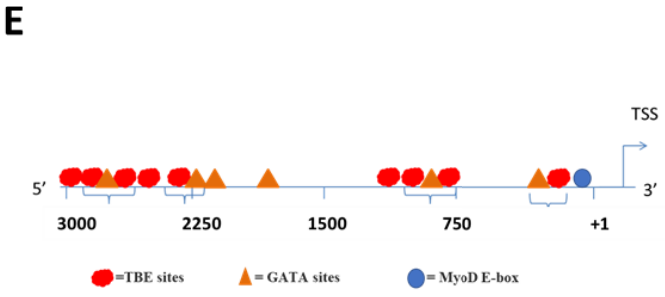




C

Selected functional pathways of Tbx5a target genes	Differentially expressed genes in pathway
Cell cycle	Ccne1/2, Cdc6, Chek1/2, Aurkb/a, Ccnb2, Ccnd2, Cdc45, Cdk6, Ccna2, Gas1
Cytoskeleton	Nusap1, Cenpf, Clic5, Marcks, Nup85, Sass6, Hook1, Ska1, Eml4, Fez1
Development	Nfatc4, Tgfbr2, Smad3, Tgfbr3, Id2, Pdlim3, Mef2c, Kcnj8
Angiogenesis	Edn1, Col18a1, Angptl6, Fgfr1, Mmp19, Ednra, Mmp14
Cell differentiation	Clu, Nrg1, Vegfc, Ngf, Socs5, Tgfb3, Jag1, Ifi204, Runx2, Mstn
Cancer	Hhip, Brca2, Jun, Rad51, Figf, Pax8, Fgf21, Kras, Wnt10b, Rarb





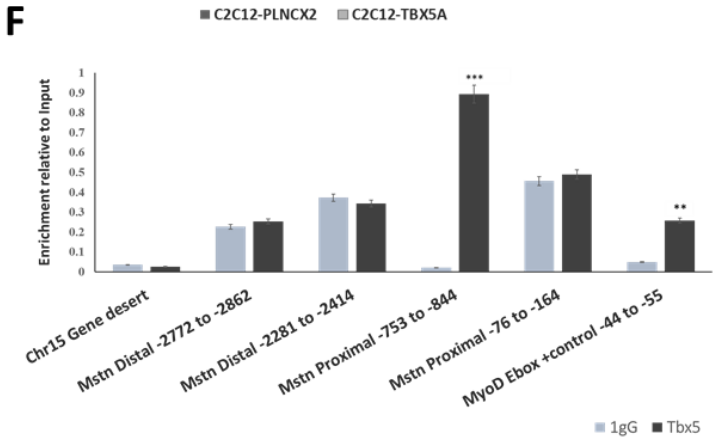


Figure 3.4 Microarray analysis identifies Myostatin as a downstream target of TBX5a in skeletal muscle cells. A. Heat map illustrating the differential gene expression of the Control C2C12 overexpressing pLNCX2 (triplicates) versus C2C12 cells overexpressing TBX5a (triplicates). **B.** TBX5a target genes in C2C12 classified by functional pathways. **C**. List of TBX5a selected target pathways/genes by IPA analysis. **D.** qPCR confirmation of selected differentially expressed genes from microarray results. Values were normalized to Gapdh and are the mean values +/- SEM * $p<0.05$, TBX5a versus PLNCX2 (control). **E.** MSTN mouse promoter mapped with predicted TBE and GATA binding sites (all with matrix similarity >0.8). Note: The curly brackets highlight sites for which chip oligos are generated. **F.** ChIP analysis demonstrating TBX5 binding sites on the Myostatin promoter ***$p<0.001$, **$p<0.01$ IgG ChIP negative control Vs. TBX5a ChIP.

Since expression of MSTN highly correlated to TBX5a; we examined whether MSTN is a downstream target of TBX5a. Bioinformatics analysis of the Mus musculus MSTN promoter by Genomatix (genomatix.de) located 9 conserved TBE motifs within the first 3000 base pairs from the transcription start site (**Figure 3.4 E**). Chromatin immunoprecipitation of TBX5a in the C2C12 cell line demonstrated that TBX5a binds in vivo to the proximal TBE motif (-753 to -844), suggesting that MSTN is a direct target of TBX5a in skeletal muscle cells (**Figure 3.4 F**).

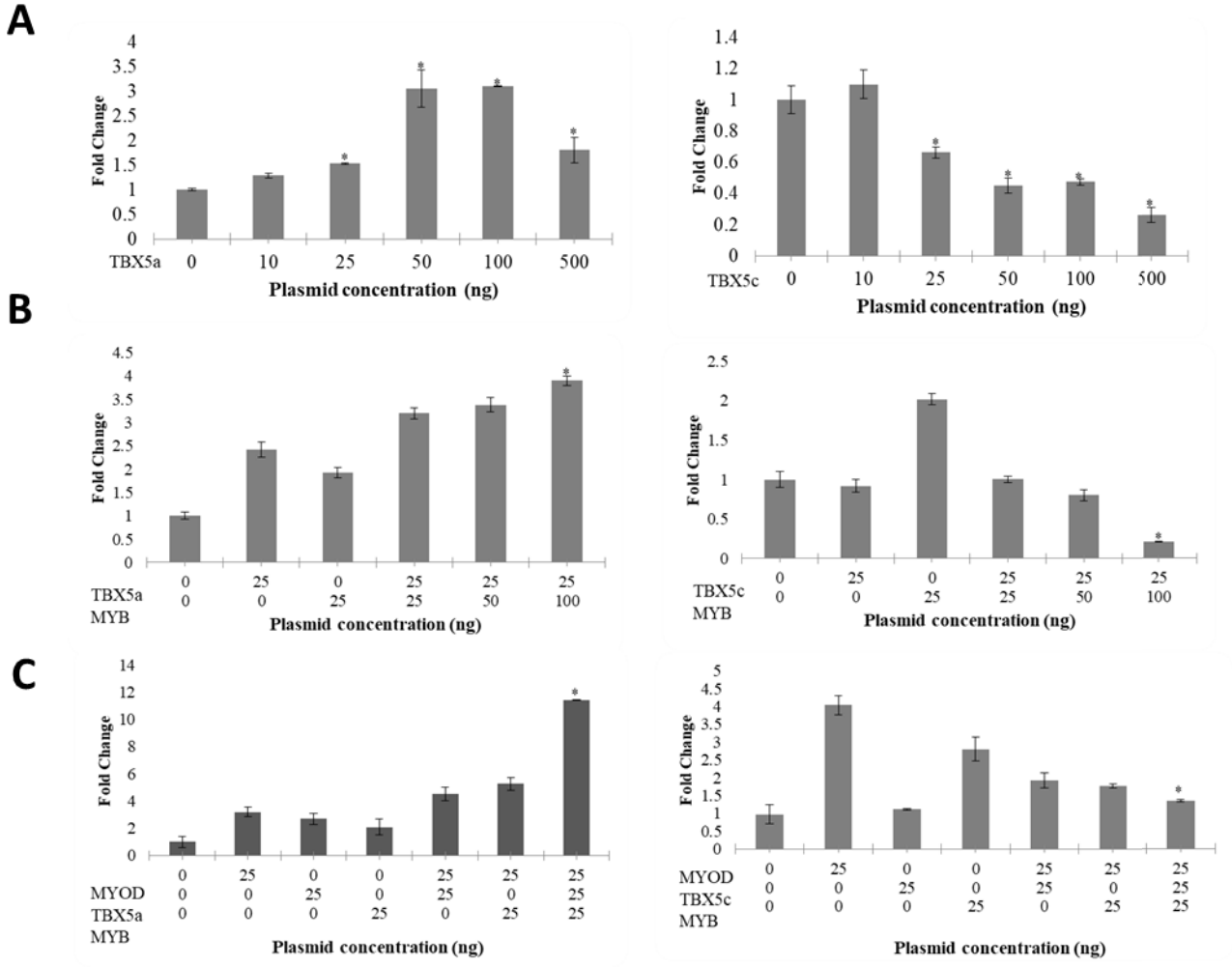
TBX5a and MYBBP1a synergistically upregulate MSTN transcriptional activity

Given the synergistic inhibition of Myogenin (pro-differentiation transcription factor) transcription by TBX5a and MYBBP1a, we tested whether MYBBP1a can modify TBX5 effect on the MSTN promoter. Firstly, we carried out luciferase assays to observe effect of TBX5a/c on the MSTN promoter and showed that TBX5a activates the promoter in a dose dependent manner whereas TBX5c inhibits MSTN transcription (**Figure 3.5 A**). Further luciferase assays were carried out using co-transfection of TBX5a or TBX5c with or without MYBBP1a with the MSTN proximal promoter reporter into C2C12 cells. TBX5a but not

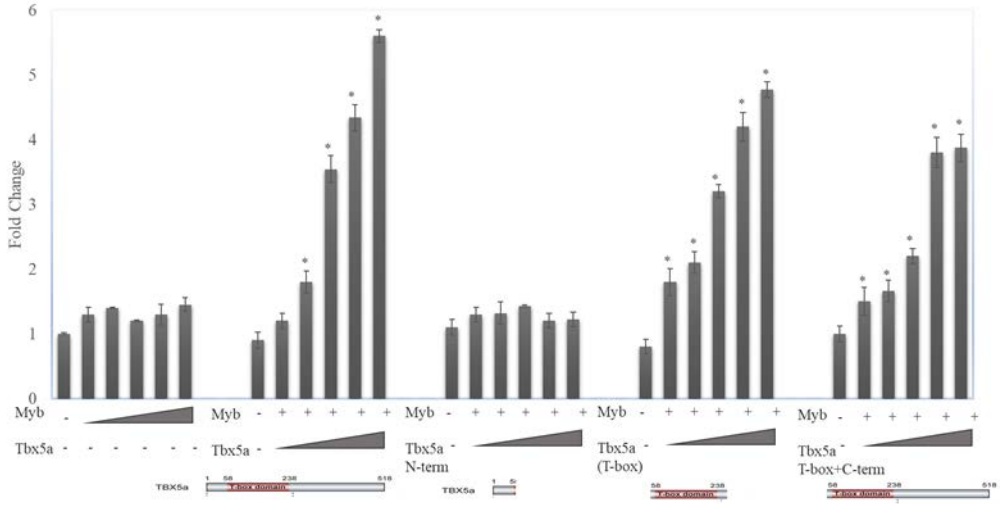
TBX5c activated the MSTN promoter and addition of MYBBP1a resulted in synergistic promoter activation. Interestingly, the same amount of MYBBP1a resulted in significant inhibition of MSTN promoter activity in presence of TBX5c (**Figure 3.5 B**). It was previously shown that MSTN promoter activity is upregulated by MyoD via its binding to E-box 6 (E6) located in the proximal promoter [34]. We asked how MYBBP1a and TBX5 might affect MyoD mediated activation of the MSTN promoter.

Luciferase assays performed using co-transfection in C2C12 cells showed that the combined presence of TBX5a, MYBBP1a and MYOD resulted in synergistic activation of the MSTN promoter; in contrast, the presence of TBX5C and MYBBP1a inhibited MYOD-dependent promoter activity (**Figure 3.5 C**). Structure-function analysis showed that TBX5a driven activation was recapitulated by a short domain containing the T-box and an activation domain previously reported by our group [1] (**Figure 3.5 D**). Interestingly, TBX5c inhibitory effect is alleviated when only the TBX5c region 1-327aa (common with TBX5a) is used (**Figure 3.5 E**).

To further examine the effects of the different TBX5 isoforms on the differentiation genes studied, we tested the effects of *Tbx5a* and *Tbx5c* down regulation on *Mybbp1a*, *Mstn* and *MyoG*. To do so, siRNA specific to *Tbx5a* or *Tbx5c* or control-siRNA were transfected into C2C12 cells. Downregulation of decreased expression of *Mybbp1a* and *Mstn* but increased *MyoG* levels (**Figure 3.5 F**). Conversely, downregulation of *Tbx5c*, increased expression of *Mybbp1a* while *MyoG* levels were decreased. Together, this suggests that the variable C-terminal region of the TBX5 isoforms contributes to their differing functions.


A
Fold Change
TBX5a 0 10 25 50 100 500
Plasmid concentration (ng)
Fold Change
TBX5c 0 10 25 50 100 500
Plasmid concentration (ng)
B
Fold Change
TBX5a 0 0 25 25 25 25
MYB 0 25 0 25 50 100
Plasmid concentration (ng)
Fold Change
TBX5c 0 0 25 25 25 25
MYB 0 25 0 25 50 100
Plasmid concentration (ng)
C
Fold Change
MYOD 0 0 25 0 25 25 25
TBX5a 0 0 0 25 0 25 25
MYB 0 0 25 0 25 0 25
Plasmid concentration (ng)
Fold Change
MYOD 0 0 25 0 25 25 25
TBX5c 0 0 0 25 0 25 25
MYB 0 25 0 25 0 25 25
Plasmid concentration (ng)

D


Fold Change
Myb
Tbx5a
Tbx5a
N-term
Tbx5a
(T-box)
Tbx5a
T-box+C-term

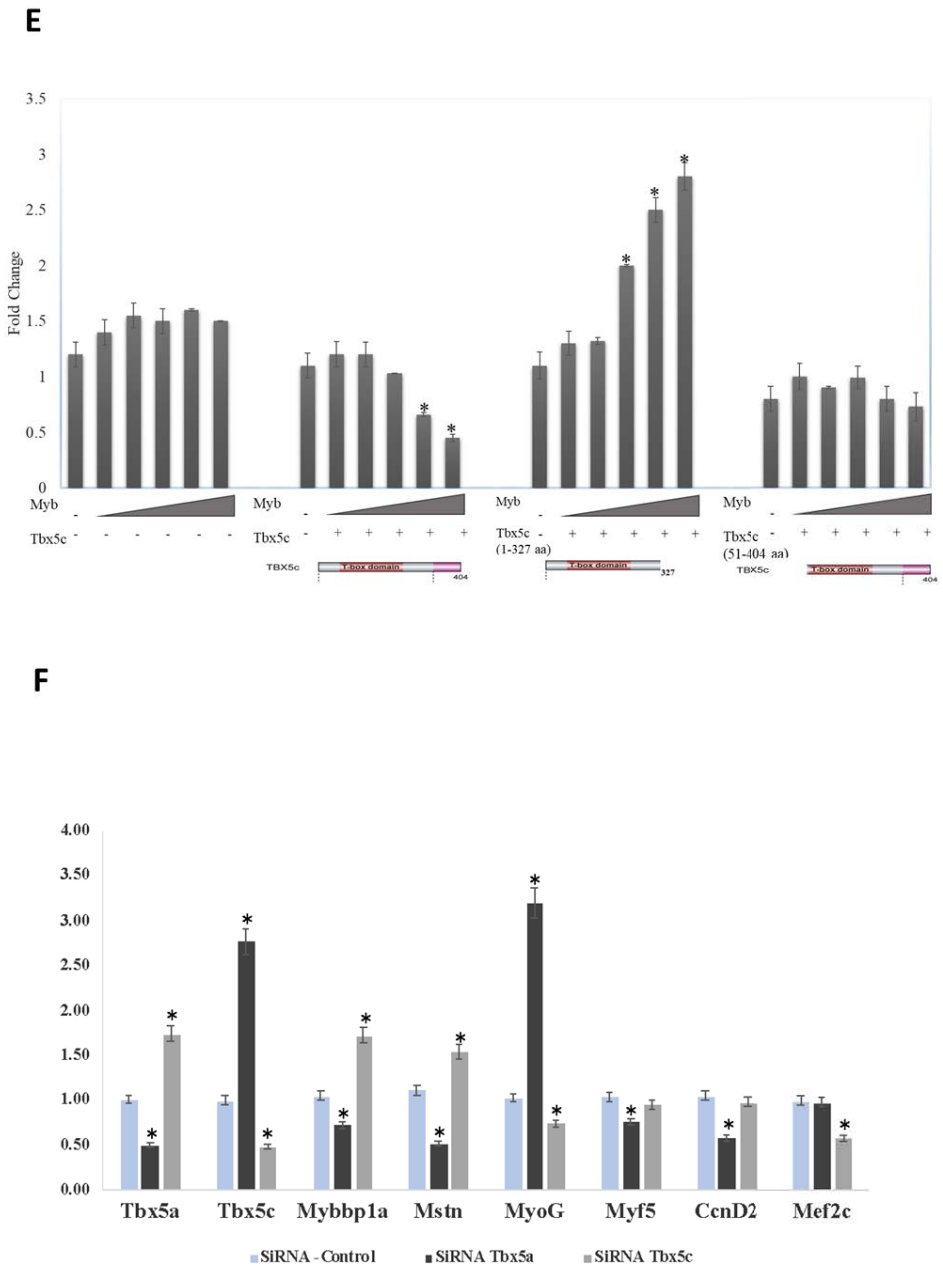


Figure 3.5 TBX5a plays an inhibitory role in differentiation of skeletal muscle cells in C2C12 cells. TBX5a and MYBBP1a synergistically upregulate Myostatin promoter while TBX5c and MYBBP1a inhibit it**. A.** Dose response assays: C2C12 cells transfected with increasing doses of TBX5a (left panel) or TBX5c (right panel) and MSTN reporter construct, * $p < 0.0005$ vs. baseline. **B.** C2C12 cells transfected with MYBBP1a and increasing doses of TBX5a (left panel) or TBX5c

(right panel) and MSTN reporter construct. **C.** C2C12 cells transfected with MyoD, TBX5a (left panel) or TBX5c (right panel), MYBBP1a or triple transfection of all three plasmids and MSTN reporter constructs. **D.** C2C12 cells transfected with increasing doses of MYBBP1a (10, 25, 50, 100, 500) and MSTN; or MYBBP1a (100 ng) and increasing doses of TBX5a (10, 25, 50, 100, 500) and MSTN reporter constructs (right panel); or MYBBP1a (100ng) and increasing doses (10, 25, 50, 100, 500) N-terminal or T-box or C-terminal of TBX5a and MSTN. **E.** C2C12 cells transfected with increasing doses of MYBBP1a (10, 25, 50, 100, 500) and MSTN; or MYBBP1a (100 ng) and increasing doses of TBX5c (10, 25, 50, 100, 500) and MSTN reporter constructs; or MYBBP1a (100ng) and increasing doses (10, 25, 50, 100, 500) of TBX5c region 1-327 aa or 327-404 (variable domain) and MSTN. Note: all representative of three experiments, * $p < 0.0005$ vs. baseline. **F.** Downregulation of Tbx5a by siRNA in C2C12 cells inhibits expression of Mybbp1a and Myostatin while increases MyoG. Tbx5c knockdown reduces expression of MyoG. *Normalized by Gapdh.

TBX5a levels in muscle myopathies

TBX5 has been shown to play a role in some cancers and muscle myopathies in addition to HOS. TBX5 is able to complex with certain oncogenes of the muscle lineage and promote progression of cancer [35]. TBX5 and its protein interactors, Tafazzin (TAZ) and Yes-Associated Protein 1 (YAP1), have shown to be elevated and form complexes in various cancers such as colon cancer promoting its survival and also in rhabdomyosarcoma where myoblast-like cells fail to differentiate into post-mitotic, multinucleated myotubes and fully mature muscle fibers [36]. Although TBX5 is known to be essential in the formation of the forelimb bud and it is expressed in all the major stages of forelimb development [9,37,38], but its implications in muscle diseases have not been as extensively studied as consequent cardiac disorders. To explore TBX5a role in muscle pathology and to physiologically, we studied TBX5a expression in both Forelimb (FL) and Hindlimb (HL) of WT mice in various stages of embryonic development (using homemade TBX5a C-terminal Ab) by Immunohistochemistry. It was seen that TBX5a is expressed in mouse forelimb at E10.5 and its expression is increased by E13.5 and by E17.5 TBX5a expression around is increased especially in interdigit region of forelimbs (**Figure 3.6 A**). Expression of TBX5a in HL

starts at E11.5 and increases by E17.5 (**Figure 3.6 B**). We were further interested to examine whether *TBX5* levels were affected in a commonly studied muscle disorder, Duchenne muscular dystrophy (DMD); where the dystrophin gene is mutated effecting the structural stability of the skeletal muscle tissue resulting in muscle wasting and weakness. Mechanistically, the muscle cells lose their ability to fully differentiate into mature myotubes and over time the muscle loses its regenerative ability [39]. MDX mice are a popular animal model used for studying this muscle disorder. QPCR analysis was performed on RNA extracts of skeletal muscle tissue (forelimb and hindlimb) of MDX and WT mice to determine whether *Tbx5* levels are altered in muscle diseases. Levels of *Tbx5a* were significantly increased in muscle tissues of MDX mice compared to the wildtypes (**Figure 3.6 C**). Furthermore, the expression of some other genes such as *Myog* (maker of differentiation) showed decreased levels in skeletal muscle of MDX mice compared to the WT, indication of lowered differentiation in the muscle tissues of these mice. But the expression of *Mstn* was increased in MDX mice compared to WTs indicating less differentiation in these tissues (**Figure 3.6 C**). We carried out immunofluorescence (IF) on forelimb tissues of WT and MDX mice (age 6-9 weeks) and observed an increased expression of TBX5a and PAX7 (marker of satellite cells) (**Figure 3.6 D**). Further staining was done using IF assays showing an increased expression of MYBBP1a, KI67 (proliferation marker) and MSTN in MDX forelimb tissue compared to WT (**Figure 3.6 E-F**). This suggests that there is a correlation between raised TBX5a levels, and the proliferative state the dystrophic muscle tissues experience as a result of dystrophin defects. This could potentially imply the use of TBX5a as a diagnostic marker in earlier stages of DMD.

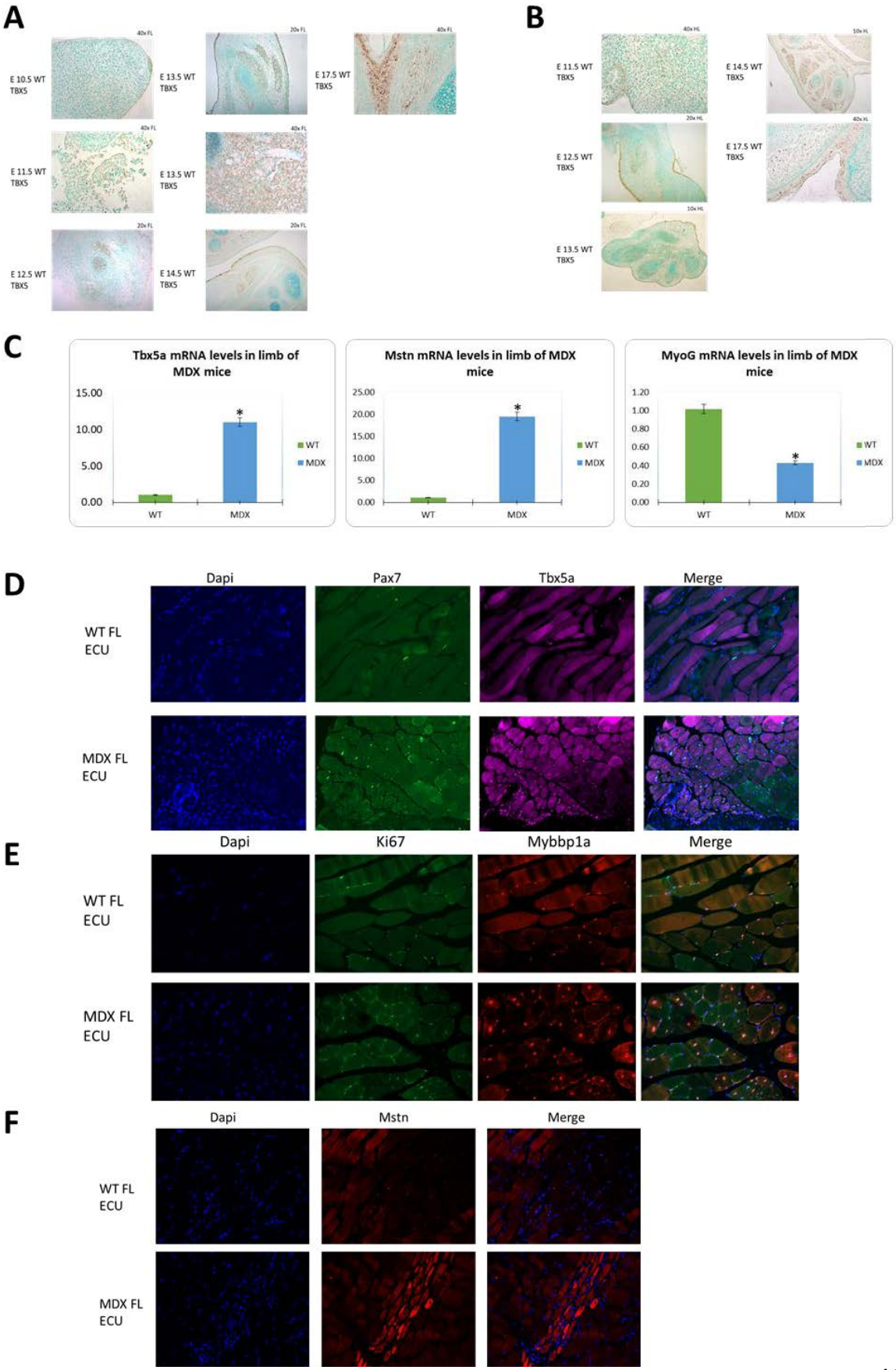

A
E 10.5 WT TBX5
E 13.5 WT TBX5
E 17.5 WT TBX5
E 11.5 WT TBX5
E 13.5 WT TBX5
E 12.5 WT TBX5
E 14.5 WT TBX5
B
E 11.5 WT TBX5
E 14.5 WT TBX5
E 12.5 WT TBX5
E 17.5 WT TBX5
E 13.5 WT TBX5
C
Tbx5a mRNA levels in limb of MDX mice
15.00
10.00
5.00
0.00
WT
MDX
Mstn mRNA levels in limb of MDX mice
25.00
20.00
15.00
10.00
5.00
0.00
WT
MDX
MyoG mRNA levels in limb of MDX mice
1.20
1.00
0.80
0.60
0.40
0.20
0.00
WT
MDX
D
Dapi
Pax7
Tbx5a
Merge
WT FL ECU
MDX FL ECU
E
Dapi
Ki67
Mybbp1a
Merge
WT FL ECU
MDX FL ECU
F
Dapi
Mstn
Merge
WT FL ECU
MDX FL ECU

Figure 3.6 TBX5a expression profile increases in MDX mice a model of DMD compared to adult WT. A. TBX5a expression pattern in forelimb of WT mice at different embryonic stages E10.5-17.5. **B.** TBX5a expression pattern in hindlimb of WT mice at different embryonic stages E11.5-17.5. Immunohistochemistry staining using TBX5a antibody (blue dapi and brown indicating positive staining). **C.** TBX5a, Myostatin levels show to have increased expression in MDX limb tissues compared to WT. Whereas, Myogenin levels are decreased in MDX mice. *Normalized to Gapdh. N=3 for WT and MDX mice. **C.** Immunofluorescence on forelimb Extensor Carpi Ulnaris (ECU) tissue of WT and MDX stained for TBX5a (Purple), PAX7 (Green), Dapi (Blue). **D.** Immunofluorescence on forelimb Extensor Carpi Ulnaris (ECU) tissue of WT and MDX stained for Ki67 (Green), MYBBP1a (Red), Dapi (Blue). **E.** Immunofluorescence on forelimb Extensor Carpi Ulnaris (ECU) tissue of WT and MDX stained for MSTN (Red) and Dapi (Blue) (images taken at 20x magnification).

3.10 Discussion

Over the past 20 years, *TBX5* mutations were linked to abnormal cardiac and limb development in human and in animal models of disease [4,5]. However, the molecular mechanisms underlying TBX5 function and its precise role in limb development remain incompletely elucidated. Although various studies have shown a role for TBX5 in the initiation of limb bud outgrowth, most studies have focused on its skeletal expression and function. Little is known about TBX5 in skeletal myogenesis. In this report, we show that TBX5 plays a key role in regulating skeletal myogenesis and identify several novel interacting proteins that modulate its transcriptional activity. Moreover, we show that protein-protein interactions differentially modulate the activity of TBX5 isoforms and their regulation of target genes.

TBX5 is a dosage sensitive developmental regulator and subtle variations in protein levels or activity can dysregulate important pathways and lead to disease. Therefore, it is essential to identify TBX5 regulators and modifiers in all expressing cell types, including in skeletal muscle cells. Using protein immunoprecipitation and mass spectroscopy analysis, we have

identified novel TBX5 protein partners in C2C12 myoblast cells. This unbiased approach allows identification of unknown or unsuspected interactors and pathways. Using the actively proliferating C2C12 cell line, we isolated TBX5a protein complexes allowing identification of TBX5 interacting proteins in skeletal myoblasts. Several of the interactors obtained were previously identified components of general transcription machinery but novel proteins, such as MYBBP1a were also isolated from TBX5 complexes. Given our previous study showing the inhibitory role of TBX5a on C2C12 cell differentiation [6] and the reported negative regulation of myogenesis by MYBBP1a [25], we fully analyzed the physical and functional consequences of TBX5a and MYBBP1a interaction. MYBBP1a, a transcriptional coregulator and suppressor of C2C12 myoblast differentiation, potentiates TBX5a inhibition of MYOD-dependent Myogenin transcription. It is known that at the onset of muscle differentiation, MYBBP1a expression is downregulated resulting in co-repressor complex destabilization and the subsequent gene activation [40]. Our results suggest that TBX5a might be a component of the co-repressor complex (which also includes HDACs), through MYBBP1a interaction. Moreover, MYBBP1a was shown to be associated with tumor suppression in some cancers by repressing the activity of proto-oncogenes [41]. It would be interesting to test whether MYBBP1a interaction might be relevant to TBX5 function in some cancers such as rhabdomyosarcomas.

TBX5 exists as different splice isoforms including in C2C12 cells where TBX5a and TBX5c exert opposing effects on myogenesis. We found that both isoforms, which differ in the C-terminal domain outside the T-box, interact with MYBBP1a, but that the resulting interaction differentially regulate transcription of target TBX5 genes. This suggests that TBX5-MYBBP1a interactions recruit additional co-activators or co-repressors to the

transcriptional complexes via the TBX5 C-terminal domain. We have previously shown the existence of three important functional domains in TBX5, trans-activating domain (TAD), in addition to the T-box domain [1]. TBX5c which is truncated in the C-terminal lacks one of the TADs located between aa 400 and 518 which might affect its interactions and function with downstream targets. Moreover, our group had previously observed a similar growth factor dependent distribution of TBX5a and TBX5b in C2C12 myoblasts; where TBX5a was dominant in proliferating myoblasts while TBX5b-which has a truncated C-terminal, was present at a higher level in differentiated myotubes [1]. Misexpression of TBX5b in the proliferating C2C12 cells caused growth arrest and apoptosis [1]. Similarly, our attempts to generate a C2C12 stable cell line overexpressing TBX5c led to growth arrest. Further analysis of the role of TBX5c and TBX5b in cell cycle arrest and survival will require inducible Tbx5b-expressing transfectants. Notwithstanding the mechanistic uncertainties, the fact that TBX5 isoforms can bind similar co-factors further point to the importance of isoform ratios during myogenesis.

Of the several genes whose expression is regulated by TBX5a, activation of the *Mstn* gene (also known as *Gdf8*) is noteworthy. Myoblast differentiation involves multiple stages such as withdrawal from the cell cycle, activation of the myocyte differentiation program and myocyte maturation (elongation, alignment and fusion of myoblasts into myotubes and formation of myofibers) [42]. MSTN is a hormone produced by muscle cells which acts to inhibit muscle cell differentiation and growth [43]. The *Mstn* promoter harbors binding sites for muscle specific and inducible transcription factors such as CREB, MEIS1, FOXO1 and MEF2 which regulate MSTN transcription during myogenesis [44]. The upregulation of *Mstn* has been shown to associate with sarcopenia, a muscle degenerative condition resulting in

loss of skeletal muscle mass and strength that is observed in ageing and also other many other muscle degenerative diseases such as cancer, muscular dystrophies and pulmonary and cardiac disorders.[45]. Here we show that TBX5a can directly bind to this promoter through TBE sites present in the proximal region. TBX5a activates the promoter and this is further potentiated by the presence of MYBBP1a. Thus, the TBX5a-MYBBP1a-MSTN pathway could be a potential target for pharmacologic development for muscle wasting diseases. To date, various attempts to block or alter *Mstn* transcription such as using competitive inhibitors like follistatin, or using unprocessed MSTN, have suffered specificity challenges[45]. Cell-specific targeting might overcome off target effects in other tissues.

Since, *TBX5* is known to be associated with HOS, cancers, and some muscle myopathies, we examined its expression in muscular dystrophies which are characterized by impaired muscle differentiation and continuous muscle regeneration [30]. Examples of these inherited diseases are Duchene Muscular Dystrophy (DMD), Becker Muscular Dystrophy (BMD), and Facioscapulohumeral Muscular Dystrophy (FSHD) amongst the other more than 30 muscular dystrophies [30]. Studies have shown that impaired differentiation of muscle cells which causes defective fusion of myoblasts during myogenesis is one of the important obstacles in proper muscle regeneration in dystrophic muscle [46–48]. Our results show a link between increased *Tbx5a* levels in MDX skeletal muscle tissue, a commonly used mouse model for DMD. Along with TBX5a, levels of proliferation and satellite cell markers (Ki67 and PAX7 respectively) were also increased which indicated a regenerative state in the MDX muscle tissues. Our newly identified, TBX5 protein interactor, MYBBP1a and the TBX5 direct target gene, MSTN, were also significantly increased in the MDX forelimb tissues. These observations suggest that the TBX5 pathway is associated with and indicative

of onset of proliferation and regeneration in MDX skeletal muscle tissues. Regulation of MSTN in MDX mice was shown to be beneficial in the progression of disease for enhancing muscle mass in dystrophic tissues [49], highlighting its importance as a target in therapeutic interventions. Identification of TBX5a as a MSTN regulator is important in understanding the proper functioning of *Mstn* gene.

Our work collectively shows that TBX5a correlates, *in vitro* and *in vivo*, to inhibition of skeletal muscle differentiation and myoblasts proliferation. Moreover, it identifies TBX5 as an important upstream activator of skeletal muscle genes and a co-regulator of several myogenic pathways through its interaction with key proteins. This study provides a better insight on the mechanism of action of TBX5 by identifying components of the regulatory network and circuits involved in myoblast differentiation which is required to enhance our insight into the mechanism of pathogenesis. Further characterization of the role of the TBX5 in skeletal muscle proliferation and differentiation and identifying its targets and interacting partners will help elucidate the pathogenesis of skeletal muscle problems in HOS and other diseases associated with TBX5. Identification and characterization of novel interactors and targets of this essential gene involved in muscle development will be an opportunity to better understand its manifestation in disease. Currently muscle myopathies such as DMD have no definite cure but studying associated genes and their mechanisms will help with enhanced targeting of relevant genes and factors and creating combination therapies to better tackle the current problems patients have been facing.

3.11 References

1. Georges, R., Nemer, G., Morin, M., Lefebvre, C. & Nemer, M. Distinct Expression and Function of Alternatively Spliced Tbx5 Isoforms in Cell Growth and Differentiation. *Mol. Cell. Biol.* **28**, 4052–4067 (2008).
2. Heinritz, W., Shou, L., Moschik, A. & Froster, U. G. The human TBX5 gene mutation database. *Hum. Mutat.* **26**, 397–397 (2005).
3. Fahed, A. C. & Nemer, G. M. Genetic Causes of Syndromic and Non-Syndromic Congenital Heart Disease. *Mutat. Hum. Genet. Dis.* (2012) doi:10.5772/48477.
4. Basson, C. T. *et al.* Mutations in human cause limb and cardiac malformation in Holt-Oram syndrome. *Nat. Genet.* **15**, 30–35 (1997).
5. Li, Q. Y. *et al.* Holt-Oram syndrome is caused by mutations in TBXS, a member of the Brachyury (T) gene family. **15**, 9 (1997).
6. Yamak, A. *et al.* Novel Exons in the Tbx5 Gene Locus Generate Protein Isoforms with Distinct Expression Domains and Function. *J. Biol. Chem.* **290**, 6844–6856 (2015).
7. Steimle, J. D. & Moskowitz, I. P. TBX5: A Key Regulator of Heart Development. *Curr. Top. Dev. Biol.* **122**, 195–221 (2017).
8. Plageman, T. F. & Yutzey, K. E. Microarray analysis of Tbx5-induced genes expressed in the developing heart. *Dev. Dyn.* **235**, 2868–2880 (2006).
9. Agarwal, P. *et al.* Tbx5 is essential for forelimb bud initiation following patterning of the limb field in the mouse embryo. *Development* **130**, 623–633 (2003).
10. Ahn, D., Kourakis, M. J., Rohde, L. A., Silver, L. M. & Ho, R. K. T-box gene tbx5 is essential for formation of the pectoral limb bud. *Nature* **417**, 754–758 (2002).
11. Rallis, C. *et al.* Tbx5 is required for forelimb bud formation and continued outgrowth. *Dev. Camb. Engl.* **130**, 2741–2751 (2003).
12. Liu, Y. *et al.* Six1 Regulates MyoD Expression in Adult Muscle Progenitor Cells. *PLoS ONE* **8**, (2013).
13. Gallagher, J. M., Komati, H., Roy, E., Nemer, M. & Latinkić, B. V. Dissociation of Cardiogenic and Postnatal Myocardial Activities of GATA4. *Mol. Cell. Biol.* **32**, 2214–2223 (2012).
14. Messaoudi, S. *et al.* Endothelial Gata5 transcription factor regulates blood pressure. *Nat. Commun.* **6**, (2015).
15. Wang, J. *et al.* Convergence of protein kinase C and JAK-STAT signaling on transcription factor GATA-4. *Mol. Cell. Biol.* **25**, 9829–9844 (2005).
16. Yamak, A. *et al.* Cyclin D2 rescues size and function of GATA4 haplo-insufficient hearts. *Am. J. Physiol. Heart Circ. Physiol.* **303**, H1057-1066 (2012).
17. Waldron, L. *et al.* The cardiac TBX5 interactome reveals a chromatin remodeling network essential for cardiac septation. *Dev. Cell* **36**, 262–275 (2016).
18. Hara, Y. *et al.* Molecular characterization of Mybbp1a as a co-repressor on the Period2 promoter. *Nucleic Acids Res.* **37**, 1115–1126 (2009).
19. Tavner, F. J. *et al.* Molecular cloning reveals that the p160 Myb-binding protein is a novel, predominantly nucleolar protein which may play a role in transactivation by Myb. *Mol. Cell. Biol.* **18**, 989–1002 (1998).

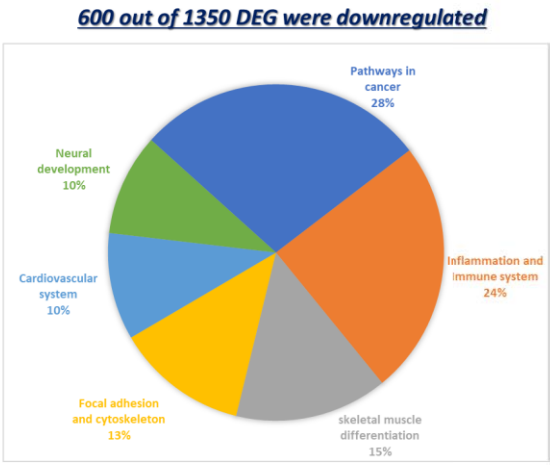
20. Fan, M. *et al.* Suppression of mitochondrial respiration through recruitment of p160 myb binding protein to PGC-1α: modulation by p38 MAPK. *Genes Dev.* **18**, 278–289 (2004).
21. Díaz, V. M. *et al.* p160 Myb-Binding Protein Interacts with Prep1 and Inhibits Its Transcriptional Activity. *Mol. Cell. Biol.* **27**, 7981–7990 (2007).
22. Owen, H. R. *et al.* MYBBP1a is a novel repressor of NF-kappaB. *J. Mol. Biol.* **366**, 725–736 (2007).
23. Jones, L. C., Okino, S. T., Gonda, T. J. & Whitlock, J. P. Myb-binding Protein 1a Augments AhR-dependent Gene Expression. *J. Biol. Chem.* **277**, 22515–22519 (2002).
24. Felipe-Abrio, B. & Carnero, A. The Tumor Suppressor Roles of MYBBP1A, a Major Contributor to Metabolism Plasticity and Stemness. *Cancers* **12**, 254 (2020).
25. Yang, C.-C. *et al.* Epigenetic silencing of myogenic gene program by Myb-binding protein 1a suppresses myogenesis. *EMBO J.* **31**, 1739–1751 (2012).
26. Ishibashi, J., Perry, R. L., Asakura, A. & Rudnicki, M. A. MyoD induces myogenic differentiation through cooperation of its NH2- and COOH-terminal regions. *J. Cell Biol.* **171**, 471–482 (2005).
27. Tapscott, S. J. The circuitry of a master switch: Myod and the regulation of skeletal muscle gene transcription. *Dev. Camb. Engl.* **132**, 2685–2695 (2005).
28. Langley, B. *et al.* Myostatin Inhibits Myoblast Differentiation by Down-regulating MyoD Expression. *J. Biol. Chem.* **277**, 49831–49840 (2002).
29. Bradley, L., Yaworsky, P. J. & Walsh, F. S. Myostatin as a therapeutic target for musculoskeletal disease. *Cell. Mol. Life Sci.* **65**, 2119–2124 (2008).
30. Saitoh, M., Ishida, J., Ebner, N., Anker, S. D. & Haehling, S. von. Myostatin inhibitors as pharmacological treatment for muscle wasting and muscular dystrophy. in (2017). doi:10.17987/jcsm-cr.v2i1.37.
31. Yarasheski, K. E., Bhasin, S., Sinha-Hikim, I., Pak-Loduca, J. & Gonzalez-Cadavid, N. F. Serum myostatin-immunoreactive protein is increased in 60-92 year old women and men with muscle wasting. *J. Nutr. Health Aging* **6**, 343–348 (2002).
32. Lenk, K. *et al.* Impact of exercise training on myostatin expression in the myocardium and skeletal muscle in a chronic heart failure model. *Eur. J. Heart Fail.* **11**, 342–348 (2009).
33. Costelli, P. *et al.* Muscle myostatin signalling is enhanced in experimental cancer cachexia. *Eur. J. Clin. Invest.* **38**, 531–538 (2008).
34. Spiller, M. P. *et al.* The myostatin gene is a downstream target gene of basic helix-loop-helix transcription factor MyoD. *Mol. Cell. Biol.* **22**, 7066–7082 (2002).
35. Rosenbluh, J. *et al.* β-catenin driven cancers require a YAP1 transcriptional complex for survival and tumorigenesis. *Cell* **151**, 1457–1473 (2012).
36. Mohamed, A. *et al.* The Hippo effector TAZ (WWTR1) transforms myoblasts and TAZ abundance is associated with reduced survival in embryonal rhabdomyosarcoma. *J. Pathol.* **240**, 3–14 (2016).
37. Hasson, P. *et al.* Tbx4 and Tbx5 acting in connective tissue are required for limb muscle and tendon patterning. *Dev. Cell* **18**, 148–156 (2010).
38. Schueller, M., Breeden, H., Vu, T. & Steadmon, K. R. Upper Limb Anomalies and Cardiac Defects: Presentation of Holt-Oram Syndrome and its Relevance to Pediatric Providers. 4.

39. Tenente, I. M. *et al.* Myogenic regulatory transcription factors regulate growth in rhabdomyosarcoma. *eLife* **6**,.
40. Kaspar, P. *et al.* c-Myb Inhibits Myoblast Fusion. *PLoS ONE* **8**, (2013).
41. Felipe-Abrio, B., Verdugo-Sivianes, E. M., Sáez, C. & Carnero, A. Loss of MYBBP1A Induces Cancer Stem Cell Activity in Renal Cancer. *Cancers* **11**, 235 (2019).
42. Mancini, A. *et al.* Regulation of myotube formation by the actin-binding factor drebrin. *Skelet. Muscle* **1**, 36 (2011).
43. Elkina, Y., von Haehling, S., Anker, S. D. & Springer, J. The role of myostatin in muscle wasting: an overview. *J. Cachexia Sarcopenia Muscle* **2**, 143–151 (2011).
44. Grade, C. V. C., Mantovani, C. S. & Alvares, L. E. Myostatin gene promoter: structure, conservation and importance as a target for muscle modulation. *J. Anim. Sci. Biotechnol.* **10**, 32 (2019).
45. White, T. A. & LeBrasseur, N. K. Myostatin and Sarcopenia: Opportunities and Challenges - A Mini-Review. *Gerontology* **60**, 289–293 (2014).
46. Jasmin, G., Tautu, C., Vanasse, M., Brochu, P. & Simoneau, R. Impaired muscle differentiation in explant cultures of Duchenne muscular dystrophy. *Lab. Investig. J. Tech. Methods Pathol.* **50**, 197–207 (1984).
47. Oexle, K. & Kohlschütter, A. Cause of Progression in Duchenne Muscular Dystrophy: Impaired Differentiation More Probable Than Replicative Aging. *Neuropediatrics* **32**, 123–129 (2001).
48. Merrick, D., Stadler, L. K. J., Larner, D. & Smith, J. Muscular dystrophy begins early in embryonic development deriving from stem cell loss and disrupted skeletal muscle formation. *Dis. Model. Mech.* **2**, 374–388 (2009).
49. Wagner, K. R., McPherron, A. C., Winik, N. & Lee, S.-J. Loss of myostatin attenuates severity of muscular dystrophy in mdx mice. *Ann. Neurol.* **52**, 832–836 (2002).

3.12 Supplementary Figures and Tables

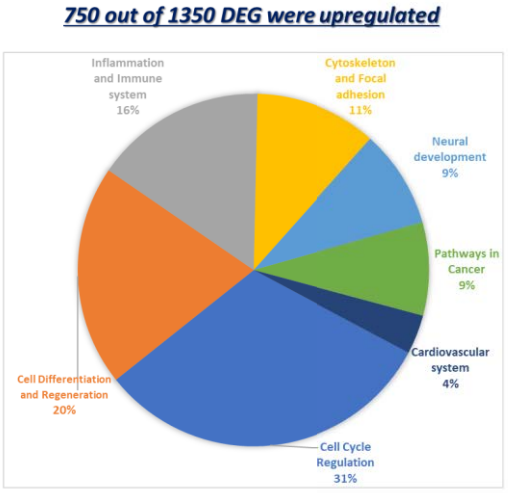
Downregulated TBX5a target genes in C2C12 classified by functional pathways

Selected functional pathways of Downregulated Tbx5a target genes	Differentially expressed genes in pathway
Pathways in Cancer	Cebpa, Smad3, Traf5, Bdkrb1, Bdkrb2, Cxcl12, Cxcr4, Dapk2, Ednra, Egfr, Fgf21, Fgfr1, Fzd8, Itga3, Jun, Lef1, Mmp2, Mmp9, Plcb4, Pdgfra, Pdgfrb, Rarb, Tgfb1, Tgfb3, Tgfbr2, Wnt10b
Inflammation and Immune system	Htr2b, Camk2a, Cyp2j12, Cyp2j13, Cyp2j6, Cyp2j9, Il1r1, Ngf, Cd40, Ets2, Nfatc4
Skeletal muscle differentiation	Jag1, Dll1, Thbd, Runx2, Gdf11, Tnnt3, Eml4, Fez1, Hook1, Map1lc3a, Srpx2, Tacc1, Ndrg2, Mmp19, Ngfr, Nrp2, Nxn, Tgfbr2, Sdc2, Tenm3, Zeb1, Sfrp4
Cytoskeleton and Focal adhesion	Nckap1l, Tiam1, Bdkrb1, Bdkrb2, Egfr, Fgf21, Fgfr1, Itga3, Itgb2, Itga10, Pdgfra, Pdgfrb, Vav3, Col1a1, Col3a1, Col5a2, Col6a1, Col6a2, Col6a3, Mmp9
Cardiovascular system	Cacna2d1, Itga3, Itga10, Sgcd, Sgcg, Tgfb1, Tgfb3, Gja1, Sgcd, Lef1
Neural Development	Ephb2, Ephb6, Cxcl12, Cxcr4, Efna2, Nfatc4, Sema7a, Sema3f, Sema4g, Ryr3, Kcnj2, Plcb4, Camk1g



Upregulated TBX5a target genes in C2C12 classified by functional pathways

Selected functional pathways of Upnregulated Tbx5a target genes	Differentially expressed genes in pathway
Cell Cycle Regulation	Bub1, Bub1b, Cks1b, Cks2, Chtf18, Ddias, Dscc1, Dsn1, E2f7, Fbxo5, Fancd2, Fanci, H2afx, Haus5, Hjurp, Mad2l1, Mis18bp1, Mis18a, Mtbp, Ndc80, Nek2, Nek3, Nsl1, Nuf2, Oip5, Racgap1, Sass6, Spc24, Spc25, Ticrr, Tpx2, Anln, Mki67, Aspm, Aurka, Aurkb, Birc5, Brca1, Brca2, Cdc20, Cdc25b, Cdc25c, Cdc45, Cdc6, Cdc7, Cdca2, Cdca3, Cdca5, Cdca8, Cenpe, Cenpt, Cep55, Chek1, Chek2, Chaf1a, Chaf1b, Cit, Clspn, Ccna2, Ccnb1, Ccnb2, Ccnd2, Ccne1, Ccne2, Cdk1, Cdk6, Ckap2, Dlgap5, Ect2, Esco2, Ercc6l, Fam64a, Fam83d, Foxm1, Gmnn, Hells, Incenp, Itgb3bp, Kif11, Kif18b, Kif20b, Kif23, Kif2c, Kifc1, Kntc1, Knstrn, Lig1, Melk, Mastl, Mcm8, Mcm2, Mcm3, Mcm4, Mcm5,Mcm6, Mcm7, Ncapd2, Ncaph, Ncapg2, Nasp, Nusap1, Nup43, Pard3b, Pttg1, Plk1, Plk3, Pmf1, Pkmyt1, Prc1, Rcc1, Rif1, Rbl1, Stk10, Sgol1, Spag5, Ska1, Ska2, Ska3, Spdl1, Smc2, Suv39h2, Timeless, Tipin, Tacc3, Ube2c, Uhrf1, Vrk1, Zwilch
Cell Differentiation and Regeneration	Mstn, Eid2, Htatip2, L1cam, Nme1, Racgap1, Btg2, Sox11, Atf3, Egr1, Mef2c, Myf5, Mylk2, Sox11, Vegfc, Gdf15, Angptl6, Cdc20, Cit, Cdk6, Ect2, Efna5, Fhl1, Gap43, Ostn, Pax8, Robo1, Sema4f, Sema3c, Sort1, Strbp, Stmn1, Suv39h2, Tenm4, Trip13, Zfp521
Inflammation and Immune system	Bub1b, Kras, Mad2l1, Ranbp1, Atf3, Cdc20, Chek1, Chek2, Ccnd2, Egr1, Xpo1, Fzd3, Pttg1, Pold1, Pole2, Pole3, Pole, Pcna, H2afx
Cytoskeleton and Focal adhesion	Cd44, Col1a2, Hmmr, Itga4, Itga6, Itgb6, Lamb3, Sv2b
Neural Development	Avpr1a, Calcrl, Gabre, Gipr, Glrb, Lpar4, Ptgir, P2ry1, Vipr2
Pathways in Cancer	Cd44, Kras, Brca1, Cdc25b, Cdc25c, Cdca5, Ccnd2, Ccne1, Ccne2, Cdk6, Ezh2, Erbb3, Ezr, Fzd3, Kif23, Marcks, Ptgs2, Stmn1, Zfpm2
Cardiovascular system	Itga4, Itga6, Itgb6



Supplementary Figure 3.1 Functional analysis of differentially upregulated and down regulated pathways between control pLNCX2 and 3xFlag TBX5a C2C12 cell lines as determined by Ingenuity Pathway Analysis (Qiagen) and literature review.

Identified Proteins	Accession Number	64 kDa band (peptides)	58kDa band (Peptides)
Non-POU domain-containing octamer-binding protein OS=Mus musculus GN=Nono PE=1 SV=3	Q99K48	61	83
Cluster of Vimentin OS=Mus musculus GN=Vim PE=1 SV=3 (P20152)	P20152 [2]	21	117
Vimentin OS=Mus musculus GN=Vim PE=1 SV=3	P20152	19	89
Desmin OS=Mus musculus GN=Des PE=1 SV=3	P31001	5	39
Neurofilament light polypeptide OS=Mus musculus GN=Nefl PE=1 SV=5	P08551	2	0
Heterogeneous nuclear ribonucleoprotein K OS=Mus musculus GN=Hnrnpk PE=1 SV=1	P61979	73	47
Prelamin-A/C OS=Mus musculus GN=Lmna PE=1 SV=2	P48678	79	11
Heterogeneous nuclear ribonucleoprotein U OS=Mus musculus GN=Hnrnpu PE=1 SV=1	Q8VEK3	37	32
Actin, aortic smooth muscle OS=Mus musculus GN=Acta2 PE=1 SV=1	P62737	19	20
Poly(U)-binding-splicing factor PUF60 OS=Mus musculus GN=Puf60 PE=2 SV=2	Q3UEB3	64	22
RNA-binding protein FUS OS=Mus musculus GN=Fus PE=2 SV=1	P56959	50	17
U4/U6 small nuclear ribonucleoprotein Prp31 OS=Mus musculus GN=Prpf31 PE=1 SV=3	Q8CCF0	15	53
Protein arginine N-methyltransferase 5 OS=Mus musculus GN=Prmt5 PE=1 SV=3	Q8CIG8	38	4
Cluster of Serotransferrin OS=Bos taurus GN=TF PE=2 SV=1 (Q29443)	Q29443 [3]	15	37
Serotransferrin OS=Bos taurus GN=TF PE=2 SV=1	Q29443 (+1)	12	28
Serotransferrin OS=Mus musculus GN=Tf PE=1 SV=1	Q921I1	3	11
Cluster of Probable ATP-dependent RNA helicase DDX5 OS=Mus musculus GN=Ddx5 PE=1 SV=2 (Q61656)	Q61656 [2]	21	15
Probable ATP-dependent RNA helicase DDX5 OS=Mus musculus GN=Ddx5 PE=1 SV=2	Q61656	20	15
Probable ATP-dependent RNA helicase DDX17 OS=Mus musculus GN=Ddx17 PE=1 SV=1	Q501J6	7	3
Plasminogen activator inhibitor 1 RNA-binding protein OS=Mus musculus GN=Serbp1 PE=1 SV=2	Q9CY58	4	24
Heterogeneous nuclear ribonucleoprotein Q OS=Mus musculus GN=Syncrip PE=1 SV=2	Q7TMK9	26	7
Plastin-3 OS=Mus musculus GN=Pls3 PE=1 SV=3	Q99K51	35	4
WD repeat-containing protein 1 OS=Mus musculus GN=Wdr1 PE=1 SV=3	O88342	12	18
Heterogeneous nuclear ribonucleoprotein L OS=Mus musculus GN=Hnrnpl PE=1 SV=2	Q8R081	26	17
Elongation factor 1-alpha 1 OS=Mus musculus GN=Eef1a1 PE=1 SV=3	P10126	15	13
26S proteasome non-ATPase regulatory subunit 3 OS=Mus musculus GN=Psmd3 PE=1 SV=3	P14685	19	13
Cluster of Histone-binding protein RBBP4 OS=Mus musculus GN=Rbbp4 PE=1 SV=5 (Q60972)	Q60972 [2]	4	25
Histone-binding protein RBBP4 OS=Mus musculus GN=Rbbp4 PE=1 SV=5	Q60972	3	20
Histone-binding protein RBBP7 OS=Mus musculus GN=Rbbp7 PE=1 SV=1	Q60973	2	13
Histone deacetylase 1 OS=Mus musculus GN=Hdac1 PE=1 SV=1	O09106	31	24
Protein RCC2 OS=Mus musculus GN=Rcc2 PE=1 SV=1	Q8BK67	2	15

Tubulin alpha-1A chain OS=Mus musculus GN=Tuba1a PE=1 SV=1	P68369 (+1)	6	15
Transcriptional repressor p66-beta OS=Mus musculus GN=Gatad2b PE=1 SV=1	Q8VHR5	20	0
Cluster of T-box transcription factor TBX5 OS=Mus musculus GN=Tbx5 PE=2 SV=2 (P70326)	P70326	10	7
T-box transcription factor TBX5 OS=Mus musculus GN=Tbx5 PE=2 SV=2	P70326	10	7
T-box transcription factor TBX4 OS=Mus musculus GN=Tbx4 PE=2 SV=3	P70325	0	1
Annexin A2 OS=Mus musculus GN=Anxa2 PE=1 SV=2	P07356	3	22
Ras GTPase-activating protein-binding protein 1 OS=Mus musculus GN=G3bp1 PE=1 SV=1	P97855	15	1
Cytoskeleton-associated protein 4 OS=Mus musculus GN=Ckap4 PE=1 SV=2	Q8BMK4	24	0
Cluster of ATP-dependent RNA helicase DDX39A OS=Mus musculus GN=Ddx39a PE=2 SV=1 (Q8VDW0)	Q8VDW0 [2]	6	11
ATP-dependent RNA helicase DDX39A OS=Mus musculus GN=Ddx39a PE=2 SV=1	Q8VDW0	6	10
26S protease regulatory subunit 4 OS=Mus musculus GN=Psmc1 PE=1 SV=1	P62192	0	27
Histone deacetylase 2 OS=Mus musculus GN=Hdac2 PE=1 SV=1	P70288	15	27
Myb-binding protein 1A OS=Mus musculus GN=Mybbp1a PE=1 SV=2	Q7TPV4	15	3
Apoptosis inhibitor 5 OS=Mus musculus GN=Api5 PE=1 SV=2	O35841	9	8
Pyruvate kinase PKM OS=Mus musculus GN=Pkm PE=1 SV=4	P52480	2	20
Stress-70 protein, mitochondrial OS=Mus musculus GN=Hspa9 PE=1 SV=3	P38647	13	0
Heterogeneous nuclear ribonucleoprotein M OS=Mus musculus GN=Hnrnpm PE=1 SV=3	Q9D0E1	10	5
26S proteasome non-ATPase regulatory subunit 4 OS=Mus musculus GN=Psmd4 PE=1 SV=1	O35226	4	15
SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily E member 1 OS=Mus musculus GN=Smarce1 PE=1 SV=1	O54941	1	16
TREMBL:Q3MHH8 (Bos taurus) Amylase, alpha 2B; pancreatic	Q3MHH8	10	8
U4/U6.U5 tri-snRNP-associated protein 2 OS=Mus musculus GN=Usp39 PE=2 SV=2	Q3TIX9	14	6
RNA-binding protein 39 OS=Mus musculus GN=Rbm39 PE=1 SV=2	Q8VH51	11	9
Ras GTPase-activating protein-binding protein 2 OS=Mus musculus GN=G3bp2 PE=1 SV=2	P97379	13	2
Cluster of Tubulin beta-5 chain OS=Mus musculus GN=Tubb5 PE=1 SV=1 (P99024)	P99024 [2]	6	15
Tubulin beta-5 chain OS=Mus musculus GN=Tubb5 PE=1 SV=1	P99024	6	14
Tubulin beta-4B chain OS=Mus musculus GN=Tubb4b PE=1 SV=1	P68372	4	11
T-complex protein 1 subunit alpha OS=Mus musculus GN=Tcp1 PE=1 SV=3	P11983	1	13
Alpha-1-antiproteinase OS=Bos taurus GN=SERPINA1 PE=1 SV=1	P34955	8	8
T-complex protein 1 subunit epsilon OS=Mus musculus GN=Cct5 PE=1 SV=1	P80316	1	13
LIM and SH3 domain protein 1 OS=Mus musculus GN=Lasp1 PE=1 SV=1	Q61792	3	11
U1 small nuclear ribonucleoprotein 70 kDa OS=Mus musculus GN=Snrnp70 PE=1 SV=2	Q62376	9	1
Cleavage and polyadenylation specificity factor subunit 7 OS=Mus musculus GN=Cpsf7 PE=1 SV=2	Q8BTV2	5	11

Pre-mRNA-processing factor 19 OS=Mus musculus GN=Prpf19 PE=1 SV=1	Q99KP6	0	13
RuvB-like 1 OS=Mus musculus GN=Ruvbl1 PE=1 SV=1	P60122	1	15
Splicing factor 3A subunit 3 OS=Mus musculus GN=Sf3a3 PE=2 SV=2	Q9D554	11	3
Paraspeckle component 1 OS=Mus musculus GN=Pspc1 PE=1 SV=1	Q8R326	17	0
Vitamin D-binding protein OS=Bos taurus GN=GC PE=2 SV=1	Q3MHN5	1	13
T-complex protein 1 subunit beta OS=Mus musculus GN=Cct2 PE=1 SV=4	P80314	0	11
Fascin OS=Mus musculus GN=Fscn1 PE=1 SV=4	Q61553	2	9
Bcl-2-associated transcription factor 1 OS=Mus musculus GN=Bclaf1 PE=1 SV=2	Q8K019	5	4
Clusterin OS=Bos taurus GN=CLU PE=1 SV=1	P17697	5	10
Recombining binding protein suppressor of hairless OS=Mus musculus GN=Rbpj PE=1 SV=1	P31266	2	8
DNA topoisomerase I, mitochondrial OS=Mus musculus GN=Top1mt PE=2 SV=1	Q8R4U6	10	0
tRNA-splicing ligase RtcB homolog OS=Mus musculus GN=Rtcb PE=2 SV=1	Q99LF4	0	12
Cluster of SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily D member 3 OS=Mus musculus GN=Smarcd3 PE=1 SV=2 (Q6P9Z1)	Q6P9Z1	3	12
SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily D member 3 OS=Mus musculus GN=Smarcd3 PE=1 SV=2	Q6P9Z1	3	12
SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily D member 1 OS=Mus musculus GN=Smarcd1 PE=1 SV=3	Q61466	0	2
Splicing factor U2AF 65 kDa subunit OS=Mus musculus GN=U2af2 PE=1 SV=3	P26369	6	6
High mobility group protein HMGI-C OS=Mus musculus GN=Hmga2 PE=1 SV=1	P52927	2	2
Microfibrillar-associated protein 1 OS=Mus musculus GN=Mfap1 PE=1 SV=1	Q9CQU1	11	0
Lamin-B1 OS=Mus musculus GN=Lmnb1 PE=1 SV=3	P14733	6	0
Myelin expression factor 2 OS=Mus musculus GN=Myef2 PE=1 SV=1	Q8C854	7	6
Cluster of DNA topoisomerase 2-alpha OS=Mus musculus GN=Top2a PE=1 SV=2 (Q01320)	Q01320	4	1
DNA topoisomerase 2-alpha OS=Mus musculus GN=Top2a PE=1 SV=2	Q01320	4	0
DNA topoisomerase 2-beta OS=Mus musculus GN=Top2b PE=1 SV=2	Q64511	3	1
Alpha-2-macroglobulin-P OS=Mus musculus GN=A2mp PE=2 SV=2	Q6GQT1	2	0
Cluster of Protein phosphatase 1B OS=Mus musculus GN=Ppm1b PE=1 SV=1 (P36993)	P36993	1	10
Protein phosphatase 1B OS=Mus musculus GN=Ppm1b PE=1 SV=1	P36993	1	10
Eukaryotic initiation factor 4A-I OS=Mus musculus GN=Eif4a1 PE=1 SV=1	P60843	5	3
TAR DNA-binding protein 43 OS=Mus musculus GN=Tardbp PE=1 SV=1	Q921F2	6	4
Beta-catenin-like protein 1 OS=Mus musculus GN=Ctnnbl1 PE=2 SV=1	Q9CWL8	3	0
Coronin-1C OS=Mus musculus GN=Coro1c PE=1 SV=2	Q9WUM4	0	8
Inosine-5'-monophosphate dehydrogenase 2 OS=Mus musculus GN=Impdh2 PE=1 SV=2	P24547	0	9
T-complex protein 1 subunit zeta OS=Mus musculus GN=Cct6a PE=1 SV=3	P80317	0	8

Replication protein A 70 kDa DNA-binding subunit OS=Mus musculus GN=Rpa1 PE=1 SV=1	Q8VEE4	8	0
Metastasis-associated protein MTA3 OS=Mus musculus GN=Mta3 PE=1 SV=1	Q924K8	7	10
Serpin A3-1 OS=Bos taurus GN=SERPINA3-1 PE=1 SV=3	Q9TTE1	3	4
26S protease regulatory subunit 6A OS=Mus musculus GN=Psmc3 PE=1 SV=2	O88685	0	7
Cytosolic purine 5'-nucleotidase OS=Mus musculus GN=Nt5c2 PE=1 SV=2	Q3V1L4	1	6
Luc7-like protein 3 OS=Mus musculus GN=Luc7l3 PE=1 SV=1	Q5SUF2	1	6
EH domain-containing protein 2 OS=Mus musculus GN=Ehd2 PE=1 SV=1	Q8BH64	7	0
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1 OS=Mus musculus GN=Rpn1 PE=1 SV=1	Q91YQ5	8	0
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 2 OS=Mus musculus GN=Rpn2 PE=2 SV=1	Q9DBG6	9	0
Heterogeneous nuclear ribonucleoprotein A3 OS=Mus musculus GN=Hnrnpa3 PE=1 SV=1	Q8BG05	2	4
Eukaryotic translation initiation factor 3 subunit D OS=Mus musculus GN=Eif3d PE=1 SV=2	O70194	5	1
T-complex protein 1 subunit theta OS=Mus musculus GN=Cct8 PE=1 SV=3	P42932	0	7
T-complex protein 1 subunit gamma OS=Mus musculus GN=Cct3 PE=1 SV=1	P80318	3	4
AP-2 complex subunit mu OS=Mus musculus GN=Ap2m1 PE=1 SV=1	P84091	0	5
TREMBL:Q1RMK2 (Bos taurus) IGHM protein	Q1RMK2	4	0
Septin-9 OS=Mus musculus GN=Sept9 PE=1 SV=1	Q80UG5	5	2
Nuclear pore complex protein Nup54 OS=Mus musculus GN=Nup54 PE=1 SV=1	Q8BTS4	0	7
SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily D member 2 OS=Mus musculus GN=Smarcd2 PE=2 SV=2	Q99JR8	9	4
H/ACA ribonucleoprotein complex subunit 4 OS=Mus musculus GN=Dkc1 PE=1 SV=4	Q9ESX5	0	5
EH domain-containing protein 1 OS=Mus musculus GN=Ehd1 PE=1 SV=1	Q9WVK4	7	0
ADP/ATP translocase 2 OS=Mus musculus GN=Slc25a5 PE=1 SV=3	P51881	3	1
Cluster of Heterogeneous nuclear ribonucleoprotein H OS=Mus musculus GN=Hnrnph1 PE=1 SV=3 (O35737)	O35737 [2]	1	6
Heterogeneous nuclear ribonucleoprotein H OS=Mus musculus GN=Hnrnph1 PE=1 SV=3	O35737	1	5
Heterogeneous nuclear ribonucleoprotein F OS=Mus musculus GN=Hnrnpf PE=1 SV=3	Q9Z2X1	0	2
Dihydropyrimidinase-related protein 2 OS=Mus musculus GN=Dpysl2 PE=1 SV=2	O08553	7	0
Antithrombin-III OS=Bos taurus GN=SERPINC1 PE=1 SV=2	P41361	0	4
60 kDa heat shock protein, mitochondrial OS=Mus musculus GN=Hspd1 PE=1 SV=1	P63038	5	2
Guanine nucleotide-binding protein G(s) subunit alpha isoforms short OS=Mus musculus GN=Gnas PE=1 SV=1	P63094 (+1)	5	0
Splicing factor 3A subunit 2 OS=Mus musculus GN=Sf3a2 PE=1 SV=2	Q62203	6	0
Transcription intermediary factor 1-beta OS=Mus musculus GN=Trim28 PE=1 SV=3	Q62318	0	7
Matrin-3 OS=Mus musculus GN=Matr3 PE=1 SV=1	Q8K310	4	0
Eukaryotic translation initiation factor 3 subunit L OS=Mus musculus GN=Eif3l PE=1 SV=1	Q8QZY1	5	0

U4/U6 small nuclear ribonucleoprotein Prp4 OS=Mus musculus GN=Prpf4 PE=1 SV=1	Q9DAW6	0	5
F-box-like/WD repeat-containing protein TBL1X OS=Mus musculus GN=Tbl1x PE=2 SV=2	Q9QXE7	0	5
Phenylalanine--tRNA ligase beta subunit OS=Mus musculus GN=Farsb PE=2 SV=2	Q9WUA2	5	0
Vitamin K-dependent protein S OS=Bos taurus GN=PROS1 PE=1 SV=1	P07224	3	2
Thyroid hormone receptor-associated protein 3 OS=Mus musculus GN=Thrap3 PE=1 SV=1	Q569Z6	2	2
Serine/arginine-rich splicing factor 7 OS=Mus musculus GN=Srsf7 PE=1 SV=1	Q8BL97	2	2
Actin-related protein 2/3 complex subunit 1B OS=Mus musculus GN=Arpc1b PE=2 SV=4	Q9WV32	1	2
Transcriptional enhancer factor TEF-1 OS=Mus musculus GN=Tead1 PE=2 SV=2	P30051	0	3
Nucleolar protein 58 OS=Mus musculus GN=Nop58 PE=1 SV=1	Q6DFW4	4	0
TGF-beta-activated kinase 1 and MAP3K7-binding protein 1 OS=Mus musculus GN=Tab1 PE=1 SV=2	Q8CF89	3	3
Regulator of chromosome condensation OS=Mus musculus GN=Rcc1 PE=1 SV=1	Q8VE37	4	2
Cluster of Chromodomain-helicase-DNA-binding protein 4 OS=Mus musculus GN=Chd4 PE=1 SV=1 (Q6PDQ2)	Q6PDQ2	0	3
Chromodomain-helicase-DNA-binding protein 4 OS=Mus musculus GN=Chd4 PE=1 SV=1	Q6PDQ2	0	3
ATP synthase subunit beta, mitochondrial OS=Mus musculus GN=Atp5b PE=1 SV=2	P56480	0	5
T-complex protein 1 subunit delta OS=Mus musculus GN=Cct4 PE=1 SV=3	P80315	0	5
Alpha-1B-glycoprotein OS=Bos taurus GN=A1BG PE=1 SV=1	Q2KJF1	3	0
Serine/threonine-protein kinase 38 OS=Mus musculus GN=Stk38 PE=1 SV=1	Q91VJ4	0	5
SAP30-binding protein OS=Mus musculus GN=Sap30bp PE=2 SV=2	Q02614	0	2
REST corepressor 1 OS=Mus musculus GN=Rcor1 PE=1 SV=2	Q8CFE3	2	0
Complement C3 OS=Bos taurus GN=C3 PE=1 SV=2	Q2UVX4	3	2
Cluster of Phosphatidylinositol-binding clathrin assembly protein OS=Mus musculus GN=Picalm PE=1 SV=1 (Q7M6Y3)	Q7M6Y3	5	0
Phosphatidylinositol-binding clathrin assembly protein OS=Mus musculus GN=Picalm PE=1 SV=1	Q7M6Y3	5	0
Nucleobindin-1 OS=Mus musculus GN=Nucb1 PE=1 SV=2	Q02819	4	0
Zinc finger Ran-binding domain-containing protein 2 OS=Mus musculus GN=Zranb2 PE=1 SV=2	Q9R020	3	0
Coronin-1B OS=Mus musculus GN=Coro1b PE=1 SV=1	Q9WUM3	3	0
Pleiotropic regulator 1 OS=Mus musculus GN=Plrg1 PE=2 SV=1	Q922V4	0	3
DNA methyltransferase 1-associated protein 1 OS=Mus musculus GN=Dmap1 PE=1 SV=1	Q9JI44	0	3
Parafibromin OS=Mus musculus GN=Cdc73 PE=2 SV=1	Q8JZM7	2	0
Aladin OS=Mus musculus GN=Aaas PE=1 SV=1	P58742	0	2
Protein disulfide-isomerase OS=Mus musculus GN=P4hb PE=1 SV=2	P09103	0	2
DNA primase large subunit OS=Mus musculus GN=Prim2 PE=1 SV=1	P33610	0	3
Probable ATP-dependent RNA helicase DDX6 OS=Mus musculus GN=Ddx6 PE=1 SV=1	P54823	0	3

Importin subunit alpha-4 OS=Mus musculus GN=Kpna3 PE=1 SV=1	O35344	0	3
Peroxisomal membrane protein PEX14 OS=Mus musculus GN=Pex14 PE=1 SV=1	Q9R0A0	3	0
Ran GTPase-activating protein 1 OS=Mus musculus GN=Rangap1 PE=1 SV=2	P46061	2	0
Polypyrimidine tract-binding protein 3 OS=Mus musculus GN=Ptbp3 PE=2 SV=1	Q8BHD7	0	5
Peptidyl-prolyl cis-trans isomerase-like 4 OS=Mus musculus GN=Ppil4 PE=1 SV=2	Q9CXG3	2	0
Thioredoxin domain-containing protein 5 OS=Mus musculus GN=Txndc5 PE=1 SV=2	Q91W90	2	0
mRNA cap guanine-N7 methyltransferase OS=Mus musculus GN=Rnmt PE=1 SV=1	Q9D0L8	0	2
Interleukin enhancer-binding factor 3 OS=Mus musculus GN=Ilf3 PE=1 SV=2	Q9Z1X4	2	0
Putative oxidoreductase GLYR1 OS=Mus musculus GN=Glyr1 PE=1 SV=1	Q922P9	2	1
Tropomyosin alpha-4 chain OS=Mus musculus GN=Tpm4 PE=2 SV=3	Q6IRU2	2	0
Protein disulfide-isomerase A3 OS=Mus musculus GN=Pdia3 PE=1 SV=2	P27773	0	2
Protein SSXT OS=Mus musculus GN=Ss18 PE=2 SV=2	Q62280	1	2
Stromal membrane-associated protein 1 OS=Mus musculus GN=Smap1 PE=1 SV=1	Q91VZ6	0	2
Complement C3 OS=Mus musculus GN=C3 PE=1 SV=3	P01027	2	0
THO complex subunit 3 OS=Mus musculus GN=Thoc3 PE=2 SV=1	Q8VE80	1	2
Cluster of Calcium/calmodulin-dependent protein kinase type II subunit gamma OS=Mus musculus GN=Camk2g PE=1 SV=1 (Q923T9)	Q923T9	0	3
Calcium/calmodulin-dependent protein kinase type II subunit gamma OS=Mus musculus GN=Camk2g PE=1 SV=1	Q923T9	0	3
Nucleosome assembly protein 1-like 1 OS=Mus musculus GN=Nap1l1 PE=1 SV=2	P28656	0	2
V-type proton ATPase catalytic subunit A OS=Mus musculus GN=Atp6v1a PE=1 SV=2	P50516	2	0
Hemopexin OS=Bos taurus GN=HPX PE=2 SV=1	Q3SZV7	0	2
Amine oxidase [flavin-containing] A OS=Mus musculus GN=Maoa PE=1 SV=3	Q64133	0	2
RNA binding protein fox-1 homolog 2 OS=Mus musculus GN=Rbfox2 PE=1 SV=2	Q8BP71 (+1)	0	2
Polyadenylate-binding protein 2 OS=Mus musculus GN=Pabpn1 PE=2 SV=3	Q8CCS6	0	2
CD2 antigen cytoplasmic tail-binding protein 2 OS=Mus musculus GN=Cd2bp2 PE=1 SV=1	Q9CWK3	0	2
Protein RTF2 homolog OS=Mus musculus GN=Rtfdc1 PE=2 SV=1	Q99K95	0	2
EH domain-containing protein 4 OS=Mus musculus GN=Ehd4 PE=1 SV=1	Q9EQP2	2	0
SWISS-PROT:P02768-1 Tax_Id=9606 Gene_Symbol=ALB Isoform 1 of Serum albumin precursor	P02768-1	0	9

Supplementary Table 3.1 Potential protein interactors of TBX5a in C2C12 cells identified by high performance liquid chromatography electrospray ionization tandem mass spectrometry (HPLC-ESI-MS/MS) according to peptide abundance. Note: both bands 64 kDa and 58 kDa were excised and sent for analysis and both contained TBX5 peptides.

Gene	TBX5Afc.logFC	TBX5Afc.adj.P.Val	Fold change
Dpt	-3.406364333	2.62E-06	0.094315302
A630076J17Rik	-3.038751667	2.69E-06	0.121687116
Unc93a	-3.027493	3.28E-05	0.122640467
Ryr3	-2.81771	1.05E-06	0.141835443
Cdh10	-2.790539	2.44E-06	0.144532015
Daam2	-2.761658333	1.35E-06	0.147454491
Car3	-2.746690333	6.90E-07	0.148992299
Gm5662	-2.695399667	4.83E-05	0.154384555
Cyp2j6	-2.632691	6.90E-07	0.161243063
Bdkrb2	-2.603715	1.61E-06	0.164514312
Prl2c2	-2.507619	7.86E-06	0.175845583
Cyp2j12	-2.487000333	1.97E-05	0.178376772
Bdkrb1	-2.446674667	1.31E-05	0.183433028
Col6a1	-2.437227	1.61E-06	0.184638204
4930451C15Rik	-2.417130333	1.44E-05	0.187228202
Gm9992	-2.400081	9.28E-06	0.189453934
Cdh11	-2.398437667	8.03E-05	0.189669858
Prl2c3	-2.360185333	2.42E-05	0.194766123
Kcnj2	-2.329095667	4.66E-06	0.199008828
Postn	-2.328468	1.30E-06	0.199095428
Grin2a	-2.231208	4.65E-05	0.212980315
Cxcr4	-2.212742	3.08E-06	0.215723911
Nppc	-2.160133	8.80E-05	0.223735641
Gm2022	-2.158357667	0.000142382	0.224011132
Phex	-2.129516	2.65E-06	0.228534519
Spon2	-2.109452667	2.61E-06	0.231734915
Mir872	-2.104023333	7.86E-06	0.232608652
Npr3	-2.087734667	5.00E-06	0.235249789
Havcr2	-2.079914333	4.15E-06	0.236528456
Gm2046	-2.063918333	0.000189441	0.239165578
Gm10554	-2.06023	1.06E-05	0.2397778
Sh2d1b1	-2.029714333	2.69E-05	0.244903563
Efemp1	-2.028957667	0.000333025	0.245032044
Prok1	-2.022260667	0.001103837	0.246172128

Adam12	-2.014197667	1.71E-05	0.247551798
Klk1b27	-2.009191333	0.001804455	0.248412326
Gm4027	-1.997647	1.71E-05	0.250408077
Hook1	-1.993929	3.45E-05	0.251054241
Enpp2	-1.993542	1.74E-05	0.251121595
Slc25a23	-1.972587	5.94E-06	0.254795729
Fmo1	-1.969494333	2.60E-06	0.255342514
Lbp	-1.965094333	1.79E-05	0.256122458
Dio2	-1.910838333	1.13E-05	0.265937967
Gm5423	-1.864248	0.003447935	0.274666335
Fbxo32	-1.862488333	1.35E-06	0.275001552
Crhr1	-1.846266667	6.60E-06	0.27811112
Ttll2	-1.837813	0.001058504	0.279745533
Igfbp3	-1.834316	5.10E-06	0.280424441
Plxdc1	-1.792562	6.76E-05	0.288658977
Cyp2j9	-1.786468667	7.89E-05	0.289880729
Caap1	-1.782352667	6.86E-06	0.290708938
Setbp1	-1.773355667	1.05E-05	0.292527535
Pdgfra	-1.765386333	1.71E-05	0.294147905
Gm21936	-1.760446333	9.96E-05	0.295156837
Enpp1	-1.750521	7.86E-06	0.297194434
Pi15	-1.72978	2.85E-05	0.30149793
Nptx1	-1.72864	8.39E-06	0.301736264
Lipg	-1.726641	2.14E-05	0.30215464
Il13ra2	-1.717599	0.000202994	0.304054322
Islr	-1.706926333	3.25E-06	0.306311973
Kcnk5	-1.690952333	7.86E-06	0.309722408
Chst11	-1.689971667	4.09E-06	0.309933012
Caap1	-1.660098	5.22E-05	0.316417654
Adamts16	-1.659157333	0.000126443	0.316624032
Fggy	-1.65744	2.61E-06	0.317001155
Mysm1	-1.640158667	7.88E-06	0.320821189
Vldlr	-1.638959667	0.000195902	0.321087929
Gpx3	-1.638055333	1.46E-05	0.321289261
Atp1b1	-1.634415333	3.38E-06	0.322100916

Angpt4	-1.633588	0.00042465	0.322285682
Piezo2	-1.631775333	1.72E-05	0.32269087
Cd53	-1.631125333	2.30E-05	0.32283629
Aebp1	-1.623573	4.17E-06	0.32453073
Tspan11	-1.608258667	7.87E-06	0.327994001
Colec12	-1.606603333	1.01E-05	0.328370554
Col1a1	-1.58834	6.80E-06	0.332553878
Htra3	-1.585648667	6.65E-06	0.333174833
Rxfp1	-1.580128	4.08E-05	0.334452214
Tgm2	-1.563829333	4.47E-06	0.33825207
Cyp2j13	-1.563675667	0.001575952	0.3382881
Add2	-1.549136667	1.52E-05	0.341714491
Slc16a11	-1.539054333	1.12E-05	0.344110941
Ndrg2	-1.534987	2.47E-06	0.345082448
Adam8	-1.525388667	5.01E-06	0.347385955
Nt5e	-1.524564333	1.58E-05	0.347584503
St8sia2	-1.524445	1.89E-05	0.347613255
Slc4a4	-1.523453667	4.17E-06	0.347852196
Atp1a3	-1.519608667	0.000125365	0.348780511
P2rx6	-1.519323	2.29E-05	0.34884958
Pdgfrb	-1.517984667	4.64E-06	0.349173344
Edil3	-1.516764	1.05E-05	0.349468905
Igfbp2	-1.516662667	7.86E-06	0.349493453
Adam19	-1.508342667	9.28E-06	0.351514799
Tek	-1.505210333	4.65E-05	0.352278826
Ngfr	-1.499232	0.000142671	0.35374165
Arhgap24	-1.471520667	6.80E-06	0.360602007
Adamts2	-1.466210333	1.89E-05	0.361931773
Kng2	-1.464461	0.00117389	0.362370897
Spock2	-1.454932333	2.56E-05	0.364772195
Slc7a2	-1.449231333	5.01E-06	0.366216492
Podnl1	-1.444306333	4.47E-06	0.3674688
Col6a2	-1.443013667	2.71E-05	0.367798202
Itga10	-1.441280333	5.36E-05	0.368240361
Gm5039	-1.426921	0.036210986	0.371923806

Eno2	-1.425535667	1.71E-05	0.372281114
P4ha3	-1.423768	2.36E-05	0.372737532
Inhbb	-1.422782333	3.26E-05	0.372992278
Nt5dc2	-1.420374667	2.69E-06	0.373615272
Smim4	-1.418644667	0.000241503	0.374063559
Cebpa	-1.417847333	4.21E-05	0.37427035
Csn3	-1.413581333	0.00016519	0.375378693
Ltbp2	-1.410743667	5.21E-05	0.376117759
Csf2rb	-1.402610333	0.00047648	0.378244148
Sema3f	-1.399904333	7.86E-06	0.37895427
Nkain1	-1.383945333	1.05E-05	0.383169509
Itgb2	-1.383094	2.25E-05	0.383395684
Gm12703	-1.378948	8.17E-05	0.384499066
Gm8979	-1.376875667	0.000116379	0.38505177
Cdk15	-1.374824667	0.000202081	0.385599566
Star	-1.373671667	0.000497738	0.38590786
Ift74	-1.373533	5.96E-06	0.385944954
Tmem86a	-1.371006333	1.52E-05	0.386621471
Tbx3	-1.369730333	5.02E-06	0.386963572
Kprp	-1.368536667	0.000329392	0.387283873
Gas1	-1.360766	8.39E-06	0.389375496
Csrp3	-1.358534667	2.25E-05	0.389978186
Aspn	-1.357966	0.000661988	0.390131934
Apln	-1.353949	1.71E-05	0.39121972
Crabp1	-1.351370667	0.000251166	0.39191952
Kcnn3	-1.347456667	0.0004101	0.392984232
Tgfbi	-1.338592	2.47E-05	0.395406365
Fez1	-1.336838667	1.05E-05	0.395887202
Anpep	-1.322248667	5.10E-06	0.399911129
Gm20412	-1.320437	0.000259347	0.400413633
Adamtsl1	-1.319972333	0.000172009	0.40054262
Ermp1	-1.311444	3.76E-06	0.402917396
Gm8979	-1.309842333	0.000348173	0.403364959
Pcdhb20	-1.299059333	6.81E-05	0.406391087
Jun	-1.296915333	5.01E-06	0.406995477

Arl4c	-1.296912667	1.05E-05	0.406996229
Evi2a	-1.287213333	4.43E-05	0.40974171
Tspan18	-1.275974667	2.03E-05	0.412946084
Igsf5	-1.270934	1.85E-05	0.414391409
Ednra	-1.270260333	6.62E-06	0.414584955
Larp1b	-1.26867	0.000104272	0.415042218
Brinp3	-1.266237333	0.000155872	0.415742651
Casp12	-1.265696333	4.43E-05	0.415898581
Gm12415	-1.264052667	1.29E-05	0.416372685
Ptpn22	-1.256751	6.86E-06	0.418485343
Col12a1	-1.255925333	1.13E-05	0.418724914
Smg6	-1.252330333	0.00038179	0.419769621
1810011O10Rik	-1.251421	7.08E-05	0.420034286
Csf2rb2	-1.251233	0.002949608	0.420089025
Rasl11b	-1.23915	0.000235708	0.42362217
Krt13	-1.230584333	1.74E-05	0.42614481
Slc1a3	-1.230436	0.000986027	0.426188627
Slc8a1	-1.227011333	0.000220741	0.427201514
Cped1	-1.210873667	0.00147013	0.432006922
Mcc	-1.195655667	3.54E-05	0.436587985
Plaa	-1.193976	1.02E-05	0.437096581
Satb1	-1.190216667	0.000185941	0.43823704
Mmp19	-1.183948	6.29E-05	0.440145371
Cap2	-1.17872	0.000902379	0.441743251
Alox5ap	-1.175809667	4.22E-05	0.442635275
Bcl6b	-1.175674	0.000767902	0.442676901
Mmp2	-1.173537667	8.67E-06	0.4433329
Gulp1	-1.172110667	3.58E-05	0.443771626
Gxylt2	-1.171246667	3.55E-05	0.444037472
Cd200	-1.168436667	7.73E-06	0.444903185
Fut11	-1.166571667	1.34E-05	0.445478692
Pid1	-1.164786667	0.003266773	0.44603021
Lce3c	-1.162975667	0.001108909	0.446590459
Gm2016	-1.162400333	0.00053561	0.44676859
Tenm3	-1.162195667	0.000235656	0.446831975

Prdm8	-1.154812667	7.38E-05	0.449124501
Rftn1	-1.153450667	4.16E-05	0.449548705
Mir5125	-1.153000667	0.000256558	0.449688948
Serpinb2	-1.152154333	0.000122466	0.449952828
Naip1	-1.150687	0.007375312	0.450410698
Dll1	-1.150327	0.000118542	0.450523105
Lmcd1	-1.147852667	0.000260817	0.451296449
Mctp1	-1.147144667	2.56E-05	0.451517977
Gm25107	-1.146706667	0.002230387	0.451655078
Rcn3	-1.14389	5.91E-05	0.452537734
Cxcl12	-1.140993333	1.51E-05	0.45344726
Gpr64	-1.140965	0.000486161	0.453456165
Pygo1	-1.140082667	0.000124026	0.453733578
Nrep	-1.135447667	0.000259002	0.455193649
Cd14	-1.134004333	3.05E-05	0.455649272
Myom1	-1.133327	0.000617227	0.455863245
Fgf21	-1.132590333	8.60E-05	0.456096077
Enox1	-1.131948333	0.000767902	0.456299085
Gfra1	-1.131829667	1.57E-05	0.456336619
Ifitm10	-1.131096	0.000103099	0.456568743
Nsg1	-1.129721667	1.34E-05	0.457003885
Il1rl1	-1.129162	0.000553679	0.457181205
Wnt10b	-1.126636	3.21E-05	0.45798238
Rrad	-1.125841333	5.73E-05	0.458234716
Klhl40	-1.12567	0.000634963	0.458289139
Camk2a	-1.123623	0.004040389	0.458939854
Gfra4	-1.12051	4.99E-05	0.459931209
Srpx2	-1.120167333	1.34E-05	0.460040464
Ephb6	-1.118030667	2.73E-05	0.4607223
Pcdhb21	-1.116545667	1.10E-05	0.461196776
Pmepa1	-1.115746	1.20E-05	0.461452482
Dapk2	-1.115571333	1.85E-05	0.461508354
Hs3st1	-1.112287667	1.80E-05	0.462559973
Prrx1	-1.110909	1.55E-05	0.463002215
Capn6	-1.105535667	4.21E-05	0.464729887

Foxred2	-1.105398	0.000661988	0.464774235
Rgcc	-1.105327	0.0005166	0.464797109
Ociad2	-1.103282	0.00012681	0.465456419
Tmod1	-1.100161333	1.59E-05	0.466464329
Dpy19l1	-1.098776	1.58E-05	0.466912462
Cfh	-1.096322667	0.001057349	0.467707132
Snx22	-1.096217333	0.001569788	0.467741281
Apob	-1.094098667	0.002437164	0.468428686
Zfp503	-1.094065	1.05E-05	0.468439618
Gm7897	-1.090064	0.000313191	0.469740536
Il1r1	-1.088837667	4.75E-05	0.470139999
Rarb	-1.086484667	0.001600845	0.470907411
Fndc4	-1.08494	0.000109465	0.471411873
Gjb4	-1.084523667	5.21E-05	0.471547933
Gm11166	-1.079185	0.003584979	0.47329612
Gm4070	-1.077908333	3.03E-05	0.473715133
Gm4070	-1.075793667	0.000124026	0.474410003
Adm	-1.071569667	0.002400931	0.475801041
Dlgap2	-1.070348	0.000281702	0.476204117
Steap3	-1.070214667	4.22E-05	0.47624813
Doc2b	-1.069801	0.001769939	0.476384705
Klhl41	-1.067997	1.62E-05	0.476980767
Slc9a2	-1.063539333	4.05E-05	0.478456831
Naalad2	-1.062473	0.000590282	0.478810601
Adamts7	-1.061465667	4.87E-05	0.479145038
Gm21319	-1.061398	0.003159921	0.479167512
Oaf	-1.059204333	3.01E-05	0.479896656
Bicc1	-1.058657667	0.000107897	0.480078534
Slc24a3	-1.057936	8.06E-05	0.480318739
Dnajc15	-1.057785333	1.52E-05	0.480368903
Dkk2	-1.054366333	9.82E-06	0.481508665
Thbd	-1.053962667	0.000101098	0.481643411
Trp53inp2	-1.050516333	9.33E-06	0.482795344
Ets2	-1.046563333	1.71E-05	0.484120022
Arhgef9	-1.046534	1.80E-05	0.484129865

Phospho1	-1.044125667	8.28E-05	0.484938713
Figf	-1.042464333	4.80E-05	0.485497465
Tmem182	-1.042308333	7.06E-05	0.485549965
Lrp4	-1.037476333	4.11E-05	0.487178937
Pcdhb18	-1.034545667	0.001536485	0.48816959
Lrfn3	-1.034330667	0.000820988	0.488242346
Chst8	-1.029359	0.003491656	0.489927779
Slc6a17	-1.028853333	2.37E-05	0.490099529
Gpc6	-1.027193333	0.00010006	0.490663774
Scn7a	-1.027183667	9.77E-05	0.490667062
Fzd8	-1.024346333	0.006949695	0.491633001
Mmp14	-1.020529667	9.33E-06	0.492935344
AU020206	-1.019771333	4.48E-05	0.493194517
Fam19a5	-1.018128333	4.92E-05	0.493756507
Bace1	-1.016202667	3.87E-05	0.494415999
Bmper	-1.014695	0.000641799	0.494932951
Lum	-1.010771333	0.005746737	0.496280841
Xdh	-1.008264	0.000414048	0.497144103
Btc	-1.007781333	0.000186228	0.497310455
Mmp9	-1.007087333	0.001604285	0.497549741
Ccl5	-1.005408	0.00422599	0.498129239
Gm23521	-1.003496	0.011979587	0.498789846
Cyp46a1	-1.002249667	0.001171908	0.499220933
Col3a1	-1.000931333	2.97E-05	0.499677329
Medag	-0.999462667	3.03E-05	0.50018626
Gm16033	-0.997464333	0.000131032	0.500879568
Gm10306	-0.996027667	0.015339297	0.501378603
Vav3	-0.993240667	1.75E-05	0.502348103
Gm26523	-0.992821667	0.004802724	0.50249402
Amot	-0.99099	0.000189441	0.503132399
Tnnt3	-0.988233333	4.08E-05	0.504094691
Ppm1j	-0.982629667	3.34E-05	0.506056486
Smarca2	-0.982432333	7.06E-05	0.50612571
Rin2	-0.979801	0.000147168	0.507049676
Nrn1	-0.977072333	0.000295038	0.5080096

Dab2	-0.975872333	1.02E-05	0.508432327
Apcdd1	-0.973216667	2.30E-05	0.509369095
Nckap1l	-0.972772	8.77E-05	0.509526116
Nfatc4	-0.971785667	0.000541725	0.509874585
Gramd1b	-0.971785333	0.000104092	0.509874703
Mmd	-0.968194333	2.85E-05	0.511145409
Rerg	-0.967380333	0.000533924	0.511433889
Rspo1	-0.965986	0.000386218	0.511928418
Pdk1	-0.963764	3.55E-05	0.512717484
Ccdc74a	-0.960231667	0.001161867	0.513974373
B3gnt9	-0.952145333	0.000159055	0.516863298
Fmo4	-0.951620333	0.00036123	0.51705142
Cox7a1	-0.951296333	0.000364571	0.517167552
Fndc1	-0.951222667	2.41E-05	0.51719396
Gdnf	-0.950945667	0.00012247	0.517293272
Pcdhb22	-0.950045333	0.000590282	0.517616197
Ahr	-0.949473	8.13E-05	0.517821582
Id2	-0.947334667	3.21E-05	0.518589655
Casp4	-0.947289667	0.000541725	0.518605831
Hspb2	-0.945641333	4.32E-05	0.519198697
Zfp52	-0.943662667	0.000249362	0.51991127
Kcnd1	-0.942941667	0.001621837	0.520171165
Mgat3	-0.942935	7.68E-05	0.520173569
LOC102638255	-0.938643667	0.003309257	0.521723142
Slc7a4	-0.936551667	0.003159921	0.522480223
Matn2	-0.936469667	3.53E-05	0.52250992
Tnip2	-0.935137333	0.00025338	0.522992683
Tmem181b-ps	-0.933600333	0.018309374	0.523550159
Slc43a1	-0.931769667	0.0018993	0.524214925
Eml4	-0.930047	1.13E-05	0.524841243
Ocln	-0.929120667	0.000805006	0.525178344
Clmp	-0.928167333	1.85E-05	0.525525497
Adora2b	-0.927480333	3.55E-05	0.525775808
Mir199a-1	-0.927433333	0.001621837	0.525792937
Runx2	-0.924837667	2.80E-05	0.526739784

Csad	-0.919572333	5.61E-05	0.528665713
Emilin1	-0.918205667	0.015313233	0.529166755
5830411N06Rik	-0.917721667	0.000542046	0.529344312
Plcb4	-0.916381	4.53E-05	0.529836449
F3	-0.91604	0.000103099	0.529961698
Gm17757	-0.914969333	0.00348637	0.530355144
Ephb2	-0.913825667	0.001190969	0.530775739
Sh3gl3	-0.912492	0.001593486	0.531266629
Whrn	-0.909597	0.000172009	0.532333772
Bcl6	-0.908486	3.99E-05	0.532743873
Gm17275	-0.908440333	0.002267976	0.532760736
Ecm2	-0.90711	0.000909211	0.53325223
Cyp2d22	-0.906036333	0.001171259	0.533649229
Hs6st1	-0.905368667	0.000414048	0.533896255
Tiam1	-0.903086	5.66E-05	0.534741667
Fyb	-0.898761667	0.002313875	0.536346906
Gata2	-0.898502333	0.000555619	0.536443326
Prkd3	-0.898235	1.65E-05	0.536542739
Tbx1	-0.896719	7.02E-05	0.53710684
Mrc2	-0.888632	9.01E-05	0.540126037
Hsd11b1	-0.887348667	0.026658105	0.540606714
Htr2b	-0.884014	0.000107768	0.541857725
Sdc2	-0.883043333	9.94E-05	0.542222418
AI480526	-0.882923333	0.00317046	0.54226752
Jade1	-0.882495	0.000156753	0.542428543
Fam132a	-0.876292333	0.000477222	0.544765659
6030419C18Rik	-0.875552333	0.000199739	0.545045157
Ptgfrn	-0.875149333	1.83E-05	0.54519743
Glrx	-0.873718	0.000209471	0.545738602
Wisp2	-0.872416667	0.000849314	0.546231089
Olfml2a	-0.872039333	0.001427771	0.546373973
Cers4	-0.871817667	0.000256558	0.546457929
Ccbe1	-0.871065667	0.000233368	0.546742842
Lrrc32	-0.866883	0.005701026	0.548330261
Cyp2f2	-0.865530333	0.000282632	0.548844615

Lphn2	-0.865499333	0.000248123	0.548856409
Rbp1	-0.862836333	0.011963076	0.549870451
Gpam	-0.861386667	6.81E-05	0.550423257
Lgr4	-0.860625333	2.43E-05	0.5507138
Slc7a7	-0.860307667	0.000122466	0.550835075
Kirrel3	-0.859403667	0.001048846	0.551180339
Fam20a	-0.858836667	0.000109619	0.551397004
Ypel2	-0.858031	0.000452343	0.551705015
Nynrin	-0.856206333	0.000352644	0.552403232
Sh3d19	-0.854799333	0.00012247	0.552942231
Elmod2	-0.85262	0.00050287	0.553778136
Adck1	-0.851594	0.000113404	0.554172106
Apobec2	-0.851183	0.009057461	0.554330003
Neurl3	-0.850637333	0.000144446	0.554539705
Map1lc3a	-0.848888	0.000200981	0.555212518
Mylip	-0.848608333	2.24E-05	0.555320156
Tapbp	-0.848425333	2.37E-05	0.555390601
Bnip3	-0.847803	7.38E-05	0.55563023
Ip6k3	-0.847234667	0.000820988	0.555849158
Nrp2	-0.844919333	0.000481964	0.556741938
Gm19589	-0.843401667	0.001135061	0.55732792
Abcb1b	-0.841804667	0.000647078	0.557945199
H2-Ke6	-0.840029333	0.000807968	0.558632211
Mb	-0.838191333	0.001531057	0.559344364
Adam23	-0.838110333	0.001559036	0.559375769
Pygm	-0.834199333	2.31E-05	0.560894238
Sulf2	-0.834107	7.06E-05	0.560930136
Fut8	-0.833777667	1.89E-05	0.561058198
Antxr2	-0.833446667	2.37E-05	0.561186937
Olfml3	-0.831422667	0.005981267	0.561974796
Sema4g	-0.829018667	0.009920201	0.56291201
LOC101055995	-0.826937333	0.001902786	0.563724692
Ppt2	-0.823358333	0.00010006	0.565124902
Uck2	-0.823119	0.00068827	0.56521866
C230013L11Rik	-0.822140667	0.005519762	0.565602081

Tgfbr3	-0.820153	0.002189789	0.566381874
Wtip	-0.819793333	0.000474354	0.566523092
Il33	-0.819138333	0.000610221	0.566780358
Gm16381	-0.819038667	0.034549352	0.566819514
Cacna2d1	-0.818823667	2.91E-05	0.566903992
Cd40	-0.817822333	0.001565286	0.5672976
Qsox1	-0.815727667	0.000119016	0.568121865
Ptgfr	-0.814137	0.002000736	0.568748602
Mir145a	-0.813998667	0.044766727	0.56880314
Kbtbd11	-0.813505	0.00078385	0.568997808
3300005D01Rik	-0.812819667	0.000277544	0.569268167
Ptplb	-0.812239333	8.74E-05	0.569497205
Abca1	-0.812161333	0.001835846	0.569527996
Six5	-0.811539	0.001052437	0.569773726
Arhgap18	-0.810879667	2.62E-05	0.57003418
Ptpn5	-0.810155667	4.22E-05	0.570320317
Eqtn	-0.809297	0.002110285	0.570659863
Galc	-0.809141667	0.000336731	0.570721308
Jag1	-0.808073	0.002509708	0.571144223
Muc13	-0.807888333	0.009005919	0.571217335
Gm17757	-0.807644667	0.003122911	0.57131382
Glul	-0.805393333	3.65E-05	0.572206054
Sgcd	-0.804797333	0.00169571	0.57244249
Camk1g	-0.803878667	0.002018757	0.572807121
Csgalnact1	-0.803623	0.000339352	0.57290864
Rgs4	-0.803421	0.025347789	0.572988862
Cnr1	-0.801243	0.004039718	0.573854542
Tap2	-0.800251	0.00648306	0.574249261
Traf5	-0.800023333	0.000165964	0.574339888
Atp6v0d2	-0.799934	0.000625548	0.574375453
Ush1g	-0.799049667	0.005984818	0.574727638
Gm25026	-0.797812667	0.032812357	0.575220634
Lancl3	-0.796941667	0.00090652	0.575568017
Mgat5	-0.795941667	0.001108909	0.575967109
Plekho2	-0.795037	3.00E-05	0.576328392

Adamts6	-0.793393	0.000536767	0.576985513
Svil	-0.792772333	0.000107768	0.577233793
2900026A02Rik	-0.791715333	0.000329392	0.577656862
Rdh5	-0.790445	0.000134729	0.578165729
Akr1c14	-0.789651333	0.025504464	0.578483881
Pim1	-0.785760333	0.000150978	0.580046179
Fgfr1	-0.785740333	0.000121327	0.58005422
Nxn	-0.784861667	0.000108616	0.580407607
Pisd-ps1	-0.783908333	0.024134964	0.580791267
n-R5s165	-0.78306	0.022272762	0.581132885
Cyp1b1	-0.781919667	0.000615313	0.581592405
Pcdh7	-0.781797333	0.000712149	0.581641723
Dhrs3	-0.781336333	7.02E-05	0.581827611
2010111I01Rik	-0.778584333	0.000337218	0.58293853
Pxdn	-0.777514667	0.000196239	0.583370902
Lhpp	-0.776939667	0.000442426	0.583603457
Snord123	-0.774524667	0.000644043	0.584581198
Ctla2a	-0.774417333	0.003484069	0.584624691
Deptor	-0.774028667	7.68E-05	0.584782212
Gdf11	-0.773503333	0.000204863	0.58499519
Efna2	-0.773411	0.000251166	0.585032631
Map3k13	-0.773100667	0.000928728	0.585158489
Vwa1	-0.772369	0.002534529	0.585455329
Maf	-0.770353	0.00024953	0.586274007
D030025P21Rik	-0.769320333	0.000147454	0.586693806
Zfp799	-0.768578333	0.000100714	0.586995629
Fam69a	-0.768374	0.000251166	0.587078773
Klhl5	-0.767698333	4.80E-05	0.587353788
Sphk1	-0.763819333	0.000337218	0.588935142
Cd33	-0.759221	0.000953828	0.590815262
Pcp4	-0.758589	0.001014033	0.591074137
Zbed6	-0.758213333	0.000155872	0.591228068
Zfp651	-0.758005667	0.003928966	0.591313178
Naglu	-0.757523333	5.27E-05	0.591510903
Mmp23	-0.757415667	0.000590282	0.591555049

Zbp1	-0.756682667	0.002492275	0.59185568
C2cd2	-0.756158333	0.000640603	0.592070824
Gm22	-0.755109333	0.000330083	0.592501482
Adarb1	-0.755060333	4.73E-05	0.592521606
Dand5	-0.754435667	0.000722166	0.592778215
Pgrmc2	-0.751033667	8.40E-05	0.594177687
Il6ra	-0.748637667	0.000338909	0.595165305
Zfp532	-0.748058667	0.002097726	0.595404212
Syngap1	-0.746171	0.000147115	0.596183767
Neat1	-0.743006667	0.026223957	0.597492841
Pla2g16	-0.741108333	0.000804826	0.598279555
Dsel	-0.740893667	0.002774815	0.598368583
Cx3cl1	-0.739320667	0.00046831	0.599021352
Fdxr	-0.739258	0.00115176	0.599047372
Tslp	-0.739077667	0.01045587	0.599122256
Emilin2	-0.738718667	0.000116379	0.59927136
1600020E01Rik	-0.737299	0.003536888	0.599861356
Grik5	-0.736845333	0.000162092	0.600050017
Enpp3	-0.736263	8.00E-05	0.600292272
Snx24	-0.735629667	0.000224491	0.600555854
Vaultrc5	-0.735629	0.000329392	0.600556131
Sgcg	-0.735386333	0.00163616	0.600657156
Lrp12	-0.735130333	0.000103099	0.600763749
Cd300lb	-0.734529	0.002207303	0.601014207
Tgfb1	-0.731892	4.73E-05	0.602113763
C330006D17Rik	-0.731507667	0.00037038	0.602274187
Dpyd	-0.730110333	0.002540536	0.602857807
Plcl2	-0.729712667	6.77E-05	0.603024003
Sh3bp2	-0.729404	0.001908393	0.603153034
Rab31	-0.72747	7.84E-05	0.603962131
Kat6b	-0.726053333	0.000713797	0.604555488
Hspa12a	-0.724025667	0.00054962	0.605405771
Mr1	-0.723939333	0.000545274	0.605442001
Gm12747	-0.721273	0.029235624	0.60656199
Snx29	-0.720551667	0.000371647	0.606865341

Bcl3	-0.720458	0.002454512	0.606904743
Pdgfrl	-0.718932667	0.00016822	0.607546751
Tcf4	-0.718523333	0.000111761	0.607719153
Ucp2	-0.717695667	0.000530351	0.608067899
Fgfrl1	-0.716012667	0.000185941	0.608777664
Chrdl1	-0.715766667	0.030128084	0.608881478
Slc6a2	-0.715161333	0.016995185	0.60913701
Tmprss11b	-0.714797	0.032443352	0.609290858
Clec2d	-0.714088333	0.003654921	0.609590222
Zfp30	-0.712931	0.001247045	0.610079433
Syt17	-0.712650333	0.00029503	0.610198131
Rbm24	-0.711748	0.000803599	0.610579899
Mtmr11	-0.709757333	0.000300528	0.611422974
Smad3	-0.709566	9.69E-05	0.611504067
Egfr	-0.709482333	0.000329708	0.611539532
Plekho1	-0.708333	0.000762741	0.612026913
Sfxn5	-0.706791333	0.001280042	0.612681276
Afap1l2	-0.706016667	0.000480523	0.613010348
Tgfb3	-0.705937	0.000217632	0.6130442
Fam131b	-0.705850333	0.00312894	0.613081028
Vps37b	-0.704946667	0.003305815	0.613465167
Sfrp4	-0.704337667	0.004331054	0.613724181
Tmem229b	-0.704182667	0.001850121	0.613790122
Zeb1	-0.704144	6.95E-05	0.613806573
Ccl8	-0.702816667	0.001801097	0.614371558
S1pr3	-0.702314333	7.15E-05	0.614585514
Igsf3	-0.700673	0.000227163	0.615285117
2310015B20Rik	-0.699855667	0.00491063	0.615633794
Gm24775	-0.698545333	0.038899716	0.6161932
Gm25518	-0.697923	0.045116848	0.616459064
Atp8b1	-0.696774333	0.003654921	0.616950081
Gm996	-0.696656333	0.001424239	0.617000544
Mir155	-0.693394333	0.00179871	0.618397189
Eya1	-0.691451	0.000346798	0.619230741
Osmr	-0.690785	0.000101206	0.619516667

1110054M08Rik	-0.688882333	0.00420359	0.620334241
Zfp319	-0.688795333	0.002392445	0.620371651
A330021E22Rik	-0.688348667	0.003885252	0.620563751
9430069I07Rik	-0.686984333	0.024390681	0.621150886
Epb4.1l1	-0.686186333	0.00010006	0.621494559
Lrrc15	-0.684756667	0.00616598	0.622110747
Gm25568	-0.684718667	0.029099674	0.622127133
Mir143hg	-0.684694333	0.00491063	0.622137626
Nme3	-0.683517	0.000162054	0.622645538
Gm10436	-0.683421	0.015382838	0.622686972
Tacc1	-0.682906333	0.000251166	0.622909149
Layn	-0.682616333	0.001951588	0.623034374
Arhgef25	-0.682030667	4.86E-05	0.623287348
Ltbp3	-0.680768333	5.27E-05	0.623832952
Col6a3	-0.680494	0.000344883	0.623951588
Itga3	-0.679709333	0.000231253	0.624291041
Tgfbr2	-0.678647667	9.69E-05	0.62475062
6430548M08Rik	-0.678277667	0.000113404	0.624910867
Desi2	-0.678139	0.001450209	0.624970934
Gm5945	-0.67804	0.041072945	0.625013822
Gm20471	-0.677153333	0.011879077	0.625398068
Arhgap26	-0.674557667	0.000861847	0.626524284
Selo	-0.674519667	9.42E-05	0.626540786
Dnase2a	-0.674330333	0.001359035	0.626623016
Col5a2	-0.67379	6.72E-05	0.62685775
Ninj1	-0.673427333	0.000295494	0.62701535
Hspa12b	-0.673227333	0.000452343	0.627102279
Amotl1	-0.670922667	0.00019796	0.628104858
Atp2a1	-0.670814667	0.000107897	0.62815188
Hdc	-0.67063	0.003987547	0.628232289
Socs5	-0.670033667	0.000142382	0.628492021
Sh3rf3	-0.668231667	0.009160438	0.62927753
Zfp105	-0.667069333	0.023039812	0.629784723
Pepd	-0.666775333	7.86E-05	0.629913077
Ranbp3l	-0.665963667	0.003131237	0.630267568

Usp54	-0.665185	0.000462043	0.630607835
Sil1	-0.665059667	0.000280019	0.630662621
Cxcl5	-0.665051333	0.000710446	0.630666264
BC034090	-0.664424333	0.000984609	0.630940413
Rnf150	-0.662599667	0.001103343	0.631738908
Trim44	-0.661784667	0.00012247	0.632095887
Scn5a	-0.660516333	0.001683606	0.632651834
P4ha1	-0.660385667	0.00025291	0.632709136
D130052B06Rik	-0.660351333	0.005281944	0.632724194
Neat1	-0.659683333	0.009521806	0.633017227
Mmp10	-0.659235333	0.013078289	0.633213828
Ank	-0.659069333	0.000204508	0.633286692
Tmem204	-0.658158333	0.004266764	0.633686711
Gm14635	-0.657821333	0.020789235	0.633834752
Slc35f5	-0.657677667	0.000647078	0.633897873
Neat1	-0.657056667	0.008974018	0.63417079
Fam69b	-0.655904	0.001133066	0.634677674
Ngf	-0.655446667	0.000150004	0.634878899
Slc43a2	-0.655397333	0.003084981	0.634900609
Sema7a	-0.655097	0.000204451	0.635032793
Bace2	-0.654434333	0.012326263	0.635324547
Gm10615	-0.654006667	0.044223444	0.635512908
Lef1	-0.653882667	0.001467064	0.635567533
Loxl1	-0.653240333	0.000205342	0.635850571
Gja1	-0.653163	0.001412102	0.635884655
Fmo2	-0.652686667	0.034935793	0.636094639
Pxylp1	-0.652398667	0.022243868	0.636221633
Gm2a	-0.652268667	0.000127161	0.636278965
Gm9982	-0.650715	0.001976088	0.636964556
Galnt18	-0.650323333	0.005985507	0.637137504
Cda	0.551904667	0.005118812	1.466019879
Scara3	0.552184333	0.004071673	1.466304095
Prpf31	0.553024	0.001045597	1.467157751
Cep83	0.553125667	0.004679102	1.467261145
Efcab3	0.55314	0.001450209	1.467275723

Ppp1r14a	0.553498	0.001565286	1.467639867
Gins4	0.554013	0.00068827	1.468163865
Xpo1	0.554358667	0.000824409	1.468515677
Mtfr2	0.554777	0.01107496	1.468941559
Arntl2	0.554823667	0.012004775	1.468989075
Htatip2	0.555335667	0.00662611	1.469510499
Erbb3	0.555436667	0.001244069	1.46961338
Fstl3	0.555945667	0.000647404	1.470131969
Mthfd2	0.557071667	0.000212195	1.471279831
Clhc1	0.558369	0.00836757	1.472603464
LOC102643284	0.558819333	0.004835941	1.473063205
Ano2	0.558952333	0.033313857	1.47319901
Shmt2	0.55896	0.000213624	1.473206839
4930523C07Rik	0.559190333	0.003825579	1.473442063
Fyn	0.559415	0.000150577	1.473671535
Tyms	0.560795	0.001634214	1.47508184
Hist1h2bq	0.561419667	0.008005783	1.475720668
Hist1h2bq	0.561419667	0.008005783	1.475720668
Pole3	0.562264333	0.001514217	1.476584924
A930003A15Rik	0.562309	0.014312713	1.47663064
Fam149a	0.564271	0.004548882	1.478640157
Syne2	0.56499	0.024756788	1.479377255
Cnksr3	0.565441	0.00016519	1.479839794
Rnd3	0.565933	0.000865701	1.480344548
LOC102641333	0.566832333	0.030715469	1.481267639
Aknad1	0.567019333	0.020072796	1.481459651
Isg15	0.567481667	0.00309515	1.481934483
Fam213a	0.568224333	0.006579929	1.482697545
Cacna1b	0.568449333	0.001867832	1.482928802
Ppp1r15a	0.569440333	0.000834361	1.483947789
Rps12	0.56967	0.00209954	1.484184042
Add3	0.569854667	0.00032515	1.484374031
Exosc8	0.570860667	0.000371733	1.485409455
Sdpr	0.571153333	0.000590282	1.485710817
Tmem45a	0.571309667	0.008261684	1.485871821

Madd	0.572320667	0.013476594	1.486913443
Gm6921	0.572560333	0.01609904	1.487160476
Nop58	0.572736667	0.000445908	1.487342255
Asns	0.572871333	0.000302692	1.487481096
Stard5	0.573177333	0.000894833	1.487796628
Ada	0.574745667	0.022488706	1.48941487
Tonsl	0.575661	0.000631457	1.490360145
Naaa	0.575779333	0.000229036	1.490482393
Angptl6	0.576439333	0.037010903	1.491164411
Pcdh1	0.577704333	0.003240829	1.492472484
Gm15091	0.579233	0.006600206	1.494054732
Mrto4	0.579390667	0.004227603	1.494218021
Gm26225	0.579857667	0.04977134	1.494701777
Cox6a2	0.580887	0.000901938	1.495768597
Khdrbs3	0.581961667	0.001129863	1.496883213
Rif1	0.582700667	0.00211112	1.497650167
Rnu1b1	0.583133	0.003137007	1.498099036
Rnu1b1	0.583133	0.003137007	1.498099036
Rnu1b1	0.583133	0.003137007	1.498099036
Rnu1b1	0.583133	0.003137007	1.498099036
Rnu1b1	0.583133	0.003137007	1.498099036
Rnu1b1	0.583133	0.003137007	1.498099036
Rnu1b1	0.583133	0.003137007	1.498099036
Rnu1b1	0.583133	0.003137007	1.498099036
Rnu1b1	0.583133	0.003137007	1.498099036
Gm22303	0.583247333	0.000497886	1.498217764
4930500J02Rik	0.584553333	0.002459125	1.49957464
Xylb	0.585307333	0.005514305	1.500358572
2010310C07Rik	0.585607	0.031514682	1.500670249
Smpdl3a	0.587731333	0.009797027	1.502881577
Myrfl	0.588223	0.0166145	1.503393842
Ptcd3	0.588971	0.004227603	1.504173515
Gm23468	0.589027333	0.000775906	1.50423225
Lrrc4c	0.589739333	0.002097726	1.504974803
Snord1c	0.590937	0.041081151	1.506224691

Trim16	0.591051667	0.000326131	1.506344412
Efhd1	0.591792667	0.000469594	1.507118302
Cdca3	0.592077	0.042516474	1.507415362
Snhg1	0.592969	0.001810028	1.508347666
Wfs1	0.593909333	0.00133682	1.509331111
Ccdc77	0.593950667	0.00127441	1.509374354
Slfn3	0.594543	0.006164717	1.509994192
Selp	0.594574	0.022482383	1.510026638
Jakmip1	0.59657	0.032034019	1.512117239
Anp32b	0.596934333	0.002905719	1.512499152
Gm16352	0.597153667	0.028531724	1.512729115
Sema3c	0.597282333	0.000445199	1.512864034
Frem1	0.597819333	0.003286694	1.513427257
Rpa1	0.598751	0.000301925	1.514404917
Cit	0.598852667	0.000389178	1.51451164
Ddx39	0.599256667	0.001782056	1.514935811
Nup160	0.600589333	0.001486906	1.516335855
Snord1c	0.601267667	0.019770701	1.517048981
Mamdc2	0.601722667	0.001152532	1.517527506
Neto2	0.602059333	0.047296892	1.517881677
Mylk2	0.602449333	0.003596921	1.518292058
Mdn1	0.603031	0.016125157	1.518904327
Col18a1	0.603477	0.007721092	1.519373959
Rpl32	0.604206667	0.004588842	1.520142602
Pop4	0.6043	0.000122731	1.520240949
Strbp	0.604473667	0.002161066	1.520423961
Mef2c	0.604643	0.006973847	1.520602428
Mfap3l	0.606249333	0.000217064	1.522296449
Nek3	0.606403667	0.002488876	1.522459306
Btbd19	0.607625333	0.000442933	1.523749063
Gm8894	0.608845667	0.023792157	1.525038502
Gm24089	0.609030333	0.018509719	1.525233722
Gm24089	0.609030333	0.018509719	1.525233722
Rpl32	0.609076667	0.000633018	1.525282707
Rad18	0.609421333	0.000197005	1.525647147

Ctps	0.610905667	0.001566921	1.527217635
Trdn	0.611643	0.006515425	1.527998365
Adamtsl3	0.611749	0.00634696	1.528110637
Mtbp	0.613128	0.000620068	1.52957198
Ska2	0.613255667	0.002115233	1.52970734
Gpnmb	0.615127	0.001486135	1.531692826
BC055324	0.615961	0.001915476	1.53257853
Acy3	0.616398667	0.000660779	1.533043535
BC030867	0.616735333	0.005242547	1.533401327
Nup107	0.618888333	0.000230166	1.5356914
Crabp2	0.618967667	0.00053565	1.53577585
Fam111a	0.619200333	0.000237328	1.536023548
Nefl	0.619305	0.000324241	1.53613499
Hjurp	0.619683667	0.001662498	1.536538234
Col1a2	0.620482667	0.000219751	1.537389443
Tbx15	0.620573	0.000283893	1.537485708
Acer3	0.620868333	0.001303853	1.537800478
Etv1	0.621250667	0.002310181	1.53820807
Ccdc18	0.621402667	0.008396669	1.538370142
Runx1t1	0.621676667	0.021960594	1.53866234
Nup205	0.622698667	0.000222097	1.539752709
Cyb5r1	0.623125667	0.001826525	1.540208503
Cenpa	0.623549	0.000601943	1.540660517
Lrp11	0.624264667	0.020773784	1.54142497
Nuak2	0.626296	0.007757851	1.543596845
Rpf2	0.626343	0.000161056	1.543647133
Fat1	0.627945	0.001275018	1.545362184
Il22ra1	0.628291	0.003660535	1.545732851
Rasgrp3	0.628645	0.03423119	1.546112181
Zfp820	0.630242667	0.02980939	1.547825322
Epha1	0.630331333	0.010166071	1.547920452
Snora68	0.630945	0.035884146	1.548579018
1810008I18Rik	0.631553	0.023046797	1.549231779
Plat	0.631652667	0.00090472	1.549338809
Epb4.1	0.631880333	0.000968879	1.549583324

Hecw2	0.633126	0.00107554	1.550921859
Tamm41	0.634448667	0.002345451	1.5523444
Taf9b	0.635129667	0.00095716	1.553077331
Usp29	0.636912	0.00300626	1.554997219
Asap2	0.637117667	0.001194306	1.555218911
Gm22269	0.640178	0.016586235	1.558521438
Nme1	0.641729333	0.000122466	1.560198221
Rpl34	0.642321333	0.005085139	1.560838569
Thop1	0.642724333	0.000136461	1.561274632
1500012F01Rik	0.643832333	0.001010529	1.562474163
Tipin	0.644958333	0.016206024	1.563694124
Ifi44	0.645443667	0.00578204	1.564220251
Pabpc4	0.646137333	0.000381076	1.56497253
Gtpbp8	0.646868333	0.002532859	1.565765687
Dmxl2	0.647715667	0.001751324	1.566685573
Pdss1	0.648158	0.011332375	1.567165996
Anxa3	0.648423333	0.000113898	1.567454248
Snord1b	0.648582333	0.021811668	1.567627007
Zfp521	0.648984667	0.036320245	1.568064242
Hat1	0.650672	0.000193373	1.569899277
Eldr	0.650744667	0.003832421	1.569978352
Rnaseh2b	0.651148333	0.000336731	1.570417694
Fat4	0.65299	0.003209475	1.572423685
Mrps27	0.655287667	0.001908393	1.574929956
Gm25544	0.655503	0.045116848	1.575165043
Dkc1	0.656069	0.000545274	1.575783136
Prim2	0.656665667	0.000286206	1.576434979
Dsn1	0.656781667	0.000536767	1.576561738
Cdc25b	0.657171	0.000476642	1.576987254
Sgol2	0.657227	0.025488883	1.577048468
Tube1	0.657819333	0.001563085	1.577696097
Mcm6	0.658028333	0.000125305	1.577924671
Btg2	0.658074333	0.00047648	1.577974983
Tmem194	0.658578	0.001602322	1.578525974
Ezh2	0.660764	0.000205373	1.580919601

Ank2	0.661435	0.00211112	1.58165506
Sort1	0.661732	0.002076587	1.581980701
Gm10110	0.662285667	0.000339352	1.582587938
Cks1b	0.662418	0.002750774	1.58273311
Cks2	0.662583333	0.000487533	1.582914502
Ptgir	0.663611	0.001061823	1.584042452
Rnf144b	0.663722	0.001565286	1.584164332
Map3k7cl	0.663793333	0.001411235	1.584242662
Scarna8	0.664265	0.02231931	1.58476069
Dock11	0.664934333	0.000442933	1.585496105
Ppa1	0.666148	0.000168157	1.586830464
Sulf1	0.668313667	0.000108153	1.589214285
Cep128	0.670622333	0.008102093	1.591759454
B3gnt5	0.671086667	0.000672348	1.592271847
Ezr	0.672014667	0.000252582	1.59329639
Pask	0.672972333	0.000359582	1.594354378
Mpp6	0.673341667	0.000407916	1.594762588
Cep41	0.675170333	0.001470188	1.596785288
Mns1	0.676345333	0.009659425	1.598086316
Rgs7	0.676676	0.007554505	1.59845264
Sv2b	0.678054333	0.002090742	1.599980512
Dnph1	0.678109	0.003831402	1.60004114
Slc18b1	0.678208667	0.000256681	1.600151681
Cenpm	0.680026667	0.000934284	1.602169369
Rin3	0.681641333	0.001209969	1.603963524
Pcyt1b	0.682475	0.017873929	1.604890648
Lsm5	0.682488667	0.016876522	1.604905851
Hdac9	0.684154333	0.000286206	1.606759869
Pold1	0.684728667	0.000146737	1.607399643
Sass6	0.684818	7.15E-05	1.607499178
Creb5	0.685592667	0.001125018	1.60836257
Rab11fip1	0.686153	0.002201167	1.608987368
Rad54b	0.687020333	0.002073489	1.609954966
Pola2	0.688410333	6.74E-05	1.611506864
D830031N03Rik	0.688579333	0.0020491	1.61169565

Fam60a	0.689674	0.000149697	1.612919012
Nup85	0.690369667	0.000445199	1.613696949
Ccne1	0.690678	0.001936798	1.614041865
Panx1	0.690702	0.000238604	1.614068716
Dnajc12	0.69171	0.001735138	1.615196848
Pclo	0.691929	0.010726844	1.615442052
LOC102640619	0.692623	0.001014033	1.616219338
Pttg1	0.692696333	0.001284612	1.616301493
Peg3	0.693995333	0.000135065	1.617757464
Nxf3	0.694280667	0.000452343	1.618077452
Pdlim3	0.694743333	0.000150004	1.618596447
AK129341	0.695596333	0.020938644	1.619553732
Rrm1	0.695858667	0.000106056	1.619848251
Eogt	0.697026667	0.000101102	1.621160205
Rpa3	0.697649333	0.005053846	1.621860048
5031410I06Rik	0.697667333	0.000817905	1.621880284
Acot7	0.698248667	0.000112385	1.622533951
F11r	0.699488	0.000170415	1.623928372
Calml4	0.699703	0.008211461	1.624170399
Tmem209	0.700288667	0.000127026	1.62482987
Gatm	0.701151333	0.000166259	1.625801736
Fbxo16	0.701909	0.007044168	1.626655789
Gnpnat1	0.702134333	0.003940944	1.626909875
Ranbp1	0.703561	0.000106488	1.628519506
Rps11	0.705127667	0.014986024	1.630288926
Rcc1	0.705661667	0.001288871	1.630892473
Bcat1	0.705885333	0.000530351	1.631145337
Syt12	0.706176333	0.002361372	1.631474381
Gm22303	0.706657667	8.59E-05	1.632018789
Tmem8	0.706910333	0.000924342	1.632304638
Atp10d	0.707012	0.004759053	1.63241967
Ly75	0.707162667	0.001600845	1.63259016
Slc35d2	0.707286	0.003253803	1.632729733
Inadl	0.707464333	0.002998279	1.632931569
Klf4	0.708332	0.000125305	1.633913943

D430020J02Rik	0.708484667	0.008065782	1.634086854
Reep6	0.712011667	0.003192735	1.638086642
Slfn9	0.713389	0.041944219	1.639651262
Kras	0.713391	0.000538825	1.639653535
Nlrc5	0.714181	0.000287215	1.640551632
Lamb3	0.715572	0.000338909	1.642134162
Cd44	0.718568333	7.65E-05	1.645548255
Gm23130	0.718719333	0.000701333	1.645720496
Tnfrsf10b	0.719903667	7.15E-05	1.647072051
Gm8074	0.722402	0.001842597	1.649926777
Nqo1	0.723262333	0.000732803	1.650910984
Syk	0.723522	0.000423965	1.651208154
Acox2	0.723786333	0.002631662	1.651510719
Donson	0.724156333	0.004320427	1.651934327
Samd9l	0.724638333	0.007038306	1.652486326
Marcks	0.724738333	0.000141342	1.652600871
Mis18bp1	0.724806333	0.000903652	1.652678767
Ero1l	0.724860667	0.000517182	1.652741009
Slc38a3	0.725318	0.002057959	1.65326501
Gm25432	0.725727667	0.008396669	1.653734537
Mcm7	0.726725667	2.91E-05	1.654878921
Ska1	0.72691	0.003192735	1.655090379
Gm5424	0.727087667	0.00317046	1.655294215
Traip	0.727197333	0.001360463	1.655420047
D6Wsu163e	0.728279667	0.000116379	1.656662436
Rfc5	0.732888333	0.001208759	1.66196308
Dscc1	0.737683667	0.000398093	1.667496423
Cyp2d40	0.738798667	0.018826471	1.668785661
Fam132b	0.739593	0.00149418	1.66970473
Slc20a1	0.740731	0.000105	1.671022316
Ssu72	0.741256	0.030874828	1.671630515
Sorbs2	0.742800333	6.84E-05	1.673420871
Hist1h2bl	0.743304	0.006283144	1.674005189
Rad51c	0.743444333	0.001014033	1.67416803
Dennd2d	0.744638	0.004227603	1.675553788

Egr1	0.745006667	0.000116379	1.675982014
BB557941	0.745539333	0.027957566	1.676600928
Haus5	0.746705667	0.000122777	1.677956909
Pmfl	0.747064667	5.01E-05	1.678374503
Pcna	0.748231333	8.40E-05	1.679732306
Zfp704	0.748431667	0.000131032	1.679965571
Gm16494	0.748751667	0.036210986	1.68033824
Rrm2	0.749648	5.20E-05	1.681382544
Gpatch4	0.749932333	0.000379849	1.681713951
Adh7	0.750227333	0.00018518	1.682057861
Peg3os	0.750371333	0.00105509	1.682225761
Nap1l2	0.750560333	0.006815527	1.682446155
Fam83b	0.750767333	0.016646345	1.682687572
Spire2	0.750972333	0.000605205	1.682926691
Itgb3bp	0.752331667	0.000601943	1.684513122
Gpsm2	0.752421	7.02E-05	1.684617432
Vipr2	0.755229333	0.000899304	1.687899882
Zgrf1	0.757082667	0.00683245	1.690069607
Clic5	0.757293	0.004297343	1.690316024
G630090E17Rik	0.757521667	0.001284612	1.69058396
Atp2a3	0.758096667	0.010184729	1.691257892
Tnfaip3	0.7582	0.001488871	1.691379033
Cenpt	0.758384333	0.000432249	1.691595155
2700099C18Rik	0.762041333	0.001482967	1.695888516
Ifi204	0.765693	0.049239131	1.700186488
Slc1a5	0.768163667	0.000248123	1.703100613
Mastl	0.768366333	5.44E-05	1.703339878
Gm7361	0.770518	0.033736268	1.705882171
5730508B09Rik	0.770582333	0.002261556	1.705958242
Orc1	0.770850333	0.002147613	1.706275176
Dgkg	0.771254333	0.017423815	1.706753054
Vrk1	0.771306333	6.81E-05	1.706814573
Sema4f	0.771357333	0.005479994	1.70687491
Fanci	0.771408	0.000110272	1.706934856
1700025F24Rik	0.774708333	0.000406745	1.710844139

Spc25	0.775446667	0.000672348	1.711719928
Hist1h2bn	0.775582333	0.000631278	1.7118809
Sgk3	0.775833333	0.001695835	1.712178759
Thyn1	0.776075333	0.000182967	1.712465987
Psph	0.776900667	0.000107768	1.71344593
Edem1	0.779244	7.83E-05	1.716231299
Cdc7	0.779316	0.001771194	1.716316952
1500012F01Rik	0.779788667	0.000163069	1.716879357
Ckap2	0.781892333	0.001112264	1.719384652
Atad5	0.782152333	0.004630882	1.719694544
AV320801	0.784085	0.001778442	1.72199983
Alg8	0.785599333	0.000119016	1.723808286
Il1rn	0.785780667	0.018358305	1.724024966
Cks2	0.788356333	4.38E-05	1.727105645
Cks2	0.788356333	4.38E-05	1.727105645
Cks2	0.788356333	4.38E-05	1.727105645
Gm15706	0.788737	0.018696397	1.727561416
Cenpf	0.790399	0.002264566	1.729552731
Ap1s3	0.791081	0.000905549	1.73037053
Gm14029	0.793785333	0.010212721	1.733617153
Terc	0.793871667	0.003686989	1.733720899
Fhl1	0.794955667	0.000124307	1.735024057
Topbp1	0.795588667	6.75E-05	1.735785487
n-R5s136	0.796221	0.01061959	1.736546449
Hist1h2an	0.796377	0.003940944	1.736734233
Cyp3a13	0.796502667	0.000875136	1.736885519
Rpa3	0.797901667	0.000722166	1.738570616
Dnajc9	0.798431	0.000106649	1.739208625
LOC102640775	0.799171333	5.79E-05	1.740101346
Ablim1	0.800182333	2.97E-05	1.741321188
Fam83d	0.801048667	0.001305335	1.742367159
Lsm3	0.802394667	0.000118298	1.743993504
Clcn3	0.805023667	0.000543449	1.747174453
Pfkp	0.805158	0.000106488	1.747337145
Prrg4	0.805179	4.29E-05	1.74736258

Gm6594	0.805897333	0.002778756	1.748232827
Chd11	0.807604667	0.002588818	1.750302968
Slc7a11	0.807708667	0.002649193	1.750429147
Arhgap11a	0.807888333	0.001175214	1.750647152
Gpr133	0.808532667	0.020938644	1.751429196
Rorb	0.809165	0.000328849	1.752197016
Efcab11	0.809447667	3.50E-05	1.752540357
Tmcc3	0.810471	0.000438064	1.753783911
Gprc5a	0.811472333	7.33E-05	1.755001585
Klf2	0.812051	0.000619179	1.755705659
Eid2	0.812464333	0.007916454	1.756208742
Kif14	0.813250333	0.001623686	1.75716581
Slc37a3	0.815921	0.000141997	1.760421626
Gm19980	0.816388333	0.008394562	1.760991973
Leprel1	0.817701	0.000424912	1.762594978
Podxl	0.819339	0.001979861	1.764597321
Plaur	0.819404	2.85E-05	1.764676826
Klf12	0.820205667	0.00826534	1.765657682
H2afx	0.82329	0.000324301	1.769436514
Ncapg	0.824076667	2.91E-05	1.770401608
A630033H20Rik	0.829516667	0.005558992	1.7770899
Ndc1	0.831537	9.80E-05	1.779580259
Plk4	0.834179	2.84E-05	1.782842181
Abcg2	0.835466333	0.000121327	1.784433742
Prc1	0.840212333	7.04E-05	1.790313618
Fabp5	0.840324333	0.000997665	1.79045261
Nudt12	0.841406	0.013020257	1.791795512
Aig1	0.843863333	7.08E-05	1.794850067
Sirpa	0.844091333	2.56E-05	1.795133743
Akap12	0.845399333	7.46E-05	1.796762015
Utp20	0.846137333	0.000204508	1.79768137
Rnf125	0.846698667	0.000590282	1.79838096
Sdr39u1	0.846861333	7.86E-05	1.798583743
Nasp	0.847419	0.000233368	1.799279111
Stxbp2	0.847490333	4.80E-05	1.799368077

Actg2	0.847728667	8.28E-05	1.799665358
AI662270	0.847951	0.013496008	1.799942725
Ncaph	0.848380333	5.67E-05	1.800478452
Krt19	0.848862667	0.002579748	1.801080503
Tbc1d9	0.850114333	0.00343661	1.802643779
Enpp5	0.850619667	0.00099418	1.803275302
Chek2	0.852085	0.000220741	1.805107804
Dut	0.852572667	6.74E-05	1.805718079
Serpinb6b	0.853301	0.000271805	1.806629911
Ccna2	0.853407667	0.002090278	1.806763491
Atf3	0.853791	2.42E-05	1.807243623
Gm12695	0.854552667	0.006274602	1.808198004
Gabre	0.855241	0.00325626	1.809060931
Osr1	0.855327	0.001304682	1.809168773
Anxa8	0.856075	0.000206619	1.810107024
Akna	0.856458	0.00220412	1.810587626
Prg4	0.857287	0.011032382	1.811628323
Slc7a3	0.864984667	0.002601524	1.82132031
Lmnb1	0.86509	0.000286006	1.821453292
Hist1h3g	0.867618	0.000929481	1.824647779
Elovl7	0.868947	3.63E-05	1.826329405
Cdk6	0.869705667	2.83E-05	1.827290065
Stk10	0.869746	0.00021989	1.827341152
Tacc3	0.869824667	0.000477455	1.827440795
Plag1	0.870997667	0.000970671	1.828927221
Lmbr1	0.872189667	0.000335563	1.830438963
Pde8a	0.872218667	4.80E-05	1.830475757
Tspan33	0.872452	0.000671593	1.830771832
Depdc1b	0.873422333	0.001553359	1.832003594
Diap3	0.873463	2.58E-05	1.832055235
Ptn	0.873842667	0.002147613	1.832537431
Gm20939	0.874559333	0.001600845	1.83344798
2700029M09Rik	0.87639	0.000288232	1.835775958
Gm6445	0.876792667	0.005684987	1.836288408
Itga6	0.877102	3.55E-05	1.836682175

Gzmd	0.877498667	0.017359141	1.837187238
Serpinb9f	0.877953667	0.016220724	1.837766745
Pole2	0.878690667	0.000179301	1.838705807
Knstrn	0.879185	0.000133626	1.839335939
Nup43	0.885784667	8.60E-05	1.847769331
Wdr62	0.887722	0.000308403	1.850252288
Calcrl	0.889375	6.22E-05	1.85237347
Vegfc	0.891596333	0.000172009	1.855227787
Rab39b	0.892331667	0.000675072	1.856173627
Troap	0.893063	4.88E-05	1.8571148
Mthfd1l	0.893161667	4.22E-05	1.857241813
Racgap1	0.893844333	3.50E-05	1.858120847
Nuf2	0.895125	0.000162404	1.859771015
Hist1h2af	0.895735	0.011879077	1.860557529
Dna2	0.895810333	0.001769939	1.860654685
Snx10	0.896497333	0.000407848	1.861540925
Cdca5	0.896510667	0.000249314	1.861558129
Ostn	0.896632	0.007239948	1.861714696
Cdc25c	0.897040667	0.000316725	1.862242132
Fzd3	0.898136667	0.005252356	1.863657395
Mug2	0.898279	0.00638317	1.863841268
Suv39h2	0.903463667	0.002995134	1.870551474
Nek2	0.903702	0.000372752	1.870860515
Lrrn1	0.907287	3.36E-05	1.875515258
Cdca2	0.911798667	6.64E-05	1.881389642
Avpr1a	0.913040333	0.000312683	1.883009572
Cpa4	0.914017333	0.001613531	1.884285187
Gnpnat1	0.914465667	0.000667955	1.88487084
Pkia	0.914699	4.88E-05	1.885175713
Cdc45	0.915201	0.000166259	1.885831792
Ube2c	0.916282667	0.000463536	1.887246233
Haus6	0.918110667	2.77E-05	1.889639027
Bmpr1b	0.918766667	0.009855253	1.89049845
Hist2h2ab	0.920337	0.000122466	1.892557325
Brca2	0.920945667	0.000736739	1.893355955

Kif4	0.921091333	0.000112804	1.893547134
Kif18a	0.921713	0.000696764	1.894363252
Wars	0.924987	2.81E-05	1.898667133
Gm24357	0.926491667	0.000147356	1.900648391
Kif20a	0.926679333	4.32E-05	1.900895645
Gm773	0.927592	0.000618883	1.902098555
Lig1	0.927628	0.000510354	1.90214602
Unc93b1	0.930212	0.001590544	1.905555992
Gm10021	0.931071667	0.037010903	1.906691804
Rragd	0.931538667	3.21E-05	1.9073091
Snhg12	0.932444333	0.000820988	1.908506809
Plk1	0.935005667	0.00010738	1.911898145
Nusap1	0.935021333	4.80E-05	1.911918907
Aspm	0.937643667	0.00038179	1.915397291
Stmn1	0.938706333	0.001108909	1.916808663
Rpa2	0.939342	8.95E-05	1.917653415
Angptl4	0.941595333	4.16E-05	1.920650922
Adamts1	0.943435333	0.000590282	1.923102065
Plk3	0.943898667	0.000102565	1.923719785
Kif22	0.950135667	0.000254575	1.932054334
Dync1i1	0.951022	0.00010006	1.933241674
Fam64a	0.951171	0.000252582	1.933441348
Chchd10	0.951628	0.000602501	1.934053898
Lpar4	0.953816333	0.00320665	1.936989768
Cdh5	0.960789	0.000736739	1.946374062
Spc24	0.961601667	1.51E-05	1.947470759
Depdc1a	0.962471333	0.00049615	1.948645062
Tmem144	0.971116667	5.05E-05	1.960357353
Arl6ip1	0.971557667	1.42E-05	1.960956682
Rad51ap1	0.976125333	0.008544461	1.967175038
Cdca8	0.976517667	2.77E-05	1.967710074
Pkmyt1	0.976572	3.59E-05	1.967784181
Scara5	0.980047	0.006812389	1.972529669
Ccnd2	0.980343333	7.08E-05	1.972934873
Soga3	0.98097	0.003137007	1.973792047

Alox5	0.981124	1.45E-05	1.974002751
Cdc20	0.982233667	0.000348173	1.975521663
Anln	0.983781667	0.000162082	1.977642519
Tpx2	0.984206	7.02E-05	1.97822428
Mcm8	0.984786333	0.000264285	1.979020193
Ptprk	0.985556333	3.59E-05	1.980076725
P2ry1	0.986167333	0.000701333	1.98091549
Hist1h2bn	0.987388	0.000725708	1.982592255
Kifc1	0.989327333	0.0012445	1.985259134
Ubash3b	0.989539667	1.13E-05	1.985551343
Rbl1	0.990626	6.49E-05	1.987047004
Dtl	0.990756333	8.00E-05	1.987226522
Gm22805	0.992373333	0.009987929	1.989455092
Armc2	0.992560667	0.000883285	1.989713439
Myf5	0.994176333	4.32E-05	1.991942956
Mapk11	0.994911667	9.28E-05	1.992958497
Npas2	0.995440667	0.000319152	1.993689399
Gtse1	0.997806	0.001377604	1.996960782
Cdca7	0.999460333	0.000109465	1.999252003
Arhgap19	1.000965333	0.000319385	2.001338684
Atad2	1.002412	0.000308403	2.003346539
Zfp462	1.006622667	0.000362031	2.00920207
Uchl1	1.016538333	0.000660466	2.023058914
Cenpq	1.020103333	0.001488871	2.028064215
Mphosph6	1.022566	2.23E-05	2.031529058
Serpinb9b	1.023351333	0.000122466	2.032635225
Aldh1l2	1.025009333	0.000349255	2.034972549
Fbxo48	1.025898	0.000798238	2.036226431
Lmnb2	1.026111667	0.000415101	2.036528024
Pcbd1	1.026614	0.000245796	2.037237248
Mcm4	1.027347667	1.83E-05	2.038273526
Slc43a3	1.029632333	4.38E-05	2.041503914
Paqr3	1.030511333	0.002140752	2.042748133
Fen1	1.031084	2.18E-05	2.043559147
Spdl1	1.034501667	0.000182554	2.048405967

Gmnn	1.034737	0.000248123	2.048740131
Gdf15	1.034869333	2.12E-05	2.048928064
Dlgap5	1.035076	6.08E-05	2.049221595
Aurka	1.035609667	2.31E-05	2.049979761
Tenm4	1.038282667	0.000699984	2.053781449
Ppl	1.039359	0.000640603	2.055314259
Nsl1	1.039734667	0.002018761	2.055849517
Brip1	1.041424667	0.000693846	2.058259189
Hmmr	1.042761	0.000601159	2.060166587
Parpbp	1.042981667	4.22E-05	2.060481723
Iqgap3	1.045259	0.000197005	2.063736818
Gm7903	1.045359333	0.000328849	2.063880347
Gm7903	1.045359333	0.000328849	2.063880347
Acot2	1.04708	7.02E-05	2.066343355
Cep55	1.047459333	3.55E-05	2.066886738
LOC102640004	1.053069667	0.000371527	2.074940068
Glrb	1.054497	0.000541725	2.07699393
Colec10	1.056493667	0.006065045	2.079870446
Steap1	1.058760667	8.73E-05	2.08314125
Blm	1.058943333	4.88E-05	2.083405024
2810417H13Rik	1.059893667	0.003512449	2.084777858
Neil3	1.060895333	0.00024382	2.086225827
Spag5	1.061091333	2.83E-05	2.086509274
Espl1	1.064088333	0.000106056	2.090848215
Ccnb2	1.067106667	0.000108176	2.095227161
Fanca	1.067227667	0.000174139	2.095402896
Robo1	1.068416667	0.000479113	2.097130539
Ar	1.068946333	0.000971832	2.097900614
Mis18a	1.069940333	1.71E-05	2.099346541
Shmt1	1.073805	0.000104656	2.104977774
Kif2c	1.074343	2.97E-05	2.105762894
Clspn	1.075881	0.000371647	2.108008961
Gen1	1.076023667	0.00016712	2.108217431
Ercc6l	1.076444667	7.89E-05	2.10883273
Top2a	1.078621667	6.74E-05	2.112017321

Ska3	1.080306	4.38E-05	2.114484522
Cdk1	1.081562	7.73E-06	2.116326179
Rfc4	1.081801667	5.28E-05	2.116677781
Kif23	1.081970333	4.72E-05	2.116925258
Bub1b	1.084338333	0.000101922	2.120402774
Ttk	1.085463333	0.000122731	2.122056889
Tigit	1.085717	0.000322114	2.12243004
Prr11	1.086244	5.08E-05	2.123205481
Casc5	1.087851667	0.001112771	2.125572793
Hddc2	1.090173333	2.14E-05	2.128996139
Chaf1a	1.091682	0.000259122	2.131223654
Aurkb	1.091717667	0.000227896	2.131276344
Pola1	1.091722	9.28E-06	2.131282745
Foxm1	1.091929333	0.000126879	2.131589059
Gins2	1.092103333	0.000150004	2.131846161
Mybl1	1.092976	0.000697488	2.133136076
Sox11	1.093736667	3.50E-05	2.134261077
Recql4	1.096297333	0.000107768	2.138052581
Elmo1	1.100144667	0.000104532	2.143761881
Gm15023	1.100921	0.000290517	2.144915778
Ckap2l	1.101656	7.97E-06	2.146008812
9430076C15Rik	1.102869333	0.003685986	2.147814404
2610318N02Rik	1.104455667	0.012513552	2.150177359
Mki67	1.104601667	0.001108709	2.150394967
Pxdc1	1.105218333	2.14E-05	2.15131433
Kif15	1.10579	7.97E-06	2.152166955
Nefm	1.109597667	0.000214821	2.157854615
Rnf113a2	1.112212	0.001164177	2.161768447
Efna5	1.116605	1.83E-05	2.168361054
Hmgb2	1.117590667	9.77E-05	2.169843011
Zim1	1.119059	3.77E-05	2.172052538
Mad2l1	1.119635333	1.74E-05	2.172920412
Gm23258	1.119734	0.009273736	2.173069024
Timeless	1.126135667	1.79E-05	2.182733002
Tk1	1.126408333	5.28E-05	2.183145573

Incenp	1.129926333	2.66E-05	2.188475652
Grb14	1.130468333	7.15E-05	2.189297986
Myo1d	1.131556667	7.88E-06	2.190950161
Gins1	1.133210667	2.97E-05	2.19346345
Naip2	1.13387	0.00184394	2.194466125
Dpep2	1.136913667	0.008272441	2.199100697
Oip5	1.143560667	0.00013452	2.209256099
Ncapg2	1.147856333	6.70E-05	2.215844026
Cth	1.150866667	0.001337152	2.220472443
Pax8	1.154816667	0.000696438	2.226560274
Hist1h2ag	1.157669	0.000285481	2.230966731
Igsf11	1.158417667	3.43E-05	2.232124761
C330027C09Rik	1.16163	3.36E-05	2.237100392
Strip2	1.163587	0.000743473	2.240137053
Cenpi	1.166143	5.36E-05	2.244109386
Vgll3	1.169167667	7.87E-06	2.248819185
Pdgfd	1.170871333	5.22E-05	2.251476366
Birc5	1.176076667	4.04E-05	2.259614505
Adh1	1.180787667	0.001826525	2.267005147
Ect2	1.181314	1.16E-05	2.267832362
Wdhd1	1.18159	0.000407916	2.268266259
Ncapd2	1.183223	1.72E-05	2.270835184
Brca1	1.187428667	0.000452343	2.277464659
Cenpe	1.190128	8.46E-05	2.281729864
Cenpk	1.193721667	0.000127161	2.287420601
Tspan7	1.194859	7.87E-06	2.289224576
Kif20b	1.195150333	0.000248123	2.289686901
Hist1h2bk	1.199175333	0.001010529	2.296083858
Kif18b	1.202497	2.58E-05	2.301376461
Hist1h2ab	1.203339667	0.000357277	2.302721069
Mms22l	1.205631667	0.000359582	2.306382294
Cyp4v3	1.209377	4.58E-05	2.312377597
Chtf18	1.210334667	0.000147454	2.313913072
Ccnb1	1.210885	3.43E-05	2.31479591
Uhrf1	1.210885	9.96E-05	2.31479591

Dck	1.211478	3.50E-05	2.31574757
Cd200r4	1.215119667	0.005318887	2.32160039
Gm5593	1.221983667	0.027442625	2.33267233
Pbk	1.22357	0.000196239	2.33523866
Rad54l	1.223883333	0.000583893	2.335745896
Kntc1	1.22514	0.0001569	2.337781346
Slamf9	1.231891667	0.000543449	2.348747566
Fancd2	1.235920333	7.50E-05	2.355315513
Pfn2	1.237133667	0.000203686	2.357297211
Ddah1	1.237512667	1.13E-05	2.357916561
Adamts5	1.238856	7.65E-05	2.360113104
Hist1h1b	1.240342333	0.000427207	2.362545859
Rad51	1.241687	0.000636741	2.364748901
Mfsd6	1.242508333	8.38E-06	2.366095547
Ndc80	1.253136667	9.51E-06	2.383590936
Smc2	1.253206	3.66E-05	2.38370549
Dhfr	1.253391333	0.000157407	2.384011728
Kif11	1.258847	6.76E-05	2.393044127
Zwilch	1.260686	8.39E-06	2.39609648
Tom1l1	1.263448	0.000218478	2.400688135
Dgat2	1.267220667	0.000249314	2.406974181
Sgol1	1.267345667	2.56E-05	2.407182739
Ddias	1.271367333	0.000181735	2.41390238
Esco2	1.278034667	2.80E-05	2.42508391
Trib3	1.27908	1.21E-05	2.426841689
AU018091	1.281912	9.63E-05	2.431610241
Fabp5	1.288013	3.01E-05	2.441915029
Mcm10	1.289649333	6.96E-05	2.444686269
Asf1b	1.292902333	8.71E-06	2.450204786
Tpd52l1	1.296820333	0.000275693	2.456867975
Cenph	1.304044667	0.000177636	2.469201661
Shcbp1	1.307277667	5.20E-05	2.47474121
Slc16a2	1.307431333	0.00017537	2.475004817
Stil	1.312244	3.54E-05	2.483274939
Fabp5	1.314040667	3.62E-05	2.486369423

Trip13	1.314548667	3.07E-05	2.487245074
Fam71f1	1.315167	0.00139185	2.488311326
Tlr3	1.320010667	0.000882955	2.496679557
2410012E07Rik	1.320647667	8.13E-05	2.497782171
Chek1	1.322283	5.91E-05	2.500615079
Gm13415	1.323977667	0.000407916	2.503554161
Fignl1	1.325456	0.000204155	2.506120875
Exo1	1.330369333	0.000116309	2.514670428
Polq	1.330777	0.001553359	2.515381107
Hells	1.331733667	7.06E-05	2.517049636
Melk	1.343766333	3.06E-05	2.538130649
Prim1	1.345602	1.01E-05	2.54136219
Serping1	1.347972	3.68E-05	2.545540466
AU021092	1.350394667	1.75E-05	2.549818693
Hephl1	1.352351333	2.59E-05	2.553279251
Tmlhe	1.356337	3.05E-05	2.56034283
Tinagl1	1.356762667	2.86E-06	2.56109837
Gpr141	1.362149333	0.000291251	2.570678752
Fbxo5	1.368139	2.69E-05	2.581373678
Plk2	1.382557667	8.71E-06	2.607301944
Lgals9	1.385055	1.49E-05	2.611819143
Pla2g4a	1.386493333	1.12E-05	2.614424364
Prkg1	1.394718	0.000239105	2.629371513
Bub1	1.396572333	1.11E-05	2.632753286
Rbpms	1.399913333	1.38E-05	2.638857293
Cenpn	1.406491	5.10E-06	2.650916081
L1cam	1.411632667	0.000227163	2.66038062
E2f7	1.421070667	3.44E-05	2.677841678
Cenpu	1.426493333	0.000147357	2.687925838
Serpinb9c	1.427776667	0.000345365	2.690317916
Ticrr	1.445645667	0.00028035	2.723846996
Mal	1.457372	9.16E-06	2.746076849
Itgb6	1.461312333	0.000182967	2.753587271
Gm24084	1.465829	0.000590606	2.762221476
Gap43	1.468465	1.71E-05	2.767273044

Abra	1.481425667	1.72E-05	2.792245258
Mcm2	1.485945	8.71E-06	2.801005857
Chaf1b	1.486172333	9.13E-06	2.801447261
Mybl2	1.490857667	1.16E-05	2.810560103
Nkain2	1.496134667	0.001527448	2.820859218
Pole	1.500756	5.91E-05	2.829909663
Tmc3	1.507775	6.74E-05	2.843711287
Cxadr	1.511708333	8.01E-06	2.8514749
Krt80	1.515043	5.42E-06	2.858073464
Mcm3	1.520308333	2.69E-06	2.868523492
Cdc6	1.521701	7.86E-06	2.871293881
Tnfsf15	1.539588333	0.000144084	2.907115384
Tubb3	1.565580333	2.58E-05	2.959965444
Gipr	1.577618667	8.39E-06	2.984767728
Mcpt8	1.588788333	1.45E-05	3.007966153
Nrg1	1.601755	5.42E-06	3.035123034
Mcm5	1.614291333	9.96E-05	3.06161173
Akr1b7	1.622221333	0.010253683	3.078486695
Fam198b	1.626655667	4.32E-05	3.087963429
Ccne2	1.683548	7.86E-06	3.212169449
Edn1	1.686047333	9.05E-05	3.217739053
Stk39	1.699587333	1.35E-06	3.248080376
Serpinb9e	1.729253333	0.000846189	3.31556177
Nov	1.768384667	1.30E-05	3.406723047
Fhdc1	1.776027667	4.71E-05	3.424818816
Itga4	1.81211	1.67E-05	3.511554923
Pard3b	1.818625667	5.50E-05	3.527450081
Gfpt2	1.841475333	0.000257609	3.58376325
Hhip	1.856319667	2.84E-05	3.620828051
Zdhhc15	1.870326333	3.03E-05	3.656152718
Lpl	1.943283667	6.86E-06	3.845799811
Cd74	1.962701333	3.58E-05	3.897911495
Prelp	1.975833333	5.10E-06	3.933553848
Tmem158	1.980468333	1.13E-05	3.946211645
Stmn2	2.035961667	2.69E-06	4.100959985

Gm13227	2.083904333	0.000525654	4.239529996
Slc18a1	2.107844333	1.74E-05	4.310467456
Sned1	2.119675333	2.51E-05	4.345961317
Hey1	2.132168667	1.80E-06	4.383759545
Ppp1r14c	2.148692667	5.91E-05	4.434257856
Ptgs2	2.181701333	2.14E-06	4.536882617
Stc1	2.211535333	5.21E-05	4.631679202
Pet2	2.211891333	4.07E-05	4.632822258
Zfpm2	2.354274667	4.64E-06	5.113370866
Cd24a	2.441994333	2.58E-05	5.433923795
Avil	2.555825	4.47E-06	5.880036049
Dthd1	2.588791333	2.47E-06	6.015944816
Cdsn	2.601579333	1.16E-05	6.069506983
Clu	2.701528667	8.27E-06	6.504908062
Mest	2.721640667	2.61E-06	6.596225257
Car6	2.817597667	7.86E-06	7.049874947
Gzme	2.819771667	0.000354643	7.060506425
Vgll2	2.851796667	8.22E-07	7.218988304
Kcnj8	2.891079	5.42E-06	7.418250574
Cldn1	3.02064	6.34E-06	8.115275091
Mstn	3.115740667	5.10E-06	8.668249423
Pappa	3.28207	6.28E-06	9.727506236
Plxdc2	3.583842667	4.77E-08	11.9906891
Tbx5	5.972114333	1.71E-07	62.77483095

Supplementary Table 3.2 All differentially expressed downstream genes of TBX5a in C2C12 cells from microarray analysis. Note: Transcriptomic analysis identified over 1200 genes which were considered significantly regulated if Log2 fold change ≥0.6 or ≤-0.6 and P≤0.05, Mann–Whitney test with P value adjusted—Benjamini and Hochberg—for multiple comparisons) with altered expression between the C2C12 pLNCX2 and C2C12 3xFlag TBX5 cell lines. Blue indicates downregulated target (listed from beginning of table up to the Cda gene) and red is upregulated (from Cda gene onwards).

4. Chapter III: Impaired Skeletal Muscle Repair in Mice Haploinsufficient for *Tbx5*

Massomeh Sheikh Hassani, Hiba Komati and Mona Nemer*

Molecular Genetics and Cardiac Regeneration Laboratory, Department of Biochemistry, Microbiology, and Immunology, University of Ottawa, Ottawa, Ontario, Canada.

*Corresponding author:
Dr. Mona Nemer
University of Ottawa
451 Smyth Road, Room 4522
Ottawa, Ontario K1H 8M5
Tel: 613 562 5800 ext. 3995
Email: mnemer@uottawa.ca

4.1 Statement of the Manuscript

The manuscript "Impaired skeletal muscle repair in mice haploinsufficient for *Tbx5*" is ready to be submitted (February 2020).

4.2 Contribution Statement

In this manuscript, MSH performed all experiments, analyzed results and wrote the draft manuscript. HK helped analyze and interpret results. MN designed experiments, interpreted results and edited the manuscript.

4.3 Acknowledgements

The authors recognize the technical support of Megan Fortier and Janie Beauregard. We thank members of the Nemer lab for helpful discussions, the uOttawa Histology Core Facility for in vivo staining and the Ottawa Hospital Research Institute for completion of the microarray experiments.

4.4 Sources of Funding

This work was made possible by grants provided by the Canadian Institutes for Health Research awarded to MN and by uOttawa tuition fee scholarships awarded to MSH.

4.5 Disclosures

None

4.6 Abstract

Skeletal muscle tissue is known for its ability to regenerate and the loss of this ability leads to debilitating disorders. Mutations in *TBX5* cause Holt-Oram Syndrome (HOS) characterized by limb and heart abnormalities. The function of TBX5 in heart and skeletal development is well investigated but its role in the skeletal muscle compartment has received little attention. We previously showed that TBX5 is abundantly expressed in C2C12 myoblasts where it has pro-proliferative activity. In this study we show that TBX5 is a direct activator of several genes involved in cell cycle progression including *Ccnd2*. Furthermore, we demonstrate cooperative interaction between TBX5a and GATA4 in promoting skeletal myoblast proliferation through regulation of common target genes. Importantly, we show that mice with heterozygous deletion of *Tbx5* have impaired skeletal muscle repair in response to Cardiotoxin (CTX)-induced injury, in part due to decreased number of muscle progenitor cells (PAX7+) at steady state. Together, the data suggest an important role for TBX5 in skeletal muscle regeneration. These findings provide molecular insights into skeletal muscle repair and may have implications for individuals with HOS or other muscle disorders.

4.7 Introduction

Skeletal muscle is a very organized and complex structure which serves multiple important functions in the organism, including body movement and metabolic regulation [1–3]. Skeletal muscle has a superior capacity of regenerating itself and maintaining homeostasis after injury, by partial recapitulation of embryonic muscle development [4]. Therefore, studies of

muscle regeneration have not only provided insights into the study of developmental biology but have been essential in understanding skeletal muscle pathogenesis in various muscular disorders. The process of muscle regeneration is governed by key regulators of the myogenic lineage and various molecular signaling pathways. The primary subset of cells which maintain skeletal muscle plasticity are the satellite cells located beneath the basal lamina of the muscle fiber [5]. Satellite cells are normally maintained in a quiescent state but become activated as a result of damage to the muscle. Once satellite cells are activated, they enter the cell cycle and expand a few rounds, before they differentiate and fuse with the pre-existing muscle fibers to reconstruct the injured area [6]. Satellite cells express Paired-box transcription factor 7 (PAX7) in both quiescent and activated state; whereas activated cells additionally express Myogenic differentiation factor 1 (MYOD); eventually proliferating myoblasts will either differentiate into myotubes with the downregulation of PAX7 and upregulation of MYOG or will exit the cell cycle and return to quiescence through MYOD downregulation [7,8]. Studies have demonstrated that there is a dynamic regulation involved in making the choice between proliferation and self-renewal or differentiation [9]. In addition to the PAX7, MYOD and MYOG, there are other factors involved in regulating muscle regeneration including growth factors and hormones (IGF-l, TGFβ, FGF, Vasopressin...) [3], and transcription factors such as Myogenic factor 5 (MYF5) and Myogenic regulatory factor 4 (MRF4) important for muscle proliferation and differentiation, respectively [10]. There are also several key regulators which are involved in promoting progression of the cell cycle and progenitor cell expansion such as cyclins and cyclin-dependent kinases; while other factors such as CDK inhibitors and pocket protein family like RB, p107 and p130, trigger irreversible interruption of the cell cycle leading to myogenic cell commitment to

differentiation [10,11]. Fibroblast growth factor 10 (FGF10), Insulin growth factor 1 (IGF1), Bone morphogenetic protein 4 (BMP4) and Notch signaling are amongst the signaling pathways that stimulate satellite cell proliferation and expansion [4].

It has been well established that the T-box transcription factor 5 (TBX5) is a major player in limb development in part through regulation of Fibroblast growth factor 10 (FGF10) expression [12]. Misexpression studies have also suggested a role for TBX5 in regulating limb identity and outgrowth [13]. Importantly, *TBX5* mutations cause Holt-Oram Syndrome (HOS), a disorder characterized by various heart defects and musculoskeletal abnormalities. Inactivation of one *Tbx5* allele in mice recapitulates the heart and limb defects observed in HOS patients [14,15]. Studies of TBX5 role in limb development have largely focused on the skeletal component. Previously, we found that TBX5 is present as multiple splice isoforms in C2C12 myoblasts and that TBX5a, abundantly expressed in myoblasts, promotes proliferation at the expense of differentiation [16]. TBX5a activates pro-proliferative signals involved in skeletal muscle development and regeneration such as *Igf1*, Fgf10, and *Bmp4* [16]. This pro-proliferative effect of TBX5a also extends to cardiac myocytes [16–19].

Given these findings and the importance of coordinated musculoskeletal development in limb morphogenesis, we further investigated the role and mechanism of action of TBX5 in skeletal muscle. Here we show that TBX5a promotes skeletal myoblast proliferation cooperatively with GATA4, by directly activating important cell cycle regulators like Cyclin D2 (*Ccnd2*). We show that TBX5 is present in postnatal skeletal muscle and its level is increased in the highly proliferative limbs of MDX mice. Importantly, we show that the

heterozygous loss of *Tbx5* in mice results in fewer PAX7+ satellite cells and impairs skeletal muscle regenerative ability in response to Cardiotoxin (CTX) induced injury. These results suggest that TBX5a is important in maintaining and promoting muscle progenitor expansion and proliferation and that TBX5a function in limb development involves coordinated regulation of both muscle and skeletal compartments.

4.8 Materials and Methods

Plasmids

The *Tbx5a* promoter-pCMV and *Gata4* promoter-pCMV constructs have been previously described [16,17]. To produce the retroviral pLNCX2 plasmid containing 3xFlag and V5-His tags, the 3xFlag sequence was subcloned into the pLNCX2 vector (Clontech, 631503) using the BglII and BamHI sites. Likewise, a V5His tag was subcloned between the BamHI and HindIII sites. To produce 3xFlag-TBX5a plasmid, the complete mouse TBX5a sequence was subcloned from the TBX5a-pcgn plasmid [16], and subcloned into the BamHI site between the 3xFlag and V5-His tags of the pLNCX2 construct described above. To produce the 3xFlag GATA4 pLNCX2 plasmid, the complete rat GATA4 sequence was subcloned from the GATA4-pcgn plasmid and subcloned into the BamHI in frame with the 3xFlag and V5-His constructs as described above. To produce the *Ccnd2* promoter-pXP2 construct, the first 1100 base pairs upstream from the transcription start site of the murine *Ccnd2* gene were cloned from genomic DNA and inserted into the pXP2 construct using the BamHI site. The *Ccnd2* promoter from mouse gDNA between -1100 and -1 bp from the transcription start site

using the primers and subcloned into the pXP2 plasmid at the BamHI site. All constructs were confirmed by sequencing. Primers available upon request.

Cell Culture, Transfections, Transductions and Differentiations

NIH 3T3 and C2C12 cells were maintained in DMEM supplemented with 10% fetal bovine serum and penicillin/streptomycin. To transfect cells using Calcium Phosphate, cells were plated 24 hours prior to transfection at a confluency of 50,000 cells/well in 12-well dishes and transfected as previously described [16]. Luciferase assays were completed as previously described [16]. To produce the C2C12 pLNCX2 (control) and 3xFlag TBX5a stable cell lines, the 3xFlag pLNCX2 and 3xFlag TBX5a pLNCX2 plasmids were transfected via Calcium Phosphate into AD293 cells. 48 hours post-transfection, media containing retroviral particles was collected and added to 80% confluent C2C12 cells in the presence of 10 uM hexadimethrine bromide (Sigma, TR-1003-G). 48 hours post-transduction, media was replaced with DMEM supplemented with 500 ug/ml G418 and was replaced every 48 hours (Sigma G8168). Overexpression of the 3xFlag-tagged construct was confirmed one week post-transduction by QPCR, western blot and immunofluorescence.

Microarray

Microarray analysis of the 3xFlag TBX5a, 3xFlag GATA4 and the respective control cell line was completed as previously described [20]. Functional analysis was completed using Ingenuity Pathway Analysis (IPA) by Qiagen (www.qiagen.com/ingenuity).

Western Blot

Western blot of nuclear and cytoplasmic extracts from C2C12 myoblasts overexpressing GATA4 was performed as previously described [16]. Western blots of nuclear extracts from cardiac myocytes or other cell lines overexpressing various GATA4 proteins were performed

as previously described. Anti-GAPDH (Abcam, Cambridge, UK, ab8245) and anti-GATA4 (Santa Cruz, sc-25310) were used at 1/1000 dilution.

Chromatin Immunoprecipitation

ChIP assays were completed as previously described [21,22]. 3 ug of antibody was used per 25 ug of chromatin. The antibodies used are anti-IgG mouse (Abcam ab18413) and anti-TBX5 (Proteintech 13178-1-AP). The primers used are sense 5'-TCCTCCCCATCTGTGTCATC-3' and antisense 5'-GGATCCATCACCATCAATAACC-3' for the Chromosome 15 Gene Desert negative control, sense 5'-CTCCACGCACGTGGCTCGGGGCGG-3', and antisense 5'- TAGGGGAACCCACAAACCCCATGG-3'for proximal GATA4 site of the *Ccnd2* promoter as positive control (-247 to -585 bp from TSS), sense 5'-TTGCAAGCCTCCGAAGTTAG-3' and antisense 5'-GATAGGGGAACCCACAAACC-3'for the proximal *Ccnd2* promoter TBE binding site (-244 to -325 bp from TSS), sense 5'-CTTTATGCCCCCATGGTATG-3' and antisense 5'-GGATAGGGGAACCCACAAAC-3'for the proximal *Ccnd2* promoter TBE binding site (-751 to -891 bp from TSS), sense 5'-TATTCGGGAAACCCATCTTG-3' and antisense 5'-TGCGAACTTAAGACGGAACC-3'for the *Ccnd2* promoter TBE binding site (-1855 to -1954 bp from TSS), sense 5'-ACACTCGTGCTTCTGCACAC-3' and antisense 5'-AGCTAGCACATGGGAAAAGC-3'for the distal Ccnd2 promoter TBE binding site (-2772 to -2862 bp from TSS).

QPCR and RNA analysis

Total RNA was isolated from cells using Trizol reagent (Invitrogen). QPCR was completed as previously described [16].Quantitative PCR analysis was performed on mouse tissues as described previously [16]. Primers are available upon request.

Immunohistochemistry and Immunofluorescence assays

Immunohistochemistry was done on mouse tissues at different embryonic stages as described previously [16]. Homemade anti-TBX5a antibody was used at a concentration of 1/2300 [16]. Hybridoma PAX7 antibody (AB_528428 DSHB) used at a concentration of 1/100.

Immunofluorescence was completed on mouse tissues as previously described [16]. Tissues were fixed in 4% PFA for 24 hours. Paraffin sections were stained with primary antibodies against TBX5a 1/2300 (homemade), PAX7 1/100 (Developmental Studies Hybridoma Bank AB_528428), KI67 1/200 (Thermo Scientific, RM-9106-SO) and 1/20 (Abcam, ab8191) , MYBBP1a 1/20 (Abcam, ab89121), MSTN 1/200 (ProteinTech, 19142-1-AP), MYOD 1/100 (Santa Cruz, sc-377460), MYOG 1/100 (Developmental Studies HBank, F5D-s), eMHC 1/100 (Developmental Studies Hybridoma B F1.652), followed by fluorescent secondary antibodies obtained from Molecular Probes. Pictures were acquired with a Carl Zeiss AxioVision Observer M2. An average of 20 random fields were counted.

In Vivo Studies

All animal experiments were completed in accordance to University of Ottawa guidelines for animal care and were approved by Institutional Animal Care and Use Committee (IACUC). 6 to 9 weeks old Wild Type (C57Bl/6) and MDX mice (C57BL/10ScSn-DmdMdx from Jackson Laboratories) kindly provided from Dr. Bernard Jasmin (University of Ottawa). Age matched mice heterozygous for a null *Tbx5* allele were previously described and compared to Wild Type littermates [15]. At end points, mice were perfused in 4% paraformaldehyde, embedded and sectioned for immunofluorescence assays or frozen for RNA extractions.

Cardiotoxin-induced muscle regeneration

6-9 weeks old animals were used for cardiotoxin injection experiments. Cardiotoxin (Cedarlane, L8102) was diluted to a 10 μM in phosphate-buffered saline (PBS) and 25 μl was injected intramuscularly into the *Tibialis anterior* muscles. Muscles were collected at Day 0 (resting) and 4 and 7 days after cardiotoxin injection, perfused in 4% paraformaldehyde, embedded and cross sectioned at 4 μm for histological observation. Alternatively, some samples were flash frozen in liquid nitrogen for RNA analysis. Control (WT) and *Tbx5* heterozygous mice were harvested at same time points.

Quantification of muscle morphology

Myofiber diameter and central nuclei measurements were carried out and quantified using Zeiss AxioVision Rel.4.8 and Fiji software. Measurements were taken using images of Hematoxylin/Eosin stained muscle sections at 20x magnification. Muscle sections from at least four separate planes of the injected muscle were counted in total and averaged to create single average data points for each sample and the average of these measures from all animals per time point or condition was then used to determine the average myofiber diameter and the percentage of myofibers with central nuclei. For Cardiotoxin experiments comparing Wild Type and *Tbx5* heterozygous mice, 3 muscles were averaged per condition.

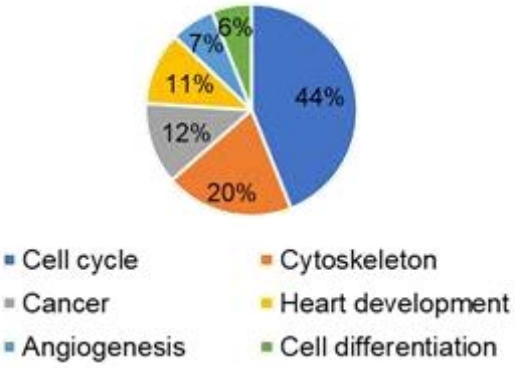
4.9 Results

Cyclin D2 (*Ccnd2*) is a direct target of TBX5a in skeletal myoblasts

To identify downstream targets of TBX5a in skeletal myoblasts, RNA extracts of the C2C12 cells overexpressing TBX5a were collected and microarray analysis was performed.

Transcriptomic analysis of the C2C12 pLNCX2 and C2C12 3xFlag TBX5a cell lines was completed using the Affymetrix Gene Chip Mouse Gene 2.0 ST Array. Ingenuity Pathway Analysis (IPA) Software (QIAGEN) analysis revealed over 1300 differentially expressed genes ($p<0.05$) with 750 upregulated and 600 downregulated genes, while the majority of the differentially expressed genes belonged to cell cycle pathways (e.g. *Ccnd2, Ccne1/2, Cdc6, Mcm2, Ccna2, Myf5, Aurka/b*) (**Figure 4.1 A-B**). qPCR validation on select genes associated with proliferation, including *Ccnd2, Ccne1/2, Cdc6, Mcm2, Ccna2, Myf5* and *Aurka/b,* confirmed the microarray results (**Figure 4.1 C**). We focused on, CCND2 whose activity is required for cell cycle G1/S transition [23], results and qPCR validations (**Figure 4.1 C**). Previous work in our lab has shown that *Ccnd2* is a direct transcriptional target of GATA4 [24]. CCND2 is also required for myoblast proliferation and is an inhibitor of myoblast differentiation making it an interesting target for therapeutic development in degenerative muscle diseases like DMD [25,26]. Bioinformatics analysis of the *Ccnd2* promoter by Genomatix (genomatix.de) located 4 conserved TBE motifs within the first 3000 base pairs from the transcription start site including the distal and proximal sites (**Figure 4.1 D**). A significant enrichment of TBX5 on the predicted TBE site in the proximal promoter of *Ccnd2* (-244 to -325) was observed using ChIP analysis (**Figure 4.1 E**). Interestingly, this region contains both TBE and GATA sites. These results demonstrate that TBX5a directly interacts with the proximal *Ccnd2* promoter through the TBE site.

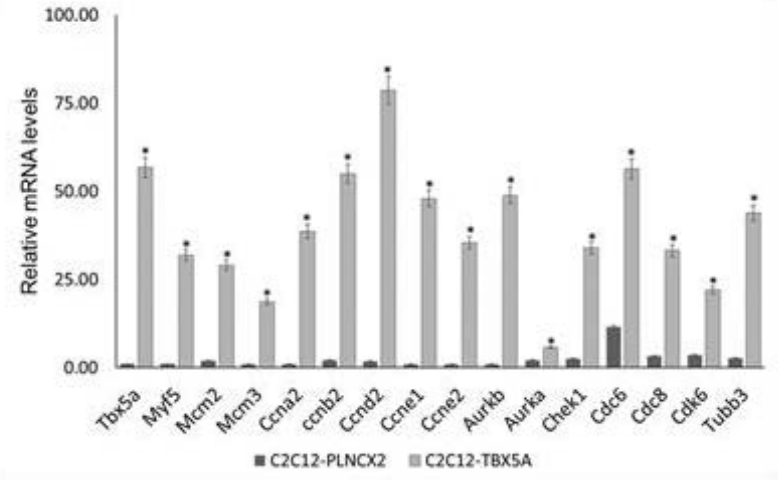
A



B

Gene	Gene Name	Fold Change
Ccne2	cyclin E2	3.212169
Tubb3	tubulin, beta 3	2.959965
Cdc6	cell division cycle 6	2.871294
Mcm3	minichromosome maintenance deficient 3	2.868523
Mcm2	minichromosome maintenance deficient 2 mitotin	2.801006
Fbxo5	F-box protein 5	2.581374
Chek1	checkpoint kinase 1	2.500615
Zwilch	Zwilch, kinetochore associated	2.396096
Mki67	antigen identified by monoclonal antibody Ki 67	2.150395
Pola1	polymerase (DNA directed), alpha 1	2.131283
Aurkb	aurora kinase B	2.131276
Chaf1a	chromatin assembly factor 1, subunit A (p150)	2.131224
Bub1b	budding uninhibited by benzimidazoles 1 beta	2.120403
Cdk1	cell division cycle 2 homolog A	2.116326
Ccnb2	cyclin B2	2.095227
Cep55	centrosomal protein 55	2.066887
Aurka	aurora kinase A	2.04998
Gmnn	geminin	2.04874
Rbl1	retinoblastoma-like 1 (p107)	1.987047
Mcm8	minichromosome maintenance deficient 8	1.97902
Cdc20	cell division cycle 20	1.975522
Ccnd2	cyclin D2	1.972935
Cdca8	cell division cycle associated 8	1.96771
Stmn1	stathmin 1	1.916809
Cdk6	cyclin-dependent kinase 6	1.82729
Ccna2	cyclin A2	1.806763
Ccne1	cyclin E1	1.614042
Mcm6	minichromosome maintenance deficient 6visiae)	1.577925

C



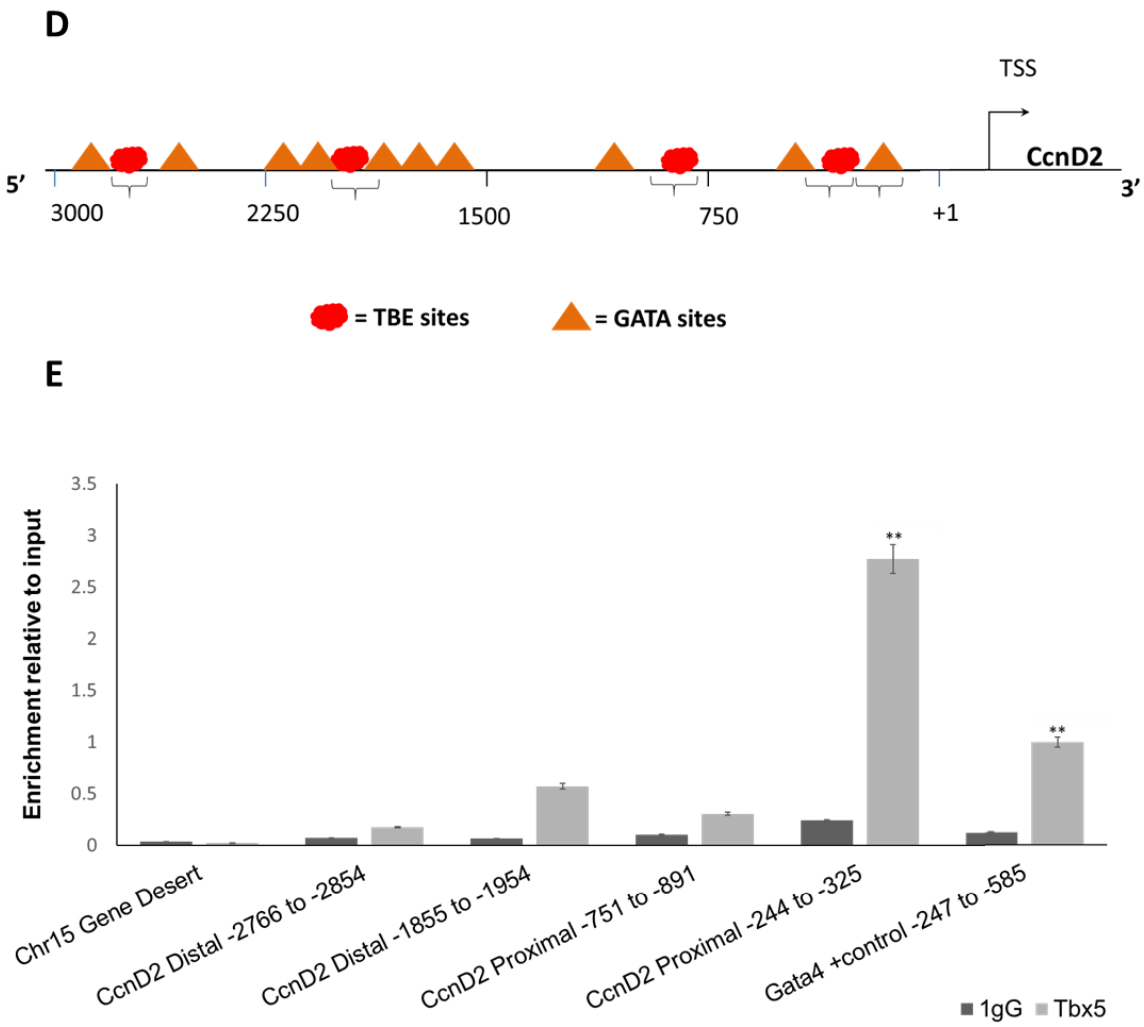
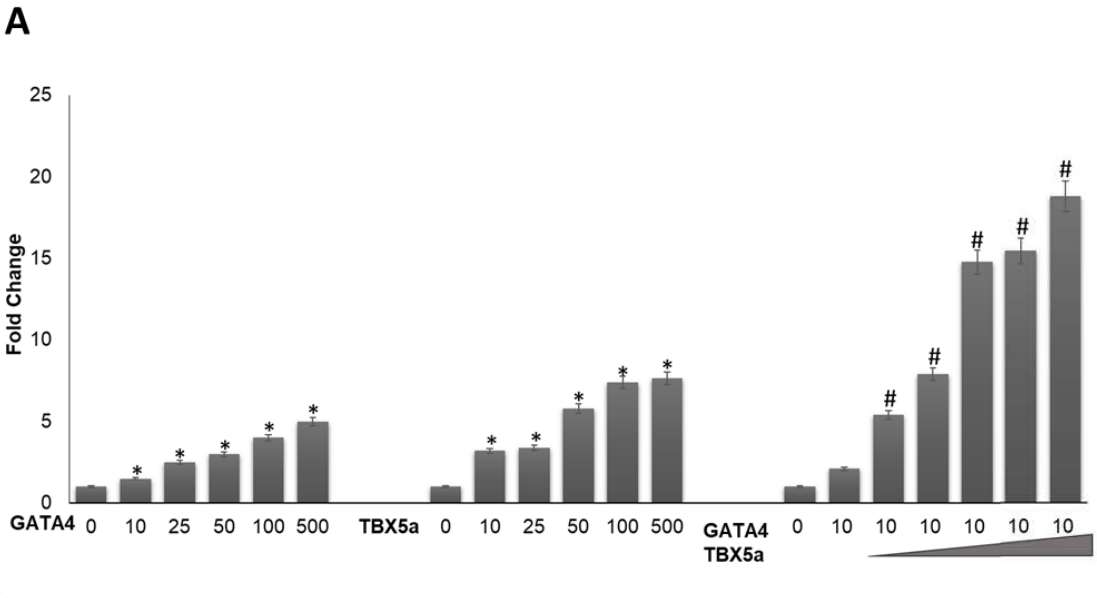


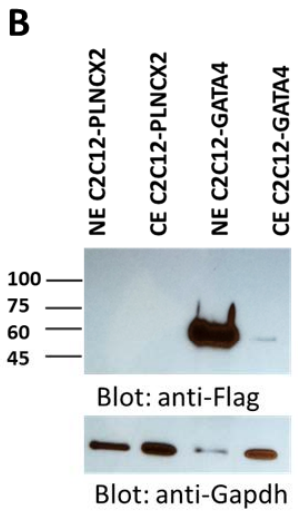
Figure 4.1 Microarray analysis identifies *Ccnd2* as a direct downstream target of TBX5a in skeletal myoblasts. A. TBX5a target genes in C2C12 myoblasts classified by functional pathways with the majority of differentially expressed genes in cell cycle pathways. **B**. Examples of significantly upregulated TBX5a genes in C2C12 myoblasts using IPA analysis. **C.** qPCR confirmation of selected differentially expressed genes as revealed in the microarray analysis. Values were normalized to Gapdh and are the mean values +/- SEM * $p<0.05$, TBX5a versus PLNCX2 (control). **D.** *Ccnd2* mouse promoter mapped with predicted TBE and GATA binding sites (all with matrix similarity >0.8). Note: The curly brackets highlight sites for which oligos are generated. **E.** ChIP analysis demonstrating endogenous TBX5 binding on the *Ccnd2* promoter **$p<0.01$ IgG ChIP negative control Vs. TBX5a ChIP.

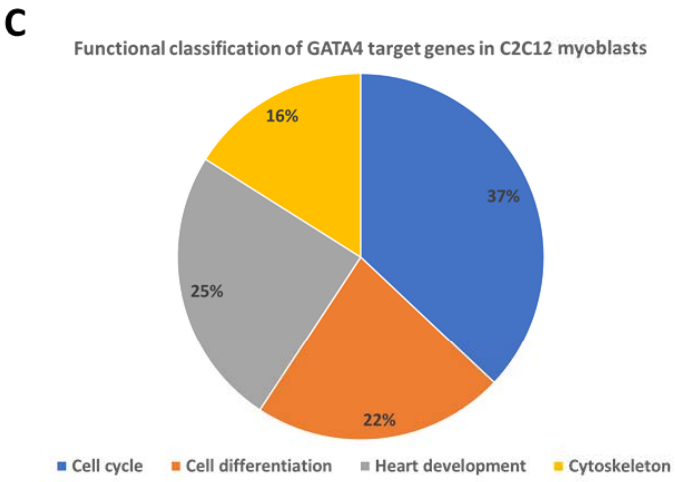
TBX5 cooperates with GATA4 to promote myoblast proliferation

Since GATA4, a known TBX5 partner in cardiomyocytes, is known to be a transcriptional activator of *Ccnd2*, we performed luciferase assays to test the effect of TBX5a on the *Ccnd2* promoter in presence or absence of GATA4. For this, TBX5a and GATA4 mammalian expression plasmids were either individually transfected or co-transfected with the *Ccnd2* proximal promoter reporter plasmid into C2C12 myoblasts. As shown in **Figure 4.2 A**, TBX5a is able to activate the *Ccnd2* promoter, similar to GATA4, in a dose dependent manner. Moreover, GATA4 and TBX5a synergistically activated the *Ccnd2* promoter also in a dose dependent manner. These results prompted us to test the effect of GATA4 on myoblast proliferation. We had previously observed that TBX5a and GATA4 both inhibit C2C12 myoblast differentiation [16]. To determine if GATA4 and TBX5 have convergent targets in myoblasts, RNA extracts of the C2C12 cells overexpressing GATA4 (GATA4 overexpression confirmed by western blots as shown in **Figure 4.2 B**) were collected and microarray analysis was performed. Transcriptomic analysis of the C2C12 pLNCX2 (Control) and C2C12 3xFlag GATA4 cell lines was completed using the Affymetrix Gene Chip Mouse Gene 2.0 ST Array. Ingenuity Pathway Analysis (IPA) Software (QIAGEN) analysis revealed over 500 differentially expressed genes with 270 upregulated and 250 downregulated genes functionally classified into cell cycle, cell differentiation, cytoskeleton and heart development pathways (**Figure 4.2 C**). Comparison of differentially expressed genes by TBX5a or GATA4 in C2C12 myoblasts revealed 156 common genes (**Figure 4.2 D**). Examples of common downregulated genes include Insulin-Like Factor Binding Protein 3 (*Igfbp3*) known to promote myoblast differentiation [27], Cellular Retinoic Acid Binding Protein 1 (*Crabp1*) related to family member Crabp2 known to regulate myoblast

differentiation [28], Semaphorin 7a (*Sema7a*) involved in osteoblast differentiation [29], and Dickkopf Wnt Signaling Pathway Inhibitor 2 (*Dkk2*) required for terminal osteoblast differentiation [30]. Interestingly, amongst the 34 common upregulated target genes of TBX5a and GATA4 in C2C12 myoblasts, *Ccnd2* was consistently seen to be significantly upregulated. Examples of other upregulated common genes are Checkpoint Kinase 1 (*Chek1*) involved in S-G2 phases of the cell cycle [31], Krupple-Like Factor (*Klf4*) a negative regulator of smooth muscle cell differentiation [32], Histone Deacetylase 9 (*Hdac9*) a suppressor of myoblast differentiation [33], Mini-Chromosome Maintenance 4 (*Mcm4*) which is essential for the initiation of DNA replication [34], and Chromatin Assembly Factor 1 Subunit A (*Chaf1a*) which promotes cell proliferation [35]. Several of these genes displayed multiple GATA and TBX binding motifs in their upstream sequences. As an example, in silico analysis of the *Klf4* promoter by Genomatix (genomatix.de) located 4 conserved TBE and 6 GATA motifs within the first 3000 base pairs from the transcription start site including distal and proximal cluster sites. Thus, TBX5a and GATA4 appear to be important transcriptional regulator of myoblast expansion by regulating common genes involved in skeletal myogenesis.


A
Fold Change
25
20
15
10
5
0
GATA4 0 10 25 50 100 500
TBX5a 0 10 25 50 100 500
GATA4
TBX5a 0 10 10 10 10 10 10


B
NE C2C12-PLNCX2
CE C2C12-PLNCX2
NE C2C12-GATA4
CE C2C12-GATA4
100
75
60
45
Blot: anti-Flag
Blot: anti-Gapdh


C
Functional classification of GATA4 target genes in C2C12 myoblasts
16%
37%
25%
22%
Cell cycle
Cell differentiation
Heart development
Cytoskeleton

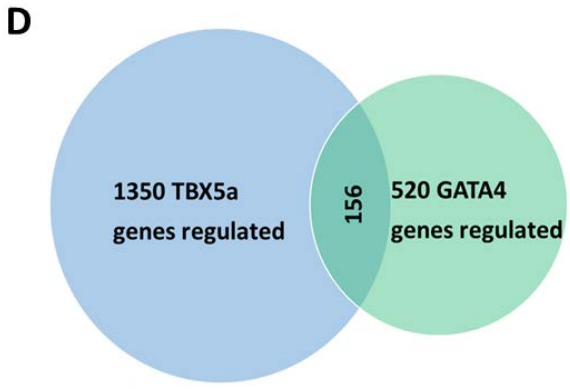
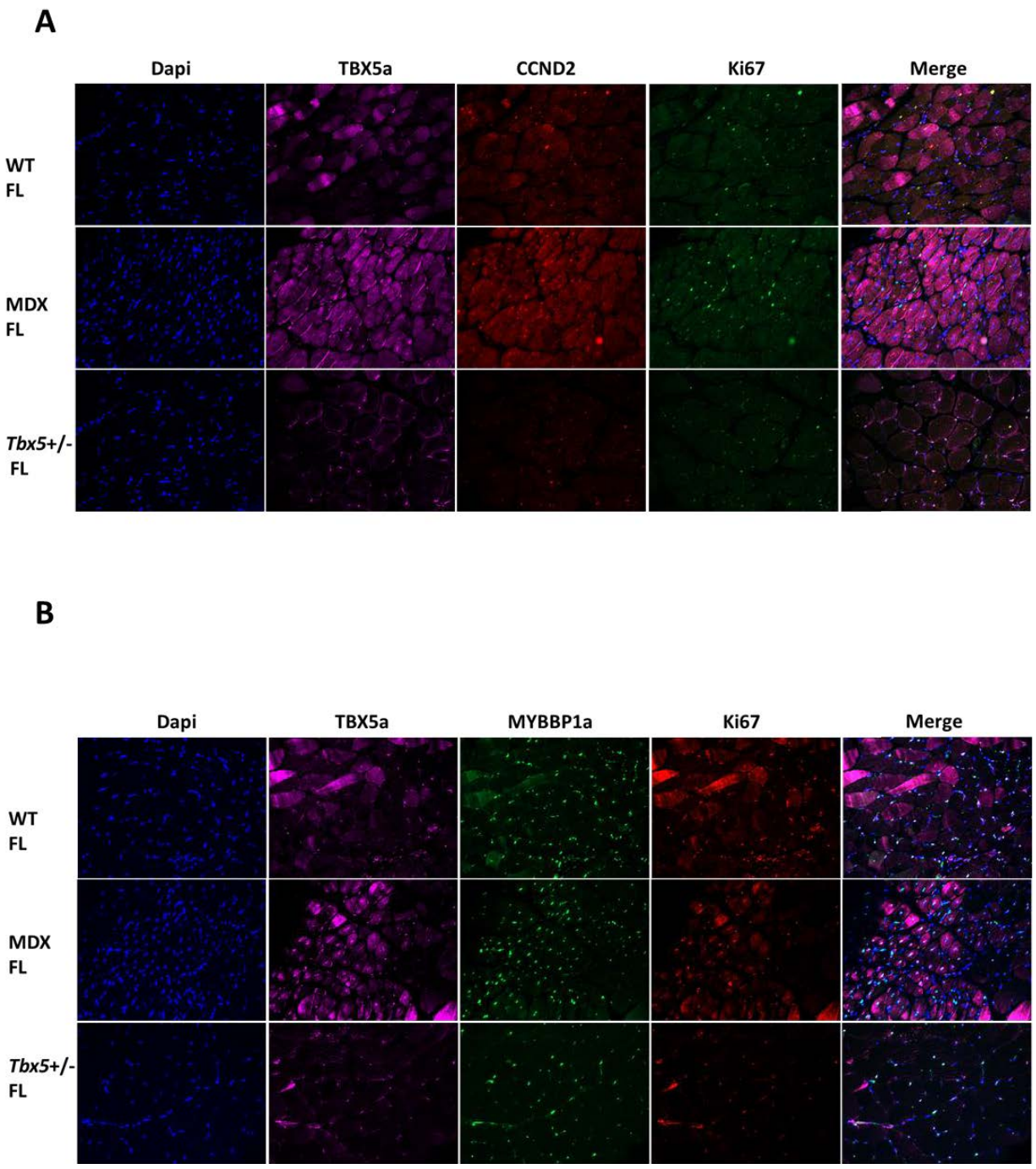

D
1350 TBX5a
genes regulated
156
520 GATA4
genes regulated

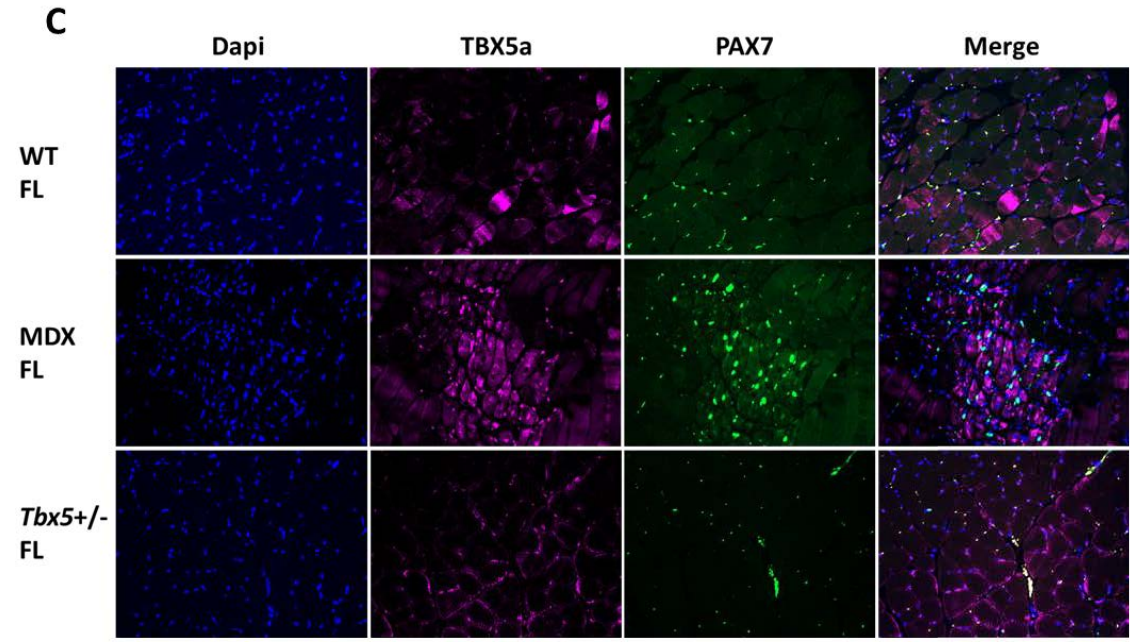
Figure 4.2 TBX5a and GATA4 promote myoblast proliferation cooperatively. A. TBX5a and GATA4 synergistically increase *Ccnd2* promoter activity. C2C12 cells transfected with GATA4 (left), TBX5a (middle), or GATA4 (10 ng) and increasing doses of TBX5a (0, 10, 25, 50,100, 500 ng) (right) and the *Ccnd2* proximal promoter reporter construct. Note: all representative of three experiments, * $p< 0.05$ vs. baseline, # $p< 0.05$ vs. GATA4 at 500 ng. **B.** Western blot representing NE (nuclear extracts) and CE (cytoplasmic extracts) of C2C12 stable cell line overexpressing GATA4 versus the control (empty vector-pLNCX2) cells. **C.** GATA4 target genes in C2C12 myoblasts classified by functional pathways with the majority of differentially expressed genes in cell cycle pathways. **D.** Microarray analysis reveals 156 common target genes for TBX5a and GATA4 in C2C12 myoblasts as analyzed by IPA.

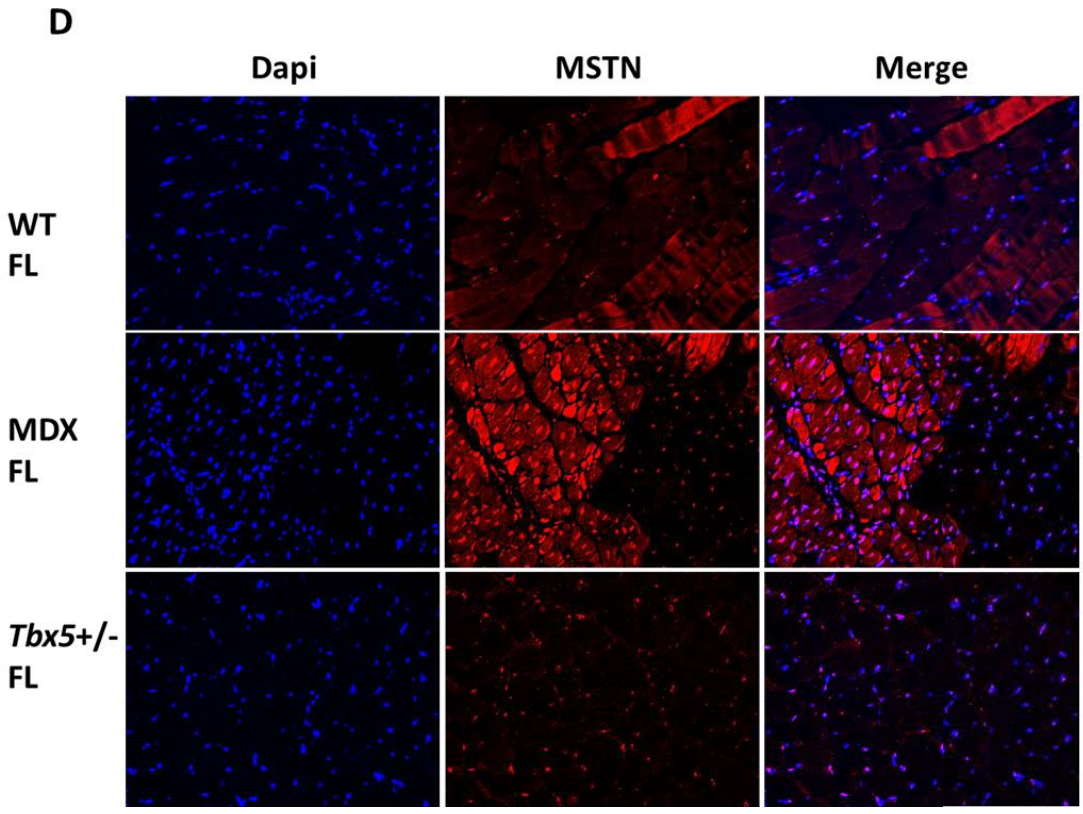
Impaired regeneration capacity of *Tbx5* heterozygote mice

As stated above, TBX5a promotes proliferation of myoblasts [16], and TBX5a levels are increased in the skeletal muscle tissues of MDX mice (previous chapter and **Supplementary Figure 4.1**). No muscle abnormalities have been reported in *Tbx5* heterozygote mice. However, given our data, we examined whether loss of a *Tbx5* allele in mice impacts skeletal muscle regenerative capacity by effecting the number and proliferative capacity of satellite cells. First, we used immunofluorescence assays on the forelimb tissues of WT, MDX and *Tbx5* heterozygote mice to assess the expression of TBX5a, CCND2 and the proliferation marker Ki67 (**Figure 4.3 A**). Studies have shown that the MDX skeletal muscle tissues display excessive proliferation but as they age the regenerative ability of the tissue declines due to exhaustion of the satellite cell pool [5]. Consistent with this, TBX5 and CCND2 levels were increased in MDX forelimbs compared to controls. In contrast to the MDX tissues, both markers as well as Ki67 were significantly decreased in the limbs of mice lacking one *Tbx5* allele. This shows that *in vivo*, decreased *Tbx5* levels are associated with decreased expression of proliferation markers in post-natal limb muscles.

Given the above, we tested the expression of other markers of myoblast proliferation and differentiation such as PAX7, MYBBP1a and MSTN in WT, MDX and *Tbx5* heterozygote mice to see if the lack of one *Tbx5* allele is associated with broader proliferation/regeneration and differentiation markers. Immunofluorescence assays on the forelimb tissues of the various mouse lines showed that MYBBP1a (pro-proliferation marker) and Ki67 were upregulated in the MDX tissues and downregulated in the *Tbx5*+/- mice compared to WT (**Figure 4.3 B**). Interestingly, levels of PAX7 were increased in the MDX tissues but decreased in the *Tbx5*+/- mice compared to WT (**Figure 4.3 C**). Since PAX7 marks satellite cells, the results suggested that *Tbx5* haploinsufficiency may be associated with decreased satellite cells and impaired regenerative capacity. Downregulation of MSTN (inhibitor of differentiation and downstream target of TBX5a in skeletal myoblasts) was also observed in the *Tbx5* heterozygote mice tissue (**Figure 4.3 D**). qPCR analysis carried out on these tissues show that *Tbx5a*, *Myf5* (marker of proliferating myoblasts), *Ccnd2* and *Pax7* were downregulated in the *Tbx5*+/- compared to both MDX and WT mice (**Figure 4.3 E**). In contrast, Myogenin, a marker of skeletal muscle differentiation, was not significantly changed across the different lines. These results suggest that TBX5a may be involved in progenitor expansion and muscle regeneration.


A
Dapi
TBX5a
CCND2
Ki67
Merge
WT FL
MDX FL
Tbx5+/- FL
B
Dapi
TBX5a
MYBBP1a
Ki67
Merge
WT FL
MDX FL
Tbx5+/- FL


C
Dapi
TBX5a
PAX7
Merge
WT
FL
MDX
FL
Tbx5+/-
FL


D
Dapi
MSTN
Merge
WT
FL
MDX
FL
Tbx5+/-
FL

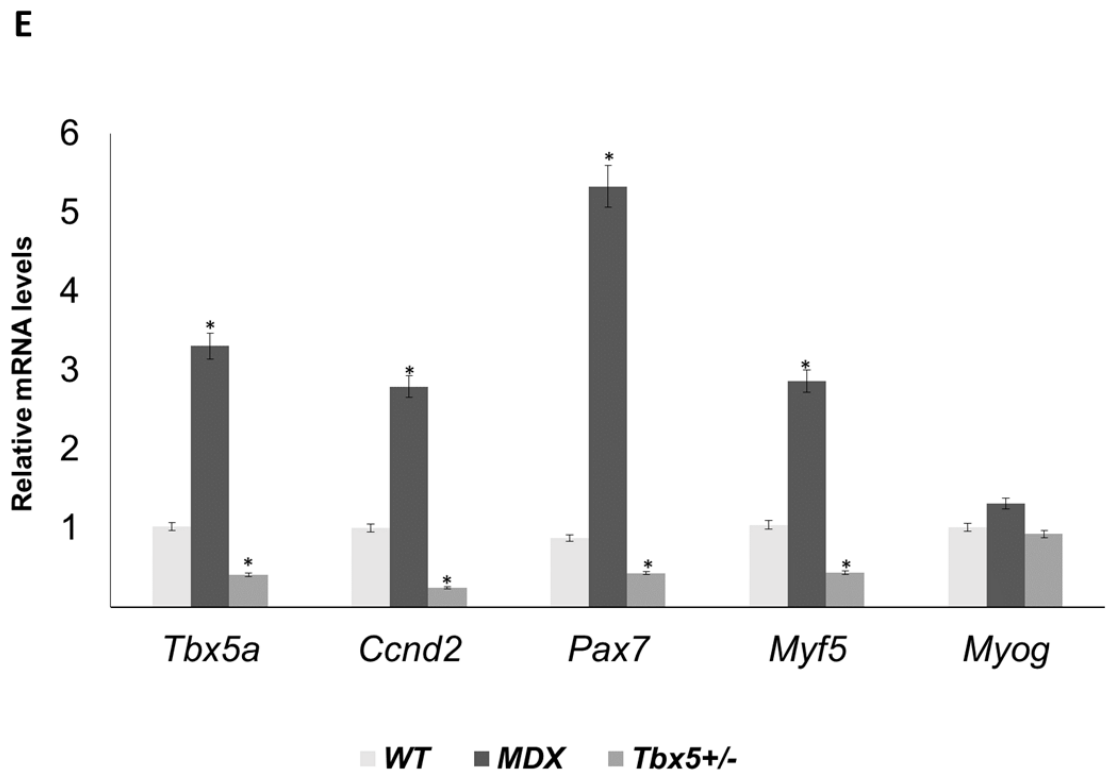
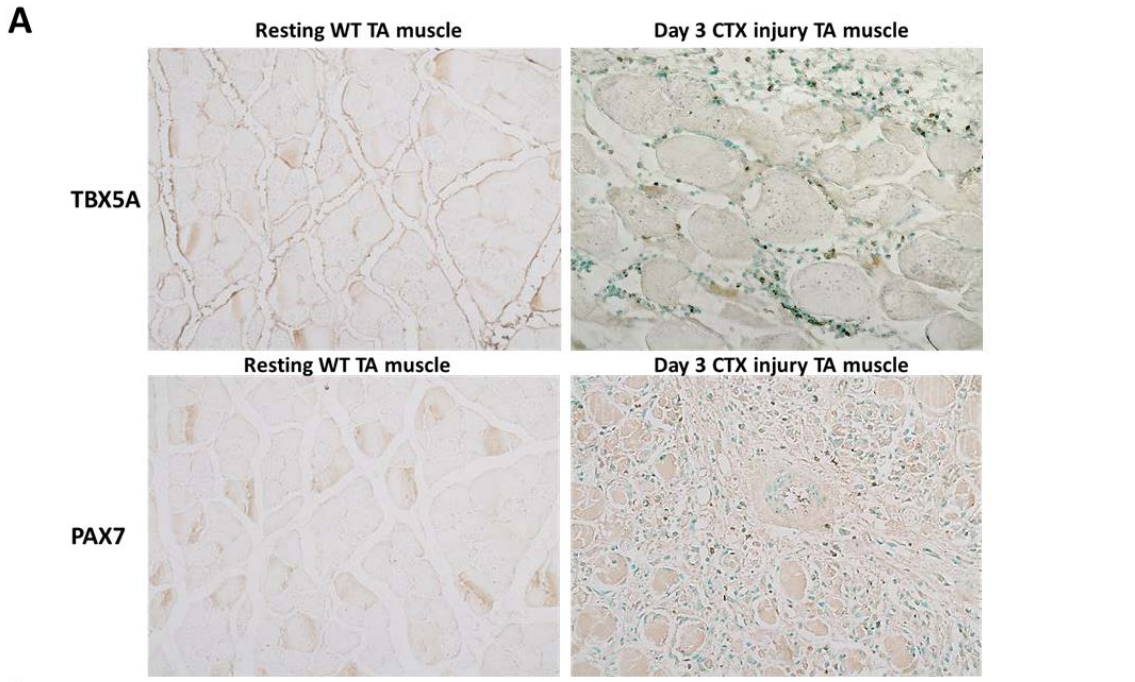


Figure 4.3 Decreased proliferation markers in the limbs of *Tbx5* heterozygous mice. A. Immunofluorescence assays on forelimb, Extensor Carpi Ulnaris (ECU), tissue of WT, MDX and *Tbx5*+/- mice stained for TBX5a (Purple), CCND2 (Red), Ki67 (Green), Dapi (Blue). **B.** Immunofluorescence assays on forelimb Extensor Carpi Ulnaris (ECU) tissue of WT, MDX and *Tbx5*+/- mice stained for TBX5a (Purple), MYBBP1a (Green), Ki67 (Red), Dapi (Blue). **C.** Immunofluorescence assays on forelimb Extensor Carpi Ulnaris (ECU) tissue of WT, MDX and *Tbx5*+/- mice stained for TBX5a (Purple), PAX7 (Green), and Dapi (Blue). **D.** Immunofluorescence on forelimb Extensor Carpi Ulnaris (ECU) tissue of WT, MDX and *Tbx5*+/- mice stained for MSTN (Red) and Dapi (Blue) (images taken at 20x magnification). **E.** qPCR analysis of *Tbx5a, Ccnd2, Pax7, Myf5* and *Myog* gene expressions in WT, MDX and *Tbx5*+/- mice forelimb tissues n=3, Normalized to *Gapdh*. *p value <0.05 compared to WT.

To test this possibility, we first analyzed the levels of TBX5a in response to skeletal muscle injury using Cardiotoxin (myonecrotic agent) [36]. TBX5 is known to be expressed in the developing forelimb but we have also observed TBX5 expression, postnatally, in the muscle tissues of the hindlimbs and the diaphragm in addition to the forelimb (shown in **Supplementary Figure 4.2 A-B**). Immunostaining of the hindlimb tissues of adult wildtype mice injected or not with Cardiotoxin (CTX) showed increased expression of TBX5a along with PAX7 (in lower panel) in the cardiotoxin treated limbs after 3 days (**Figure 4.4 A**). To determine whether TBX5 and PAX7 staining colocalize, we carried out immunofluorescence assays which also showed that TBX5 and PAX7 immunoreactivity remained elevated at day 4 post injection, when regeneration and proliferation of muscle cells is taking place (**Figure 4.4 B**). Moreover, PAX7 and TBX5a colocalized in some but not all cells, located on the fibers (**Figure 4.4 B**). The results suggest a role for TBX5a in satellite/progenitor expansion as well as in other cells involved in skeletal muscle response to injury.


A
Resting WT TA muscle
Day 3 CTX injury TA muscle
TBX5A
Resting WT TA muscle
Day 3 CTX injury TA muscle
PAX7

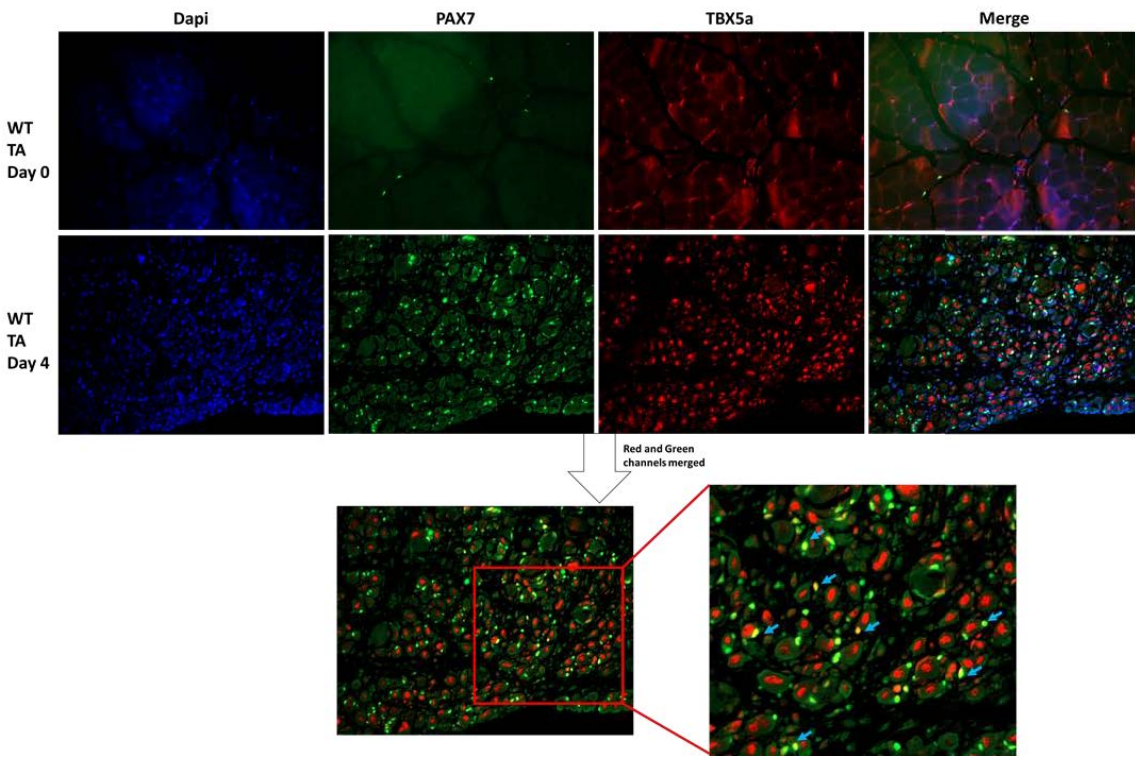
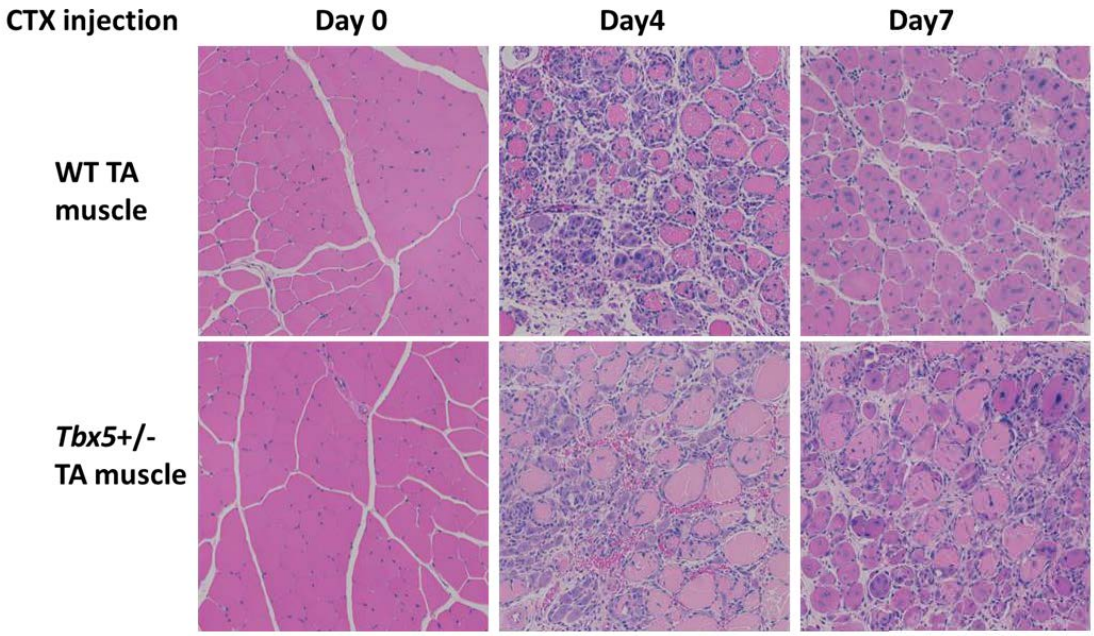

B
Dapi
PAX7
TBX5a
Merge
WT
TA
Day 0
WT
TA
Day 4
Red and Green
channels merged

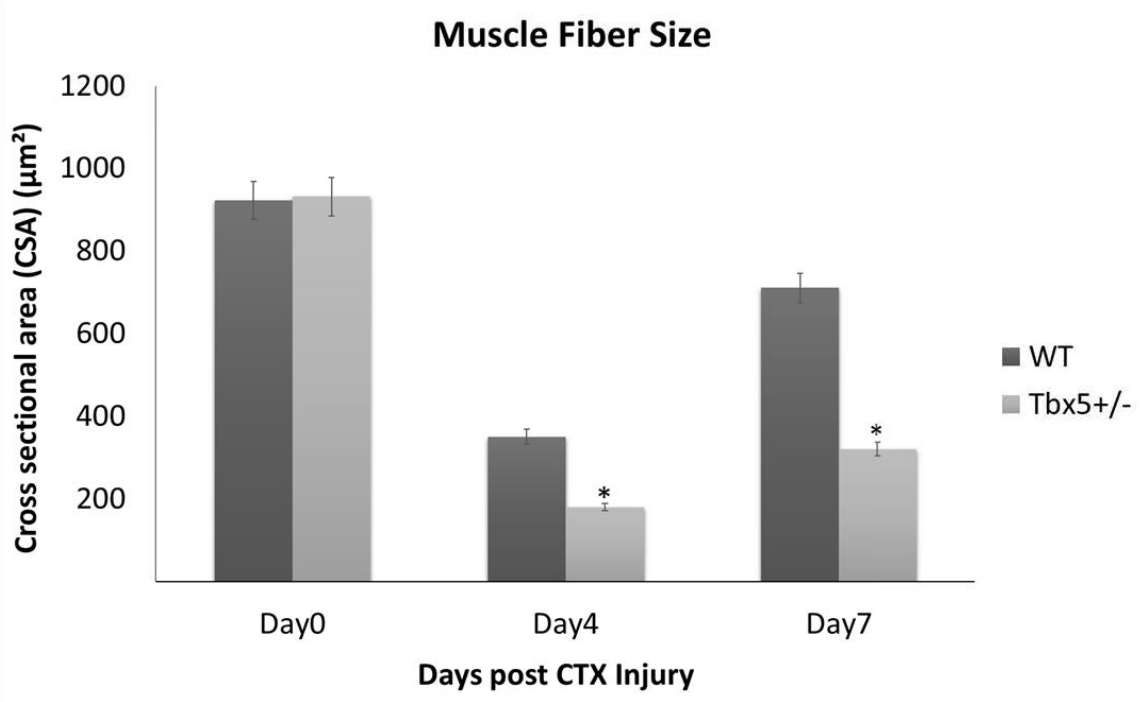
Figure 4.4 TBX5a is increased after cardiotoxin injection in hindlimb of WT mice. A. Immunohistochemistry on WT mouse tibialis anterior (TA) muscle tissue prior injection (Day 0) and Day 3 after CTX injection stained for TBX5a and PAX7 (brown nuclei) and counterstained with Dapi (blue). **B.** Immunofluorescence on TA muscle tissue of WT and Day 4 CTX injection stained for TBX5a (Red), PAX7 (Green), and Dapi (Blue) (20x magnification). **Note:** Region of interest highlighted in the illustration below where the green and red channels are merged showing colocalization of PAX7 and TBX5a in some satellite cells (yellow) indicated with blue arrows.

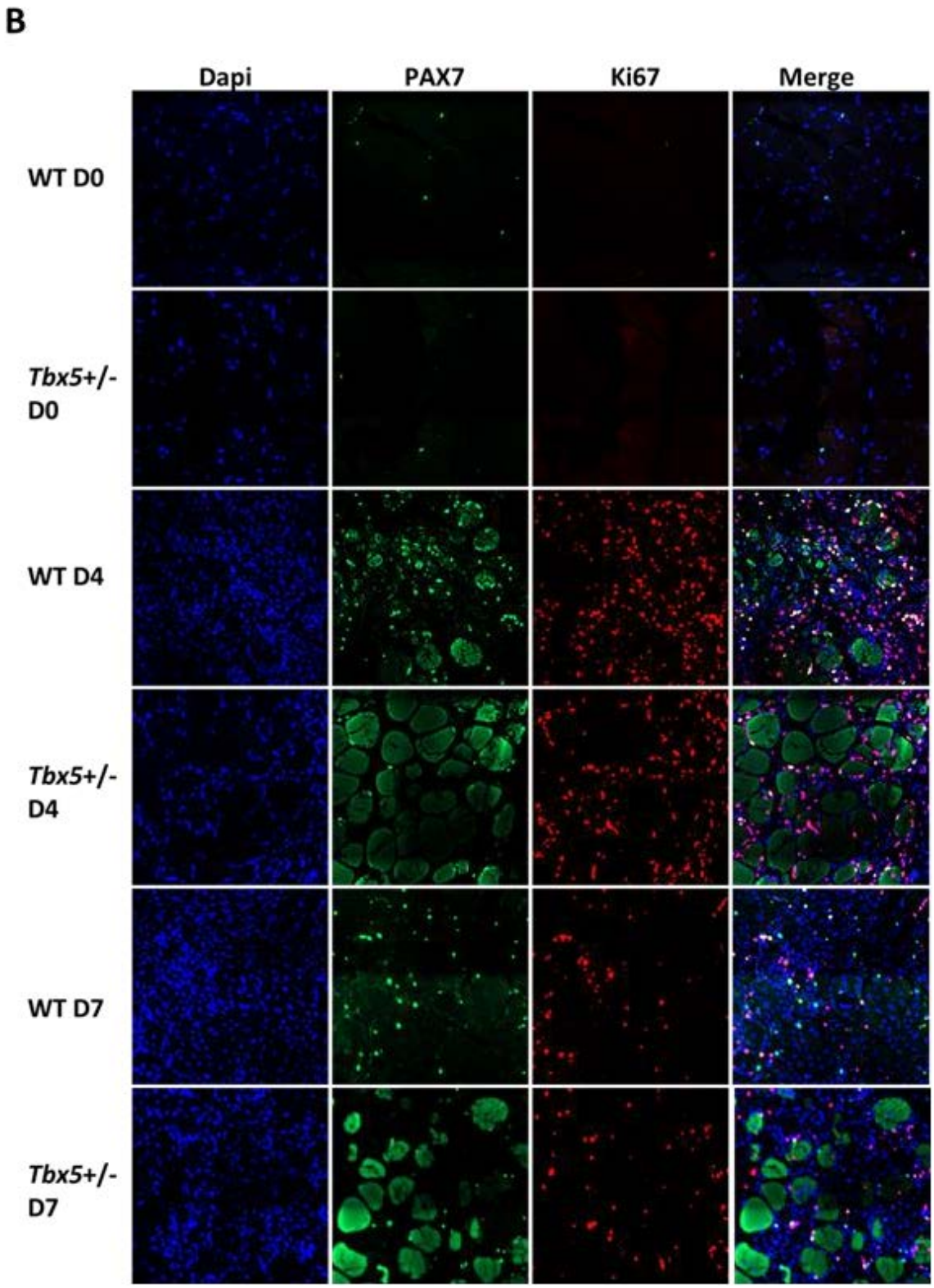
Next, we used *Tbx5*+/- mice to directly evaluate whether reduced levels of *Tbx5* affect skeletal muscle regeneration. Muscle injury was similarly induced by injection of CTX and histological analyses were carried out at 3 time points of the regeneration process. Day 0 (known as the resting state) is the baseline where there is no injection, by Day 4 post-injection, structural damage and degenerated myofibers are visible while the regeneration process is triggered and peaks which is indicated by presence of amplifying satellite cells (myoblasts), and by Day 7 post-injection, fusion and differentiation of myoblasts occurs to form new regenerated myofibers with centrally located nuclei and inflammation starts to progressively decrease [37]. Hematoxylin and eosin (H&E) staining, at day 0 (resting state) show that, there was no physical difference observed in the hindlimb Tibialis Anterior (TA) muscle tissues of the WT and *Tbx5*+/- mice prior to injury, where both tissues appeared to have similar sized muscle fibers (**Figure 4.5 A**). At 4 days post-injection, there was significant inflammatory infiltration along with necrotic muscle fibers and this was more pronounced in the *Tbx5*+/- mice. Moreover, at day 4, when regeneration peaks and formation of myofibers with centrally located nuclei occurs, the *Tbx5*+/- mice appear to form fewer and smaller regenerating fibers compared to WT mice. By day 7, when fiber differentiation and maturation starts to occur, the nascent myofibers progressively increased in diameter affiliated while immune cell infiltrate was reduced in the WT mice. In contrast, the *Tbx5*+/-

mice still exhibited necrotic fibers at this stage and the diameter of regenerating fibers remained smaller compared to the WT mice (**Figure 4.5 A**). These results are indicative of decreased regenerative ability and delayed removal of necrotic tissue in the limbs of *Tbx5*+/- mice. We further carried out immunofluorescence assays on the TA tissues of the WT and *Tbx5*+/- mice at the same timepoints, to evaluate the expression of various regeneration/proliferation and differentiation markers. At day 0 (baseline), *Tbx5*+/- mice appeared to have less satellite cells, marked by PAX7, compared to WT. At both day 4 and 7, PAX7 and Ki67 expression was significantly higher in the WT mice compared to *Tbx5*+/- mice, suggesting fewer number of activated satellite cells, less proliferation and impaired regeneration in the *Tbx5*+/- mice (**Figure 4.5 B)**. This was confirmed by quantification and shown in **Figure 4.5 D**. The expression of early and later markers of differentiation, MyoD and MyoG, respectively, were also reduced throughout the regenerative phases in the *Tbx5*+/- mice (**Figure 4.5 D and Supplementary Figure 4.3 A-B)**. Furthermore, *Tbx5*+/- mice appear to form significantly less nascent fibers stained by eMHC (marker of newly formed fibers) at both day 4 and 7, compared to WT muscle tissues (**Figure 4.5 C-D**). Last, although *Tbx5*+/- mice seem to contain more MyoD+ cells at baseline, MyoD levels are reduced after injury (**Figure 4.5 D)**. These results suggest that the loss of a *Tbx5* allele significantly impairs skeletal muscle reservoir of satellite cells and impairs muscle regeneration in response to injury.

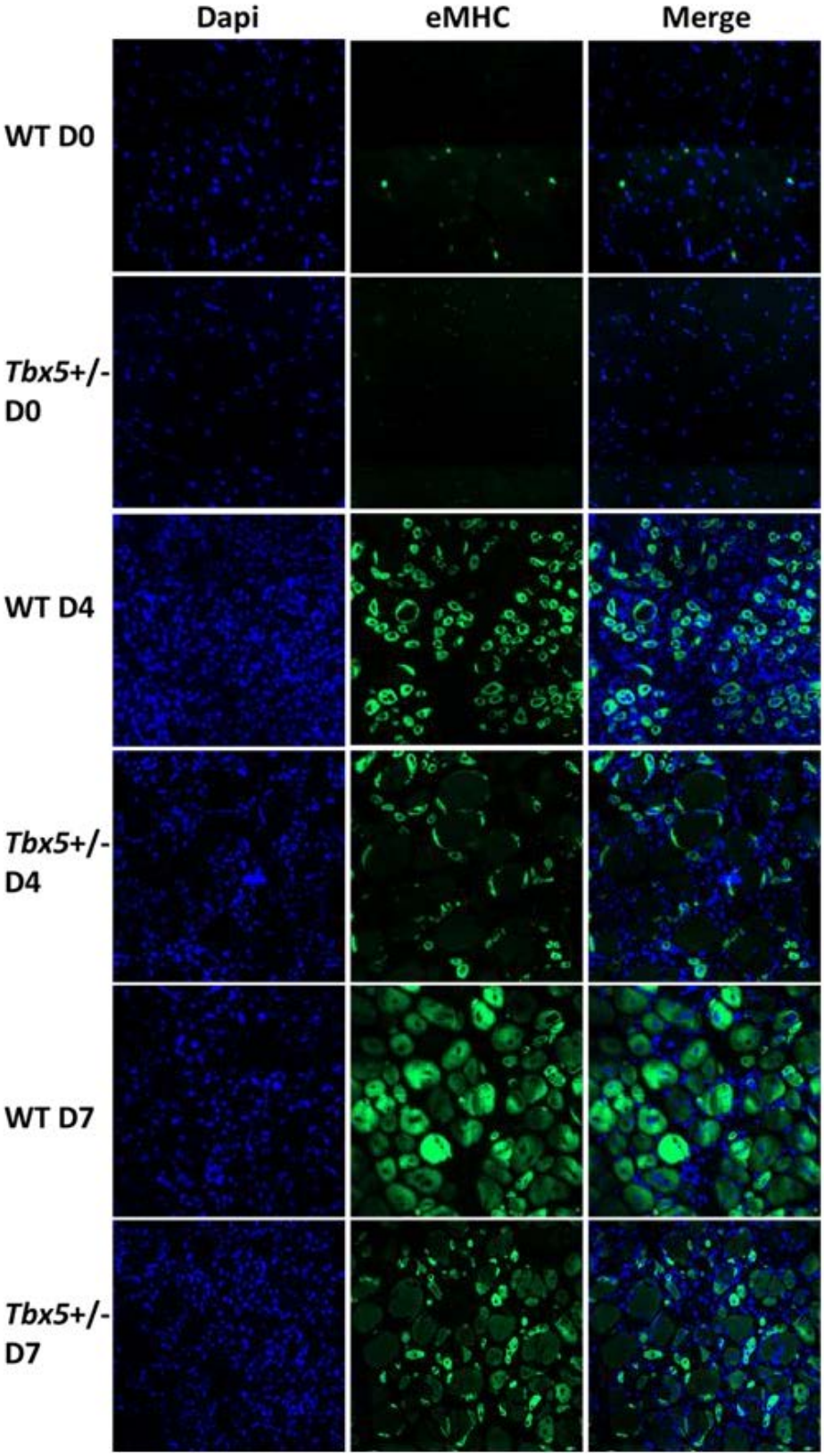
A


CTX injection
Day 0
Day4
Day7
WT TA muscle
Tbx5+/- TA muscle


Muscle Fiber Size
Cross sectional area (CSA) (µm²)
1200
1000
800
600
400
200
Day0
Day4
Day7
Days post CTX Injury
WT
Tbx5+/-
*
*


B
Dapi
PAX7
Ki67
Merge
WT D0
Tbx5+/- D0
WT D4
Tbx5+/- D4
WT D7
Tbx5+/- D7

C



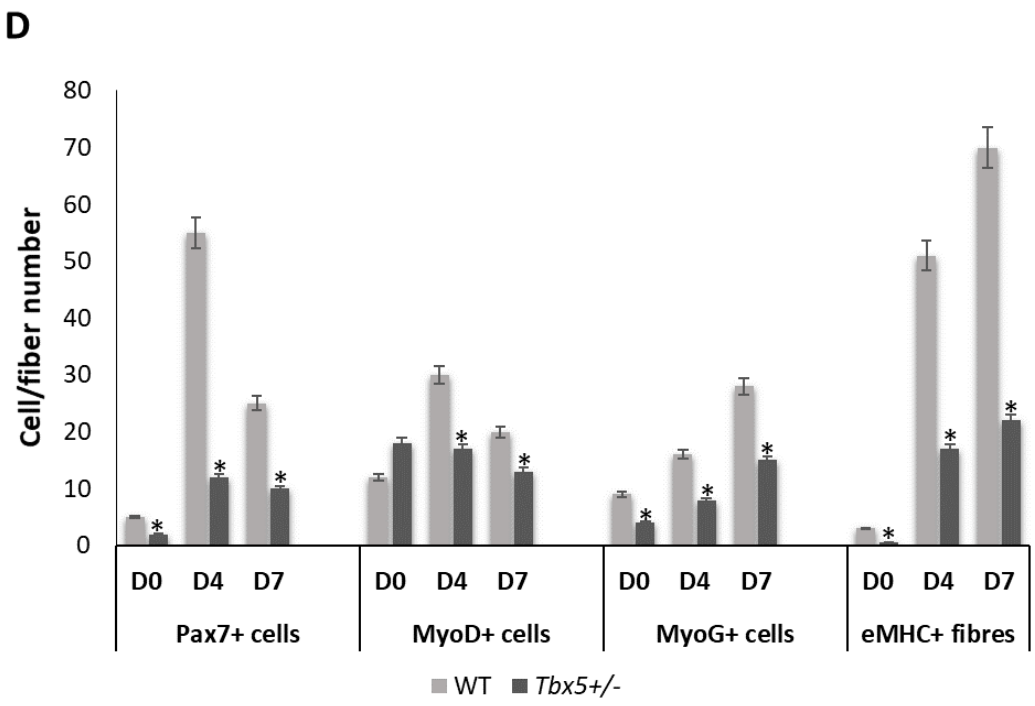


Figure 4.5 *Tbx5*+/- mice exhibit impaired skeletal muscle regeneration. A. Representative H&E staining of WT and *Tbx5*+/- mice TA muscle tissues at Day 0 (no injury), 4 and 7 days after injury (pink stains for the muscle and blue the nuclei). Regenerating (central nucleated) muscle fiber cross-sectional area (CSA) was significantly reduced in *Tbx5*+/- at Day 4 and 7 post injury (*P < 0.05) represented in graph. **B.** Representative Immunofluorescence images on WT and *Tbx5*+/- TA muscle at D0, D4 (4 days after injury) and D7 (7 days after injury) stained for PAX7 (Green) and Ki67 (Red). **C.** Representative Immunofluorescence images on WT and *Tbx5*+/- TA muscle at D0, D4, and D7 stained for eMHC (Green). Images are representative of 3 mice per condition and are taken at 20x magnification. **D.** Quantification of PAX7, MYOD, MYOG and eMHC positive cells/fibers 0, 4, and 7 days post injury in TA muscle sections of WT and *Tbx5*+/- mice. Positive cells in 20 randomly selected fields from the TA muscle sections of WT and *Tbx5*+/- mice, were counted under a fluorescent microscope with 40x objective lens. The results are presented as mean ± SEM. *P < 0.05 relative to WT.

4.10 Discussion

TBX5 is known to have essential roles in limb development as evidenced from gain and loss of function studies in various animal models [15,38], and the phenotype of individuals with *TBX5* mutations [14]. Despite the importance of musculoskeletal development in limb

formation, studies have largely focused on TBX5 skeletal role. Our previous studies have shown a correlation between TBX5a and myoblast proliferation in the skeletal muscle C2C12 cell line [16,17]. However, TBX5 function and mechanism of action in skeletal muscle proliferation and regeneration remained undetermined. In particular, little is known about TBX5 downstream targets and regulators in myoblast proliferation and whether its pro-proliferative role has implications in skeletal muscle regeneration and muscle disorders. In this report, we have shown that TBX5a plays a key role in regulating myogenesis by promoting proliferation of muscle progenitor cells, in part by directly activating cell cycle associated genes such as *Ccnd2*. Moreover, we identify GATA4 as a TBX5a cofactor in myoblast proliferation and show synergistic activation of *Ccnd2* promoter by cooperative TBX5a and GATA4 action. Studies show that GATA4 potentiates second heart field proliferation in cardiac septation [39] and is a cofactor and transcriptional activator of CCND2 in cardiogenesis [23,24]. Moreover studies have shown GATA4 to be an important regulator of osteoblast genes [40]. The synergistic growth promoting role of TBX5 and GATA4 extends to skeletal myoblasts by mediating transcriptional activation of cell cycle associated gene *Ccnd2*. Upstream regulators of cell cycle associated genes are critical in controlling the balance in cell growth in many tissues and pathways. Moreover, we have identified other common target genes of TBX5 and GATA4 in myoblasts that are known to be involved in the regulation of proliferation and differentiation of myoblasts and have opened up interesting questions to pursue; this will help shed light on the regulatory mechanisms governing skeletal myoblast proliferation and the pathogenesis of skeletal muscle disorders.

We further show that *Tbx5* heterozygote mice exhibit decreased levels of *Ccnd2* and other proliferation markers, whereas an increase in these genes is observed in the MDX mice

skeletal muscle, which are undergoing continued regeneration. We further show that the *Tbx5* heterozygote mice exhibit decreased expression of PAX7 (marker of satellite cells) compared to WT skeletal muscle tissues, raising the possibility that TBX5 haploinsufficiency may lead to decreased regeneration capacity. Indeed, colocalization of TBX5a in PAX7+ satellite cells on the myofibers, after CTX-injury, is consistent with a role in the expansion of the regenerating muscle progenitor cells. We show for the first time TBX5 involvement in the process of muscle regeneration using the CTX injury model. In addition to the previously known subtle skeletal defects in *Tbx5* heterozygote mice limb (elongated first digit and hypoplastic bones), we show that these mice have lower regenerative capacity in response to CTX-induced injury compared to WT mice. Our results suggest that the heterozygous loss of *Tbx5* impairs adult muscle regeneration as evidenced histologically and at the molecular level by the downregulation of important regeneration markers. Thus, TBX5 may have an important function in maintaining the regenerative capacity of skeletal muscle cells. Most skeletal muscle disorders exhibit reduced muscle regeneration capacity due to exhaustion of satellite cell pools, which progressively leads to muscle wasting and reduced mobility. TBX5 could potentially be a useful prognostic marker for muscle disorders. TBX5a upregulation in satellite cells might also act to boost muscle regeneration by coordinately regulating growth-promoting genes.

Lastly, while HOS patients exhibit musculoskeletal abnormalities due to mutations in the *TBX5* gene, skeletal muscle phenotypes and physiology in HOS individuals have received limited attention. Interestingly, an earlier study reported muscular involvement in HOS with 13 patients from unrelated families demonstrating hypoplasia of discrete muscle as found by MRI, neurological and radiological investigations [41]. Our data, together with the suggested

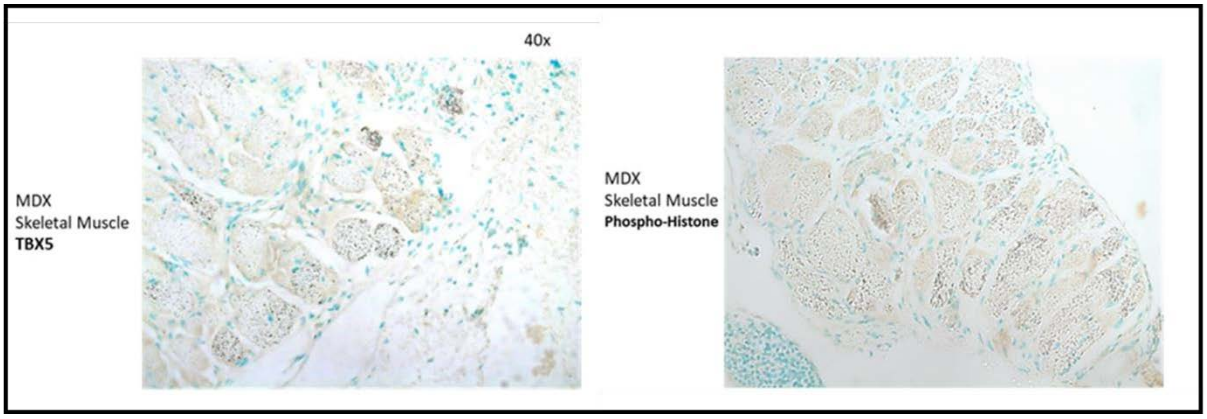
role of TBX4/TBX5 in muscle connective tissues [42], raise the intriguing possibility that TBX5 might coordinately regulate musculoskeletal development through its role in bones, muscle and tendons. This has potential implications for the management of HOS patients.

4.11 References

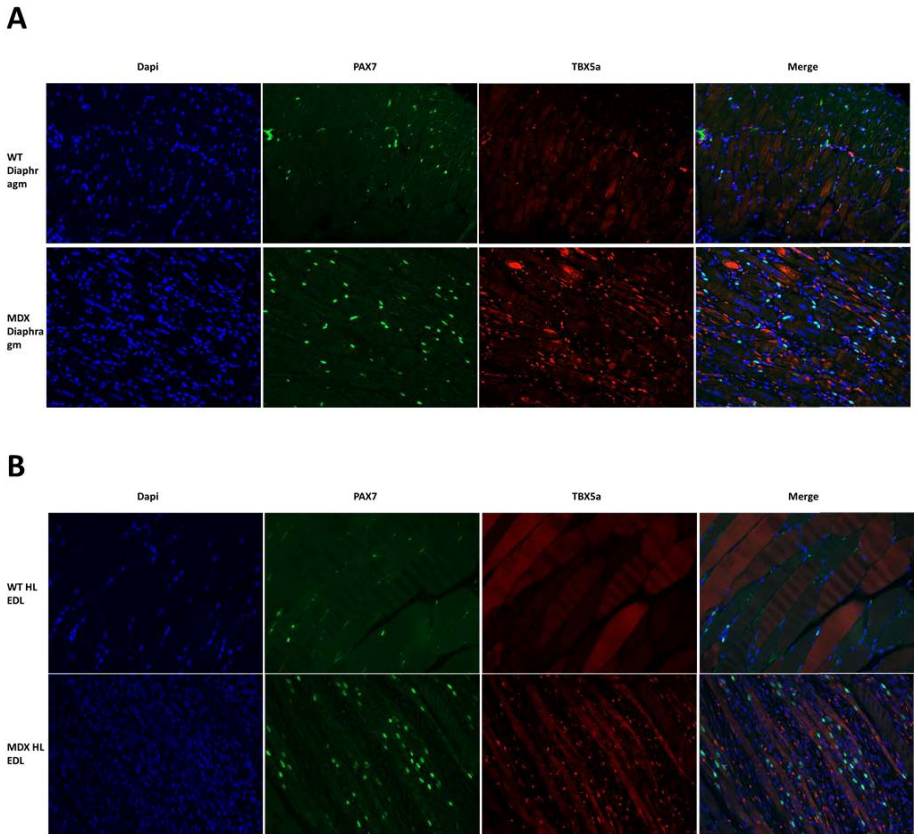
1. Bentzinger, C. F., Wang, Y. X. & Rudnicki, M. A. Building Muscle: Molecular Regulation of Myogenesis. *Cold Spring Harb. Perspect. Biol.* **4**, (2012).
2. Tajbakhsh, S. Skeletal muscle stem cells in developmental versus regenerative myogenesis. *J. Intern. Med.* **266**, 372–389 (2009).
3. Trovato, F. M., Imbesi, R., Conway, N. & Castrogiovanni, P. Morphological and Functional Aspects of Human Skeletal Muscle. *J. Funct. Morphol. Kinesiol.* **1**, 289–302 (2016).
4. Musarò, A. The Basis of Muscle Regeneration. *Adv. Biol.* **2014**, 1–16 (2014).
5. Kuang, S. & Rudnicki, M. A. The emerging biology of satellite cells and their therapeutic potential. *Trends Mol. Med.* **14**, 82–91 (2008).
6. Seale, P. *et al.* Pax7 Is Required for the Specification of Myogenic Satellite Cells. *Cell* **102**, 777–786 (2000).
7. Angione, A. R., Jiang, C., Pan, D., Wang, Y.-X. & Kuang, S. PPARδ regulates satellite cell proliferation and skeletal muscle regeneration. *Skelet. Muscle* **1**, 33 (2011).
8. Zammit, P. S. *et al.* Muscle satellite cells adopt divergent fates a mechanism for self-renewal? *J. Cell Biol.* **166**, 347–357 (2004).
9. Conboy, M. J., Karasov, A. O. & Rando, T. A. High Incidence of Non-Random Template Strand Segregation and Asymmetric Fate Determination In Dividing Stem Cells and their Progeny. *PLoS Biol.* **5**, e102 (2007).
10. Milewska, M., Grabiec, K. & Grzelkowska-Kowalczyk, K. [Interactions of proliferation and differentiation signaling pathways in myogenesis]. *Postepy Hig. Med. Doswiadczalnej Online* **68**, 516–526 (2014).
11. Bentzinger, C. F., Wang, Y. X., Dumont, N. A. & Rudnicki, M. A. Cellular dynamics in the muscle satellite cell niche. *EMBO Rep.* **14**, 1062–1072 (2013).
12. Khan, P., Linkhart, B. & Simon, H.-G. Different Regulation of T-Box Genes Tbx4 and Tbx5 during Limb Development and Limb Regeneration. *Dev. Biol.* **250**, 383–392 (2002).
13. Rodriguez-Esteban, C. *et al.* The T-box genes Tbx4 and Tbx5 regulate limb outgrowth and identity. *Nature* **398**, 814–818 (1999).
14. Basson, C. T. *et al.* Mutations in human TBX5 [corrected] cause limb and cardiac malformation in Holt-Oram syndrome. *Nat. Genet.* **15**, 30–35 (1997).
15. Bruneau, B. G. *et al.* A Murine Model of Holt-Oram Syndrome Defines Roles of the T-Box Transcription Factor Tbx5 in Cardiogenesis and Disease. *Cell* **106**, 709–721 (2001).
16. Yamak, A. *et al.* Novel Exons in the Tbx5 Gene Locus Generate Protein Isoforms with Distinct Expression Domains and Function. *J. Biol. Chem.* **290**, 6844–6856 (2015).
17. Georges, R., Nemer, G., Morin, M., Lefebvre, C. & Nemer, M. Distinct Expression and Function of Alternatively Spliced Tbx5 Isoforms in Cell Growth and Differentiation. *Mol. Cell. Biol.* **28**, 4052–4067 (2008).
18. Hatcher, C. J. *et al.* TBX5 transcription factor regulates cell proliferation during cardiogenesis. *Dev. Biol.* **230**, 177–188 (2001).
19. Goetz, S. C., Brown, D. D. & Conlon, F. L. TBX5 is required for embryonic cardiac cell cycle progression. *Dev. Camb. Engl.* **133**, 2575–2584 (2006).

20. Messaoudi, S. *et al.* Endothelial Gata5 transcription factor regulates blood pressure. *Nat. Commun.* **6**, (2015).
21. Wang, J. *et al.* Convergence of protein kinase C and JAK-STAT signaling on transcription factor GATA-4. *Mol. Cell. Biol.* **25**, 9829–9844 (2005).
22. Yamak, A. *et al.* Cyclin D2 rescues size and function of GATA4 haplo-insufficient hearts. *Am. J. Physiol. Heart Circ. Physiol.* **303**, H1057-1066 (2012).
23. Yamak, A., Latinkic, B. V., Dali, R., Temsah, R. & Nemer, M. Cyclin D2 is a GATA4 cofactor in cardiogenesis. *Proc. Natl. Acad. Sci.* **111**, 1415–1420 (2014).
24. Rojas, A. *et al.* GATA4 Is a Direct Transcriptional Activator of Cyclin D2 and Cdk4 and Is Required for Cardiomyocyte Proliferation in Anterior Heart Field-Derived Myocardium. *Mol. Cell. Biol.* **28**, 5420–5431 (2008).
25. Zhu, W., Zhao, M., Mattapally, S., Chen, S. & Zhang, J. CCND2 Overexpression Enhances the Regenerative Potency of Human Induced Pluripotent Stem Cell–Derived Cardiomyocytes: Remuscularization of Injured Ventricle. *Circ. Res.* **122**, 88–96 (2018).
26. Khanjyan, M. V., Yang, J., Kayali, R., Caldwell, T. & Bertoni, C. A high-content, high-throughput siRNA screen identifies cyclin D2 as a potent regulator of muscle progenitor cell fusion and a target to enhance muscle regeneration. *Hum. Mol. Genet.* **22**, 3283–3295 (2013).
27. Foulstone, E. J., Savage, P. B., Crown, A. L., Holly, J. M. P. & Stewart, C. E. H. Role of insulin-like growth factor binding protein-3 (IGFBP-3) in the differentiation of primary human adult skeletal myoblasts. *J. Cell. Physiol.* **195**, 70–79 (2003).
28. Yuan, J., Tang, Z., Yang, S. & Li, K. CRABP2 Promotes Myoblast Differentiation and Is Modulated by the Transcription Factors MyoD and Sp1 in C2C12 Cells. *PLoS ONE* **8**, (2013).
29. Delorme, G., Saltel, F., Bonnelye, E., Jurdic, P. & Machuca-Gayet, I. Expression and function of semaphorin 7A in bone cells. *Biol. Cell* **97**, 589–597 (2005).
30. Morvan, F. *et al.* Deletion of a Single Allele of the Dkk1 Gene Leads to an Increase in Bone Formation and Bone Mass. *J. Bone Miner. Res.* **21**, 934–945 (2006).
31. Kaneko, Y. *et al.* Cell cycle-dependent and ATM-independent expression of human Chk1 kinase. *Oncogene* **18**, 3673–3681 (1999).
32. Prosdocimo, D. A., Sabeh, M. K. & Jain, M. K. Kruppel-Like Factors in Muscle Health and Disease. *Trends Cardiovasc. Med.* **25**, 278–287 (2015).
33. Zhang, Z. *et al.* Increase in HDAC9 suppresses myoblast differentiation via epigenetic regulation of autophagy in hypoxia. *Cell Death Dis.* **10**, 1–15 (2019).
34. Bruck, I., Perez-Arnaiz, P., Colbert, M. K. & Kaplan, D. L. Insights into the Initiation of Eukaryotic DNA Replication. *Nucleus* **6**, 449–454 (2015).
35. Xia, D. *et al.* Over-expression of CHAF1A in Epithelial Ovarian Cancer can promote cell proliferation and inhibit cell apoptosis. *Biochem. Biophys. Res. Commun.* **486**, 191–197 (2017).
36. Mahdy, M. A. A., Lei, H. Y., Wakamatsu, J.-I., Hosaka, Y. Z. & Nishimura, T. Comparative study of muscle regeneration following cardiotoxin and glycerol injury. *Ann. Anat. - Anat. Anz.* **202**, 18–27 (2015).
37. Guardiola, O. *et al.* Induction of Acute Skeletal Muscle Regeneration by Cardiotoxin Injection. *J. Vis. Exp. JoVE* (2017) doi:10.3791/54515.
38. Garrity, D. M., Childs, S. & Fishman, M. C. The heartstrings mutation in zebrafish causes heart/fin Tbx5 deficiency syndrome. *Development* **129**, 4635–4645 (2002).
39. Zhou, L. *et al.* Gata4 potentiates second heart field proliferation and Hedgehog signaling for cardiac septation. *Proc. Natl. Acad. Sci.* **114**, E1422–E1431 (2017).
40. Güemes, M. *et al.* GATA4 is Essential for Bone Mineralization via ERα and TGFβ/BMP Pathways. *J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res.* **29**, 2676–2687 (2014).
41. Spranger, S. *et al.* Muscular involvement in the Holt-Oram syndrome. *J. Med. Genet.* **34**, 978–981 (1997).
42. Hasson, P. *et al.* Tbx4 and Tbx5 acting in connective tissue are required for limb muscle and tendon patterning. *Dev. Cell* **18**, 148–156 (2010).

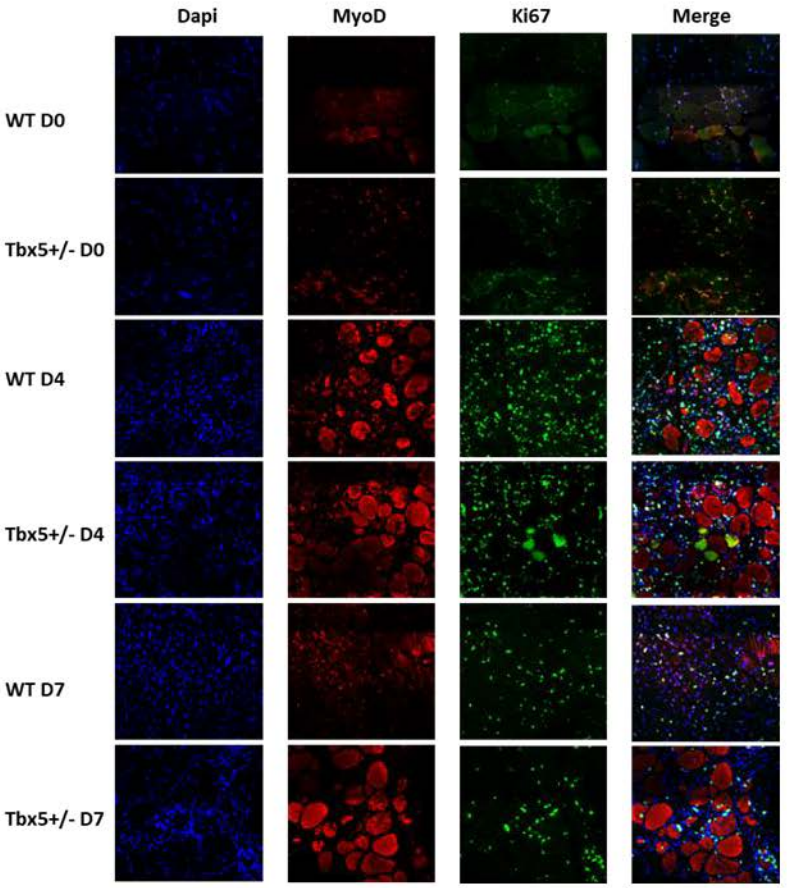
4.12 Supplementary Figures

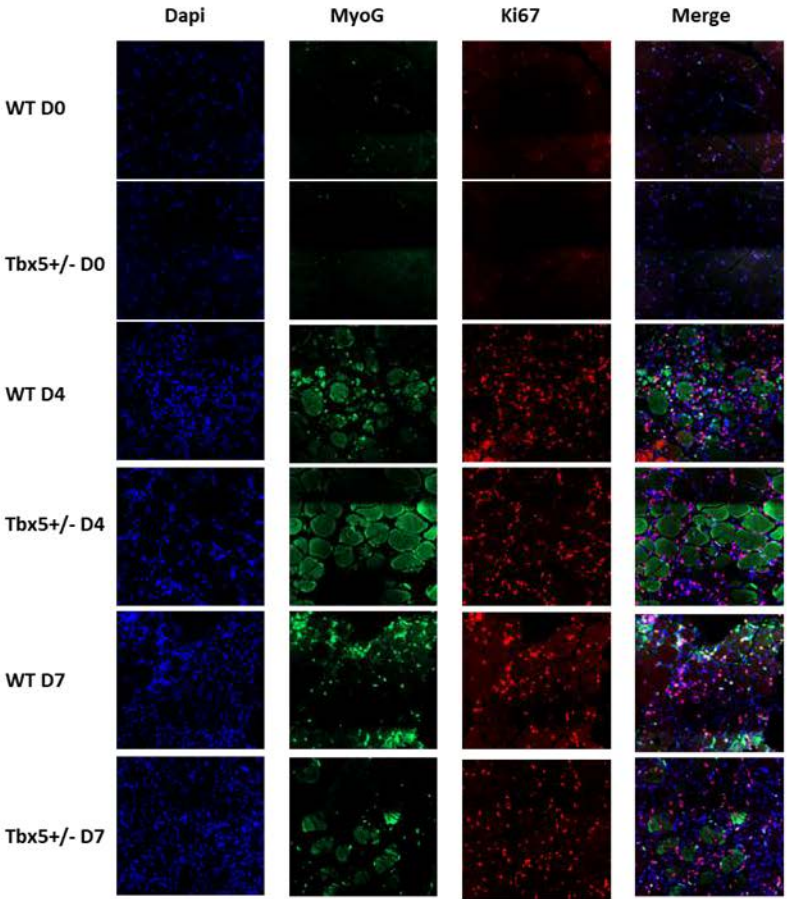


Supplementary Figure 4.1 Increased TBX5a expression in MDX skeletal muscle tissue. Immunohistochemistry staining for TBX5a in MDX skeletal muscle cells. The right side of the panel is MDX skeletal muscle (SM) stained for TBX5a and the left side stained for PH3 (proliferation marker) as indicated.



Supplementary Figure 4.2 TBX5 expression increased in MDX mice diaphragm and hindlimb tissues. A. Representative Immunofluorescence images on WT and MDX Diaphragm stained for PAX7 (Green) and TBX5a (Red) and Dapi (Blue). **B.** Representative Immunofluorescence images on WT and MDX Hindlimb EDL muscle stained for PAX7 (Green) and TBX5a (Red) and Dapi (Blue).


Dapi
MyoD
Ki67
Merge
WT D0
Tbx5+/- D0
WT D4
Tbx5+/- D4
WT D7
Tbx5+/- D7


Dapi
MyoG
Ki67
Merge
WT D0
Tbx5+/- D0
WT D4
Tbx5+/- D4
WT D7
Tbx5+/- D7

Supplementary Figure 4.3 *Tbx5*+/- mice exhibit impaired skeletal muscle regeneration. **A.** Representative Immunofluorescence images on WT and *Tbx5*+/- TA muscle at D0, D4 and D7 stained for MYOD (Red) and Ki67 (Green). B. Representative Immunofluorescence images on WT and *Tbx5*+/- TA muscle at D0, D4, and D7 stained for MYOG (Green) and Ki67 (Red).

5. General discussion

Prior to the work presented in this thesis, TBX5 had been established as a dosage sensitive regulator of heart and limb development. *TBX5* mutations have been directly associated with Holt Oram Syndrome (HOS), a disease characterized by a variety of cardiac defects and musculoskeletal abnormalities. In the heart, the role of TBX5 has been extensively studied by various groups and has led to the identification of several upstream pathways, protein partners, and gene targets. Collectively, these factors allow TBX5 to mediate its wide variety of roles therein [1–6]. In vertebrate embryogenesis, skeletal muscle is formed from progenitor cells originating in the somites from the paraxial mesoderm (the dorsal part of the somite called the dermomyotome), migrating to the skeletal muscles of the body and the limbs where they express myogenic determination factors and differentiate into skeletal muscle [7]. The ventral part of the somite (the sclerotome) gives rise to the other components of the skeletal system such as the bone, cartilage, vertebral columns and ribs [7]. TBX5 is also required musculoskeletal development in limb formation, but considerably less is known about its mechanism of action in skeletal muscle development [8–10]. The research presented in this thesis aims to help address this discrepancy by studying the role of TBX5 in skeletal muscle myoblasts. In doing so, we have uncovered the presence of 5 novel protein splice isoforms that regulate distinct cell fates, elucidated the roles of TBX5a/c in skeletal muscle myogenesis, and highlighted the importance of TBX5 in maintaining the regenerative capacity of skeletal muscle cells and enhancing muscle progenitor cell expansion. These critical roles in skeletal muscle homeostasis and function are made possible by specific protein-protein interactions that result in activation of distinct downstream gene targets,

genetic programs and cell phenotypes. Collectively, this work has provided novel insights into the molecular function and regulation of TBX5 in skeletal muscle cells and has potential clinical implications for inherited and acquired skeletal muscle disorders.

5.1 TBX5 is an essential regulator of skeletal myoblast development and function

Skeletal muscle development involves the cooperative interaction between many different transcription factors. Some of these transcription factors are muscle specific such as PAX3/7, MYF5, MYOD, MYOG and MRF4. These factors collectively regulate the myogenic cell lineage. However, myogenesis also requires the contribution of several non-muscle specific transcription factors as well. Disruption of these regulatory pathways leads to various types of muscular diseases. In this work, along with previous studies completed by our research group, we have identified TBX5 as an important regulator of myogenesis. TBX5 is not muscle-specific and has been well studied in both the heart and limb formation. In heart, it regulates several important processes including, cardiac specification, cardiomyocyte proliferation and cardiac conduction system function, as a few examples [5,6]. Despite the importance of TBX5 in cardiac cells in addition to the skeletal system described above, its role in skeletal muscle cells remains unclear. This work has elucidated for the first time the mechanistic role of TBX5 in skeletal muscle cells. The formation of skeletal muscular tissue during myogenesis consists of various stages including muscle progenitor cell activation and proliferation, myoblast proliferation, differentiation of myocytes into myotubes and fiber maturation [11]. Each stage is crucial for the development, function and integrity of the myogenic cell lineage. Previous studies by our group have shown that the

TBX5a isoform is predominantly expressed in proliferating C2C12 myoblasts whereas TBX5b is more abundant in differentiated myotubes [12]. Our work has now expanded on these findings and has identified TBX5a as a central regulator of myoblast proliferation that suppresses their differentiation into myotubes. Furthermore, conversely, we have also reported for the first time that the novel TBX5c isoform plays an opposing role to TBX5a, promoting differentiation from myoblast into myotube. Synchronized cell proliferation and differentiation is critical for both heart, limb and skeletal muscle morphogenesis thus a balanced ratio of all of the identified isoforms is likely required for all of these processes. In skeletal muscle, this balanced ratio would maintain sufficient progenitor expansion and differentiation while avoiding problems in the myogenic cell lineage. Currently, as the isoforms have been recently discovered, we do not know how each is expressed in its particular spatiotemporal manner. However, this information will be critical to the understanding of TBX5 functions in all tissues in which it is expressed. The discovery of these isoforms has also raised the interesting possibility that they promote or antagonize each other's function. Previous work by our group has shown that TBX5a and TBX5b play opposing roles in myogenesis where TBX5a correlates with growth stimulation and TBX5b expression is associated with myoblast growth arrest and cell death [12]. In addition, we have now shown that during myogenesis, TBX5a expression is involved with early stages associated with proliferation whereas TBX5c is involved with later stage differentiation. It is possible that they regulate each other's function at the level of the gene by transactivating or repressing each other's expression but it is also possible that they interact at the protein level as modifiers (a topic that will be discussed in the next section). Moreover, when the *TBX5* gene incurs mutations, they could significantly impact the expression and function of only

specific isoforms depending on the location of the mutation. This could impact their interactions with one another and with other protein partners and downstream gene targets (discussed in the next section). The importance of studying the skeletal muscle-specific roles of the TBX5 isoforms will lend greater insight into the function and regulation of the TBX5 protein in a spatiotemporal manner. As such, we will gain a greater understanding of normal myogenesis as well as the pathogenesis of musculoskeletal defects such as those associated with HOS and other inherited or acquired limb defects.

5.2 TBX5 modifiers influence skeletal myoblast formation and function

To improve our understanding of the involvement of TBX5 in normal muscle development and function as well as in musculoskeletal disorders, it is essential to study how TBX5 is regulated and influenced by other factors resulting in variable skeletal muscle phenotypes. HOS is linked to mutations in the *TBX5* gene and likewise, DMD is associated with mutations in the Dystrophin (*DMD*) gene. That said, the wide variety of phenotypes and levels of penetrance shown by individuals diagnosed with either condition suggest that the molecular mechanisms governing these phenotypes are complex and do not arise by the action of a single gene. Structural or functional problems with the proteins that mediate these pathways such as TBX5 (often as a result of mutations to the genes that encode them) could potentially interfere with a whole host of protein-protein interactions, post-translational modifications and transactivation of target genes. Likewise, mutations to genes encoding proteins that interact with TBX5 (referred to as modifiers) could directly impact its function, even if TBX5 itself is produced and functions normally. This could explain the variable

phenotypes and penetrance among humans and mice with diseases impacting skeletal muscle. To date, knockdowns or mutations of known TBX5 interacting protein partners that modify its function in cardiac cells have been extensively studied (listed in **Table 1.1**). Examples include interactions between TBX5 and two major cardiac transcription factors NKX2.5 and GATA4 which when diminished, reduce the expression of downstream cardiac gene targets such as *Nppa* resulting in cardiac defects [13–15]. In the limb as well as the heart, TBX5 is known to interact with and be tightly co-expressed with a PDZ-LIM protein composed of one N-terminal PDZ and three C-terminal LIM domains (named LMP4), which shows to be important in the regulation of heart and limb development [16]. However, despite its importance in the musculoskeletal system, no modifiers of TBX5 have been identified in skeletal muscle cells until this work. Our mass spectrometry approach used to identify TBX5a protein interactors in skeletal myoblasts uncovered many potential protein partners involved with a wide range of biological processes. Many of the identified TBX5a protein interactors were categorized into cell cycle and proliferation pathways which correlates to the known pro-survival and proliferative role of TBX5a established in cardiomyocytes [3]. Moreover, our earlier data suggested that TBX5a is more abundantly expressed in proliferating C2C12 myoblasts as compared to differentiated C2C12 myotubes and other cardiac cell lines. This correlates with many of the TBX5a protein interactors identified in our mass spectrometry results that are known to have similar roles in skeletal muscle cells. For example, our analysis identified various histone deacetylases including HDAC1 and 2. HDAC family members are known to both modify TBX5 activity in cardiac cells and suppress Myogenin in myocytes and thus reduce muscle cell differentiation [17,18]. For example, HDAC3 physically interacts with and deacetylates TBX5 and as a result,

cardiomyocyte lineage-specific genes whose expression are TBX5-dependent are reduced [17]. Our mass spectrometry results also included several other known modifiers of TBX5 in cardiomyocytes, suggesting a parallel function for this protein in both cell types. For example, we identified members of the NURD complex which has previously been established to interact with TBX5 to repress non-cardiac gene programs during cardiac development [19]. This interaction is important in cardiac development to a point where *TBX5* mutations disrupt this interaction and cause cardiac septal defects [19]. Similarly, the mass spectrometry analysis also identified members of the (SWI/SNF) chromatin-modifying complex, amongst which SMARCD3 (also known as BAF60c) is known to be essential in activating both cardiac and skeletal muscle programs [20,21]. Given the identification of potential TBX5a modifiers with known roles in both skeletal muscle and cardiac cells, our results suggest that there are common gene targets and pathways in the development of both cell types that hinge on TBX5 expression and function. However, similar proteins in the two systems might be regulated differently via different upstream signaling cascades and specific interactions with other modifiers and as such, lead to different target gene activation or repression. The common TBX5 protein interactors in heart and skeletal muscle cell lineages suggest a conserved TBX5 function central to the development and function of the two tissues. However, our data suggests that TBX5 expression is also modulated by different modifiers aside from those common to both cardiomyocytes and skeletal muscle cells. This contributes to the variable expression and function of TBX5 in the different tissues.

Amongst our identified TBX5 protein interactors are other members of the TBX family, such as TBX4, which would also be interesting to see whether they can modify TBX5 skeletal muscle activity. TBX5 and TBX4 are known to be co-expressed in many common

stages of both heart and limb development and seem to share common protein interactors (as the one discussed above). Therefore, it would be interesting to elucidate this interaction in skeletal myoblasts, demonstrating their resulting effect on the downstream pathways involved in skeletal muscle function. The different TBX genes may be able to interact and as a result modulate one another's activity by compensating for each other's function in distinct cell types. An example is the two family members, GATA4 and GATA6, which can compensate for each other's action in the cardiovascular system [22].

Our list of TBX5 protein interactors provides a wealth of data which can be further studied to investigate other potential modifiers of TBX5 in skeletal muscle cells which can potentially have roles in the development of musculoskeletal diseases. But for the purpose of this study, it was interesting to see which partners are involved in skeletal muscle proliferation or differentiation as we see opposing roles of the two TBX5a and TBX5c isoforms. This work has identified MYBBP1a, known to inhibit skeletal muscle differentiation [18], as a protein interactor of TBX5a and TBX5c. MYBBP1a is the first leucine zipper motif-like protein reported to interact with TBX5 and the first TBX5 physical interactor confirmed in skeletal myoblasts. Both TBX5 isoforms expressed in skeletal myoblasts, TBX5a and TBX5c, physically interact with MYBBP1a; however, the interactions play opposing roles. The interaction between TBX5a and MYBBP1a synergistically inhibit the skeletal muscle pro-differentiation factor Myogenin, whereas they activate the differentiation suppressor Myostatin. Conversely, the interaction between TBX5c and MYBBP1a leads to the transcriptional activation of Myogenin while inhibiting Myostatin. Protein isoforms from the same gene may have different interacting protein partners or interact with the same partner differently from one tissue to another [23]. Our

results suggest that MYBBP1a acts as a repressor with one isoform and an activator with another, likely activating some different target genes but also the same targets and pathways in different ways. The outcome of this interaction might be cell-type specific and dependent on the existence of other regulatory proteins or factors. The intron and exon alterations in the TBX5 isoforms may lead to various protein partners being recruited to the novel sites leading to variable activation of downstream targets and pathways. Likewise, some of the partners could be the same but since expression of TBX5a/c is spatially distinct, these partners preferentially interact with one or the other because its more readily available in a particular cell type. The impact of modifiers on the specific TBX5 isoforms can also help to explain how this protein can have such wide ranging and important roles in the body. Each TBX5 isoform will only be able to interact with modifiers present with them in the same cellular contexts. Given the unique spatiotemporal expression patterns of the different TBX5 isoforms, they will be able to interact with different modifiers and activate different downstream pathways in cardiac and skeletal muscle cell types. As well, their different protein domains would also provide specificity to their particular interacting partners. Our work has also suggested that all the isoforms can physically interact with the TBX5a protein suggesting that they could even act as modifiers to each other's function. The fact that these isoforms show differential activity in skeletal muscle cells suggests that they likely also have different binding partners, gene targets and some variability in binding motif recognition; and even so likely to be stimulated by different upstream signaling cascades.

Mutations in TBX5 modifiers will help us understand how and why mutations in HOS and other skeletal muscle disorders associated with *TBX5* lead to partially penetrant and heterogonous phenotypes. TBX5 is a dosage sensitive regulator of heart and limb

development, even minor changes to its expression or function are associated with the occurrence of cardiac and musculoskeletal defects. Thus, mutations in its modifiers can potentially modulate TBX5 function and expression affecting the downstream signaling pathways and ultimately disease. Although we have identified many potential TBX5 interactors in skeletal myoblasts, but there remain many unidentified protein interactors which their identification will aid us in better elucidating the genetic basis and the variable genotype-phenotype relations in TBX5 associated muscular diseases.

5.3 TBX5 target genes elucidate its function in skeletal muscle proliferation and differentiation

The limited understanding of TBX5 downstream target genes in skeletal myoblasts and the low incidence of HOS have delayed full understanding and investigation of its role in skeletal muscle development [24]. To date, multiple direct targets of TBX5 have been identified in cardiomyocytes such as *Nppa*, *Scn5a* and *Gja5* that are implicated in cardiac proliferation, maturation and function; however no direct targets of TBX5 have been reported in skeletal muscle cells [25]. In this work, we have identified novel TBX5 downstream targets that play important roles in skeletal muscle development and function. Our microarray analysis identified over 1200 potential TBX5a target genes in skeletal muscle myoblasts, functionally classified into various pathways indicating the extensive regulatory function of TBX5 on various types of genes networks and pathways. The majority of target genes seem to be categorized in DNA replication and cell cycle pathways confirming the proliferative role of TBX5a in myoblasts. Amongst the gene targets are various microRNAs (miRs) detected that are differentially expressed. miRs are now

commonly used as biomarkers in disease detection and have also shown to regulate TBX5 function in cardiomyocytes [26,27]. Patients with Rhabdomyosarcoma show elevated serum levels of various miRs such as *miR-1*, *miR-133a*, *miR-133b*, and *miR-206*, miRs that are known to display enriched expression in skeletal muscle [28,29]. It would be interesting to further analyze the effect of TBX5a on the detected miR genes and to identify whether they are dysregulated in HOS patients or other TBX5-associated skeletal muscle disorders. Also, among our microarray results were some previously identified TBX5 target genes in cardiac cells such as *Vegfc, Id2, Gja1* and *Mef2c*. The *Id2* promoter has been previously shown to be regulated by TBX5 and NKX2.5, resulting in the specification and coordination of the ventricular conduction system [30]. Likewise, the Connexin Family Component of Gap Junction (*Gja1*), known to be essential in regulating cell death, proliferation and differentiation in cardiomyocytes, was also amongst our identified TBX5 targets in skeletal muscle myoblasts. As well, we identified upregulation of members of the *Mef2* (Myocyte Enhancer Factor-2) family in our microarray results. In the heart, *Mef2c* is involved with cardiomyocyte differentiation and post-natal hypertrophy, whereas in skeletal muscle, it is involved with myoblast differentiation along with members of the MyoD family [31–33]. These interesting identified targets provide us a wealth of data to investigate but for the purpose of this study we focus on two significantly modified genes: Myostatin (*Mstn*) and Cyclin D2 (*CCnD2*). The activation of *Mstn* transcriptional activity by TBX5a provides evidence that the inhibition of myotube formation by TBX5a may happen through the myostatin negative regulatory pathway [34]. The *Mstn* gene network is known to regulate muscle proliferation and differentiation and is now widely being studied as a target to develop treatments for muscle wasting [35]. Similarly, TBX5a coordinately regulates progenitor cell expansion and myotube

formation through directly activating pro-proliferative gene targets such as *Ccnd2* and potentially other cyclins known to be involved in cell cycle. This pro-proliferative role in myoblasts seems to be synergistically coordinated with GATA4 leading to transcriptional activation of cell cycle associated gene *Ccnd2*, while sharing many other identified common genes in myoblasts, known to be involved in the regulation of proliferation and differentiation of myoblasts. Myoblast proliferation is also promoted by TBX5 directed activation of *Mstn* which inhibits differentiation of myoblasts into myotubes. Interestingly, as discussed in the previous section, TBX5c seems to have the opposing effect on the *Mstn* gene and promotes myotube formation and terminal differentiation. Identification of direct targets of TBX5, namely *Mstn* and *Ccnd2*, that regulate skeletal muscle proliferation and differentiation might unravel new pathways and networks governing skeletal muscle function. These findings can be useful for identifying novel or combinatory treatment strategies for skeletal muscle disorders which will be discussed below.

5.4 TBX5 pathways as therapeutic targets for skeletal muscle disorders

HOS patients or individuals with other types of disorders involving skeletal muscle might harbor a variety of mutations to TBX5 (including in the novel domains of the new isoforms) or to a TBX5 modifier. This could lead to dysregulation of downstream pathways and thus variable disease expressivity and phenotypes observed in musculoskeletal diseases. Therefore, it is crucial to understand downstream TBX5 regulatory networks in skeletal muscle cells in order to have a better understanding of how and why we observe the different disease phenotypes that often do not have a genotypic explanation. Spatio-temporal

expression and regulation of TBX5 in different stages of skeletal muscle myogenesis appears to be an important regulator in normal skeletal muscle generation and might be important in muscle affected by disease where there is imbalance of myoblast generation versus mature muscle formation. Muscular dystrophies are an example of muscular disorders where proliferating myoblasts are not able to sufficiently differentiate into myotubes and form enough mature muscle fibers to compensate for muscle loss due to the loss of Dystrophin expression resulting in muscle structural disorganization [36]. Our findings show a link between elevated TBX5a levels in MDX (DMD mouse model) skeletal muscle tissue and suggest association of the TBX5 pathways with proliferation and regeneration in MDX skeletal muscle tissues. As such, TBX5a could serve as a useful diagnostic tool in dystrophic tissue to detect increased myoblast proliferation along with other markers. Also, improved understanding of the pathways involving TBX5a, its modifiers and its downstream targets may allow for the identification of novel treatment strategies for those affected by DMD and other musculoskeletal disorders. For example, repressors of Myostatin are currently being used to treat DMD progression and to increase muscle mass in such muscle degenerative diseases. Our knowledge of novel regulators of this process may yet identify more targets. These new strategies could be implemented at earlier stages of the disease, where the muscle progenitor cell pool is still sufficient and muscle progenitor cells are still able to proliferate adequately.

Moreover, regardless of the cause, genetic or physical, diseased skeletal muscle tissue usually involves reduction and exhaustion of the myogenic progenitor and satellite cell pool. This is due to excessive degeneration and regeneration cycles that reduce the regenerative ability of the muscle tissue and ultimately lead to muscle wasting [37,38]. A genetic example is

the muscle fiber necrosis that results from defective Dystrophin gene and the gradual loss of the muscle regenerative ability in DMD patients [39]. It is important to understand the mechanism of impaired muscle regeneration in muscle dystrophies, cancers affecting the muscle and even where normal muscle wasting occurs with aging or in relation to other diseases or treatments [40]. We show in this work that TBX5a seems to be involved in the regeneration process by expansion of muscle progenitor cells. TBX5a expression was elevated in the first days after induction of acute muscle injury by Cardiotoxin and was also raised in the MDX mice that undergo chronic regeneration. Chemical induction of muscle injury to mice with heterozygous deletion of *Tbx5* impairs the regenerative capacity of the skeletal muscle tissue following degeneration. This is suggested by observing a decrease in markers of progenitor cells (satellite cells), proliferation markers and other myogenic markers necessary for proper skeletal muscle fiber formation. A significant decrease in muscle fiber formation in the skeletal muscle tissue of the *Tbx5* heterozygous mice suggests that TBX5 is important in maintaining skeletal muscle regenerative capacity. TBX5 function seems to be essential in proliferation and expansion of myogenic progenitor cells giving rise to myoblasts and would be important in muscular dystrophy patients where the muscle tissue progressively loses its regenerative ability resulting in muscle wasting and ultimately death. The identification of TBX5 targets important in myoblast proliferation such as *Ccnd2* allow for TBX5-mediated regulation of satellite cell and myoblast proliferation in the MDX mouse model. For example, a potential approach could be to use viral gene therapy to upregulate TBX5 downstream pro-proliferation targets to enhance muscle repair and regeneration in skeletal muscle disorders. Currently, various cell-based therapies are being produced that still face obstacles in the treatment of muscle degenerative diseases. These therapeutic

options in development include such promising examples as intramuscular delivery of myogenic cells, genetic correction of isolated satellite cells and transplantation and grafting of satellite stem cells [41]. We hope that elucidating TBX5 role in expansion of skeletal muscle progenitor cells and its novel downstream targets in skeletal muscles could help enhance current therapeutic approaches in myopathies such as DMD where muscle regeneration is insufficient or impaired. Moreover, our findings could provide further insight on the musculoskeletal problems in HOS patients (who incur TBX5 mutations), raising the possibility that TBX5 regulates musculoskeletal development through its function in all components of the skeletal system including muscle, bone and tendons.

Overall, the identification of TBX5 modifiers and targets regulating skeletal muscle differentiation and proliferation as described in the past two sections is a step towards better understanding the mechanistic network involved in TBX5-mediated skeletal muscle networks and pathways in muscle development and function; enabling its utilization as a diagnostic tool for broader mutational screenings and a prognostic tool by targeting TBX5 pathways in various types of muscle diseases.

5.5 Contributions to Knowledge

The pathways underlying skeletal muscle development and post-natal function are complex and are not completely understood to date. The incomplete understanding of regulatory networks and pathways involved in skeletal muscle development and function limits us in developing proper diagnostics and treatments for muscle disorders. By identifying its novel modifiers and downstream gene targets, the work presented in this thesis has identified novel

mechanisms governing cell and stage-specific cell fates in skeletal muscle proliferation and differentiation. This work provides insights into diagnostic and treatment strategies that could potentially be utilized for congenital and acquired skeletal muscle defects. However, this work has opened interesting questions to be answered and challenged in the field to fill the knowledge gap on this topic.

1) TBX5 isoforms as additional diagnostic tools in HOS screenings

Our work has discovered that TBX5 is highly regulated through alternative splicing and exists as five distinct isoforms with differential expression, especially in skeletal muscle cells. The novel isoforms present overlapping but also specific expression patterns in heart, limb and skeletal muscle. The variable domains of each isoform could potentially recruit different sets of protein interactors and cause variable transactivation of downstream genes and pathways. This allows for the independent, tissue-specific and unique functions of the isoforms. Currently, a major clinical challenge in HOS is that there are a wide range of associated TBX5 mutations as well as variable disease penetrance and heterogenicity. As such, associations between genotype and phenotype remain unclear [3]. TBX5 mutations could result in very different phenotypes depending on the isoforms impacted by the mutation. This would in turn affect the ratio between isoforms in different tissues and could also impact protein-protein interactions between the isoforms themselves and with other modifiers. Furthermore, the novel exons in the *TBX5* locus identified from our work may be applicable to mutational screenings for HOS patients where no mutations are found in the previously described TBX5a exons and may even provide insights on its involvement in other types of musculoskeletal disorders.

2) **Further insights from our mass spectrometry and microarray approach: potential protein interactors and target genes of TBX5 in skeletal muscle**

The unbiased mass spectrometry-based method used in this study to identify novel TBX5a interactors in skeletal muscle cells revealed over 180 different potential protein partners. One of these proteins, MYBBP1a, was studied in detail and showed to physically interact with TBX5a, inhibiting Myogenin, transcriptional activity while activating Myostatin. MYBBP1a and TBX5c interaction resulted in opposite action in both muscle promoters. The remaining list of TBX5 interactors in C2C12 cells provides a significant amount of data that could be studied to identify other mechanisms and pathways involved in skeletal muscle development and function. The microarray analysis of potential TBX5 downstream genes has also provided many interesting potential targets in skeletal muscle (over 1200 genes) that are worth further investigation. The advantage of these types of technologies is that they are fast and very informative. However, these protein partners and downstream genes must be confirmed via more direct approaches to confirm their interaction with TBX5a, their role in skeletal muscle cell fates and how they are impacted by disease. We aim to take this approach with several of the proteins and genes remaining from our list that will be discussed in the proceeding future directions section.

3) **Testing linkage between TBX5 interactors/targets and muscle diseases**

Our work has identified many potential TBX5 protein interacting partners and downstream targets; based on our work it would be interesting to determine whether mutations therein are

associated with muscle disorders such as Muscular Dystrophy. Usually in mutational screenings for skeletal muscle or cardiac disorders, a limited number of previously characterized and associated genes are sequenced. The mutations in the undiscovered muscle-associated genes will be missed in this case. The recent advancements in high throughput sequencing methods will enable entire genome sequencing of skeletal muscle disease patients allowing us to search for causative mutations in the newly discovered partners and targets of TBX5. This type of information will help elucidate the genotype-phenotype relationship in various TBX5-associated diseases by providing insight on the contribution of TBX5-mediated pathways in skeletal muscle development and function.

5.6 Future directions

Our work has opened many interesting and challenging questions to be followed up on with regards to the role of TBX5 and its gene regulatory network and pathways in skeletal muscle myocyte generation and function and its implications in skeletal muscle diseases. Some areas of interest that warrant future research are highlighted below:

1) **Investigation of other protein-protein interactors and target genes of TBX5a in skeletal muscle myoblasts**

In order to study the various identified protein interactors identified by our mass-spectrometry approach, we will observe their roles in skeletal muscle cells in culture and then use microarray or RNA sequencing to identify their downstream targets in C2C12 cells. The cell-specific roles of TBX5 partners can then be further explored *in vivo* by creating tissue specific knock out models. Promising candidates for further research include histone

deacetylases, specifically, HDAC1 and HDAC2. Given their involvement with the regulation of TBX5 activity in cardiomyocytes during cardiogenesis, we believe that they may play central roles in the development and function of skeletal muscle as well [17]. Also, MYBBP1a suppression of muscle differentiation is mediated through association with HDAC1/2 which stabilizes the co-repressor complex and maintains gene silencing of the downstream muscle-specific genes. Based on this knowledge it will be interesting to see if HDAC1/2 also complex with TBX5a and MYBBP1a to mediate the transactivation of the Myogenin or Myostatin promoters. This will help elucidate the bigger picture of what other TBX5a modifiers lead to its pro-proliferative and anti-differentiation role in myoblasts. The HDACs might have a role in the differential function of the two isoforms in skeletal muscle cells. This will be investigated by confirming this interaction by pulldown experiments and performing Co-IP experiments to see if they form a complex. Further analysis will be performed on muscle-specific genes using co-transfection experiments and luciferase assays to observe the effect of this complex. Then microarray or RNA sequencing could be carried out to identify the downstream targets of the TBX5 partners to further elucidate the pathway involved in myocyte differentiation to enhance our current understanding of TBX5 mechanism of action in these muscle types.

There are also some microRNAs (miRs) identified from our list of TBX5a downstream target genes which we think are interesting to further study. As previously mentioned, miRs have shown to be essential in regulation of gene expression in skeletal muscle development and disease. *miR155* has shown to be upregulated in RNA sequencing analysis of aged muscle tissue [42]. Interestingly, *miR155* is downregulated in our microarray results of TBX5a targets in myoblasts. TBX5a could be regulating the expression of such miRs to maintain the

myoblast proliferation characteristic of young, healthy skeletal muscle. This possibility would be further investigated by performing ChIP experiments to identify if this miR is a direct target of TBX5a. Then knock-down experiments can be performed to observe effect of its attenuation on other myogenic factors or partners of TBX5 by siRNA experiments. If we observe that m*iR155* is a direct target of TBX5a and that its downregulation would affect factors involved in skeletal myoblast differentiation and proliferation, then we can look further into how reduced levels of *miR155* correlate with myoblast proliferation in a disease context. Identification of such novel miRs and their mechanism of action in relation to key genes would be useful to understanding their contribution to human muscular diseases.

2) Investigation of TBX5c protein partners in skeletal muscle cells and its pro-differentiation role in Muscular Dystrophy and other skeletal muscle disorders

Our work has determined that the TBX5c isoform has an opposing role compared to TBX5a in skeletal muscle myocytes [43]. However, there still remains much to be learned about the role of this isoform with regards to skeletal muscle differentiation. This would be done by performing a similar type of mass spectrometry approach where TBX5c is overexpressed in C2C12 cells to identify TBX5c specific protein interactors and to see if there are different interacting partners and variable target gene activation/repression compared to TBX5a. It is recommended that TBX5c expression in the C2C12 cell line is transient since our experience shows that production of a stable cell with TBX5c overexpression is not viable, due to the isoforms ability to cause premature differentiation and growth arrest. Furthermore, TBX5c specific knock-out mice could be obtained by targeting exon 10 to see physiological effects *in vivo*. The pro-differentiation role of TBX5c might be a new window in alleviating and

decreasing effect of anti-differentiation factors that are raised in MDX mice and DMD patients such as Myostatin. Currently many inhibitors of myostatin are being studied and used for treatment of Muscular Dystrophies but none of which have shown to be fully efficient in treating Muscular Dystrophy symptoms [44]. Stage-specific modulation of TBX5a and TBX5c in skeletal muscle disorders might be a potential therapeutic approach used with other combinatory therapies.

5.7 References

1. Hatcher, C. J. & Basson, C. T. Holt-Oram Syndrome and the TBX5 Transcription Factor in Cardiogenesis. in *Molecular Genetics of Cardiac Electrophysiology* (eds. Berul, C. I. & Towbin, J. A.) 297–315 (Springer US, 2000). doi:10.1007/978-1-4615-4517-0_19.
2. Hatcher, C. J., Goldstein, M. M., Mah, C. S., Delia, C. S. & Basson, C. T. Identification and localization of TBX5 transcription factor during human cardiac morphogenesis. *Dev. Dyn.* **219**, 90–95 (2000).
3. Hatcher, C. J. *et al.* TBX5 transcription factor regulates cell proliferation during cardiogenesis. *Dev. Biol.* **230**, 177–188 (2001).
4. Bruneau, B. G. *et al.* A Murine Model of Holt-Oram Syndrome Defines Roles of the T-Box Transcription Factor Tbx5 in Cardiogenesis and Disease. *Cell* **106**, 709–721 (2001).
5. Arnolds, D. E. *et al.* TBX5 drives Scn5a expression to regulate cardiac conduction system function. *J. Clin. Invest.* **122**, 2509–2518 (2012).
6. Goetz, S. C., Brown, D. D. & Conlon, F. L. TBX5 is required for embryonic cardiac cell cycle progression. *Development* **133**, 2575–2584 (2006).
7. Buckingham, M. *et al.* The formation of skeletal muscle: from somite to limb. *J. Anat.* **202**, 59–68 (2003).
8. Agarwal, P. *et al.* Tbx5 is essential for forelimb bud initiation following patterning of the limb field in the mouse embryo. *Development* **130**, 623–633 (2003).
9. Rallis, C. *et al.* Tbx5 is required for forelimb bud formation and continued outgrowth. *Development* **130**, 2741–2751 (2003).
10. Hasson, P. *et al.* Tbx4 and Tbx5 acting in connective tissue are required for limb muscle and tendon patterning. *Dev Cell* **18**, 148–156 (2010).
11. Chal, J. & Pourquié, O. Making muscle: skeletal myogenesis *in vivo* and *in vitro*. *Development* **144**, 2104–2122 (2017).
12. Georges, R., Nemer, G., Morin, M., Lefebvre, C. & Nemer, M. Distinct Expression and Function of Alternatively Spliced Tbx5 Isoforms in Cell Growth and Differentiation. *Mol Cell Biol* **28**, 4052–4067 (2008).
13. Hiroi, Y. *et al.* Tbx5 associates with Nkx2-5 and synergistically promotes cardiomyocyte differentiation. *Nature Genetics* **28**, 276 (2001).
14. Garg, V. *et al.* GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. *Nature* **424**, 443–447 (2003).

15. Boogerd, C. J. J., Moorman, A. F. M. & Barnett, P. Expression of Muscle Segment Homeobox Genes in the Developing Myocardium. *The Anatomical Record* **293**, 998–1001 (2010).
16. Krause, A. *et al.* Tbx5 and Tbx4 transcription factors interact with a new chicken PDZ-LIM protein in limb and heart development. *Developmental Biology* **273**, 106–120 (2004).
17. Lewandowski, S. L. *et al.* Histone deacetylase 3 modulates Tbx5 activity to regulate early cardiogenesis. *Hum Mol Genet* **23**, 3801–3809 (2014).
18. Yang, C.-C. *et al.* Epigenetic silencing of myogenic gene program by Myb-binding protein 1a suppresses myogenesis. *EMBO J* **31**, 1739–1751 (2012).
19. Waldron, L. *et al.* The cardiac TBX5 interactome reveals a chromatin remodeling network essential for cardiac septation. *Dev Cell* **36**, 262–275 (2016).
20. Puri, P. L. & Mercola, M. BAF60 A, B, and Cs of muscle determination and renewal. *Genes Dev* **26**, 2673–2683 (2012).
21. Lickert, H. *et al.* Baf60c is essential for function of BAF chromatin remodelling complexes in heart development. *Nature* **432**, 107–112 (2004).
22. Xin, M. *et al.* A threshold of GATA4 and GATA6 expression is required for cardiovascular development. *Proceedings of the National Academy of Sciences* **103**, 11189–11194 (2006).
23. Alternative proteins encoded by the same gene have widely divergent functions in cells. *ScienceDaily* https://www.sciencedaily.com/releases/2016/02/160211141339.htm.
24. Pizard, A. *et al.* Connexin 40, a Target of Transcription Factor Tbx5, Patterns Wrist, Digits, and Sternum. *Mol Cell Biol* **25**, 5073–5083 (2005).
25. Steimle, J. D. & Moskowitz, I. P. TBX5: A Key Regulator of Heart Development. *Curr Top Dev Biol* **122**, 195–221 (2017).
26. Guzzolino, E. *et al.* miR-182-5p is an evolutionarily conserved Tbx5 effector that impacts cardiac development and electrical activity in zebrafish. *Cell. Mol. Life Sci.* (2019) doi:10.1007/s00018-019-03343-7.
27. Wang, F. *et al.* miR-10a and miR-10b target the 3'-untranslated region of TBX5 to repress its expression. *Pediatr Cardiol* **35**, 1072–1079 (2014).
28. Suzuki, T. & Springer, J. MicroRNAs in muscle wasting. *J Cachexia Sarcopenia Muscle* **9**, 1209–1212 (2018).
29. Siracusa, J., Koulmann, N. & Banzet, S. Circulating myomiRs: a new class of biomarkers to monitor skeletal muscle in physiology and medicine. *J Cachexia Sarcopenia Muscle* **9**, 20–27 (2018).
30. Moskowitz, I. P. G. *et al.* A Molecular Pathway Including Id2, Tbx5, and Nkx2-5 Required for Cardiac Conduction System Development. *Cell* **129**, 1365–1376 (2007).
31. Akazawa Hiroshi & Komuro Issei. Roles of Cardiac Transcription Factors in Cardiac Hypertrophy. *Circulation Research* **92**, 1079–1088 (2003).
32. Taylor, M. V. & Hughes, S. M. Mef2 and the skeletal muscle differentiation program. *Semin. Cell Dev. Biol.* **72**, 33–44 (2017).
33. Liu, N. *et al.* Requirement of MEF2A, C, and D for skeletal muscle regeneration. *Proceedings of the National Academy of Sciences* **111**, 4109–4114 (2014).
34. Langley, B. *et al.* Myostatin Inhibits Myoblast Differentiation by Down-regulating MyoD Expression. *J. Biol. Chem.* **277**, 49831–49840 (2002).
35. Bradley, L., Yaworsky, P. J. & Walsh, F. S. Myostatin as a therapeutic target for musculoskeletal disease. *Cell. Mol. Life Sci.* **65**, 2119–2124 (2008).
36. Oexle, K. & Kohlschütter, A. Cause of Progression in Duchenne Muscular Dystrophy: Impaired Differentiation More Probable Than Replicative Aging. *Neuropediatrics* **32**, 123–129 (2001).
37. Khan, P., Linkhart, B. & Simon, H.-G. Different regulation of T-box genes Tbx4 and Tbx5 during limb development and limb regeneration. *Dev. Biol.* **250**, 383–392 (2002).
38. Sacco, A. *et al.* Short Telomeres and Stem Cell Exhaustion Model Duchenne Muscular Dystrophy in mdx/mTR Mice. *Cell* **143**, 1059–1071 (2010).

39. Luz, M. A. M., Marques, M. J. & Santo Neto, H. Impaired regeneration of dystrophin-deficient muscle fibers is caused by exhaustion of myogenic cells. *Braz J Med Biol Res* **35**, 691–695 (2002).
40. Jang, Y. C., Sinha, M., Cerletti, M., Dall'Osso, C. & Wagers, A. J. Skeletal Muscle Stem Cells: Effects of Aging and Metabolism on Muscle Regenerative Function. *Cold Spring Harbor Symposia on Quantitative Biology* **76**, 101–111 (2011).
41. Bentzinger, C. F., Wang, Y. X., von Maltzahn, J. & Rudnicki, M. A. The emerging biology of muscle stem cells: Implications for cell-based therapies. *Bioessays* **35**, 231–241 (2013).
42. Kim, J. Y. *et al.* Genome-wide profiling of the microRNA-mRNA regulatory network in skeletal muscle with aging. *Aging (Albany NY)* **6**, 524–544 (2014).
43. Yamak, A. *et al.* Novel Exons in the Tbx5 Gene Locus Generate Protein Isoforms with Distinct Expression Domains and Function. *J Biol Chem* **290**, 6844–6856 (2015).
44. Saitoh, M., Ishida, J., Ebner, N., Anker, S. D. & Haehling, S. von. Myostatin inhibitors as pharmacological treatment for muscle wasting and muscular dystrophy. in (2017). doi:10.17987/jcsm-cr.v2i1.37.

6. Curriculum Vitae

NAME **Massomeh Sheikh-Hassani**

LANGUAGES English & Persian (fluent) and learning French

EDUCATION/DEGREES

2003-2004 Diploma in Biological Sciences from Wellington East Girl's College, Wellington – New Zealand.

2004-2007 Double Degree in Biomedical Sciences: "Human Genetics" and "Molecular Pathology", Victoria University of Wellington - New Zealand.

2008-2010 Masters of "Biology" (Plant Biosystematics) from Tarbiat Moallem University of Tehran-Iran.

2012-2013 Guest PhD student in Medical Genetics at Tehran University of Medical Sciences, Tehran-Iran.

2014-2020 PhD in Biochemistry- Specialization of Human & Molecular Genetics from the University of Ottawa, Ottawa-Canada.

RESEARCH EXPERIENCE

2008-2010 Master thesis on "Biosystematic investigation of some Olive (Olea europaea L.) genotypes across different regions of the mountain ranges of Zagros of Iran using Morphological and Molecular markers" in the Tarbiat Moallem University of Tehran, Faculty of Bilogical Sciences, Tehran, Iran.

2008-2010 The National project of "Investigation and Identification of Iran's Olive Germplasm" in the "National Institute of Genetic Engineering and Biotechnology

(NIGEB)", Molecular Genetics Department, Olive Biotechnology Research Team, Tehran, Iran.

2010 Olive Genotyping using SSR Markers in the Institute of Plant Genetics, "Istituto di Genetica Vegetale" (CNR-IGV), Perugia, Italy. Supervised by Professor Luciana baldoni (Head of the Olive team of the Institute).

2011-2012 Graduate researcher at the Cardiovascular Devices Division (CVD) of University of Ottawa Heart Institute (UOHI). I worked on a DNA sequencing project with my supervisor Dr. Tofy Mussivand (inventor of an Artificial Cardiac Pump) with the objective `to develop a rapid, accurate, portable, cost-effective device for DNA sequencing`.

2012 Research and studies on metabolism of lipids and lipoproteins (Fatty liver) with my co-supervisor Dr. Zemin Yao, Professor and chairman in the Department of Biochemistry, Microbiology, and Immunology and the Department of Pathology and Laboratory Medicine at the University of Ottawa.

2014-now Research on Molecular Genetics of Heart Development in Dr. Mona Nemer's lab (Thesis Supervisor- former Vice President of Research of uOttawa and current Chief Science Advisor of Canada) in the Department of Biochemistry, Microbiology, and Immunology and the Department of Pathology and Laboratory Medicine at the University of Ottawa.

PUBLICATIONS/ABSTRACTS

1) **Sheikh-Hassani, M.**, Torkzaban, B., Ataei, S., Zeinanloo, A., Ghahremaninejad, F., Hosseini-Mazinani, M., 2010, *Morphological evaluation of ancient olive varieties across the Zagros Mountains of Iran*, International Horticulture Congress (IHC) Portugal-Lisbon August.

2) **Sheikh-Hassani M.**, Torkzaban B., Hosseini Mazinani M., Baldoni L., Pandolfi S., Cultrera N.G.M., Mariotti R., 2010, *Evaluation of the Iranian Olive germplasm by SSR and chloroplast markers*, Proceedings of the 54th Italian Society of Agricultural Genetics Annual Congress, Matera, Italy – 27/30 September, ISBN 978-88-904570-0-5.

3) Hosseini-Mazinani M, Mariotti R, Torkzaban B, **Sheikh-Hassani M**, Ataei S, Cultrera NG, et al. High genetic diversity detected in olives beyond the boundaries of the Mediterranean Sea. PLoS One 2014 Apr 7;9(4):e93146.

4) Yamak A, Georges RO, **Sheikh-Hassani M**, Morin M, Komati H, Nemer M. Novel exons in the tbx5 gene locus generate protein isoforms with distinct expression domains and function. J Biol Chem 2015 Mar 13;290(11):6844-6856.

BOOKS

- Contributor and English editor: **Massomeh Sheikh-Hassani**, Authors: Mehdi HosseiniMazinani and Bahareh Torkzaban (2013), Iranian olive catalogue "morphological and molecular characterization of Iranian olive germplasm", NIGEB , ISBN: 978-964-8516-23-4, 220 pages.

ACTVITIES and PRESENTIONS

- Coordinator and organizer of the "Olive morphological and molecular identification" workshop in the National Institute of Genetic Engineering and Biotechnology of Iran, 2010.
- Attendance to the 5th International ROYAN Stem cell Congress in Iran, 2009.
- Attended the 1st National Biotechnology Congress of Iran, 2010.
- National MedOliva congress in Arezzo-Italy, 2010.
- Involved in the Medical Device conference in Ottawa Heart Institute, August 2011.
- Attended professional development activities for teaching assistants offered by the Centre for University Teaching (CUT) of the University of Ottawa in winter 2012.
- Contributed in organizing and planning the third annual Medical devices summit 2012 (held by the Medical Devices Innovation Institute, MDI2), as the treasurer and adviser of the organizing committee, in the spring/summer of 2012.

- Grant writings for research projects at the Medical Devices Innovation Institute as a graduate researcher (i.e. for Natural Sciences and Engineering Research Council of Canada (NSERC)), 2011-2012.
- Attended the Third Iranian Medical Genetics Congress, Children's Medical Center, Tehran University of Medical Sciences, 15-16th May, 2013.
- Poster presentation on research project "Characterization of Novel Tbx5 Protein isoforms" in BMI poster day at the University of Ottawa, May 2014
- Seminar presentation on research project "Characterization of Novel Tbx5 Protein isoforms" in BMI Seminar day at the University of Ottawa, March 2015.
- Poster presentation in the "3rd Ottawa International Conference on Neuromuscular Biology, Disease and Therapy 2015". Poster title: "Title: A novel *Tbx5* protein isoform has a critical role in muscle differentiation".
- Seminar presentation for the department on research project "Role of Novel Tbx5c isoform in skeletal muscle" in BMI Seminar day at the University of Ottawa, March 2017.
- Poster presentation on "Role of Tbx5 in proliferation and differentiation of skeletal muscle cells" in University of Ottawa Heart Institute (UOHI) annual Research Day, May 2017.
- Poster presentation on "Role of Tbx5 in proliferation and differentiation of skeletal muscle cells" in CSMB-Celebrating Canadian Molecular Biosciences – from Organelles to Systems Biology, May 2017.
- Poster presentation on "Role of Tbx5 in proliferation and differentiation of skeletal muscle cells" in 4th Ottawa International Conference on Neuromuscular Disease Biology, September 2017.
- Poster presentation on "Role of Tbx5 in proliferation and differentiation of skeletal muscle cells" in 2019 OISB Scientific Conference, May 2019.

WORK EXPERIENCE

- Undergraduate Laboratory experience in the field of human genetics and molecular pathology (In Victoria University of Wellington-NZ).
- Graduate lab experience in the fields of molecular genetics: DNA Extraction, PCR,

Capillary Electrophoresis, Genotyping using SSR markers (In National Institute of Genetic Engineering & Biotechnology of Iran, & "Istituto di Genetica Vegetale" (CNRIGV), Perugia, Italy).

- Teaching laboratory techniques (Molecular Genetics) in workshops.
- Familiar with statistical softwares such as: SPSS, GenAlex, Genemapper, Mega4, etc.
- Research assistant at the Medical Devices Innovation Institute (MDI2) from May 2011 to September 2012.
- Research assistant at the from May 2011 to September 2012.
- Research assistant in Molecular Genetics of Heart Development in Dr. Mona Nemer's lab 2014-present.

AWARDS

- Travel and research scholarship from the Ministry of Agriculture and the National Institute of Genetic Engineering and Biotechnology (NIGEB) of Iran, for a project collaboration between Iran and Italy, during my master's degree (2010).
- Travel reimbursements for conferences by University of Ottawa (2015-2019).
- An admission scholarship of $9,000 per year for my first twelve full-time sessions of Ph.D. by the Faculty of Graduate and Postdoctoral Studies (FGPS) of the University of Ottawa, in addition to a minimum of $17,850 per year awarded by my academic unit for the duration of the scholarship (2011-2017).

INTERESTS

- Cellular & Molecular Biology: Stem cell research and novel therapeutics.
- Human Genetics: areas concerning genetic diseases and linkage analysis.
- Cellular medicine: Experimental pathology & Disease identification and treatment using animal models.
- Learning languages, Hiking, Biking, bungee jumping, skiing and outdoor activities, Instruments (Piano), Communications and networking, passionate for plants and gardening.